Optimal Tax Salience

Jacob Goldin*

March 2012

Abstract

Recent empirical work suggests that how an individual responds to a tax depends at least in part on the tax’s salience. The more salient a tax is, the more taxpayers adjust their demand in response to changes in the taxed good’s after-tax price. If tax salience affects behavior, a natural question follows: How salient should a government’s revenue collection system be? I investigate this question by considering the problem faced by a benevolent government choosing between high- and low-salience commodity taxes to meet a revenue constraint. I show that low-salience taxes introduce two offsetting welfare effects: on the one hand, they reduce the excess burden traditionally associated with distortionary taxation by dampening consumers’ substitution away from the taxed good; on the other hand, low-salience taxes introduce new welfare costs by causing consumers to make optimization errors when deciding how much of each good to purchase. Under certain conditions, I show that governments can utilize a combination of high- and low-salience commodity taxes to achieve the first-best welfare outcome, even without employing a lump-sum tax. I also derive a simple and intuitive formula that characterizes the optimal mix between high- and low-salience taxes needed to obtain this outcome. Under the optimal policy, the low-salience tax is strictly non-zero, and the ratio of low- to high-salience taxes is 1) increasing in the compensated own-price elasticity of demand for the taxed good, 2) decreasing in the income-sensitivity of the taxed good, and 3) decreasing in the taxed good’s share of the budget. Finally, high-salience taxes tend to be efficient when consumption of the taxed good generates negative externalities.

1. Introduction

The subject of optimal commodity taxation is at the heart of public finance. Traditional approaches to the issue attempt to determine the most efficient way to levy taxes across goods when lump-sum taxes are not available. For example, the canonical Ramsey-Boiteux formula provides useful insights for choosing which

---

*Princeton University and Yale Law School. For helpful comments, I am grateful to Raj Chetty, Mikhail Golosov, Jonah Gellbach, Nikolaj Harmon, Tatiana Hononoff, Ilyana Kuziemko, Yair Listokin, David Lee, Alex Mas, Harvey Rosen, and Dean Spears, as well as participants in the Prospectus Workshop in Labor Economics at Yale and the Public Finance Working Group at Princeton. All errors are my own.
goods to tax and how much to tax them. In contrast, questions of tax design have traditionally been perceived as second-order and have not received the same level of theoretical attention.

Recent empirical findings suggest a need to reconsider this emphasis. A series of papers suggest that the design of a tax—in particular the tax’s salience—has important effects on consumer behavior: the more prominent the after-tax price of a good, the more consumers respond to changes in the tax rate on that good.¹

The presence of such “salience effects” suggests an additional margin through which policymakers can affect the efficiency of a tax.² In many cases, policymakers exercise some measure of control over the salience of the taxes used to raise revenue. For example, road tolls can be collected manually by operators or automatically through an EZ-Pass system (Finkelstein 2009). Similarly, negative income tax programs such as the Earned Income Tax Credit can be implemented as a payroll tax (so that it is appears on each paycheck) or folded into taxpayers’ general income tax returns (Chetty and Saez 2009). In the context of commodity taxation, governments may manipulate tax salience in several ways. First, policymakers may adopt general tax-inclusive pricing regulations, which require retailers to include the full amount of consumption taxes in the prices displayed to consumers, for all goods. Such regulations are common in Europe but are quite rare in the United States. Similarly, governments may require tax-inclusive pricing for a particular good. For example, the United States recently adopted legislation requiring airline ticket websites to include state and federal airline taxes in the initial price displayed to consumers. Finally, policymakers can manipulate salience by deciding whether a tax is included in a good’s posted price or added at the register (when the consumer checks out of the store). Given evidence that the former are more salient than the latter (Chetty et al. 2010), governments can alter a tax’s salience by adjusting the degree to which it relies on these two tax designs.³

Thus although policymakers typically lack perfect control over a given tax’s salience, they frequently face a choice between relying on high-versus low-salience means of raising revenue.

Because so many routine tax design choices affect salience, policymakers must make decisions about salience—whether they intend to or not—whenever they levy a new tax. Given that such decisions are soubiquous, it is striking that the topic has not received systematic theoretical attention. Indeed, as Congdon, Kling, and Mullainathan (2009) conclude in their review, “[T]he theoretical literature has yet to

¹See, for example, Chetty, Looney, and Knoft (2009) (grocery store customers reduce demand for goods when the sales tax inclusive price is posted; beer consumption declines more in response to excise tax changes [high-salience] than to sales tax changes [low-salience]); Finkelstein (2009) (drivers’ behavior becomes less sensitive to tolls upon adoption of EZ-Pass systems); Cabral and Hoxy (2010) (property taxes lower in jurisdictions that allow the tax to be collected in less salient ways); Gallagher and Muehlegger (2008) (income tax incentives to purchase fuel-efficient vehicles were most effective in the quarters most closely following income tax payment); Fochmann and Weimann (2011) (field experiment participants exerted more effort in response to a constant net wage when taxed than when untaxed). A related literature documents that income tax filers tend to systematically misperceive marginal tax rates, e.g., Lieberman and Zeckhauser (2001).

²Throughout, I employ the word salience to refer to the prominence of the taxed good’s tax-inclusive price. Thus an excise tax included in a good’s posted price is “high-salience” even though consumers may not be able to identify how much of what they pay is tax as opposed to the producer price.

³One can imagine other methods by which governments could manipulate tax salience. For instance, a government might combine high-salience taxes with lower-salience rebates. Many income tax deductions— which reduce individuals’ tax liability based on their expenditures—would likely fall into this category. Related concerns have been raised in recent debates over the desirability of a payroll tax cut as opposed to income tax rebates.
yield the type of rules of thumb with respect to optimal tax salience that translate into practical policy recommendations."

