A Characteristics Approach to Optimal Taxation: Line Drawing and Tax-Driven Product Innovation*

August 2015
Keywords: Tax design, notches
JEL Code: H21

Abstract

Real-world tax legislation assigns goods to different tax rate categories on the basis of observable characteristics, allowing the tax system to handle a constantly evolving set of available goods. We recast the theory of optimal taxation in the language of characteristics and show how to optimally draw lines that delineate tax rate regimes. Such lines are associated with notches in tax liability as a function of characteristics, creating incentives to introduce goods with new characteristics combinations in order to reduce tax liability. With a restricted set of tax instruments, such notches are in general part of the second-best optimal tax system.

Christian Gillitzer
Reserve Bank of Australia
Sydney, NSW 2000, Australia
gillitzerc@rba.gov.au

Henrik Jacobsen Kleven
Department of Economics
London School of Economics and Political Science
London WC2A 2AE, United Kingdom
h.j.kleven@lse.ac.uk

Joel Slemrod
Stephen M. Ross School of Business
University of Michigan
Ann Arbor, MI 48109, United States
jslemrod@umich.edu

*We thank Alan Auerbach, Tim Besley, Robin Boadway, Peter Fredriksson, Louis Kaplow, Wojciech Kopczuk, Dan Shaviro, David Weisbach, Shlomo Yitzhaki, and anonymous referees for helpful comments. We also thank Yulia Paramonova for valuable research assistance. Disclaimer: Christian Gillitzer: Opinions expressed should not be attributed to the Reserve Bank of Australia.

I Introduction

Optimal tax theory in general prescribes different tax rates on each commodity, but in practice consumption tax systems feature a small number of distinct rates. A non-capricious commodity tax system must have procedures for distinguishing among goods subject to the different tax rates. Real-world tax systems do that by appealing to the characteristics of the commodities. For example, American states' retail sales taxes often exempt food purchases but not restaurant meals, requiring the tax code to draw a line between the two categories. This is done by appealing to a set of characteristics of a restaurant meal, and the line can be fine such as when grocery stores sell pre-prepared meals that may or may not be eaten on the premises, or set up in-store salad bars. The retail sales tax in the Canadian province of Ontario exempts basic food items such as flour but applies to other processed foods such as chocolate bars, requiring lines to be drawn, including one that subjects to tax "biscuits or wafers specifically packaged and marketed to compete with chocolate bars." Several European countries provide a subsidy for certain kinds of consumer services (e.g., cleaning, gardening, and house repair) based on a Ramsey-type justification that such services compete with untaxed home production. This requires the classification of services eligible for the subsidy based on observable characteristics.

Although line drawing is a ubiquitous issue in real-world tax systems and a pervasive point of contention among tax lawyers, there is little economic analysis of the issue. We note that a "line" shares many attributes of a "notch" in tax schedules, which refers to a discontinuity in the function of how tax liability relates to the tax base, and which have attracted much recent attention among economists as a means of identifying behavioral response elasticities (e.g. Kleven and Waseem 2013). Indeed, a line creates a notch in characteristics space, because the tax liability changes discontinuously when the characteristics vector of a good crosses the statutory line. As long as a continuum of tax rates is administratively infeasible, notches in characteristic space are an unavoidable feature of tax systems, not an idiosyncrasy.

The prominent role of characteristics in commodity tax systems is due to several factors. First, using observable characteristics is a natural and intuitive way to distinguish among different goods,

or different groups of goods, and assign them to tax-rate categories. The alternative that the theory of optimal commodity taxation implies—classifying goods according to compensated elasticities—is infeasible, both because these elasticities are notoriously difficult to estimate precisely and because this kind of rule would not be intuitive to either policy makers, voters or consumers in the way that characteristics-based rules are. Second, a shared characteristic plausibly signals something about the relative substitutability of the goods, and so may serve as a more readily measurable indicator of the ideal, but not observable, determinants of the appropriate tax rate.[1] Third, modern economies produce a vast amount of different goods, and the set of available goods is constantly evolving. If tax laws were specified literally in terms of goods and their associated elasticities, then there would be no natural way to assign a new good to a tax category and the law would have to be re-specified to explicitly deal with the new good. In contrast, a characteristics-based rule for assigning tax rates to goods naturally handles the creation of new goods by limiting the tax policy choice to which characteristic-based category the new good falls in.

In this paper we reformulate optimal commodity tax theory in the language of characteristics so that it matches up more easily with real tax systems. To do so we make use of the idea developed by Gorman (1980) and Lancaster (1966, 1975) that there exists a mapping of each good into characteristics space, and that it is the characteristics of goods, not the goods themselves, that generate utility.[2] In the standard optimal tax model, addressing the creation of new goods is not tractable, because a change in the set of available goods must be associated with a new utility function (with new arguments) and therefore a new optimal tax problem. In the Gorman-Lancaster approach, on the other hand, as long as the set of characteristics that consumers value is stable, the utility function is robust to the introduction of new goods and product innovation can then be incorporated into the optimal tax problem. In this paper we augment the standard model of optimal commodity taxation by developing an analysis of optimal line drawing and notches in characteristic space.

[1] Kleven and Slemrod (2009) addresses this problem with a completely distinct modeling approach, formalizes the relationship between characteristics, substitutability and optimal tax rates, and shows that the closer two goods are in characteristics space, the smaller the optimal tax rate differential.

[2] Although Gorman's paper did not appear in a journal until 1980, it was originally written in 1956 and therefore predates Lancaster's work.

Reformulating optimal tax theory in this way allows us to address an important aspect of reality that has been ignored by the literature on optimal taxation, namely tax-driven product innovation.[3] By this term we refer to the creation of new products, i.e., new characteristics combinations, which are introduced in the market in response to the tax system. For example, the prevalence of salad bars and cafes inside supermarkets may be in part a response to the differential tax treatment of restaurant meals and food purchased in grocery stores. In developing countries that impose higher taxes on automobiles than on other types of vehicles, industries emerge that produce low-tax vehicles that share many characteristics with cars. For example, the preferential tax treatment of motorcycles in Indonesia led to the creation of a new type of motorcycle with three wheels and long benches at the back seating up to eight passengers—car-like but not so car-like as to be taxed as cars. When Chile imposed much higher taxes on cars than on panel trucks, the market soon offered a redesigned panel truck that featured glass windows instead of panels and upholstered seats in the back.[4] Arrow manufacturers in the United States avoided an excise tax on arrow shafts measuring 18 inches in length or longer by making 17.9 inch long arrows.[5] Even "fashion" responds to tax incentives: Abba's *The Official Photo Book* reveals that "...the band's style was influenced in part by laws that allowed the cost of outfits to be deducted against tax—so long as the costumes were so outrageous they could not possibly be worn on the street" (Bowers, 2014).

We show that notches in characteristic space are generally optimal when there is a restricted set of tax instruments available (i.e., in all real-world applications). Existing results on the undesirability of notches crucially depend on the existence of a complete set of tax instruments. For example, Mirrlees (1971) shows that the optimal income tax schedule does not feature marginal tax rates above 100 percent when tax liability can be arbitrarily non-linear (which rules out notches as they create implicit marginal tax rates of more than 100 percent over an interval). Other work has argued in favor of notches, but not clarified the dependence on the set of tax instruments available: Blinder and Rosen (1985) show that, when non-linear commodity taxes are infeasible, notches pro-

[3] We also discuss tax-driven product destruction.
[4] These examples are taken from Harberger (1995).
[5] Arrows have notches!

vide a means to effectively target a favored behavior, but leave after-tax prices unchanged for those away from the notch. Dharmapala et al. (2011) show that discontinuous tax treatment in firm size is optimal when there are firm-level administrative costs.

Our model is related to a literature studying optimal income taxation with restricted instruments, in particular work on piecewise linear income taxes (see e.g., Slemrod et al. 1994). A conceptual difference is that we study an optimal tax system featuring *notches*, whereas the income tax literature focuses on *kinks*—discontinuities in *marginal* tax rates—as tax liability is assumed to be continuous. It is sensible to rule out notches in a piecewise linear income tax system, because Mirrlees (1971) showed that in general such discontinuities should be avoided. But when considering commodity taxes as a function of characteristics, notches are inevitable if tax differentiation across goods is to be achieved with real-world restrictions on available tax instruments.

Our paper contributes to the large literature on optimal commodity taxation and proposes a framework that has implications for optimal income taxation and the theory of tax avoidance and evasion more generally.[6] Related to the Gorman-Lancaster characteristics approach analyzed here, Kleven (2004) considers a Becker household production approach where the factor shares of goods and household time in different household activities represent the characteristics that govern optimal taxation. Within the standard optimal commodity tax model, Gordon (1989), Weisbach (1999, 2000), Belan and Gauthier (2004, 2006), and Belan, Gauthier, and Laroque (2008) have studied a question related to line drawing: how to group goods into a limited set of tax categories. This set of papers offers rules for grouping goods based on compensated demand elasticities and possibly distributional weights. Similarly, Yitzhaki (1979) and Wilson (1989) analyze how to draw the line between a set of taxed and untaxed goods in a world where uniform taxation is optimal, but where expanding the tax base is associated with administrative costs.

As far as we are aware, none of the earlier literature addresses the salient features of real-world tax systems that we explore: characteristics-based tax rules, line drawing and notches in charac-

[6]For recent surveys of the literature on optimal commodity and income taxation, we refer to Auerbach and Hines (2002) and Salanié (2003). The literature on tax avoidance and evasion has been surveyed by, e.g., Slemrod and Yitzhaki (2002) and Shaw, Slemrod, and Whiting (2010).

teristics space, and tax-driven product innovation. Although we address these issues in the context of a Ramsey-style optimal consumption tax framework, we believe that they are a ubiquitous feature of all forms of taxation as well as many regulatory provisions. This includes income taxation where different forms of income are treated differently, requiring lines to be drawn based on the characteristics of different income forms and where new types of compensation may be introduced in order to facilitate tax avoidance.[7]

II Model Setup

Consumers

A type-ω consumer derives utility

$$u(x, c_1, c_2; \omega) \tag{1}$$

from consumption of a numeraire good x, and from consumption of a characteristics pair (c_1, c_2). By assumption, good x cannot be taxed, and in most applications corresponds to leisure. The characteristics pair (c_1, c_2) is acquired from the consumption of a single unit of a good embodying this particular mix of characteristics. For example, the characteristics pair (c_1, c_2) may represent the characteristics of a car valued by consumers, with c_1 carrying capacity and c_2 style. The decision problem that we model is what type of good to buy in characteristics space as opposed to how much of a given type of good to buy. We consider a 2-dimensional rather than n-dimensional characteristics vector because this is the most parsimonious and tractable way to model characteristics. The model can be extended to $n > 2$ characteristics dimensions, but this would add complexity with relatively little additional insight. A model with $n = 1$ in which there is a one-to-one mapping between goods and characteristics effectively corresponds to the standard consumer model that underlies the existing optimal tax theory.

Consumers are heterogeneous in the parameter ω, which controls their preferences over the

[7] Slemrod and Gillitzer (2014) provide more examples and place this set of questions in a larger context of the study of tax systems.

two characteristics. We assume that the marginal rate of substitution of c_2 for c_1 is increasing in ω. Thus, at any set of prices, the larger is ω the greater is a consumer's preference for characteristic c_2 relative to c_1. (A concrete example with nested CES preferences is provided in an online appendix.) The distribution of consumer types is given by the cumulative distribution function $F(\omega)$. All consumers are assumed to have the same exogenous full-income endowment Y.

Production

Perfectly competitive firms produce a potential continuum of goods, each embodying whatever mix of characteristics (c_1, c_2) is demanded by consumers. For analytical simplicity, firms are assumed to operate a linear production technology: the cost to produce, and therefore the pre-tax cost to consume, a good embodying the characteristics pair (c_1, c_2) is given by $z \equiv z(c_1, c_2) = \sum_j q_j c_j$, where the exogenous pre-tax "price" of characteristic c_j is q_j.

In keeping with the optimal tax theory literature, the goods-generating technology does not depend on the tax system. The analysis therefore does not deal with the potential effect of taxation on technology-changing innovations that allow previously infeasible characteristics combinations to be produced ("technology-driven product innovation"). It deals instead with the effect of the tax system on product innovations that consist of a re-packaging of characteristics within an already feasible set in order to reduce tax liability ("tax-driven product innovation"). The evidence discussed in the introduction suggests that this is an empirically important phenomenon.

