Attending to Inattention: Identification of Deadweight Loss under Non-Salient Taxes

Giacomo Brusco and Benjamin Glass*†

March 14, 2018

Abstract

Taxes create deadweight loss by distorting consumer choice, so to the extent that consumers perceive taxes to be lower than they really are, deadweight loss declines. Deadweight loss is a convex function of perceived taxes, so its aggregate magnitude depends not only on the average tax misperception, but also on its heterogeneity among consumers. Aggregate data cannot reveal this heterogeneity, yet one can infer lower and upper bounds on deadweight loss relying solely on properties of aggregate demand. Sufficiently rich individual-level data permit identification of deadweight loss even with heterogeneous tax misperceptions. Under strong assumptions on the joint distribution of tax salience and preferences, survey data illustrate that tax salience heterogeneity can yield deadweight loss twice as large as one would calculate under the assumption of a homogeneous perceived price. Relaxing these assumptions, even slightly, yields much more destructive results: the unconstrained upper bound of deadweight loss is more than fifty times larger than the lower bound one would compute assuming homogeneous perceptions of price.

1 Introduction

Taxing a good results in a loss of economic efficiency whenever it distorts equilibrium behavior away from the Pareto optimum. If each person faced a different tax rate when buying a certain good, understanding the welfare effects of such a tax would require us to study not only aggregate demand responsiveness, but the demand responsiveness of every individual. Surely, imposing a high tax on low elasticity individuals and a low tax on high elasticity individuals would have a very different effect on welfare than doing the opposite.

A similar reasoning applies when all agents face the same tax rate, but perceive

*University of Michigan

†The work in the following pages would have been impossible without the thought-provoking conversations we had with James Hines, Joel Slemrod, Dmitry Taubinsky, Alex Rees-Jones, Jeremy Fox, Tilman Börgers, Ying Fan, William Boning, Luis Alejos, Tejaswi Velayudhan, Aristos Hudson, the participants of the Public Finance and Theory seminars at the economics department of the University of Michigan, and the participants of the 2017 NTA conference. We warmly thank all of them for their invaluable advice.

different tax rates, as noted in Chetty et al. (2009). Correctly assessing the welfare effects of taxation requires us to understand how each person reacts to the tax based on both their preferences and their perception of the tax. The possibility of heterogeneous attention intrinsically changes the way we should estimate deadweight loss. One can wildly miscalculate the magnitude of deadweight loss if they fail to account for the heterogeneous perception of taxes. Suppose, e.g., that aggregate tax responsiveness were half of sticker price responsiveness. To a second order approximation, deadweight loss is quadratic in the perceived tax. If all consumers noticed only half of the tax, deadweight loss would be 25% of what it would be under full attention; if instead half of consumers did not notice the tax at all, while the other half noticed it perfectly, deadweight loss would be 50% of what it would be under full attention. Besides varying across individuals, attention may also endogenously co-vary with the tax rate.

Tax non-salience leaves consumers even worse off because they fail to avoid the tax by reducing consumption. Yet to what degree the tax does not distort consumer behavior, deadweight loss is reduced. In fact, we find that increasing an existing tax, while still increasing deadweight loss, can make consumers better off when it discourages consumption that would already be avoided if agents paid full attention.

Suppose the econometrician only has aggregate data: that is, data on aggregate consumption of a good across a number of markets, differentiated by time or geographic location. These data may indicate average responsiveness to taxes and sticker prices, but provide no information on the heterogeneity of demand responsiveness to taxes, nor about its correlation with sticker price responsiveness. Thus, aggregate data cannot precisely identify deadweight loss. However, we can provide bounds for deadweight loss even with aggregate data. These bounds are tight, in the sense that for each bound there is a distribution of salience under which that bound accurately describes deadweight loss.

The lower bound for deadweight loss is the calculation one would perform in the case of a representative consumer. Since the loss in efficiency is a convex function of the perceived tax rate, the calculation of deadweight loss from one perceived tax-inclusive price consistent with aggregate demand will generically underestimate deadweight loss. Intuitively, heterogeneity in tax salience creates heterogeneity in perceived net-of-tax prices. This creates an allocative inefficiency across consumers: it is no longer true that the people who value the good the most are the ones who end up buying it. As the calculation with a representative consumer only accounts for inefficiency from *aggregate* foregone consumption due to the tax, it will under-estimate excess burden.[1] However, in the case in which all agents pay the same amount of attention to the tax, there is no allocative inefficiency between consumers, and so performing the calculation as with a representative consumer yields the correct value for deadweight loss.

We obtain an upper bound for deadweight loss by assuming that tax salience has support on a known bounded non-negative interval. The upper bound comes from maximizing perceived price heterogeneity, again exploiting the convexity of deadweight loss with respect to the perceived tax. Additionally to assuming tax salience is either zero or maximal, a distribution yielding the upper bound for deadweight loss assigns high

[1] While we came up with our lower bound independently, we thank Taubinsky and Rees-Jones (2018) for inspiring this intuitive insight by making a related point as it pertained to their model. They note in proposition 7 of their appendix that deadweight loss can vary between lower and upper bounds when the choice set is binary.

tax salience to those agents whose particular preferences yield maximal deadweight loss from that agent relative to the change in consumption of that agent. In other words, this distribution allocates high tax salience to those agents who have more convex demand curves, keeping the aggregate change in quantity demanded constant. Finally, any agents with multiple optimal decisions at the perceived price consume the highest amount consistent with their preferences when they perceive low prices, whereas they consume the lowest amount consistent with their preferences when they perceive high prices. This illustrates the fact that heterogeneity in perceived prices permits different equilibria with the same sticker price, tax rate, and aggregate consumption, yet yielding different values of deadweight loss due to different distributions of consumption among individuals.

To obtain an understanding of the possible magnitude of the uncertainty in deadweight loss estimated with aggregate data, we perform an empirical exercise with experimental data from Taubinsky and Rees-Jones (2018). For the good we observe, we estimate an average deadweight loss from sales taxes on that good across the country of between 0.38 cents and 20.79 cents per consumer. This means the upper bound of deadweight loss is around 55 times as large as the lower bound for deadweight loss. In fact, the upper bound for deadweight loss is substantially larger than if all agents were to perfectly account for the sales tax, in which case deadweight loss would be 1.87 cents.

When able to observe consumption choices at the individual level with long panel data, one can point-identify deadweight loss. The easiest way to see this is to imagine one could observe the choice made by individual consumers for infinitely many periods, each time under a different tax and sticker price regime. Then, one would be able to infer sticker price and tax responsiveness for each individual, and would therefore be able to compute deadweight loss for each single agent, and thus for the entire population.

If we are willing to impose linear structure on the choice behavior of agents, then one need not observe each individual responsiveness to sticker prices and taxes: it will suffice to know their distribution in the population. For this, one need not follow the same agents across time; cross-sectional data with sufficiently rich variation in taxes and sticker prices will allow the identification of the distribution of responsiveness, which in turn would allow us to integrate for expected deadweight loss.[2]

An alternative to observing individual choices is to restrict the choice set that the agent is facing. Estimating random coefficients discrete choice models is common practice, and can be achieved with aggregate data on market shares. The estimation of such a model allows the econometrician to then compute expected deadweight loss. This is because in the case of choices over a discrete set, aggregate data reveals the distribution of consumption on an individual level. For instance, if the good in question is a house, and assuming everyone either buys one house or does not buy a house at all, knowing the amount of houses purchased and the size of the population of perspective buyers allows us to calculate the probability that an agent buys a house. However, the econometrician still requires linearity assumptions. In this context, these assumptions are equivalent to assuming that tax salience is independent of both the tax rate and the sticker price.

[2]See for instance Beran and Hall (1992), Beran, Feuerverger, and Hall (1996), and Hoderlein, Klemelä, and Mammen (2010), all of which would require full support of sticker prices and sales taxes.

This paper complements a growing literature in public finance on non-salient taxes. Rosen (1976) does not find evidence of limited tax salience, but Chetty et al. (2009), Finkelstein (2009), Gallagher & Muehlegger (2011), Goldin & Homonoff (2013), and Taubinsky & Rees-Jones (2018) all find strong evidence of dramatically limited tax salience. For instance, Chetty et al. (2009) provide a summary estimate that agents perceive six percent of sales tax variation using a log-log specification. We estimate a linear specification using their data and find a summary estimate of 27 percent.

Recent theoretical papers on the efficiency implications of non-salient taxes include Chetty et al. (2007), Farhi and Gabaix (2015), and Taubinsky and Rees-Jones (2018). Our theoretical model comes from Gabaix (2014), which differs slightly from Chetty et al. (2009) in how it handles income effects. In our model, a sales tax of any positive salience creates deadweight loss.[3] In contrast, Goldin (2015) demonstrates that a model developed by Chetty et al. (2009) generically yields zero deadweight loss for some positive salience for a sales tax on a normal good.[4] We show that the Gabaix (2014) model provides a general description of behavior when agents have convex preferences, and proceed to use it for welfare analysis.

Taubinsky and Rees-Jones (2018) is most similar to this paper in spirit. They make similar points about the inability to identify deadweight loss with aggregate data due to the role played by heterogeneous attention. They also find lower and upper bounds for a second order approximation to deadweight loss with a binary choice set. We find similar bounds without imposing restrictions on the choice set.

This paper proceeds as follows. In section 2, we develop the decision theory model and theoretically derive deadweight loss. In section 3, we illustrate how heterogeneity in attention prevents identification of deadweight loss with aggregate continuous choice data. We then show how we can still identify lower and upper bounds to deadweight loss under weak assumptions. We provide positive point identification results in section 4 using individual-level or binary choice data.[5] We perform an empirical calculation in section 5 before concluding in section 6. We relegate proofs and details of empirical work to the appendix.

2 Theoretical Derivation of Deadweight Loss

This section describes the theoretical model and results that underlie the rest of this paper. The main modeling challenge in dealing with misperceived prices is to allow for the misperception of prices while keeping agents solvent. Chetty et al. (2007, 2009) get around this issue by having a single good "absorb" all optimization mistakes. Gabaix (2014), instead, has agents conjecture themselves a certain income such that they end up consuming on their true budget constraint when presented with the relative prices they perceive. We begin with a general model of decision-making problem for a single agent that encompasses both the Chetty et al. (2007, 2009) and Gabaix (2014). The

[3]This claim assumes an undistorted economy without perfectly inelastic supply or demand.

[4]This result depends on a two-good framework in which the untaxed good absorbs any budget shortfalls. In particular, Goldin's (2015) result relies on using the ratio of uncompensated demand responses to tax and sticker price variation. In our model, tax salience corresponds to a ratio of compensated demand responses.

[5]It may appear strange that Taubinsky and Rees-Jones (2018) find a negative identification result with binary choice data, whereas we find a positive identification result. The distinction comes from whether one assumes that tax salience does not depend on the sticker price or tax rate.

choice behavior of the agent is modeled in a general manner, so that the agent correctly optimizes for the goods without non-salient taxes, but has an arbitrary (continuous) consumption function for the taxed good. We then show that the choice behavior from such a model can be represented by the model from Gabaix (2014), which dramatically simplifies exposition. In the process, we define compensating variation and analytically characterize it using the Gabaix (2014) model. Finally, we aggregate over agents and account for the change in government revenue in order to derive aggregate deadweight loss due to the tax.

2.1 Choice under Misperceived Taxes

We consider a decision problem over two goods, one with a non-salient sales tax. We generalize to multiple taxed and non-taxed goods in the appendix. The agent has closed consumption set $X = X^T \times \mathbb{R}_+ \subseteq \mathbb{R}^2_+$. The agent has choice function for the taxed good $q(\bar{p}, p^{NT}, \tau, W)$, with $(\bar{p}, p^{NT}) \in \mathbb{R}^2_{++}$, where $\bar{p}$ and p^{NT} are the sticker price of the taxed and non-taxed good, respectively, $\tau \in \mathbb{R}$ is the sales tax on the taxed good, and W is the income of the agent.[6] We express taxes as if they were specific, so that $\bar{p} + \tau$ is the tax-inclusive price of the taxed good.

The agent has continuous and strictly monotonic preferences $\succeq$ with continuous utility representation $u(q, q^{NT})$, where q denotes generic consumption of the taxed good. The choice vector function $\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W) = (q(\bar{p}, p^{NT}, \tau, W), q^{NT}(\bar{p}, p^{NT}, \tau, W)) \in X$ satisfies two requirements. One, the agent always spends all available income:

$$(\bar{p} + \tau) q(\bar{p}, p^{NT}, \tau, W) + p^{NT} q^{NT}(\bar{p}, p^{NT}, \tau, W) = W$$

In the appendix, we generalize this point so that agents optimally choose q^{NT} given their choice of q. Two, the agent correctly optimizes in the choice of all consumption bundles when there is no tax:

$$\boldsymbol{q}(\bar{p}, p^{NT}, 0, W) \in \underset{\{\bar{p}*q + p^{NT}*q^{NT} \leq W\}}{\arg\max} u(q, q^{NT})$$

We now want some measure of the incidence of the tax on the consumer. For concreteness, we consider the compensating variation due to the tax with complete pass-through, defined as:[7]

$$\Delta CS \equiv \inf\{\Delta W | u(\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W + \Delta W)) \geq u(\boldsymbol{q}(\bar{p}, p^{NT}, 0, W))\}$$

In words, the change in consumer surplus is the greatest lower bound of the amount of money we must provide the agent so that the agents achieves the utility from before the imposition of the tax. One can verify that compensating variation is always non-negative for any non-negative tax when preferences are locally non-satiated.[8]

[6]We implicitly restrict consideration to sticker prices, taxes, and income such that q$(\bar{\boldsymbol{p}}, \tau, W)$ is well-defined at those values.

[7]We impose assumptions in subsection 2.2 that would allow us to define compensating variation using a minimum, rather than an infimum.

[8]For any $\Delta W < 0$, $(\bar{p} + \tau, p^{NT}) * \boldsymbol{q}(\bar{p}, p^{NT}, \tau, W + \Delta W) = W + \Delta W < W = (\bar{p} + \tau, p^{NT}) * \boldsymbol{q}(\bar{p}, p^{NT}, 0, W)$, and so $u(\boldsymbol{q}(\bar{\boldsymbol{p}}, \tau, W + \Delta W)) < u(\boldsymbol{q}(\bar{\boldsymbol{p}}, 0, W))$.

2.2 Gabaix Representation of Choice Behavior

We now provide sufficient conditions under which the above model of choice can be written in the framework of Gabaix (2014) without loss of generality. In this model, the agent perceives price p^s for the non-saliently taxed good, correctly perceives price p^{NT} for the non-taxed goods, and conjectures an income W^s so that the agent's true budget constraint is satisfied while the agent optimizes subject to the perceived budget constraint. Formally:[9]

$$\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W) \in \underset{p^s q + p^{NT} q^{NT} \leq W^s}{\arg\max} u(q, q^{NT})$$
$$\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W^s) * (\bar{p} + \tau, p^{NT}) = W$$

To find a Gabaix representation for a choice, we want to find a budget line through the consumption bundle such that all strictly preferred bundles lie strictly above this line. This is analogous to finding an equilibrium to a single-agent endowment economy. As in the Second Welfare Theorem, we require some sense of continuous convexity of preferences. However, we do not want to assume that the choice set is convex, to later allow for treatment of discrete choice sets. Instead, we impose convexity via the utility representation:[10]

Proposition 1. *Suppose we can extend u to $\mathbb{R}^2$ such that u is continuous and quasi-concave. Then for any $\bar{p}$, p^{NT}, τ, and W on which $\boldsymbol{q}$ is defined, there exist scalar values p^s and W^s such that:*

$$\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W) \in \underset{p^s q + p^{NT} q^{NT} \leq W^s}{\arg\max} u(q, q^{NT})$$
$$\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W^s) * (\bar{p} + \tau, p^{NT}) = W$$

The proof idea is that the set of points $\{\boldsymbol{q'} \in \mathbb{R}^2 | u(\boldsymbol{q'}) > u(\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W))\}$ is open and convex. Since it's convex, there is a budget line separating it from $\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W)$. Since it is open, that set can never touch this budget line. Finally, we can always shift the budget line down if need be so that it goes through the point $\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W)$.