This paper aims to remedy this hole in the literature by considering the question of optimal commodity tax salience: how should a benevolent government choose between high- and low-salience taxes on a particular good in order to raise some required amount of revenue? That is, given the empirical finding that consumers under-react to low-salience taxes, should governments rely only on high-salience taxes, only on low-salience taxes, or instead utilize a combination of the two?

Addressing this question through the lens of economic theory is complicated by the conceptual and methodological challenges that arise in the context of behavioral welfare analysis. To analyze these thorny issues, I follow Chetty, Looney, and Kroft (2009) and adopt a “refinement” approach to welfare analysis, drawing on insights from Bernheim and Rangel (2009) and Chetty (2009b). As I elaborate below, this framework allows one to conduct welfare analysis while remaining agnostic about the exact mechanism driving consumer under-reaction to low-salience taxes, at least within a broad range of plausible models.

The analysis highlights two distinct mechanisms by which tax salience affects consumers’ well-being. On the one hand, low-salience taxes dampen the excess burden traditionally associated with non-lump-sum taxation; because consumers are less prone to substitute away from goods subject to low-salience taxes, such taxes are less distortionary for a given amount of revenue raised. On the other hand, low-salience taxes drive consumers to make optimization errors, reducing their welfare by causing them to misallocate their income among consumption goods. The government’s decision between high- and low-salience taxes trades off between these two distortions.

My results suggest an important role for salience considerations in the design of tax policy. When policymakers lack access to lump-sum taxation, commodity taxes generate an excess burden by distorting consumption decisions between taxed and untaxed goods. However, I show that when the government has access to tax instruments of varying salience, it can employ those instruments in combination to reach the first-best welfare outcome, even without relying on a lump-sum tax (at least under certain conditions). The assumptions I rely on to reach this result are exactly those adopted by Chetty et al. (2009).

Solving the government’s tax-design problem yields a simple and intuitive formula that characterizes the optimal mix between high- and low-salience taxes. Most notably, the optimal level of the low-salience tax is non-zero. Although low-salience taxes drive consumers to make optimization errors, the welfare costs of those errors is second-order for small values of the tax. In contrast, even small values of a low-salience tax may raise substantial revenues, allowing the government to reduce distortionary high-salience taxes while still meeting its budget constraint. Additionally, I show that the optimal ratio of high- to low-salience commodity taxes depends on the properties of the good being taxed. For luxury goods, or for goods that make up a large share of consumer expenditures, governments should (all else equal) rely more heavily on high-salience
taxes. In contrast, the presence of many close substitutes for a taxed good implies that low-salience taxes will be more efficient. Finally, high-salience taxes tend to be more efficient when consumption of the taxed good is associated with negative externalities.

Apart from the empirical work cited above, the theoretical papers closest to the current analysis are Chetty et al. (2009) and Chetty (2009a). Those authors derive formulas for measuring the excess burden of a new tax that is less than fully salient. Although intimately related to the current analysis, such formulas answer a fundamentally different question than what I address here. In particular, knowing the excess burden of a low-salience tax is not sufficient to determine what the optimal mix between those taxes should be. Computing the optimal mix also requires comparing the low-salience tax’s excess burden against the excess burden that would be generated by a posted tax, accounting for differences in revenue generation between the two tax types.

The remainder of this paper proceeds as follows. Section 2 elaborates the refinement approach to behavioral welfare analysis. Section 3 develops the model and derives the welfare results. In Section 4, I characterize the optimal combination of high- and low-salience taxes in terms of elasticities. Section 5 extends the analysis to goods that generate positive or negative externalities. Section 6 concludes.

2. Framework for Behavioral Welfare Analysis

The empirical literature on salience suggests that decision-makers fail to perfectly optimize when taxes are less than fully-salient.\(^4\) Adjusting decision-making models to account for departures from rationality can have important implications for welfare analysis, but one quickly runs into the problem that a large (potentially very large) number of models may be able to explain some observed pattern of decision-making, yet each do so in a way that implies a different welfare conclusion.

One approach to addressing this problem is to attempt to differentiate competing positive behavioral models empirically, waiting to draw welfare conclusions until the “true” decision-making model has been identified. However, the observationally-distinct qualities of competing decision-making models can be subtle at best, and the large number of potential models available may mean that confidently identifying the true underlying model may be a Sisyphean task.\(^5\)

Instead, I pursue a “sufficient statistics” approach to the problem (discussed in Chetty 2009b). In particular, I adopt the same two assumptions relied on by Chetty et al. (2009):

(A1) Consumption is a sufficient statistic for utility: tax design affects utility only through its effect on

\(^4\)Of course, the literature discussed in Section 1 does not prove that low-salience taxes induce consumers to behave sub-optimally, even putting aside questions of econometric validity. After all, any observed choice behavior can be rationalized by sufficiently relaxing one’s assumptions concerning the content of the preferences being observed. The observed behavioral discrepancy between high- and low-salience taxes only constitutes evidence of irrationality if one assumes that the form of a tax is irrelevant to the taxpayer’s true preferences. See (A1) below.

\(^5\)For example, Chetty et al. (2007) develop several positive behavioral models that generate salience effects.
the agent’s consumption.

(A2). When tax-inclusive prices are fully-salient, the agent chooses the same allocation as an optimizing agent.

Assumption (A1) states that if one holds an agent’s consumption bundle fixed, changing the tax rate or design does not affect the agent’s welfare. Although it is usually implicit, (A1) appears in most standard public finance models; writing an agent’s utility in terms of consumption alone implies that other parameters (such as the tax rate) do not directly enter the utility function. To understand the assumption, consider an example in which it fails: an agent would violate (A1) if she preferred facing a register tax to a posted tax on political grounds, perhaps because the amount being paid to the government is more transparent under the former relative to the latter. An important limitation of (A1) is that it rules out a class of interesting models in which decision-makers experience a psychic cost when faced with a low-salience tax, independent of the ultimate effect of the tax on their consumption. Consequently, if low-salience taxes generate substantial disutility for consumers independent of their effect on consumption decisions, the results presented here will over-state the benefits of low-salience taxes by neglecting such costs. However, as Chetty et al. (2007) demonstrate, even relatively small cognitive costs generate substantial under-reaction to a tax; as a result, omitting such costs from the model may not be as misleading as one might otherwise believe.