Tax Instruments

It is administratively infeasible to set a unique tax rate for each of the infinitely many goods in characteristic space (c_1, c_2). There is a limited set of tax instruments that permits assigning goods embodying all possible characteristic pairs to a finite set of tax categories. We consider for simplicity a case with two tax categories: the tax authority can delineate a "line" $\bar{c}_2$ in (c_1, c_2) space, and set a tax rate τ_1 on goods for which $c_2 \leq \bar{c}_2$ and τ_2 on goods for which $c_2 > \bar{c}_2$.[8] The tax

[8] Kleven and Slemrod (2009) define tax regions using a ray rather than a horizontal line at $\bar{c}_2$ in characteristics space. Our formulation using a line is much more tractable and practically relevant: tax categories are typically defined by the

policy is thus given by $\{\tau_1, \tau_2, \bar{c}_2\}$. The tax is assumed to be ad-valorem: the cost to a consumer of purchasing a good with pre-tax cost $z(c_1, c_2)$ is equal to $(1+\tau_k) z(c_1, c_2)$, given tax rate $\tau_k \equiv 1(c_2 \leq \bar{c}_2)\tau_1 + 1(c_2 > \bar{c}_2)\tau_2$, where the indicator function $1(\cdot)$ takes the value one whenever the argument is satisfied and zero otherwise; the implicit consumer "price" of characteristic j is $p_j \equiv (1+\tau_k) q_j$.

III Consumer Responses to Characteristics Notches

A type-ω consumer maximizes utility $u^\omega \equiv u(x, c_1, c_2; \omega)$ subject to the budget constraint

$$Y = x + z(c_1, c_2) + T(c_1, c_2, \tau_1, \tau_2, \bar{c}_2), \tag{2}$$

where $z(c_1, c_2) \equiv q_1 c_1 + q_2 c_2$ is the pre-tax cost of purchasing the characteristics pair (c_1, c_2), and

$$T(c_1, c_2, \tau_1, \tau_2, \bar{c}_2) = 1(c_2 \leq \bar{c}_2) \cdot \tau_1 z(c_1, c_2) + 1(c_2 > \bar{c}_2) \cdot \tau_2 z(c_1, c_2) \tag{3}$$

is the tax liability when the consumer purchases the characteristics pair (c_1, c_2) and the tax policy is given by $\{\tau_1, \tau_2, \bar{c}_2\}$.

A Graphical Treatment

In the panels of Figure 1 we illustrate the effect on consumer behavior of a notched tax system, using two-dimensional graphs to illustrate a three-dimensional choice problem. First consider the left-hand panel, which applies when goods corresponding to each characteristics pair are taxed at the same rate τ_1, i.e. with no notch. (For clarity, the figure is drawn, without loss of generality, assuming that $q_1 = q_2 = 1$.) The diagonal black line intersecting the axes is a purely hypothetical baseline, indicating (c_1, c_2) pairs available to a consumer when none of their exogenous income,

absolute amount of a given characteristic (e.g., weight of a vehicle) rather than by a ratio of characteristics (e.g., weight of a vehicle relative to horsepower). With our formulation, tax arbitrage is possible if multiple goods with different quantities of each characteristic can be combined. We therefore focus on the case where consumers purchase a single good.

Y, is spent on the numeraire good, x, say leisure.[9] But in general some leisure will be consumed. The solid grey line lying below the diagonal black line indicates characteristics pairs chosen by consumers with different values of ω, and the vertical and horizontal distance from the solid grey line to the diagonal black line indicates the amount of income spent on leisure. The larger is ω the greater is a consumer's marginal rate of substitution of c_1 for c_2, and the further to the north-west on the solid grey line is the consumer's chosen characteristics pair. The curvature (if any) of the solid grey line indicates substitution between leisure and taxed characteristics as ω changes. We assume that the marginal utility of leisure is increasing in consumption of c_2 relative to c_1, so consumption of leisure rises with c_2 and the solid grey line curves away from the diagonal black line as ω rises.

[Figure 1 about here]

Now consider the right-hand panel of Figure 1, which illustrates the case when goods embodying characteristic pairs (c_1, c_2) with $c_2 > \bar{c}_2$ are taxed at a rate $\tau_2 > \tau_1$. In this case we can distinguish between three types of consumers illustrated in the figure. First, consumers with $\omega \leq \underline{\omega}$ are unaffected by the imposition of a notch at $\bar{c}_2$, because their chosen good in the absence of a notch features $c_2 \leq \bar{c}_2$. The good represented by point A is chosen by the consumer of type $\underline{\omega}$, who is the highest ω-type who prefers the same characteristic mix whether or not a higher tax rate applies above $\bar{c}_2$.

Second, a set of consumer types $\bar{\omega} > \omega \geq \underline{\omega}$ who purchase goods featuring $c_2 > \bar{c}_2$ in the absence of a notch, distort their purchases to goods featuring $c_2 = \bar{c}_2$ when the notch is introduced. In the figure, these are the consumers located on the line segment A-B without a notch and on the line segment A-D with the notch. Hence there is bunching of consumers at the characteristics notch $\bar{c}_2$. The consumer of type-$\bar{\omega}$ is the marginal bunching individual: this consumer purchases the good at point B in the absence of a notch, and is exactly indifferent between the goods at points C and D under a notched tax system; the latter good avoids the higher tax rate but requires purchasing

[9] Thus the lines in Figure 1 discussed here and below do not represent budget sets faced by any taxpayer, and therefore it is not meaningful to show indifference curves on these diagrams. In principle we could show both budget sets and indifference curves of a particular taxpayer, but only in a three-dimensional diagram that would not be visually helpful.

a good with a much more drastic change in characteristic mix. Formally, points B, C and D are characterized as follows:

Definition (Points B, C and D in Figure 1): *Let $\bar{\omega}$ be the consumer type who, after the imposition of a notch at $\bar{c}_2$ with $\tau_2 > \tau_1$, is indifferent between purchasing a good featuring $c_2 > \bar{c}_2$ and a good featuring $c_2 = \bar{c}_2$. Points $C = \left(c_1^C, c_2^C\right)$ and $D = \left(c_1^D, \bar{c}_2\right)$ show the pair of goods that, in the presence of a notch, give equal utility for the type-$\bar{\omega}$ consumer: $u\left(c_1^C, c_2^C; \bar{\omega}\right) = u\left(c_1^D, \bar{c}_2; \bar{\omega}\right)$. Point B shows the good that the same type of consumer would purchase in the absence of a notch (i.e., when all goods are taxed at rate τ_1).*

The precise location of point D along the horizontal line depends on the substitutability among x, c_1 and c_2.

Lemma 1 (Location of Point D in Figure 1): *Ceteris paribus, point D in Figure 1 lies further to the right the more complementary is consumption of the second characteristic with the numeraire, x, than is consumption of the first characteristic.*
Proof: In the appendix. □

Third and finally, consumers with $\omega > \bar{\omega} > \underline{\omega}$ have sufficiently strong preferences for c_2 that they continue to purchase goods featuring $c_2 > \bar{c}_2$ after a notch has been imposed at $\bar{c}_2$. While they do not bunch at the notch, the higher tax rate on goods induces them to substitute toward leisure and to change the characteristics of the goods they buy. In the figure, these consumer types are located on the line segment containing point C in the presence of the notch (rather than on the dashed grey line above point B absent a notch).[10] In the equilibrium with a characteristics notch, no consumers purchase goods between point C (a good with $\hat{c}_2$ units of characteric 2) and point D (a good with $\bar{c}_2$ units of characteric 2), creating a hole in the characteristics distribution between $\bar{c}_2$ and $\hat{c}_2$.

[10]Note that the line segment containing point C is steeper than the line segment above point B (as in the example in the figure) whenever characteristic c_2, which features more prominently in goods purchased by high-ω type consumers, is more complementary with x than is characteristic c_1.

A Formal Treatment

Having illustrated the effect of a notch on consumer behavior graphically, we now describe the consumer's maximization problem formally. A type-ω consumer's Lagrangian is

$$L = u(x, c_1, c_2; \omega) + \alpha^{\omega} \left[Y - x - z(c_1, c_2) - T(c_1, c_2, \tau_1, \tau_2, \bar{c}_2) \right], \tag{4}$$

where the Lagrange multiplier α^{ω} is equal to the marginal utility of exogenous income for a type-ω consumer. The consumer's first-order condition for x is

$$\frac{\partial L}{\partial x} = \frac{\partial u^{\omega}}{\partial x} - \alpha^{\omega} = 0, \tag{5}$$

and their implicit first-order condition for c_1 is

$$\frac{\partial L}{\partial c_1} = \frac{\partial u^{\omega}}{\partial c_1} - \alpha^{\omega} \left[1(c_2 \leq \bar{c}_2)(1+\tau_1) q_1 + 1(c_2 > \bar{c}_2)(1+\tau_2) q_1 \right] = 0. \tag{6}$$

(Because characteristics are not acquired separately, but via consumption of goods embodying a particular characteristics mix, we refer to the consumer's *implicit* demand for characteristics). For all consumers of type $\omega \leq \underline{\omega}$ (who consume characteristic pairs with $c_2 \leq \bar{c}_2$) and $\omega \geq \bar{\omega}$ (who consume characteristic pairs with $c_2 \geq \hat{c}_2 > \bar{c}_2$), their implicit choice of c_2 has an interior optimum solution to the first-order condition

$$\frac{\partial L}{\partial c_2} = \frac{\partial u^{\omega}}{\partial c_2} - \alpha^{\omega} \left[1(c_2 \leq \bar{c}_2)(1+\tau_1) q_2 + 1(c_2 > \bar{c}_2)(1+\tau_2) q_2 \right] = 0. \tag{7}$$

Bunching consumers, who purchase goods featuring $c_2 = \bar{c}_2$, do not have an interior optimum for $\bar{c}_2$.

Implications: Tax-Driven Product Innovation and Destruction

The introduction of a notch causes *tax-driven product innovation* if it induces the production of new *types* of goods. Each ray through the origin in (c_1, c_2) space represents a unique characteristic mix and type of good, because it is possible to move along a ray by adjusting the quantity of an existing good. For example, if coarseness and nutrition content are the defining characteristics of flour, then two one-pound packs of flour are the same good as one two-pound pack of flour with the same coarseness and nutrition content.

For a notch to cause tax-driven product innovation there must be some characteristics rays that are not represented in the pre-notch equilibrium (i.e., if every conceivable good in characteristics space already exists, production innovation is of course not possible). The absence of certain product types (rays) may arise either because some rays represent peculiar characteristic mixes for which there is no demand absent a notch, or because some characteristic mixes are inherently more costly to produce. As an example of the former, consumers demand exercise machines and strong whiskey separately, but no product that combines the characteristics of the two. In our framework, such non-existing goods can be captured by allowing for a non-continuous support in the distribution of the consumer preference parameter ω. Figure 2 shows a pre-notch equilibrium in which there is no demand for goods on rays between O-E and O-F. (The other line segments in Figure 2 are the same as in Figure 1.) The introduction of a notch at $\bar{c}_2$ induces consumption (and production) of goods along the horizontal line segment in Figure 2. Because there were no goods on rays through this line segment in the pre-notch equilibrium, the introduction of the notch induced tax-driven product innovation.

[Figure 2 about here]

The real-world is replete with goods designed to fall just on the low-tax side of a tax-price line (Harberger 1995; Slemrod 2013), embodying peculiar characteristic mixes that would not be observed absent discontinuous tax treatment of otherwise similar goods. An example is the three-wheeled cars marketed at one time in Indonesia, providing car-like features but taxed preferentially

relative to cars. The three-wheeled car can be thought of as a good lying on the horizontal line segment in Figure 2: relative to a world without taxes, it features an unusually low level of one characteristic (e.g., wheels/stability, on the y-axis) relative to its other characteristic (e.g., passenger capacity, on the x-axis). Generally, whether or not the introduction of a notch induces tax-driven product innovation depends on the distribution of consumer preferences over characteristic pairs (c_1, c_2) and the location of the line, $\bar{c}_2$.