It is interesting to note that proposition 1 does not rule out alternative explanations for consumer behavior. For instance, Chetty, Looney, and Kroft (2009) have a model in which $W^s = W$. The consumer first buys the taxed good knowing the available total income, then re-optimizes upon discovering how little income remains after purchasing the taxed good. Instead, the proposition notes that any such model satisfying minimal conditions is observationally equivalent to the Gabaix model. Intuitively, this model states that the agent doesn't notice some fraction of the tax-inclusive price, instead simply figuring that the extra amount spent on the taxed good was never in the bank in the first place. This story may appear unlikely, but one need not take it literally. Instead, it is a representation of consumer behavior from a general model. Still, the

[9] Technically, for the subjective price p^s, Gabaix (2014) also wanted to choose W^s maximally amongst all values of W^s that would satisfy these two equations. The idea is for the agent to conjecture income so as to optimally satisfy the true budget constraint. This distinction is moot if the taxed good is weakly normal, since then there is at most one conjectured income that satisfies the true budget constraint.

[10] In the appendix, we generalize this result with a description of properties of the choice set and preferences. The key idea is that there is an open and convex set containing all the bundles strictly preferred to $\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W)$.

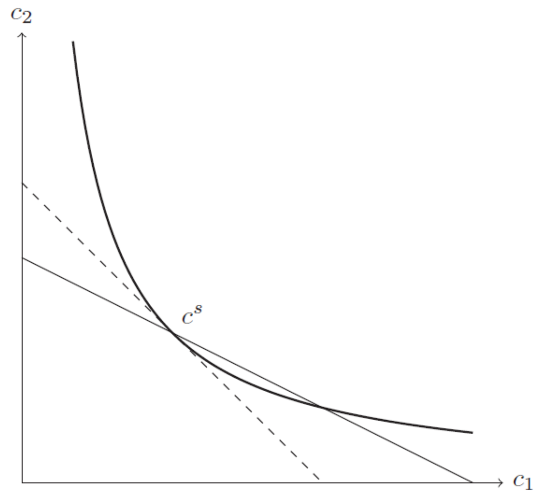


Figure 1: Consumer's decision under limited attention in Gabaix's model. The solid line represents the consumer's real budget constraint. For any chosen point c^s on a convex indifference curve, one can always find a supporting hyperplane, represented by the dotted line, that rationalizes the point as a budget line. Source for image: Gabaix (2014)

intuition helps us understand our characterization of the compensating variation due to the tax. In the following proposition, we introduce terminology with some arguments suppressed for ease of reference.

Proposition 2. *Let $e(p)$ and $h(p)$ denote the expenditure function and compensated demand for the taxed good respectively at price p for the taxed good and price p^{NT} for the other good, so that the agent is minimally compensated so as to achieve utility of at least $u(\boldsymbol{q}(\bar{p}, p^{NT}, 0, W))$.*[11] *Then compensating variation due to the tax satisfies:*

$$\Delta CS = (\bar{p} + \tau - p^s)q(\bar{p}, p^{NT}, \tau, W + \Delta CS) + e(p^s) - e(\bar{p}) \tag{1}$$

where p^s is a Gabaix representation of the perceived price with the tax.[12]

Proof. Letting W^s denote a Gabaix representation of conjectured wealth when facing tax τ, local non-satiation of preferences implies that:

$$(p^s, \bar{p}^{NT}) * \boldsymbol{q}(\bar{p}, p^{NT}, \tau, W + \Delta CS) = W^s = e(p^s)$$

In words, total perceived expenditures equal perceived wealth, which must be exactly the wealth the agent would need under perceived prices to achieve the utility from before the tax. Plugging in and using the fact that $h(p^s) = q(\bar{p}, p^{NT}, \tau, W + \Delta CS)$ yields:

$$(\bar{p} + \tau - p^s)q(\bar{p}, p^{NT}, \tau, W + \Delta CS) = [(\bar{p} + \tau, \bar{p}^{NT}) - (p^s, \bar{p}^{NT})] * \boldsymbol{q}(\bar{\boldsymbol{p}}, \tau, W + \Delta CS)$$
$$(\bar{p} + \tau - p^s)q(\bar{p}, p^{NT}, \tau, W + \Delta CS) = W + \Delta CS - e(p^s)$$

[11] Formally, $e(p) = \min\{W' | u(d(p, p^{NT}, W'), d^{NT}(p, p^{NT}, W)) \geq u(\boldsymbol{q}(\bar{p}, p^{NT}, 0, W))\}$, which is well-defined by continuity of u and connectedness of the choice set.

[12] The same choice model may have multiple Gabaix representations, particularly if the taxed good is inferior.

Rearranging and again using local non-satiation yields:

$$\Delta CS = e(p^s) - W + (\bar{p} + \tau - p^s)h(p^s) = e(p^s) - e(\bar{p}) + (\bar{p} + \tau - p^s)h(p^s)$$

□

This result has a natural interpretation. The integral represents the amount the consumer would have to be compensated if the tax-inclusive price were *actually* p^s. But instead, the agent is paying an extra $(\bar{p} + \tau - p^s)$ per unit of the taxed good consumed, and so must be compensated for that "lost" income. Furthermore, we can define $h(p)$ as the compensated demand with income $e(p)$. The expenditure function is concave, and so has well-defined derivatives almost everywhere. By Shephard's Lemma, these derivatives are Hicksian demand. By the Fundamental Theorem of Calculus:

$$e(p^s) - e(\bar{p}) = \int_{\bar{p}}^{p^s} h(p)dp$$

Thus, if $h(p^s)$ is well-defined, we can express the change in consumer surplus as:

$$\Delta CS = (\bar{p} + \tau - p^s)h(p^s) + \int_{\bar{p}}^{p^s} h(p)dp \tag{2}$$

From the representation of deadweight loss, we can immediately derive a couple of insights. First, suppose for a moment that $\tau > 0$ and $p^s \in [\bar{p}, \bar{p} + \tau]$, so that the agent does not fully notice the increase in the tax-inclusive price. We can then confirm that the non-salience of the tax weakly exacerbates the loss of consumer surplus:

$$\begin{aligned}\Delta CS &= (\bar{p} + \tau - p^s)h(p^s) + \int_{\bar{p}}^{p^s} h(p)dp \\ &\geq \int_{p^s}^{\bar{p}+\tau} h(p)dp + \int_{\bar{p}}^{p^s} h(p)dp \\ &= \int_{\bar{p}}^{\bar{p}+\tau} h(p)dp\end{aligned}$$

This reflects the fact that consumers only partially respond to the tax. If they completely understood how the tax was affecting the tax-inclusive price, they would protect themselves from this tax burden by reducing consumption of the taxed good. As they fail to fully take the tax into account, they end up worse off than if they did.

Second, much empirical work uses linear demand functions. We will also consider linear choice functions in section 3. As in the standard decision-making model, the calculation of the change in consumer surplus calculated using sticker price and tax derivatives naïvely as if the choice function q were linear in these arguments turns out to be a second order approximation to the true change in consumer surplus. We now demonstrate this claim formally.

Assume that h is continuously differentiable with respect to its own price. Assume p^s is continuously differentiable with respect to τ so that we can define $m \equiv \frac{\partial p^s}{\partial \tau}|_{(\bar{p},0)}$.[13]

[13] Formally, the claim is that there is a Gabaix representation that has $\frac{\partial p^s}{\partial \tau}$, where the derivative is taken while the consumer is being compensated. If $\frac{\partial h}{\partial p}(\bar{p}) \neq 0$, then the Inverse Function Theorem implies that $\frac{\partial p^s}{\partial \tau} = \frac{\frac{\partial q}{\partial \tau} + \frac{\partial q}{\partial W}\frac{\partial \Delta CS}{\partial \tau}}{\frac{\partial h}{\partial p}}$. If $\frac{\partial h}{\partial p} = 0$ in a neighborhood around $\bar{p}$, then $\frac{\partial p^s}{\partial \tau}|_{\tau=0} = 0$ and $\frac{\partial \Delta CS}{\partial \tau} = -\frac{\frac{\partial q}{\partial \tau}}{\frac{\partial q}{\partial W}}$.

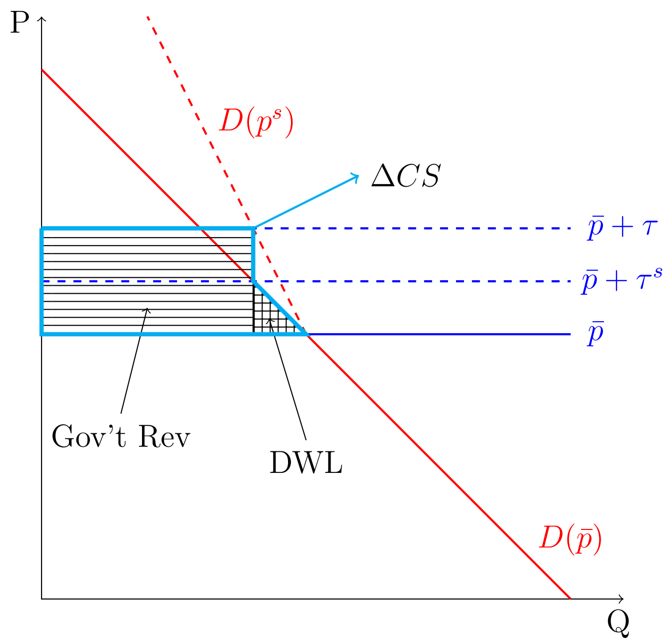


Figure 2: Welfare effects from the imposition of a non-salient tax.

Taking a second order approximation around $\tau = 0$ then yields:[14]

$$\Delta CS \approx h(\bar{p})\tau + \big[2(1-m)\frac{\partial h}{\partial p}|_{\bar{p}} + m\frac{\partial h}{\partial p}|_{\bar{p}}\big]\tau^2/2$$

Rearranging yields:

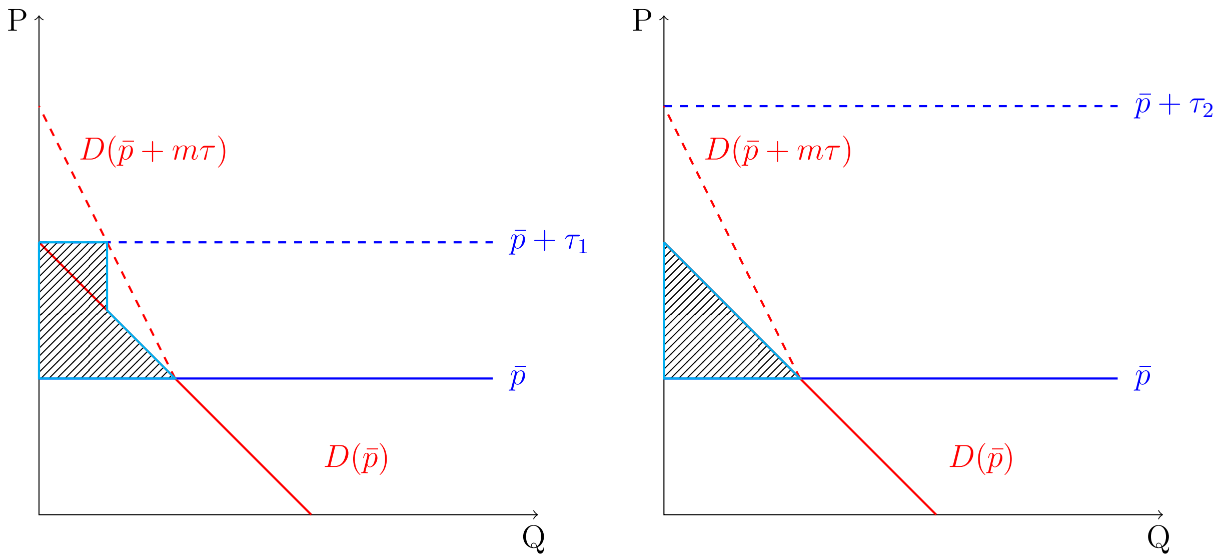
$$\Delta CS \approx \left[h(\bar{p})\tau + \left(m\frac{\partial h}{\partial p}|_{\bar{p}}\right)\tau^2/2\right] + (1-m)m\frac{\partial h}{\partial p}|_{\bar{p}}\tau^2/2$$

The first term in this calculation, $h(\bar{p})\tau + (m\frac{\partial h}{\partial p}|_{\bar{p}})\tau^2/2$, is the change in consumer surplus one would naïvely expect from using tax (compensated) responsiveness as if the agent were fully attentive to the tax. But this evaluation does not account for the discrepancy between the tax rate and the true marginal value of the good to the consumer. The second term, $(1-m)m\frac{\partial h}{\partial p}|_{\bar{p}}\tau^2/2$, reflects how the agent effectively loses income to the lump-sum portion of the tax $(\bar{p}+\tau-p^s)$, so that further increasing the tax motivates the consumer to reduce consumption and so mitigate the income lost to non-salient taxation. More formally, note that the amount of income the agent loses to the tax without noticing is:

$$(1-m)q^s\tau$$

When the tax marginally increases from zero, it causes demand to decrease marginally by $m\frac{\partial h}{\partial p}$. Integrating $(1-m)m\frac{\partial h}{\partial p}\tau$ from 0 to τ yields the term $(1-m)m\frac{\partial h}{\partial p}\tau^2/2$ (when ignoring the curvature of Hicksian demand as in the second order approximation). Thus, the second term is an approximation around $\tau = 0$ of the income no longer "misplaced" due to the tax. It is an *internality* in that it is a benefit to the consumer

[14]This derivation appears to generalize Proposition 22 from Gabaix (2014).



(a) Under τ_1 the consumer loses more than total CS (b) Under τ_2 the consumer loses exactly total CS

Figure 3: The consumer in (a) is worse off than in (b), although she is subject to a lower tax

for which the consumer does not account. This measure of the internality is generically substantial. In fact, it can overwhelm the direct harm from the tax at the margin. If the non-salient tax is already sufficiently large, then it is possible for a marginal increase in the tax to increase consumer welfare, as illustrated in Figure (3).[15]

To see this result, suppose for a moment that demand d were linear and without income effects. Still suppressing the price of the other good, the loss of consumer surplus is then:

$$\Delta CS = d(\bar{p}, W)\tau + m\frac{\partial d}{\partial p}\tau^2/2 + (1-m)m\frac{\partial d}{\partial p}\tau^2/2$$

The rate of change of the loss of consumer surplus is:

$$\frac{\partial \Delta CS}{\partial \tau} = d(\bar{p}, W) + m\frac{\partial d}{\partial p}\tau + (1-m)m\frac{\partial d}{\partial p}\tau$$

Linear demand implies that quantity is positive whenever $d(\bar{p}, W) + m\frac{\partial d}{\partial \tau}\tau > 0$. So from the Law of Demand and if $m \in (0,1)$:

$$d(\bar{p} + m\tau, W) = 0 \Rightarrow \frac{\partial \Delta CS}{\partial \tau} < 0$$

From continuity, we can conclude that the loss of consumer surplus decreases whenever the tax is sufficiently high such that consumption is sufficiently small (but positive).

[15]It may appear strange that we were previously discussing a second order approximation around $\tau = 0$, but now discuss tax changes ”sufficiently large”. The possibility of a further tax increase benefiting a consumer holds without reference to any particular functional form, but note that there is no need to speak of approximations if q is actually linear with respect to $\bar{p}$ and τ.

