

UC Berkeley

UC Berkeley Previously Published Works

Title

Tax Base Diversification as an Enforcement Tool

Permalink

<https://escholarship.org/uc/item/0b02x0pn>

Journal

American Law and Economics Review

ISSN

1465-7252

Authors

Galle, Brian

Gamage, David

Kuchumova, Yulia

Publication Date

2026

DOI

10.1093/aler/ahaf022

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Tax Base Diversification as an Enforcement Tool

Brian Galle,^a David Gamage,^b Yulia Kuchumova^c

July 10, 2024

a: Georgetown University Law Center. 600 New Jersey Avenue NW, Washington, DC, USA 20001. brian.galle@georgetown.edu (**corresponding author**).

b: Indiana University-Bloomington Maurer School of Law. 211 S. Indiana Avenue, Bloomington, IN, USA 47405. dgame@indiana.edu

c: Department of Economics, Higher School of Economics in St. Petersburg. St. Petersburg, 3 Kantemirovskaya Ulitsa, Korpus 1, Russia. jul.paramanova@gmail.com

Abstract

We examine when it is optimal to employ sales or VAT-type taxes as complements to a labor income tax. We find a Ramsey-type result in which each tax instrument should be imposed in inverse proportion to the combined elasticity of real and avoidance responses to the respective tax. Contrary to some prior results, we find that sales-type taxes are optimally non-zero across a variety of settings, and in particular when the (weighted) elasticity of taxable income with respect to the wage tax is greater than the *cross-elasticity* of taxable income with respect to the sales tax. We argue this parameter should be of key interest for empirical research, and conduct some exploratory studies in which we estimate it using changes in California local sales taxes and EU VAT rates. Our estimates generally suggest non-zero consumption taxes would be efficient.

Key Words: tax avoidance, tax enforcement, sales tax, VAT, optimal taxation

JEL Codes: H21, H26, H71, K42

Nontechnical Summary

Economic theory holds that taxes on wages are, with certain limited exceptions, generally preferable to taxes based on transactions, such as the U.S. sales tax or global value-added taxes. If we set aside savings, all income earned is also spent, so that taxing earnings and taxing spending can be economically equivalent. But if taxes on sales have exceptions or otherwise impose different rates on different items, as they almost always do in the real world, the sales tax is less efficient, because it distorts consumer choices in ways the tax on wages would not. At the same time, some authors have recognized that this general proposition may no longer hold once tax avoidance is taken into account. For example, development consultants often observe that a VAT can help nations with weak wage-tax institutions collect revenue. The intuition behind these claims, as we understand it, is that a sales-type tax is effectively another way of collecting wage taxes. If the two are economically equivalent, but a sales tax is harder to avoid than the income tax, it might be preferable to collect the “wage” tax at the cash register.

In this paper, we develop this intuition more formally. When, if ever, will these enforcement gains ever exceed the economic distortion caused by taxing some sales differently than others? We develop a model that allows us to identify the “sufficient statistics” or economic variables a policy maker would need to observe to decide whether to impose a sales tax. Generally speaking, we show that a

sales tax is more likely to be a good policy choice when what we call the cross-elasticity of taxable income with respect to the sales tax (“XETI”) is smaller in absolute magnitude than the elasticity of taxable income with respect to the wage tax (“ETI”). That is, if sales taxes distort taxpayers’ reported wages less than a revenue-equivalent wage tax, the sales tax is more likely to be a good choice. Since prior work has not identified this important role for XETI, there are no prior estimates of it. We therefore offer some preliminary evidence, based on taxpayer responses to California municipal sales tax elections and EU VAT changes, on whether XETI in fact is smaller in magnitude than ETI. Our estimates of XETI suggest that governments should likely collect at least some revenue through a sales tax or VAT.

1 Introduction

Development consultants often observe that a VAT can help nations with weak wage-tax institutions collect revenue.¹ E.g., Benitez et al. (2023); Ebrill et al. (2002). But this general intuition has little formal support, making it hard for policy makers to specify exactly when it should be deployed or what tradeoffs it entails. In addition, tax avoidance is a significant concern in developed nations, as well. It would be valuable to know whether there are potential welfare gains from using a VAT or sales taxes (which for brevity we hereafter refer to as “sales tax”) as a supplement to more traditional enforcement efforts.² ? and ? show that when wage taxes can be evaded but sales taxes cannot, it will often be welfare-improving to also employ a sales tax. In what circumstances does this edge case hold if both wage and sales taxes are subject to evasion or avoidance?

Thus, in this paper we attempt to identify when sales taxes can be justified as supplements to a labor income tax. To develop this argument, we present a model of the optimal mix of sales and wage taxes allowing for flexible-form interactions between wage- and sales-tax avoidances.³ We find that sales taxes are optimal to use when the (weighted) elasticity of

¹By wage tax we mean labor income tax, and use the two terms interchangeably throughout.

²The key defining feature of a “sales tax” in our model is that it is not measured and collected identically to a wage tax. Thus, we rule out consumption taxes that are implemented through a tax on wages with deductions for savings, but otherwise would broadly consider our model to extend to any “indirect” tax such as VAT, sales, or excise taxes, or economically-equivalent instruments such as product-specific redistributive legal rules. See Gillis, Mieszowski, & Zodrow (1995) for additional discussion of the distinctions between these instruments.

³We emphasize that we focus on the wage tax, not the income tax, in order to set to one side the appropriate treatment of capital taxation.

taxable wage with respect to the wage tax is greater than the *cross-elasticity* of taxable wages with respect to the sales tax, which we call XETI. This cross-elasticity is a new but important statistic because it captures the effect that sales taxes produce on the wage tax base. This effect arises not only due to real price effects, but also due to the sales tax’s influence on wage tax avoidance. We argue that XETI should be a key parameter for future empirical estimates. We then offer some exploratory estimates using two sets of data: from California and from the EU. Finally, we apply our estimates to test whether the data support the use of a positive sales tax in California and European settings.

We build on Keen and Slemrod (2017), who show that employing multiple tax enforcement “instruments,” such as investments in auditing, information reporting, and withholding, can improve the efficiency of an income-tax system. In a one-period model without avoidance or evasion, if all wage income is used to fund consumption, wage and sales taxes are generally equivalent (Bastani & Kohne 2022).⁴ We relax this equivalence by allowing for the possibility that avoidance techniques that escape one instrument may fail to reduce tax in the face of another. In that setting, the sales tax can be thought of as an additional potential enforcement instrument for collecting taxes on a wage base. Essentially, the sales tax offers an alternative mechanism for measuring and collecting a tax on labor income, in which standard taxpayer techniques for minimizing the wage tax may fail to minimize the sales tax. We thus extend Keen and Slemrod (2017) to consider the case in which an available enforcement tool may result in distortions in consumer choices.

Intuitively, if most taxpayer response to the labor income tax is in the form of avoidance, as in ?, and sales and labor income tax have distinct avoidance technologies, the first dollar of revenue raised through a sales tax will likely cause less deadweight loss than if labor income taxes were raised by the same amount (??). An increase in sales tax rate may even increase taxable income. Among other potential explanations, this could happen when higher sales tax rates increase the productivity of sales-tax avoidance, shifting taxpayer budgets towards that instrument and away from income-tax avoidance. More generally, in a world where sales-tax and wage-tax avoidances are variform and interdependent, avoidance of one may complement or substitute for avoidance of the other.

⁴It is widely accepted that under certain key assumptions a tax on wages is perfectly equivalent to an undifferentiated consumption tax as shown in ?. Of course, a long literature addresses cases in which the key assumptions do not hold, such as when some goods are necessary for work (?), there is uncertainty about individual wages (?), agents are heterogeneous in dimensions other than earning ability (??), or there are different underlying production technologies (?). Additionally, consumption taxes may in practice offer a broader base and so allow for lower rates (?). Differentiated tax rates may be appropriate for goods with externalities. ? and ? provide useful overviews of many of these arguments.

Accordingly, we present a model of the optimal mix of sales and wage taxes in which we allow flexibly for possible interactions between wage- and sales-tax avoidance. Avoidance of one may sometimes complement avoidance of the other, but in other cases they may simply be substitutes. As Keen and Slemrod (2017) do, we employ a Ramsey-type model that abstracts away from distributive effects. This assumption allows us to simplify the model down to a representative-agent model with a linear labor income tax and both sales-tax avoidance and wage-tax avoidance.

To help fix intuitions, consider two examples Gordon and Nielsen (1997) offer. In the “underground economy,” individuals may choose to be paid in cash in order to avoid reporting their wages to the tax authority. This cash may then be used to purchase goods from merchants, who in turn avoid reporting the sales transaction; ? finds evidence of this effect from rollout of electronic payments in the United States. If there were only a single margin for decision, entering the underground or not, then the division of tax burden between income and consumption taxes would be irrelevant. ? offers a similar argument. Alternately, Gordon and Nielsen (1997) suggest, cross-border work may also facilitate VAT evasion, so that the joint savings motivate the single decision about where to work. ? similarly argue that cross-border shopping likely does not justify consumption taxation.

In fact, though, there are many margins, each differently affected by the mix of the two taxes. Workers might accept only a portion of their wages in cash, or might take a part-time side gig that pays additional cash, so that the “underground” decision is a continuous, not discrete, one. So, too, each consumer transaction may be in cash or not. Perhaps the quality of products available with cash differs from those that may be purchased legally, and that as a result the marginal returns to consumption-tax avoidance diminish rapidly. At some point the worker will use her earnings for non-cash purchases, even if her earnings were in cash. Knowing this, her incentive to work off-books falls, in a way it would not have if she were solely subject to the income tax. We could say something similar about cross-border shopping, where in addition to discrete choices about each purchase decision workers may face wage and price effects caused by differential avoidance opportunities. ? considers general equilibrium effects of this kind in his analysis of the impact of sales taxes on the informal economy.⁵

Of course, even if sales taxes mitigate tax avoidance, they may also distort consumer

⁵In our case, we do so by assuming that tax avoidance opportunities and consumption choices are uncorrelated with earning potential (? considers the opposite case where avoidance can serve as a tag for income). We thus can focus our analysis on when there can be efficiency gains from optimal avoidance, leaving for subsequent development the possibility that any gains might be traded off against distributive concerns.

choices. Using a sales tax to collect wage taxes creates a potential tradeoff in which there may be reduced deadweight loss from avoidance and evasion but additional losses caused by differentially taxing different products. It is not obvious from prior work when one of these effects predominates over the other, or indeed whether potential enforcement gains can ever justify use of a differentiated sales tax. We model this tradeoff by allowing households to incur costly substitutions between taxed products and untaxed products. We find that even with potentially complex interdependence of avoidances, to judge whether sales taxes should supplement a wage tax we only need to know XETI.

We find a wide range of settings under which it is likely optimal to impose greater than zero sales tax, even if that tax distorts consumer choices among differently-taxed products. Specifically, we report a Ramsey-type result in which each of wage and sales taxes should be imposed in inverse proportion to the combined elasticity of real and avoidance responses to the respective tax. We also identify the necessary conditions for it to be optimal that sales taxes are non-zero. In addition, we consider a pair of extensions that take into account other significant real-world factors. In the first of these, we consider the possibility that a portion of private avoidance expenses are transferred back to the government, such as in the case where avoidance results in fines or shifts taxable income from a higher-taxed to lower-taxed base. In the second, we consider the administrative costs of enforcing both wage and sales taxes.

We find that sales taxes are optimally non-zero when the (weighted) elasticity of taxable wage with respect to the wage tax is greater than the *cross-elasticity* of taxable wages with respect to the sales tax, where the relevant weights depend on the relative sizes and rates of the two tax bases. What matters is that sales tax affects both the sales tax base and the wage tax base. This effect arises not only due to real price effects, but also due to the sales tax's influence on wage tax avoidance. Our proposition shows that if the impact of raising the sales tax on the wage tax base is large enough (which is measured by the cross-elasticity of taxable income) compared to the impact of raising the wage tax rate on the wage tax base then the sales tax rate should be zero. This result is identical if avoidance costs are partially transferrable to government, and holds with minor modifications in the presence of administrative costs that are continuous and increasing in both revenue and avoidance costs.