Assumption (A2) represents a weakening of the usual instrumental rationality assumption underlying the revealed preference approach to welfare analysis. Rather than assume that all of a decision-maker’s choices reflect their true preferences, (A2) posits rationality only for the subset of observed choices made when taxes are fully salient. One can understand (A2) as a “refinement” within the Bernheim and Rangel (2009) framework for behavioral welfare analysis: because we have reason to be skeptical about the quality of choices made when taxes are less than fully-salient, we privilege the choices revealed when such conditions are not present.

Although this approach to behavioral welfare analysis requires imposing assumptions on the content of consumers’ preferences, the payoff to that assumption is substantial. I am able to derive simple and intuitive formulas for optimal tax salience by specifying only that consumers under-react to some taxes relative to others, without having to make assumptions about exactly why they under-react the way in which they do.

3. The Model and Results

Consider the problem faced by a government that must raise $R$ revenue from taxes on some good $x$. The government has two tax designs to choose from: a high-salience tax $t_h$ and a low-salience tax $t_l$. Society
is composed of a representative consumer, who divides her income between \( x \) and a composite of all other goods, \( y \). I assume that utility is concave and smooth in both \( x \) and \( y \). Consistent with A1, tax salience does not affect utility apart from its effect on the agent’s final consumption.

The consumer’s budget constraint takes the form

\[
y + (p + t_h + t_l)x = I
\]

where \( p \) represents the pre-tax price of \( x \), \( I \) is income, and the price of \( y \) is normalized to 1. Production of \( x \) is characterized by constant returns to scale technology, so that the pre-tax price of \( x \) is fixed at its (constant) marginal cost. Taking income as fixed, demand for \( x \) and \( y \) can be expressed as a function of the two tax rates: \( x = x(t_h, t_l, p) \) and \( y = y(t_h, t_l, p) \). Indirect utility is thus given by \( V(t_h, t_l) = U(x(t_h, t_l, p), y(t_h, t_l, p)) \).

Consistent with the behavioral evidence described in the introduction, the model takes as its starting point the observation that consumers adjust their demand for the taxed good more strongly in response to changes in high salience taxes relative to changes in low salience ones. As in CLK, I define a tax’s salience by how much consumers respond to a change in the tax relative to a change in the posted price: \( \frac{\partial x}{\partial t_h} = \theta_H \frac{\partial x}{\partial p} \) and \( \frac{\partial x}{\partial t_l} = \theta_L \frac{\partial x}{\partial p} \), where \( \theta_H > \theta_L \). To illustrate, a tax that appeared as part of a good’s posted price (e.g. an excise tax) would be fully-salient (i.e. \( \theta = 1 \)). In contrast, a tax that consumers completely ignore would have \( \theta = 0 \). I focus on the case in which \( 0 \leq \theta_L < \theta_H \leq 1 \), although the approach here can be generalized to cases when a tax’s salience falls outside of this range.\(^7\) Although I motivate and discuss the model in terms of salience, \( \theta \) can be any characteristic of the tax that causes taxpayers to systematically err in a particular direction.

**Proposition 1**

*Under \((A1)\) and \((A2)\), the optimal combination of high- and low-salience taxes achieves the same welfare outcome as a non-distortionary lump-sum tax.*\(^8\)

**Proof of Proposition 1**

The government’s budget constraint is given by \( R(t_h, t_l) \equiv x(t_h, t_l, p)(t_h + t_l) = \mathcal{R} \). A government may manipulate the balance between high- and low-salience taxes, but it must do so in a revenue-neutral way in order to meet this constraint. Because consumers are more sensitive to increases in high salience taxes, the revenue brought in by a high-salience tax is less than the revenue raised by a low-salience tax of the same

\(^7\)The salience of the available tax instruments is ultimately an empirical question. The problem I consider takes as its starting point the availability of tax instruments with differing (but fixed) salience. As discussed below, additional feasibility considerations arise if a tax’s salience is endogenously related to the tax’s magnitude.

\(^8\)Proposition 1 assumes that policymakers may utilize both taxes and subsidies to raise the required amount of revenue (neither \( t_h \) nor \( t_l \) is constrained to be positive). Proposition 4 derives conditions for when this result will hold if tax rates are constrained to be positive.
\[ \frac{\partial R}{\partial t_l} = \theta L \frac{\partial x}{\partial p}(t_h + t_l) + x = \frac{\partial R}{\partial t_h} \]

Consequently, a revenue-neutral increase in \( t_l \) accommodates a greater than one-for-one reduction in \( t_h \):

\[ \frac{\partial t_h}{\partial t_l} \bigg|_\pi = -\frac{\theta L \frac{\partial x}{\partial p}(t_h + t_l) + x}{\theta H \frac{\partial x}{\partial p}(t_h + t_l) + x} < -1 \]  (2)

Totally differentiating the indirect utility function reveals that a revenue-neutral shift towards the low salience tax will benefit consumers if and only if

\[ \left. \frac{dV}{dt_l} \right|_\pi \geq 0 \iff U_x(x,y) \left. \frac{\partial x}{\partial t_l} \right|_\pi + U_x(x,y) \left. \frac{\partial x}{\partial t_h} \right|_\pi + U_y(x,y) \left. \frac{\partial y}{\partial t_l} \right|_\pi + U_y(x,y) \left. \frac{\partial y}{\partial t_h} \right|_\pi \geq 0 \]  (3)