Given the potentially large and qualitatively important correction to welfare due to the internality from the tax, we do not recommend ignoring it.[16] Instead, we calculate deadweight loss while taking this internality into account. Deadweight loss from the tax is the difference between the income necessary to compensate the consumer from the tax and the revenue that would be raised from the tax (with the consumer compensated):[17]

$$dwl \equiv \Delta CS - \tau q(\bar{p}, p^{NT}, \tau, W + \Delta CS) = e(p^s) - e(\bar{p}) - (p^s - p)q(\bar{p}, p^{NT}, W + \Delta CS)$$

Note that one calculates deadweight as if with perceived tax $\tau^s \equiv p^s - \bar{p}$. With well-defined and sufficiently smooth compensated demand around $\bar{p}$:

$$dwl = [e(p^s) - e(\bar{p})] - (p^s - \bar{p})h(p^s) = \int_{\bar{p}}^{p^s} h(p)dp - (p^s - \bar{p})h(p^s) \qquad (3)$$

Taking a second order approximation around $p^s = \bar{p}$ yields:

$$dwl \approx -\frac{1}{2}m^2 \frac{\partial h}{\partial p}\tau^2$$

2.3 Aggregate Deadweight Loss from a Non-Salient Tax

In this section we derive aggregate deadweight loss, and discuss some of its properties. In a model of misperceived prices aggregating single agents' deadweight loss is far from trivial, as different joint distributions of perceived prices and preferences could yield the same aggregate demand but result in extremely different deadweight loss. We start with a generic environment with a non-taxed good, priced at p^{NT}, and a taxed good, with sticker price $\bar{p}$ and specific tax τ. Let $i \in \mathcal{I}$ index consumers. Each consumer is characterized by her perception of the price of the taxed good, p_i^s, type, θ_i, standing in for her preferences $\succeq_{\theta_i}$ & income W_{θ_i}, and tie-breaking parameter ζ_i. These consumer-specific parameters are distributed according to $F^*_{p^s,\theta,\zeta}$. Each agent has a choice function for the taxed good, satisfying:

$$\begin{aligned} q(\bar{p}, p^{NT}, \tau, W_{\theta_i}; \theta_i, \zeta_i) \in \mathcal{Q}_{\theta_i} \equiv \{q | \exists\, q^{NT} \in X^{NT} : p_i^s * q + p^{NT} * q^{NT} \leq W_i^s, \\ (q, q^{NT}) \succeq_{\theta_i} \boldsymbol{q'} \forall \boldsymbol{q'} \in X : (p_i^s, p^{NT}) * \boldsymbol{q'} \leq W^s\} \end{aligned}$$

where W_i^s is endogenously determined as in the Gabaix model, with corresponding expenditure function $e(p_i^s; \theta_i)$.

Again, ζ_i acts as a tie-breaker among bundles that could all have been chosen: choices do not necessarily reflect true preferences when agents misperceive prices, and agents might appear indifferent between choices that do not actually yield the same ex-post utility. This is in sharp contrast with the neo-classical model, where the actual choice that one selects among indifferent bundles has no impact on consumer surplus. But this reasoning overlooks how, in the Gabaix model, different values of conjectured income might yield different choices, even given the same preferences, perceived prices, and true income. For example, suppose agents faced the problem of spending a fixed income $W > 2$ on a binary taxed good $q \in \{0, 1\}$, with a sticker price $\bar{p} = 1$ and a

[16] One might dismiss the internality as second-order. But even in the case of fully salient taxation, all efficiency loss is second-order.

[17] We maintain the convention that deadweight loss is generically positive.

Table 1: Compensating variation and Tax Revenue for each type in each scenario. Total deadweight loss is the weighted average of compensating variation in excess of tax revenue.

	Type	$\boldsymbol{q}$ pre-tax	$\boldsymbol{q}$ post-tax	W^s post-tax	ΔCS	Tax Revenue	Total DWL
Scenario 1	Type 1	$(1, W-1)$	$(1, W-2)$	$W-1$	1	1	0.5
	Type 2	$(1, W-1)$	$(0, W)$	W	1	0	
Scenario 2	Type 1	$(1, W-1)$	$(0, W)$	W	0	0	0
	Type 2	$(1, W-1)$	$(1, W-2)$	W	1	1	

tax of $\tau = 1$, and a non-taxed good $q^{NT} \in \mathbb{R}_+$, priced at $p^{NT} = 1$. There are two types; half of agents maximize utility $u_1(\boldsymbol{q}) = q + q^{NT}$, and perceive $p_1^s = 1$, so they fail to see the tax and are indifferent as to buying q before and after its imposition; the other half of agents maximize $u_2(\boldsymbol{q}) = 2q + q^{NT}$, and $p_2^s = 2$, so agents of type 2 perfectly notice the tax, and are indifferent as to buying the good after the tax but would always buy q in its absence. Let us suppose everyone buys q before the tax is imposed.[18] Consider two possible scenarios occurring after the imposition of the tax. In the first scenario all agents of type 1 buy q and no agents of type 2 buy q. In the second scenario, no agents of type 1 buy q and all agents of type 2 buy q. We work out the details in the appendix, but we summarize results for aggregate deadweight loss in table (1), so that one can check that deadweight loss will differ depending on the final choices of agents who are "indifferent". This is because when agents misperceive prices, they can appear indifferent between choices that they actually value differently: in the example, agents of type 1 are actually happier when they don't buy the good (as if they perceived the tax), but in the model their final choice depends on whether they conjecture an income $W_1^s = W$, and pick $\boldsymbol{q} = (0, W)$, or $W_1^s = W - 1$, and pick $\boldsymbol{q} = (1, W-2)$. Which choice they end up making is governed by the parameter ζ_i.

Note that, although the Gabaix model does differ from the neoclassical consumption model in a number of ways, the introduction of this notation does not challenge our intuitive understanding of compensated demand. Consider for instance two values l and h such that for any p_i^s and θ_i:

$$[q(p_i^s;\theta_i,l), q(p_i^s;\theta_i,h)] \supseteq \{q | \exists\, q^{NT} \in X^{NT} : p_i^s * q + p^{NT} * q^{NT} \leq W^s, (q, q^{NT}) \succeq_{\theta_i} \boldsymbol{q}' \\ \forall \boldsymbol{q}' \in X : (p_i^s, p^{NT}) * \boldsymbol{q}' \leq W^s\}$$

So $q(p_i^s;\theta_i,l)$ and $q(p_i^s;\theta_i,h)$ are the smallest and largest amounts, respectively, of the taxed good that the agent could choose to consume. With that in mind, lemma 1 shows that the Law of Compensated Demand holds in this framework.

Lemma 1. *For any agent i with type θ_i and any two prices p and p':*

$$p < p' \Rightarrow q(p';\theta_i,h) \leq q(p;\theta_i,l)$$

We prove this result in the appendix.

For the rest of this section, we assume well-defined and sufficiently smooth compensated demand, so that $\mathcal{Q}_{\theta_i}$ is always single-valued. Suppressing ζ, deadweight loss

[18] Note that the punchline of this example does not depend on this, as either way consumers of type one obtain utility W before the tax is imposed

with complete pass-through takes the form:

$$DWL = \int_{p_i^s,\theta_i} [e(p_i^s;\theta_i) - e(\bar{p};\theta_i)] - (p_i^s - \bar{p})q(p_i^s;\theta_i)dF^*_{p^s,\theta}(p_i^s,\theta_i)$$

In line with most existing analyses of deadweight loss, we now consider a second-order approximation, which allows us to characterize our object of interest in terms of first derivatives.

$$DWL \approx -\frac{1}{2}\int_{p_i^s,\theta_i} m_i^2 \frac{\partial h(\bar{p};\theta_i)}{\partial p} dF^*_{p^s,\theta}(p_i^s,\theta_i)\tau^2$$

Note that neither aggregate price responsiveness $\int_{p_i^s,\theta_i} \frac{\partial h(\bar{p};\theta_i)}{\partial p} dF^*_{p^s,\theta}(p_i^s,\theta_i)$ nor aggregate tax responsiveness $\int_{p_i^s,\theta_i} m_i \frac{\partial h(\bar{p};\theta_i)}{\partial p} dF^*_{p^s,\theta}(p_i^s,\theta_i)$ are sufficient statistics for deadweight loss as a function of the tax rate. This illustrates the challenge of calculating deadweight loss from observable data.

Finally, we wish to consider the impact on aggregate deadweight loss of allocative inefficiency. So far we have usually thought about a perfectly elastic supply, so that the post-tax sticker price remains unchanged at $\bar{p}$. In the case of an arbitrary differentiable aggregate supply function Q^{supply}, deadweight loss has second order approximation:[19]

$$DWL \approx -\frac{1}{2}\left[\int_{m_i,\theta_i} m_i^2 \frac{\partial h_i}{\partial p}|_{(\bar{p};\theta_i)} - \frac{(\sum_i m_i \frac{\partial h_i}{\partial p}|_{(\bar{p},\theta_i)})^2}{\sum_i \frac{\partial h_i}{\partial p} - \frac{\partial Q^{supply}}{\partial p}} dF_{m,\theta}(m_i,\theta_i)\right]\tau^2$$

Consider for simplicity the case in which there are no income effects. Then:

$$\begin{aligned} DWL &\approx -\frac{1}{2}\int_{m_i,\theta_i}\left[m_i^2\frac{\partial d_i}{\partial p} - \frac{(\sum_i m_i \frac{\partial d_i}{\partial p})^2}{\sum_i \frac{\partial d_i}{\partial p} - \frac{\partial Q^{supply}}{\partial p}} dF_{m,\theta}(m_i,\theta_i)\right]\tau^2 \\ &= -\frac{1}{2}\int_{m_i,\theta_i}\left[m_i^2\frac{\partial q_i}{\partial p}dF_{m,\theta}(m_i,\theta_i) - \frac{(\int_{m_i,\theta_i} m_i \frac{\partial q_i}{\partial p})^2}{\int_{m_i,\theta_i}[\frac{\partial q_i}{\partial p} - \frac{\partial Q^{supply}}{\partial p}]dF_{m,\theta}(m_i,\theta_i)}\right]\tau^2 \end{aligned}$$

The additional term due to incomplete pass-through of the tax is entirely determined by aggregate tax responsiveness, aggregate price responsiveness, and aggregate price responsiveness of supply. So the problem of calculating deadweight loss in a general setting reduces to the problem of calculating deadweight loss with complete pass-through. This motivates our principal concern with determining deadweight loss from observable data while assuming complete pass-through.

However, consider the case of perfectly inelastic supply, i.e. $\frac{\partial Q^{supply}}{\partial p} = 0$. Then:

$$DWL \approx -\frac{1}{2}\int_{m_i,\theta_i}\left[m_i^2\frac{\partial q_i}{\partial p}dF_{m,\theta}(m_i,\theta_i) - \frac{(\int_{m_i,\theta_i} m_i \frac{\partial q_i}{\partial p})^2}{\int_{m_i,\theta_i}\frac{\partial q_i}{\partial p}dF_{m,\theta}(m_i,\theta_i)}\right]\tau^2 \geq 0$$

Note that deadweight loss is zero when $m_i = m \ \forall i$.[20] The imposition of the tax forces the sticker price to decrease so that aggregate consumption remains constant. If there

[19]We formally demonstrate this result with a finite population in the appendix. To make this result general, one need only assume that the price derivative of compensated demand is uniformly bounded on the support of (p^s,θ).

[20]Technically, the previous expression only shows that the second order approximation for deadweight loss is zero in the case of homogeneous attention with perfectly inelastic supply. We demonstrate in the appendix that deadweight loss is precisely zero in this case.

is heterogeneity in attention, then different consumers perceive different prices even while facing the same tax. Some agents perceive a higher price due to the tax, while others perceive a lower price as a result. Those who perceive a higher price reduce their consumption, with that foregone consumption going to those perceiving a lower price. This creates allocative inefficiency, as units of the good are not necessarily going in the hands of the people who value them the most.[21]

Now that we have taken a more extensive look at how deadweight loss works in aggregate when taxes are misperceived, we next tackle the problem of identifying it with data. In particular, section 3 deals with identification in the case where the econometrician can only observe aggregate demand across a number of markets; we will find that while point-identification is impossible, one can still bound deadweight loss. In section 4 we instead deal with the case where the econometrician can observe demand across a number of individuals, and discuss under which assumptions one might be able to point-identify deadweight loss.

3 Non-Identification with Aggregate Continuous Choice Data

This section discusses to what degree one can infer deadweight loss from aggregate choice data. Most of the results from this section apply generally. For simplicity, we will assume that the econometrician has already determined the distribution of preference types. In this case, the econometrician can provide a lower bound for deadweight loss by assuming that all agents perceive the same tax-inclusive price, i.e. assume there is no attention heterogeneity. Alternatively, one can derive an upper bound for deadweight loss by supposing maximal attention heterogeneity. Since the data do not reveal the variance in tax salience, one cannot point identify deadweight loss from aggregate data.[22] In the case when the consumption function is linear in the sticker price and tax rate, deadweight loss can take on any value between the upper and lower bounds. The results in this section are described as if all agents face the same sales tax, but none of our results depend on that assumption.

Since we are considering the problem of identification with aggregate demand, we assume there are no income effects. This is because even the standard model requires strong restrictions on income effects in order to achieve identification with aggregate data. Suppressing income and price for the non-taxed good, we denote the consumption function for agent i with type θ_i and perceived tax-inclusive price p_i^s for the taxed good by $q(p_i^s; \theta_i, \zeta_i)$.[23] To ensure integrability, we assume the econometrician knows that $F_{p^s}^*$ has support bounded above zero, and so only considers marginal distributions of subjective prices bounded above zero. The econometrician observes aggregate demand:

$$\int_{p_i^s, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}^*(p_i^s, \theta_i, \zeta_i)$$

[21] While we believe this particular result to be novel as it pertains to our model, we credit Taubinsky and Rees-Jones (2017) for the insight that taxation still generically yields allocative inefficiency even when supply is perfectly inelastic.

[22] This claim holds generically, with an exception if the tax does not alter aggregate consumption.

[23] However, all of these results follow if one reinterprets $q(p_i^s; \theta_i, \zeta_i)$ as the compensated choice of agent i.

Deadweight loss for an individual i is a function of their expenditure function $e(p)$ and prices via:

$$dwl(p_i^s;\theta_i,\zeta_i) = e(p_i^s;\theta_i) - e(\bar{p};\theta_i) - (p_i^s - \bar{p})q(p_i^s;\theta_i,\zeta_i)$$

We are interested in aggregate deadweight loss:

$$DWL \equiv \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

The problem of identification is to find conditions for which any joint distribution $F_{p^s,\theta,\zeta}$ of (p^s,θ,ζ) (as a function of observable variables) satisfying these conditions and such that (for any observed values of observable variables):

$$\int_{p_i^s,\theta_i\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta_i}(p_i^s,\theta_i,\zeta_i) = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

yields the same value for deadweight loss:

$$\int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) = DWL$$

The main message of this section will be the failure of such a result to obtain under plausible restrictions arising from aggregate data alone.

The upper and lower bounds use the following lemma:

Lemma 2. *For any agent i with type θ_i and any two pairs (p,ζ_i) & (p',ζ_i'):*

$$dwl(p';\theta_i,\zeta_i) \geq dwl(p;\theta_i,\zeta_i') - (p-\bar{p})(q(p';\theta_i,\zeta_i') - q(p;\theta_i,\zeta_i))$$

This lemma comes entirely from the compensated law of demand (CLD), and can be confirmed with a simple graph, as we show in figure 4. This lemma indicates the convexity of deadweight loss with respect to the perceived price.