Our model suggests that XETI should be a key empirical parameter for tax planners, but there are no existing estimates of it to our knowledge.⁶ As a preliminary exploration, we

⁶Prior literature has estimated the cross-elasticity of leisure with respect to changes in the price of commodities, as in ? or ?. These estimates do not capture any of the tax avoidance outcomes or interactions that are our focus. Similarly, although ? and ? estimate the elasticity of leisure with respect to certain con-

estimate the cross-elasticity using data from California. California defines its sales tax base uniformly at the state level, and imposes a statewide sales tax, but localities are permitted to impose additional tax of their own on the same base. These taxes must be approved by local referendum. To obtain an exogenous source of variation, we exploit a wrinkle of California election law under which the timing of some elections is set about a decade in advance, so that it cannot be driven by local economic or political conditions. Alternately, we exploit another wrinkle under which similar tax-increase ballot propositions are sometimes required to pass by fifty-percent vote and sometimes by two-thirds majority, such that for initiatives with vote shares in between passage is essentially random. We then employ public ZIP-level data on taxes paid to estimate the impact of changes in the sales tax rate on income taxes due. While our confidence intervals are relatively wide, we can reject the optimality of zero sales tax rates at a 10% level of confidence for most of our estimates. For our preferred specifications, we can reject zero sales-tax rates at a traditional 5% level of confidence.

Although these California changes are relatively well-identified in the sense that they can plausibly be attributed to the sales tax change and not endogenous local conditions, they rely for identification on somewhat small changes in rates. To get a sense of whether our results are robust to larger changes, we also estimate the cross-elasticity using longitudinal nation-level VAT and income-tax data from the EU, where we have some changes that are ten times larger than our mean California change. We reject zero sales-tax rates with a .1% level of confidence, though admittedly the EU tax changes are less clearly exogenous than in the California setting.

In short, we offer new formal theory and some preliminary evidence in favor of the popular belief that sales taxes can be an effective enforcement tool for collecting taxes on the wage base, and extend earlier work on the efficiency of multiple tax bases by formalizing the trade-off between improved enforcement and distortions of both labor and consumption choices. This finding potentially has some relevance to the longstanding question about whether differentiated consumption taxes are ever optimal when an income tax is available. Despite its considerable generality and social importance, the question of when consumption-type taxes may be optimal has, in the words of the Mirrlees Report, “received surprisingly little formal attention” (?).⁷ While we do not address distributive concerns, which are arguably first order,

sumption taxes, they again do not estimate the overall cross-elasticity of the taxable wage base. ? estimate the cross-elasticity of labor income with respect to tax rates on capital, which has certain resemblances to our model, but because most of the cross-effects are caused by recharacterizing labor income as capital or vice-versa, which is generally not possible for the wage and sales tax bases, it does not offer any clear lessons for the cross-elasticity we estimate.

⁷This basic observation has a number of wide-ranging implications, as explored thoroughly in ?, ? and ?.

we note that any unwanted distributional effects of relying on the sales tax can be offset by adjustments to the income-tax schedule (Kaplow 2008), by refunding portions of sales tax revenues, or similar mechanisms. We also note that empirically many sales-tax imposing jurisdictions, such as U.S. states, impose income taxes that are not graduated, diminishing any distributional differences between the two instruments.

As we already mentioned, we are modeling only the static case. We therefore do not address the role of consumption taxes in a long-run dynamic model, such as the model explored in ?, ?, and more recently in ?; ? extends this model to include the possibility of consumption taxes. Accordingly, we have no occasion to consider the realism of the assumptions, such as long-run foresight by economic agents, infinite lives, and zero evasion, that these models arguably depend upon (? explore whether consumption taxes are optimal in a dynamic setting with evasion).

The paper proceeds as follows. Section 2 sets out our model, beginning with a base Case 1 in which there are no “multi-instrument” responses, and then a series of three extensions. Section 3 reports our exploratory California and EU empirical estimates. We then conclude.

2 Model

In this section, we will explore the optimal wage-sales tax structure in a representative consumer model. Case 1 begins with the assumption, similar to that in Gordon and Nielsen (1997), that wage and sales avoidance technologies are independent of each other. Case 2 extends our analysis by considering “multi-instrument” avoidance technologies, so that wage- and sales-tax avoidances are interdependent.

To begin, consider a representative individual who selects her labor supply (hours worked) L and the amount of consumption X .⁸ The consumer’s utility depends on leisure, consumption, and a public good G . The consumer cares about the amount of the public good, but has no direct control over it. The government can increase the amount of the public good only by increasing taxes on reported income and/or reported consumption. Specifically, the government imposes tax on labor income at rate t_Y and tax on consumer purchases at rate t_X . In so doing, however, there will be behavioral consequences, as taxes generate avoidance

Many of these are seemingly remote from questions of tax system design. ?, for instance, observe that one key function of the American Medicaid program is to shift financing of health care for the poor from what is essentially a sales tax to general federal revenues, which are primarily income-derived. Whether this shift is welfare-improving depends in part on the optimality of strict income-tax financing (?, ?).

⁸To highlight the importance of the avoidance decision to our result, we assume that all of the key assumptions of Atkinson & Stiglitz (1976), such as weak separability of labor and consumption, are satisfied.

expenditures, which we discuss in detail further. Define the utility of the representative consumer as $U(G, H(X, L))$ with $U'_G > 0$, $U'_H > 0$, $H'_X > 0$ and $H'_L < 0$. The price of the good without a tax is 1. The wage rate is 1. Hence, labor income $Y = L$.

To reduce wage and consumption taxes, a taxpayer can avoid some amount of taxable income, A_Y , at some cost and some amount of taxed sales, A_X , at some cost. We now turn to describe the cost associated with those avoidance activities.

Case 1. Costs of labor income and sales-tax avoidances are independent of each other. We start by assuming that the cost associated with wage tax avoidance, $C_Y(A_Y)$, depends only on the amount of labor income avoidance. Also, the cost associated with sales tax avoidance, $C_X(A_X)$, depends only on the amount of taxed sales avoidance. That is, in this baseline scenario, there are no “multi-instrument” avoidance technologies. Relative price effects are still allowed. Thus, sales-tax avoidance can still affect labor/leisure decisions by changing the after-tax returns to labor, and each form of avoidance can affect the other via the household budget constraint.

We assume that both $C_Y(A_Y)$ and $C_X(A_X)$ are increasing and convex. Also we assume that avoidance costs produce no utility. Intuitively, a reader can think of these costs as representing the difference in marginal productivity between taxed and untaxed expenditures or labor. Thus, we model the household’s choice between two goods: taxed sales and untaxed sales. This represents, in our model, the additional distortion from imposing a differentiated consumption tax that is emphasized in ?. For simplicity, in this base case we model avoidance as riskless, following the well-known approach of ?.

Let us now turn to the budget constraint. First, fix labor supply at a value L , which yields income Y . The amount of reported income is $Y - A_Y$. The amount of unreported income is A_Y . Given an income tax rate t_Y and avoidance expenditures $C(A_Y)$ and $C(A_X)$, the amount of money available for consumption is $(1 - t_Y)(Y - A_Y) + A_Y - C(A_Y) - C(A_X)$. Second, fix consumption at a value X . The amount of reported consumption is $X - A_X$. The amount of unreported consumption is A_X . Given a sales tax t_X , to consume X units of the good, the consumer pays the following $(1 + t_X)(X - A_X) + A_X$.

Now putting all things together, the consumer’s problem is to maximize her utility:

$$\max_{(X, Y, A_X, A_Y)} U(G, H(X, Y)) \quad (1)$$

subject to the budget constraint:

$$(1 + t_X)(X - A_X) + A_X + C_X(A_X) = (1 - t_Y)(Y - A_Y) + A_Y - C_Y(A_Y). \quad (2)$$

The government, in its turn, chooses two tax rates (t_X, t_Y) to maximize individual utility subject to the government budget constraint, which is

$$G = t_X(X - A_X) + t_Y(Y - A_Y) \quad (3)$$

Let us analyze the consumer's problem first. Because they only affect the consumer's budget constraint and not the consumer's utility directly, we first determine first the amounts of avoidance A_X and A_Y . . Consequently, the solution of the consumer's problem will be broken down into two steps: (1) how to spend on avoidance and (2) how many labor hours to supply (and thus to consume).

We proceed to the first step. Intuitively, for any consumption/labor income pair (X, Y) ultimately selected by the consumer, the consumer wants to maximize the income available for consumption and minimize the price of that consumption, recognizing she must spend real resources avoiding taxes. More technically, this means that the optimal amounts of avoidance A_X^* and A_Y^* are such that

$$A_X^*(t_X) = \underset{A_X}{argmax} - (1 + t_X)(X - A_X) - A_X - C_X(A_X), \quad (4)$$

$$A_Y^*(t_Y) = \underset{A_Y}{argmax} (1 - t_Y)(Y - A_Y) + A_Y - C_Y(A_Y). \quad (5)$$

According to these equations, the marginal benefit of increasing income tax avoidance is t_Y (the savings on income taxing). The marginal benefit of increasing sales tax avoidance is t_X . The consumer balances these benefits against the marginal cost of avoidance. This trade-off is reflected by the following FOCs, which determine the optimal amounts of avoidance, $A_X^*(t_X)$ and $A_Y^*(t_Y)$:

$$t_X = \frac{\partial C_X(A_X^*)}{\partial A_X}. \quad (6)$$

$$t_Y = \frac{\partial C_Y(A_Y^*)}{\partial A_Y}. \quad (7)$$

We will assume throughout that $C_X' > 0$, $C_X'' > 0$ and $C_Y' > 0$, $C_Y'' > 0$, and that the taxpayer chooses strictly positive avoidance amounts when $t_X > 0$ and $t_Y > 0$. Note that when $t_X = 0$, it is optimal to have $A_X^* = 0$ as a corner solution (for which equation (6) should

be modified to the inequality “less than or equal to”) and when $t_Y = 0$, it is optimal to have $A_Y^* = 0$ as a corner solution (for which equation (7) should be modified to the inequality “less than or equal to”).

As equation (6) and (7) indicate, the optimal avoidance expenditures only depend on the tax rates and the avoidance costs. These expenditures do not depend on the amount of labor supply or the amount of consumption chosen. Hence, our second step of the solution of the consumer’s problem will reduce to a standard one.

And now we move on to the second step. Given optimal avoidance choices, the consumer selects labor supply (hours worked) and consumption so that their marginal rate of substitution equals their relative prices. This trade-off is identified by the following FOC:

$$\frac{H'_Y}{H'_X} = -\frac{1 - t_Y}{1 + t_X}. \quad (8)$$

The optimal consumption X^* and income Y^* are determined by this FOC and the individual budget constraint (2).⁹ Plugging the optimal X^* and Y^* into the “inner” utility function $H(X, Y)$ gives us the “inner” indirect utility function $\tilde{H}(t_X, t_Y) = H(X^*(t_X, t_Y), Y^*(t_X, t_Y))$.