Drawing on our behavioral assumptions about responsiveness to the tax designs and the consumer’s budget constraint, we can rewrite (3) as:

\[ \left. \frac{dV}{dt_l} \right|_\pi \geq 0 \iff -x \left( 1 + \left. \frac{\partial t_h}{\partial t_l} \right|_\pi \right) U_y(x,y) + \left. \frac{\partial x}{\partial p} \right|_\pi \left( \frac{\theta L}{\theta H} \left. \frac{\partial x}{\partial t_h} \right|_\pi \right) (U_x(x,y) - (p + t_h + t_l) U_y(x,y)) \geq 0 \]  (4)

Finally, using (2) and a little algebra, it is straightforward to show that (4) simplifies to:

\[ \left. \frac{dV}{dt_l} \right|_\pi \geq 0 \iff U_x(x,y) - pU_y(x,y) \geq 0 \]  (5)

Note that under the optimal policy, the government’s choice of \( t_h \) and \( t_l \) satisfies (5) and (2) with equality. Note further that the well-behaved nature of the the utility function and revenue constraint guarantee a unique solution to the government’s maximization problem. That is, Equations (5) and (2) uniquely determine the taxes by pinning down the taxpayer’s consumption of \( x \) and \( y \).

To prove Proposition 1, consider the effect on consumption of a (fully-salient) lump-sum tax of size \( \pi \). Because the lump-sum tax is fully-salient (A2) implies that we can find the consumption bundle induced by the tax by solving the ordinary consumer maximization problem subject to the lump-sum tax budget constraint:

\[ p x + y = I - \pi \]  (6)

which yields the first-order condition:

\[ U_x(x,y) - pU_y(x,y) = 0 \]  (7)

Assume that demand for the taxed good is not so sensitive that levying a new posted tax would actually reduce revenue: \( \frac{\partial x}{\partial p}(t_h + t_l) + x > 0 \). Note that this implies \( \tau e < 1 \).
Comparing the optimal tax salience conditions (1, (2) and (5)) with the lump-sum tax conditions ((6) (7)) reveals that both tax instruments induce the same consumption bundle. Because (A1) guarantees that consumption is a sufficient statistic for welfare, the two tax instruments achieve the same welfare result as well.

**Discussion of Proposition 1**

Proposition 1 states that when policymakers can manipulate the design of distortionary commodity taxes to modulate consumer responsiveness, they can do so in a way that replicates the efficiency effects of a non-distortionary lump-sum tax.

To understand the intuition behind the result, it is helpful to consider a stylized example. Suppose that the government is choosing between a fully-salient tax on $x (\theta^H = 1)$ and a fully-hidden one ($\theta^L = 0$). The consumer’s pre-tax budget constraint is given by AB in Figure 1, and consumption ($x_0$) is characterized by the tangency of the consumer’s indifference curve ($IC_0$) with AB. Because any feasible choice of tax rates must raise $R$, the taxpayer’s final consumption will lie somewhere on the line CD, which is simply AB shifted down by the vertical distance $R$.

If the government relied entirely on $t_h$, the consumer’s budget constraint will shift to AE; consumption of the taxed good ($x_h$) is the value that induces tangency with the consumer’s new indifference curve ($IC_h$). Like normal, the tax generates excess burden by driving consumers to substitute away from the taxed good.

In contrast, if the government relied solely on $t_l$, consumption of the taxed good does not change: $x_l = x_0$. Although consumers do not substitute away from the taxed good, the tax still generates an excess burden because consumers fail to adjust their consumption to account for the tax’s (non-distortionary) income effect.\(^{10}\)

Finally, under a lump-sum tax, there are no relative price changes. Consequently, consumption of the taxed good ($x_{LST}$) only falls in response to the tax’s income effect. Graphically, $x_{LST}$ is characterized by the point at which the consumer’s indifference curve is tangent to the new budget constraint (CD). Assuming that $x$ is a normal good, it is easy to see that $x_h < x_{LST} < x_l$.\(^{11}\) Because the lump-sum tax represents the first-best welfare outcome, the optimal policy lies somewhere between full reliance on either $t_h$ or $t_l$. Intuitively, by shifting the balance between the high- and low-salience tax, the government can move consumption along CD until it achieves $x_{LST}$.

**Corollary 1.1**

*Suppose that $\theta^H = 1$. Then under (A1) and (A2), the optimal size of the low-salience tax is non-zero, $t_l \neq 0$.*

\(^{10}\)see CLK on this point.

\(^{11}\)This condition that may fail when we do not have $\theta^H = 1$ and $\theta^L = 0$. In such cases, the government will need to rely on a combination of taxes and subsidies to reach $x_{LST}$. If subsidies are unavailable, the optimal policy takes the form of a corner-solution, see Proposition 4 below.
Proof of Corollary 1.1

By contradiction. Suppose that the optimal combination of high- and low-salience taxes entailed $t_l = 0$. Under (A2), this policy implies that the taxpayer behaves as a fully-optimizing agent. Because the taxpayer’s budget constraint is given by $(p + t_h) x + y = I$, consumption satisfies the standard first-order condition for an interior maximum:

$$U_x(x, y) - U_y(x, y) (p + t_h) = 0$$ (8)

However, we also know that at the optimal combination of high- and low-salience taxes, consumption is such that 5 holds with equality. Because $R > 0$ implies $t_h > 0$, Equations 5 and 8 may not be satisfied simultaneously. Hence we may conclude that $t_l \neq 0$.

Discussion of Corollary 1.1

Corollary 1.1 highlights a somewhat surprising result. Even when policymakers have access to a fully-salient tax instrument – one that induces no consumer mistakes – consumers are actually better off when the government utilizes a less than fully-salient tax. Intuitively, the optimal tax salience problem reflects a basic tension between high- and low-salience taxes. On the one hand, the less salient a tax is, the more it mutes consumer substitution away from the taxed good, thereby reducing the deadweight loss typically associated with non-lump-sum taxes. On the other hand, by causing consumers to depart from optimal decision-making, low salience taxes drive consumers to make optimization errors when making their purchasing decisions.