Our model also explains *tax-driven product destruction*, where the tax system drives out from the market certain characteristic pairs, creating a *hole* in the distribution of characteristics produced. In the right-hand panel of Figure 1, the notch eliminates products on rays between O-C and O-D. Regardless of the distribution of consumer preferences, a notch induces tax-driven product destruction, unless it is placed in an empty region where there is no demand for goods featuring c_2 near $\bar{c}_2$. For a consumer who purchased a good featuring c_2 slightly greater than $\bar{c}_2$ prior to the introduction of the notch, substituting to a good that features $c_2 = \bar{c}_2$ has a negligible (second-order) effect on maximized utility, but results in a discrete (first-order) reduction in tax liability.

IV Social Planner's Problem

Having examined the effect of a notched tax system on consumer behavior, and the equilibrium set of goods produced, we now consider the social planner's problem. The benevolent social planner maximizes a concave and individualistic social welfare function $W\left[v^{\omega}\left(\tau_1, \tau_2, \bar{c}_2\right)\right]$, subject to an exogenous revenue requirement G, where $v^{\omega}\left(\tau_1, \tau_2, \bar{c}_2\right)$ is a type-ω consumer's indirect utility function. The planner can choose the height of the tax-price line $\bar{c}_2$, and set the tax rate applying to goods above and below the line. Formally, the Lagrangian for the planner's problem is given by

$$L = \int_{\omega^{min}}^{\omega^{max}} W\left[v^{\omega}\left(\tau_1, \tau_2, \bar{c}_2\right)\right] dF\left(\omega\right) + \lambda \left[\tau_1 \int_{\omega^{min}}^{\underline{\omega}} z\left(c_1^{\omega}, c_2^{\omega}\right) dF\left(\omega\right) + \tau_1 \int_{\underline{\omega}}^{\bar{\omega}} z\left(c_1^{\omega}, \bar{c}_2\right) dF\left(\omega\right) + \tau_2 \int_{\bar{\omega}}^{\omega^{max}} z\left(c_1^{\omega}, c_2^{\omega}\right) dF\left(\omega\right) - G\right], \tag{8}$$

where λ is the Lagrange multiplier on the aggregate budget constraint. The first term in the expression for the aggregate budget constraint is revenue raised from consumers on the low-tax side of $\bar{c}_2$, who are unaffected by the notch; the middle term is revenue raised from the mass of consumers who bunch on the low-tax side of the line at $\bar{c}_2$; and the final term is revenue raised from consumers on the high-tax side of $\bar{c}_2$. The maximization problem is expressed assuming that the optimal tax policy features $\tau_2 \geq \tau_1$. This is without loss of generality because the model is symmetric about $\bar{c}_2$: if instead the optimal tax policy featured $\tau_1 \geq \tau_2$ then a hole would arise below rather than above $\bar{c}_2$, and the propositions we state below would be modified in an intuitive manner.

The Desirability of a Notched Tax System

Before characterizing the solution to the planner's problem, we consider the desirability of a notched tax system in a hypothetical world with a complete set of tax instruments (i.e., one that permits a possibly unique tax rate for each of the infinitely many goods defined by characteristic pairs). A notch makes tax liability non-linear and discontinuous in the amount of the second characteristic consumed, whereas a complete set of tax instruments permits tax liability to be arbitrarily non-linear without requiring it to be discontinuous. If consumer preferences are continuous, the welfare-maximizing planner will not introduce a notch (a discontinuity in tax liability) given a complete set of tax instruments:

Remark 1 (Notch is Suboptimal with Complete Instruments): *Assume that consumers have continuous preferences, and that the planner has a complete set of tax instruments permitting arbitrary tax liability on each of the infinitely many potential goods in characteristic space. Then the welfare-maximizing tax policy does not feature a discontinuity in tax liability (a notch).*
Proof: In the appendix. □

The intuition for this result is the following. The locally infinite marginal tax rate at a notch creates a hole in the distribution of characteristics, distorting consumer choice and raising no revenue on characteristic mixes in the hole. The discrete jump in tax rates at the notch can be smoothed

over the range of the hole by a continuous tax function. A sufficiently steep tax function leaves consumer choice (and tax revenue raised) above and below the hole unchanged relative to the notched tax system, but presents bunching consumers with a lower tax-rate schedule in the region of the hole, inducing them to move into the hole. Because previously bunching consumers pay more tax (as tax rates in the hole are larger than at the notch point) *and* are better off by revealed preference (as goods at the notch point is still available at an unchanged tax), social welfare is strictly higher when the hole is removed.

With at most two tax rates, it is optimal to differentiate the tax rates, and thus introduce a notch, if, under certain conditions, the tax base elasticity differs above and below any arbitrarily chosen tax-price line $\bar{c}_2$. At least some differentiation in tax rates between the regions is optimal because the efficiency cost of taxation is lower when a disproportionately large share of revenue is raised from the less elastic tax base. Proposition 1 provides a formal statement of this result.

Proposition 1 (Notch is Optimal with Incomplete Instruments): *Assume that the social planner has a restricted set of tax instruments that only permits setting a different tax rate on goods featuring c_2 greater or less than some arbitrary level $\bar{c}_2$. In this case the optimal tax policy features a tax-price notch in characteristic space (except in a knife-edge case). In the special case of no equity concerns (constant social marginal utilities of income across consumers), the optimal tax policy features a notch whenever the tax base elasticity above and below $\bar{c}_2$ differs.*
Proof: In the appendix. □

To provide some intuition for Proposition 1, we next discuss the optimal relative tax rates in the regions above and below a (not necessarily optimally chosen) tax-price line $\bar{c}_2$.

The Optimal Tax System

In the standard model, in which c_1 and c_2 are distinct goods that can be purchased and taxed separately, rather than characteristics embodied in a single good, the optimal choice of tax rates satisfies the Corlett and Hague (1953-54) rule: the optimal tax rate is highest on the good that is

most complementary with the numeraire, typically leisure, which by assumption cannot be taxed. Intuitively, taxing the good most complementary with leisure at a high rate provides an indirect means of taxing leisure. For example, suppose the first taxed good is skis, and the second taxed good is a work uniform. Taxing skis at a higher rate than uniforms mitigates the substitution away from taxed goods to untaxed leisure.

In our restricted tax instrument framework, in which goods embodying characteristics can be taxed, but not the underlying characteristics themselves, the optimal choice of tax rates applying to goods above and below the line $\bar{c}_2$ has a familiar Corlett-Hague interpretation: goods featuring more prominently the characteristic more complementary with the (untaxed) numeraire are taxed at a higher rate than goods featuring more prominently the good less complementary with the numeraire. Proposition 2 describes the optimal relative tax rates in the region above and below $\bar{c}_2$.

Proposition 2 (Optimal Tax Rates): *Assuming the social marginal utility of income, γ, is constant across consumers, the optimal tax policy satisfies the following Corlett-Hague type rule for any given height of the tax-price line $\bar{c}_2$:*

$$\frac{\tau_1/(1+\tau_1)}{\tau_2/(1+\tau_2)} = \Gamma \frac{1/[1-F(\bar{\omega})]\int_{\bar{\omega}}^{\omega^{max}} (z^{\omega}/z^{u}) \left[\theta_1^{\omega}\varepsilon_{10}^{c} + \theta_2^{\omega}\varepsilon_{20}^{c}\right] dF(\omega)}{1/F(\bar{\omega})\int_{\omega^{min}}^{\bar{\omega}} \left(z^{\omega}/z^{l}\right) \left[\theta_1^{\omega}\varepsilon_{10}^{c} + \theta_2^{\omega}\varepsilon_{20}^{c}\right] dF(\omega)}, \tag{9}$$

where ε_{j0}^{c} is the compensated demand elasticity between characteristic j and the numeraire, assumed to be the same for all consumers, $\theta_j^{\omega} \equiv \left(q_j c_j^{\omega}\right)/z^{\omega}$ is the pre-tax cost share of characteristic j in the good bought by a type-ω consumer, with $\sum_k \theta_k^{\omega} = 1$, while $z^{u} \equiv 1/[1-F(\bar{\omega})]\int_{\bar{\omega}}^{\omega^{max}} z^{\omega} dF(\omega)$ and $z^{l} \equiv 1/F(\bar{\omega})\int_{\omega^{min}}^{\bar{\omega}} z^{\omega} dF(\omega)$ are the mean (pre-tax) value of expenditure on the characteristics good by consumers who purchase goods above and below the line $\bar{c}_2$, respectively. Finally,

$$\Gamma \equiv \frac{(1-\gamma/\lambda) - \left[\Delta R(\bar{\omega})/z^{l}\right]\left[f(\bar{\omega})/F(\bar{\omega})\right]\left[\partial\bar{\omega}/\partial\tau_1\right]}{(1-\gamma/\lambda) - \left[\Delta R(\bar{\omega})/z^{u}\right]\left[f(\bar{\omega})/1-F(\bar{\omega})\right]\left[\partial\bar{\omega}/\partial\tau_2\right]} \tag{10}$$

measures the social value of revenue spillovers across tax bases (the regions above and below $\bar{c}_2$), with $\Delta R(\bar{\omega}) \equiv \left(\tau_2 z\left(c_1^{\bar{\omega}}, c_2^{\bar{\omega}}\right) - \tau_1 z\left(c_1^{\bar{\omega}}(\bar{c}_2), \bar{c}_2\right)\right)$ being the change in revenue when a type-$\bar{\omega}$ consumer moves from bunching at $\bar{c}_2$ to the high-tax region.

Proof: In the appendix. □

The numerator of (9) measures the average compensated substitution among consumers purchasing goods featuring $c_2 > \bar{c}_2$ between the taxed good and the untaxed numeraire. The ratio (z^{ω}/z^{u}) is a type-ω consumer's expenditure on the taxed good relative to other consumers in the region above $\bar{c}_2$, and $\left[\theta_1^{\omega}\varepsilon_{10}^{c} + \theta_2^{\omega}\varepsilon_{20}^{c}\right]$ is compensated substitution from the taxed good to untaxed leisure, for a type-ω consumer; it is larger the more prominent is the characteristic most substitutable with leisure. The denominator of (9) has the same interpretation for consumers in the region below $\bar{c}_2$. If the second characteristic is more complementary with the numeraire than the first characteristic, $\varepsilon_{20}^{c} < \varepsilon_{10}^{c}$, the numerator of (9) is smaller than the denominator, because consumers in the region above $\bar{c}_2$ tend to purchase goods featuring a high share of the second characteristic compared to consumers in the region below $\bar{c}_2$. Hence, it is optimal to tax goods featuring $c_2 > \bar{c}_2$ at a higher rate than goods featuring $c_2 < \bar{c}_2$: $\tau_2 > \tau_1$.

The larger is $\tau_2 - \tau_1$, the greater is the consumption of goods that fall on the low-tax side of $\bar{c}_2$, reducing revenue collections and increasing the distortion in goods consumed. Expression (10) measures the welfare consequences of bunching behavior, which moderates the optimal degree of tax rate differentiation between the high- and low-tax regions. The term $(1-\gamma/\lambda) > 0$ in (10) is the social cost (measured in dollars) of raising a marginal dollar of revenue. It is assumed that, for any tax-system policy vector under consideration, the social planner assigns marginal social welfare weights such that γ is the same for all consumers. Abstracting from distributional considerations does not affect our qualitative conclusions, but simplifies the resulting expressions, allowing us to focus on the efficiency considerations that are relevant to our analysis. The term $\Delta R(\bar{\omega}) > 0$ is the revenue gain when a marginal buncher (a type-$\bar{\omega}$ consumer who is indifferent between consuming at or above the notch) moves from the low-tax region to the high-tax region. The ratio $[f(\bar{w})/F(\bar{\omega})]$ measures the density of marginal bunchers relative to the share of consumers in the region below $\bar{c}_2$, and analogously for the ratio $[f(\bar{w})/(1-F(\bar{\omega}))]$; the higher is the line $\bar{c}_2$, the fewer is the number of consumers above $\bar{c}_2$, and the larger is $F(\bar{\omega})$. When $\tau_2 > \tau_1$, a marginal increase in τ_1 reduces the incentive to bunch at $\bar{c}_2$, reducing $\bar{\omega}$, and thus implying $\partial\bar{\omega}/\partial\tau_1 < 0$; by similar reasoning,

$\partial\bar{\omega}/\partial\tau_2 > 0$. Hence, $\Gamma > 1$ and the optimal degree of differentiation between τ_1 and τ_2 is lower than were there no bunching behavior.

