# Is Taxing the Externality Enough?*

Christopher Conlon
NYU Stern School of Business
cconlon@stern.nyu.edu

Nirupama L. Rao
University of Michigan Ross School of Business
nirurao@umich.edu

November 4, 2026

**Abstract**

A central prescription in public economics is to correct externalities by taxing the externality-generating attribute: for example, ethanol content for alcohol, THC for cannabis, sugar for sweetened beverages, or $CO_2$ emissions for vehicles. This prescription presumes that every externality-generating product can be taxed on its own damage. When some substitutes are untaxed or taxed on a coarser base, as in every market we examine, the marginal external damage from a product depends not only on its externality content, but also on the external damage generated by the substitutes its buyers move to. We illustrate the conditions under which a simple attribute-based tax is likely to be sufficient and characterize when it fails. We then present case studies for distilled spirits and automobiles that quantify the welfare cost of ignoring within-market substitution.

**JEL Classification:** H21, H23, L13, L66.
**Keywords:** Excise tax, Pigouvian tax, externality correction, differentiated products, diversion ratio, corrective taxation.

*Preliminary draft. Comments welcome.

## 1 Introduction

Governments routinely levy corrective excise taxes on products that generate negative externalities. A standard prescription, dating to **?**, is to levy the tax in proportion to the externality-generating attribute, whether ethanol in alcoholic beverages, THC in cannabis, grams of sugar in sweetened drinks, or $CO_2$ per mile for vehicles. When the government seeks to raise revenue in addition to correcting the externality, **?** shows that the optimal tax has a Ramsey component that depends on demand elasticities, with less elastic goods taxed more heavily, and a Pigouvian correction equal to marginal damage. **?** shows that this "additivity property" holds under relatively general conditions.

The additivity result does not require that demands be independent across goods. It does require that every externality-generating good can be taxed directly on its own damage. In the classical treatment, and much applied work since, the externality-generating good is modeled as a broad product category or composite good such as "alcohol" or "cigarettes" or "sugar-sweetened beverages", with the corrective tax set according to marginal external damage.[1] Consumers, however, do not purchase composite goods but *products*, which in most of these markets are highly differentiated. A bottle of whiskey contains some amount of externality-generating ethanol, but also differs in quality, branding, and flavor, attributes that a planner correcting the externality would have no reason to tax unless the planner also sought to raise revenue.

It is well understood that products within a category can differ substantially in the external damage they generate. Beverages differ in sugar, cannabis products in THC, and vehicles in $CO_2$ per mile. A social planner maximizing a standard welfare function would set corrective taxes that vary with these differences, but in practice governments often rely on coarser instruments. As we show later in **??**, most jurisdictions tax sugar-sweetened beverages by volume rather than sugar content, and while the federal government taxes spirits by ethanol content, most states tax distilled spirits by volume alone.

This paper asks how a standard corrective-tax prescription changes when consumers choose among differentiated products and some products are untaxed or taxed on a coarser basis. Once products are substitutes, the reduction in external damage from raising the price of one product depends not only on that product's externality content but also on where its buyers go and how much damage their alternatives generate. Consider two 80-proof spirits,

[1] **?** state the zero-cross-elasticity assumption explicitly; **?** provide a modern treatment of the single-good corrective tax. **?** allow substitution from the taxed good to untaxed harmful goods, but at the level of broad categories rather than products within a market; the composite-good treatment, as in **??** on cigarettes, remains standard in the applied literature on tobacco, alcohol, and sugary-drink taxes reviewed by **?**. **?** and **?** study indirect remedies when the externality-generating good cannot itself be taxed, with clean related goods.

both taxed on ethanol content. If raising the price of one sends most of its marginal buyers to the other, the price increase changes overall ethanol consumption very little. If instead its marginal buyers leave the spirits market entirely, the same price increase reduces ethanol consumption nearly one for one. The same tax applied to the same ethanol content can have very different effects on aggregate external damage depending on substitution patterns. When a product's substitutes are *dirtier* than the product itself, raising its price can increase total external damage, so the corrective policy may call for a subsidy rather than a tax.

None of this conflicts with additivity. If every product can be taxed directly on its own external damage, substitution among taxed products does not enter the Pigouvian correction. The issue arises when coverage is incomplete. Some substitutes may be untaxed, as with the illicit cannabis market or the used vehicle fleet. Others may be taxed on a base that does not track their external damage, as with beer and wine, which are taxed by volume rather than alcohol content in the U.S. The markets we examine here and many others have some form of incomplete coverage.

We begin by writing down a standard welfare function and deriving the tax that maximizes it in two polar cases. In the first case, the planner chooses the tax on every product in the market, and the welfare-maximizing tax on each product depends only on its own damage relative to the outside option, as additivity implies. In the second case, the tax on a single product is adjusted while the others are held fixed. The tax then depends on the product's damage relative to the diversion-weighted average damage of its substitutes, which we call the substitution-adjusted marginal externality.

We next ask when a tax schedule based only on a product's own externality-generating attribute can reproduce the welfare-maximizing tax. This depends on how well the average damage of a product's substitutes can be approximated by an affine function of its own damage, which we capture with a simple regression. An attribute-based schedule performs well either when substitutes' damage is nearly the same for every product (the degenerate case) or when the relationship between average substitute damage and own damage is approximately affine, as it is when substitution stays within categories of similar products.

We then use estimated substitution patterns in two markets, distilled spirits and automobiles, to assess how well taxes based on a product's own external damage approximate the welfare-maximizing tax. For distilled spirits, we use the estimated demand system of **?**, while for automobiles we use a large second-choice survey together with the demand system of **?**.

The two applications generate quite different patterns. In spirits, nearly every product has similar ethanol content and the damage of a product's substitutes is essentially flat in its own damage. When the tax on a single spirit is adjusted, the welfare-maximizing

tax rises with ethanol content at nearly the full Pigouvian rate. Taxes on volume and on ethanol content deliver nearly identical welfare. What the attribute cannot capture is the exit margin. When the price of a budget product rises, more than 60 percent of its marginal buyers leave the spirits market, compared with only 40 percent for premium products, a difference unrelated to proof that a tax on ethanol content cannot reach.

In automobiles, emissions vary tenfold across the new-car fleet, but substitution follows well-defined categories, so a vehicle's own emissions predict its substitutes' emissions reasonably well across the fleet. A vehicle that emits 100 g/mile more than another sends its marginal buyers to substitutes that emit about 68 g/mile more, on average. As a result the welfare-maximizing tax on a single vehicle rises with its emissions at about a third of the naive Pigouvian rate. When taxes are levied within particular categories, however, own emissions can be a poor guide to the welfare-maximizing tax. The Ford F-Series, the best-selling vehicle in the United States and one of the highest-emitting, carries a welfare-maximizing tax of about $18,000 when every vehicle is taxed, but a small subsidy of about $1,400 when only its own tax is adjusted, because its buyers would otherwise choose even dirtier pickups. For clean-vehicle rebates restricted to the plug-in vehicles that qualified for rebates in 2015, a schedule written on emissions attains 68 percent of the attainable welfare gain at a high cost of public funds, while one written on the substitution-adjusted damage attains 95 percent. Our point is not that we can or should base taxes directly on substitution-adjusted damage, but rather that, in practical applications with incomplete coverage, we should be careful to pay attention to substitutes (particularly untaxed ones).

Throughout, statements about the "optimal" or "welfare-maximizing" tax refer to the tax that maximizes the specific objective in **??**: consumer surplus plus tax revenue, net of monetized external damage, with market power and a revenue motive treated as extensions. We characterize what such a tax looks like and how far simple schedules depart from it, but take no position on whether that objective is the right one for any particular government. **??** presents the framework, **????** apply it to distilled spirits and automobiles, **??** considers cannabis, sugar-sweetened beverages, and nicotine products, and **??** discusses practical lessons.

## 2 Framework

We begin with a simple setting meant to illustrate that the forces we describe arise in even the most basic environment. We first present the baseline under perfect competition and without a revenue motive, then allow for markups and a revenue requirement in **??**.

### 2.1 Baseline Model

Consumers have quasilinear preferences and make a discrete choice among $j = 1, \ldots, J$ differentiated products, and an outside option, indexed by $j = 0$. Let $q_j(\mathbf{p})$ denote the quantity of product $j$ demanded at prices $\mathbf{p}$. We fix the size of the market at $N$, so that $q_0 = N - \sum_{j=1}^{J} q_j$.

Each product, potentially including the outside option, carries an externality-generating attribute $e_j$, such as ethanol, sugar, THC or emissions. We assume that the externality is *atmospheric*. That is, total external damage depends only on the aggregate quantities consumed, not which consumers buy which products: $E = \sum_{j=0}^{J} e_j q_j$. Substituting for $q_0$, external damages can be written as $E = e_0 N + \sum_{j=1}^{J} (e_j - e_0) q_j$. The first term is constant, so the relevant externality content of inside product $j$ is its damage relative to the outside option. We define:

$$\tilde{e}_j \equiv e_j - e_0.$$

We also assume that markets are *perfectly competitive* and marginal costs are constant, so that prices $p_j$ are equal to a constant marginal cost $c_j$ plus a per-unit tax $\tau_j$, $p_j = c_j + \tau_j$, which means that taxes are fully passed through into prices $\frac{\partial p_j}{\partial \tau_j} = 1$ and $\frac{\partial p_k}{\partial \tau_j} = 0$ for $k \neq j$.

We make these strong assumptions not because we think they are necessarily an accurate depiction of how markets with externalities work in practice, but rather because we want to demonstrate that even in our highly simplified world, it is important to consider how consumers substitute across products when calculating the marginal external damage and corrective taxes.[2]

The government (or "social planner") chooses per-unit taxes (or subsidies) $\tau_j$ on the $J$ inside products to maximize the sum of consumer surplus and tax revenue, net of external damage.[3]

$$W \;=\; CS(\mathbf{p}) \;+\; \textstyle\sum_{j=1}^{J} \tau_j q_j \;-\; \lambda_e \left( \sum_{j=0}^{J} e_j q_j \right). \tag{1}$$

[2]For example, we assume perfect competition in order to make clear that substitution across products matters when addressing externalities even in the absence of market power. The atmospheric assumption rules out externalities that depend on who consumes the product or on non-linear interactions. In reality, young smokers lose more quality-adjusted life years than old smokers, or heavy drinkers may generate more external damage per unit of ethanol than light drinkers because the external damage function is convex rather than linear (see **?** for the latter). If the externality is non-atmospheric, substitution patterns will still mediate the ultimate marginal external damage but with the added complication that differences in substitution patterns across individuals with different propensities to create the external damage will also matter (**?**).

[3]Including a revenue target adds a Ramsey component (scaled by a cost of public funds $\lambda_r > 0$). **??** derives the extended first-order conditions.

where $\lambda_e$ is the dollar value of a unit of external damage. Because firms are perfectly competitive, producer surplus is zero in this environment.

Differentiating with respect to $\tau_j$ gives[4]:

$$\frac{dW}{d\tau_j} = \sum_{k=1}^{J} \Big(\tau_k - \lambda_e \, \tilde{e}_k\Big) \frac{\partial q_k}{\partial p_j}, \qquad \tilde{e}_k \equiv e_k - e_0. \tag{2}$$

The first-order condition has a simple interpretation. Increasing $\tau_j$ changes demand for product $j$ and its substitutes. The welfare effect of each quantity response is determined by the wedge $(\tau_k - \lambda_e \tilde{e}_k)$, which measures how far the existing tax on product $k$ is from the money-metric external damage generated by $k$, measured relative to the outside option.

To understand the role of substitution, we consider two polar cases for **??**. One where all inside products are taxed and a second where the planner can only adjust the tax on a single product.

**Case A: All dirty products are fully taxed.** Suppose all dirty products are taxed, and the rates are set jointly. Then we set every wedge in **??** to zero, yielding:

$$\tau_j^\star = \lambda_e \, (e_j - e_0) \qquad \text{for every } j. \tag{3}$$

The planner levies a tax on each product $j$ based on its external damage relative to the outside option $(e_j - e_0)$. Substitution is irrelevant because each destination product is already taxed for its own external damage. This is the classic *additivity* result of **??**, where the price of each good reflects its full marginal social cost.

**Case B: Change the rate on one product.** Suppose instead we fix the tax rates of substitutes $\bar{\tau}_k$, and consider how the planner would set $\tau_j$ holding all else fixed. Let

$$D_{j \to k}(p) \equiv \frac{\partial q_k / \partial p_j}{|\partial q_j / \partial p_j|}.$$

This diversion ratio denotes the share of consumers leaving product $j$ who substitute to product $k$.[5] The set of destinations includes the outside option, so the diversion ratios sum

[4]We provide a complete step-by-step derivation in **??**, where we exploit the quasilinearity assumption to cancel the consumer-surplus derivative $\frac{\partial CS}{\partial p_j} = -q_j$ against the mechanical revenue term $q_j$. We also exploit perfect competition to swap $\frac{\partial q_k}{\partial \tau_j} = \frac{\partial q_k}{\partial p_j}$, and the discrete choice assumption to normalize $e_j - e_0$ relative to the outside option and change the starting point of the summation from $j = 0$ to $j = 1$.

[5]Mathematically, the diversion ratio is defined as: $D_{j \to k}(p) = (\partial q_k / \partial p_j)/|\partial q_j / \partial p_j|$ or using the own and cross-elasticities as: $D_{j \to k}(p) = \frac{\mathcal{E}_{jk}}{|\mathcal{E}_{jj}|} \cdot \frac{p_j}{p_k}$. The diversion ratio is similar to the cross-elasticity in that it

to one across all $k \neq j$.

Setting the first-order condition **??** equal to zero, rewriting in terms of diversion ratios (details in **??**), and solving for the single tax rate $\tau_j$ gives:

$$\tau_j^\star(\mathbf{p}) \;=\; \lambda_e \underbrace{\left(e_j - \bar{e}_j(\mathbf{p})\right)}_{\equiv M_j^e(\mathbf{p})} \;+\; \sum_{k \neq j} D_{j \to k}(\mathbf{p})\, \bar{\tau}_k, \tag{4}$$

where $\bar{e}_j(\mathbf{p}) = \sum_{k \neq j} D_{j \to k}(\mathbf{p}) \cdot e_k$ is the diversion-weighted mean external damage of $j$'s substitutes. We refer to $M_j^e(\mathbf{p})$ as the **substitution-adjusted marginal externality**: it measures how the marginal damage of product $j$ compares to the alternatives its marginal buyers move to when the price of product $j$ rises.[6] Much of the empirical analysis in **??** asks how well this substitution-adjusted marginal externality $M_j^e$ can be approximated by an affine function of the externality-generating attribute $e_j$.[7]

Note that $\tau_j^\star(\mathbf{p})$ in **??** is not the jointly optimal vector of tax rates. It is the welfare-maximizing tax rate on product $j$ when the planner can only adjust $\tau_j$ and all other rates must remain fixed at $\bar{\tau}_k$. Unlike the case where all tax rates are set jointly and prices reflect the marginal damage of each product, here raising $\tau_j$ may lead to leakage through substitution to products whose prices do not fully reflect their external damages. The relevant externality effect of raising $\tau_j$ must also account for this leakage. $M_j^e(\mathbf{p})$ reflects this leakage by netting out the external damage of options chosen by $j$'s marginal buyers, weighted by diversion shares.

Both $M_j^e$ and $\bar{e}_j$ depend on the prices of substitutes $\mathbf{p}$ and **??** is an implicit solution because they move with $\tau_j$ and $\bar{\tau}_k$. The tax in **??** also depends on the diversion-weighted average of its substitutes' taxes $\bar{\tau}_k$ (because the planner trades off revenue against consumer surplus), with more heavily taxed substitutes pushing up $\tau_j^\star$.