The key insight is that when taxes on $x$ are close to fully-salient, the former effect will be large relative to the latter. To see why, recall that by making the tax a little less salient, the government can raise the same amount of revenue while reducing the tax’s distortionary effects on consumption, thereby reducing the traditional source of excess burden. Although the reduction in salience does drive consumers to accidentally over-consume $x$ relative to $y$, the utility cost of that optimization error is trivial when the tax is close to fully salient; because consumers facing a fully-salient tax align the marginal utility of expenditures on $x$ and $y$, consuming a little too much $x$ relative to the optimum will not generate much less utility than if the consumer had purchased $y$ instead. Put differently, because the optimization error engendered by the low-salience tax depends on the difference in marginal utilities between $x$ and $y$, the welfare cost of that error is small when the consumer is near the optimum, (i.e. when the tax is close to fully salient).

More practically, Corollary 1.1 highlights that whenever taxes on a particular good are fully-salient, welfare can be improved by a small revenue-neutral shift towards a less-salient tax instrument.
4. Characterizing Optimal Salience

Thus far I have shown that the optimal combination of high- and low-salient taxes achieves the first-best welfare outcome. In this section, I characterize that optimal combination and highlight the factors that shape the relative desirability of high- and low-salience taxation for a particular good. The formula I derive is intuitive, and depends on empirically observable elasticities.

Let \( \omega_x \equiv \frac{(p+t_h+t_l)x}{x} \) denote the budget share of expenditures on \( x \). Let \( \eta_{x,t} \equiv \frac{\partial x}{\partial t} \) denote \( x \)'s income elasticity, \( \varepsilon_{x,p} \equiv -\frac{\partial x}{\partial p} \frac{(p+t_h+t_l)}{x} \) denote the own-price elasticity of \( x \), and \( \varepsilon_{x,t} \equiv -\frac{\partial x}{\partial p} \frac{(p+t_h+t_l)}{x} \) denote the compensated own-price elasticity of \( x \), where all elasticities have been defined to be positive. Finally, let \( \tau \equiv \frac{t_h+t_l}{p+t_h+t_l} \) denote taxes as a share of \( x \)'s tax-inclusive price.

**Proposition 2**

Suppose that utility is additively separable in \( x \) and \( y \). Let \( \rho \) denote the fraction of taxes on \( x \) that are low-salient: \( \rho \equiv \frac{t_l}{t_h+t_l} \). Then under (A1) and (A2), the optimal combination of high- and low-salience taxes is given by the value of \( \rho \) that solves \( \rho \varepsilon^L + (1 - \rho) \varepsilon^H = \theta^* \), where \( \theta^* \equiv 1 - \frac{\varepsilon_{x,p}}{\varepsilon_{x,p}[1 - \tau \omega_{x,t}]} \).

Proposition 2 states that for a given taxed good, there exists an optimal level of salience, \( \theta^* \), the value of which depends upon the nature of demand for the good in question. When the government can choose from tax instruments of varying salience, it should choose among them by reference to \( \theta^* \).

**Proof of Proposition 2**

Under additive separability, utility is given by \( U(x,y) = u(x) + v(y) \), so that indirect utility can be written as \( V(t_h,t_l) = u(x(t_h,t_l)) + v(y(t_h,t_l)) \). Additionally, we can rewrite (7) as:

\[
\frac{dV}{dt_l} \bigg|_R \geq 0 \iff v'(y)(t_h + t_l) + [u'(x) - (p + t_h + t_l)v'(y)]
\]

Let \( (x^*, y^*) \) denote the consumption bundle the consumer would choose if both the high- and low-salience taxes were included in the posted price of \( x \) (i.e. if both taxes were fully salient). That is, define \( x^*(t_h,t_l,p) \equiv x(0,0,p + t_h + t_l) \) and \( y^*(t_h,t_l,p) \equiv y(0,0,p + t_h + t_l) \). Under (A2), consumers behave optimally when the tax-inclusive price of \( x \) is fully salient. Consequently, \( (x^*, y^*) \) satisfies the first order condition for an interior solution from the consumer's standard utility maximization problem: \( u'(x^*) = (p + t_h + t_l)v'(y^*) \). Taking first-order Taylor approximations of \( u'(x) \) and \( v'(y) \) around \( x^* \) and \( y^* \) respectively yields

\[
u'(x) - (p + t_h + t_l)v'(y) \approx (x - x^*)\gamma
\]

All elasticities in the formulas are evaluated from a full-salience baseline, i.e. when all taxes that are not fully salient are set to 0. For example, this holds when \( x \) is only subject to posted taxes, or when \( x \) is left entirely untaxed. With additional assumptions, one can estimate these quantities when \( x \) is subject to a less than fully-salient tax, as long as one only identifies the parameter using variation in the posted price of \( x \).
where $\gamma = u''(x^*) + (p + t_h + t_t)^2 v''(y^*)$.

Additionally, using the definition of $x^*$, we can rewrite

$$x(t_h, t_t, p) - x^*(t_h, t_t, p) = x(t_h, t_t, p) - x(0, 0, p+t_h+t_t) = x(t_h, t_t, p) - x(0, 0, p) + x(0, 0, p) - x(0, 0, p+t_h+t_t)$$

Taking Taylor approximations and cancelling terms allows us to conclude:

$$x(t_h, t_t, p) - x^*(t_h, t_t, p) \approx -\frac{\partial x}{\partial p}[h(t_h(1 - \theta_H) + t_t(1 - \theta_L))]$$

(11)

Substituting (10) and (11) into (9) yields

$$\frac{dV}{dt_t} \geq 0 \iff v'(y)(t_h + t_t) \geq \gamma \frac{\partial x}{\partial p}[h(t_h(1 - \theta_H) + t_t(1 - \theta_L))]$$

(12)

To prove the result, we will need to express (12) in more familiar terms. First, note that we can approximate

$$v'(y) \approx v'(y^*) + (p + t_h + t_t) \frac{\partial x}{\partial p}[h(t_h(1 - \theta_H) + t_t(1 - \theta_L))v''(y^*)]$$

(13)

Additionally, implicitly differentiating the consumer’s first order condition and budget constraint at $(x^*, y^*)$ yields $\frac{\partial x}{\partial \gamma} = \frac{(p+t_h+t_t)v''(y^*)}{\gamma}$ and $\frac{\partial x}{\partial p} = \frac{v'(y^*)}{\gamma}$, where $x^*$ represents the compensated (Hicksian) demand for $x$.