Proof. Note from the definition of the expenditure function and optimal compensated consumption vectors $\boldsymbol{q}$ & $\boldsymbol{q'}$ for price vectors (p,p^{NT}) and (p',p^{NT}) respectively:

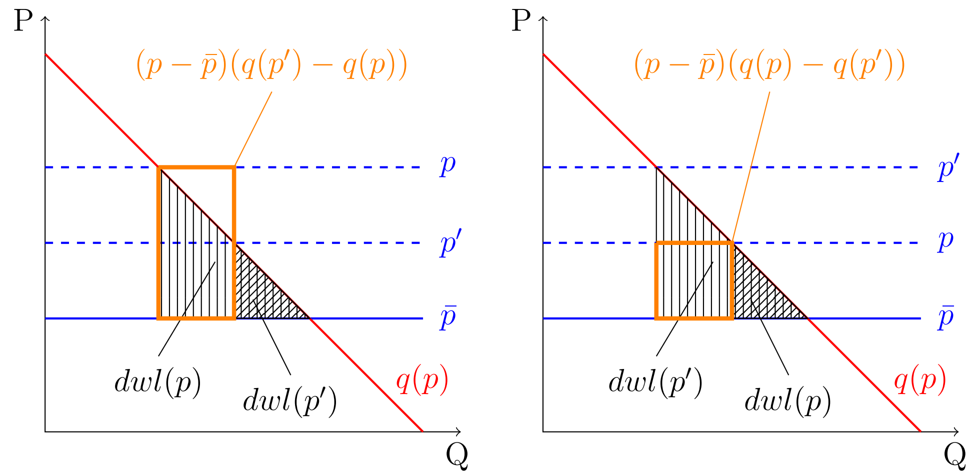
$$e(p')-e(p) = (p',p^{NT})*\boldsymbol{q'}-(p,p^{NT})*\boldsymbol{q} \geq (p',p^{NT})*\boldsymbol{q'}-(p,p^{NT})*\boldsymbol{q'} = (p'-p)q(p';\theta_i,\zeta_i')$$

Plugging in yields:

$$\begin{aligned}
dwl(p';\theta_i) &= [e(p') - e(\bar{p})] - (p'-\bar{p})q(p';\theta_i,\zeta_i') \\
&= [e(p') - e(p)] + [e(p) - e(\bar{p})] - [(p'-p) + (p-\bar{p})]q(p';\theta_i,\zeta_i') \\
&\geq (p'-p)q(p';\theta_i,\zeta_i') + e(p) - e(\bar{p}) - [(p'-p) + (p-\bar{p})]q(p';\theta_i,\zeta_i') \\
&= e(p) - e(\bar{p}) - (p-\bar{p})q(p';\theta_i,\zeta_i') \\
&= dwl(p) + (p-\bar{p})q(p;\theta_i,\zeta_i) - (p-\bar{p})q(p';\theta_i,\zeta_i') \\
&= dwl(p;\theta_i) - (p-\bar{p})(q(p';\theta_i,\zeta_i') - q(p;\theta_i,\zeta_i))
\end{aligned}$$

□

Finally, we impose regularity conditions to rule out ill-defined integrals. Formally, we insist that the econometrician only consider distributions that satisfy the *integrability conditions*, described below.



(a) $p > p'$

(b) $p' > p$

Figure 4: A graphical illustration of Lemma 2. As long as demand is weakly decreasing, $dwl(p')$ cannot be smaller than $dwl(p)$ minus (plus) the orange rectangle.

Definition 1. *A distribution $F_{p^s,\theta,\zeta}$ satisfies the integrability conditions if:*

1. *q and dwl are integrable on any measurable set*
2. *$q(p;\theta_i,z)$ is integrable on any subset of the support of θ for any $p > 0$ and any z in the range of ζ*

For instance, all distributions with a finite support of (p^s,θ,ζ) satisfy the above conditions.

3.1 Lower Bound on Deadweight Loss

Even though we cannot point identify deadweight loss with aggregate data, we can identify upper and lower bounds. For the lower bound, consider arbitrary $\bar{p}$, p^{NT}, and τ. For arbitrary $F_{p^s,\theta,\zeta}$ that could generate the data, we can choose a price $\hat{p}^s$ that could also rationalize the data if perceived by everyone.

Proposition 3. *For any $F_{p^s,\theta,\zeta}$ that yields integrable aggregate demand, $\exists \hat{p}^s$ such that for some distribution $F'_{\theta,\zeta}$ such that $F'_\theta = F_\theta$:*

$$\int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta}(p_i^s,\theta_i,\zeta_i) = \int_{\theta_i,\zeta_i} q(\hat{p}^s;\theta_i,\zeta_i)dF'_{\theta,\zeta}(\theta_i,\zeta_i)$$

We can always rationalize the data with a joint distribution of (p^s,θ) in which θ has marginal distribution F^*_θ, whereas $p^s = \hat{p}^s$ with probability one. We now show that such a joint distribution provides a generic underestimate to the possible values of deadweight loss.

Theorem 1. *Consider any joint distributions $F_{p^s,\theta,\zeta}$ and $F_{\theta,\zeta}$ with corresponding value $\hat{p}^s$ such that:*

$$\int_{\theta,\zeta} q(\hat{p}^s;\theta_i,\zeta_i)dF_{\theta,\zeta}(\theta_i,\zeta_i) = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \tag{4}$$

Then the following inequality obtains:

$$\int_{\theta_i,\zeta_i} dwl(\hat{p}^s;\theta_i,\zeta_i)dF_{\theta,\zeta}(\theta_i,\zeta_i) \leq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

The intuition behind this result is that introducing heterogeneity in perceived prices facilitates gains from trade by having agents trade with each other after making their consumption decisions. If one person perceives a higher price than another, then the two agents will have different marginal valuations of the good. If they could exchange with each other, the one who perceived the higher price could purchase some of the good from the other agent, making both agents better off. Thus, ruling out perceived price heterogeneity eliminates the possibility of an allocative inefficiency of the amount of the taxed good consumed in aggregate.

Proof. From lemma 2:

$$\int_{p_i^s,\theta_i,\zeta_i} [dwl(p_i^s;\theta_i,\zeta_i)+(\hat{p}^s-\bar{p})q(p_i^s;\theta_i,\zeta_i)]dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$
$$\geq \int_{p_i^s,\theta_i,\zeta_i} [dwl(\hat{p}^s;\theta_i,\zeta_i)+(\hat{p}^s-\bar{p})q(\hat{p}^s;\theta_i,\zeta_i)]dF'_{\theta,\zeta}(\theta_i,\zeta_i)$$

But note from the rationalizability of the data that:

$$(\hat{p}^s-\bar{p})\int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) = (\hat{p}^s-\bar{p})\int_{\theta_i,\zeta_i} q(\hat{p}^s;\theta_i,\zeta_i)dF'_{\theta,\zeta}(\theta_i,\zeta_i)$$

Thus, we can conclude that:

$$\int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \geq \int_{\theta_i,\zeta_i} dwl(\hat{p}^s;\theta_i,\zeta_i)dF_{\theta,\zeta}(\theta_i,\zeta_i)$$

□

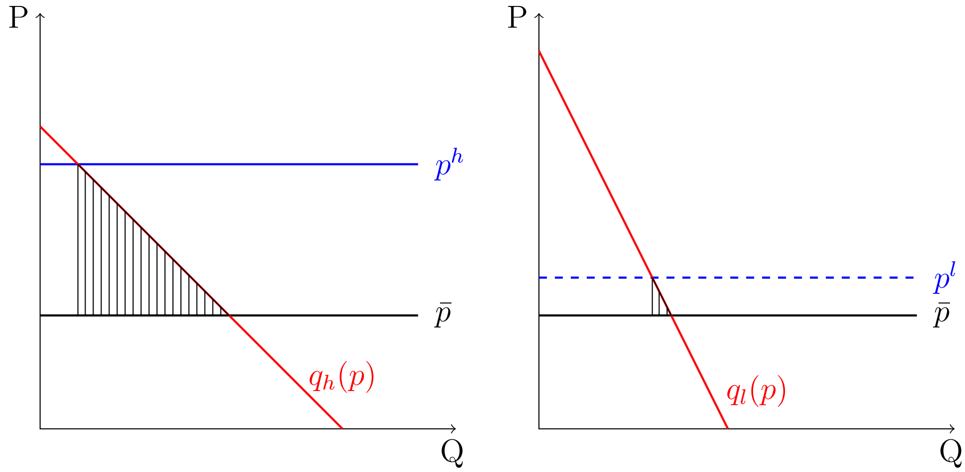
Theorem 1 points out that for any distribution that rationalizes the data, one can alternatively rationalize the data with a homogeneous perceived price that yields (weakly) less deadweight loss. From this, we can make two conclusions. One, we generally cannot identify deadweight loss because we could always alternatively rationalize the data with a homogeneous perceived price.[24] This holds even if we already knew the distribution of preference types F^*_θ. Two, if there is a minimum value of deadweight loss that is consistent with the data, that value of deadweight loss comes from a distribution with no heterogeneity in tax salience.

3.2 Upper bound on deadweight loss

The upper bound comes from an assumption on the limits to tax salience:

Assumption 1. *There is some known value $\bar{m} \geq 0$ such that $F^*_{p^s}$ yields p^s with support contained entirely in $\mathcal{P} \equiv [\bar{p}, \bar{p}+\bar{m}\tau]$*

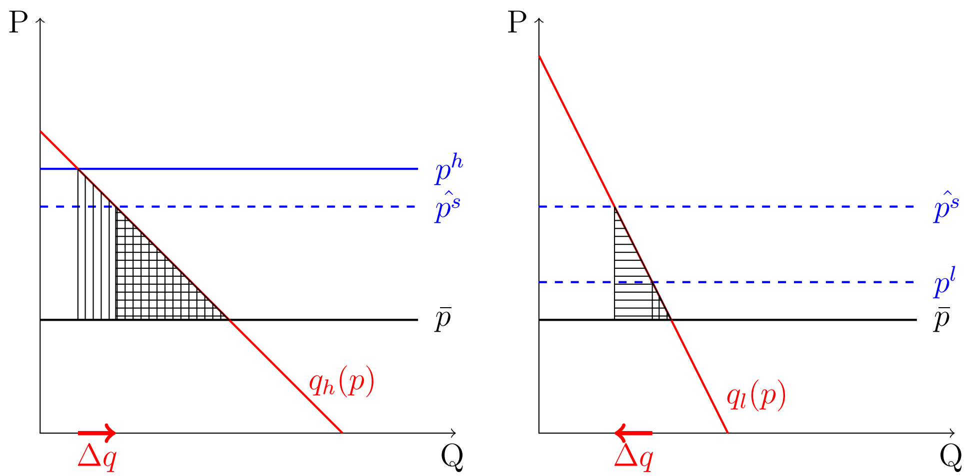
[24] This claim holds generically, but would not hold, for instance, if there was no heterogeneity in tax salience.



(a) High perceived price

(b) Low perceived price

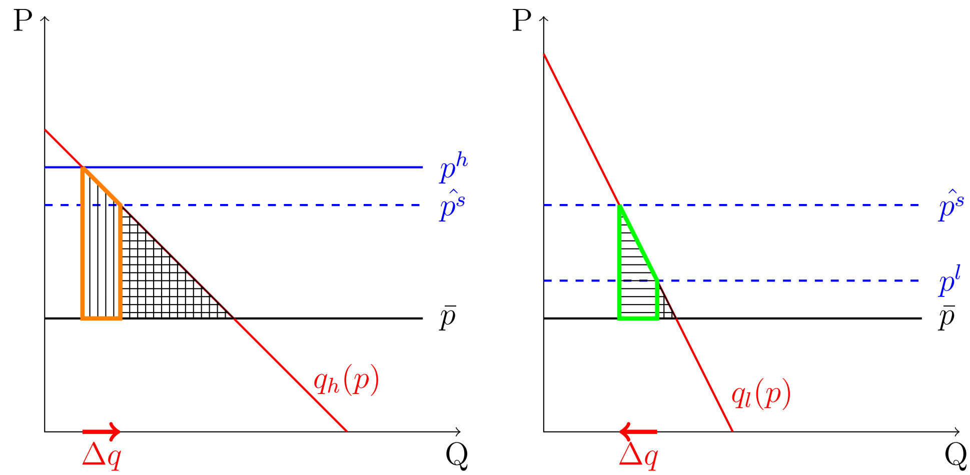
Figure 5: Pre-slide for figure 6



(a) High perceived price

(b) Low perceived price

Figure 6: Pre-slide for figure 7



(a) High perceived price

(b) Low perceived price

Figure 7: A graphical illustration of Theorem 1. When one picks $\hat{p}^s$ as to make the change in demand equal for the consumer in (a) and in (b), the decrease in *dwl* for the consumer in (a) (orange) must be at least as large as the increase in *dwl* for the consumer in (b) (green).

This assumption says that agents must perceive a non-negative tax τ^s no greater than fraction $\bar{m}$ of the true tax.[25] For instance, setting $\bar{m} = 1$ would be to assume that agents never over-react to a tax rate. Imposing that $\tau^s \geq 0$ with probability one already ensures that deadweight loss is no greater than the original consumer surplus.[26] But the interval restriction implies any distribution yields no more deadweight loss than a distribution with "binary" perceived prices, i.e. where p^s can only take on values in $\{\bar{p}, \bar{p} + \bar{m}\tau\} \equiv \partial\mathcal{P}$.

Before we prove this result in theorem 2, we show that one can always construct a binary distribution that rationalizes the data. For that, consider any $F_{p^s,\theta,\zeta}$ that rationalizes the data. If $F_{p^s,\theta,\zeta}$ puts no mass on $(\bar{p}, \bar{p} + \bar{m}\tau) \equiv int(\mathcal{P})$, then the claim holds trivially. So let us consider only distributions such that

$$lim_{m \to \bar{m}^-} F_{p^s}(\bar{p} + m\tau) - F_{p^s}(\bar{p}) > 0 \tag{5}$$

pick $\tilde{p}^s \in \partial\mathcal{P}$ and a corresponding $p^b(p_i^s) \equiv \bar{p} + \mathbb{I}(p_i^s > \tilde{p}^s)\bar{m}\tau$ such that:

$$\begin{aligned} \int_{p_i^s \in int\mathcal{P}, \theta_i} q(p^b(p_i^s); \theta_i, l) dF_{p^s,\theta}(p_i^s, \theta_i) &\leq \int_{p_i^s \in int\mathcal{P}, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i) \\ &\leq \int_{p_i^s \in int\mathcal{P}, \theta_i} q(p^b(p_i^s); \theta_i, h) dF_{p^s,\theta}(p_i^s, \theta_i) \end{aligned} \tag{6}$$

In words, for any distribution that puts mass on $int(\mathcal{P})$, we pick a $\tilde{p}^s$ that acts as a divide: people below it get assigned to a group that does not perceive the tax at all,

[25] This description implicitly assumes that $\tau > 0$.