### 2.2 Comparing Complete Coverage and Single-Rate Reform

We expect that most realistic examples will not live at the extremes of either Case A or B, but rather somewhere in the middle. In many cases, some but not all products will be reachable with taxes, while in others, products will be reachable only with imperfect tax instruments that hit some attribute correlated with $e_j$, but not $e_j$ itself. Nonetheless, these two polar cases are useful because they provide easy-to-characterize solutions for the corrective tax.

quantifies substitution across products, but has the advantage that it sums to one across all substitutes.

[6] Because diversion ratios sum to one, both $M_j^e$ and $\bar{e}_j$ can be interpreted as expectations (or weighted averages).

[7] **??** derive the same object from the first-order conditions of a planner who sets product-specific prices subject to revenue and externality constraints; the full pricing rule there also includes Ramsey and market-power components.

The general logic of **????** is that the welfare-maximizing corrective tax is levied on the external damage of a product *relative* to the untaxed or mis-taxed alternative. Under Case A, all inside goods are taxed, so the relevant alternative is the outside option and the tax depends on $(e_j - e_0)$. Under Case B, only $\tau_j$ is adjusted, so the relevant alternative is the diversion-weighted set of products chosen by product $j$'s marginal buyers, and the tax depends on $(e_j - \bar{e}_j)$.[8]

In both cases, a dirtier under-taxed alternative means that increasing the tax on $j$ reduces the overall external damage $E$ by less, so the optimal corrective tax $\tau_j^\star$ is lower. In some cases, if a product is cleaner than its (under-taxed) substitutes, increasing its tax may push consumers to dirtier alternatives and actually increase the overall level of external damage. The optimal corrective tax may even be negative (a subsidy).

Real tax systems commonly feature two main departures from the full coverage benchmark of Case A. First, some externality-generating products may not be taxed. Second, some taxed products may be taxed based on attributes other than $e_j$, such as volume, or price. In alcohol markets, for example, the federal government taxes the ethanol content of distilled spirits, in fact tracking $e_j$, but taxes beer and wine by volume (albeit at different rates). Most states also use volumetric taxes with different rates for beer, wine and spirits. Taxes based on volume do not vary across products within the same category that are sold in the same container size but differ in their alcohol content. Cannabis markets face both issues with the illicit market going wholly untaxed and legal products often taxed on price or quantity rather than THC content. For example, Alaska taxes cannabis by weight, Illinois uses a hybrid system with tax rates that vary by product type and potency, and Connecticut's current THC-based tax is scheduled to be replaced by a gross-receipts tax in October 2026.[9]

### 2.3 When Does the Conventional Wisdom Fail?

The conventional prescription is to tax each product on its "dirty" attribute $e_j$ alone (e.g. alcohol, THC, sugar, or $CO_2$ content), without accounting for the externalities incurred when consumers substitute to other dirty products. One goal of this article is to understand if and when such Pigouvian taxes are sufficient to address the externality, and when they fall short of the optimal corrective tax.

With that in mind, we consider whether any attribute tax of the "feebate" form $\tau_j =$

[8] This similarity also carries over to welfare changes. For Case A, when we introduce a small tax, the change in welfare is given by: $dW/d\tau_j = \lambda_e \left(e_j - e_0\right) |\partial q_j / \partial p_j|$. (Recall that Case A starts from fully corrected taxes $\bar{\tau}_k = \lambda_e (e_k - e_0)$.) For Case B, when $\bar{\tau}_k = 0$, the marginal change in welfare is $dW/d\tau_j = \lambda_e \left(e_j - \bar{e}_j\right) |\partial q_j / \partial p_j|$. See **??** for full results.

[9] See **?**.

$c_0 + c_1 e_j$ is able to replicate $\tau^\star$. For Case A, this is trivially true; to match $\tau^\star$ set $c_1 = \lambda_e$ and $c_0 = -\lambda_e e_0$ to produce **??** exactly. The intercept allows the feebate to account for the outside option, while a pure attribute tax, $\tau_j = c \cdot e_j$, is sufficient only when the outside option is clean, $e_0 = 0$.

For scenarios outside of Case A, a feebate only works if the optimal corrective tax is well described by an intercept and the attribute $e_j$, or more specifically, we need $\tau^\star$ to lie in the span of $[1, e_j]$. Consider the simple example of Case B with no other taxes, such that $\tau^\star = \lambda_e M_j^e = \lambda_e(e_j - \bar{e}_j)$. The term $e_j$ is already in the feebate's span. The question is whether the substitute term $\bar{e}_j$ is also spanned by the attribute and an intercept, which we can assess by regressing $\bar{e}_j$ on $e_j$:

$$\bar{e}_j \;=\; \alpha + \beta\, e_j + u_j. \tag{5}$$

When $\bar{e}_j$ is well approximated by an affine function of $e_j$, a feebate on the attribute itself, $\tau_j = c_0 + c_1 e_j$, can approximate the optimal tax, $\tau^\star$. The simplest case is the degenerate case, where $\mathrm{Var}(\bar{e}_j) \approx 0$. Average substitute damage is then nearly constant across products, even if their own damage differs. Most distilled spirits, for example, are 40% alcohol by volume, so conditional on remaining in the spirits market, marginal buyers tend to substitute toward products with similar external damage regardless of which spirits product they leave. A nearly constant $\bar{e}_j$ can also arise because there is very little variation in substitution $D_{j\to k}$ across products $j$ (including to the outside option), as in many simple parametric models.[10] A special case of degeneracy occurs when the outside option is as dirty as the average substitute, $\bar{e}_j = e_0$ for all $j$. Then the formulas for $\tau^\star$ in Cases A and B, **????**, coincide exactly. When substitution is almost entirely to the outside option ($D_{j\to 0} \approx 1$), again $\bar{e}_j \approx e_0$, so the Case B tax approaches the complete-coverage tax in Case A.

More generally, taxing the attribute may also be sufficient if $\bar{e}_j$ varies across products but is still well predicted by $e_j$. This can happen when substitution mainly occurs within a group of products that have similar external damage. For example, compact cars have much lower greenhouse gas emissions per mile than pickup trucks, but emissions are similar within each vehicle category. If compact-car buyers mostly substitute to other compact cars, and pickup buyers mostly substitute to other pickups, then own external damage $e_j$ will closely predict the average external damage of its substitutes, $\bar{e}_j$. Importantly, "well predicted" does not mean that $\bar{e}_j$ must be close in level to $e_j$. What matters is that the relationship between them is approximately affine, allowing the effect of substitute damage on the corrective tax

[10] For example, under CES $D_{j\to k} \propto q_k$ (independent of $j$). Likewise the IIA property of the plain multinomial logit gives $D_{j\to k} = \frac{q_k}{N - q_j}$ for all products, which is roughly independent of $j$ when $q_j \ll N$. Both cases lead to $\bar{e}_j \approx \sum_k e_k\, q_k / \sum_k q_k$ — the quantity-weighted market average — for all $j$.

to be captured by a tax schedule affine in $e_j$.

Consider Case B without other taxes ($\bar{\tau}_k = 0$), so that $\tau_j^\star = \lambda_e(e_j - \bar{e}_j)$. Substituting **??** for $\bar{e}_j$ in **??** gives:

$$\begin{aligned}\tau_j^\star &= \lambda_e\big(e_j - \alpha - \beta e_j - u_j\big), \\ &= \underbrace{-\lambda_e\alpha}_{\text{intercept}} + \underbrace{\lambda_e(1-\beta)}_{\text{slope}}\, e_j - \quad \underbrace{\lambda_e\, u_j}_{\text{tagging error}} \quad . \end{aligned} \tag{6}$$

The intercept and slope in **??** are the part of $\tau_j^\star$ that can be implemented with a feebate, $c_0 + c_1 e_j$. The corresponding coefficients are $c_0 = -\lambda_e\alpha$ and $c_1 = \lambda_e(1-\beta)$. The regression intercept $\alpha$ captures the level of substitute damage that does not vary with own damage: an increase of $\Delta\alpha$ shifts the corrective tax schedule downward by $\lambda_e\Delta\alpha$. The slope depends on how substitute damage varies with own damage: a one-unit increase in $e_j$ is associated with a $\beta$-unit increase in $\bar{e}_j$, so the fitted corrective tax increases by $\lambda_e(1-\beta)$. When $\bar{e}_j$ is constant across products, $\beta = 0$, and the coefficient on own damage remains $\lambda_e$. Under Case A, when all products are taxed, destination products are corrected through their own tax rates and the slope is simply $\lambda_e$.

For example, the relationships $\bar{e}_j = e_j$ and $\bar{e}_j = \frac{1}{2}e_j$ are both exact affine relationships, but imply different corrective tax schedules. In the first, taxing product $j$ alone does not reduce aggregate external damage because its marginal buyers move to alternatives with the same average damage. The Case B correction is therefore zero. In the second, average substitute damage is half of own damage, so the Case B correction is $\frac{1}{2}\lambda_e e_j$. Both corrective tax schedules can be represented as functions of own damage, but the coefficient on own damage differs. More generally, with an exact affine relationship, Case B tax rates rise with own damage when $\beta < 1$, are constant at $-\lambda_e\alpha$ when $\beta = 1$, and fall with own damage when $\beta > 1$.

The residual, or tagging error, captures the gap between $\tau_j^\star$ and the fitted affine tax schedule in $e_j$. In **??**, it appears as $-\lambda_e u_j$, where $u_j$ is the residual from projecting $\bar{e}_j$ onto the span of $[1, e_j]$. When two products have the same $e_j$, but different values of $\bar{e}_j$, an attribute-based tax will assign them the same tax rate even though their optimal corrective taxes differ. We report the coefficient $\beta$, the RMSE of $u_j$, and the $R^2$ of a regression of $M_j^e$ on $e_j$ in our empirical case studies to characterize how closely a tax schedule affine in $e_j$ approximates $\tau_j^\star$. The coefficient $\beta$ determines the slope of the fitted corrective tax, the RMSE measures the magnitude of the approximation error in external-damage units, and

the $R^2$ measures the share of variation in $M_j^e$ captured by an affine function of $e_j$.[11]

Our regression idea resembles **?**, but those authors ask a different question. In their setting the planner cannot write a tax on the damage $e_j$ itself, only on some observable attribute $f_j$. They ask how well an affine schedule in $f_j$ approximates a tax schedule based on $e_j$, and show that the sales-weighted $R^2$ from regressing $e_j$ on $f_j$ and an intercept equals the share of the first-best welfare gain, beyond a uniform tax, that the attribute recovers.[12] Their setup resembles Case A in that every dirty product is taxed, firms are competitive, and they assume away substitution to the outside option, so $e_0$ is irrelevant. Here, the only way an attribute fails is by mismeasuring damage. Our point is that coverage matters even when $e_j$ is measured perfectly. When the tax does not reach every product (e.g. a rebate on clean vehicles only) or does not fully correct every substitute (e.g. flat volumetric taxes on beer and wine), the target $\tau_j^\star$ itself moves, as Case B shows, because part of the damage the tax removes from product $j$ reappears at its substitutes. Because we regress the single-reform target $M_j^e$ rather than the rates a full-coverage planner would set, our $R^2$ is interpreted as a tagging diagnostic, not the welfare share.

The caveat is that a good fit is necessary, but not sufficient. When a tax reaches only some products, substitution among those products is corrected at the destination and drops out of the target. The attribute $e_j$ of the taxed products must predict leakage to the untaxed ones. A market can pass the market-wide regression and still fail this test, as the clean-vehicle subsidies in **??** demonstrate.

### 2.4 Extensions: Market Power and a Revenue Motive

The baseline model above abstracts from two features of real-world markets that affect the optimal corrective tax but do not change the underlying logic of when substitutes will matter. First, firms may have market power, setting prices above marginal cost and creating a pre-existing wedge before any corrective tax is imposed. Second, governments may value revenue, which can push tax rates above the levels implied by external damage alone.

With markups $\eta_j = p_j - c_j - \tau_j$ and a marginal cost of public funds $\lambda_r \geq 0$ that values

[11] Regressing $M_j^e$ or $\bar{e}_j$ on $e_j$ gives residuals of equal magnitude but opposite sign, so the two regressions have the same RMSE. The former has slope $1 - \hat{\beta}$. We report the $R^2$ of the $M_j^e$ regression rather than the $R^2$ from **??**, because it scales the residual variation relative to variation in the corrective tax target.

[12] In their automobiles application $f_j$ is the EPA fuel-consumption rating (gallons per mile) and $e_j$ is lifetime $CO_2$, the rating times lifetime miles. The attribute falls short because lifetime mileage varies across models with the same rating.

revenue at $(1+\lambda_r)$ per dollar, the complete-coverage optimum of Case A becomes:

$$\boldsymbol{\tau}^{\star\star} = \frac{1}{1+\lambda_r}\Big( \underbrace{\lambda_e \tilde{\mathbf{e}}}_{\text{corrective}} - \underbrace{\boldsymbol{\eta}}_{\text{markup}} \Big) + \frac{\lambda_r}{1+\lambda_r} \underbrace{\widetilde{\Omega}^{-1}\mathbf{q}}_{\text{Ramsey}}, \qquad \tilde{\mathbf{e}} \equiv \mathbf{e} - e_0\mathbf{1}. \tag{7}$$

Here $\widetilde{\Omega} \equiv -\left(\partial \mathbf{q}/\partial \boldsymbol{\tau}'\right)'$ is the $J \times J$ matrix of demand responses to the tax vector, so $\widetilde{\Omega}^{-1}\mathbf{q}$ is the multi-product analogue of the inverse-elasticity term. See **??** for the full derivation.

For Case B, holding the taxes on substitutes fixed at $\bar{\tau}_k$, define $\widetilde{D}_{j\to k} \equiv \frac{\partial q_k/\partial \tau_j}{|\partial q_j/\partial \tau_j|}$. As above, the set of destinations includes the outside option with $\eta_0 = \bar{\tau}_0 = 0$. The single-rate first-order condition gives

$$\tau_j^{\star\star} = \frac{1}{1+\lambda_r}\left[\underbrace{\lambda_e\left(e_j - \sum_{k\neq j}\widetilde{D}_{j\to k}e_k\right)}_{\text{corrective}} - \underbrace{\left(\eta_j - \sum_{k\neq j}\widetilde{D}_{j\to k}\eta_k\right)}_{\text{markup}}\right] + \underbrace{\sum_{k\neq j}\widetilde{D}_{j\to k}\bar{\tau}_k}_{\text{substitute taxes}} + \frac{\lambda_r}{1+\lambda_r}\underbrace{\frac{q_j}{|\partial q_j/\partial \tau_j|}}_{\text{Ramsey}} \tag{8}$$

The externality and markup wedges are both substitution-adjusted: the relevant comparison is between product $j$ and the alternatives chosen by its marginal buyers. The fixed taxes on those alternatives enter through their diversion-weighted average, while the revenue motive adds the single-rate Ramsey term.

Market power enters as a pre-existing wedge with the opposite sign from the externality. Markups mean that consumers already pay more than marginal cost, leaving less of the externality for the tax to correct. When the markup exceeds the external damage, the corrective component of the tax becomes a subsidy (**?**). Because markups reflect elasticities and portfolio pricing rather than externality content, they are unlikely to be correlated with $e_j$.[13] Under imperfect competition, $\widetilde{\Omega}$ in **??** and the diversion ratios $\widetilde{D}_{j\to k}$ in **??** reflect the equilibrium response to the tax, including rivals' repricing. In the automobile analysis of **??**, we capture rivals' repricing directly by re-solving the Bertrand equilibrium as taxes change.