Substituting this result and (13) into (12), and writing in terms of elasticities yields:

$$\frac{dV}{dt_t} \geq 0 \iff \frac{t_h(1 - \theta_H) + t_t(1 - \theta_L)}{t_h + t_t} \leq \frac{\varepsilon_{x,p}}{\varepsilon_{x,p}[1 - \tau\omega_x\eta_{x,t}]}$$

(14)

Finally, by rearranging terms and introducing the notation $\rho \equiv \frac{t_t}{t_h + t_t}$, we can express (14) as:

$$\frac{dV}{dt_t} \geq 0 \iff \rho \theta_L + (1 - \rho)\theta_H \geq 1 - \frac{\varepsilon_{x,p}}{\varepsilon_{x,p}[1 - \tau\omega_x\eta_{x,t}]} \equiv \theta^*$$

(15)

At $\rho^*$, (15) implies a (local) optimum: consumer welfare cannot be improved by a revenue neutral shift towards either $t_h$ or $t_t$. 

---

Notes:

13 Note that $\frac{\partial x}{\partial p}$ is evaluated at $(t_h, t_t, p) = (0, 0, p)$. As is standard in this literature (e.g. Auerbach (1985) and Chetty et al. (2009)), I ignore second-order terms in this approximation.

14 Recall $\omega_x = \frac{(p+t_h+t_t)x}{x}$ is the budget share of $x$, $\eta_{x,t} \equiv \frac{\partial x}{\partial y}$ is the income elasticity of $x$, $\varepsilon_{x,p} \equiv -\frac{\partial x}{\partial p} \frac{(p+t_h+t_t)}{x}$ is the compensated own-price elasticity of $x$ (defined to be positive), and $\tau \equiv \frac{t_h+t_t}{p+t_h+t_t}$. Note that $(t_h + t_t) \frac{\partial x}{\partial p} = \tau\omega_x\eta_{x,t}$. 
Discussion of Proposition 2

(a) The optimal tax is fully-hidden if and only if demand for the taxed good is entirely insensitive to income \( \theta^* = 0 \iff \eta_{x,t} = 0 \).

It is easy to see that this result follows mechanically from the optimal salience formula, once we restrict attention away from the (uninteresting) case in which \( \omega_x = 0 \). To understand the intuition, consider a fully-hidden tax \( \theta = 0 \), i.e., a tax that consumers entirely ignore when making decisions about what to buy. Let \((x_0, y_0)\) represent the initial consumption of \(x\) and \(y\). Suppose the government raises the tax to \( t^1 = t^0 + \Delta \).

Because \( \theta = 0 \), consumers buy the same amount of \(x\) as before the tax increase, leaving them with \( \Delta x_0 \) less income to spend on other goods.

When \( \eta_{x,t} = 0 \), the inattentive consumer’s response to the tax exactly matches what a fully-optimizing agent would do. Because the optimal amount of \(x\) does not depend on income, the consumer has nothing to gain by reconsidering her expenditures on \(x\) after a decline in income. In contrast, when \( \eta_{x,t} > 0 \), the consumer who fails to adjust her consumption of \(x\) in response to a tax increase is worse off for failing to do so. 15

(b) Optimal salience is lower for goods with close substitutes, \( \frac{d\theta^*}{d\tilde{\varepsilon}_{x,p}} < 0 \)

Mechanically, this result follows from totally differentiating the optimal salience formula with respect to the compensated elasticity of demand.16 To interpret this observation, recall the canonical public finance result that taxes generate welfare losses over and above the revenue they collect by driving consumers to alter their consumption decisions to avoid tax liability. The greater the compensated elasticity of demand for the taxed good \( (\tilde{\varepsilon}_{x,p}) \), the larger the excess burden associated with the tax (Auerbach (1985)). Consequently, the more readily consumers can substitute away from \(x\), the greater the welfare gain will be from reducing that substituting behavior.

(c) Optimal salience is (typically) higher for luxury goods, \( \frac{d\theta^*}{d\eta_{x,t}} > 0 \), and for goods that constitute a large share of consumers’ budgets, \( \frac{d\theta^*}{d\omega_x} > 0 \).

As noted above, the magnitude of the optimization error caused by low-salience taxes depends in part on the size of the income effect generated by the tax increase. In particular, \( \omega_x \eta_{x,t} \) measures the extent to which rational consumers reduce their consumption of \(x\) in response to the reduction in their purchasing power caused by the tax increase. The more sensitive demand for \(x\) is to income, and the greater the share

---

15When demand for the taxed good is insensitive to income, a fully-hidden tax is equivalent to a lump-sum tax. Under a lump-sum tax of \(t\), the consumer has \(t\) less income to spend, but otherwise allocates her income between goods as she did before the tax was imposed. Similarly, when a low-salience tax is fully hidden, the consumer allocates her income as she did before the tax was imposed; in the special case that \(\eta_{x,t} = 0\), that allocation will continue to be optimal.