The optimal tax rates τ_1 and τ_2 depend on the location of the line $\bar{c}_2$. A shift of the line $\bar{c}_2$ affects the set of consumers in each tax region, and thus the average value of the (compensated) elasticity of substitution between taxed characteristics and the untaxed good in each region: the lower is $\bar{c}_2$, the greater is the extent to which the average elasticity for the region above $\bar{c}_2$ reflects the population-average elasticity, and conversely for the region below $\bar{c}_2$. The location of the line also affects the importance of marginal bunchers relative to the set of consumers in each tax category: the larger is $f(\bar{\omega})/F(\bar{\omega})$ the more important are marginal bunchers for the region below $\bar{c}_2$ and, analogously, the larger is $f(\bar{\omega})/(1-F(\bar{\omega}))$ the more important are marginal bunchers for the region above $\bar{c}_2$.

When the location of the line $\bar{c}_2$ is such that it groups together consumers with similar compensated elasticities of substitution between taxed characteristics and the untaxed good, the difference in average compensated elasticities between the two groups is relatively large, and thus so is the difference in the optimal tax rates τ_1 and τ_2. The benefit of grouping together consumers with most similar elasticities must be balanced against socially costly bunching behavior induced by a large difference between τ_1 and τ_2. The larger is the difference in tax rates τ_1 and τ_2, the greater is the number of consumers who distort their consumption toward characteristic goods that fall just on the low-tax side of $\bar{c}_2$. The marginal social cost of the distortion for each bunching consumer is equal to the difference between the marginal utility of consumption for c_2 and the shadow cost of purchasing a characteristics good with a marginally higher level of c_2: $u_2^{\omega} - \alpha^{\omega} p_2 \geq 0$; it is largest for a type-$\bar{\omega}$ marginal buncher, and smallest for a type-$\underline{\omega}$ consumer.

Proposition 3 (Optimal Placement of Line/Notch): *Without loss of generality, assume that an optimum features $\tau_2 > \tau_1$. At an optimum, the height of the line $\bar{c}_2$ satisfies the following first-order necessary condition:*

$$\int_{\underline{\omega}}^{\bar{\omega}} \frac{W'}{\lambda}\left(u_2^{\omega} - \alpha^{\omega} p_2\right) dF(\omega) = \Delta R(\bar{\omega}) f(\bar{\omega}) \frac{\partial \bar{\omega}}{\partial \bar{c}_2} - \int_{\underline{\omega}}^{\bar{\omega}} \tau_1 \left(q_1 \frac{\partial c_1^{\omega}}{\partial \bar{c}_2} + q_2\right) dF(\omega). \tag{11}$$

Proof: In the appendix. □

The term in parentheses on the left-hand-side of (11) is the reduction in the consumption distortion for each bunching consumer from a marginal increase in the height of the line $\bar{c}_2$: a marginal increase in $\bar{c}_2$ allows bunching consumers to purchase a characteristics good with a slightly less constrained level of c_2 and still avoid the high-tax region. At an optimum, the reduction in the distortion from bunching is equal to the marginal revenue consequences of an increase in $\bar{c}_2$, shown by the terms on the right-hand-side of Equation (11). The first term on the right-hand-side of Equation (11) captures the revenue consequences of bunching behavior. The term $\partial\bar{\omega}/\partial\bar{c}_2$ is positive, because a marginal increase in the height of $\bar{c}_2$ induces some consumers who previously consumed at $\hat{c}_2$ in the high-tax region to bunch at $\bar{c}_2$ in the low-tax region. There is a loss of $\Delta R(\bar{\omega})$ in tax revenue on each marginal buncher who, following a marginal increase in $\bar{c}_2$, chooses to bunch at $\bar{c}_2$ in the low-tax region, and there is a mass $f(\bar{\omega})$ of marginal bunchers. Ceteris paribus, it is optimal to place the notch where few individuals will be induced to jump (i.e., where the density of types $f(\omega)$ is small); this might be at a particularly unattractive characteristic mix. The term in parentheses on the right-hand-side of (11) is the rise in the value of purchases by bunching consumers when $\bar{c}_2$ is marginally increased, on which taxes are collected at rate τ_1, partly offsetting the revenue loss from increased bunching.

In an online appendix, we consider a special case of this problem when the consumer has nested CES preferences. We show that if the expenditure share on the untaxed good does not vary across consumers in the absence of taxes, the social marginal utility of income is constant across consumers, and the revenue requirement is sufficiently small, then a notch is optimal, except in the special case where each characteristic has the same elasticity of substitution with the untaxed good.

V Conclusion

Any tax system imposing selective commodity taxation must have procedures for assigning different goods to tax rate categories, and real-world tax legislation does this on the basis of observable characteristics. Writing tax laws in terms of characteristics is intuitive and allows a tax system to handle a constantly evolving set of available goods. In this paper, we recast the theory of optimal taxation in the language of characteristics using the Gorman-Lancaster model of consumer behavior, and develop a theory of tax-driven product innovation and optimal line drawing. Such lines are associated with unavoidable discontinuities, or notches, in tax liability as a function of characteristics, and can create incentives to introduce new goods that are similar to the existing goods as well as qualitatively different goods that—in order to avoid the high tax rate—have characteristics just on the low-tax side of the line. These goods are socially inferior (but privately optimal) in characteristics space, and therefore represents a distortion in the set of available goods.

We show that, under certain conditions, when characteristics differ in their substitutability with untaxed leisure it is optimal to tolerate at least some distortion in the set of available goods. Drawing a line in characteristic space permits setting a relatively high tax rate on goods that feature most prominently the characteristic least substitutable with untaxed leisure. This reduces substitution from taxed characteristics to untaxed leisure, and thus the excess burden of taxation. Ceteris paribus, the larger is the difference in the tax base elasticity for goods lying above and below the line, the larger is the optimal difference in tax rates applying to those goods. But the larger is the difference in tax rates above and below the line, the greater is the distortion caused by the creation (and consumption) of goods that fall just on the low-tax side of the line. The optimal location of the line and the difference in tax rates above and below the line are chosen to balance these distortions and minimize the total excess burden of taxation.

The analysis is framed in terms of the classic Ramsey model of second-best efficient revenue collection, because we feel this is a natural starting point for establishing a characteristics-based theory of optimal taxation and line drawing. In practice, there are of course several other reasons for tax differentiation such as externalities, internalities, or distributional equity. The application

to externalities is of particular interest. To simplify, consider the case where there is no Ramsey reason to differentiate the tax rate on goods, but different goods embody different amounts of a negative externality whose marginal social cost is known. If the quantity of the externality can be measured, then a first-best solution can be reached by applying the appropriate Pigouvian tax to each good, and adding that to the appropriate uniform tax that raises the required revenue net of what is raised by the Pigouvian tax. Now suppose that, for example because of administrative reasons, such a Pigouvian tax cannot be levied, but only broad (say two) categories of goods can be differentiated, say cars with engines that exceed a certain displacement, and only one class can be taxed. In our model, think of the two car characteristics as being power, which is correlated with the release of pollutants and measured by engine displacement, and a composite other characteristic that includes style, comfort, etc. In general the optimal policy would be to draw a line at some engine displacement above which a "gas guzzler" tax applies, with the optimum determined by the factors in expression (11). In the event some consumers will sacrifice the otherwise-privately-optimal car model to purchase one with an engine displacement just below the notch. We would expect to see bunching of cars with an engine displacement just below this threshold, as Sallee and Slemrod (2012) demonstrate occurs in the United States under the multi-notch gas guzzler tax, and who discuss the issues that arise in applying this type of framework to environmental taxes. Whatever the motivation for tax differentiation, a tax system has to assign goods to tax rate categories by drawing lines in characteristics space, which give rise to the kind of effects we have explored in this paper. Hence, we believe that the basic insights presented here are widely applicable, regardless of the reason for tax rate differentiation.

The existing literature does not address the aspects of real-world tax systems that we explore: characteristics-based tax rules, line drawing and notches in characteristics space, and the tax-driven product innovation this generates. Although we have addressed these issues in the context of a consumption tax, they arise in all forms of taxation. For example, the Nordic-style dual income tax must differentiate—i.e., draw a line—between labor and capital income, especially with regard to self-employment income, any income tax must differentiate between debt and equity financing,

and many countries' income tax systems must distinguish between employees, on whose behalf employers must withhold and remit tax liability, and independent contractors, for whom the employer does not have these, legal responsibilities.[11] Because of the ubiquity of this issue, we hope that our analysis will illuminate how to think rigorously about this set of issues.

References

Auerbach, A. J. and Hines, J. R. (2002), Taxation and Economic Efficiency, In Auerbach, A. J. and Feldstein, M. (eds.), *Handbook of Public Economics*, Volume 3, Chapter 21, pp. 1347–1421.

Belan, P. and Gauthier, S. (2004), Optimal Commodity Grouping in a Partial Equilibrium Framework, *Economics Letters* 83(1): 49–54.

Belan, P. and Gauthier, S. (2006), Optimal Indirect Taxation With a Restricted Number of Tax Rates, *Journal of Public Economics 90*: 1201–1213.

Belan, P., Gauthier, S., and Laroque, G. (2008), Optimal Grouping of Commodities for Indirect Taxation, *Journal of Public Economics* 92(7): 1738–1750.

Blinder, A. S. and Rosen, H. S. (1985), Notches, *American Economic Review* 75(4): 736–747.

Bowers, S. (2014), Abba Admit Outrageous Outfits Were Worn to Avoid Tax, *The Guardian*, Monday 17 February. Available at: www.theguardian.com/music/2014/feb/16/abba-outfits-tax-deduction-bjorn-ulvaeus.

Corlett, W. and Hague, D. (1953-54), Complementarity and the Excess Burden of Taxation, *Review of Economic Studies* 21(1): 21–30.

Dharmapala, D., Slemrod, J., and Wilson, J. D. (2011), Tax Policy and the Missing Middle: Optimal Tax Remittance with Firm-Level Administrative Costs, *Journal of Public Economics* 95(9-10): 1036–1047.

Diamond, P. A. (1975), A Many-Person Ramsey Tax Rule, *Journal of Public Economics* 4(4): 335–342.

Gordon, J. P. F. (1989), Tax Reform via Commodity Grouping, *Journal of Public Economics* 39(1): 67–81.

Gorman, W. M. (1980), A Possible Procedure for Analysing Quality Differentiation in the Egg Market, *Review of Economic Studies* 47(5): 843–856.

[11] Johannesen (2014) discusses the use of hybrid debt instruments that allow multinational companies to take advantage of the fact that different countries draw the line in characteristic space differently.

Harberger, A. (1995), Tax Lore for Budding Reformers, In Dornbusch, R. and Edwards, S. (eds.), *Reform, Recovery, and Growth: Latin America and the Middle East*, Chapter 9, pp. 291–310. University of Chicago Press and NBER, Chicago, IL.

Johannesen, N. (2014), Tax Avoidance with Cross-Border Hybrid Instruments, *Journal of Public Economics 112*: 40–52.

Kleven, H. J. (2004), Optimum Taxation and the Allocation of Time, *Journal of Public Economics 88*: 545–557.

Kleven, H. J. and Slemrod, J. (2009), A Characteristics Approach to Optimal Taxation and Tax-Driven Product Innovation, Working Paper, University of Michigan.

Kleven, H. J. and Waseem, M. (2013), Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan, *Quarterly Journal of Economics 128*: 669–723.

Lancaster, K. J. (1966), A New Approach to Consumer Theory, *Journal of Political Economy* 74(2): 132–157.

Lancaster, K. J. (1975), Socially Optimal Product Differentiation, *American Economic Review* 65(4): 567–585.

Mirrlees, J. A. (1971), An Exploration in the Theory of Optimal Income Taxation, *Review of Economic Studies* 38(2): 175–208.

Salanié, B. (2003), *The Economics of Taxation*. Cambridge MA: MIT Press.