A revenue motive changes the target by scaling the corrective and markup wedges by $1/(1+\lambda_r)$ while leaving their bases unchanged (**?**). It also adds the Ramsey term, which depends on the substitution structure rather than on $e_j$ and is also unlikely to be correlated with $e_j$.

[13] **?** find that higher quality products have higher markups but lower $M_j^e$, so this works against the externality correction; **?** find that ignoring market power in SSB taxation reduces welfare gains by 40 percent.

In the perfect-competition baseline, $M_j^e(\mathbf{p})$ and the diversion ratios $D_{j\to k}(\mathbf{p})$ evaluated at prevailing prices are exact for the first-order effect of a marginal tax on a single good.[14] With market power, the corresponding first-order condition instead depends on the equilibrium quantity response to the tax, as captured by $\widetilde{D}_{j\to k}$ above. Large tax changes, especially those that adjust the prices of many goods at once, require re-solving the demand system at the prices the schedule induces, as we do in the equilibrium exercises in **??**.

## 3 Case Study: Distilled Spirits

Distilled spirits provide a benchmark case where the externality-generating attribute is nearly homogeneous across products, but diversion patterns, particularly to the outside option, are not. Ethanol is the ingredient associated with both health consequences (cirrhosis, cancer) and safety issues (domestic violence, drunk driving) and most spirits are bottled at 40% ABV (80 proof) with flavored products (20–30% ABV) and overproof products (50%+) as the main exceptions. The outside option is not necessarily clean as spirits consumers may substitute to beer or wine, where $e_0 > 0$.[15]

To compute $\bar{e}_j$ and $M_j^e$, we use demand estimates from our prior work on distilled spirits, **?**. The estimation uses product-level sales and pricing, individual purchase data by income, and a state excise tax change for identification. The resulting demand estimates give us $\widehat{D}_{j\to k}$ for every pair of products and the outside option, from which we compute $\bar{e}_j$ and $M_j^e$ for each product.

**??** plots destination proof $\bar{e}_j$ against own proof $e_j$. The top panel assumes every consumer who leaves a product switches to another spirit. Whichever product a marginal buyer leaves, the destination spirits are roughly 80 proof. The estimated slope is small, $\hat{\beta} = 0.08$, with an RMS residual of only 1.3 proof points. This is the *degenerate case* of **??**. Even for Dubra Vodka, the cheapest product in the market and the one whose buyers are most likely to exit, marginal buyers who switch to other spirits choose products with roughly the same average proof as marginal buyers leaving other products.

The bottom panel adds the exit margin, treating the outside option as clean ($e_0 = 0$). The slope stays near zero, $\hat{\beta} = 0.04$, but the dispersion more than doubles, with an RMS residual of 3.7 proof points. Since $\bar{e}_j \approx (1 - D_{j\to 0}) \times 80$, the added spread comes from differences in exit rates, not from differences in destination products among marginal buyers

[14] This is the approximation that **?** rely on for their $R^2$ interpretation.

[15] Our sample predates the rise of non-alcoholic spirits. The obvious fix is to include beer, wine, and non-alcoholic alternatives in the market and measure substitution to them directly, though this is rarely done: **??????** include beer and wine in the outside option for spirits, while **?** model choices among broad categories (rum, white wine, red wine) rather than brands. **?** find the heaviest drinkers consume primarily spirits and limited cross-purchasing across the major categories (beer, wine, spirits).

who continue to purchase spirits. Bulldog Gin stays near 50 proof because relatively few of its marginal buyers switch to the outside good, while Dubra falls to 30 because more than 60% of its marginal buyers leave for the outside good. Only the lowest-proof products, such as Malibu Coconut Rum, have destination products with higher proofs than their own. For them $M_j^e < 0$, and raising their price increases ethanol consumption on net.

The exit margin is the main source of the variation in $M_j^e$, which we illustrate in **??**. Budget products have diversion to the outside good as high as 60%, while premium products (Belvedere, Grey Goose) are as low as 40% because their marginal buyers are more likely to substitute to other premium spirits. Outside diversion is unrelated to ethanol content ($R^2 = 0.00$) and U-shaped in price.[16] The consequence is substantial dispersion in $M_j^e$ *conditional on* $e_j$. Dubra and Bulldog are both 80 proof, but the marginal reduction in ethanol from taxing Dubra is 60% larger. A uniform ethanol tax, even if set correctly on average, overtaxes premium products and undertaxes budget products relative to the optimum. **?** find that the main source of counterfactual welfare gains is taxing the low end of the market while reducing markups on expensive products.

For tax design, the key result is that $\hat{\beta} \approx 0$, meaning that the external damage of a product's substitutes does not rise with $e_j$. With little to correct in the slope, the single-reform slope in **??** is essentially the full Pigouvian slope $\lambda_e e_j$. The intercept does most of the work. There is little difference between taxing volume and ethanol content when there is little dispersion in $\bar{e}_j$, and neither tax base captures the dispersion in $D_{j\to 0}$. This is consistent with the welfare results in **?**, which find nearly identical welfare estimates for taxes based on volume and ethanol content (and good news for state governments that tax spirits by volume). However, the dispersion in $\bar{e}_j$ conditional on $e_j$ (the $u_j$ term in **??**) is not reachable by a tax on $e_j$, and requires instruments that move with the exit margin itself, such as minimum unit prices (**??**), which can be effective at targeting the low-priced products that are most substitutable to the outside option. Allowing for a dirty outside option (**??**, **??**) compresses the gap between the outside option $e_0$ and inside substitutes $e_k$. Dispersion in $D_{j\to 0}$ then generates less dispersion in $\bar{e}_j$, the $R^2$ rises from 0.82 to 0.91, and a tax on ethanol alone performs better than it does with a clean outside option. In this sense, a dirty outside option lowers the stakes of ignoring substitution patterns.

## 4 Case Study: Automobiles

Vehicles provide a second quantitative application in which the substitution problem is the reverse of the spirits case. In spirits, most products have similar levels of ethanol, while

[16] All fits here and below are sales-weighted, because welfare aggregates over quantities and there is a long tail of seldom purchased products.

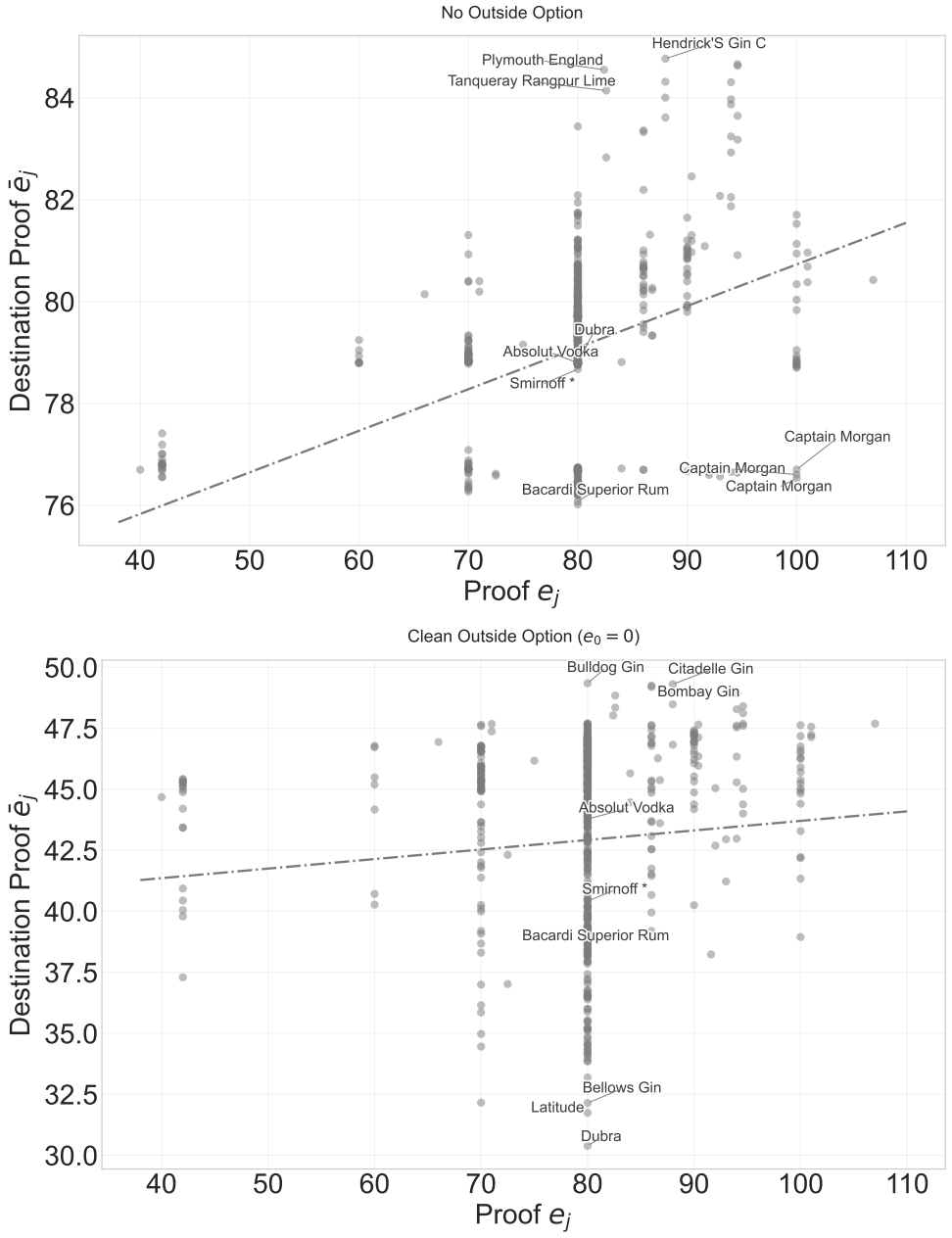



Figure 1: Destination proof $\bar{e}_j$ against own proof $e_j$ for spirits (2013 Q3); note the vertical scales differ across panels. Top: conditional on remaining in the market ($\hat{\beta} = 0.08$, RMS residual 1.3 proof points). Bottom: unconditional, with a clean outside option ($e_0 = 0$; $\hat{\beta} = 0.04$, RMS 3.7). The sales-weighted $R^2$ of the $M_j^e$ regression is 0.82. **??** repeats the bottom panel with a dirty outside option.

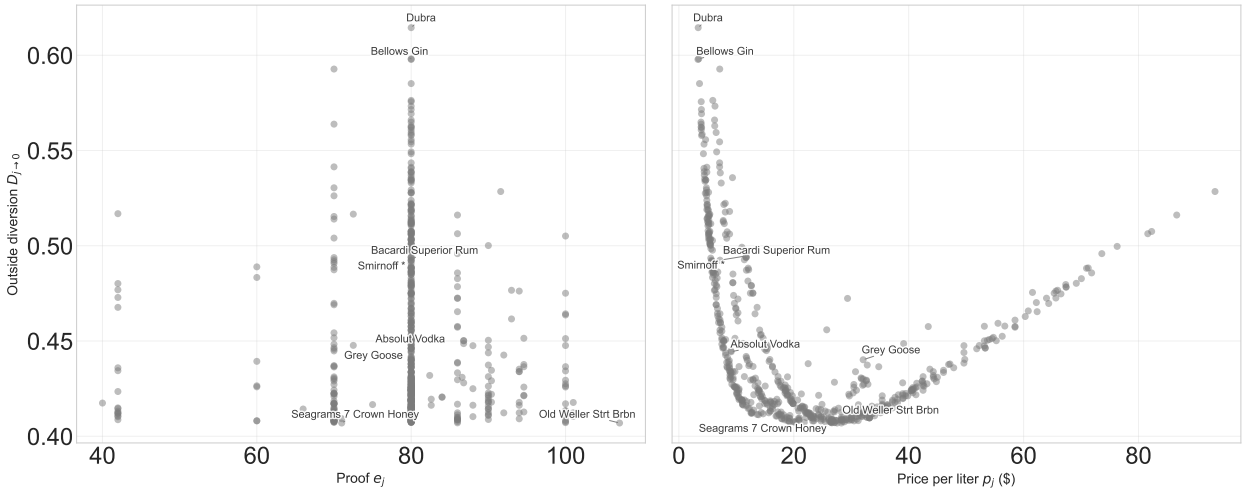



Figure 2: Estimated outside good diversion $D_{j\to 0}$ against proof (left) and price per liter (right) for spirits. The sales-weighted $R^2$ of $D_{j\to 0}$ on proof is 0.00. The distinct curves in the second panel correspond to bottle sizes. Source: **?**.

diversion to the outside good varies substantially. In vehicles, diversion to the outside good is modest, while emissions vary widely within the new-car fleet, from near-zero for electric vehicles to very high for large trucks and SUVs. The question is not so much whether consumers leave the market, but which new vehicle they substitute to when a given vehicle is taxed. Vehicles are not close to the degenerate case because both own emissions and destination damage $\bar{e}_j$ vary widely. Instead, vehicles are closer to the "well-predicted" case of **??**. Substitution largely follows well-defined categories, with truck buyers diverting to other trucks and hybrid buyers to other hybrids, so $\bar{e}_j$ rises with $e_j$, and $\hat{\beta} = 0.68$.

When measuring the externality, we focus only on $CO_2$ emissions, using the EPA's standard conversion from miles per gallon to $CO_2$ and monetizing emissions at \$28.50 per g/mile, which corresponds to a \$190 per ton social cost of carbon over 150,000 lifetime miles. This calibration is detailed in **??** and is intentionally narrow. Our goal is to highlight the role that substitution to untaxed goods with externalities plays, rather than to provide a complete environmental analysis.

The outside option is harder to characterize for automobiles. A consumer priced out of the new-car market may keep an older and dirtier vehicle, in which case $e_0$ is high, or may drive less, in which case $e_0$ is closer to zero. **?** show that when the outside option is the used fleet, the scrappage margin creates interactions that a naive emissions tax misses. We take the outside option to be a retained used car and set $e_0 = 450$ g/mile, about the median of the 2015 on-road fleet, as described in **??**. For all battery electric vehicles (BEVs), we

assume a constant $e_j = 110$ g/mile.[17]

Policies aimed at vehicle emissions work through the purchase margin in different ways. Fuel taxes are closest to a tax on driving, but higher expected fuel costs also steer buyers toward more efficient vehicles. CAFE standards work through firms by changing the compliance value of each model. Feebates and purchase taxes change relative prices at the point of sale. One close example is Norway's registration tax, which rises with $CO_2$ and curb weight while largely exempting battery-electric vehicles. All of these policies shift which vehicles consumers buy, so their effect on emissions depends on diversion across models.[18]

The actual U.S. policy in place during our 2015 sample period was not a tax on $CO_2$ but a subsidy to plug-in vehicles. From 2010 through 2022 the federal clean-vehicle credit (IRC §30D) paid up to $7,500 for a battery-electric or plug-in hybrid vehicle, scaled with battery capacity. Conventional hybrids were excluded once their own credit lapsed after 2010.[19] Applying the statutory credit amounts to the 2015 sales in our sample implies a program cost of roughly $0.66B, in line with fiscal estimates of $0.5–0.7B per year over this period.[20] Flat, class-based subsidies of this kind are exactly the simple instruments our framework is designed to evaluate.