16Note that solving for the total derivative of \(\theta^*\) with respect to \(\tilde{\varepsilon}\) requires accounting for the fact that \(\tau\) is endogenous as well. Mechanically, this entails totally differentiating the government’s revenue constraint in addition to the optimal salience formula.
of consumers’ budgets devoted to x, the more consumers err by failing to fully substitute consumption away from the taxed good. Consequently, the optimal ratio of high- to low-salience taxes will generally depend positively on $\eta_{x,I}$ and $\omega_x$.\textsuperscript{17}

(d) Numerical Illustrations

To provide a better sense of what Proposition 3 implies for the optimal salience of commodity taxation, this section computes optimal tax salience for a range of demand parameters. The computations assume that taxes constitute 5 percent of the taxed good’s after-tax price ($\tau = 0.05$). For context, CLK estimate that the sales tax has a salience of 0.06 with respect to beer purchases.

<table>
<thead>
<tr>
<th>Own-price elasticity ($\varepsilon_{x,p}$)</th>
<th>Income-elasticity ($\eta_{x,I}$)</th>
<th>Budget Share ($\omega_x$)</th>
<th>Optimal Salience ($\theta^*$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>0.1</td>
<td>0.10</td>
</tr>
<tr>
<td>0.5</td>
<td>1.0</td>
<td>0.1</td>
<td>0.20</td>
</tr>
<tr>
<td>0.5</td>
<td>2.0</td>
<td>0.2</td>
<td>0.80</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 1 shows that optimal salience is relatively small for a range of goods with typical demand elasticities. In particular, optimal salience is quite low even for luxury goods, unless those goods constitute a sizable share of the budget.

**Corollary 2.1**

Suppose that only taxes (and not subsidies) are available, i.e. $t_h \geq 0$ and $t_l \geq 0$. Then under (A1), (A2), and additive separability:

(a) the government can achieve the first-best efficiency outcome if and only if $\theta^L \leq \theta^* \leq \theta^H$.

(b) When $\theta^L > \theta^*$, the optimal policy is to tax $x$ using only low salience taxes, $\rho^* = 1$.

(c) When $\theta^H < \theta^*$, the optimal policy is to tax $x$ using only high salience taxes, $\rho^* = 0$.

**Proof of Corollary 2.1**

For (a), recall from 15 that a revenue-neutral shift towards the low-salience tax benefits consumers if and only if $\rho \theta^L + (1 - \rho) \theta^H \geq \theta^*$. When taxes are constrained to be positive, then $\rho \equiv \frac{t_l}{t_h + t_l} \in [0, 1]$. Consequently, the first-best welfare outcome will be achieved (i.e., 15 will only hold with equality) iff $\theta^L \leq \theta^* \leq \theta^H$.

\textsuperscript{17}An exception to the positive relationship between $\theta^*$ and $\eta_{x,I}\omega_x$ occurs when either $\tau$ or $\eta_{x,I}\omega_x$ is particularly large. Specifically, we can sign of the total derivative $\frac{d\theta^*}{d\eta_{x,I}\omega_x} > 0 \iff 1 - \tau (\varepsilon_{x,p} + \omega_x \eta_{x,I}) > 0$. To understand why this reversal occurs, recall from (7) that the magnitude of the positive welfare effect of a shift towards the low-salience tax depends on the marginal utility of the untaxed good, $v'(y)$. Reducing the salience of a tax accommodates a reduction in the total tax rate on $x$, generating a positive welfare effect by freeing up income that was being spent on $x$ for use on other goods. From (13), we can see that when the tax is a large fraction of the after-tax price of $x$, the optimization error pushes the marginal utility of $y$ higher than it was otherwise, $v'(y) > v'(y^*)$, raising the welfare gain from lowering the tax’s salience. For most goods in the United States, taxes are a small enough share of the good’s posted price so that the above condition will hold.
To prove (b), suppose that $\theta^L > \theta^*$ and $\rho < 1$. Because $\theta^H > \theta^L$ by assumption, and $\rho \in [0, 1)$, it must be the case that $\rho \theta^L + (1 - \rho) \theta^H > \theta^*$. Then by (15), the government can improve consumer welfare through a revenue-neutral shift from $t_h$ to $t_l$. Part (c) follows from analogous reasoning.

**Discussion of Corollary 2.1**

Corollary 2.1 states that when taxes are constrained to be positive, governments will only be able to achieve the first-best welfare outcome if the optimal degree of salience lies in between the salience of the available tax instruments. When $\theta^*$ lies outside the range of salience for the available tax instruments, the optimal policy takes the form of a corner solution, in which the government relies fully on the tax whose salience is closest to the optimal level.

**5. Tax Salience in the Presence of Externalities**

Section 4 derived a baseline formula for optimal tax salience and showed that the optimal degree of salience is relatively low for most goods. In Section 5, I highlight one factor that can increase optimal salience above the levels suggested by the previous section: the presence of negative externalities associated with the taxed good.

**Proposition 3**

When consumption of $x$ generates a negative (positive) externality, the optimal salience for taxes on $x$ is higher (lower) than if no externality was present.

**Proof of Proposition 3**

Suppose that consuming $x$ reduces social welfare $W$ by some positive quantity $\phi(x)$, $W(t_h, t_l) = V(t_h, t_l) - \phi(x(t_h, t_l))$, where $\phi' > 0$ and $\phi'' > 0$. The effect on social welfare of a revenue-neutral shift towards low-salience taxes is given by

$$
\left. \frac{dV}{dt_l} \right|_{\rho} \geq 0 \iff \rho \theta^L + (1 - \rho) \theta^H \geq 1 - \frac{\epsilon_{x,p}}{\epsilon_{x,p}(1 - \tau \omega_x \eta_{x,I})} + \frac{\phi'(x)}{(1 - \tau \omega_x \eta_{x,I})(t_h + t_l)} \gamma \frac{\partial x}{\partial p}
$$
where \( \gamma \equiv u''(x^*) + (p + t_h + t_l)\gamma v''(y^*) \). Interpreting the optimal degree of salience \( (\theta^*) \) as in Theorem 1, we thus have

\[
\theta^* = \theta^*_{\text{NoExternality}} + \frac{\phi'(x)}{(1-\tau_{x,I})(t_h+t_l)\gamma} \Theta > \theta^*_{\text{NoExternality}},
\]

where \( \theta^*_{\text{NoExternality}} \equiv 1 - \frac{\tau_{x,p}}{\varepsilon_{x,p}(1-\tau_{x,I})} \) is the optimal degree of salience when \( x \) does not generate any externalities. The proof is analogous for the case in which \( x \) generates a positive externality.