We now turn to the government problem. By increasing tax rates, the government balances the utility gain due to an increase in the public good against the utility loss due to a raise in taxes which reduces disposable income. This trade-off is reflected by the FOCs for the government’s problem, which for an inner solution are

$$U'_G \frac{\partial G}{\partial t_X} + U'_H \frac{\partial \tilde{H}}{\partial t_X} = 0, \quad (9)$$

$$U'_G \frac{\partial G}{\partial t_Y} + U'_H \frac{\partial \tilde{H}}{\partial t_Y} = 0, \quad (10)$$

In equations (9) and (10), $\frac{\partial G}{\partial t_X}$ and $\frac{\partial G}{\partial t_Y}$ are determined by the government budget constraint $G = t_X(X^* - A_X^*) + t_Y(Y^* - A_Y^*)$, by differentiating which we obtain

$$\frac{\partial G}{\partial t_X} = X^* - A_X^* + t_X \left(\frac{\partial X^*}{\partial t_X} - \frac{\partial A_X^*}{\partial t_X} \right) + t_Y \frac{\partial Y^*}{\partial t_X}, \quad (11)$$

$$\frac{\partial G}{\partial t_Y} = Y^* - A_Y^* + t_Y \left(\frac{\partial Y^*}{\partial t_Y} - \frac{\partial A_Y^*}{\partial t_Y} \right) + t_X \frac{\partial X^*}{\partial t_Y}. \quad (12)$$

As equation (11) shows, an increase in sales tax rate raises the amount of public goods

⁹To ensure that the sufficient conditions for the consumer problem’s optimum are satisfied, we will assume that $U''_{HH} < 0$, $H''_{XX} < 0$, $H''_{YY} < 0$, and $U''_{HH}H'_X H'_Y + U'_H H''_{XY} > 0$.

through a mechanical effect (the $X^* - A_X^*$ term) but this increase is mitigated by the fact that the tax distorts the choice of consumption and labor supply (the $\frac{\partial X^*}{\partial t_X}$ and $\frac{\partial Y^*}{\partial t_X}$ terms) and induces an increase of sales tax avoidance (the $\frac{\partial A_X^*}{\partial t_X}$ term). Similar logic lies behind equation (12).

By the envelop theorem, we have

$$\frac{\partial \tilde{H}}{\partial t_X} = -\frac{\lambda}{U'_H} (X^* - A_X^*), \quad (13)$$

$$\frac{\partial \tilde{H}}{\partial t_Y} = -\frac{\lambda}{U'_H} (Y^* - A_Y^*), \quad (14)$$

where λ is the Lagrange multiplier in the consumer problem. That is, by the envelop theorem, an increase in taxes reduces the “inner” indirect utility only through an income effect.

Using equations (9) - (14), we obtain that at the optimum, the tax rates (t_X, t_Y) should satisfy

$$\frac{-t_X \left(\frac{\partial X^*}{\partial t_X} - \frac{\partial A_X^*}{\partial t_X} \right) - t_Y \frac{\partial Y^*}{\partial t_X}}{X^* - A_X^*} = \frac{U'_G - \lambda}{U'_G}, \quad (15)$$

$$\frac{-t_Y \left(\frac{\partial Y^*}{\partial t_Y} - \frac{\partial A_Y^*}{\partial t_Y} \right) - t_X \frac{\partial X^*}{\partial t_Y}}{Y^* - A_Y^*} = \frac{U'_G - \lambda}{U'_G}. \quad (16)$$

These equations reflect that by increasing taxes, the government causes distortions in consumption and labor supply choices and additionally encourages consumers to increase spending on avoidance, which is privately and socially costly. Taxes can be cost-justified nonetheless because these resources are necessary to provide the public good — a good the consumer desires. More specifically, for instance, equation (15) reflects the government’s trade-off from a small uptick in the sales tax. On the benefit side (the right-hand side of equation (15)), a marginal increase in the sales tax rate boosts the amount of the public good, benefiting consumers. On the cost side, a marginal increase in the sales tax rate distorts consumption (the $\frac{\partial X^*}{\partial t_X}$ term); it increases sales tax avoidance (the $\frac{\partial A_X^*}{\partial t_X}$ term); and it makes leisurer more attractive relative to labor (the $\frac{\partial Y^*}{\partial t_X}$ term). This is the familiar “double distortion” argument, as in Kaplow & Shavell (1994).

Introducing the notation for taxable income, $Z_Y = Y - A_Y$, and taxable purchases, $Z_X = X - A_X$, we can rewrite the conditions for the optimal tax rates, (t_X, t_Y) , in a more

common way as

$$\frac{t_X}{1-t_X} \left[E(Z_X, 1-t_X) + \frac{t_Y Y}{t_X Z_X} E(Y, 1-t_X) \right] = \frac{U'_G - \lambda}{U'_G}, \quad (17)$$

$$\frac{t_Y}{1-t_Y} \left[E(Z_Y, 1-t_Y) + \frac{t_X X}{t_Y Z_Y} E(X, 1-t_Y) \right] = \frac{U'_G - \lambda}{U'_G}, \quad (18)$$

where $E(Z_Y, 1-t_Y) = \frac{1-t_Y}{Z_Y} \frac{\partial Z_Y}{\partial 1-t_Y}$ is the elasticity of taxable income with respect to $(1-t_Y)$ and $E(Z_X, 1-t_X) = \frac{1-t_X}{Z_X} \frac{\partial Z_X}{\partial 1-t_X}$ is the elasticity of taxable purchases w.r.t. $(1-t_X)$.

Proposition 1. At an optimum, the optimal wage tax rate and the optimal sales tax rate are determined by equations (17) and (18).

It is important to note that the formula for the optimal wage tax rate in our setting differs from the formula for the optimal wage tax rate in the case when only wage taxes (and not sales tax) is used. Specifically, compare equation (18) to equation (7) from ?. As can be seen, in Keen and Slemrod (2017) the optimal income tax rate depends only on the elasticity of taxable income, $E(Z_Y, 1-t_Y)$.¹⁰ In our case, when both wage and sales taxes are imposed, it is not only the elasticity of taxable income that matters but also the elasticity $E(X, 1-t_Y)$, which we will call the *cross-elasticity* of real consumption, that reflects how the wage tax rate affects household consumption. This effect arises because both wage and sales taxes affect the relative prices of consumption and labor. That is, the sales tax rate changes the benefit of avoidance. Thus, it changes the amount of consumption subject to the tax. The sales tax rate also changes the relative price of consumption to labor income (or leisure). Thus, it impacts the choice of how much to work. This latter, additional, cost that arises when the tax enforcement mechanism is implemented via the taxpayer's consumption choices is not modeled in Keen and Slemrod (2017).

Additionally, this result differs from Gordon and Nielsen (1997), where they artificially restrict their solution to the case where minor perturbations in tax rates cause no change in the relative prices of consumption (X) and labor (Y) and consequently the derived tax rates have no effect on household consumption or labor supply decisions, as well as no cross-effects.

We now can turn to our main question of interest: when should the government impose any sales tax in addition to the labor income tax in this scenario? To derive the corresponding

¹⁰Although Keen & Slemrod (2017) use the terminology of an “income” tax, nothing in their model depends on the government choice whether to tax capital in addition to labor income, and they do not model any inter-temporal choices by taxpayers. We thus understand their “income” tax to be equivalent to our use of a wage tax.

condition, we consider when a marginal increase in t_X would increase utility starting from $t_X = 0$. Thus, we need to replace the FOC (9) by

$$\left[U'_G \frac{\partial G}{\partial t_X} + U'_H \frac{\partial \tilde{H}}{\partial t_X} \right]_{t_X=0} > 0.$$

Utilizing then equations (13), (11) and (18), we obtain that for the sales tax rate to be positive it is necessary that¹¹

$$E(Y, 1 - t_X)|_{t_X=0} < \frac{X}{(1 - t_Y)Y} E(Z_Y, 1 - t_Y)|_{t_Y=0}, \quad (19)$$

where t_Y is determined by equation 18 with $t_X = 0$.

Corollary 1. A necessary condition for the optimal sales tax rate to be positive is that the cross-elasticity of income w.r.t. $(1 - t_X)$, $E(Y, 1 - t_X)|_{t_X=0}$, should be smaller than the elasticity of *taxable* labor income w.r.t. $(1 - t_Y)$, $E(Z_Y, 1 - t_Y)|_{t_Y=0}$, multiplied by $\frac{X}{(1 - t_Y)Y}$.

In general, Corollary 1 suggests that sales taxes are more likely to be optimal when real labor supply is less responsive to the sales tax rate than taxable labor income is to the wage-tax rate. In addition, holding all else constant, sales taxes are more desirable when the legal wage base is smaller than the legal sales-tax base.¹²

We will postpone further discussion of this result until we consider the more general overlapping (interdependent) avoidance technologies (case 2) and generalize the expression 19 accordingly.

Case 2. Costs of labor income and sales-tax avoidances are overlapping (interdependent).

We assume now that the cost of wage-tax avoidance depends on both the amount of avoidance and the amount of sales-tax avoidance, $C_X(A_X, A_Y)$. Similarly, the cost of sales-tax avoidance

¹¹The conditions required for the second-order sufficient condition for the maximum to be satisfied are $U''_{GG} < 0$, $U''_{HH} < 0$, $U''_{GH} > 0$, as well as $\frac{\partial^2 G}{\partial t_X^2} < 0$, $\frac{\partial^2 G}{\partial t_Y^2} < 0$, $\frac{\partial^2 G}{\partial t_X \partial t_Y} > 0$, $\frac{\partial^2 \tilde{H}}{\partial t_X^2} < 0$, $\frac{\partial^2 \tilde{H}}{\partial t_Y^2} < 0$, $\frac{\partial^2 \tilde{H}}{\partial t_X \partial t_Y} > 0$, and $\frac{\partial^2 U}{\partial t_X^2} \frac{\partial^2 U}{\partial t_Y^2} - \left(\frac{\partial^2 U}{\partial t_X \partial t_Y} \right)^2 > 0$.

¹²We use the term “legal” tax base here to distinguish between two concepts. In economic terms, both the wage tax and an undifferentiated sales tax are in effect taxes on the wage base. We say “legal wage base” to refer to a tax whose economic burden falls on labor and that is collected through payroll or similar mechanisms. We say “legal sales-tax base” to refer to a tax whose economic burden may also fall on labor but that is collected through taxes on retail sales.

depends on both the amount of wage-tax avoidance and the amount of sales-tax avoidance, $C_Y(A_X, A_Y)$. Note that this case is very general and nests several real-world stories about why the costs of avoiding sales tax would be interdependent with the costs of avoiding the labor income tax, including: 1) there is only one avoidance technology that is used to jointly avoid both wage and sales taxes; and 2) there are separate avoidance technologies, taxpayers must choose between avoiding wage or sales taxes, and demand for one avoidance technology affects the price of the other. That is, our model nests settings both in which avoidance of sales tax is a complement to avoidance of the wage tax, and those in which it is a substitute.¹³

Another example is if investments in one technology also pay off partially for the other, as with cross-border working and shopping, or the example in ? where research on tax evasion also reveals “barely legal” avoidance opportunities.¹⁴ ? also model evasion when evasion and avoidance can be complements.

The taxpayer problem and the government problem are the same as stated above, with the only modification the costs of wage-tax and sales-tax avoidance technologies we have already mentioned.

The optimal amounts of avoidance, A_X^* and A_Y^* , are now determined by the FOCs:

$$t_X = \frac{\partial}{\partial A_X} [C_X(A_X^*, A_Y^*) + C_Y(A_X^*, A_Y^*)], \quad (20)$$

$$t_Y = \frac{\partial}{\partial A_Y} [C_X(A_X^*, A_Y^*) + C_Y(A_X^*, A_Y^*)]. \quad (21)$$

In contrast to Case 1, optimal sales-tax avoidance, $A_X^*(t_X, t_Y)$, depends on both t_X and t_Y . Similar, optimal wage-tax avoidance, $A_Y^*(t_X, t_Y)$, depends on both t_X and t_Y . How exactly they depend on those tax rates is determined by their respective cost functions, which we define here quite generally.