We use data on roughly 50,000 automobile purchasers from MaritzCX (now InMoment, the same data used by **?**). Each respondent reports not only which car they purchased but which car they *would have purchased* had their first choice been unavailable. These responses provide a direct product-level measure of second-choice preferences from which we construct diversion ratios $D_{j\to k}$. We use these data in two ways. First, a *tagging* exercise

[17] We do this for lack of better data. In practice, BEVs differ in their kWh/mile efficiency, and in the emissions of the electricity used to charge them (solar, coal, natural gas, etc.).

[18] **?** show that attribute-based fuel economy regulation, where compliance depends on vehicle weight, creates incentives to distort the secondary attribute, a mechanism analogous to the markup distortions documented for distilled spirits.

[19] The §30D plug-in credit took its up-to-$7,500, battery-capacity-scaled form in 2010, following the Energy Improvement and Extension Act of 2008 and the 2009 Recovery Act, with a 200,000-vehicle per-manufacturer phaseout that no manufacturer had reached when our 2015 vehicles were sold. The earlier credit for conventional hybrids, worth up to $3,400 under the Energy Policy Act of 2005, lapsed after 2010. The Inflation Reduction Act of 2022 kept the $7,500 ceiling but split it into two $3,750 halves tied to critical-mineral and battery-component sourcing, added a North-American final-assembly requirement and MSRP and income caps, and removed the manufacturer cap, while keeping eligibility limited to battery-electric and plug-in hybrid vehicles. The credit was terminated for vehicles acquired after September 30, 2025. Several states layered flat rebates on the same plug-in classes, including California's Clean Vehicle Rebate Project. Sources: IRC §30D; IRS, *Credits for New Clean Vehicles Purchased in 2023 or After*; U.S. Department of Energy, Alternative Fuels Data Center.

[20] Statutory credits on the 95,946 plug-in vehicles in our sample: $7,500 for each BEV and the Chevrolet Volt, $2,500–$4,000 for the remaining plug-in hybrids. For comparison, individual taxpayers claimed $375 million in §30D credits in tax year 2016, and cumulative tax expenditures over FY2011–FY2017 were $2.2B. Individual claims understate program cost because the credit on a leased vehicle accrues to the lessor. Source: Congressional Research Service, *The Plug-In Electric Vehicle Tax Credit*, IF11017.

takes the diversion ratios at prevailing prices, computes $M_j^e$ for each vehicle, and asks how much of the substitution-adjusted externality a naive $CO_2$ tax recovers. This is the closest analogue to the spirits exercise. Second, an *equilibrium counterfactual* uses the mixed logit demand system estimated in **?**, backs out marginal costs from the multi-product Bertrand first-order conditions, and re-solves equilibrium under counterfactual taxes and subsidies. The equilibrium exercise is where accounting for substitution matters most, and is the focus of the final part of this section.

### 4.1 Dispersion in Emissions of Substitutes

A Pigouvian emissions tax depends on a vehicle's *own* emissions $e_j$, while $M_j^e$ or $\bar{e}_j$ depend on what the marginal buyer would otherwise have driven. In the framework above, a tax or subsidy on $e_j$ is a good approximation when average destination emissions $\bar{e}_j$ are well predicted by own emissions. **??** mirrors the spirits plot in **??**, showing destination damage $\bar{e}_j$ against own emissions $e_j$, conditional on staying within the new-car market.

The pattern for vehicles differs from what we saw in spirits. Both own emissions and destination emissions vary considerably, and the slope is relatively large, $\hat{\beta} = 0.68$. A vehicle that emits an additional 100 grams of $CO_2$ per mile has substitutes that emit, on average, 68 grams more. This means that the correct single-reform feebate slope is about a third of the naive Pigouvian one in **??**. The dispersion around the fitted line is modest, with an RMS residual of 72 g/mile, so over the fleet's wide emissions range, a feebate corrects most but not all of the external damage.

Products above the 45-degree line in **??** have $\bar{e}_j > e_j$, so $M_j^e$ is negative and the optimal single-product correction becomes a subsidy. The entire clean-vehicle category lies above the 45-degree line, but the three most prominent models differ sharply in what their marginal buyers would otherwise purchase. A marginal **Tesla Model S** buyer would otherwise choose among higher-emissions luxury sedans or crossovers, and other plug-ins, yielding $M_j^e = -465$, the largest marginal emissions reduction in the fleet. A marginal **Toyota Prius** buyer, despite choosing a conventional hybrid, would otherwise buy a mainstream gasoline car such as a VW Golf, Mazda3, or Jetta, giving $M_j^e = -174$. A marginal **plug-in Prius** buyer substitutes mostly toward other electrified cars that are already clean, so its correction is only $M_j^e = -55$. Even though the conventional hybrid is the dirtier of the Prius models, it is a more valuable subsidy target than the plug-in Prius because its marginal buyer displaces a dirtier gasoline car.

Negative $M_j^e$ is not confined to clean vehicles. The Ford F-Series, a full-size pickup truck and the best-selling vehicle in the U.S., has one of the highest emissions rates in the fleet, $e_j = 1{,}087$ g/mile, but its substitutes have even higher emissions, meaning $\bar{e}_j > e_j$ and

$M_j^e < 0$. If all vehicles were taxed, the optimal tax on the F-Series would be large and positive: $\tau_j^\star = \lambda_e(e_j - e_0) = \$18{,}200$ with $\lambda_r = 0$.[21] If the only policy instrument is a tax on the F-Series alone, the optimal single-product correction is instead a small subsidy, $\lambda_e M_j^e = -\$1{,}400$. The contrast is useful because the full-coverage and single-product calculations are polar cases. Actual tax or rebate designs, with partial coverage or imperfect instruments, generally lie somewhere in the middle. We provide more details on the F-Series and its substitutes in **??**.

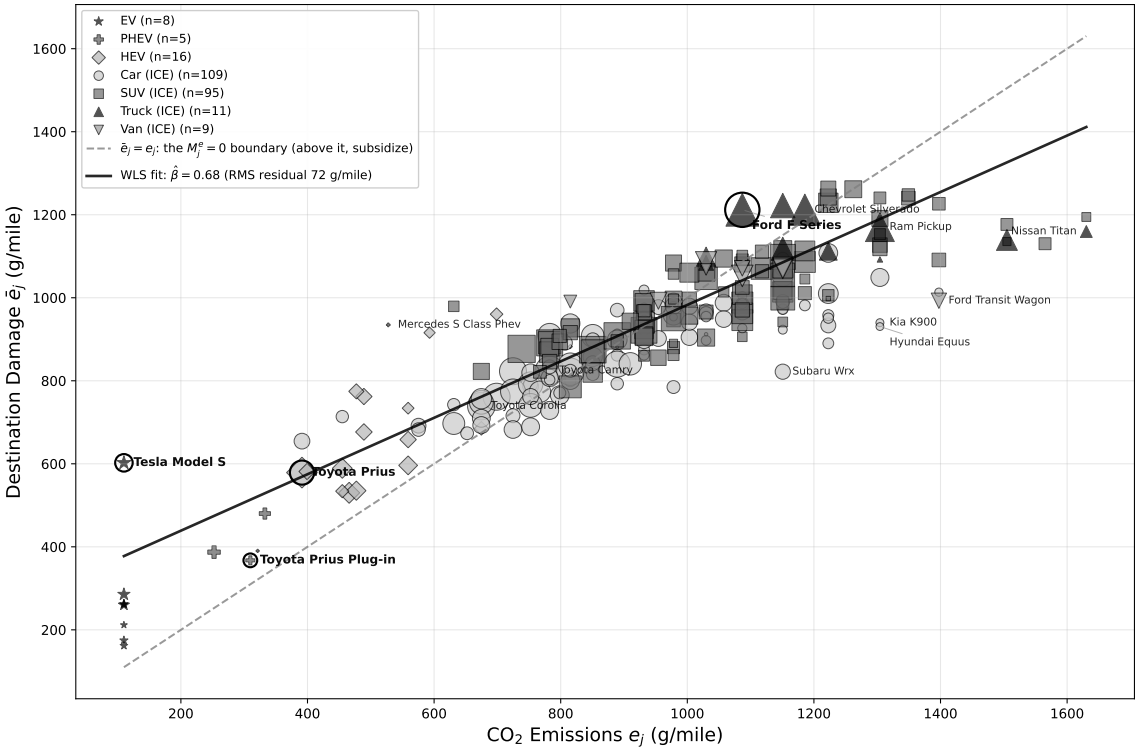



Figure 3: Where the marginal buyer goes. Destination damage $\bar{e}_j$: the diversion-weighted mean $CO_2$ of a marginal buyer of $j$'s alternatives against own emissions $e_j$ for the 2015 fleet. The dashed 45-degree line $\bar{e}_j = e_j$ is the $M_j^e = 0$ boundary, above which the optimal correction is a subsidy. Points are sized by sales and marked by vehicle category; the four vehicles discussed in the text are ringed. The sales-weighted fit has slope $\hat{\beta} = 0.68$ with RMS residual 72 g/mile; the equivalent regression of $M_j^e$ on $e_j$ (same residuals with opposite sign, slope $1 - \hat{\beta}$) has sales-weighted $R^2 = 0.47$.

Across the entire fleet, an affine function of own emissions explains 47 percent of the sales-weighted variation in $M_j^e$[22], which is why an attribute tax does well when products span a wide emissions range. **??** zooms in on the 29 clean vehicles, the segment of the market

[21] Accounting for the markup $\eta_j$ this shrinks to $\tau_j^\star = \$11{,}600$. Increasing the value of public funds raises the tax even further (because of the revenue motive).

[22] Regressing $M_j^e$ on $e_j - e_0$ is identical to regressing on $e_j$ alone, since $e_0$ enters only as a constant shift.

where U.S. emissions policy has largely focused. Among these vehicles, the fit collapses, with $R^2 = 0.04$ equal-weighted and $R^2 = 0.33$ sales-weighted[23]. On one hand, this means a subsidy based on $e_j$ is unlikely to be helpful when restricted to the clean-vehicle category. On the other hand, the simplest uniform rebate may do reasonably well because all clean vehicles merit a subsidy. A uniform subsidy (or one tied to $e_j$) performs less well as the cost of public funds increases and we have to ration spending. As the cost of public funds rises, substitution becomes more relevant, as we show in the next section.

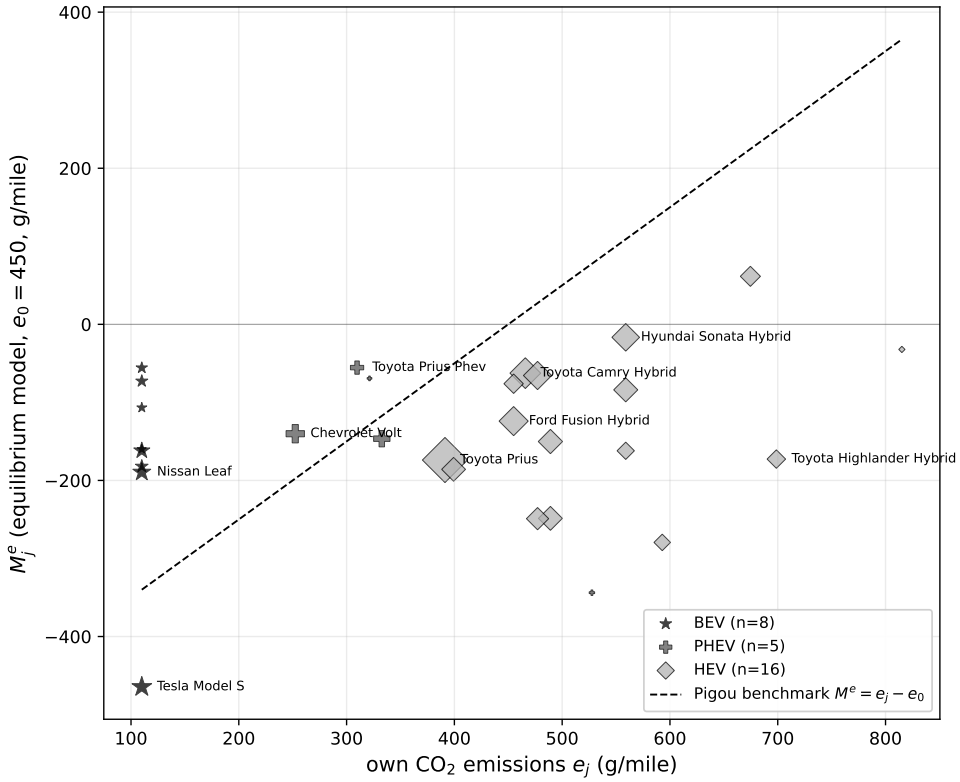



Figure 4: The substitution-adjusted externality $M_j^e$ against own $CO_2$ emissions $e_j$ for the 29 clean (BEV/PHEV/HEV) vehicles. Within-category $R^2 = 0.33$ sales-weighted, 0.04 equal-weighted (sales-weighted puts 48% of weight on three vehicles and 35% of weight on the Prius HEV alone). The dashed line is the Pigouvian benchmark $M_j^e = e_j - e_0$, and the vertical gap is the substitution correction.

### 4.2 Limited Coverage and Clean-vehicle Subsidies

Actual U.S. policy in 2015 subsidized clean vehicles by class rather than taxing emissions. Of the 254 models in the data, 8 battery-electric vehicles (BEVs) and 5 plug-in hybrids

[23]We report the fit both ways because the Toyota Prius alone is 35% of category sales and the top three products are 48%.

(PHEVs) qualified for the §30D credit. The 16 conventional hybrids (HEVs) did not, even though they accounted for three-quarters of clean-vehicle sales. A subsidy confined to the clean-vehicle category, or to part of it, is highly incomplete. With incomplete coverage, the target depends not only on own emissions but also on the emissions of substitutes. Within the covered class, the own-emissions attribute adds little information, since we assign every BEV the same $e_j = 110$ in the analysis. What varies across clean vehicles is where each vehicle's marginal buyers would otherwise go, $\bar{e}_j$.

**??** illustrates how the set of products covered by the tax or subsidy and the cost of public funds, $\lambda_r \in \{0, 0.25, 0.5\}$, affect welfare. We consider four coverage scenarios. The first allows all 254 products to be reached by the tax or subsidy. We then narrow coverage to clean vehicles (BEVs/PHEVs/HEVs), to the vehicles eligible for the §30D credit (BEVs and PHEVs), and finally to fully electric BEVs only. For each scenario, the header rows report the maximum attainable welfare from an unrestricted product-specific tax or subsidy and the corresponding cost to the government, or revenue raised. The remaining rows report the fraction of that maximum achieved by each rebate or tax scheme.

Unsurprisingly, as the tax base shrinks from full coverage to smaller sets of products, the maximum attainable welfare shrinks as well. Likewise, as the value of public funds rises, clean-vehicle rebates become less generous and generate smaller welfare gains.[24] As an example, we focus on the §30D case, which covers only the 13 BEVs and PHEVs. A single uniform rebate for all BEVs and PHEVs attains 79% of the maximum attainable welfare when the cost of public funds is low, $\lambda_r = 0$, and only 56% when the cost of public funds is high, $\lambda_r = 0.5$. We also consider what we call a "tiered feebate", which has different base rebates for BEVs and PHEVs along with a variable component that scales with either $e_j$ or $M_j^e$. (The actual §30D program rebate scaled with battery capacity.) We see that these tiered feebates do significantly better and capture $> 90\%$ of the maximum attainable welfare at low cost of public funds, with a slight advantage to the $e_j$ case. At the higher value of $\lambda_r$, the $e_j$ feebate achieves only 68% of welfare gains, while the $M_j^e$ version achieves 95%.