[26] Recall that deadweight loss equals its calculation as if taxes actually satisfied $\tau_i = p_i^s - \bar{p}$, so excess burden cannot exceed the original consumer surplus for any agent. One can show that if τ^s has support on negative values, then it's possible to have total deadweight loss substantially greater than the original total consumer surplus.

while people above it get assigned to a group that perceives it "maximally". Since demand is monotonic in p, and given our definitions of l and h, one can always pick $\tilde{p}^s$ such that the above inequalities hold weakly. Thus, it is always possible to find $\lambda \in [0,1]$ such that:

$$\lambda \int_{p_i^s \in int\mathcal{P}, \theta_i} q(p^b(p_i^s); \theta_i, h) dF_{p^s,\theta}(p_i^s, \theta_i) + (1-\lambda) \int_{p_i^s \in int\mathcal{P}, \theta_i} q(p^b(p_i^s); \theta_i, l) dF_{p^s,\theta}(p_i^s, \theta_i)$$
$$= \int_{p_i^s \in int\mathcal{P}, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i)$$

Now, define the alternative distribution $F''_{p^s,\theta,\zeta}$ so that $F''_{p^b(p^s),\theta} = F_{p^s,\theta}$, $F''_{p^s,\theta,\zeta|p^s \in \partial\mathcal{P}} = F_{p^s,\theta,\zeta|p^s \in \partial\mathcal{P}}$, and conditional on $p^s \in int(\mathcal{P})$, $\zeta = h$ with probability λ, $\zeta = l$ with probability $1-\lambda$, and $(p_i^s, \theta) \perp \zeta$. In words, we propose no change to the distribution of preference types, and propose no change at all when perceived prices are at the extremes. When perceived prices are interior, the tie-breaking parameter is then independent of perceived price and preference type. Then $F''_{p^s,\theta,\zeta}$ rationalizes the data when agents perceive subjective prices $p^b(p_i^s)$:

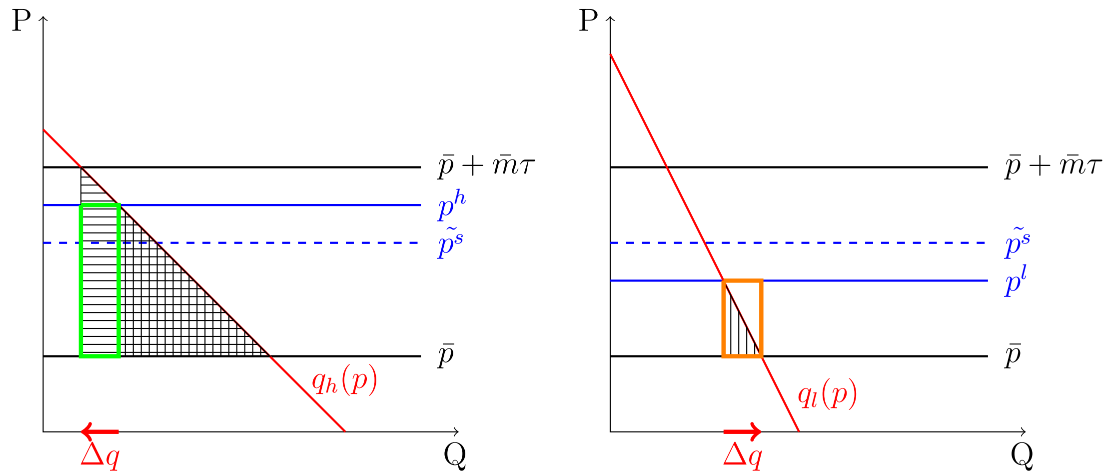
$$\begin{aligned}
\int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) &= \int_{p_i^s \in int(\mathcal{P}),\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\
&+ \int_{p_i^s \in \partial\mathcal{P},\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\
&= \int_{p_i^s \in int(\mathcal{P}),\theta_i} [\lambda q(p^b(p_i^s);\theta_i,h) + (1-\lambda) q(p^b(p_i^s),\theta_i,l)] dF_{p^s,\theta}(p_i^s,\theta_i) \\
&+ \int_{p_i^s \in \partial\mathcal{P},\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\
&= \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)
\end{aligned}$$

Furthermore, such a distribution provides a generically larger value of deadweight loss than does $F_{p^s,\theta,\zeta}$.

Theorem 2. *Under assumption 1, for any $F_{p^s,\theta,\zeta}$ and any corresponding $F''_{p^s,\theta,\zeta}$ as described above:*

$$\int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \geq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

We can obtain intuition in two ways. One is to note that the method of forcing binary perceived prices increases heterogeneity of perceived prices compared to $F_{p^s,\theta,\zeta}$. Another is by considering the case where $\bar{m} = 1$ and F^*_θ is known to be degenerate, so that all agents have the same preferences. For a given aggregate demand, deadweight loss is maximized under these preferences when some perceive price $p_i^s = \bar{p}$, while others correctly perceived the true tax rate $p_i^s = \bar{p} + \tau$. This is because for each individual agent, deadweight loss is convex in the perceived price. Hence, for a given aggregate demand, *aggregate* deadweight loss will be highest when it is as high as possible for some – namely, those who fully perceive the tax – while it is null for everybody else – as those who don't perceive the tax at all are effectively subject to a lump-sum tax.



(a) High perceived price

(b) Low perceived price

Figure 8: A graphical illustration of Theorem 2. The watershed price $\tilde{p}^s$ is picked to make the change in demand equal for the consumer in (a) and in (b). As long as we are dealing with weakly decreasing demand functions, the increase in deadweight loss for (a) is at least as big as the green box, while the decrease in deadweight loss for (b) is at most as big as the orange box. By assigning a perceived price of $\bar{p} + \bar{m}\tau$ to the consumer in (a) and $\bar{p}$ to the consumer in (b), we have increased aggregate deadweight loss holding aggregate demand constant.

The proof follows closely the proof of theorem 1. Since we are using two separate perceived prices, $\bar{p}$ and $\bar{p}+\bar{m}\tau$, instead of just one, $\hat{p}^s$, the algebra is a bit more involved, so we relegate the proof to the appendix. Nonetheless, the idea is the proof uses lemma 2 to obtain the inequality:

$$\int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \geq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$
$$- \int_{p_i^s,\theta_i,\zeta_i} (p_i^s - \tilde{p}^s)q(p_i^s;\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) + \int_{p_i^s,\theta_i,\zeta_i} (p_i^s - \tilde{p}^s)q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

Then, we use the CLD to show that the combination of terms in the second line is non-negative; because $p^b(p_i^s)$ can have two separate outputs, this is a bit more involved than in the proof of theorem 1.

Theorem 2 illustrates that for any distribution of (p^s,θ) that rationalizes the data, we can alternatively rationalize the data with a distribution with support for p_i^s on $\{\bar{p},\bar{p}+\bar{m}\tau\}$ that yields (weakly) greater deadweight loss. Again, we see that identification of deadweight loss is not generally possible even if we knew the distribution of F_θ^*, as different marginal distributions of p^s and ζ could have different implications for deadweight loss. Also, any upper bound to the possible values of deadweight loss must be generated from a distribution with support of perceived prices on $\{\bar{p},\bar{p}+\bar{m}\tau\}$.

However, not all distributions that have $p^s \in \partial\mathcal{P}$ with probability one yield the same value of deadweight loss, even when rationalizing the same data with the same distribution of preference types. We demonstrate this point in our example in table

1. Intuitively, in a model of price misperception, people can be seemingly indifferent between several choices even when they ex-post would prefer some choices over others. This is because people might conjecture themselves different incomes, leading them to believe they can't afford their most preferred bundle. This implies that when several agents can make different choices based on their tie-breaking type ζ alone, we can increase deadweight loss by transferring some consumption from people who value it more to people who value it less, holding aggregate demand constant. Given aggregate demand and knowledge of the distribution of θ, one can always find the distribution of p^s and ζ that maximizes deadweight loss.

Theorem 3. *There exists values $\Delta \in [0, \bar{m}\tau]$ and $\gamma \in [0, 1]$ such that:*

$$\int_{p_i^s,\theta_i} \tilde{q}_{\Delta,\gamma}(\theta_i) dF^*_{p^s,\theta_i} = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

where:[27]

$$\begin{aligned}\tilde{q}_{\Delta,\gamma}(\theta_i) = &[\mathbb{I}(\frac{dwl(\bar{p}+\bar{m}\tau;\theta_i,l)}{q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)} > \Delta) + \gamma\mathbb{I}(\frac{dwl(\bar{p}+\bar{m}\tau;\theta_i,l)}{q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)} = \Delta)]q(\bar{p}+\bar{m}\tau;\theta_i,l) \\ &+ [\mathbb{I}(\frac{dwl(\bar{p}+\bar{m}\tau;\theta_i,l)}{q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)} < \Delta) + (1-\gamma)\mathbb{I}(\frac{dwl(\bar{p}+\bar{m}\tau;\theta_i,l)}{q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)} = \Delta)]q(\bar{p};\theta_i,h)\end{aligned}$$

Furthermore, under assumption 1, for any $F_{p^s,\theta,\zeta}$ that rationalizes the data such that $F_\theta = F^*_\theta$:[28]

$$\int_{p_i^s,\theta_i} \frac{\tilde{q}_{\Delta,\gamma}(\theta_i) - q(\bar{p};\theta_i,h)}{q(\bar{p}+\bar{m}\tau;\theta_i,l) - q(\bar{p};\theta_i,h)} dwl(\bar{p}+\bar{m}\tau;\theta_i,l) dF^*_{p^s,\theta}(p_i^s,\theta_i) \geq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

In words, theorem 3 says that the maximal value of deadweight loss consistent with the data and knowledge of F^*_θ is given by having all agents perceive the highest possible price when doing so yields more than (and fraction γ of those for whom it's equal to) a certain value of deadweight loss per quantity reduced from the perceived price increase due to the tax. We have these same agents choose the lowest quantity consistent with their preferences and perceived budget, whereas we have the other agents choose the largest quantity consistent with their preferences and perceived budget.

We relegate this proof to the appendix, but the intuition is straightforward. The econometrician observes the reduction in aggregate demand due to the tax. In searching for the explanation of that reduction in demand that maximizes deadweight loss, one should assign the reduction in quantity demanded to those for whom that allocation yields the greatest deadweight loss. Following this procedure, there is a cutoff value Δ which describes the amount of deadweight loss obtained relative to the reduction in quantity demanded sufficient to warrant the assignment of subjective tax-inclusive price $p_i^s = \bar{p} + \bar{m}\tau$ to that agent.

3.3 Linear Special Case

We conclude this section with a discussion of the special case in which q is known to be linear in $\bar{p}$ and τ (for fixed p^{NT}). We focus on this example both because of

[27] Of course, if $q(\bar{p};\theta_i,h) = q(\bar{p}+\bar{m}\tau;\theta_i,l)$, then $\tilde{q}_{\Delta,\gamma}(\theta_i) = q(\bar{p};\theta_i,h)$.

[28] The integrand on the left-hand side is zero for any θ_i such that $q(\bar{p}+\bar{m}\tau;\theta_i,l) = q(\bar{p};\theta_i,h)$.

how frequently economists estimate linear models and because of its relationship to the second order approximation of deadweight loss. As demonstrated in section 2.3, one can express a second order approximation to deadweight loss as a function of derivatives. In the case where the choice function is linear in regressors $\bar{p}$ & τ, the second order approximation is an exact calculation of deadweight loss, and our results from the previous subsections apply.

Formally, each preference type θ_i takes the form $\theta_i = (\beta_i, \epsilon_i) \in \mathbb{R}^2$.[29] To maintain linearity in regressors, we also assume that tax salience m is constant with respect to τ. The choice function q then takes the form:

$$q_i = \alpha + \beta_i p_i^s + \epsilon_i = \alpha + \beta_i[\bar{p} + m_i\tau_i] + \epsilon_i$$

We are suppressing tie-breaking parameter ζ because $\theta_i, \bar{p}_i, m_i$, and τ_i always uniquely determine consumption. We have the parameter α so that we can assume without loss of generality that $\mathbb{E}[\epsilon] = 0$. Defining $\tilde{\beta}_i \equiv m_i\beta_i$ yields:

$$q_i = \alpha + \beta_i\bar{p} + \tilde{\beta}_i\tau_i + \epsilon_i$$

with corresponding deadweight loss per agent from equation 3:

$$dwl_i = \int_{\bar{p}}^{p^s} [\alpha + \beta_i p + \epsilon_i]dp - (p^s - \bar{p})[\alpha + \beta_i p^s + \epsilon_i] = \int_{\bar{p}}^{p^s} (p - p^s)\beta_i dp$$

$$dwl_i = \frac{1}{2}[\frac{p^{s2} - \bar{p}^2}{2} - (p^s - \bar{p})p^s]\beta_i = \frac{1}{2}(p^s - \bar{p})[(p^s + \bar{p}) - 2p^s]\beta_i = -\frac{1}{2}\tau^{s2}\beta_i$$

$$dwl_i = -\frac{1}{2}m_i^2\beta_i\tau^2$$

As always, we assume that the joint distribution of parameters remains unaffected by the specific values of $\bar{p}$ and τ. The econometrician observes for various values of regressors:

$$\mathbb{E}[q|\bar{p},\tau] \equiv \int\limits_{\beta_i,\tilde{\beta}_i,\epsilon_i} [\alpha + \beta_i\bar{p} + \tilde{\beta}_i\tau + \epsilon_i]dF^*_{\beta,\tilde{\beta},\epsilon}(\beta_i,\tilde{\beta}_i,\epsilon_i) = \alpha + \mathbb{E}[\beta]\bar{p} + \mathbb{E}[\tilde{\beta}]\tau \qquad (7)$$

where $F^*_{\beta,\tilde{\beta},\epsilon}$ is the true distribution of $(\beta, m\beta, \epsilon)$. The challenge is to use the observed values of triplets $(\bar{p}, \tau,\ \mathbb{E}[q|\bar{p},\tau])$ to infer aggregate deadweight loss, which is equivalent to its second order approximation around $\tau = 0$:

$$DWL = -\frac{1}{2}\int_{\beta_i,m_i} m_i^2\beta_i dF_{\beta,m}(\beta_i,m_i)\tau^2 = -\frac{1}{2}\mathbb{E}[m^2\beta]\tau^2 = -\frac{1}{2}\mathbb{E}[m\tilde{\beta}]\tau^2$$

The only restriction that the econometrician imposes on the distribution of tax salience m is that the support of tax salience is contained within the interval $[0, \bar{m}]$.[30] The econometrician can also use the CLD as in lemma 1, so that $\mathbb{P}[\beta \leq 0] = 1$. In fact, we

[29]Agents have quasi-linear utility $u_i = \frac{q_i^2/2-(\alpha+\epsilon_i)q_i}{\beth_i} + q_i^{NT}$. For a given p^{NT}, we define $\beta_i \equiv \frac{\beth_i}{p^{NT}}$, yielding utility representation $U_i = \frac{q_i^2/2-(\alpha+\epsilon_i)q_i}{\beta_i} + p^{NT}q_i^{NT}$.

[30]We assume that $\bar{m}$ is sufficiently small so that $q_i \geq 0$ with probability one, ruling out instances of negative consumption.

can permit the econometrician to know the entire distribution of $\theta = (\beta, \epsilon)$. It will not affect our results.