Equations (8), (9), (10), (13), and (14) are still valid here. The difference arises in equations

¹³In general, we should expect that individuals will pursue a mix of single-tax and overlapping avoidance strategies. For example, let ‘X’ (‘Y’) be a technology that reduces only sales tax (only wage tax) and let ‘XY’ be a technology that reduces both wage and sales taxes. Avoidance technologies ‘X’, ‘Y’, and ‘XY’ all provide diminishing marginal returns, which means that the corresponding cost functions (call them $c_X(A_X)$, $c_Y(A_Y)$, and $c_{XY}(A_X, A_Y)$) are convex. Taxpayers choose an optimal mix of ‘X’, ‘Y’, and ‘XY’ technologies. Unless ‘XY’ is exceptionally productive, it is likely that at least some expenditures on both ‘X’ and ‘Y’ will be part of the optimal avoidance portfolio. The cost functions of avoidance that we introduced above, C_X and C_Y , in this example can be calculated as $C_X(A_X, A_Y) = \min_{(A_X^1, A_X^2, A_Y^1, A_Y^2)} [c_X(A_X^1) + c_Y(A_Y^1) + c_{XY}(A_X^2, A_Y^2)]$ s.t. $A_X = A_X^1 + A_X^2$, $A_Y = A_Y^1 + A_Y^2$.

¹⁴Since we assume tax reduction is riskless, we do not separately model avoidance and evasion. We offer the two here simply as examples of multiple possible tax reduction technologies.

(11) and (12), which now become

$$\frac{\partial G}{\partial t_X} = X^* - A_X^* + t_X \left(\frac{\partial X^*}{\partial t_X} - \frac{\partial A_X^*}{\partial t_X} \right) + t_Y \left(\frac{\partial Y^*}{\partial t_X} - \frac{\partial A_Y^*}{\partial t_X} \right), \quad (22)$$

$$\frac{\partial G}{\partial t_Y} = Y^* - A_Y^* + t_Y \left(\frac{\partial Y^*}{\partial t_Y} - \frac{\partial A_Y^*}{\partial t_Y} \right) + t_X \left(\frac{\partial X^*}{\partial t_Y} - \frac{\partial A_X^*}{\partial t_Y} \right). \quad (23)$$

Hence, the conditions for the optimal tax rates, (t_X, t_Y) , become

$$\frac{t_X \left(\frac{\partial X^*}{\partial t_X} - \frac{\partial A_X^*}{\partial t_X} \right) + t_Y \left(\frac{\partial Y^*}{\partial t_X} - \frac{\partial A_Y^*}{\partial t_X} \right)}{X^* - A_X^*} = \frac{\lambda - U'_G}{U'_G}, \quad (24)$$

$$\frac{t_Y \left(\frac{\partial Y^*}{\partial t_Y} - \frac{\partial A_Y^*}{\partial t_Y} \right) + t_X \left(\frac{\partial X^*}{\partial t_Y} - \frac{\partial A_X^*}{\partial t_Y} \right)}{Y^* - A_Y^*} = \frac{\lambda - U'_G}{U'_G}. \quad (25)$$

The interpretation of these equations stays basically the same as of equations (15) and (16). In addition to all the above-mentioned effects, there is one new effect: an increase in the sale tax rate induces not only a change in sales tax avoidance but also a change in wage tax avoidance (similarly, an increase in the wage tax rate induces not only a change in wage tax avoidance but also a change in sales tax avoidance).

Using the notation for taxable income, $Z_Y = Y - A_Y$, and taxable consumption, $Z_X = X - A_X$, we can rewrite the conditions for the optimal tax rates, (t_X, t_Y) , as

$$\frac{t_X}{1 - t_X} \left[E(Z_X, 1 - t_X) + \frac{t_Y Z_Y}{t_X Z_X} E(Z_Y, 1 - t_Y) \right] = \frac{U'_G - \lambda}{U'_G}, \quad (26)$$

$$\frac{t_Y}{1 - t_Y} \left[E(Z_Y, 1 - t_Y) + \frac{t_X Z_X}{t_Y Z_Y} E(Z_X, 1 - t_X) \right] = \frac{U'_G - \lambda}{U'_G}. \quad (27)$$

Proposition 2. At an optimum, the optimal wage tax rate and the optimal sales tax rate are determined by equations (26) and (27).

As can be seen from equation (27), the optimal wage tax rate depends not only on the elasticity of taxable income w.r.t. $(1 - t_Y)$, $E(Z_Y, 1 - t_Y)$, but also on the elasticity of taxable consumption w.r.t. $(1 - t_Y)$, $E(Z_X, 1 - t_Y)$. Compared to the previous case (see equation (18)), the second term in the brackets in equation (27) accounts not only for the effect of the change in tax rate t_Y on real consumption, X , but also on the amount of sales-tax

avoidance, A_X . Similar considerations apply to the condition that determines the optimal sales-tax rate (equation 26). We therefore have a Ramsey-type result in which each tax base should be used in inverse proportion to its impact on the combination of real responses and avoidance. This is also parallel to the result in ?, who find that individual components of the consumption tax base should face rates inversely proportional to the combined elasticity of all real and avoidance responses.

Again, we are particularly interested in when it would be optimal to have a non-zero sales tax rate. As explained in case 1, to derive the corresponding condition, we need to replace the FOC (9) by

$$\left[U'_G \frac{\partial G}{\partial t_X} + U'_H \frac{\partial \tilde{H}}{\partial t_X} \right]_{t_X=0} > 0.$$

Utilizing then equations (13), (22) and (27), we obtain that for the optimal sales tax rate to be positive it is necessary that

$$E(Z_Y, 1 - t_X)|_{t_X=0} < \frac{Z_X}{(1 - t_Y)Z_Y} E(Z_Y, 1 - t_Y)|_{t_X=0}, \quad (28)$$

where t_Y is determined by equation (27) with $t_X = 0$. This leads us to the following result.

Corollary 2. A necessary condition for the optimal sales tax rate to be positive is that the cross-elasticity of taxable income w.r.t. $(1 - t_X)$, $E(Z_Y, 1 - t_X)|_{t_X=0}$, should be higher than the elasticity of taxable income w.r.t. $(1 - t_Y)$, $E(Z_Y, 1 - t_Y)|_{t_X=0}$, multiplied by $\frac{Z_X}{(1 - t_Y)Z_Y}$.

In words, the use of a sales tax depends on a comparison of two key statistics. The first of these is the familiar elasticity of taxable (labor) income with respect to the (labor) income tax rate, $E(Z_Y, 1 - t_Y)$, which here should be evaluated when the sales tax is set to zero. The second and less familiar is the extent to which reported taxable income responds to the sales tax rate, $E(Z_Y, 1 - t_X)$, which we might call XETI for the *cross-elasticity* of taxable income. Sales taxes are less likely to be optimal when XETI is large relative to ETI. A convenient feature of this relationship is that its two components are readily observable, although to our knowledge there are no prior estimates of XETI.

Intuitively, what matters is that sales tax affects not only the legal sales tax base but also the wage tax base. In case 1 this effect arises only through one channel, the effect of the sales tax on real prices. Here there is also another channel: sales taxes influence wage tax avoidance. Corollary 2 shows that if the impact of raising the sales tax on the wage

tax base is small enough (compared to the impact of raising the wage tax rate on the wage base) then the sales tax rate should be positive. If XETI is sufficiently large, however, sales taxes should be zero. This situation appears in the Gordon and Nielsen (1997) multiple-instruments hypothetical in which opportunities for consumption-tax avoidance impact the income tax base as much or more than income tax avoidance would.

The ratio of the respective (reported) sales and wage tax bases is also important. A sales tax is less likely to be optimal, all else equal, when the reported wage-tax base is larger than the reported sales-tax base. With a small sales base, taxes raise little revenue at any given rate, making it more difficult for the sales tax to be net welfare-improving. Lastly, sales taxes become less desirable when wage-tax rates are low. Again, this is quite intuitive; at low rates, the marginal excess burden of the wage tax is small, so the sales tax would have to be relatively efficient in order to replace any portion of it.

Recall that our model is very general about possible forms and interdependence of avoidances, nesting cases both in which avoidance of sales tax is a complement to avoidance of the wage tax, and those in which it is a substitute. However, in all these settings, to evaluate whether sales tax should be used or not, we only need to know the XETI in addition to the ETI.

Note that the sign of the XETI could be positive or negative depending on whether sales-tax avoidance is a complement or substitute to wage-tax avoidance. If the sign of the XETI is negative then the condition for the sales tax rate to be positive defined by equation (28) would be trivially satisfied (because the ETI is positive and thus the right-hand side of the equation (28) is positive). This could occur when higher sales tax rates increase the productivity of sales-tax avoidance, shifting taxpayer budgets towards that instrument and away from wage-tax avoidance.

In Appendix A, we also consider two further extensions of the model. Case 3 considers the possibility that a portion of private avoidance expenses are transferred back to the government. Also, in Case 4, we add the administrative costs of enforcing compliance with two distinct collection mechanisms.

Overall, in all considered cases, our first core result is, simply speaking, that two distortions might be better than the one outlined in equation (28). Indeed, start with no sales tax. This pushes the consumer towards consumption rather than leisure. Yet with zero sales tax, the government must tax reported income at a high rate to provide public goods. The tax on reported income induces increased effort to conceal income. As the government increases the sales tax from 0, it can decrease the income tax and offer the same level of the public good. Further, starting from zero, the impact of a small increase in sales tax on the reported

labor income (i.e. on the labor supply and wage tax avoidance) might be quite small. The government prefers to have two (presumably) moderate taxes on different activities rather than one large tax on a single activity, as previously argued informally in Gamage (2014, 2015). By formalizing this claim, we provide some basis for determining the optimal point at which further increases in the sales tax would no longer produce welfare gains.

Additionally, another core result is that the *cross-elasticity* of taxable income (XETI) is a critical parameter to evaluate the optimality of the sales tax used. That is, it is the effect of sales tax rate on reported income (and not the effect of sales tax rate on reported consumption) that determines whether the sales tax rate should be zero or positive. Thus, while XETI has not been yet recognized and evaluated, an empirical investigation of the cross-elasticity of taxable income with respect to the sales tax seems to be an important agenda.

3 Exploratory Estimates of XETI

In this section, we attempt to provide an empirical sense of the extent to which our model would judge the use of real-world sales taxes to be efficient. In essence, we test Corollary 2 using data from California and the European Union. To do this, we first will estimate the key term from Proposition 2, the cross-elasticity of taxable income. Then, drawing on equation 28, we compute the critical value for this cross-elasticity required for a sales tax to be optimally non-zero. We then test whether our estimate can reject the null of zero sales tax.

Our empirical analysis relies on one or two key assumptions. Because Corollary 2 depends on elasticities when the sales-tax rate is close to zero, our exercise requires us to assume that the elasticities we measure are not sensitive to changes in rates. To the extent that XETI at the rates we observe differs from what XETI would be at rates closer to zero, we expect that if anything elasticities are likely to be larger in magnitude when rates are higher (for example, due to salience effects, as in ?). Thus our data are likely less conducive to finding

that sales taxes are optimally non-zero than data drawn from near-zero rates would be.¹⁵ In addition, while our model focuses on a tax on labor earnings, our data report income from both labor and investments. We do not have any basis for believing that results using this broader measure of income will be either more or less favorable to the efficiency of a sales tax than a narrower measure. We note, however, that our European data are concentrated on low- and middle-income taxpayers, for whom the labor share of income is very high.

3.1 The California Data

California provides a potentially useful setting in which to estimate the key term from Corollary 2, XETI, that is $E(Z_y, 1 - t_x)$. Although California imposes a statewide retail sales tax, city and county governments within California are also allowed to adopt their own increments to the tax. The definition of the tax base, such as whether certain items are exempt as food or services, is uniform throughout the state.¹⁶ Thus, unlike in a typical cross-jurisdictional setting, there is variation only with respect to rates. This crucial feature allows us to identify effects caused by rate changes alone, without having to be worried that this effect is confounded by difficult-to-observe changes in tax base. The rate changes here are relatively small in terms of $1 - t_x$, though, with the typical local increment falling in the range of one-quarter to one-half of a percentage point on top of the statewide rate, which was 7.5% for most of our sample period.

Another useful feature of California’s sales tax is that it offers several potential sources of random variation. This is critical because we might expect that political and bureaucratic correlates of a sales-tax change might confound any efforts to identify its impact on the income tax. For instance, jurisdictions might increase sales taxes when they become aware of looming threats to their income-tax base, or willingness to vote for sales-tax increases might be correlated with preferences for income-tax compliance.