Sallee, J. M. and Slemrod, J. (2012), Car Notches: Strategic Automaker Responses to Fuel Economy Policy, *Journal of Public Economics* 96(11): 981–999.

Shaw, J., Slemrod, J., and Whiting, J. (2010), Administration and Compliance, In Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., and Poterba, J. (eds.), *Dimensions of Tax Design: The Mirrlees Review*, pp. 1100–1162. Oxford University Press for the Institute of Fiscal Studies, Oxford.

Slemrod, J. (2013), Buenas Notches: Lines and Notches in Tax System Design, *eJournal of Tax Research* 11(3): 259–283.

Slemrod, J. and Gillitzer, C. (2014), *Tax Systems*. MIT Press, Cambridge, MA.

Slemrod, J. and Yitzhaki, S. (2002), Tax Avoidance, Evasion, and Administration, In Auerbach, A. J. and Feldstein, M. (eds.), *Handbook of Public Economics*, Volume 3, pp. 1423–1470.

Slemrod, J., Yitzhaki, S., Mayshar, J., and Lundholm, M. (1994), The Optimal Two-Bracket Linear Income Tax, *Journal of Public Economics* 53(2): 269–290.

Weisbach, D. A. (1999), Line Drawing, Doctrine, and Efficiency in the Tax Law, *Cornell Law Review 84*: 1627–1681.

Weisbach, D. A. (2000), An Efficiency Analysis of Line Drawing in the Tax Law, *Journal of Legal Studies 29*: 1196–1206.

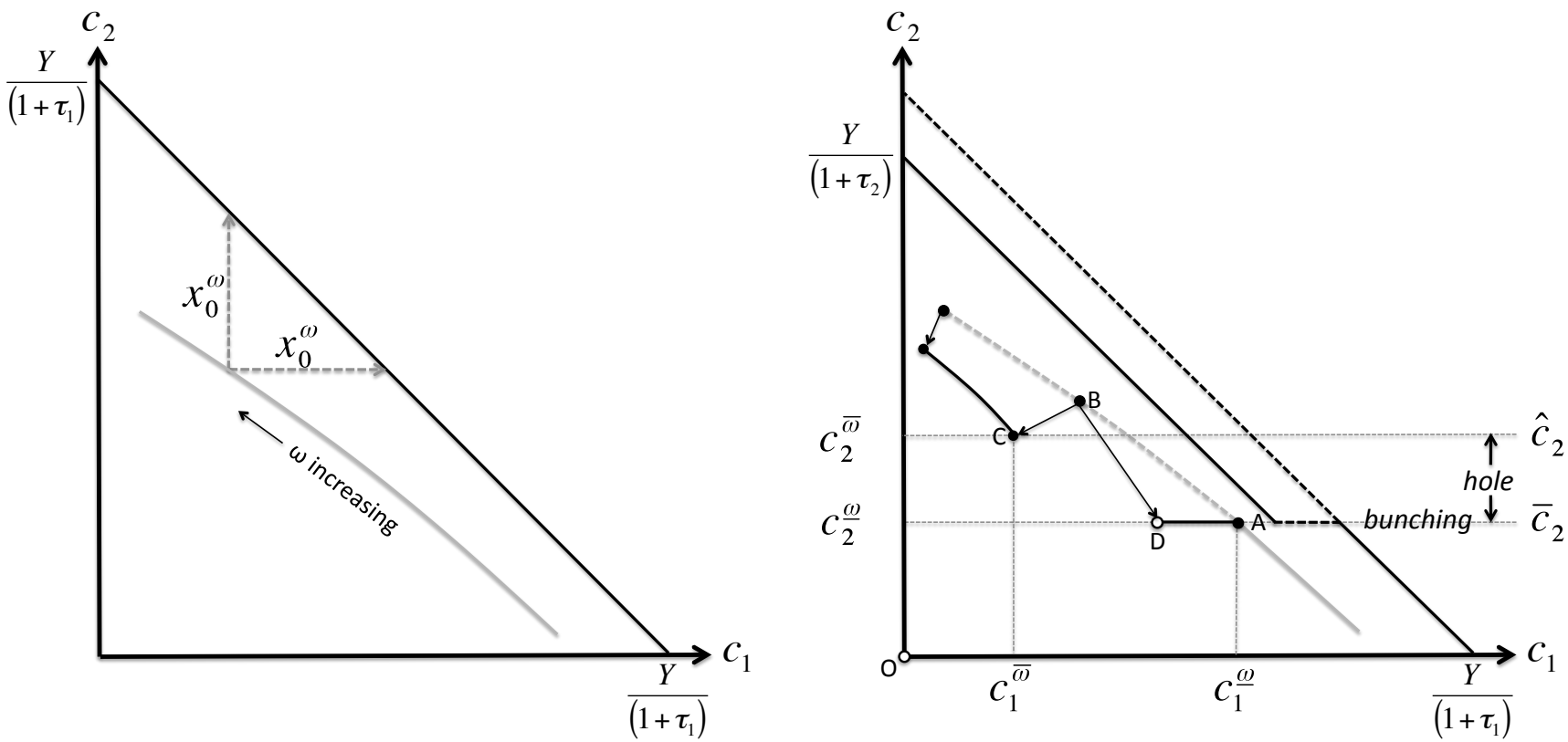
Wilson, J. D. (1989), On the Optimal Tax Base for Commodity Taxation, *American Economic Review* 79(5): 1196–1206.

Yitzhaki, S. (1979), A Note on Optimal Taxation and Administrative Costs, *American Economic Review* 69(3): 475–480.

Figure 1: The Effect of a Notched Tax System on Consumer Behavior

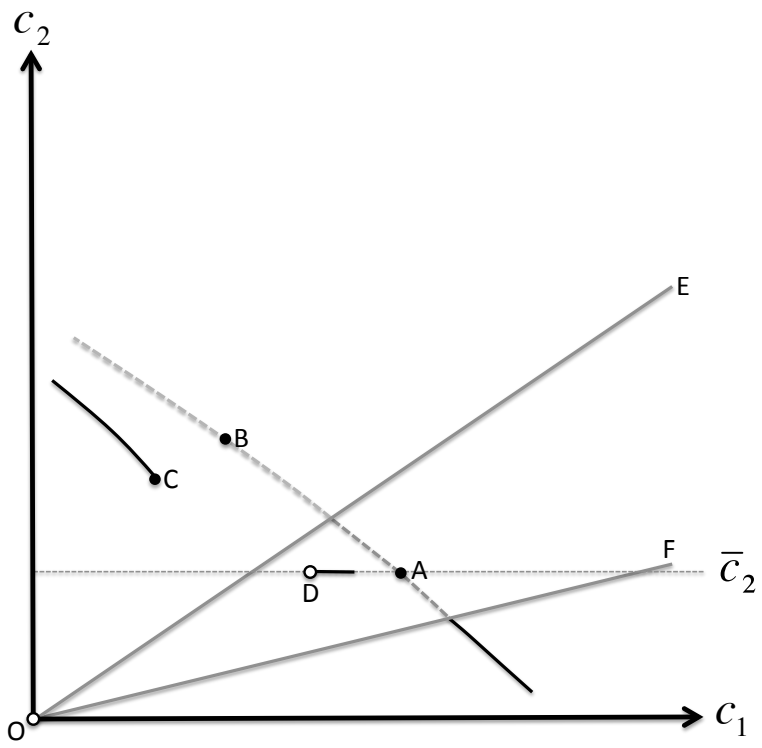
(a) Without Tax-Price Notch

(b) With Tax-Price Notch at $\bar{c}_2$: $\tau_2 > \tau_1$



Notes: The left-hand panel applies when goods corresponding to each characteristics pair are taxed at the same rate τ_1. The diagonal black line intersecting the axes indicates (c_1, c_2) pairs when the consumer spends none of their exogenous income, Y, on the numeraire good, x. The solid grey line indicates hypothetical characteristics pairs chosen by consumers; the amount spent on the numeraire good by each consumer is indicated by the vertical and horizontal distance to the diagonal black line. The right-hand panel applies when characteristic pairs (c_1, c_2) with $c_2 > \bar{c}_2$ are taxed at a rate $\tau_2 > \tau_1$: consumer types $\omega \leq \underline{\omega}$ are unaffected by the imposition of a notch at $\bar{c}_2$; consumer types $\omega > \bar{\omega}$ continue to consume a characteristics pair with $c_2 > \bar{c}_2$; and consumer types $\underline{\omega} < \omega < \bar{\omega}$ bunch at $\bar{c}_2$ and consume a characteristics pair on the horizontal line A-D. A type-$\bar{\omega}$ consumer, who chooses point B when all characteristics pairs are taxed at the same rate τ_1, is indifferent between characteristic pairs C and D following the introduction of the notch; the precise location of point D relative to point B is discussed in the main text. For simplicity, this figure assumes that producer prices q_1 and q_2 are equal to unity.

Figure 2: Tax-Driven Product Innovation



Notes: The line segments shown and the points marked correspond to those in the right-hand panel of Figure 1. Here, the distribution of consumer preferences, $F(\omega)$, is assumed to be such that, without a notch, there is no demand for goods lying on rays between O-E and O-F. Each ray through the origin represents a unique product type. The introduction of a notch at $\bar{c}_2$ generates *tax-driven product innovation*: the notched tax system induces consumers to purchase goods lying on the horizontal line segment between the rays O-E and O-F, on rays that featured no goods in the pre-notch equilibrium.

A Appendix

Proofs Omitted from the Main Text

Lemma 1 (Location of Point D in Figure 1): *Ceteris paribus, point D in Figure 1 lies further to the right the more complementary is consumption of the second characteristic with the numeraire, x, than is consumption of the first characteristic.*

Proof:

Consider the consumer who, prior to the introduction of an increased tax rate for characteristic pairs featuring $c_2 > \bar{c}_2$, chooses point B in Figure 1. Next, suppose this consumer is subject to the restriction $c_2 = \tilde{c}_2$. The maximization problem for this consumer is:

$$\max_{c_1} u\left(Y - (1+\tau_1)\left[c_1 + \tilde{c}_2\right], c_1, \tilde{c}_2\right), \tag{A.1}$$

and the associated first-order condition is

$$-(1+\tau_1)\, u_0\left(x, c_1, \tilde{c}_2\right) + u_1\left(x, c_1, \tilde{c}_2\right) = 0. \tag{A.2}$$

The restriction $c_2 = \tilde{c}_2$ is not binding when $\tilde{c}_2$ coincides with the value of c_2 indicated by point B in Figure 1. Next, beginning at point B, consider reducing the value of $\tilde{c}_2$ toward $\bar{c}_2$. By total differentiation of the first-order condition under the restriction $c_2 = \tilde{c}_2$ (expression A.2), it can be shown that, in the neighborhood of point B,

$$\frac{dc_1}{d\tilde{c}_2} = \frac{-(1+\tau_1)^2 u_{00} + (1+\tau_1)\left[u_{10} + u_{02}\right] - u_{12}}{(1+\tau_1)^2 u_{00} - 2(1+\tau_1) u_{10} + u_{11}}. \tag{A.3}$$

Provided the consumer's maximization problem (given by A.1) has a unique global maximum, the denominator on the right-hand-side of (A.3) is strictly negative. Thus, the sign of $dc_1/d\tilde{c}_2$ depends (negatively) on the sign of the numerator of the right-hand-side of (A.3). With some

further manipulation, using expression (A.2), it can be shown that

$$\begin{aligned}
\frac{dc_1}{d\tilde{c}_2} &= -\frac{1}{\mathbf{D}}\left\{\left[u_{12}-(1+\tau_1)\,u_{02}\right]+(1+\tau_1)\left[(1+\tau_1)\,u_{00}-u_{10}\right]\right\} && \text{(A.4)}\\
&= -\frac{1}{\mathbf{D}}\left\{\left[u_{21}-(1+\tau_1)\,u_{20}\right]+(1+\tau_1)\left[(1+\tau_1)\,u_{00}-u_{10}\right]\right\} && \text{(A.5)}\\
&= -\frac{1}{\mathbf{D}}\left\{\left[\frac{\partial u_2}{\partial c_1}-(1+\tau_1)\frac{\partial u_2}{\partial x}\right]+(1+\tau_1)\frac{\partial\left[(1+\tau_1)\,u_0-u_1\right]}{\partial x}\right\}, && \text{(A.6)}
\end{aligned}$$

where $\mathbf{D}\equiv(1+\tau_1)^2\,u_{00}-2\,(1+\tau_1)\,u_{10}+u_{11}<0$ is the denominator on the right-hand-side of (A.3). Hence, *ceteris paribus*, point D in Figure 1 lies further to the right the more complementary is x with the second characteristic than is consumption of the first characteristic: $(1+\tau_1)\,\partial u_2/\partial x > \partial u_2/\partial c_1$. □