Three patterns hold across the panels of **??**. First, under full coverage the emissions feebate is nearly sufficient, capturing 93%–96% of the optimum at every $\lambda_r$, and the substitution-adjusted base adds nothing. When every product is taxable, the planner can reach the dirty substitutes directly, as in Case A. Second, when public funds are free ($\lambda_r = 0$), the emissions feebate performs at least as well as other schemes (within 0.01) in every panel. Getting the average subsidy level right is then nearly all that matters, and there is little cost to over-subsidizing some clean vehicles. Third, as the value of public funds rises, the substitution

[24] In the full-coverage tax and rebate scenario, increasing the value of public funds raises taxes on dirty vehicles and increases government revenue.

Table 1: Welfare by Instrument Coverage and the Cost of Public Funds

| | $\lambda_r = 0$ | $\lambda_r = 0.25$ | $\lambda_r = 0.5$ |
|---|---|---|---|
| **Full coverage: taxes on all 254 vehicles** | | | |
| attainable welfare $W_{\lambda_r}$ (\$B) | 33.9 | 54.7 | 97.9 |
| revenue at the optimum (\$B) | 4.7 | 138.8 | 198.3 |
| feebate $c_0 + c_1 e_j$ | 0.95 | 0.96 | 0.93 |
| feebate $c_0 + c_1 M_j^e$ | 0.65 | 0.81 | 0.89 |
| **Clean-category rebates: 29 BEV/PHEV/HEV** | | | |
| attainable welfare $W_{\lambda_r}$ (\$B) | 12.6 | 3.9 | 0.8 |
| revenue at the optimum (\$B) | −55.2 | −20.5 | −6.2 |
| uniform rebate | 0.69 | 0.67 | 0.52 |
| tiered BEV/PHEV/HEV | 0.91 | 0.89 | 0.69 |
| tiered feebate $c_0 + c_1 e_j$ | 0.96 | 0.95 | 0.75 |
| tiered feebate $c_0 + c_1 M_j^e$ | 0.91 | 0.91 | 0.93 |
| **Plug-in-only rebates: 13 BEV/PHEV (the §30D coverage)** | | | |
| attainable welfare $W_{\lambda_r}$ (\$B) | 6.6 | 1.8 | 0.3 |
| revenue at the optimum (\$B) | −32.2 | −10.4 | −2.6 |
| uniform rebate | 0.79 | 0.76 | 0.56 |
| tiered feebate $c_0 + c_1 e_j$ | 0.97 | 0.94 | 0.68 |
| tiered feebate $c_0 + c_1 M_j^e$ | 0.92 | 0.92 | 0.95 |
| **BEV-only rebates: 8 vehicles** | | | |
| attainable welfare $W_{\lambda_r}$ (\$B) | 3.4 | 0.8 | 0.1 |
| revenue at the optimum (\$B) | −18.2 | −5.3 | −1.0 |
| uniform rebate (= feebate $c_0 + c_1 e_j$) | 0.91 | 0.84 | 0.43 |
| feebate $c_0 + c_1 M_j^e$ | 0.92 | 0.92 | 0.98 |

*Notes:* Schedule rows report the welfare of the optimal schedule in each class as a share of the product-specific optimum on the same covered set; the header rows give that optimum's welfare $W_{\lambda_r}$ and its tax revenue in \$B (negative entries are subsidy outlays), both relative to the no-tax equilibrium, with revenue valued at $1 + \lambda_r$. Products outside the covered set are untaxed. Every BEV has the same grid-based $e_j = 110$ g/mile, so in the BEV-only panel the emissions feebate coincides with the uniform rebate. The tiered feebate rows (clean and §30D panels) give each covered class its own intercept and a single common slope on $M_j^e$. $e_0 = 450$ g/mile.

adjustment becomes more valuable because the schedule must ration the subsidy toward the vehicles whose buyers would otherwise drive something dirty. Within the clean category the ranking reverses at $\lambda_r = 0.5$. The tiered substitution-adjusted feebate recovers 93% of the category optimum, the tiered emissions feebate 75%, the tiered rebate 69%, and the uniform rebate 52%. On BEVs alone, there is no variation in the emissions base. Every BEV has the same $e_j$, so the emissions feebate *is* the uniform rebate and recovers 43% of the BEV-only optimum, against 98% for the substitution-adjusted feebate. In the clean-category and §30D panels the crossover occurs between $\lambda_r = 0.25$ and 0.5. On BEVs alone the substitution-adjusted feebate leads at every $\lambda_r$ (**??** traces a finer grid for the clean category). Since most tax credits, rebates, and standards apply only to selected vehicle classes, these narrower coverage cases may be more relevant for policy.

The optimal tiered rates for the clean category at $\lambda_r = 0.5$ are \$6.4k for BEVs, \$4.5k for PHEVs, and \$2.6k for HEVs. The best uniform rebate is \$2.8k per vehicle. Displaced buyers come primarily from other new vehicles rather than from the outside good, so clean-vehicle subsidies mostly reshuffle the new-vehicle fleet rather than expand it. The largest emissions reductions come from conventional hybrids, which are cheaper, more numerous, and closer substitutes for the dirty cars they displace. Conventional hybrids were notably excluded from the plug-in-only federal credit. The actual credit was also small relative to the optimal subsidies for the vehicles it covered. The \$0.66B statutory outlay on 2015 plug-in sales is roughly a quarter of the \$2.6B spent under the optimal plug-in-only schedule at $\lambda_r = 0.5$, suggesting that higher values of $\lambda_r$, or lower values of $\lambda_e$, might implicitly be the policy-relevant ones.

One limitation of our analysis is that it holds the set of vehicles fixed and does not account for the innovation margin, where firms change the vehicles they offer, or their attributes, in response to the tax or subsidy system.

## 5 Other Markets

The framework in **??** applies more broadly to markets where policy does not achieve complete coverage. **??** describes the two case studies above together with three additional policy-relevant markets in terms of **??**. For each market, we consider the outside option, how much the attribute varies among substitutes, whether destination damage is predicted by the attribute, and which relevant substitutes an attribute tax fails to reach. In none of the five markets does policy achieve complete coverage (Case A), because some relevant substitutes are either untaxed or taxed on a coarser basis. The two case studies above illustrate the extremes. In spirits, destination damage is flat in own damage ($\hat{\beta} = 0.04$) and

the difficulties are in the exit margin. In automobiles, destination damage rises with own damage ($\hat{\beta} = 0.68$) and within-category substitution is strong.

Table 2: Attribute-Based Taxation Across Five Markets

| **Market** | **Outside option** $e_0$ | **Dispersion in** $e_k$ | **Fit of** $\bar{e}_j$ **to own damage** $e_j$ | **Where the tax leaks** |
|---|---|---|---|---|
| Distilled spirits† | Beer/wine ($e_0 > 0$; $e_0 = 0$ conservative) | Low (most 80 proof) | $\bar{e}_j$ nearly constant; $\hat{\beta} = 0.04$; $R^2(M_j^e$ on $e_j) = 0.82$ | Exit margin ($D_{j\to 0}$: 0.41–0.61, unrelated to proof) |
| Cannabis | Illicit market (dirtier) | High (THC 15–90%); taxes mostly on price | Likely poor | Exit to illicit supply; potency switching |
| Sweetened beverages | Juice ($e_0 > 0$) vs. water/diet ($e_0 = 0$) | High in sugar; taxes on volume | Partial (diet products break the link) | Exempt juice; diet–regular switching |
| E-cigarettes / nicotine | Combustibles (far dirtier) | Harm tracks delivery, not nicotine | Poor (wrong attribute) | Exit to combustibles |
| Vehicles / emissions† | Used fleet ($\approx$ 450 g/mi) or none | Very high (110–1,100 g/mi) | Mixed: $\hat{\beta} = 0.68$, $R^2(M_j^e$ on $e_j) = 0.47$ across the fleet; fails within the clean category | Within-category residual; used fleet |

† Detailed case study in ????.

### 5.1 Cannabis

Cannabis presents several problems for a simple attribute tax. The outside option may generate even greater externalities than legal cannabis, while both externality content and substitution patterns vary sharply across products.

The outside option to legal cannabis may not be "abstinence" but instead the illicit market, and potentially alcohol, tobacco, and other substances. The illicit market is untaxed and may carry higher externalities than legal cannabis because unregulated products pose health risks and illicit distribution may generate violence and incarceration costs. **?** also document substantial substitution between legal cannabis and alcohol, finding that cannabis

legalization reduces alcohol demand by approximately 15%. **?** show that excessive cannabis taxes push consumers back toward illicit markets. The externality effect of taxing legal cannabis critically depends on where marginal consumers go when legal prices rise. When $e_0$ is large, $M_j^e$ can be very small or even *negative*. Raising legal cannabis prices may then increase total harm by diverting consumers to worse alternatives.

Externality content also varies substantially across cannabis products. THC content varies by a factor of five or more across product categories. Low-potency flower contains roughly 15% THC, high-potency flower approximately 30%, and concentrates 60–90% THC. Yet few jurisdictions tax THC directly. As **??** shows, most states levy excise taxes as a percentage of price and a few tax by weight. Only a handful of states tax by potency. Illinois applies higher rates to higher-THC products, while New York repealed its per-milligram THC tax in 2024 and Connecticut's is scheduled to be replaced by a 10.75% gross-receipts tax (**?**).[25] A price-based tax falls on quality, branding, and retail markups rather than on THC. Because concentrates deliver THC at a lower price per milligram, such a tax can tilt consumers toward higher-potency products. **?** study potency-based taxation in Washington State and find that demand for potency is relatively inelastic, so a naive THC tax has limited behavioral effect. A tax that, for example, pushes consumers away from lower-potency flower and toward higher-potency concentrates could reduce legal cannabis purchases while doing little to reduce, and possibly even increase, aggregate THC consumption.

Diversion patterns are also likely to differ across products. **?** show that entry restrictions in legal cannabis markets create substantial market power, with firms behaving as local monopolists. As in spirits, the resulting markups need not align with externality content. Firms may set high markups on premium products and low markups on budget products, even when THC content is similar. The illicit market may also serve as a better substitute for certain kinds of products, such as loose flower, than for others, such as gummies. This markup structure likely affects substitution patterns. Consumers of high-end legal products may be more likely to substitute to other legal cannabis products, while consumers of low-end products may be more likely to substitute to the illicit market. Diversion to the outside good may also differ across product forms if, for example, the illicit market is a closer substitute for loose flower than for gummies or other processed products.

In cannabis, a welfare-maximizing tax schedule might involve *lower* taxes on some legal products than a naive THC tax would prescribe when $e_0$ is high, in order to keep consumers in the regulated market rather than push them toward illicit alternatives. It might also involve

[25] New York's potency tax, in force from the start of retail sales in December 2022, was repealed effective June 1, 2024 by the FY2025 state budget and replaced by a 9% wholesale excise. Connecticut's Public Act 26-8 replaces its per-milligram THC tax with a flat 10.75% excise on gross receipts effective October 1, 2026 (Connecticut Department of Revenue Services, *Cannabis Tax Information*).

Table 3: Recreational Cannabis Tax Designs Across States

| State (sales began) | Structure | Base | Potency | Form | Rate |
|---|---|---|---|---|---|
| *Price-based (ad valorem)* | | | | | |
| Washington (2014) | Retail | Price | No | No | 37% of retail price |
| Colorado (2014) | Two-tier | Price[c] | No | No | 15% wholesale + 15% retail |
| Oregon (2015) | Retail | Price | No | No | 17% |
| Nevada (2017) | Two-tier | Price[c] | No | No | 15% wholesale + 10% retail |
| California (2018) | Retail | Price | No | No | 15% of gross receipts |
| Massachusetts (2018) | Retail | Price | No | No | 10.75% |
| Michigan (2019) | Two-tier | Price | No | No | 10% retail + 24% wholesale (2026) |
| Arizona (2021) | Retail | Price | No | No | 16% |
| Montana (2022) | Retail | Price | No | No | 20% |
| New Mexico (2022) | Retail | Price | No | No | 12%, rising 1 pt/yr to 18% by 2030 |
| Vermont (2022) | Retail | Price | No | No | 14% |
| New York (2022) | Two-tier | Price | No[a] | No | 9% wholesale + 13% retail |
| Rhode Island (2022) | Retail | Price | No | No | 13% |
| Missouri (2023) | Retail | Price | No | No | 6% |
| Maryland (2023) | Retail | Price | No | No | 9% (12% from 2026) |
| Ohio (2024) | Retail | Price | No | No | 10% |
| Delaware (2025) | Retail | Price | No | No | 15% |
| Minnesota (2025) | Retail | Price | No | No | 10% of gross receipts (15% from 2026) |
| *Weight-based or hybrid* | | | | | |
| Alaska (2016) | Per unit | Weight | No | Yes | $50/oz flower; $25/oz immature; $15/oz trim |
| Maine (2020) | Hybrid | Price + weight | No | Yes | 10% retail + $335/lb flower, $94/lb trim |
| New Jersey (2022) | Per unit | Weight[d] | No | No | $2.50/oz social-equity fee |
| *Potency-based* | | | | | |
| Illinois (2020) | Tiered | Price | Yes | Yes | 7% wholesale + 10% (≤35% THC), 25% (>35% THC), 20% (infused) |
| Connecticut (2023)[b] | Hybrid | Price + THC | Yes | Yes | 3% retail + $0.00625/mg THC (flower), $0.0275/mg (edibles), $0.009/mg (other) |

*Notes:* State excise taxes on recreational cannabis as of 2026; general sales taxes and local-option taxes (up to 3–7% in several states) are excluded. "Base" is what the excise is levied on. "Potency" indicates whether the rate varies with THC content; "Form" whether it varies by product type (flower, concentrates, edibles). Of the 23 states with legal retail markets, 18 tax price alone; only Illinois and Connecticut tax potency, and Connecticut is scheduled to abandon it. [a]New York's per-milligram THC tax, in force from the start of retail sales, was repealed effective June 1, 2024 and replaced by the wholesale excise. [b]Public Act 26-8 replaces the THC tax with a flat 10.75% excise on gross receipts effective October 1, 2026. [c]The wholesale tier applies the rate to a state-set average market price per pound ($658 in Colorado, $1,296 in Nevada), so it operates as a per-weight tax. [d]Levied per usable ounce on cultivators and reset annually by the Cannabis Regulatory Commission, which may raise it to as much as one-third of the average retail price. Sources: excise structures, rates, and local-option ranges from **?**; years of first retail sales from state regulators and contemporaneous press; footnote *a* from the New York FY2025 budget, the rate in footnote *b* from the Connecticut Department of Revenue Services, and the fee mechanics in footnote *d* from the New Jersey Division of Taxation.

different tax rates across product forms, with higher taxes on baked goods and gummies that are less substitutable to the illicit market. THC content alone is not enough to determine the corrective tax. The optimal corrective tax also depends on where consumers go when legal prices rise. Higher legal-market prices can push consumers to higher-potency products or a potentially more harmful illicit market, making the relevant policy question not just how much THC a product contains, but how total external harm changes when its price rises.

### 5.2 Sugar-Sweetened Beverages

Sugar-sweetened beverages raise a different design problem. The relevant externality-generating attribute is straightforward. The concern is sugar, and the case for corrective taxation has been extensively studied (**???**). Yet many existing taxes are not levied directly on sugar content, but rather on volume. The tax base often excludes beverages that are close substitutes and contain substantial sugar. This is the *incomplete instrument* problem we describe in **??**. A growing number of jurisdictions have implemented SSB taxes, but their designs vary in both structure and tax base, as summarized in **??**.