Our main source of identification attempts to avoid such confounding by exploiting quirks

¹⁵We also assume, in interpreting the empirical findings, that in the real world there is a weak separability of labor and consumption. To the extent that sales taxes or VATs burden leisure, that effect could partially confound our efforts to identify XETI under the assumptions of our model, where we have assumed weak separability. But of course if weak separability fails then it will generally be efficient to impose consumption taxes.

¹⁶California’s retail sales tax has numerous major exceptions. Among others, it does not reach services. While in theory purchases outside California are taxable under a complementary “use” tax at the same rates, this tax was effectively unenforceable for most transactions during our sample period. And even for tangible goods, many items, such as food, are untaxed. Quite similar items can receive different tax treatment, as in the often-reported legal distinction between large marshmallows (taxable snacks) and small marshmallows (untaxed baking supplies).

of California’s voting rules. Every city or county sales-tax increase must be approved by voters in the affected jurisdiction. Some elections are relatively randomly timed, in that sales-tax initiatives often boost rates for a limited time, such as ten years, with a renewal election at the end of the period. While the date of the renewal election is of course known in advance, it is random in the sense that the ballot initiative’s original sponsors do not know what economic or political conditions will look like at the time of the renewal vote, making these elections less likely to be correlated with those background conditions. In our main specifications, therefore, we limit the sample to the five-year windows around renewal elections.

In addition, the threshold for success for tax proposals arbitrarily changes from fifty percent to a two-thirds threshold depending on relatively unimportant features of the proposal. Specifically, if the proponents of a tax increase (who need not be public officials) want the ballot seen by voters to include a statement that the tax proceeds will be used for any particular purpose, the proposal must receive two-thirds of the vote in order to pass. Only fifty percent is required otherwise. Since tax funds are largely fungible, these statements are not in practice binding, and as best we could determine there has never been a reported lawsuit seeking to enforce one. In short, there is only a small difference in labeling between the two kinds of tax proposals, although that difference is concededly important enough to some sponsors that they are willing to face the higher two-thirds hurdle. For propositions receiving votes between fifty percent and two-thirds of the vote, therefore, whether the vote results in a tax increase is somewhat random.¹⁷ Therefore, to further decouple the elections we study from the choices of the sponsors, in alternative specifications (reported in the Appendix) we limit the analysis to elections in the five-year window surrounding elections whose vote share falls in this range.

For example, consider the divergent election results in two representative cities. In 2014, a tax-increase proposal for Humboldt county passed with a 55.9% vote share. Sales tax rates in Eureka, a city in Humboldt county, increased from 8.25 to 8.75%. In 2016, Marin county’s renewal election failed with a vote share of 63.3%, and the sales tax rate in the city of San Rafael, located in Marin, fell from 9.25 to 9%. In our main specifications, we would include the Marin zips for the years 2014-2017 (our data end in 2017), while in our alternate specification we would include those as well as the Eureka zips for the years 2012-2016.

We scrape voting results from the California Secretary of State web site, which reports the results of all city and county elections, back to 1999. We then combine the vote information

¹⁷Results using a regression-discontinuity design, in which the discontinuity is the relevant vote threshold, are qualitatively similar to those we report in more detail. Due to the small number of observations close to fifty percent, and the even smaller number close to a two-thirds vote, these estimates are highly imprecise.

Table 1: **Summary Statistics**

	All Zips		Zips in Renewal Window	
Variable	Mean	SD	Mean	SD
Vote Share	.602	.120	.618	.112
Sales Tax Rate	8.100	.609	8.441	.680
City Population	49694	70676	170692	427646
CA Tax Liability	21.0*	49.1*	31.0*	67.1*
CA Returns	6207	8223	7706	8810
CA Tax Per Filer	4870	70807	5996	42965
US Tax Liability	96.2*	165.9*	131.0*	209.8*
US Returns	8601	8998	10667	9205
US Tax Per Filer	11337	18721	13459	22828

Notes: Dollar values in 2017 dollars. *: millions.

with zip-code level data on aggregate California income tax liability from the California Franchise Tax Board, as well as similar zip-level aggregates of U.S. tax liability from the I.R.S. Statistics of Income Division. Finally, of course, we include zip-level measures of the applicable California retail sales tax in a given zip-year. Our data cover a statewide panel ranging in time from 1999 to 2017, though federal zip-level tax data are missing values for the 1999, 2000, 2003, and 2006 tax years. We deflate dollar amounts to 2017 values using chained CPI. Table 1 provides summary statistics. Renewal zips are larger, report more tax liability per filer, and have a higher vote share in favor of sales taxes than other zips.

In order to obtain a clean comparison between winning and losing renewal elections, we limit our observations to zip-years in affected zip codes during the five-year windows around renewal elections. We drop observations that fall within the five-year window of more than one election, regardless of whether the second or additional elections were renewals.¹⁸ We also omit zip codes that cross city or county boundaries, so that zips are nested within cities which in turn are nested within counties. Because year-over-year changes in zip-level tax liability exhibit extreme swings at the tails of the distribution, we trim the sample to exclude zip-years in the top and bottom five percent by this measure.¹⁹ As shown in Figure 1, a simple binned plot of observations within this trimmed sample (controlling for year effects)

¹⁸We retain observations near the beginning of our data because we are able to observe the existence of prior elections through 1997.

¹⁹Results are essentially unchanged if we instead trim only the top and bottom one percent. Additionally, we repeat the analysis using a winsorization procedure in which we calculate the tax liability totals that would obtain if zip-level changes outside the fifth and ninety-fifth percentiles were to change by the fifth or ninety-fifth percentile rate. Results using this procedure are qualitatively similar to those reported but with rather wider confidence intervals.

../../../../Downloads/rev binscatter OMT v CA total tax.eps

Figure 1: Twenty Bins of One Minus Sales Tax vs. CA Income Tax Liability

Notes: This figure plots twenty bins of the log of one minus the sales tax rate in each zip-year. The y axis is logged CA income tax liability, aggregated to the zip-code level. Values are residuals after controlling for year fixed effects. Observations are limited to those in the five-year window around a vote on local sales tax rates, where the relevant vote share for passage falls between fifty percent and two-thirds of the vote. We also omit observations with extreme changes in income-tax liability.

suggests a negative relationship between $\log 1 - t_x$ and log tax liability.

To explore this relationship more formally in a regression framework, we then estimate fixed-effects panel regressions, with zip code as the panel variable, in which the treatment effect is a failed renewal election. To obtain an elasticity estimate, we interact this treatment with the difference in $\log 1 - t_x$ between the year after and the year before the election.²⁰ As ? explain, although the elasticity of the income-tax base is often referred to as ETI, where some taxpayer responses can include shifting between income brackets, the more direct measure is tax liability, not taxable income. Thus, our outcome variables are either California or federal tax liability, aggregated to the zip code level. We include controls for year and for city-by-year trends. Since local economic conditions may be an outcome of the sales tax regime, we omit economic controls.

Equation (29) summarizes our approach.

$$I_{it} = \alpha_i + \beta_1 Post * \Delta_{post-pre} (1 - t_x)_{it} + \lambda_t + \Phi_m \lambda_t + \epsilon_{it} \quad (29)$$

In this equation, $Post * \Delta_{post-pre} (1 - t_x)$ refers to the change, in the time from the year before to the year after each election, in logs of one minus the sales-tax rate. The coefficient β_1 can thus be interpreted as our key coefficient of interest, the cross-elasticity of taxable income with respect to one minus the sales-tax rate. λ_t is a year fixed effect, while $\Phi_m \lambda_t$ is a municipality-by-year trend. ϵ_{it} is the error term. The α_i term reflects zip-code level fixed

²⁰During the sample period, statewide sales tax rates changed several times. Since these changes affected all our observed jurisdictions simultaneously, we can account for them simply by including a control for year effects. All other observed rate changes were the result of local elections.

effects. We cluster errors at the city level, as that is the level of treatment.²¹

With respect to variable I_{it} we allow two possibilities: aggregate income and income per tax filer. Conceptually, it is not clear whether the appropriate measure of taxable income is aggregate income or instead income per tax filer. In theory, inter-jurisdictional mobility and local economic conditions are both potential outcomes of a change in the sales-tax rate, which could in turn affect the number of filers. It seems implausible to us, however, that sales-tax changes of this magnitude would cause measurable short-run relocations. Whether they could potentially affect the decision to file a return is less clear. Thus, we present both results where tax liability per filer is the outcome (essentially ruling out relocation and filing decisions as outcomes of the tax increase) and those for which the outcome is total zip tax liability. For each, we present estimates weighted by the number of filing households in the jurisdiction as well as unweighted.

Admittedly, while the timing of the renewal election is relatively independent of local conditions, the vote share in favor of retaining the prior tax increase is not. To provide a sense of whether differences in voting may be correlated with our outcomes of interest, we plot in Figure 2 the trends in tax liability for zips where renewal passed against those where it failed. It appears that three years before an election, there is a small increase in tax liability, on average, among zips where renewal fails (i.e., our treatment group), but no observable changes in the two years leading up to the election. Reported income tax liability dips in these zips after the sales-tax measure fails, which again is consistent with a small negative relationship between sales-tax rates and income-tax liability. We observe similar patterns in an event-study analysis, as shown in Appendix Figure 4.

Our results are presented in Table 2. As might be expected given the relatively small changes in rates we observe, our estimates for the cross-elasticity are somewhat imprecise. For California taxable income estimates, we find 95% confidence intervals that range from about 1.2 to as low as -7, with point estimates that are consistently negative. Federal tax estimates are somewhat smaller in magnitude but more precisely measured. Since our estimand is one minus the sales tax rate, these results imply that increases in the sales tax rate are correlated with greater taxable incomes. Among other possible explanations, this could be consistent with a hypothesis that higher sales tax rates increase the productivity of sales-tax avoidance, shifting taxpayer budgets towards that instrument and away from income-tax avoidance. While the point estimates are sizable, they are small in economic terms. The median rate change is one-quarter point, about a .27 percent change in terms of one minus the rate. Thus

²¹Results are essentially unchanged if we instead cluster by county. Estimates are rather more precisely estimated if we cluster by zip code.

../../../../Downloads/election trends renewals zoomed.eps

Figure 2: Per Capita CA Income Tax by Years from a Sales-Tax Election

Notes: The figure plots log CA income-tax liability, aggregated to the zip-code level, by the number of years each zip-year is removed from a local sales-tax election. We plot zips where the election resulted in a successful renewal of a prior sales-tax increase against those where the renewal failed. Zip-years with extreme changes in income-tax liability are omitted.

our point estimate in Table 2, Column One implies an accompanying change in total CA income tax of about .83 percent.

Of greater interest, our estimates allow us to pin down whether XETI is consistent with an efficient non-zero sales-tax rate, as defined in Corollary 2 and Equation 28. Using sample means and prior literature, we compute a critical value for XETI such that a non-zero rate would be efficient if the critical value exceeds our estimated XETI. We then report a test statistic for whether our estimate falls below the critical value, i.e., whether it is consistent with non-zero sales taxes.

NBER taxsim data report that the mean federal income tax rate in California (net of deductions for California or other state taxes) is .2131 over the latter period of our sample, while the mean California rate we observe in our data is .0635. This yields a value for $1 - t_y$ of .7234. Our review of the literature on the elasticity of taxable income suggests a range of possible values, ranging from a low of .12 reported in the survey by ? to a high of 1.2 in ?. ? report a top-range estimate of .4, and this is also close to other recent estimates in ? and ?, albeit the latter with German data. We therefore calculate three critical values for equation (28), using ETI of .12, .4, and 1.2. These critical values are .1785, .5950, and 1.785, respectively.