First, we can find a homogeneous perceived price that rationalizes the data for any τ. In particular, a linear regression of aggregate demand on sticker prices and taxes may permit identification of $\hat{\beta} \equiv \mathbb{E}[\beta]$ and $\hat{\tilde{\beta}} \equiv \mathbb{E}[\tilde{\beta}]$, respectively.[31] We define a measure of central tendency of tax salience:[32]

$$\hat{m} \equiv \frac{\hat{\tilde{\beta}}}{\hat{\beta}}$$

Then the homogeneous perceived price that rationalizes the data is $\hat{p}^s = \bar{p} + \hat{m}\tau$. To see this, note that assuming all agents have tax salience $m_i = \hat{m}$ yields aggregate demand as in equation 7:

$$\int_{\beta_i,\epsilon_i} [\alpha + \beta_i \hat{p}^s + \epsilon_i] dF^*_{\beta,\epsilon}(\beta_i, \epsilon_i) = \alpha + \bar{p} \int_{\beta_i,\epsilon_i} \beta_i dF^*_{\beta}(\beta_i) + \hat{m}\tau \int_{\beta_i} \beta_i dF^*_{\beta}(\beta_i)$$
$$= \alpha + \hat{\beta}\bar{p} + \hat{m}\hat{\tilde{\beta}}\tau = \alpha + \hat{\beta}\ \bar{p} + \hat{\tilde{\beta}}\ \tau$$

Thus, the agent cannot rule out all agents perceiving the same price $\hat{p}^s$, and so cannot rule out $m_i = \hat{m}\ \forall i$. For tax τ, this would yield deadweight loss, which by theorem 1 is a lower bound:

$$DWL_{low} = -\frac{1}{2}\hat{m}\hat{\tilde{\beta}}\tau^2$$

Alternatively, the econometrician cannot rule out the perceived tax τ^s having support in $\{0, \bar{m}\tau\}$. To see this, consider $\mathbb{P}(p^s = \bar{p} + \bar{m}\tau) = \frac{\hat{m}}{\bar{m}}$ and $\mathbb{P}(p^s = \bar{p}) = 1 - \frac{\hat{m}}{\bar{m}}$ independently of other parameters and regressors.[33] This will rationalize aggregate demand:

$$\int_{\beta_i,\epsilon_i} [\alpha + \beta_i \bar{p}_i + \frac{\hat{m}}{\bar{m}} \beta_i \bar{m} \tau_i + \epsilon_i] dF_{\beta,\epsilon}(\beta_i, \tilde{\beta}_i, \epsilon_i) = \alpha + \mathbb{E}[\beta]\bar{p} + \hat{m}\mathbb{E}[\beta]\tau = \alpha + \mathbb{E}[\beta]\bar{p} + \mathbb{E}[\tilde{\beta}]\tau$$

This yields deadweight loss for tax τ:

$$DWL_{high} = -\frac{1}{2}\frac{\hat{m}}{\bar{m}}\mathbb{E}[\beta]\bar{m}^2\tau^2 = -\frac{1}{2}\hat{m}\hat{\beta}\bar{m}\tau^2 = -\frac{1}{2}\bar{m}\hat{\tilde{\beta}}\tau^2$$

For instance, if $\bar{m} = 1$, then the value of deadweight loss under a homogeneous perceived price is fraction $\hat{m}$ of the above calculation of deadweight loss.

Proceeding from theorem 2 in the previous subsection, we noted that there is a specific distribution of perceived prices on $\{\bar{p}, \bar{p} + \bar{m}\tau\}$ that maximizes deadweight loss. We describe that distribution in theorem 3, noting that it involves assigning high or low perceived prices based on the ratio of per-person deadweight loss to the change in consumption for that individual. But in this context:

$$\frac{dwl_i}{q_i(\bar{p}) - q_i(p^s)} = \frac{\tau^s}{2}$$

[31] Such identification requires exogenous & non-collinear variation in sticker prices & taxes. If the econometrician cannot identify these terms, so much the worse for identifying aggregate deadweight loss.

[32] If $\hat{\beta} = 0$, then let $\hat{m} = 0$.

[33] In the true distribution, it must be that $\hat{m} \in [0, \bar{m}$. Alternatively, one could consider checking whether $\hat{m} \in [0, \bar{m}]$ as a weak test of the null hypothesis that tax salience is bounded within that interval.

Thus, the distribution of tax salience independent of all other parameters and regressors in which $\mathbb{P}(m = \bar{m}) = \frac{\hat{m}}{\bar{m}}$ and $\mathbb{P}(m = 0) = 1 - \frac{\hat{m}}{\bar{m}}$ maximizes deadweight loss. More generally, the econometrician cannot rule out this maximal value of deadweight loss so long as they cannot rule out the possibility of some distribution F with $F_\beta = F^*_\beta$ such that $supp(m) \in \{0, \bar{m}\}$ with:

$$\mathbb{P}_F(m = \bar{m})\mathbb{E}_F[\tilde{\beta}|m = \bar{m}] = \hat{m}\hat{\beta} = \hat{\tilde{\beta}}$$

Mathematically, one can check that such a distribution rationalizes the data and yields the maximal value of deadweight loss:

$$\alpha + \mathbb{E}[\beta]\bar{p} + \mathbb{E}_F[\tilde{\beta}]\tau = \alpha + \hat{\beta}\bar{p} + \mathbb{P}_F[m = \bar{m}]\mathbb{E}_F[\tilde{\beta}|m = \bar{m}]\tau = \alpha + \hat{\beta}\bar{p} + \hat{\tilde{\beta}}\tau$$

$$-\frac{1}{2}\mathbb{E}_F[m^2\beta]\tau^2 = -\frac{1}{2}\mathbb{P}_F[m = \bar{m}]\bar{m}\mathbb{E}_F[\tilde{\beta}|m = \bar{m}]\tau^2 = -\frac{1}{2}\bar{m}\hat{\tilde{\beta}}\tau^2 = DWL_{high}$$

More intuitively, once one knows $\hat{\beta}$ and $\hat{\tilde{\beta}}$, one can rationalize the aggregate data. Since the ratio of deadweight loss to the change in quantity is constant, the relationship between tax salience and preferences doesn't matter upon attaining the observed aggregate demand.

Finally, consider a distribution with $m \perp (\beta, \epsilon)$ with $supp(m) \subseteq \{0, \hat{m}, \bar{m}\}$, $\mathbb{P}(m = \hat{m}) = \lambda$ and $\mathbb{P}(m = \bar{m}|m \neq \hat{m}) = \frac{\hat{m}}{\bar{m}}$. Varying λ from zero to one yields:

$$DWL \in [-\frac{1}{2}\hat{m}\hat{\tilde{\beta}}\tau^2, -\frac{1}{2}\bar{m}\hat{\tilde{\beta}}\tau^2]$$

We can conclude from this result that one cannot even identify a second order approximation of deadweight loss with aggregate data alone.[34] Imposing structure on preferences to facilitate identification of F^*_θ still only permits interval identification. Nonetheless, we can use aggregate data to obtain bounds, or at least $\hat{m}$, which gives us a sense of the uncertainty over the possible values of deadweight loss.

We showcase such an application in section B of the appendix, based on data on aggregate beer consumption from CLK (2009). We replicate the regressions they ran to estimate tax salience, but using a linear (rather than log-log) specification to match this subsection. Taking the ratio of averages across regressions of measures of aggregate demand responses to sales and excise tax variation, we estimate $\hat{m} \approx 0.27$. This estimate suggests that for $\bar{m} \geq 1$, the upper bound of deadweight loss is more than four times the lower bound.

4 Point Identification with Individual-Level Data

The previous section established that attention heterogeneity prevents point identification of deadweight loss using aggregate data, even if we already know the distribution of preferences. To facilitate identification, we require more granular data.[35] If we use (repeated) cross-sectional data, we also require additional structure on tax salience or the consumption set. Alternatively, we can use (long) panel data to identify deadweight loss without any structural assumptions.

[34] One can identify a first order approximation trivially; it is zero.

[35] Technically, one could simply impose a distribution of p^s. For instance, one could assume homogeneous perceived tax-inclusive prices. We do not recommend this.

4.1 Cross-Sectional Data

We maintain the linear structure on preferences, as well as the assumptions that m is constant with respect to τ, as in section 3.3. Formally:

$$q = \alpha + \beta \bar{p} + \tilde{\beta}\tau + \epsilon$$

Here α is a constant whose identification we generally assume.[36] The econometrician observes the distribution of q conditional on $(\bar{p}, \tau)$ for all values in $supp(\bar{p}, \tau)$. The data comes from an underlying data-generating process, which we assume yields a well-defined and finite value for deadweight loss. The data identifies expected deadweight loss from a (non-zero) tax τ if any underlying distributions of $(\beta, \tilde{\beta}, \epsilon)$ across the population consistent with the observed distribution of $(\bar{p}, \tau)$ and conditional distributions of q yield the same value of $\mathbb{E}[\frac{\tilde{\beta}^2}{\beta}]$. One approach is to identify the joint distribution of $(\beta, \tilde{\beta})$ across individuals, then integrate to obtain the desired expected value. To that end, we can use the following lemma from Masten (2017):

Lemma 3. *If $supp(\bar{p}, \tau)$ contains an open ball in $\mathbb{R}^2$, then the joint distribution of $(\beta, \tilde{\beta}, \epsilon)$ is identified if and only if:*

1. *The joint distribution of $(\beta, \tilde{\beta}, \epsilon)$ is determined by its moments.*
2. *All absolute moments of $(\beta, \tilde{\beta}, \epsilon)$ are finite.*

Proof. Apply lemma 2 from Masten (2017). □

It follows as an immediate corollary that the conditions of the lemma also identify deadweight loss, since one can integrate out the marginal distribution of $(\beta, \tilde{\beta})$ and then calculate $\mathbb{E}[\frac{\tilde{\beta}^2}{\beta}]$. The intuition is that one can find the moments of the random coefficients by running increasingly higher-order regressions of the form:

$$\mathbb{E}[q_i^n|\bar{p}, \tau] = ... + \mathbb{E}[\beta^n]\bar{p}^n + ... + \mathbb{E}[\tilde{\beta}^n]\tau^n + ...$$

The enumerated conditions are satisfied, for instance, when $(\beta, \tilde{\beta})$ has finite support. However, one might want to be able to identify deadweight loss without making assumptions on the distribution of $(\beta, \tilde{\beta})$, beyond assuming that $\mathbb{E}[\frac{\tilde{\beta}^2}{\beta}]$ is well-defined and finite. The following theorem does not impose such assumptions, but does require unbounded support in regressors.

Theorem 4. *Suppose that for every pair $(\lambda_1, \lambda_2) \in (\mathbb{R} \setminus \{0\}) \times \mathbb{R}$, there is a sequence $(\bar{p}_k, \tau_k)_{k=1}^{\infty}$ contained within the support of $(\bar{p}, \tau)$ such that:*

1. $\lim_{k\to\infty} ||(\bar{p}_k, \tau_k)|| = \infty$
2. $\lim_{k\to\infty} \frac{\tau_k}{\bar{p}_k} = \frac{\lambda_2}{\lambda_1}$

In addition, suppose that there is another sequence $(\bar{p}_k, \tau_k)_{k=1}^{\infty}$ contained within the support of $(\bar{p}, \tau)$ such that:

3. $\lim_{k\to\infty} |\bar{p}_k| < \infty$
4. $\lim_{k\to\infty} \tau_k = \infty$

[36] It is identified if the econometrician observes $\bar{p} = \tau = 0$ or non-collinear variation in $\bar{p}$ and τ.

Then $\mathbb{E}[dwl_i]$ *is identified.*

We make particular note of a special case in which these results apply. Consider a binary choice problem, so that $X^T \in \{0,1\}$. Each agent i has quasi-linear utility u_{i1} for consuming the good, so that for tax-inclusive price p:

$$u_{i1} = \epsilon_i - p$$

This is an expression of the agent's preference for the taxed good scaled so that an additional dollar yields one unit of utility. We do not assume that $\mathbb{E}[\epsilon] = 0$. However, we normalize utility in the absence of the taxed good to zero:

$$u_{i0} = 0$$

For a price increase from $\bar{p}$ to p^s, the change in the expenditure function is:

$$e(p^s) - e(\bar{p}) = \begin{cases} 0 & \epsilon_i - \bar{p} \leq 0 \\ \epsilon_i - \bar{p} & p^s \geq \epsilon_i > \bar{p} \\ p^s - \bar{p} & \epsilon_i - p^s > 0 \end{cases}$$

Each agent i has a tax salience m_i independent of the tax rate. The agent perceives utility from buying the taxed good:

$$u_{i1}^s = \epsilon_i - \bar{p} - m_i \tau \tag{8}$$

They purchase the good if $u_{i1}^s > 0$. This yields deadweight loss for that agent of:

$$dwl_i = e(p^s) - e(\bar{p}) - [p^s - \bar{p}]\mathbb{I}(\alpha - p^s + \epsilon_i > 0) = \begin{cases} 0 & \epsilon_i - \bar{p} \leq 0 \\ \epsilon_i - \bar{p} & p^s \geq \alpha + \epsilon_i > \bar{p} \\ 0 & \epsilon_i - p^s > 0 \end{cases}$$

Let $F^*_{m,\epsilon}$ denote the true distribution of (m, ϵ). For simplicity, we assume that $\mathbb{P}(u_1^s = 0)$ and $\mathbb{P}(u_1)$ are always zero. We assume the econometrician observes only aggregate data, so that the only values observed are triplets $(\bar{p}, \tau, \mathbb{P}[u_1^s > 0|\bar{p}, \tau])$. The challenge is to infer deadweight loss, which takes the form:

$$DWL = \int_{m_i, \epsilon_i} \mathbb{I}(u_1 \geq 0 > u_1^s)[\epsilon_i - \bar{p}] dF^*_{m,\epsilon}(m_i, \epsilon_i) \tag{9}$$

Fox 2017 shows that if $\bar{p}$ has full support, then one can use aggregate demand in infer the CDF of $\epsilon - m\tau$, allowing one to apply Masten (2017). Thus, we can identify deadweight loss even with aggregate data if we are willing to assume that tax salience does not depend on the tax rate and the choice set is binary.

One might hope that the assumption that $m \perp (\bar{p}, \tau)$ it not required for identification, as it is generally a rather strong assumption. Unfortunately, one cannot do away entirely with this restriction. For instance, consider a population in which $\mathbb{P}(\epsilon = 2) = \mathbb{P}(\epsilon = 3) = 0.5$ and $\mathbb{P}(\zeta = l) = 1$. Consider rationalizing data in two ways. In the first case, $m = 0.5$ with probability one. This yields aggregate demand:

$$D = \begin{cases} 1 & \bar{p} + 0.5\tau \leq 2 \\ 0.5 & 2 < \bar{p} + 0.5\tau \leq 3 \\ 0 & 3 < \bar{p} + 0.5\tau \end{cases}$$

Aggregate deadweight loss takes the form:

$$DWL = \begin{cases} 0 & \bar{p} + 0.5\tau \leq 2 \\ 0.5 & 2 < \bar{p} + 0.5\tau \leq 3 \\ 1.5 & 3 < \bar{p} + 0.5\tau \end{cases}$$

In the second case, suppose for a moment that $\bar{p} \in (0,1)$. Then agents with $\epsilon = 2$ have tax salience $m_2(\bar{p}) = \frac{2-\bar{p}}{6-2\bar{p}}$, while agents with $\epsilon = 3$ have tax salience $m_3(\bar{p}) = \frac{3-\bar{p}}{4-2\bar{p}}$. If $\bar{p} \notin (0,1)$, then set $m_2(\bar{p}) = m_3(\bar{p}) = 0.5$. One can check that this yields the same aggregate demand, yet aggregate deadweight loss with $\bar{p} \in (0,1)$ now takes the form:

$$DWL = \begin{cases} 0 & \bar{p} + 0.5\tau \leq 2 \\ 0.75 & 2 < \bar{p} + 0.5\tau \leq 3 \\ 1.5 & 3 < \bar{p} + 0.5\tau \end{cases}$$

4.2 Panel Data

Now suppose that one can follow specific individuals for a long period of time. For every individual i, the econometrician observes triplets $(\bar{p}_i, \tau_i, q_i)$ with full support for $\bar{p}$ when $\tau = 0$. Then the econometrician can identify the demand function $q_i(p; 0)$. Assuming there are no income effects, deadweight loss for an individual i with perceived price p^s takes the form:

$$dwl_i = \int_{\bar{p}}^{p^s} q(p;0)dp - \tau d(p^s)$$

If one can observe tax τ, then one can determine deadweight loss using the inferred demand function via:[37]

$$p^s = d^{-1}(q_i(\bar{p}, \tau))$$

For instance, with binary choice data, demand for agent i is $\mathbb{I}(\alpha - p_i^s + \epsilon_i)$. This yields deadweight loss for agent i of:[38]

$$dwl_i = \mathbb{I}(p_i^s > \alpha + \epsilon_i > \bar{p})[\alpha - \bar{p}\epsilon_i]$$

This yields aggregate deadweight loss:

$$DWL = \int_{p_i^s, \epsilon_i} \mathbb{I}(p_i^s > \alpha + \epsilon_i > \bar{p})[\alpha - \bar{p} + \epsilon_i] dF^*_{p^s,\epsilon}(p_i^s, \epsilon_i)$$

5 Empirical Calculation

We apply our theoretical results to experimental data from Taubinsky and Rees-Jones (2018). We use a subset of their data in which subjects are randomly divided into two groups of about 1,000 test subjects. For each of twenty goods, one group reports their maximal willingness to pay for that good, while the other group reports the maximal sticker price at which they would be willing to purchase the good subject

[37] If demand is a correspondence, then one can find p^s as the price for which $q_i(\bar{p}, \tau)$ is an element of the demand correspondence at p^s. By lemma 1, this value is unique.