Existing SSB taxes differ substantially in what they cover. Every jurisdiction exempts 100% fruit juice, despite the fact that orange juice contains comparable sugar per ounce to many taxed sodas. Most jurisdictions also exempt milk-based drinks and sweetened coffee or tea purchased at cafes. These exemptions are a clear case of *incomplete coverage*. Close substitutes with comparable sugar loads, such as fruit juice, sweetened milk, and cafe coffee drinks, are untaxed. A consumer priced out of a taxed soda can divert to an untaxed sugary drink without reducing their sugar intake. **?** confirm these substitution patterns empirically.

Diet sodas and other artificially sweetened drinks raise a different issue. Most jurisdictions exclude them from the tax base, treating them as a "clean" substitute ($e_k \approx 0$). Philadelphia is a notable exception, taxing diet drinks at the same rate as regular soda. In our framework, if diet sodas are both clean ($e_k \approx 0$) and highly substitutable for sugar-sweetened sodas ($D_{j\to k}$ is large), then taxing them can increase external or internal damage by discouraging substitution toward a cleaner product, although taxing them may still be useful for raising revenue. This issue arose in Cook County, Illinois' short-lived attempt to include diet sodas in the tax base.[26]

Sugar content varies substantially across beverages, but many taxes are levied on volume

[26] Cook County's penny-per-ounce tax on sweetened beverages, which included diet drinks, took effect in August 2017 and was repealed by a 15–1 vote of the County Board on October 11, 2017, effective December 1, after a retailer- and beverage-industry-funded "Can the Tax" campaign and public complaints that the tax was a revenue measure rather than a health measure. See WTTW News, "Cook County Soda Tax Repealed," October 11, 2017, https://news.wttw.com/2017/10/11/cook-county-soda-tax-repealed.

Table 4: Sugar-Sweetened Beverage Taxes Across Jurisdictions

| **Jurisdiction** | **Structure** | **Diet** | **100% Juice** | **Milk** | **Rate** |
|---|---|---|---|---|---|
| Berkeley, CA (2015) | Flat volumetric | No | No | No | 1¢/oz |
| Philadelphia, PA (2017) | Flat volumetric | Yes | No | No | 1.5¢/oz |
| Cook County, IL (2017)* | Flat volumetric | Yes | No | No | 1¢/oz |
| Boulder, CO (2017) | Flat volumetric | No | No | No | 2¢/oz |
| Albany, CA (2017) | Flat volumetric | No | No | No | 1¢/oz |
| Oakland, CA (2017) | Flat volumetric | No | No | No | 1¢/oz |
| San Francisco, CA (2018) | Flat volumetric | No | No | No | 1¢/oz |
| Seattle, WA (2018) | Flat volumetric | No | No | No | 1.75¢/oz |
| Washington, DC (2019) | Sales-tax surcharge | Yes | No | No | 8% (vs. 6%) |
| Santa Cruz, CA (2025) | Flat volumetric | No | No | No | 2¢/oz |
| UK (SDIL, 2018) | Tiered by sugar | No | No | No | 20.8–27.8p/L |
| Mexico (2014; reformed 2026) | Flat volumetric | Reduced | No | No | 1.5–3.1 pesos/L |
| South Africa (2018) | Per gram of sugar | No | No | No | 2.31c/g >4g/100mL |

* Repealed after four months.

*Notes:* "Diet" indicates whether artificially sweetened beverages are taxed. "100% Juice" and "Milk" indicate whether these categories are included in the tax base. UK rates from April 2026 (uprated each April); lower tier applies at 5–8g sugar/100mL, upper tier at 8g+/100mL. South African rate since April 2022. Philadelphia and Cook County are the only US excise taxes to have included diet drinks; Washington, DC instead applies a higher general sales-tax rate to sweetened beverages, diet included, at the register. All other US taxes are per-ounce excises levied on distributors. Santa Cruz is the first California city to enact a tax since the state's 2018 preemption of local beverage taxes; a 2026 trial-court ruling upheld it. Not shown: the Navajo Nation's 2% sales-tax surcharge on unhealthy foods and sweetened beverages (2015, made permanent in 2020).

alone. Most U.S. SSB excise taxes have used flat volumetric designs. A 12oz Coca-Cola (39g sugar) and a 12oz Gatorade (21g sugar) pay the same tax, rather than being taxed directly on sugar content. **?** argue that per-gram-of-sugar taxes are better targeted, and South Africa's design comes closest to a tax proportional to the externality-generating attribute, $\tau_j \propto e_j$. The U.K.'s tiered Soft Drinks Industry Levy takes an intermediate approach with two sugar thresholds and led to substantial reformulation before the tax even took effect. But even South Africa's design ignores the substitution structure. It does not account for whether consumers leaving a taxed soda switch to untaxed fruit juice ($e_0 > 0$) or to water or diet sodas ($e_0 = 0$).

Diversion patterns are also unlikely to be uniform across consumers, products, or locations. Budget consumers may have higher price elasticity and be more likely to switch away from a taxed product, but they may also substitute toward cheap sugary alternatives rather than water. Consumers may substitute across jurisdictions rather than across products. **?** study Philadelphia's tax and find a 46% demand reduction in the taxed area, but cross-shopping at stores outside city limits offsets more than half of this reduction. **?** embed this substitution structure in an optimal tax framework with market power and find that concentrated brand ownership in the U.K. drinks market creates additional distortions that interact with tax design.

The SSB example shows that the tax base can matter as much as the tax rate. Taxes based on sugar content may be better targeted than taxes based on beverage volume, but targeting sugar content alone does not solve the problem if close substitutes with similar sugar loads remain untaxed. The tax design question is not only how much sugar is in the taxed product, but also whether the tax induces consumers to substitute toward lower-sugar alternatives, untaxed sugary drinks, or stores outside the taxed jurisdiction.

### 5.3 E-Cigarettes and Nicotine Products

Nicotine products present a different problem for an attribute tax. The relevant externality-generating attribute is not well captured by nicotine content alone. Health damage from nicotine products depends overwhelmingly on the delivery mechanism. A combustible cigarette and an e-cigarette may deliver similar amounts of nicotine but differ substantially in health harm. Taxes based on nicotine content alone will get the external damage wrong.

As in cannabis, a tax on e-cigarettes may push consumers to dirtier outside options. For e-cigarettes, those dirty substitutes may include combustible cigarettes where $e_0 \gg e_j$, but consumers may also substitute toward non-combustible products like Zyn or nicotine gum, which may entail much lower health harms. Where consumers substitute, and how to measure $\bar{e}_j$ and $M_j^e$ are the key empirical questions of the policy debate. Taxing vaping

products aggressively, as many states now do, or banning them outright, may push consumers toward the far more harmful combustible alternative. This is similar to cannabis, where overtaxing legal products can divert consumers to a worse outside option, but with arguably stronger evidence on the relative harms.

Diversion is also likely to differ across products. Budget disposable vapes and premium refillable devices may have very different diversion patterns to combustible cigarettes and may also entail possibly greater non-nicotine health costs. Diversion may also differ across flavors and product formats, especially if flavored disposable products are closer substitutes for cigarettes for some consumers than refillable devices or tobacco-flavored products. A uniform nicotine tax across all product forms could increase total health damage by discouraging substitution away from combustion.

Massachusetts, California, and the District of Columbia prohibit the sale of all flavored tobacco products, including menthol. New Jersey and New York prohibit flavored e-cigarettes, while Rhode Island, Utah, and Maryland restrict flavored e-cigarettes but exempt menthol. More than 420 local jurisdictions have enacted flavored-tobacco restrictions of their own.[27]

Understanding the impacts of taxes on e-cigarettes requires measures of $M_j^e$.[28] For example, in June 2022 the FDA ordered the products of JUUL, then the leading e-cigarette maker, off the market.[29] If the main response among teens to the JUUL ban was to never start using nicotine products, this would have been a public health success. If JUUL users instead largely diverted to combustible cigarettes or lower-quality e-cigarettes, the net effect might be to increase rather than decrease the external damage. Our framework treats bans and very high taxes as extreme price changes. Their effects depend on whether consumers quit, switch to less harmful products, or move toward more harmful substitutes.

The proposed federal ban on menthol cigarettes raises a related issue. Menthol cigarettes account for more than a third of U.S. cigarette sales and are used by roughly four in five Black smokers. The FDA proposed banning menthol as a characterizing flavor in 2022, but the rule was never finalized and was withdrawn in January 2025. Massachusetts and California have implemented state-level bans. The welfare effects of a ban depend on diversion patterns. The FDA projected that a ban would cause 923,000 additional smokers to quit, but evidence

[27] State laws as of 2026 per the Campaign for Tobacco-Free Kids; the local count is the Truth Initiative's tally as of December 31, 2025. San Francisco also banned the sale of e-cigarettes lacking FDA premarket authorization in 2019 while leaving combustible cigarettes on sale. Following the FDA's subsequent authorization of several products, that ordinance now applies only to unauthorized products.

[28] A nice feature of diversion $D_{j\to k}$ is that it can be measured using second-choice data as well as small price changes. See **?**.

[29] The order was stayed within two weeks, rescinded in June 2024, and effectively reversed in July 2025 when the FDA authorized JUUL's tobacco- and menthol-flavored pods. The example considers the counterfactual in which the order had stood.

from Massachusetts shows substantial diversion to non-menthol cigarettes and cross-border purchasing in neighboring states (**?**). A ban is equivalent to an infinite tax on one product variant, and if most of the diversion is to non-menthol cigarettes ($e_k \approx e_j$), the intensive margin externality reduction is near zero. If some diversion is to illicit menthol products ($e_0 \geq e_j$), the ban could increase total harm on that margin. The Canadian federal menthol ban in 2017 provides another opportunity to measure these responses. **?** find limited illicit market growth, though some switching to non-menthol variants.

The implication is that the welfare effects of nicotine taxes and bans depend on relative harm and diversion, not on nicotine content alone. A tax or ban that pushes consumers away from combustion can reduce harm even if nicotine consumption changes little. But if consumers instead move from lower-harm nicotine products toward combustible cigarettes or illicit alternatives, the same policy can increase harm.

## 6 Policy Implications

Four practical lessons emerge from this analysis.

**1. Outside options matter.** When the outside option is "dirty," the welfare-maximizing tax rates are lower than naive Pigouvian rates. This is most relevant for cannabis, where the illicit market poses greater risks than the legal market, but also applies to alcohol when the outside option includes beer and wine, SSBs when untaxed sugary alternatives are close substitutes, and vehicles when the outside option is an old, high-emitting car. In the extreme, $M_j^e$ can be negative, implying that the optimal *unilateral* correction might involve *subsidizing* the product. More generally, a tax that appears well targeted based on a product's own external damage may be much less effective if it pushes consumers toward untaxed, illicit, older, or otherwise more harmful alternatives.

**2. $M_j^e$ is not necessarily the target, but how well the tax base proxies for $M_j^e$ still matters.** Product-specific taxes are often administratively infeasible, and the welfare cost of using a simple attribute tax depends on how well $M_j^e$ can be approximated by an affine function of $e_j$. When that approximation fits well, an attribute-based feebate can closely approximate the product-specific corrective schedule, with the fitted slope determining how the tax should vary with $e_j$. When the residual from that approximation is large, or is concentrated in products that share another observable characteristic such as a low price, instruments that target that characteristic, such as minimum price floors, may outperform attribute taxes. In spirits, the overall fit is good ($R^2 = 0.83$), but the residual comes from the exit margin, which is highest among budget products. Minimum unit pricing, as adopted

in Scotland (**?**), raises prices most for the cheapest products, which in our estimates tend to have the highest $M_j^e$. Similar price floors have been discussed in cannabis regulation. The practical point is that it matters not only whether the statutory tax base captures the externality-generating attribute of a product, but also whether it is a good proxy for the substitution-adjusted reduction in external damage when the product's price rises.

**3. Market power can work against the externality correction.** Firms with market power set markups based on demand elasticities rather than external damage. When the resulting markups are poorly aligned with external damage, market power can move prices in the wrong direction. In spirits, for example, premium products have higher dollar markups, while the lowest-price products tend to have the highest $M_j^e$. A markup pattern that does not track external damage can make the welfare gap between a simple attribute tax and the optimal product-specific correction larger than under perfect competition. More generally, markups create wedges between price and private marginal cost that reflect demand conditions rather than marginal external damage. Corrective taxes that ignore these pre-existing distortions can reinforce, rather than offset, the misalignment between private prices and social costs.

**4. Constructing Estimates of $D_{j\to k}$ and $M_j^e$ is easier than it looks.** For our distilled spirits example, we used the estimated demand system from **?** to recover $D_{j\to k}$. In many cases, the data will not be sufficient to estimate a full demand system. For automobiles, we relied instead on second-choice surveys to estimate $D_{j\to k}$. The broader lesson is that the relevant substitution patterns can be measured in more than one way. Estimated demand systems, second-choice data, border-crossing evidence, and product-level responses to tax changes can all help determine whether consumers leave the market, switch to cleaner substitutes, or move toward more harmful alternatives.

## 7 Conclusion

The conventional prescription to correct an externality by taxing the externality-generating attribute relies on complete coverage. When every product can be taxed directly on its own damage, substitutes are corrected at their destination and the tax that maximizes welfare depends only on each product's own damage relative to the untaxed outside option. With incomplete coverage, the reduction in external damage from raising a product's price also depends on where its buyers go and how much damage the substitutes generate. We derive the welfare-maximizing correction when the tax on a single product is adjusted and show that it depends on a product's damage relative to the diversion-weighted damage of its

substitutes. The same logic applies more broadly when tax instruments do not reach all relevant substitutes. A simple regression of the damage of a product's substitutes on its own damage then tells us how well a tax based on the externality-generating attribute can approximate the welfare-maximizing tax.

Attribute-based taxes remain a natural starting point. In many settings they outperform cruder instruments such as ad valorem taxes or regulations that raise prices by increasing market power, and an intercept (a feebate) absorbs a dirty outside option. Simple attribute taxes approximate the welfare-maximizing tax closely when the damage of a product's substitutes is nearly constant, as in distilled spirits, and reasonably well when it is approximately an affine function of the product's own damage, as it is across the automobile fleet. They work less well in other cases, such as the clean-vehicle category, where every battery-electric vehicle has similar emissions but their buyers displace very different alternatives, and in markets such as cannabis and nicotine, where the untaxed alternative may be more harmful than the taxed product. When policymakers choose to tax or subsidize just a subset of relevant products, the external damage generated by untaxed substitutes matters for getting the policy right.