The bottom panel of Table (2) reports a p-value for a one-sided test that the estimated coefficient for $E(Z_y, 1 - t_x)$ is less than the critical value (offered with the usual caveats about one-sided tests). If ETI is at least as high as the middle value in the literature, .4, then all our estimates reject zero sales tax with 94% confidence, and most reject with 99% or greater confidence. When we weight by filers to obtain results for the average filer, rather than the average zip code, we reject zero sales tax with at least 98% confidence for all three

Table 2: **Cross-Elasticity of Income Tax Liabilities w.r.t. Sales Tax Rates**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	CA Tax	CA Tax (Weighted)	CA Tax Per Filer	CA Tax Per Filer (Weighted)	U.S. Tax	U.S. Tax (Weighted)	U.S. Tax Per Filer	U.S. Tax Per Filer (Weighted)
$Post \times \Delta Sales$	-3.063	-3.452**	-2.041	- 3.036***	-2.311**	-1.740**	-1.420*	-1.554**
	(2.041)	(1.369)	(1.662)	(1.136)	(0.999)	(0.802)	(0.786)	(0.741)
Observations	8,503	8,503	8,503	8,503	4,672	4,672	4,672	4,672
R-squared	0.399	0.738	0.386	0.688	0.822	0.883	0.810	0.871
One-sided p-Value at ETI:								
.12	.944	.996	.909	.998	.993	.998	.991	.979
.4	.963	.998	.943	.999	.998	.999	.998	.995
1.2	.991	.999	.989	.999	.999	.999	.999	.999

Notes: This table reports the cross-elasticity of reported California and U.S. income tax with respect to changes in one minus the local sales tax rate. Fixed-effects OLS panel estimates with zip code as panel variable. Standard errors clustered by city in (parenthesis). Weights by number of filing households in zip code. $Post \times \Delta Sales$ is an interaction term between an indicator for the period after an election and a continuous measure of the difference in the sales-tax rate between the year after and the year before the election. One-sided p-Value panel reports test statistic for a one-sided test that the estimated coefficient on $Post \times \Delta Sales$ is less than the critical value implied by the listed ETI. Observations are limited to years surrounding renewal elections. Zip-years with extreme changes in income-tax liability are omitted.

values of ETI. This is not to say that California’s actual rate is efficient, only that it would be efficient to impose a non-zero rate. Further, as ? find for ETI, we cannot reject equal XETI of California and federal tax (although point estimates for California XETI are larger in absolute magnitude), notwithstanding the likelihood that California’s tax base offers more margins of response.

These results are robust to several alternative specifications. As we noted, the outcomes of elections may be endogenous to other household tax responses, notwithstanding that we see little in the way of pre-trends. Accordingly, in the Appendix, we provide estimates employing another quasi-random treatment approach based on the arbitrary vote threshold needed for passage, as described above. We also repeat this method, but limiting only to renewal elections. In this third specification, both timing and passage are arguably as good as random. All three approaches yield similar point estimates, although as expected the smaller sample for the third method results in wider confidence intervals.

3.2 The EU Data

Because the California rate changes are relatively small, we next attempt to identify XETI using larger but less-cleanly identified rate changes from the EU. Drawing on the EC’s VAT Rates publication, we assemble all VAT rate changes adopted in the EU over the period 2000 to 2015.²² In addition to overall rate changes, which are generally modest (aside from a three-year period in Hungary in which rates went from twenty to twenty-five percent and then back), we also observe a number of instances in which one or more product categories move from the full VAT rate to the country’s reduced rate (which is typically ten percentage points or more lower), or vice-versa.

To translate these product-level policies into their effect on the overall VAT rate, we weight the product-level change in rate by the overall share of the consumer’s basket of goods that product represents. To ensure this fraction is exogenous to the tax changes, we use the weight that would hold in 1999 in each country. For instance, if the standard French VAT rate is 20%, the reduced rate 5%, and dairy products represent 10% of mean 1999 French consumer purchases, we would count a change to the reduced rate for dairy products as a $15\% * 10\% = 1.5\%$ overall reduction. After this adjustment, the median VAT change is .9 percentage points, a 1.1% change in terms of $1 \text{ minus } t_x$. Ten percent of the changes we observe are 5 percentage points or larger in absolute magnitude.

We combine these data with country-level estimates of income-tax payments and rates

²²We include member states joining in 2004 and 2007, but omit Croatia, which joined in 2013.

from the EC’s Structure of Earnings Survey, as reported on Eurostat. We identify the marginal income-tax rate and mean overall (real, euro-denominated) income-tax liability reported for thirteen representative household types in each country-year pair.²³ Because income-tax responses may differ sharply for households with negative tax, our main specifications omit households that would face negative effective marginal rates or negative liabilities.²⁴ Since our outcome variable is tax liability, not taxable income, we include controls for both the income-tax rate and its square. This reflects both the mechanical effect of rates on total tax, and the opposite-signed behavioral effect of rates on the taxable base.

As in ?, we then estimate a fixed-effects regression in which we generalize an event-study design with multiple periods, products, and countries. Specifically, we estimate:

$$\Delta \log(T_{icth}) = \beta_0 \Delta \log(1 - t_{ict}^x) + \sum_{k=-20, k \neq 0}^{k=20} \beta_k \Delta \log(1 - t_{ic(t+k)}^x) + \Delta \lambda_t + \gamma \Delta X_{ct} + \theta \Delta \log(1 - t_{icth}^y)^{1,2} + \Delta \epsilon_{icth}, \quad (30)$$

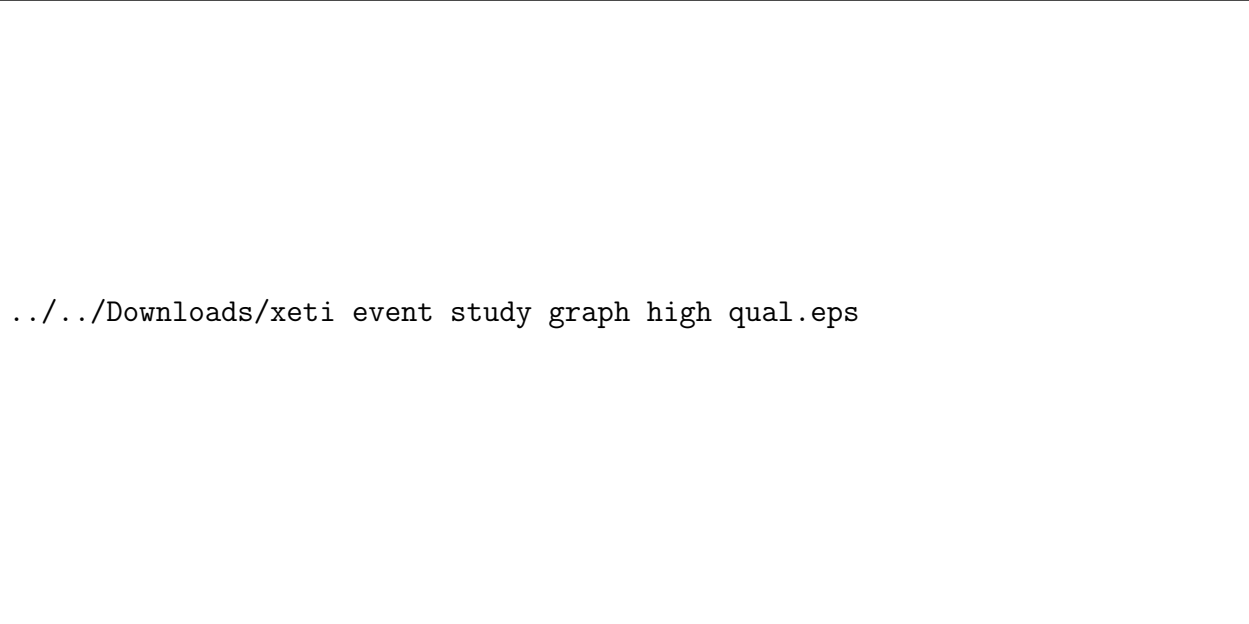
where the outcome T_{icth} is the household’s income-tax liability, t_{ict}^x is the VAT rate, λ_t represents time fixed effects, and t_{icth}^y is the marginal income-tax rate for each household type. ϵ_{icth} is the error term. X_{ct} is a vector of controls for each country’s nominal interest rate, GDP per capita, and unemployment rate. We index by i the product, c the country, t the year-month, and h the representative household type. The Δ symbol indicates first-differences; that is, for each variable X the Δ represents the value $X_t - X_{t-1}$. Our main variables of interest are the set of interaction terms $\beta_k \Delta \log(1 - t_{ic(t+k)}^x)$, which measure the effect of the monthly VAT-rate change in the twenty months before and after each actual change.²⁵

Results are fairly similar to those we obtain using the California data. While in some post-change months we estimate positive point estimates for XETI, as Figure 3 illustrates, estimated values are grouped tightly around zero. The top of the confidence interval for

²³Marginal rates are not drawn directly from survey reports, but instead are computed by Eurostat with a tax simulation calculator, based on survey responses about household characteristics. Each of the representative households represents one set of inputs into the calculator for each country. Representative households are structured by marriage status, number of children, and earnings percentile; for example, one such household is defined as two adults, two children, and one earner with median wages. Tax liabilities reported are mean values from surveyed households falling in each category.

²⁴Results are robust to including negative-tax households separately with an indicator for negative-tax status.

²⁵Because tax liability is measured yearly while our VAT rates are measured monthly, we drop observations in which there could be no change in income-tax liability. Thus, the estimate for two months post-treatment is the mean effect, in the next calendar year, of VAT changes adopted in November, while the estimate for eight months pre-treatment is the mean change from the prior calendar year for VAT changes adopted in August, and so forth.



.../.../Downloads/xeti event study graph high qual.eps

Figure 3: Event-Study Estimate of XETI from EU-VAT Changes

Notes: The figure plots estimated coefficients in a generalized event study combining all product-level VAT rate changes in the EU between 2000 and 2015. The coefficients plotted correspond to XETI, or the cross-elasticity of income-tax liability with respect to (one minus) the VAT rate. The underlying regression includes controls for the income-tax rate and country-level economic conditions.

Table 3: Cross-Elasticity of Income Tax w.r.t. EU VAT Rates

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Log (1-VAT Rate)	0.0318 (0.0254)	0.0411 (0.0339)	0.0483 (0.0433)	0.0304 (0.0246)	0.0366 (0.0323)	0.0410 (0.0415)
Observations	4,330,039	4,330,039	4,330,039	4,105,604	4,105,604	4,105,604
R-squared	0.743	0.739	0.739	0.391	0.394	0.393
Weighted by Product Share		X	X		X	X
Product FE			X			X
Instrumented for ITR				X	X	X
One-sided p-test at ETI:						
.12	.999	.999	.997	.999	.999	.999
.4	.999	.999	.999	.999	.999	.999
1.2	.999	.999	.999	.999	.999	.999

Notes: This table reports estimates of the mean cross-elasticity of income-tax liability with respect to one minus the VAT rate for thirteen representative household types in the EU between 2000 and 2015. Estimates are fixed-effects OLS using country as the panel variable. All columns include year fixed effects and controls for nominal interest rate, GDP per capita, and unemployment rate. Columns 2 and 6 also include product-category fixed effects. Standard errors clustered by country are in (parentheses). Weights are by each product's 1999 share of consumer bundle in the respective country. Instruments for income-tax rate are the first (annual) lags of the rate. One-sided p-test panel reports test statistic for a one-sided test that the estimated coefficient on Log (1-VAT Rate) is less than the critical value implied by the listed ETI. Values greater than .9995 reported as .999.

XETI exceeds .1785, the test statistic for a non-zero VAT at the lowest assumed value of ETI, in only three post-change months. Point estimates are uniformly below .1, round to less than zero in five of the twenty post-change months observed, and sum to a bit less than zero. There are no discernible pre-trends.

To give a further sense of any possible longer-run effects resulting from persistent changes, we also present simple fixed-effects estimates in which we regress log real tax liability on (log of) one minus the VAT rate, (log of) one minus the income-tax rate, and economic controls. We include year and country and, in alternative specifications, product-level fixed effects, so that identification comes from within-country and product changes in VAT rates. Since marginal income-tax rates may be endogenous to tax avoidance, we also include specifications in which we instrument for the income-tax rate and its square with their lags, as is standard in the ETI literature. As with our California data, we additionally test whether the resulting estimates of XETI would be consistent with a non-zero VAT.