Remark 1 (Notch is Suboptimal with Complete Instruments): *Assume that consumers have continuous preferences, and that the planner has a complete set of tax instruments permitting arbitrary tax liability on each of the infinitely many potential goods in characteristic space. Then the welfare-maximizing tax policy does not feature a discontinuity in tax liability (a notch).*

Proof:

With a complete set of tax instruments arbitrary tax liability can be assigned to each characteristic pair. Suppose that the discontinuity in tax liability at $\bar{c}_2$ under the notched tax system is "smoothed" using the following tax function:

$$\tau(c_2)=\begin{cases}\tau_1 & ; \quad c_2<\bar{c}_2\\ \tilde{\tau}(c_2) & ; \quad \bar{c}_2\le c_2\le\hat{c}_2\\ \tau_2 & ; \quad c_2>\hat{c}_2,\end{cases} \tag{A.7}$$

where $\tilde{\tau}(c_2)$ is a continuous and non-decreasing function, with $\tilde{\tau}(\bar{c}_2)=\tau_1$, and $\tilde{\tau}(\hat{c}_2)=\tau_2$. The tax rates τ_1 and τ_2 are the rates under the notched tax system, and $\hat{c}_2$ is the level of c_2 consumed by a type-$\bar{\omega}$ consumer (who is indifferent between consuming at $\bar{c}_2$ and $\hat{c}_2$) under the notched tax system. The dotted line in Figure A.1 depicts an example of the function $\tilde{\tau}(c_2)$.

A type-$\underline{\omega}$ consumer (whose first-order condition for c_2 is satisfied at $\bar{c}_2$) is unaffected when the function $\tau(c_2)$ is used to remove the notch: they would choose a good featuring $\bar{c}_2$ if a uniform tax rate τ_1 applied to all goods; consumer types $\omega < \underline{\omega}$, who prefer less c_2 than a type-$\underline{\omega}$ consumer at any set of post-tax prices, are similarly unaffected.

Let $(\hat{x}, \hat{c}_1, \hat{c}_2)$ denote the choice of a type-$\bar{\omega}$ consumer under the notched tax system. Smoothing the notch with the function $\tau(c_2)$ does not change the bundle chosen by a type-$\bar{\omega}$ consumer if

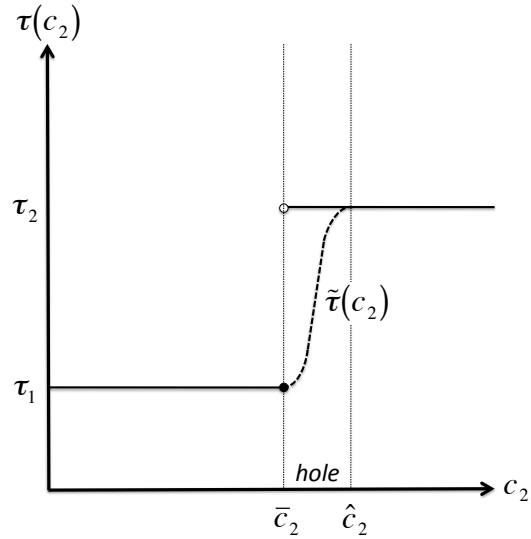
$$u(\hat{x}, \hat{c}_1, \hat{c}_2; \bar{\omega}) \geq u(Y - [1 + \tau(c_2)](c_1 + c_2), c_1, c_2; \bar{\omega}) \tag{A.8}$$

for all (c_1, c_2). We need only check the conditions on $\tilde{\tau}(c_2)$ required for (A.8) to hold because bundles featuring $c_2 < \bar{c}_2$ and $c_2 > \hat{c}_2$ were available at the same post-tax prices under the notched tax system and not chosen. Let $c_1(c_2; \tau, \bar{\omega})$ represent the optimum choice of c_1 as a function of c_2 for a type-$\bar{\omega}$ consumer, given the tax function $\tau(c_2)$. Differentiation of (A.8) with respect to c_2 indicates that a type-$\bar{\omega}$ consumer chooses the same good (c_1, c_2) under the notched tax system and under the tax system $\tau(c_2)$ if the function $\tilde{\tau}(c_2)$ is sufficiently steep:

$$\tilde{\tau}'(c_2) \geq \left(\frac{1}{c_1 + c_2}\right)\left[\frac{u_1^{\bar{\omega}}}{u_0^{\bar{\omega}}} c_1' + \frac{u_2^{\bar{\omega}}}{u_0^{\bar{\omega}}} - [1 + \tau]\left(c_1' + c_2\right)\right]. \tag{A.9}$$

Any $\tau(c_2)$ satisfying (A.9) leaves consumer choices below $\bar{c}_2$ and above $\hat{c}_2$ unchanged. The smallest tax-rate schedule $\tau(c_2)$ satisfying (A.9) induces bunching consumers to purchase goods featuring $c_2 \varepsilon(\bar{c}_2, \hat{c}_2)$. Goods at the notch point are still available at an unchanged tax rate so, by revealed preference, bunching consumers who move into the hole are better off under the smoothed tax function. Because tax rates in the hole are larger than at the notch point, tax revenue collected on previously bunching consumers rises. Thus, welfare is strictly higher when the hole is removed. □

Figure A.1: Smoothing the Difference in Tax Rates over the Range of the Hole



Proposition 1 (Notch is Optimal with Incomplete Instruments): *Assume that the social planner has a restricted set of tax instruments that only permits setting a different tax rate on goods featuring c_2 greater or less than some arbitrary level $\bar{c}_2$. In this case the optimal tax policy features a tax-price notch in characteristic space (except in a knife-edge case). In the special case of no equity concerns (constant social marginal utilities of income across consumers), the optimal tax policy features a notch whenever the tax base elasticity above and below $\bar{c}_2$ differs.*

Proof:

The value of the tax base below the line $\bar{c}_2$ is

$$\begin{aligned} Z_1 &\equiv \int_{\omega^{min}}^{\bar{\omega}} z^{\omega} dF(\omega) \\ &\equiv \int_{\omega^{min}}^{\bar{\omega}} q_1 c_1^{\omega} + q_2 c_2^{\omega} dF(\omega), \end{aligned} \tag{A.10}$$

and above the line $\bar{c}_2$ is

$$\begin{aligned} Z_2 &\equiv \int_{\bar{\omega}}^{\omega^{max}} z^{\omega} dF(\omega) \\ &\equiv \int_{\bar{\omega}}^{\omega^{max}} q_1 c_1^{\omega} + q_2 c_2^{\omega} dF(\omega). \end{aligned} \tag{A.11}$$

The tax base elasticity below the line $\bar{c}_2$, with respect to a marginal change in τ_1, is a weighted average elasticity for each consumer below $\bar{c}_2$:

$$\begin{aligned} \varepsilon(Z_1, \tau_1) &\equiv \int_{\omega^{min}}^{\bar{\omega}} \left(\frac{z^{\omega}}{Z_1}\right) \varepsilon(z^{\omega}, \tau_1) dF(\omega) \\ &= \int_{\omega^{min}}^{\bar{\omega}} \left(\frac{z^{\omega}}{Z_1}\right) \left(\frac{\tau_1}{z^{\omega}} \frac{\partial z^{\omega}}{\partial \tau_1}\right) dF(\omega) \\ &= \int_{\omega^{min}}^{\bar{\omega}} \frac{\tau_1}{Z_1} \frac{\partial z^{\omega}}{\partial \tau_1} dF(\omega). \end{aligned}$$

Similarly, the tax base elasticity above the line $\bar{c}_2$, with respect to a marginal change in τ_2, is

$$\begin{aligned} \varepsilon(Z_2, \tau_2) &\equiv \int_{\bar{\omega}}^{\omega^{max}} \left(\frac{z^{\omega}}{Z_2}\right) \varepsilon(z^{\omega}, \tau_2) dF(\omega) \\ &= \int_{\bar{\omega}}^{\omega^{max}} \left(\frac{z^{\omega}}{Z_2}\right) \left(\frac{\tau_2}{z^{\omega}} \frac{\partial z^{\omega}}{\partial \tau_2}\right) dF(\omega) \\ &= \int_{\bar{\omega}}^{\omega^{max}} \frac{\tau_2}{Z_2} \frac{\partial z^{\omega}}{\partial \tau_2} dF(\omega). \end{aligned}$$

Assume that the combined revenue raised across the two tax bases satisfies the exogenous revenue requirement:

$$G = \tau_1 Z_1 + \tau_2 Z_2. \tag{A.12}$$

For $\tau_1 = \tau_2$ (without a notch), total differentiation of this condition gives

$$\begin{aligned} dG &= \left[Z_1 + \tau_1 \frac{\partial Z_1}{\partial \tau_1} + \tau_2 \frac{\partial Z_2}{\partial \tau_1}\right] d\tau_1 + \left[Z_2 + \tau_1 \frac{\partial Z_1}{\partial \tau_2} + \tau_2 \frac{\partial Z_2}{\partial \tau_2}\right] d\tau_2 && \text{(A.13)} \\ &= [1+\varepsilon(Z_1,\tau_1)] Z_1 d\tau_1 + [1+\varepsilon(Z_2,\tau_2)] Z_2 d\tau_2 && \text{(A.14)} \\ &\quad + (\tau_1 - \tau_2) z^{\bar{\omega}} f(\bar{\omega}) \left[\frac{\partial \bar{\omega}}{\partial \tau_1} d\tau_1 + \frac{\partial \bar{\omega}}{\partial \tau_2} d\tau_2\right] \\ &= [1+\varepsilon(Z_1,\tau_1)] Z_1 d\tau_1 + [1+\varepsilon(Z_2,\tau_2)] Z_2 d\tau_2. && \text{(A.15)} \end{aligned}$$

Beginning at $\tau_1 = \tau_2$, suppose the tax rates τ_1 and τ_2 are perturbed to keep revenue raised unchanged: $dG = 0$ and

$$d\tau_2 = -\left[\frac{1+\varepsilon(Z_1,\tau_1)}{1+\varepsilon(Z_2,\tau_2)}\right]\left[\frac{Z_1}{Z_2}\right] d\tau_1. \tag{A.16}$$

The envelope theorem implies that the welfare effect for consumers of a small tax rate perturbation is equal to the marginal mechanical burden (i.e., the behavioral response to the change in tax rates has no effect on welfare). Hence, the change in social welfare from a revenue-neutral tax rate perturbation is

$$\begin{aligned} \Delta W &= \int_{\omega^{min}}^{\bar{\omega}} \gamma^{\omega} z^{\omega} d\tau_1 dF(\omega) + \int_{\bar{\omega}}^{\omega^{max}} \gamma^{\omega} z^{\omega} d\tau_2 dF(\omega) && \text{(A.17)} \\ &= d\tau_1 \left[\int_{\omega^{min}}^{\bar{\omega}} \gamma^{\omega} z^{\omega} dF(\omega) - \left(\frac{1+\varepsilon(Z_1,\tau_1)}{1+\varepsilon(Z_2,\tau_2)}\right)\left(\frac{Z_1}{Z_2}\right) \int_{\bar{\omega}}^{\omega^{max}} \gamma^{\omega} z^{\omega} dF(\omega)\right], && \text{(A.18)} \end{aligned}$$

where γ^{ω} is the social marginal utility of income for a type-ω consumer. Unless the bracketed expression on the right-hand-side of (A.18) is equal to zero, at least some arbitrarily small revenue-neutral differentiation in tax rates above and below $\bar{c}_2$ is optimal: if the expression in brackets is negative then an optimum features $\tau_2 > \tau_1$, and vice-versa. For the special case in which the social marginal utility of income is the same for all consumers, $\gamma^{\omega} = \gamma$, the introduction of a notch has

pure efficiency consequences; expression (A.18) can be re-expressed as follows:

$$\begin{aligned}
\Delta W &= \gamma\left[\int_{\omega^{min}}^{\bar{\omega}} z^{\omega} dF(\omega) - \left(\frac{1+\varepsilon(Z_1,\tau_1)}{1+\varepsilon(Z_2,\tau_2)}\right)\left(\frac{Z_1}{Z_2}\right)\int_{\bar{\omega}}^{\omega^{max}} z^{\omega} dF(\omega)\right] d\tau_1 && \text{(A.19)}\\
&= \gamma\left[Z_1 - \left(\frac{1+\varepsilon(Z_1,\tau_1)}{1+\varepsilon(Z_2,\tau_2)}\right)\left(\frac{Z_1}{Z_2}\right) Z_2\right] d\tau_1 && \text{(A.20)}\\
&= \gamma\left[1 - \left(\frac{1+\varepsilon(Z_1,\tau_1)}{1+\varepsilon(Z_2,\tau_2)}\right)\right] Z_1 d\tau_1 && \text{(A.21)}\\
&= \gamma\left[\left(\frac{\varepsilon(Z_2,\tau_2)-\varepsilon(Z_1,\tau_1)}{1+\varepsilon(Z_2,\tau_2)}\right)\right] Z_1 d\tau_1. && \text{(A.22)}
\end{aligned}$$