## A Appendix: Deriving the Optimal Taxes in the Two Cases

This appendix derives the first-order conditions for the planner's problem in **??** from **??**. The planner chooses per-unit taxes $\tau_1, \ldots, \tau_J$ on the inside products to maximize:

$$W \;=\; CS(\mathbf{p}) \;+\; \sum_{j=1}^{J} \tau_j q_j \;-\; \lambda_e \sum_{k=0}^{J} e_k q_k, \tag{??}$$

under the maintained assumptions of **??**: (1) perfect competition: $p_j = c_j + \tau_j$, so pass-through is one-for-one ($\partial p_j/\partial \tau_j = 1$ and $\partial p_k/\partial \tau_j = 0$ for $k \neq j$) and producer surplus is identically zero (which is why **??** has no profit term); (2) Roy's identity for aggregate demand: $\partial CS/\partial p_j = -q_j$; (3) discrete choice with a fixed market size: $q_0 = N - \sum_{k=1}^{J} q_k$, so that:

$$\frac{\partial q_0}{\partial \tau_j} \;=\; -\sum_{k=1}^{J} \frac{\partial q_k}{\partial \tau_j}. \tag{A1}$$

**Step 1: The generic first-order condition**

Differentiate **??** with respect to each $\tau_j$:

$$\frac{dW}{d\tau_j} = \underbrace{-q_j}_{\partial CS/\partial p_j} + \underbrace{q_j}_{\text{revenue gain}} + \underbrace{\sum_{k=1}^{J} \tau_k \frac{\partial q_k}{\partial \tau_j}}_{\text{revenue substitution}} - \lambda_e \sum_{k=1}^{J} e_k \frac{\partial q_k}{\partial \tau_j} - \lambda_e e_0 \frac{\partial q_0}{\partial \tau_j}.$$

The mechanical revenue gain $q_j$ exactly offsets the consumer-surplus loss $-q_j$: a small tax increase acts as a transfer from consumers to the government, so only substitution to other goods (and their externalities) matters. Plugging in **??** to eliminate $\partial q_0/\partial \tau_j$ lets us redefine externalities *relative to the outside option* $\tilde{e}_k \equiv (e_k - e_0)$, and perfect competition swaps $\partial q_k/\partial \tau_j$ for $\partial q_k/\partial p_j$ (full pass-through):[30]

$$\frac{dW}{d\tau_j} = \sum_{k=1}^{J} \left(\tau_k - \lambda_e \tilde{e}_k\right) \frac{\partial q_k}{\partial p_j}. \tag{A2}$$

The welfare impact of moving *each rate* depends on *every* product's wedge $\tau_k - \lambda_e \tilde{e}_k$, weighted by how much substitution is sent to $k$.

**Case A: Complete Coverage**

When the planner can set each element $\tau_j$ simultaneously for $j = 1, \ldots, J$, so that $\nabla_{\boldsymbol{\tau}} W = \mathbf{0}$, then the FOCs from **??** define a system of equations:

$$\Omega' \left(\boldsymbol{\tau} - \lambda_e \, \tilde{\mathbf{e}}\right) = \mathbf{0}, \qquad \tilde{\mathbf{e}} \equiv \mathbf{e} - e_0 \mathbf{1},$$

where $\Omega \equiv -\partial \mathbf{q}/\partial \mathbf{p}'$ is the $(J \times J)$ demand Jacobian. For any non-singular $\Omega$, the solution becomes the usual additive/Pigouvian correction (net of the outside good):[31]

$$\tau_j^\star = \lambda_e \left(e_j - e_0\right) \qquad \text{for every } j. \tag{A3}$$

**Case B: Single Taxed Product**

If the planner can only set a single $\tau_j$ while other rates are fixed at $\bar{\tau}_k$, then we solve **??** for a single $\tau_j$ in isolation, beginning by splitting the summation in **??** into own and cross terms

[30] Note that we can equivalently start the summation at $k = 0$ because $\tau_0 = 0$ (the outside good is untaxable) and $\tilde{e}_0 = 0$ (by construction).

[31] A singular $\Omega$ would require perfect substitutes. Under discrete choice with an outside good, $\Omega$ is strictly diagonally dominant column by column — $\sum_{k \neq j} |\Omega_{kj}| = \Omega_{jj} \left(1 - D_{j \to 0}\right) < \Omega_{jj}$ whenever every product loses some demand to the outside good — and strictly diagonally dominant matrices are non-singular.

(inclusive of the outside good):

$$\Big(\tau_j - \lambda_e \tilde{e}_j\Big)\frac{\partial q_j}{\partial p_j} \;+\; \sum_{k\neq j}\Big(\bar{\tau}_k - \lambda_e \tilde{e}_k\Big)\frac{\partial q_k}{\partial p_j} \;=\; 0.$$

If we divide through by $|\partial q_j/\partial p_j|$ (recall $\partial q_j/\partial p_j < 0$) then we can rewrite in terms of diversion ratios $D_{j\to k} = \big(\partial q_k/\partial p_j\big)/\big|\partial q_j/\partial p_j\big|$:

$$\begin{aligned}
\tau_j^\star(\mathbf{p}) \;&=\; \lambda_e\Bigg(\tilde{e}_j - \sum_{k\neq j} D_{j\to k}\,\tilde{e}_k\Bigg) \;+\; \sum_{k\neq j} D_{j\to k}\,\bar{\tau}_k, \\
&=\; \underbrace{\lambda_e\Bigg(e_j - \sum_{k\neq j} D_{j\to k}\,e_k\Bigg)}_{\equiv M_j^e(\mathbf{p})} \;+\; \sum_{k\neq j} D_{j\to k}(\mathbf{p})\,\bar{\tau}_k. \qquad \text{(A4)}
\end{aligned}$$

The second line follows because $\sum_{k\neq j} D_{j\to k} = 1$, so the $e_0$'s in the damage term cancel. It is important to point out that this is an implicit characterization: $\tau_j^\star(\mathbf{p})$ depends on prices (through diversion ratios), which in turn depend on taxes.

In the case where substitutes are untaxed ($\bar{\tau}_k = 0$), **??** collapses to $\tau_j^\star = \lambda_e M_j^e$. Likewise, evaluated at $\boldsymbol{\tau} = \mathbf{0}$, the marginal change in welfare with respect to the tax is $dW/d\tau_j = \lambda_e\, M_j^e(\mathbf{p})\,|\partial q_j/\partial p_j|$, the expression reported in **??**.

**Comparison of the two alternatives**

We can see that the two cases agree when we substitute the full coverage rates $\bar{\tau}_k = \lambda_e \tilde{e}_k$ from **??** for $\bar{\tau}_k$ in **??**:

$$\begin{aligned}
\tau_j^\star(\mathbf{p}) \;&=\; \lambda_e\Bigg(\tilde{e}_j - \sum_{k\neq j} D_{j\to k}\,\tilde{e}_k\Bigg) \;+\; \sum_{k\neq j} D_{j\to k}\,\lambda_e\tilde{e}_k \\
&=\; \lambda_e\tilde{e}_j = \lambda_e(e_j - e_0).
\end{aligned}$$

Thus when all other substitutes are corrected at the complete coverage/Pigouvian rate, the best single-product response is to set $\tau_j^\star(\mathbf{p}) = \lambda_e(e_j - e_0)$ — also at the complete coverage/Pigouvian rate.

Otherwise, set aside the $\bar{\tau}_k$ correction in **??**, which pushes the tax rate up as the tax rates of substitutes rise (because tax revenue is traded off against consumer surplus) and compare the external damage correction terms. When all goods are taxed, we compare $e_j$ to $e_0$, the untaxed outside option, while when alternatives are untaxed we compare $e_j$ to $\bar{e}_j \equiv \sum_{k\neq j} D_{j\to k} e_k$, the diversion-weighted average externality of the substitutes. In the

special case of "independent demands," $D_{j\to 0} = 1$: consumers substitute only to the outside good, and $e_0 = \bar{e}_j$, so the two comparisons coincide.

**Remark: adding a revenue motive**

A straightforward extension asks how things change if we include a tax revenue target $\sum_j \tau_j q_j \geq \overline{R}$, or equivalently a marginal cost of public funds $\lambda_r > 0$, so that tax revenue in **??** is valued at $(1 + \lambda_r)$ per dollar. This adds an additional term and **??** becomes:

$$\frac{dW}{d\tau_j} \;=\; \lambda_r\, q_j \;+\; \sum_{k=1}^{J} \Big( (1+\lambda_r)\,\tau_k - \lambda_e \tilde{e}_k \Big) \frac{\partial q_k}{\partial \tau_j}. \tag{A5}$$

The optimal complete-coverage tax also adds a Ramsey/revenue term:

$$\boldsymbol{\tau}^\star \;=\; \tilde{\lambda}_e\, \tilde{\mathbf{e}} \;+\; \frac{\lambda_r}{1+\lambda_r}\, \Omega^{-1}\mathbf{q}, \qquad \tilde{\lambda}_e \equiv \frac{\lambda_e}{1+\lambda_r}. \tag{A6}$$

The corrective component survives product-by-product and hits the attribute $e_j$ only, but at the discounted rate $\tilde{\lambda}_e$ (**?**). The Ramsey component $\Omega^{-1}\mathbf{q}$ is the multi-product equivalent of the inverse-elasticity rule and depends on the entire substitution structure $\Omega$ rather than $j$ or $e_j$ alone.

This is just the differentiated products analog of the additivity results from **?** and **?**: address the external damage with a correction that depends only on the externality generated by that good — and the (untaxed) outside alternative: $(e_j - e_0)$, and rely on a Ramsey rule to raise revenue. The revenue motive changes the corrective *rate* ($\tilde{\lambda}_e$ vs $\lambda_e$) while leaving the corrective *base* $(e_j - e_0)$ untouched.

The Case B/single-taxed product statistic is modified in much the same way: at $\boldsymbol{\tau} = \mathbf{0}$, **??** gives $dW/d\tau_j = \lambda_r q_j + \lambda_e \left|\partial q_j/\partial p_j\right| M_j^e$, a revenue term plus the same substitution-adjusted externality as before. Our applications value revenue at $(1 + \lambda_r)$ with $\lambda_r$ calibrated in **??**; **??** traces the schedule comparisons across $\lambda_r$.

**Remark: adding market power**

With imperfect competition, producer surplus $\Pi = \sum_k (p_k - c_k - \tau_k)\, q_k$ enters welfare and prices respond to taxes through an equilibrium pass-through matrix $\Psi \equiv \partial \mathbf{p}/\partial \boldsymbol{\tau}$. Differentiating, the pass-through terms in $\partial CS/\partial \tau_j = -\sum_k q_k \Psi_{kj}$ and in the mechanical part of $\partial \Pi/\partial \tau_j$ cancel exactly—incidence between consumers and firms is a transfer—leaving $\partial(CS + \Pi)/\partial \tau_j = -q_j + \sum_k \eta_k\, \partial q_k/\partial \tau_j$ with $\eta_k = p_k - c_k - \tau_k$. The extra term is a markup harvest: taxing $j$ reshuffles quantities across goods priced above cost. Carrying it through

?? gives

$$\frac{dW}{d\tau_j} \;=\; \lambda_r q_j \;+\; \sum_{k=1}^{J} \Big( (1+\lambda_r)\,\tau_k + \eta_k - \lambda_e \tilde{e}_k \Big) \frac{\partial q_k}{\partial \tau_j},$$

so the markup rides along as a pre-existing wedge valued at one rather than $(1+\lambda_r)$, and pass-through never appears—only the equilibrium quantity response $\widetilde{\Omega} \equiv -(\partial \mathbf{q}/\partial \boldsymbol{\tau}')'$ does.

Setting all $J$ derivatives to zero gives **??**. Solving the $j$-th equation alone with the other tax rates held fixed at $\bar{\tau}_k$ gives **??**, with diversion ratios constructed from the columns of $\partial \mathbf{q}/\partial \boldsymbol{\tau}'$.

## B Appendix: Spirits with a Dirty Outside Option

**??** repeats the bottom panel of **??** with the outside option carrying ethanol at the beer/wine calibration ($e_0 = 0.15$ liters of ethanol per liter, i.e. 30 proof). Every product's destination damage shifts up by $D_{j\to 0} \times 30$ proof, and the exit-margin dispersion compresses (RMS residual 2.4 proof points against 3.7): the exit-margin penalty scales with the content gap between inside destinations and the outside option.

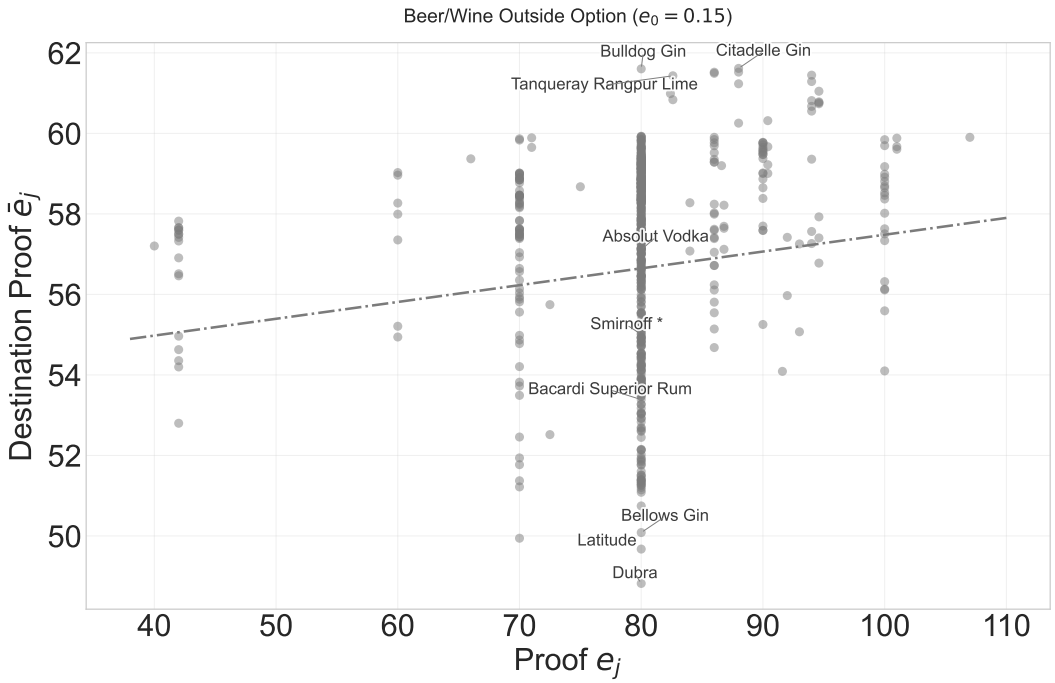



Figure A1: Destination proof under the beer/wine outside calibration ($e_0 = 0.15$; $\hat{\beta} = 0.04$, RMS residual 2.4 proof points).

## C Appendix: Raw-Survey Diversion Figures

The figure below repeats **??** using the raw second-choice diversion ratios in place of the CMS-fitted values, as a robustness check on the fleet tagging exercise. Regressing $M_j^e$ on own emissions (conditional on an inside purchase) gives $R^2 = 0.42$ from the raw survey diversion and 0.47 once it is smoothed through the CMS model; the body reports the equilibrium-model value at $e_0 = 450$ ($R^2 = 0.47$).