Across our various specifications, we consistently reject the efficiency of a zero VAT rate

at a .1% level of confidence, as summarized in Table 3. Our point estimates of XETI are positive, but very close to zero, with point estimates in the .03 to .05 range and the apex of the 95% confidence interval around .12, at or below the smallest ETI estimates in the literature.

We interpret both the event-study and fixed-effects results cautiously. In this setting, it is impossible to rule out endogeneity of VAT and income-tax policy. Survey data likely measure true tax liabilities with some error, potentially attenuating our measured coefficients, which here would bias us against rejecting our null hypothesis. What we can say is that our measures of XETI are broadly similar to those we find in the California data in one key respect: they imply little or no increase in income-tax avoidance in response to sales-tax increases.

In short, across our two data sources we find suggestive evidence that even a highly differentiated sales-type tax, such as the California sales tax or EU VAT, may be efficient, at least if measured elasticities are representative of elasticities at rates close to zero. While the California results likely have fewer sources of potential endogeneity than the EU findings, some likely remain, and we acknowledge that the relatively small rate changes in California may limit the external validity of our estimates. We therefore view both sets of estimates as tentative, and aimed at encouraging further exploration of XETI in better-identified settings. We also have not considered the more realistic but less readily estimable cases in which government faces enforcement costs and taxpayers incur transferable compliance costs.

4 Conclusion

Prior formal explorations of the optimality of consumption taxes have relied on relatively extreme sets of assumptions regarding tax avoidance opportunities, such as the assumption in ? that the consumption tax cannot be avoided. We generalize these findings and show the range of cases under which it is optimal to impose sales or VAT-type consumption taxes in addition to a labor income tax.

These findings raise the natural further question of the optimal mix of consumption and income taxes. While we find evidence that sales taxes should optimally be non-zero at observed real-world elasticities, it does not necessarily follow that they should be imposed at observed rates. Answering these questions will require further empirical investigation not only of the cross-elasticity of taxable income with respect to the sales tax, but also the cross-elasticity of taxable sales with respect to the income tax. We also have left aside other

potentially important considerations, such as the interactions of income and consumption taxes with any existing business tax.

Although they present econometric challenges, these further issues warrant careful study because of their sweeping potential implications, as Kaplow (2008) describes. For example, in Kaplow & Shavell (1994), the efficiency of taxation over redistributive legal rules turns on an assumption that the cross-elasticity of real labor responses to a legal rule will always equal or exceed the own-elasticity of labor to a labor income tax. Our results illustrate the importance of further considering the possibility of both real and avoidance responses. We do not mean to suggest that our analysis represents the only basis for determining the optimal mix of income and consumption taxes. What we offer is another possible theoretical basis for departing from exclusive reliance on the income tax.

Acknowledgments

The authors are grateful for helpful comments gathered during presentations at Indiana University-Bloomington, the American Law & Economics Association Annual Meeting, ComplianceNet, the Conference on Empirical Legal Studies, and the Annual Meeting of the National Tax Association.

Appendix

A Model extensions

Case 3. Avoidance cost includes component that is transferable to the government. So far we have assumed that taxpayers can risklessly reduce tax, albeit at some cost. In addition, as ? emphasizes, optimality results often depend on whether behavioral changes in response to tax are pure deadweight loss or instead comprise transfers between agents. We attempt to capture both these features by now explicitly accounting in our model for the fact that taxpayers may incur a risk of sanction, such that some avoidance cost might be transferable to the government. Specifically, assume government imposes a transferable penalty, F , (i.e., a fine) for tax avoidance. Taxpayer chooses A_Y (and A_X) to account for the penalty cost, $F_Y(\cdot)$ (and $F_X(\cdot)$), which is some probability-weighted cost of being fined for avoidance. Assume that penalty cost functions have the following form:

$$F_X(A_X, t_X) = F \cdot \alpha_X(A_X) \cdot t_X \cdot A_X,$$

$$F_Y(A_Y, t_Y) = F \cdot \alpha_Y(A_Y) \cdot t_Y \cdot A_Y,$$

where $\alpha_X(A_X)$ and $\alpha_Y(A_Y)$ are the probability functions of being caught for sales- and wage-tax avoidance correspondingly. That is, the penalty cost is proportional to the fine, F , to the probability of being caught, and to the avoided tax ($t_X \cdot A_X$ or $t_Y \cdot A_Y$). See ? for more detailed discussion of standard models of tax avoidance under a penalty regime. For simplicity we will build on the model of Case 2, but as will be clear these results also carry straightforwardly to Case 1.

Define $f_X(A_X) = F_X(A_X, t_X)/t_X$ and $f_Y(A_Y) = F_Y(A_Y, t_Y)/t_Y$. Then, the taxpayer problem should be reformulated as

$$\max_{(X, Y, A_X, A_Y)} U(G, H(X, Y)) \tag{A.1}$$

$$s.t. (1+t_X)(X-A_X)+A_X+t_X f_X(A_X)+C_X(A_X, A_Y) = (1-t_Y)(Y-A_Y)+A_Y-t_Y f_Y(A_Y)-C_Y(A_X, A_Y). \tag{A.2}$$

The government problem is to choose two tax rates (t_X, t_Y) to maximize individual utility subject to the government budget constraint, which is now

$$G = t_X(X - A_X + f_X(A_X)) + t_Y(Y - A_Y + f_Y(A_Y)) \tag{A.3}$$

The optimal amounts of avoidance, $A_X^*(t_X, t_Y)$ and $A_Y^*(t_X, t_Y)$, are now determined by the FOCs:

$$t_X \left(1 - \frac{\partial f_X(A_X^*)}{\partial A_X}\right) = \frac{\partial}{\partial A_X} [C_X(A_X^*, A_Y^*) + C_Y(A_X^*, A_Y^*)], \quad (\text{A.4})$$

$$t_Y \left(1 - \frac{\partial f_Y(A_Y^*)}{\partial A_Y}\right) = \frac{\partial}{\partial A_Y} [C_X(A_X^*, A_Y^*) + C_Y(A_X^*, A_Y^*)]. \quad (\text{A.5})$$

Now we can following similar solution steps as in Case 1. Specifically, equations (8), (9), and (10) are still valid here. The difference arises in equations (11) and (12), which now become

$$\frac{\partial G}{\partial t_X} = X^* - A_X^* + t_X \left(\frac{\partial X^*}{\partial t_X} - \frac{\partial A_X^*}{\partial t_X} + \frac{\partial f_X(A_X^*)}{\partial t_X} \right) + t_Y \left(\frac{\partial Y^*}{\partial t_X} - \frac{\partial A_Y^*}{\partial t_X} + \frac{\partial f_Y(A_Y^*)}{\partial t_X} \right), \quad (\text{A.6})$$

$$\frac{\partial G}{\partial t_Y} = Y^* - A_Y^* + t_Y \left(\frac{\partial Y^*}{\partial t_Y} - \frac{\partial A_Y^*}{\partial t_Y} + \frac{\partial f_Y(A_Y^*)}{\partial t_Y} \right) + t_X \left(\frac{\partial X^*}{\partial t_Y} - \frac{\partial A_X^*}{\partial t_Y} + \frac{\partial f_X(A_X^*)}{\partial t_Y} \right). \quad (\text{A.7})$$

By the envelop theorem, we have

$$\frac{\partial \tilde{H}}{\partial t_X} = -\frac{\lambda}{U'_H} (X^* - A_X^* + f_X(A_X^*)), \quad (\text{A.8})$$

$$\frac{\partial \tilde{H}}{\partial t_Y} = -\frac{\lambda}{U'_H} (Y^* - A_Y^* + f_Y(A_Y^*)), \quad (\text{A.9})$$

where λ is the Lagrange multiplier in the individual problem.

Using equations (9), (10) and equations (A.6) - (A.9), we obtain the conditions for the optimal tax rates, (t_X, t_Y) :

$$\frac{t_X \frac{\partial(X^* - A_X^* + f_X(A_X^*))}{\partial t_X} + t_Y \frac{\partial(Y^* - A_Y^* + f_Y(A_Y^*))}{\partial t_X}}{X^* - A_X^* + f_X(A_X^*)} = \frac{\lambda - U'_G}{U'_G}, \quad (\text{A.10})$$

$$\frac{t_Y \frac{\partial(Y^* - A_Y^* + f_Y(A_Y^*))}{\partial t_Y} + t_X \frac{\partial(X^* - A_X^* + f_X(A_X^*))}{\partial t_Y}}{Y^* - A_Y^* + f_Y(A_Y^*)} = \frac{\lambda - U'_G}{U'_G}. \quad (\text{A.11})$$

If we introduce (let us call it) *adjusted* taxable income, $\tilde{Z}_Y = Y - A_Y + f_Y(A_Y^*)$, and *adjusted* taxable consumption, $\tilde{Z}_X = X - A_X + f_X(A_X^*)$, we can rewrite the conditions for the optimal tax rates, (t_X, t_Y) , as

$$\frac{t_X}{1 - t_X} \left[E(\tilde{Z}_X, 1 - t_X) + \frac{t_Y \tilde{Z}_Y}{t_X \tilde{Z}_X} E(\tilde{Z}_Y, 1 - t_X) \right] = \frac{U'_G - \lambda}{U'_G}, \quad (\text{A.12})$$

$$\frac{t_Y}{1-t_Y} \left[E(\tilde{Z}_Y, 1-t_Y) + \frac{t_X \tilde{Z}_X}{t_Y Z_Y} E(\tilde{Z}_X, 1-t_Y) \right] = \frac{U'_G - \lambda}{U'_G}. \quad (\text{A.13})$$

Thus, we see that once we change taxable income to adjusted taxable income and taxable sales to adjusted taxable sales, the conditions for the optimal tax rates, (t_X, t_Y) , look exactly the same as in Case 2. Hence, Proposition 2 and Corollary 2 readily extend here.

Case 4. The government has bureaucratic (administrative) costs. Another unrealistic component of the model thus far is that we have ignored the likelihood that adopting an additional collection method will result in added administrative costs. If administrative costs from collecting taxes on sales and on wages are fixed, the optimal conditions could be derived but are not very informative. However the intuitions are straightforward, as discussed in ? and in the analogous case for use of non-rate instruments in Slemrod and Gillitzer (2013:117). A sales tax would only be welfare improving if the net utility gains from adding the tax exceed the fixed costs of administration.

It is likely, though, that administrative costs are not wholly fixed but instead have a relationship to rates and base. Let us return to the assumptions of Case 2 and additionally suppose that the government has bureaucratic (administrative) costs of collecting sales and wage taxes, $B_X(A_X, X)$ and $B_Y(A_Y, Y)$. For tractability, we assume these are each continuous and increasing in both arguments' functions, although we acknowledge that in reality administrative costs can often be discontinuous (?).