[38] We assume that demand is well-defined with probability one.

to the sales tax. The test subjects' city of residence determines the size of the sales tax.[39]

Data from the experiment allow us to infer aggregate demand for people in the sample. We compare the standard tax arm to the no-tax arm for a single good.[40] The tax arm allows us to consider what aggregate demand is at a tax of τ, equal to the sales tax in vigour in each subject's city of residence. Then we turn to the no-tax arm, which allows us to observe people's willingness to pay in the absence of taxation. We interpret this as the subjects' *true* willingness to pay, meaning that if their declared willingness to pay is ϵ, then they will buy the good if the price they *perceive* is higher than ϵ. In turn, this allows us to find the perceived prices $\hat{p}^s$ from equation 4 and $\tilde{p}^s$ from equation 6 that allow us to apply theorems 1 and 3.

Specifically, we compute aggregate demand under the tax, d^*, by finding the fraction of people in the tax arm who declare a higher willingness to pay than the true sticker price for the object (taken from Amazon). We then proceed to find the homogeneous perceived price that would equalize the demand for the no-tax arm to d^*. Letting ϵ_i indicate the willingness to pay for subject i in the no-tax arm, we then find $(\hat{p}, \hat{\lambda})$ that solve

$$d^* = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(\epsilon_i > \hat{p}) + \hat{\lambda} \mathbb{I}(\epsilon_i = \hat{p})$$

Where $\hat{\lambda}$ is introduced to deal with edgewise cases and satisfy the equation with equality. One could interpret $\hat{\lambda}$ as the probability that $\zeta = h$ in the no-tax sample, or the fraction of indifferent people who end up buying. This, in turn, allows us to find the lower bound of deadweight loss as in equation 9:[41]

$$DWL^l = \frac{1}{N} \sum_{i=1}^{N} \big(\mathbb{I}(\bar{p} < \epsilon_i < \hat{p}) + (1 - \hat{\lambda}) \mathbb{I}(\epsilon_i = \hat{p}) \big) (\epsilon_i - \bar{p})$$

Finding the upper bound of deadweight loss follows a very similar procedure. The problem is then to find $(\tilde{p}, \tilde{\lambda})$ such that

$$d^* = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(\bar{p} \leq \epsilon_i < \tilde{p}) + \tilde{\lambda} \mathbb{I}(\epsilon_i = \tilde{p})$$

Demand is rationalized by having all agents with ϵ_i less than $\tilde{p}$ perceived price $\bar{p}$, and all agents with $\epsilon_i > \tilde{p}$ perceive a price equal to the supremal value of the support of ϵ. The upper bound for deadweight loss is then what we would obtain if all the subjects who value the good most perceived a price so high that it would dissuade them from buying the good.

$$DWL^u = \frac{1}{N} \sum_{i=1}^{N} \big(\mathbb{I}(\epsilon_i > \tilde{p}) + (1 - \tilde{\lambda}) \mathbb{I}(\epsilon_i = \tilde{p}) \big) (\epsilon_i - \bar{p})$$

[39] We describe only the data that we use, which is not all of the data from the experiment. For instance, Taubinsky and Rees-Jones (2018) have a second module in which they ask both groups their willingness to pay in the absence of any sales taxes.

[40] The good we consider is RainStoppers 68-inch oversize windproof golf umbrella.

[41] We are calculating deadweight loss as if there were no other taxes in the economy.

Table 2: Lower and Upper bounds of deadweight loss, as inferred from the data in Taubinsky and Rees-Jones (2018). We consider one lower bound and four different upper bounds. UB1 is the upper bound one would obtain assuming that salience is independent of preferences and that maximal salience is 1 (that is, that people do not over-react to the tax). UB2 is the upper bound one would obtain only assuming that salience is independent of preferences. UB3 is the upper bound one would obtain assuming only that maximal salience is 1. UB4 is the highest possible upper bound. We also report what deadweight loss would be if we assumed that people in the non-taxed sample accounted for taxes perfectly, and the naive calculation the econometrician would obtain if they assumed the observed reaction to taxes in the taxed sample was induced by full salience. For comparison, the implied average tax revenue from demand under the tax is \$0.1772, and the implied total pre-tax consumer surplus is \$0.4866.

	Full Salience	Naive	LB	UB1	UB2	UB3	UB4
Dollar value	0.0187	0.0266	0.0038	0.0068	0.0441	0.0100	0.2097
Percent of revenue collected	10.55%	15.01%	2.14%	3.84%	24.89%	5.64%	118.34%

The calculations are very similar when we assume that $\bar{m} = 1$, except in that case anyone with $\epsilon_i > \tau$ is assumed to always buy, thus resulting in a smaller upper bound. We also compute the implied value of deadweight loss one would obtain assuming that preferences and salience are independent. In the case where the econometrician does not impose any assumptions on what $\bar{m}$ might be, so that anyone who perceives the high price will not buy the good, the problem reduces to picking α so that

$$d^* = \frac{1}{N}\sum_{i=1}^{N} \alpha \mathbb{I}(\epsilon_i \geq \bar{p})$$

And the resulting upper bound on deadweight loss will be

$$DWL^u_{m\perp\epsilon} = (1-\alpha)\frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(\epsilon \geq \bar{p})(\epsilon_i - \bar{p})$$

We summarize our results in table 2.[42] As we can see, the lower and upper bounds are quite far apart, implying that the exact co-distribution of preferences and perceived prices can have widely different implications for welfare. In particular, one could find an analogue for $\hat{m}$ by dividing the lower bound by the upper bound in the case of $\bar{m} = 1$. This would yield $\hat{m} = 0.38$, similar to $\hat{m} \approx 0.27$ as in subsection 3.3.

Our empirical results suggest two takeaways. One, we require strong assumptions to get an upper bound for deadweight loss close to the lower bound. To rule out deadweight loss multiple times greater than the deadweight loss with a homogeneous perceived price, the econometrician must assume both that tax salience is bounded between zero and one, and is also distributed independently of the willingness to pay.

[42]One limitation of our result is that we treat the sample as if it was the population in question. We conjecture that one could obtain standard errors for estimates by using asymptotic normality from the Generalized Method of Moments estimator. We intend to prove this claim and include standard errors in future versions of this paper.

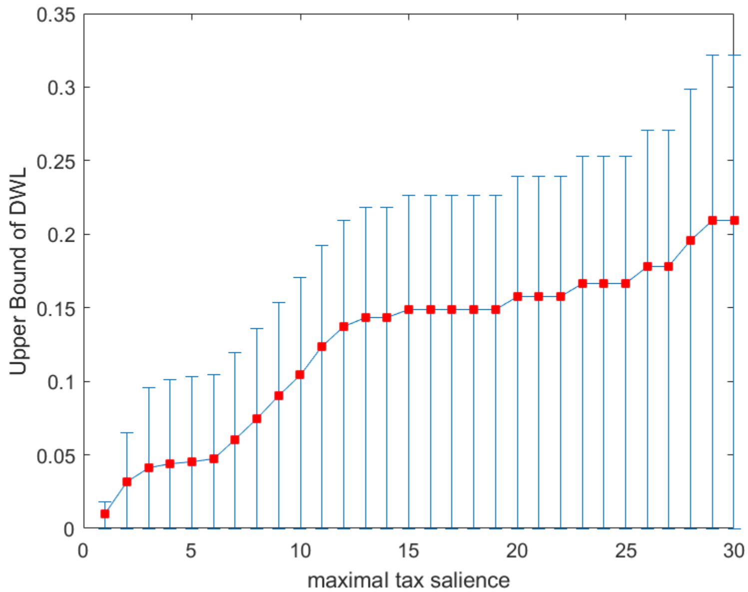


Figure 9: Estimated upper bound depending on the econometrician's choice of $\bar{m}$. Error bars are based on the 5^{th} and 95^{th} percentiles of 10,000 bootstrap replications.

Two, comparing UB1 to UB2 and UB3 suggests that that the potential dependent statistical relationship between tax salience and willingness to pay does not facilitate deadweight loss nearly as much as the possibility of agents perceiving a tax rate greater than the true tax rate. This motivates us to investigate how the upper bound of deadweight loss varies for different choices for maximal salience $\bar{m}$, reported in figure 9.

Note that the interval of 5^{th} and 95^{th} percentiles in 10,000 bootstrap replications always includes zero. This is because we cannot reject the null hypothesis that maximal sticker price under the tax arm is equal to the maximal sticker price under the no-tax arm. For this reason, any 95% confidence interval for deadweight loss will include the possibility that $DWL = 0$, as we do not allow for the possibility of negative salience.

The high variation in empirical estimates suggests that precisely inferring deadweight loss from aggregate data would require strong assumptions. On the other hand, for most reasonable values of $\bar{m}$, deadweight loss seems relatively small. In turn, this suggests that attempting to raise the same revenue with a less distortive tax scheme might not be worth it, if it incurs high enough administrative costs.[43]

6 Conclusion

In this paper, we considered identification of deadweight loss with non-salient taxes, as in the Gabaix (2014) model. We justify considering price misperception by demonstrating that a general model of consumer behavior with weak convexity and continuity assumptions on preferences is observationally equivalent to the Gabaix (2014) model.

[43]Formally, if the administrative cost of implementing a more efficient tax that collected the same revenue was greater than around 32 cents per consumer in the market, one could conclude with high confidence that the change in tax regime could not yield a Pareto improvement with transfers.

We first consider identification of deadweight loss using aggregate data. We show that deadweight loss cannot be point-identified with aggregate demand data. Nonetheless, we provide bounds for deadweight loss consistent with aggregate demand and the distribution of preference parameters across individuals. The lower bound holds for any distribution; the upper bound relies on the assumption that tax salience has support contained in a known non-negative interval. If the econometrician could not infer the distribution of preference parameters or did not know a non-negative interval that contained the support of tax salience, then the interval of possible values of deadweight loss may be even larger.

We provide context for these theoretical results with empirical findings using experimental data from Taubinsky and Rees-Jones (2018). We calculate the upper and lower bounds of deadweight loss due to sales taxes in the United States on a particular good with a binary choice set. We find an upper bound approximately fifty five times larger than the lower bound. Interestingly, this upper bound is substantially larger than the deadweight loss if all agents perfectly accounted for sales taxes. In other words, deadweight loss with imperfectly perceived taxes can be greater than in the standard model of consumer behavior. This result arises from the possibility that some agents that value the good a lot perceive a prohibitively high price, whereas other agents who value the good at barely more than the sticker price still purchase the good because they fail to notice the tax at all. Thus, the behavioral model does not rule out deadweight loss greater than in the standard model even though aggregate consumption is less distorted in the behavioral model. The allocative inefficiency swamps the distortion in aggregate consumption. Indeed, even when assuming that no agents respond more to sales taxes than sticker price changes, we find for the binary and continuous choice settings we study that deadweight loss can be roughly as large as two & a half and four times as large as when assuming homogeneous tax salience, respectively.

With individual-level cross-sectional data, identification arises from assuming choice functions that are linear in sticker prices and taxes, and that tax salience is independent of sticker prices and taxes. Under these assumptions, one can identify the joint distribution of parameters of demand responsiveness to sticker prices and taxes, and use this to compute aggregate deadweight loss. With long panel data, we can identify deadweight loss without these assumptions. Identification then arises from direct calculation of the lost surplus from agents facing the tax compared to when there was no tax.

In future work, we will explore the performance of estimators of deadweight loss using individual-level cross-sectional data. The assumptions on preferences and tax salience facilitate identification, but leave open two questions. One, our identification argument in theorem 4 does not inform us about rates of convergence. Two, the structural assumptions may not hold in practice, and we would like to have some sense of the magnitude of the bias that these assumptions introduce.

We also hope to determine to what degree one can identify deadweight loss in the intermediate case of individual-level cross-sectional data in which tax salience depends arbitrarily on the tax rate and preferences, but does not depend on the sticker price. This assumption of independence of the tax salience from the sticker price seems like the most reasonable non-trivial independence assumption one could impose on tax salience.

Finally, future work should inform what reasonable ex-ante restrictions on tax salience one can impose when observing aggregate data. Our empirical results show widespread uncertainty as to what deadweight loss might be with aggregate data in the absence of restrictions on tax salience. In fact, we still assume that tax salience is non-negative with probability one. Relaxing this assumption could yield an even greater upper bound for deadweight loss.

References

Rudolf Beran and Peter Hall. Estimating coefficient distributions in random coefficient regressions. *The annals of Statistics*, pages 1970–1984, 1992.

Rudolf Beran, Andrey Feuerverger, Peter Hall, et al. On nonparametric estimation of intercept and slope distributions in random coefficient regression. *The Annals of Statistics*, 24(6):2569–2592, 1996.

Raj Chetty, Adam Looney, and Kory Kroft. Salience and taxation. *NBER Working Paper N. 13330*, 2007.

Raj Chetty, Adam Looney, and Kory Kroft. Salience and taxation: Theory and evidence. *The American Economic Review*, 99(4):1145–1177, 2009.

Emmanuel Farhi and Xavier Gabaix. Optimal taxation with behavioral agents. *NBER Working Paper N. 21525*, 2015.

Amy Finkelstein. Ez-tax: Tax salience and tax rates. *Quarterly Journal of Economics*, 2009.

Jeremy T Fox. A note on nonparametric identification of distributions of random coefficients in multinomial choice models. *NBER Working Paper No. 23621*, 2017.

Xavier Gabaix. A sparsity-based model of bounded rationality. *Quarterly Journal of Economics*, 129(4):1661 1710, 2014.

Kelly Sims Gallagher and Erich Muehlegger. Giving green to get green? incentives and consumer adoption of hybrid vehicle technology. *Journal of Environmental Economics and Management*, 61(1):1–15, 2011.

Jacob Goldin. Optimal tax salience. *Journal of Public Economics*, 131:115–123, 2015.

Jacob Goldin and Tatiana Homonoff. Smoke gets in your eyes: cigarette tax salience and regressivity. *American Economic Journal: Economic Policy*, 5(1), 2013.

Stefan Hoderlein, Jussi Klemelä, and Enno Mammen. Analyzing the random coefficient model nonparametrically. *Econometric Theory*, 26(3):804–837, 2010.

Harvey S Rosen. Tax illusion and the labor supply of married women. *The Review of Economics and Statistics*, pages 167–172, 1976.

Dmitry Taubinsky and Alex Rees-Jones. Attention variation and welfare: theory and evidence from a tax salience experiment. *forthcoming at Review of Economic Studies*, 2018.

Appendix

A Additional Results and Proofs

A.1 Additional results and proofs from section 2

We first demonstrate a generalization of the Gabaix representation, in which multiple goods may be taxed. We consider a general setting with N goods, consumption set $X = X^T \times X^{NT} \subseteq \mathbb{R}^N_+$, with consumption vector $\boldsymbol{q} = (\boldsymbol{q^T}, \boldsymbol{q^{NT}}) \in X$. Here X^T is the consumption set for taxed goods, while X^{NT} is the consumption set for non-taxed goods. We assume that either $X^{NT} \subseteq \mathbb{R}_+$ or X^{NT} is convex.