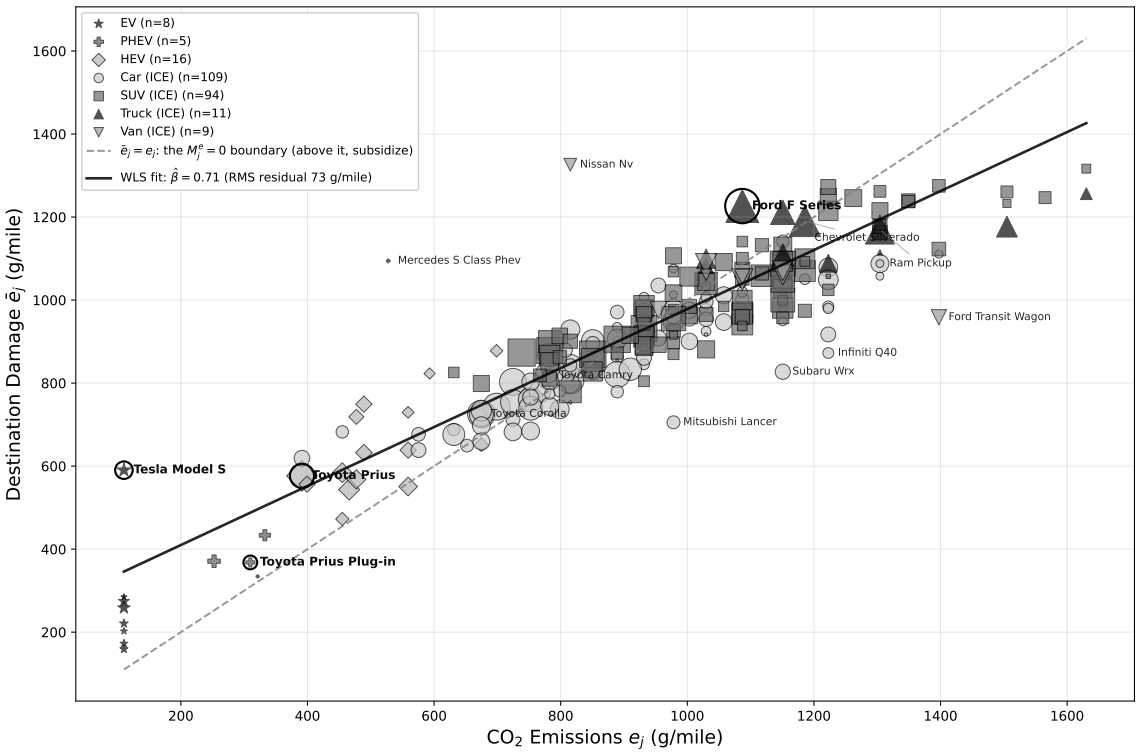



Figure A2: Where the marginal buyer goes, raw survey diversion. Destination damage $\bar{e}_j$ against own emissions $e_j$, computed directly from the observed MaritzCX second-choice diversion ratios (conditional on an inside purchase; 252 vehicles), paralleling **??**. The sales-weighted fit has slope $\hat{\beta} = 0.71$ with RMS residual 73 g/mile; the equivalent regression of $M_j^e$ on $e_j$ has sales-weighted $R^2 = 0.42$.

## D Appendix: The Ford F-Series Decomposition

This appendix decomposes the Ford F-Series substitution-adjusted externality from the second-choice survey (conditional on an inside purchase): the observed diversion in **??** sums to $M_j^e = -140$ and the CMS-fitted diversion to $-125$. In the equilibrium model with a dirty outside option ($e_0 = 450$) used in **??**, the roughly 9% of buyers who leave for the used-vehicle outside good contribute a positive $D_{j\to 0}(1{,}087 - 450) \approx +59$, which does not flip the sign: the net substitution-adjusted externality is $\approx -49$ g/mile.

**??** illustrates the mechanism using the Ford F-Series, the best-selling vehicle in the United States. On an attribute basis, the F-Series looks like a natural target for a high emissions tax. It is among the highest-emitting vehicles in the sample, with $e_j = 1{,}087$ g/mile. But its marginal externality depends on what F-Series buyers would purchase instead.

Nearly 90% of F-Series diversion stays within the pickup segment, and at 18 MPG the F-Series is *cleaner* than most of its closest substitutes. The Ram (15 MPG), Tundra (13 MPG), and Titan (12 MPG) are all dirtier. The Ram and Tundra alone contribute $-108$ of the total $M_j^e = -140$. Consumers who leave the pickup segment entirely, for crossovers, cars, or electric vehicles, contribute positively to emissions reductions, but they account for less than 6% of diversion. As a result, once substitution is taken into account, the vehicle with perhaps the strongest claim to a high emissions tax on attribute grounds is actually a vehicle whose marginal buyer *reduces* fleet emissions relative to their next-best alternative.

To be clear, this does not mean that a $CO_2$-proportional tax raises emissions through the F-Series. Under such a tax, the F-Series' dirtier substitutes also pay more, so the resulting relative prices steer pickup buyers toward the F-Series. That within-segment steering is much of what an attribute tax gets right. The problem is that an attribute-proportional schedule still taxes the F-Series as if its marginal buyer were moving from the F-Series to a clean alternative. In the data, the opposite is true: the marginal F-Series buyer reduces fleet emissions by 140 g/mile relative to their next-best alternative. The optimal schedule would tax the Ram and Tundra more heavily relative to the F-Series, while setting the F-Series' own rate far below its attribute-proportional level.

Table A1: Where Do Ford F-Series Buyers Go? A Decomposition of $M_j^e$

| **Substitute** | $D_{j\to k}$ **(%)** | $D_{j\to k}^{CMS}$ **(%)** | **MPG** | $e_j - e_k$ **(g/mi)** | $D_{j\to k}(e_j - e_k)$ |
|---|---|---|---|---|---|
| Ram Pickup | 24.6 | 21.4 | 15 | −217 | −53.5 |
| GMC Sierra | 20.3 | 23.3 | 17 | −64 | −13.0 |
| Chevrolet Silverado | 15.6 | 16.9 | 16 | −99 | −15.4 |
| Toyota Tundra | 13.0 | 11.2 | 13 | −418 | −54.3 |
| Toyota Tacoma | 6.3 | 5.9 | 17 | −64 | −4.0 |
| Chevrolet Colorado | 4.6 | 4.8 | 19 | +57 | +2.7 |
| GMC Canyon | 2.3 | 1.8 | 19 | +57 | +1.3 |
| Nissan Frontier | 1.6 | 1.9 | 16 | −136 | −2.2 |
| Jeep Wrangler | 1.6 | 1.2 | 17 | −64 | −1.0 |
| Nissan Titan | 0.7 | 0.7 | 12 | −543 | −3.8 |
| SUVs ($n = 25$) | 3.5 | 3.5 | 16 | −116 | −4.1 |
| Crossovers (CUVs) ($n = 68$) | 2.8 | 2.9 | 20 | +91 | +2.6 |
| Cars ($n = 108$) | 2.6 | 3.6 | 22 | +141 | +3.6 |
| Vans ($n = 9$) | 0.3 | 0.3 | 18 | −15 | −0.0 |
| Electric ($n = 31$) | 0.2 | 0.5 | 43* | +718 | +1.6 |
| **Total** ($= M_j^e$) | **100.0** | **100.0** | – | – | −139.6 |

*Notes:* Second choices of Ford F-Series buyers from the MaritzCX survey, conditional on inside-good purchase. The F-Series (median trim, 2015) is rated at 18 MPG combined, or $e_j$ = 1,087 g $CO_2$/mile. $D_{j\to k}$ is the observed diversion ratio; $D_{j\to k}^{CMS}$ is the fitted value from the CMS estimator ($I = 38$, 1SE rule). $e_j - e_k$ is the $CO_2$ gap (F-Series minus substitute, g/mile); negative entries indicate a dirtier substitute. The final column uses observed diversion and sums to $M_j^e = -139.6$; using CMS fitted diversion instead gives −125.0. Segment rows consolidate the remaining substitutes, with MPG and $CO_2$ gaps averaged using $D_{j\to k}$ weights. *The Electric segment includes hybrid, plug-in hybrid, and battery-electric models; its MPG averages EPA combined ratings for hybrids with MPGe-based figures for battery-electrics, while the $CO_2$ gap instead assigns battery-electrics grid-based emissions of 110 g/mile and plug-in hybrids a utility-factor blend, so the $CO_2$ column is not the MPG-implied value for those products.

## E Appendix: Robustness to the Cost of Public Funds

The body reports the clean-category schedule comparison at $\lambda_r \in \{0, 0.25, 0.5\}$. **??** reports the full ladder across $\lambda_r \in [0, 0.6]$, including the pooled (single-intercept) feebates omitted from **??**: each feasible schedule's welfare as a share of the product-specific (Ramsey) clean optimum, whose level—the ceiling $W^{\star}_{\lambda_r}$—falls as public funds become more expensive. The substitution-adjusted feebate's advantage over the emissions feebate grows with $\lambda_r$: the emissions feebate is best at $\lambda_r = 0$ and declines, while the substitution-adjusted feebate rises monotonically, overtaking it between $\lambda_r = 0.4$ and 0.5 and reaching 0.950 by $\lambda_r = 0.6$ against the emissions feebate's 0.352; the tiered versions cross between $\lambda_r = 0.3$ and 0.4. A costly subsidy must be rationed, and the substitution correction is what pins the rationing. **??** gives the optimal uniform and per-category (BEV/PHEV/HEV) rebates across the same range.

Table A2: Clean-category schedules across the cost of public funds (share of the Ramsey ceiling)

| **Schedule** | $\lambda_r$=0 | $\lambda_r$=0.1 | $\lambda_r$=0.2 | $\lambda_r$=0.3 | $\lambda_r$=0.4 | $\lambda_r$=0.5 | $\lambda_r$=0.6 |
|---|---|---|---|---|---|---|---|
| Ceiling $W^{\star}_{\lambda_r}$ (\$B) | 12.58 | 8.07 | 5.04 | 2.98 | 1.61 | 0.76 | 0.34 |
| Uniform | 0.687 | 0.687 | 0.678 | 0.659 | 0.617 | 0.516 | 0.247 |
| Tiered (BEV/PHEV/HEV) | 0.907 | 0.906 | 0.900 | 0.882 | 0.832 | 0.690 | 0.316 |
| $c_0 + c_1 e$ | 0.954 | 0.952 | 0.945 | 0.924 | 0.872 | 0.730 | 0.352 |
| $c_0 + c_1 M^e_j$ | 0.727 | 0.734 | 0.741 | 0.753 | 0.780 | 0.844 | 0.950 |
| Tiered feebate $c_0 + c_1 e$ | **0.963** | **0.963** | **0.958** | **0.940** | 0.890 | 0.746 | 0.357 |
| Tiered feebate $c_0 + c_1 M^e_j$ | 0.909 | 0.912 | 0.914 | 0.914 | **0.916** | **0.929** | **0.961** |
| Combined $c_0 + c_1 e + c_2 M^e_j$ | 0.961 | 0.962 | 0.962 | 0.959 | 0.955 | 0.955 | 0.968 |

*Notes:* Each entry is a feasible schedule's welfare $W_{\lambda_r}$ as a share of the product-specific clean optimum at that $\lambda_r$; the ceiling row gives the optimum's level in \$B. Bertrand equilibrium, $e_0 = 450$ g/mile. Welfare levels are indicative (the demand fit targets survey diversion, not registration shares); the across-schedule shares are the robust object.

Table A3: Optimal uniform and per-category clean-vehicle rebates across the cost of public funds

| $\lambda_r$ | **Uniform** | **BEV** | **PHEV** | **HEV** |
|---|---|---|---|---|
| 0 | 11.3 | 20.1 | 15.1 | 9.9 |
| 0.1 | 9.3 | 17.2 | 12.8 | 8.2 |
| 0.2 | 7.6 | 14.6 | 10.7 | 6.6 |
| 0.3 | 5.9 | 12.1 | 8.7 | 5.2 |
| 0.4 | 4.3 | 9.5 | 6.7 | 3.9 |
| 0.5 | 2.8 | 6.4 | 4.5 | 2.6 |
| 0.6 | 1.3 | 2.9 | 2.2 | 1.2 |

*Notes:* Optimal single-rate (uniform) and per-category rebates ($1,000s per vehicle; positive = subsidy) as a function of $\lambda_r$, $e_0 = 450$ g/mile.

# F Appendix: Calibrating the Marginal Damage Rate and the Cost of Public Funds

The corrective-tax calculations price the emissions externality at a constant marginal-damage rate $\lambda_e$, the dollar value of a unit of external damage from **??** and, where a positive marginal cost of public funds is assumed, value government revenue at $(1 + \lambda_r)$ per dollar. This appendix documents both calibrations.

## F.1 The marginal damage rate $\lambda_e$

Vehicle emissions are measured as an intensity $e_j$ (grams of $CO_2$ per mile). The lifetime marginal damage of that intensity, per vehicle sold, is the product of the social cost of carbon and expected lifetime mileage,

$$\lambda_e \;=\; \mathrm{SCC} \times \mathrm{VMT}_{\mathrm{lifetime}}.$$

We set the social cost of carbon to **$190** per metric ton of $CO_2$ (2020 dollars), the central estimate in the **?** report on the social cost of greenhouse gases, which incorporates the updated damage and discounting framework of **?** at a 2% near-term Ramsey discount rate. The earlier interagency interim value of $51 per ton at a 3% discount rate (**?**) provides a low-damage robustness case. We set expected lifetime mileage to **150,000 miles**, a round figure slightly below the National Highway Traffic Safety Administration's expected-lifetime estimates—about 152,000 miles for passenger cars and 180,000 for light trucks (**?**)—and therefore conservative for a fleet that is roughly half light trucks.

The arithmetic is transparent: one gram per mile sustained over 150,000 miles emits 0.15

metric tons of $CO_2$, so

$$\lambda_e \;=\; \$190/\text{t} \times 0.15 \text{ t per (g/mile)} \;=\; \$28.50 \text{ per (g/mile)}.$$

A product's Pigouvian rate is $\tau_j = \lambda_e\,(e_j - e_0)$, where $e_0$ is the emissions intensity of the outside option (a retained used vehicle). These rates are highly non-marginal: at $\lambda_e = \$28.50$, the dirtiest vehicles in the sample ($e_j \approx 1{,}100$ g/mile) carry Pigouvian taxes near \$30,000, a large fraction of the vehicle price. Under the low-damage social cost of carbon the rate falls proportionally to $\lambda_e = \$7.65$ per (g/mile).

We set the outside option's intensity to $e_0 = 450$ g/mile. The 2015 on-road light-duty fleet averages roughly 890 g $CO_2$/mile, but the scrappage literature finds that a forgone new-vehicle purchase is only partly offset by reduced driving—about half is recovered as extended use of a retained used vehicle (**?**)—so a marginal exit from the new-car market carries roughly half the used fleet's intensity, $\approx 450$ g/mile.

### F.2 The marginal cost of public funds $\lambda_r$

When a corrective schedule spends or raises revenue, the fiscal cost is not one-for-one: financing a dollar of subsidy (or rebating a dollar of tax) through the rest of the distortionary tax system costs $1 + \lambda_r$, where $\lambda_r$ is the marginal cost of public funds—equivalently, the marginal excess burden of taxation. The welfare criterion accordingly values net government revenue at $(1 + \lambda_r)$,

$$W_{\lambda_r} \;=\; \Delta\text{CS} + \Delta\text{PS} + (1 + \lambda_r)\,\text{Rev} - \Delta\mathcal{D},$$

where $\Delta\mathcal{D} = \lambda_e \sum_j e_j\,\Delta q_j$ is the change in monetized emissions damage.

Estimates of $\lambda_r$ vary with the method and the assumed behavioral elasticities. General-equilibrium computations for the United States place it near 0.17–0.33 (**?**) and 0.10–0.30 (**?**); incorporating tax avoidance and larger taxable-income elasticities raises it substantially, in some calculations above one (**?**). Surveys of the international evidence cluster near 0.2–0.3 (**?**). Rather than commit to a single value, we report results across $\lambda_r \in \{0, 0.25, 0.5\}$ in the body and the full range $\lambda_r \in [0, 0.6]$ in **??**. The upper end of that grid is high relative to the central estimates above, but one justification for it is the generosity of the actual §30D program. The enacted credit of up to \$7,500 per battery-electric vehicle is close to the welfare-maximizing BEV rebate of \$6,400 at $\lambda_r = 0.5$ (**??**); at $\lambda_r = 0$ the welfare-maximizing BEV rebate would be about \$20,000, and the plug-in-only program would spend \$32B rather than the roughly \$0.66B it actually cost (**??**). A high $\lambda_r$ raises the fiscal cost

of any subsidy and shrinks its optimal size, and is in that sense conservative about the scale of intervention. The role of a positive $\lambda_r$ in the design of corrective taxes connects to the literature on environmental taxation in second-best fiscal systems (**??**).