When $\gamma^{\omega} = \gamma$, $\varepsilon(Z_2,\tau_2) \neq \varepsilon(Z_1,\tau_1)$ is a sufficient condition for a notch to be optimal. □

Proposition 2 (Optimal Tax Rates): *Assuming the social marginal utility of income, γ, is constant across consumers, the optimal tax policy satisfies the following Corlett-Hague type rule for any given height of the tax-price line $\bar{c}_2$:*

$$\frac{\tau_1/(1+\tau_1)}{\tau_2/(1+\tau_2)} = \Gamma \frac{1/[1-F(\bar{\omega})]\int_{\bar{\omega}}^{\omega^{max}} (z^{\omega}/z^{u})\left[\theta_1^{\omega}\varepsilon_{10}^{c} + \theta_2^{\omega}\varepsilon_{20}^{c}\right] dF(\omega)}{1/F(\bar{\omega})\int_{\omega^{min}}^{\bar{\omega}} \left(z^{\omega}/z^{l}\right)\left[\theta_1^{\omega}\varepsilon_{10}^{c} + \theta_2^{\omega}\varepsilon_{20}^{c}\right] dF(\omega)}, \qquad \text{(A.23)}$$

where ε_{j0}^{c} is the compensated demand elasticity between characteristic j and the numeraire, assumed to be the same for all consumers, $\theta_j^{\omega} \equiv \left(q_j c_j^{\omega}\right)/z^{\omega}$ is the pre-tax cost share of characteristic j in the good bought by a type-ω consumer, with $\sum_k \theta_k^{\omega} = 1$, while $z^{u} \equiv 1/[1-F(\bar{\omega})]\int_{\bar{\omega}}^{\omega^{max}} z^{\omega} dF(\omega)$ and $z^{l} \equiv 1/F(\bar{\omega})\int_{\omega^{min}}^{\bar{\omega}} z^{\omega} dF(\omega)$ are the mean (pre-tax) value of expenditure on the characteristics good by consumers who purchase goods above and below the line $\bar{c}_2$, respectively. Finally,

$$\Gamma \equiv \frac{(1-\gamma/\lambda) - \left[\Delta R(\bar{\omega})/z^{l}\right]\left[f(\bar{\omega})/F(\bar{\omega})\right]\left[\partial\bar{\omega}/\partial\tau_1\right]}{(1-\gamma/\lambda) - \left[\Delta R(\bar{\omega})/z^{u}\right]\left[f(\bar{\omega})/(1-F(\bar{\omega}))\right]\left[\partial\bar{\omega}/\partial\tau_2\right]} \qquad \text{(A.24)}$$

measures the social value of revenue spillovers across tax bases (the regions above and below $\bar{c}_2$), with $\Delta R(\bar{\omega}) \equiv \left(\tau_2 z\left(c_1^{\bar{\omega}}, c_2^{\bar{\omega}}\right) - \tau_1 z\left(c_1^{\bar{\omega}}(\bar{c}_2), \bar{c}_2\right)\right)$ being the change in revenue when a type-$\bar{\omega}$ consumer moves from bunching at $\bar{c}_2$ to the high-tax region.

Proof:

The Lagrangian for the planner's problem is

$$L = \int_{\omega^{min}}^{\omega^{max}} W\left[v^{\omega}\left(\tau_1, \tau_2, \bar{c}_2\right)\right] dF\left(\omega\right) + \lambda \left[\tau_1 \int_{\omega^{min}}^{\underline{\omega}} z\left(c_1^{\omega}, c_2^{\omega}\right) dF\left(\omega\right) + \tau_1 \int_{\underline{\omega}}^{\bar{\omega}} z\left(c_1^{\omega}, \bar{c}_2\right) dF\left(\omega\right) + \tau_2 \int_{\bar{\omega}}^{\omega^{max}} z\left(c_1^{\omega}, c_2^{\omega}\right) dF\left(\omega\right) - G\right], \tag{A.25}$$

where λ is the Lagrange multiplier on the aggregate budget constraint, and it is assumed that consumer preferences are such that the optimal tax policy features $\tau_2 \geq \tau_1$. The tax-rate first-order conditions for the social planner's maximization problem (A.25) are

$$\frac{\partial L}{\partial \tau_1} = \int_{\omega^{min}}^{\bar{\omega}} -W' \alpha^{\omega} z^{\omega} dF\left(\omega\right) + \lambda \left[\int_{\omega^{min}}^{\bar{\omega}} z^{\omega} dF\left(\omega\right) + \tau_1 \int_{\omega^{min}}^{\bar{\omega}} q_1 \frac{\partial c_1^{\omega}}{\partial \tau_1} + q_2 \frac{\partial c_2^{\omega}}{\partial \tau_1} dF\left(\omega\right) - \Delta R\left(\bar{\omega}\right) f\left(\bar{\omega}\right) \frac{\partial \bar{\omega}}{\partial \tau_1}\right] = 0, \tag{A.26}$$

and

$$\frac{\partial L}{\partial \tau_2} = \int_{\bar{\omega}}^{\omega^{max}} -W' \alpha^{\omega} z^{\omega} dF\left(\omega\right) + \lambda \left[\int_{\bar{\omega}}^{\omega^{max}} z^{\omega} dF\left(\omega\right) + \tau_2 \int_{\bar{\omega}}^{\omega^{max}} q_1 \frac{\partial c_1^{\omega}}{\partial \tau_1} + q_2 \frac{\partial c_2^{\omega}}{\partial \tau_1} dF\left(\omega\right) - \Delta R\left(\bar{\omega}\right) f\left(\bar{\omega}\right) \frac{\partial \bar{\omega}}{\partial \tau_2}\right] = 0, \tag{A.27}$$

where $c_2^{\omega} = \bar{c}_2$ for taxpayers of type $\omega \varepsilon \left[\underline{\omega}, \overline{\omega}\right]$, and $\Delta R\left(\bar{\omega}\right) \equiv \left(\tau_2 z\left(c_1^{\bar{\omega}}, c_2^{\bar{\omega}}\right) - \tau_1 z\left(c_1^{\bar{\omega}}\left(\bar{c}_2\right), \bar{c}_2\right)\right)$ is the change in revenue when a type-$\bar{\omega}$ consumer moves from bunching at $\bar{c}_2$ to the high-tax region. The steps that follow detail intermediate steps used to re-express these two first-order conditions.

First, note that the marginal change in a type-ω consumer's implicit demand for characteristic j, with respect to a marginal change in the tax rate τ applying to the composite good, can be expressed as

$$\frac{\partial c_j^{\omega}}{\partial \tau} = \sum_k \frac{\partial c_j^{\omega}}{\partial p_k} \frac{dp_k}{d\tau} \tag{A.28}$$

$$= \sum_k \frac{\partial c_j^{\omega}}{\partial p_k} q_k, \tag{A.29}$$

recalling that $p_k = (1+\tau)\, q_k$. Making use of the Slutsky formula,

$$\frac{\partial c_j^\omega}{\partial \tau} = \sum_k \left(s_{jk}^\omega - c_k^\omega \frac{\partial c_j^\omega}{\partial Y} \right) q_k, \tag{A.30}$$

where s_{jk}^ω is the Slutsky (compensated) substitution effect for a type-ω consumer. Hence,

$$q_j \frac{\partial c_j^\omega}{\partial \tau} = q_j \sum_k \left(s_{jk}^\omega - c_k^\omega \frac{\partial c_j^\omega}{\partial Y} \right) q_k \tag{A.31}$$

$$= q_j \sum_k \left(\frac{c_j^\omega}{p_k} \varepsilon_{jk}^{\omega,c} - c_k^\omega \frac{\partial c_j^\omega}{\partial Y} \right) q_k \tag{A.32}$$

$$= q_j \sum_k \left(\frac{1}{1+\tau} \frac{c_j^\omega}{q_k} \varepsilon_{jk}^{\omega,c} - c_k^\omega \frac{\partial c_j^\omega}{\partial Y} \right) q_k, \tag{A.33}$$

where $\varepsilon_{jk}^{\omega,c} \equiv \left(p_k / c_j^\omega \right) s_{jk}^\omega$ is a type-ω consumer's compensated implicit demand elasticity for characteristic j with respect to a marginal change in the price of characteristic k. By summation, it can be shown that

$$\sum_j q_j \frac{\partial c_j^\omega}{\partial \tau} = z^\omega \left[\frac{1}{1+\tau} \sum_j \theta_j^\omega \left(\varepsilon_{j1}^{\omega,c} + \cdots + \varepsilon_{jN}^{\omega,c} \right) - \sum_j q_j \frac{\partial c_j^\omega}{\partial Y} \right], \tag{A.34}$$

where $\theta_j^\omega \equiv \left(q_j c_j^\omega \right) / z^\omega$ is the pre-tax cost share of characteristic j in the good bought by a type-ω consumer.

Next, following Diamond (1975), define

$$\gamma^\omega \equiv W' \alpha^\omega + \lambda \tau^\omega \sum_j q_j \frac{\partial c_j^\omega}{\partial Y} \tag{A.35}$$

to be the social marginal utility of income for a type-ω consumer, where τ^ω is the tax rate paid by a type-ω consumer. The first term is the gain to society from a dollar of increased consumption for a type-ω consumer, and the second term is the social value of a type-ω consumer's marginal propensity to pay taxes out of additional income. Substitution of (A.34) and (A.35) into (A.26) and (A.27), and some re-arrangement, results in the following representation for the planner's tax rate

first-order conditions:

$$\int_{\omega^{min}}^{\bar{\omega}} z^{\omega}\left[\left(\frac{\gamma^{\omega}}{\lambda}-1\right)-\left(\frac{\tau_1}{1+\tau_1}\right)\left(\theta_1^{\omega}\left[\varepsilon_{11}^{\omega,c}+\varepsilon_{12}^{\omega,c}\right]+\theta_2^{\omega}\left[\varepsilon_{21}^{\omega,c}+\varepsilon_{22}^{\omega,c}\right]\right)\right]dF(\omega)+\Delta R(\bar{\omega})f(\bar{\omega})\frac{\partial\bar{\omega}}{\partial\tau_1}=0 \tag{A.36}$$

and

$$\int_{\bar{\omega}}^{\omega^{max}} z^{\omega}\left[\left(\frac{\gamma^{\omega}}{\lambda}-1\right)-\left(\frac{\tau_2}{1+\tau_2}\right)\left(\theta_1^{\omega}\left[\varepsilon_{11}^{\omega,c}+\varepsilon_{12}^{\omega,c}\right]+\theta_2^{\omega}\left[\varepsilon_{21}^{\omega,c}+\varepsilon_{22}^{\omega,c}\right]\right)\right]dF(\omega)+\Delta R(\bar{\omega})f(\bar{\omega})\frac{\partial\bar{\omega}}{\partial\tau_2}=0, \tag{A.37}$$

where $\theta_j^{\omega}\equiv\left(q_j c_j^{\omega}\right)/z^{\omega}$ is the pre-tax cost share of characteristic j in the good bought by a type-ω consumer, with $\sum_k \theta_k^{\omega}=1$. The homogeneity of degree zero in prices of compensated demands for good x or any characteristic c_j implies that $\sum_{n=0}^{N}\varepsilon_{jn}^{\omega,c}=0$. We assume that, for any tax system policy vector under consideration, the social planner assigns welfare weights such that the social marginal utility of income is constant across consumers: $\gamma^{\omega}=\gamma$. Using this assumption, and assuming that $\varepsilon_{j0}^{\omega,c}=\varepsilon_{j0}^{c}$, expressions (A.36) and (A.37) can be combined to give (9), as desired. □