Under these assumptions, the taxpayer problem and its solution stay the same as in Case 2. The government problem is to choose two tax rates (t_X, t_Y) to maximize individual utility subject to the government budget constraint, which is now

$$G = t_X(X - A_X) - B_X(A_X, X) + t_Y(Y - A_Y) - B_Y(A_Y, Y). \quad (\text{A.14})$$

Accounting for this modification, we can follow similar solution steps as in Case 1. Specifically, equations (20), (21), (8), (9), (10), (13), and (14) are still valid here. The difference arises in equations (11) and (12), which now become

$$\frac{\partial G}{\partial t_X} = X^* - A_X^* + t_X \left(\frac{\partial X^*}{\partial t_X} - \frac{\partial A_X^*}{\partial t_X} \right) + t_Y \left(\frac{\partial Y^*}{\partial t_X} - \frac{\partial A_Y^*}{\partial t_X} \right) - \frac{\partial(B_X(A_X^*, X^*) + B_Y(A_Y^*, Y^*))}{\partial t_X}, \quad (\text{A.15})$$

$$\frac{\partial G}{\partial t_Y} = Y^* - A_Y^* + t_Y \left(\frac{\partial Y^*}{\partial t_Y} - \frac{\partial A_Y^*}{\partial t_Y} \right) + t_X \left(\frac{\partial X^*}{\partial t_Y} - \frac{\partial A_X^*}{\partial t_Y} \right) - \frac{\partial(B_X(A_X^*, X^*) + B_Y(A_Y^*, Y^*))}{\partial t_Y}. \quad (\text{A.16})$$

Combining all the equations mentioned above, we obtain the conditions for the optimal tax rates, (t_X, t_Y) :

$$\frac{t_X \frac{\partial(X^* - A_X^*)}{\partial t_X} + t_Y \frac{\partial(Y^* - A_Y^*)}{\partial t_X} - \frac{\partial(B_X(A_X^*, X^*) + B_Y(A_Y^*, Y^*))}{\partial t_X}}{X^* - A_X^*} = \frac{\lambda - U'_G}{U'_G}, \quad (\text{A.17})$$

$$\frac{t_Y \frac{\partial(Y^* - A_Y^*)}{\partial t_Y} + t_X \frac{\partial(X^* - A_X^*)}{\partial t_Y} - \frac{\partial(B_X(A_X^*, X^*) + B_Y(A_Y^*, Y^*))}{\partial t_Y}}{Y^* - A_Y^*} = \frac{\lambda - U'_G}{U'_G}. \quad (\text{A.18})$$

Applying the notation for taxable income, $Z_Y = Y - A_Y$, and taxable consumption, $Z_X = X - A_X$, we can rewrite the conditions for the optimal tax rates, (t_X, t_Y) , as

$$\frac{t_X}{1 - t_X} \left[E(Z_X, 1 - t_X) + \frac{t_Y Z_Y}{t_X Z_X} E(Z_Y, 1 - t_X) - \frac{B_X + B_Y}{t_X Z_X} E(B_X + B_Y, 1 - t_X) \right] = \frac{U'_G - \lambda}{U'_G}, \quad (\text{A.19})$$

$$\frac{t_Y}{1 - t_Y} \left[E(Z_Y, 1 - t_Y) + \frac{t_X Z_X}{t_Y Z_Y} E(Z_X, 1 - t_Y) - \frac{B_X + B_Y}{t_Y Z_Y} E(B_X + B_Y, 1 - t_Y) \right] = \frac{U'_G - \lambda}{U'_G}, \quad (\text{A.20})$$

where for brevity B_X denotes $B_X(A_X^*, X^*)$ and B_Y denotes $B_Y(A_Y^*, Y^*)$.

Proposition 3. At an optimum, the optimal wage tax rate and the optimal sales tax rate are determined by equations (A.19) and (A.20).

As we see, compared to equation (26) for the optimal tax rate t_X in Case 2, in this case with bureaucratic cost we have an additional (third) term in the left-hand side of equation (A.19), which reflects that a higher tax rate likely raises bureaucratic costs. Similar considerations apply to equation (A.20) for optimal tax rate t_Y . Note that $E(B_X + B_Y, 1 - t_X)$ and $E(B_X + B_Y, 1 - t_Y)$ are likely to be negative.

Again, we would like now to explore when it would be optimal to have the sales tax rate set non-zero. Following the same logic as in case 1, we obtain that for the sales tax rate to be non-zero it is necessary that

$$E(Z_Y, 1 - t_X)|_{t_X=0} < \frac{Z_X}{(1-t_Y)Z_Y} \left\{ E(Z_Y, 1 - t_Y)|_{t_X=0} + \frac{1-t_Y}{t_Y} \frac{B_X+B_Y}{Z_X} E(B_X + B_Y, 1 - t_X) \right\} \Big|_{t_X=0} - \frac{B_X+B_Y}{t_Y Z_Y} E(B_X + B_Y, 1 - t_Y) \Big|_{t_X=0}, \quad (\text{A.21})$$

where again for brevity B_X denotes $B_X(A_X^*, X^*)$ and B_Y denotes $B_Y(A_Y^*, Y^*)$ and t_Y is determined by equation (27) with $t_X = 0$.

Corollary 3. The necessary condition for the sales tax rate to be positive is given

by equation (A.21).

Compared to Case 2 (see equation (28)), to determine whether the sales tax rate should be greater than zero, the cross-elasticity of taxable income w.r.t. $(1 - t_X)$, $E(Z_Y, 1 - t_X)|_{t_X=0}$, should not be simply compared to the elasticity of taxable income, $E(Z_Y, 1 - t_Y)|_{t_X=0}$, multiplied by $\frac{Z_X}{(1-t_Y)Z_Y}$. The first term on the right-hand side for the two equations is the same, so that the factors we described in Case 2 continue to hold with administrative costs. The second and third right-hand side terms here capture the (weighted) incremental combined bureaucratic costs of increasing the sales and wage taxes, respectively. It is more difficult to rule out a sales tax when the second term is larger than the third. In short, then, we show that when the elasticity of combined wage- and sales-tax bureaucratic cost to (one minus the sales tax) is larger than its elasticity to (one minus the wage tax), a sales tax becomes relatively more desirable.

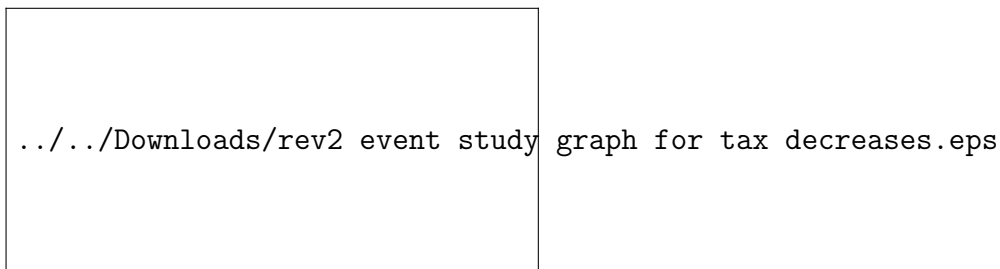


Figure 4: Sales Tax Rate Elections and Income Tax Liability: Event Study

Notes: This figure summarizes an event-study type regression examining the effects of sales-tax election results on income tax liability. The coefficients reported are for interaction terms between an indicator for an election in which tax rates declined and the number of years the observation falls from the year of the election. The outcome variable is log CA total income tax liability, aggregated to the zip-code level. Observations are limited to renewal elections. Zip-years with extreme changes in income-tax liability are omitted. Since the treatment is a tax decrease, a positive point estimates in post-treatment years is consistent with a small negative correlation between sales tax rates and income tax liability reported.

B Additional Empirical Results

Table B.1: **Cross-Elasticity of Income Tax Liabilities w.r.t. Sales Tax Rates Using Quasi-Random Vote Threshold**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	CA Tax	CA Tax (Weighted) Per Filer	CA Tax (Weighted) Per Filer	CA Tax (Weighted) Per Filer	U.S. Tax	U.S. Tax (Weighted) Per Filer	U.S. Tax (Weighted) Per Filer	U.S. Tax (Weighted) Per Filer
$Post \times \Delta Sales$	-2.549 (2.384)	-4.153 (2.305)	-2.111 (1.882)	-3.959 (2.093)	-2.764 (1.319)	-3.535 (1.284)	-2.030 (1.243)	-2.730 (1.319)
Observations	8,509	8,509	8,509	8,509	4,324	4,324	4,324	4,324
R-squared	0.355	0.558	0.339	0.581	0.729	0.822	0.675	0.772
One-sided p-Value at ETI:								
.12	.842	.929	.881	.936	.969	.975	.924	.950
.4	.879	.951	.915	.957	.983	.986	.959	.972
1.2	.949	.986	.974	.988	.998	.998	.996	.996

Notes: This table reports the cross-elasticity of reported California and U.S. income tax with respect to changes in one minus the local sales tax rate. Fixed-effects OLS panel estimates with zip code as panel variable. Standard errors clustered by city in (parenthesis). Weights by number of filing households in zip code. $Post \times \Delta Sales$ is an interaction term between an indicator for the period after an election and a continuous measure of the difference in the sales-tax rate between the year after and the year before the election. One-sided p-Value panel reports test statistic for a one-sided test that the estimated coefficient on $Post \times \Delta Sales$ is less than the critical value implied by the listed ETI. Observations are limited to years surrounding elections where the vote share in favor of passage fell between fifty percent and two-thirds of the vote. Zip-years with extreme changes in income-tax liability are omitted.

Table B.2: **Cross-Elasticity of Income Tax w.r.t. Sales Tax Rates, Using Quasi-Random Vote Threshold, Renewal Elections Only**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	CA Tax	CA Tax (Weighted)	CA Tax Per Filer	CA Tax Per Filer (Weighted)	Federal Tax	Fed Tax (Weighted)	Federal Tax Per Filer	Fed Tax Per Filer (Weighted)
$Post \times \Delta Sales$	-2.354 (2.519)	-3.072 (2.203)	-2.295 (2.096)	-3.071 (2.129)	-2.729* (1.555)	-3.079* (1.661)	-1.755 (1.345)	-2.305 (1.506)
Observations	4,384	4,384	4,384	4,384	2,799	2,799	2,799	2,799
R-squared	0.339	0.666	0.306	0.579	0.794	0.862	0.780	0.846
One-sided P value at ETI:								
.12	.873	.969	.888	.976	.987	.998	.962	.986
.4	.906	.980	.924	.985	.994	.999	.982	.994
1.2	.965	.995	.980	.997	.999	.999	.999	.999

Notes: This table reports the cross-elasticity of reported California and U.S. income tax with respect to changes in one minus the local sales tax rate. Fixed-effects OLS panel estimates with zip code as panel variable. Standard errors clustered by city in (parenthesis). Weights by number of filing households in zip code. $Post \times \Delta Sales$ is an interaction term between an indicator for the period after an election and a continuous measure of the difference in the sales-tax rate between the year after and the year before the election. One-sided p-Value panel reports test statistic for a one-sided test that the estimated coefficient on $Post \times \Delta Sales$ is less than the critical value implied by the listed ETI. Observations are limited to years surrounding elections where the vote share in favor of passage fell between fifty percent and two-thirds of the vote.

C Referee Appendix

This Appendix reports additional results described but not tabulated in the main text.

1. Main text page 20 note 11 reports that regression results are qualitatively similar to those reported if we winsorize tax liability data to values that would correspond to the fifth and ninety-fifth percentiles of year-over-year changes. Table C.1 tabulates those results.

Table C.1: Cross-Elasticity of Income Tax Liabilities w.r.t. Sales Tax Rates, Winsorized Tax Liability Values

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	CA Tax	CA Tax (Weighted)	CA Tax Per Filer	CA Tax Per Filer (Weighted)	U.S. Tax	U.S. Tax (Weighted)	U.S. Tax Per Filer	U.S. Tax Per Filer (Weighted)
$Post \times \Delta Sales$	-6.482 (4.933)	-13.35* (6.990)	-7.601* (4.477)	-13.24* (6.959)	0.637 (2.661)	-3.840** (1.706)	0.754 (2.070)	-3.008** (1.522)
Observations	9,657	9,657	9,657	9,657	6,136	6,136	6,136	6,136
R-squared	0.165	0.299	0.108	0.251	0.389	0.492	0.334	0.587

Notes: This table reports the cross-elasticity of reported California and U.S. income tax with respect to changes in one minus the local sales tax rate. Fixed-effects OLS panel estimates with zip code as panel variable. Standard errors clustered by city in (parenthesis). Weights by number of filing households in zip code. $Post \times \Delta Sales$ is an interaction term between an indicator for the period after an election and a continuous measure of the difference in the sales-tax rate between the year after and the year before the election. Observations are limited to years surrounding elections where the vote share in favor of passage fell between fifty percent and two-thirds of the vote. Zip-years with extreme changes in income-tax liability are included but winsorized.