**Discussion of Proposition 3**

Proposition 3 demonstrates that the government’s choice of tax salience raises special considerations in the context of Pigouvian taxation. Intuitively, taxes on externality-generating activities can only internalize the social costs of those activities to the extent that decision-makers account for the existence of the tax when choosing their behavior. When a Pigouvian tax increases social welfare by discouraging taxpayers from engaging in a particular activity, the government will face an additional efficiency cost to reducing the salience of that tax. Conversely, when the taxed activity generates positive externalities, the efficiency benefits to relying on low-salience taxes are greater than would otherwise be the case. Thus for goods that generate externalities, the optimal salience formula derived in Section 3 should be treated as an upper or lower bound (depending on whether the associated externality is positive or negative).

Although intuitive, Proposition 3 has important implications for policymakers concerned with bringing about behavioral changes on the part of taxpayers. For example, to reduce population weight, a number of states levy sales taxes on soda and/or candy while exempting other food purchases from the sales tax base.\(^{18}\) Although this approach may increase the true relative price of unhealthful foods, the analysis here suggests that such taxes would be more likely to generate the intended behavioral effects if they were designed in more salient ways.

**6. Conclusion**

A long literature within public finance considers how to mitigate the excess burden of distortionary taxation. Motivated by new empirical findings that a tax’s salience affects consumer behavior, I showed how policymakers can utilize a combination of high- and low-salience commodity taxes to obtain the first-best welfare outcome, even when a lump-sum tax is unavailable. I also characterized the optimal combination of high- and low-salience taxes in terms of the good-specific elasticities that determine the efficient combination of taxes on a particular good. The formula I derive is both intuitive and practical; based on empirically-observable quantities, policymakers can compute the optimal fraction of taxes on a particular good that are imposed in the high-salience way.

\(^{18}\) See Fletcher, Frisvold, and Tefft (2010) for an empirical analysis of such taxes.
One notable feature of my results is that the optimal size of the low-salience tax is strictly non-zero (the optimal degree of salience is strictly less than one). Thus even if a government can raise all of its revenue without engendering mistakes on the part of taxpayers, taxpayers are actually better off when the government designs its tax system in a way that causes taxpayers to make optimization errors. This surprising result can be readily understood as an example of the Theory of the Second Best (Lipsey and Lancaster 1956). That is, the government’s need to raise revenue through a commodity tax generates a distortion that pushes social welfare away from the first-best welfare outcome. Consequently, by creating a new distortion – taxpayer deviations from optimal decision-making – policymakers can actually increase social welfare. More practically, this result suggests that governments should typically avoid relying on fully-salient taxes, unless the purpose of the tax is to reduce consumption of the taxed good. Similarly, policymakers should be skeptical of calls for tax-inclusive price regulations.

Several important qualifications should be kept in mind when interpreting the results presented here. First, my results may be local, in two important senses. First, like other studies that characterize optimal policies in terms of elasticities, a worry is that the elasticities that go into computing the optimum are measured under current policy, which may be far from optimal. Calculating the exact optimum to take account of this effect would require either detailed knowledge of how elasticities vary as policy changes, or failing that, an iterative process in which elasticities are recalculated as governments move closer and closer to the optimum policy. Failing that, the optimal salience results should be taken as a benchmark for whether current taxes should be made more or less salient to improve social welfare.

The second sense in which my results may be local concerns the possibility that the salience of a tax depends endogenously on the tax’s size. One might imagine that if a government substantially increased its reliance on low-salience taxes, taxpayers would become more attentive (the salience of the tax would increase). If the salience of a tax depends on the tax’s magnitude, obtaining the optimal degree of salience may no longer be feasible. Instead, the optimal policy would be to combine high- and low-salience taxes in the combination that brought the weighted average of the taxes’ salience closest to the optimal level of salience, accounting for the fact that relying more heavily on the low-salience tax might end up raising that tax’s salience. Because the model in this paper does not incorporate such effects, the results presented here can be most confidently applied to relatively small changes in policymakers’ use of high- versus low-salience taxes – changes that are less likely to substantially affect the tax’s salience.

Finally, by focusing on the case of a representative consumer, I have ignored distributional consequences associated with the choice between high- and low-salience taxes. In reality, tax salience may affect the

---

19 On the other hand, the results presented in Chetty and Saez (2009) and Gallagher and Muehlegger (2011) suggest that salience effects may persist even when the amount at stake is substantial.

Similarly, long-term reliance on low-salience taxes may not be feasible if consumers gradually come to learn about and account for the tax over time. However, Chetty et al. (2009) and Goldin and Homanoff (2011) find no empirical evidence of such learning effects. In particular, Chetty et al. present survey evidence suggesting that it is the prominence of a good’s after-tax price, rather than knowledge of the tax per se, that drives salience effects.
distribution of a tax’s burden in several important ways. For example, when consumers are heterogeneous in their attentiveness to low-salience taxes, setting the salience of a tax affects the distribution of the tax’s burden between attentive and inattentive consumers. This case is explored by Goldin and Homonoff (2011), who show how governments can manipulate tax salience to reduce commodity tax regressivity when high- and low-income consumers differ in their attentiveness to low-salience taxes. Additionally, tax salience affects the incidence of a tax between consumers and producers. In particular, CLK show that reducing tax salience can increase the fraction of the tax passed on to consumers. As in other contexts, the ultimate distributional effects of the choice between high- and low-salience taxes depend upon the supply and demand for the factors employed in production of the taxed good.

References


[6]  


**Figure 1**