Proposition 3 (Optimal Placement of Line/Notch): *Without loss of generality, assume that an optimum features $\tau_2>\tau_1$. At an optimum, the height of the line $\bar{c}_2$ satisfies the following condition:*

$$\int_{\underline{\omega}}^{\bar{\omega}}\frac{W'}{\lambda}\left(u_2^{\omega}-\alpha^{\omega}p_2\right)dF(\omega)=\Delta R(\bar{\omega})f(\bar{\omega})\frac{\partial\bar{\omega}}{\partial\bar{c}_2}-\int_{\underline{\omega}}^{\bar{\omega}}\tau_1\left(q_1\frac{\partial c_1^{\omega}}{\partial\bar{c}_2}+q_2\right)dF(\omega). \tag{A.38}$$

Proof:

The first-order condition for the planner's problem (A.25) with respect to $\bar{c}_2$ is

$$\begin{aligned}\frac{\partial L}{\partial\bar{c}_2}=\int_{\underline{\omega}}^{\bar{\omega}}W'\left(u_2^{\omega}-\alpha^{\omega}(1+\tau_1)q_2\right)dF(\omega)+\lambda\Bigg[&\tau_1\int_{\underline{\omega}}^{\bar{\omega}}\left(q_1\frac{\partial c_1^{\omega}}{\partial\bar{c}_2}+q_2\right)dF(\omega)\\&-\Delta R(\bar{\omega})f(\bar{\omega})\frac{\partial\bar{\omega}}{\partial\bar{c}_2}\Bigg]=0,\end{aligned} \tag{A.39}$$

where $\Delta R(\bar{\omega})\equiv\left(\tau_2 z\left(c_1^{\bar{\omega}},c_2^{\bar{\omega}}\right)-\tau_1 z\left(c_1^{\bar{\omega}}(\bar{c}_2),\bar{c}_2\right)\right)>0$ is the gain in revenue when a type-$\bar{\omega}$ con-

sumer moves from bunching at $\bar{c}_2$ to consuming at $\hat{c}_2$. After some rearrangement (A.39) can be shown to be the same as (11). $\square$

Online Appendix to:

A Characteristics Approach to Optimal Taxation: Line Drawing and Tax-Driven Product Innovation

Christian Gillitzer, Reserve Bank of Australia

Henrik Jacobsen Kleven, London School of Economics

Joel Slemrod, University of Michigan

August 2015

Nested Constant Elasticity of Substitution Preferences

In this appendix we provide a specific example, where each consumer has a utility function of the nested constant elasticity of substitution (CES) form, as follows:

$$u(x, c_1, c_2; \omega) = \left[a_1(\omega) c_1^{\frac{\sigma-1}{\sigma}} + \left(a_0 x^{\frac{\varepsilon-1}{\varepsilon}} + a_2(\omega) c_2^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\left(\frac{\varepsilon}{\varepsilon-1}\right)\left(\frac{\sigma-1}{\sigma}\right)} \right]^{\frac{\sigma}{\sigma-1}}. \tag{1}$$

In expression (1), σ is the elasticity of substitution between characteristic one and both characteristic two and the untaxed numeraire x, and ε is the elasticity of substitution between characteristic two and the untaxed good. It is assumed that $a_2(\omega)$ is increasing in ω and that $a_1(\omega)$ is decreasing in ω, which implies that, at any set of prices, the higher is ω the greater is the consumer's implicit demand for c_2 and the lower is the consumer's implicit demand for c_1.

It is informative to consider a special case in which, at pre-tax prices, each consumer spends the same share of their income on the characteristics good (and thus also on the untaxed good). Consumers differ in the type of characteristics good they prefer, but not in their share of income spent on the characteristics good; as $a_2(\omega)$ rises, $a_1(\omega)$ declines by enough to keep expenditure at pre-tax prices on the characteristics good

constant. When the goods above and below $\bar{c}_2$ differ in their substitutability with the untaxed numeraire, it is optimal with preferences given by (1) to introduce a notch. (For analytical convenience, we maintain the assumption that the social marginal utility of income is constant across consumers, but our qualitative conclusions are unaffected by this assumption.)

Proposition 4 (CES Preferences): *If consumer preferences are given by (1), the expenditure share on the untaxed good does not vary across consumers in the absence of taxes, the social marginal utility of income is constant across consumers, and the revenue requirement is sufficiently small, then a notch is optimal for* $\sigma \neq \varepsilon$.

Proof:

Consumers have constant elasticity of substitution (CES) preferences, with a type-ω consumer having utility given by

$$u(x, c_1, c_2; \omega) = \left[a_1(\omega) c_1^{\frac{\sigma-1}{\sigma}} + \left(a_0 x^{\frac{\varepsilon-1}{\varepsilon}} + a_2(\omega) c_2^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\left(\frac{\varepsilon}{\varepsilon-1}\right)\left(\frac{\sigma-1}{\sigma}\right)} \right]^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where σ is the elasticity of substitution between characteristic one and both characteristic two and the untaxed good x, and ε is the elasticity of substitution between characteristic two and the untaxed good. The expenditure function for a type-ω consumer is

$$e(p,u) = \frac{u\left[a_0^{\varepsilon} p_0^{1-\varepsilon} + a_1^{\sigma} p_1^{1-\sigma} \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right)^{\left(\frac{\sigma-\varepsilon}{1-\varepsilon}\right)} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right]}{\left[a_1^{\sigma} p_1^{1-\sigma} \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right)^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\varepsilon-\sigma}{1-\varepsilon}\right)} + \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right)^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\varepsilon}{1-\varepsilon}\right)} \right]^{\left(\frac{\sigma}{\sigma-1}\right)}}, \tag{3}$$

and the associated Hicksian compensated demand functions are

$$x(p,u) = \frac{u a_0^{\varepsilon} p_0^{-\varepsilon}}{\left[a_1^{\sigma} p_1^{1-\sigma} \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right)^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\varepsilon-\sigma}{1-\varepsilon}\right)} + \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right)^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\varepsilon}{1-\varepsilon}\right)} \right]^{\left(\frac{\sigma}{\sigma-1}\right)}}, \tag{4}$$

$$c_1(p,u) = \frac{u a_1^{\sigma} p_1^{-\sigma} \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right)^{\left(\frac{\sigma-\varepsilon}{1-\varepsilon}\right)}}{\left[a_1^{\sigma} p_1^{1-\sigma} \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right)^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\varepsilon-\sigma}{1-\varepsilon}\right)} + \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon} \right)^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\varepsilon}{1-\varepsilon}\right)} \right]^{\left(\frac{\sigma}{\sigma-1}\right)}}, \tag{5}$$

and

$$c_2(p,u) = \frac{u a_2^{\varepsilon} p_2^{-\varepsilon}}{\left[a_1^{\sigma} p_1^{1-\sigma}\left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon}\right)^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\varepsilon-\sigma}{1-\varepsilon}\right)} + \left(a_0^{\varepsilon} p_0^{1-\varepsilon} + a_2^{\varepsilon} p_2^{1-\varepsilon}\right)^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\varepsilon}{1-\varepsilon}\right)}\right]^{\left(\frac{\sigma}{\sigma-1}\right)}}. \tag{6}$$

By differentiation of (5) and (6), the compensated cross-price demand elasticity for each characteristic with respect to a change in the price of the untaxed good can be shown to be

$$\varepsilon_{10}^{c,\omega} = \frac{\partial c_1^{\omega}(p,u)}{\partial p_0}\frac{p_0}{c_1^{\omega}} = \sigma s_0^{\omega}, \tag{7}$$

and

$$\varepsilon_{20}^{c,\omega} = \frac{\partial c_2^{\omega}(p,u)}{\partial p_0}\frac{p_0}{c_2^{\omega}} = \sigma s_0^{\omega} + (\varepsilon - \sigma)\left(\frac{s_0^{\omega}}{s_0^{\omega} + s_2^{\omega}}\right), \tag{8}$$

where $s_0^{\omega} \equiv (p_0 x^{\omega})/Y$ is the expenditure share for the untaxed good (for a type-ω consumer), and $s_j^{\omega} \equiv \left(p_j c_j^{\omega}\right)/Y$ is the expenditure share on characteristic j. Hence, recalling that we have defined $\theta_1^{\omega} = s_1^{\omega}/(s_1^{\omega} + s_2^{\omega})$ to be the expenditure share of the first characteristic in the composite good purchased by the consumer, and similarly $\theta_2^{\omega} = s_2^{\omega}/(s_1^{\omega} + s_2^{\omega})$ for the second characteristic, by combination of (7) and (8) it can be shown that

$$\theta_1^{\omega}\varepsilon_{10}^{c,\omega} + \theta_2^{\omega}\varepsilon_{20}^{c,\omega} = \sigma s_0^{\omega} + (\varepsilon - \sigma)\left(\frac{s_0^{\omega}}{1 - s_0^{\omega}}\right)\left(\frac{s_2^{\omega}}{s_0^{\omega} + s_2^{\omega}}\right). \tag{9}$$

Next, suppose $\tau_1 = \tau_2$, and that the social marginal utility of income, γ, is constant across consumers. Then (A. 36) and (A. 37) in the main text take the form

$$\frac{1}{F(\bar{\omega})}\int_{\omega^{min}}^{\bar{\omega}} Y(1 - s_0^{\omega})\left[-\left(1 - \frac{\gamma}{\lambda}\right) + \left(\frac{\tau_1}{1+\tau_1}\right)\left(\sigma s_0^{\omega} + (\varepsilon - \sigma)\left(\frac{s_0^{\omega}}{1 - s_0^{\omega}}\right)\left(\frac{s_2^{\omega}}{s_0^{\omega} + s_2^{\omega}}\right)\right)\right] dF(\omega) = 0 \tag{10}$$

and

$$\frac{1}{1 - F(\bar{\omega})}\int_{\bar{\omega}}^{\omega^{max}} Y(1 - s_0^{\omega})\left[-\left(1 - \frac{\gamma}{\lambda}\right) + \left(\frac{\tau_2}{1+\tau_2}\right)\left(\sigma s_0^{\omega} + (\varepsilon - \sigma)\left(\frac{s_0^{\omega}}{1 - s_0^{\omega}}\right)\left(\frac{s_2^{\omega}}{s_0^{\omega} + s_2^{\omega}}\right)\right)\right] dF(\omega) = 0. \tag{11}$$

Because s_2^{ω} is increasing in ω, and for a sufficiently small revenue requirement s_0^{ω} is almost constant across consumers, $\sigma s_0^{\omega} + (\varepsilon - \sigma)\left(s_0^{\omega}/1 - s_0^{\omega}\right)\left(s_2^{\omega}/s_0^{\omega} + s_2^{\omega}\right)$ is monotonically increasing or decreasing in ω if $\sigma \neq \varepsilon$. For any location of the line that has at least some consumers purchasing above and below $\bar{c}_2$, (10) and (11) cannot be simultaneously satisfied by $\tau_1 = \tau_2$. Hence, $\tau_1 \neq \tau_2$ and a tax-price notch is optimal. □

If $\sigma > \varepsilon$, in which case goods above $\bar{c}_2$ (which feature the second characteristic prominently) are less substitutable with the numeraire than are goods below $\bar{c}_2$ (which feature the first characteristic prominently), then $\tau_2 > \tau_1$, and vice versa. The social planner is willing to tolerate at least some distortion from bunching behavior to mitigate the distortion caused by substitution from taxed characteristics to the untaxed good.