The agent has preferences $\succeq$ on X. Informally, we want to assume preferences such that agents smoothly prefer moderation. To say that they prefer moderation, one generally assumes convex preferences. However, we do not want to assume a convex consumption set X. We might alternatively assume that preferences are *pseudo-convex*, in that for any $\boldsymbol{q} \in X$ and any finite n:

$$\boldsymbol{q}_k \in X, \boldsymbol{q}_k \succ \boldsymbol{q}, \lambda_k \geq 0 \ \forall k = 1, ..., n, \sum_{k=1}^{n} \lambda_k = 1, \sum_{k=1}^{n} \lambda_k \boldsymbol{q}_k \in X \Rightarrow \sum_{k=1}^{n} \lambda_k \boldsymbol{q}_k \succ \boldsymbol{q}$$

However, we also want some smoothness to preferences. More formally, we want to figure that if $\boldsymbol{q'} \succ \boldsymbol{q}$, then there is an epsilon ball around $\boldsymbol{q'}$ such that the agent would prefer any element in that epsilon ball to $\boldsymbol{q}$ if that element were also in the consumption set. Furthermore, any convex combination of points in these epsilon balls should yield a point that, if contained in X, is also strictly preferred to $\boldsymbol{q}$. We refer to this assumption on preferences as *continuous pseudo-convexity* (CPC).

Assumption 2. *For any $\boldsymbol{q} \in X$, define the set of strictly preferred allocations:*

$$\mathcal{A} \equiv \{\boldsymbol{q'} \in X | \boldsymbol{q'} \succ \boldsymbol{q}\}$$

There exists some function $\epsilon : \mathcal{A} \to \mathbb{R}_{++}$ such that for any $n \in \mathbb{N}$, for any $\lambda_1, ..., \lambda_n \geq 0$ and $\boldsymbol{q}_1, ..., \boldsymbol{q}_n \in \mathcal{A}$, if $\sum_{k=1}^{n} \lambda_k = 1$, then :

$$\exists \boldsymbol{q}'_1, ..., \boldsymbol{q}'_n \in \mathbb{R}^n : ||\boldsymbol{q}'_k - \boldsymbol{q}_k|| < \epsilon(\boldsymbol{q_k}) \ \forall k, \sum_{k=1}^{n} \lambda_k \boldsymbol{q_k}' \in X \Rightarrow \sum_{k=1}^{n} \lambda_k \boldsymbol{q_k}' \succ \boldsymbol{q}$$

We provide this description of CPC preferences to facilitate intuition, but our main result for this section comes from an equivalent, yet more geometric, expression of this description.

Lemma 4. *Preferences $\succeq$ are CPC if and only if for every $\boldsymbol{q} \in X$ with corresponding set of strictly preferred bundles $\mathcal{A}$ there is an open and convex set $\mathcal{O} \subseteq \mathbb{R}^N$ such that $\mathcal{O} \cap X = \mathcal{A}$.*[44]

Proof. For one direction, the convex hull of the union of open $\epsilon(\boldsymbol{q'})$ balls around $\boldsymbol{q'} \in \mathcal{A}$ is open, and by assumption does not contain any elements of $X \setminus \mathcal{A}$. For the other direction, for any $\boldsymbol{q'} \in \mathcal{A}$, define $\epsilon(\boldsymbol{q'})$ as a positive value such that $\boldsymbol{q''} \in \mathbb{R}^N_+ : ||\boldsymbol{q''} - \boldsymbol{q'}|| < \epsilon(\boldsymbol{q'}) \Rightarrow \boldsymbol{q''} \in \mathcal{O}$. We can do so because $\mathcal{O}$ is open. For any such $\boldsymbol{q''}$, if $\boldsymbol{q''} \in X$, then $\boldsymbol{q''} \succ \boldsymbol{q}$. □

[44] Note that $\mathcal{O}$ is open in $\mathbb{R}^N$.

Let $\boldsymbol{p} = (\boldsymbol{p^T}, \boldsymbol{p^{NT}}) \in \mathbb{R}_+^N$ denote a generic price vector, where $\boldsymbol{p^T}$ and $\boldsymbol{p^{NT}}$ are price vectors for taxed and non-taxed goods respectively. In particular, let $\bar{\boldsymbol{p}} = (\bar{\boldsymbol{p}}^{\boldsymbol{T}}, \bar{\boldsymbol{p}}^{\boldsymbol{NT}})$ denote the vector of sticker prices.

Let $\boldsymbol{\tau}$ denote the vector of taxes for taxed goods, so that $\boldsymbol{q^T}$, $\boldsymbol{q^{NT}}$, and $\boldsymbol{\tau}$ all have the same number of elements. The consumption vector $\boldsymbol{q}(\bar{\boldsymbol{p}}, \boldsymbol{\tau}) = (\boldsymbol{q}^T(\bar{\boldsymbol{p}}, \boldsymbol{\tau}), \boldsymbol{q}^{NT}(\bar{\boldsymbol{p}}, \boldsymbol{\tau}))$ satisfies the following properties:

$$\begin{aligned}
\bar{p}^{NT} * \tilde{\boldsymbol{q}}^{NT} &\leq W - \bar{\boldsymbol{p}}^{\boldsymbol{T}} * \boldsymbol{q}^T \\
(\boldsymbol{q}^T, \tilde{\boldsymbol{q}}^{NT}) &\succ (\boldsymbol{q}^T, \hat{\boldsymbol{q}}^{NT}) \ \ \forall \hat{\boldsymbol{q}}^{NT} \in X^{NT} : \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \tilde{\boldsymbol{q}}^{NT} \leq W - \bar{\boldsymbol{p}}^{\boldsymbol{T}} * \boldsymbol{q}^T \\
\boldsymbol{q}(\bar{\boldsymbol{p}}, \boldsymbol{0}) &\in \underset{\tilde{\boldsymbol{q}} \in X : \bar{\boldsymbol{p}} * \tilde{\boldsymbol{q}} \leq W}{\arg\max} \succeq
\end{aligned}$$

In words, consumption of the non-taxed goods is always optimally determined upon choosing consumption of the taxed goods, and consumption is optimally determined when the agent correctly perceives prices, i.e. when there are no taxes. We also restrict the domain of sticker prices and taxes so that expenditure on non-taxed goods is positive, i.e.:

$$\bar{\boldsymbol{p}}^{NT} * \boldsymbol{q}^{NT}(\bar{\boldsymbol{p}}, \boldsymbol{\tau}) > 0$$

The claim is that for any $\bar{\boldsymbol{p}}$ and $\boldsymbol{\tau}$ in this domain, there is a Gabaix representation for $\boldsymbol{q}(\bar{\boldsymbol{p}}, \boldsymbol{\tau})$.

Proof of Generalization of Proposition 1. Define $\boldsymbol{q} = (\boldsymbol{q^T}, \boldsymbol{q^{NT}}) = \boldsymbol{q}(\bar{\boldsymbol{p}}, \boldsymbol{\tau})$ and:

$$\mathcal{A}^e \equiv \{(\boldsymbol{q^{T\prime}}, e^{NT\prime}) | \boldsymbol{q^{T\prime}} \in X^T, \exists \boldsymbol{q^{NT\prime}} \in X^{NT} : \bar{\boldsymbol{p}}^{NT} * \boldsymbol{q^{NT\prime}} = e^{NT\prime}, (\boldsymbol{q}^{T\prime}, \boldsymbol{q}^{NT\prime}) \in \mathcal{O}\}$$

Suppose for the sake of contradiction that $(\boldsymbol{q^T}, \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \boldsymbol{q^{NT}}) \in Co(\mathcal{A}^e)$, i.e. that $\exists n \in \mathbb{N}$, $(\boldsymbol{q}_k^T, e_k^{NT}) \in \mathcal{A}^e$, and $\lambda_k \geq 0 \ \forall k = 1, ..., n$ such that $\sum_{k=1}^n \lambda_k = 1$ and:

$$\sum_{k=1}^n \lambda_k (\boldsymbol{q}_k^T, e_k^{NT}) = (\boldsymbol{q^T}, \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \boldsymbol{q^{NT}})$$

Since $(\boldsymbol{q}_k^T, e_k^{NT}) \in \mathcal{A}^e \ \forall k$, that means that:

$$\forall k \ \exists \boldsymbol{q}_k^{NT} : e_k^{NT} = \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \boldsymbol{q}_k^{NT}, \boldsymbol{q}_k \equiv (\boldsymbol{q}_k^T, \boldsymbol{q}_k^{NT}) \Rightarrow \boldsymbol{q}_k \in \mathcal{O}$$

If $X^{NT} \subseteq \mathbb{R}_+$, then $\sum_{k=1}^n \lambda_k \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \boldsymbol{q}_k^{NT} = \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \boldsymbol{q^{NT}}$ implies that $\sum_{k=1}^n \lambda_k \boldsymbol{q}_k^{NT} = \boldsymbol{q^{NT}}$ because positive non-tax expenditure requires that $\bar{\boldsymbol{p}}^{\boldsymbol{NT}} \neq 0$. In that case:

$$\sum_{k=1}^n \lambda_k \boldsymbol{q}_k = \boldsymbol{q} \Rightarrow\Leftarrow \sum_{k=1}^n \lambda_k \boldsymbol{q}_k \in \mathcal{O}$$

This is a contradiction arising from $\boldsymbol{q} \notin \mathcal{O}$.

If X^{NT} is not a subset of $\mathbb{R}_+$, then X^{NT} is convex. This means $\sum_{k=1}^n \lambda_k \boldsymbol{q}_k \in X$. Pseudo-convexity of preferences implies that:

$$\sum_{k=1}^n \lambda_k \boldsymbol{q}_k \succ \boldsymbol{q}$$

Yet the weighted average of taxed goods is the desired taxed good consumption bundle, whereas the weighted average of non-taxed goods is affordable:

$$\sum_{k=1}^{n} \lambda_k \boldsymbol{q_k^T} = \boldsymbol{q}^T$$

$$\bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \sum_{k=1}^{n} \boldsymbol{q_k^{NT}} = \sum_{k=1}^{n} e_k^{NT} = \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \boldsymbol{q}^{NT}$$

Thus, the agent could not have optimally chosen $\boldsymbol{q}^{NT}$, another contradiction. We conclude that $(\boldsymbol{q}^T, \bar{\boldsymbol{p}}^{NT} * \boldsymbol{q}^{NT}) \notin Co(\mathcal{A}^e)$.

Now, we can apply the Separating Hyperplane Theorem to say that there is a vector $(\boldsymbol{p}^{Ts}, 1)$, where $\boldsymbol{p}^{Ts}$ has as many elements as $\boldsymbol{q}^T$, such that:

$$(\boldsymbol{p}^{Ts}, 1) * (\boldsymbol{q}^T, \bar{\boldsymbol{p}}^{NT} * \boldsymbol{q}^{NT}) \leq (\boldsymbol{p}^{Ts}, 1) * (\boldsymbol{q}^{\boldsymbol{T}'}, e^{NT'}) \; \forall (\boldsymbol{q}^{\boldsymbol{T}'}, e^{NT'}) \in Co(\mathcal{A}^e)$$

Defining $\boldsymbol{p}^s \equiv (\boldsymbol{p}^{Ts}, \bar{\boldsymbol{p}}^{NT})$, this implies that for any bundle $\boldsymbol{q}' = (\boldsymbol{q}^{T'}, \boldsymbol{q}^{NT'}) \in \mathcal{O}$:

$$\boldsymbol{p}^s * \boldsymbol{q}^{'} \geq \boldsymbol{p}^s * \boldsymbol{q}$$

Since $\mathcal{O}$ is open, the above expression can never be satisfied with equality. To see this, suppose otherwise, i.e. that $\exists \boldsymbol{q}' \in \mathcal{O}$ such that:

$$\boldsymbol{p}^s * \boldsymbol{q}^{'} = \boldsymbol{p}^s * \boldsymbol{q}$$

Note that $\bar{\boldsymbol{p}}^{NT} > \boldsymbol{0}$ implies that we can choose $\boldsymbol{q}''$ within $\epsilon(\boldsymbol{q}')$ of $\boldsymbol{q}'$ by slightly reducing a component of $\boldsymbol{q}'$ for which the corresponding perceived price is positive. Thus, $\boldsymbol{q}'' \in \mathcal{O}$, yet $\boldsymbol{p}^s * \boldsymbol{q}'' < \boldsymbol{p}^s * \boldsymbol{q}$. This yields our desired contradiction. Therefore:

$$\boldsymbol{p}^s * \boldsymbol{q}' > \boldsymbol{p}^s * \boldsymbol{q} \; \forall \boldsymbol{q}' \in \mathcal{O}$$

We conclude by defining $W^s \equiv \boldsymbol{p}^s * \boldsymbol{q}$ and noting that $\forall \boldsymbol{q}' \in X$:

$$\boldsymbol{q}' \succ \boldsymbol{q} \Rightarrow \boldsymbol{q}' \in \mathcal{O} \Rightarrow \boldsymbol{p^s} > W^s$$

Therefore, the Gabaix model has rationalized consumption because no preferred consumption bundle is perceived to be affordable. □

Now that we've gone through the proof, we can make a couple of observations. One, the assumption of CPC preferences is satisfied when preferences are represented by a lower semi-continuous and quasi-concave function u on $\mathbb{R}^N$, so that:

$$\forall x, y \in X : x \succeq y \Leftrightarrow u(x) \geq u(y)$$

This makes it clear that we have, in fact, generalized Proposition 1. Also, note that it may be easier in practice to check to see that preferences have such a utility representation than to check that they satisfy continuous pseudo-convexity.

Two, it may appear strange that we needed to assume that X^{NT} is concave specifically if it has dimension greater than one. This is because a discrete grid for consumption of non-taxed goods can create a lumpy evaluation of non-tax expenditure, thwarting the existence of a separating hyperplane. For example, consider a consumption set $\mathbb{R}_+ \times \{0,1\}^2$, where there is one taxed good chosen continuously and two non-taxed goods chosen from $\{0,1\}$. The sticker price vector is $\bar{p} = (1,1,1)$. The consumer have preferences rationalized by the function:

$$u(\boldsymbol{q}) = q_1 + \min\{q_2, q_3\}$$

In words, the taxed good is perfect substitutes with the minimum consumption of the two non-taxed goods, which are perfect complements with each other. Consider the consumption bundle:

$$\boldsymbol{q} = (0,1,0)$$

If the agent perceived income $W^s \geq 2$, they could do better by consuming $(0,1,1)$. Supposing otherwise, if the agent perceives a positive tax-inclusive price of the taxed good, then optimally $q_1 > 0$ and $q_2 = q_3 = 0$. Finally, there is no optimal consumption bundle if $p_1^s \leq 0$. Thus, the consumption bundle cannot be rationalized.

Details from Table 1: Let $\bar{p} = \tau = p^{NT} = \bar{m} = 1$, $X^T = \{0,1\}$, and $u(q^T, q^{NT}; \theta_i) = \theta_i * q^T + q^{NT}$.In words, the marginal cost of the taxed good is zero, the tax-inclusive prices are one, salience is between zero & one, the choice of the taxed good is either zero or one, and the taxed & non-taxed goods are symmetric perfect substitutes. Suppose:

$$F_\theta(\theta_i) = 0.5\mathbb{I}(\theta_i \geq 1) + 0.5\mathbb{I}(\theta_i \geq 0)$$

This says that $\theta \in \{0,1\}$ with equal probability. We consider two distinct methods of distributing perceived prices on $\{0,1\}$ dependent on preference type to rationalize an aggregate demand of 0.5. First, suppose that $\theta_i = 1 \Rightarrow p_i^s = 1, \zeta_i = h$ and $\theta_i = 2 \Rightarrow p_i^s = 2, \zeta_i = l$. Then:

$$DWL = 0.5[e(1;1) - e(1;1) - 0] + 0.5[e(2;2) - e(1;2) - 0] = 0.5$$

Second, suppose that $\theta_i = 1 \Rightarrow p_i^s = 1, \zeta_i = l$ and $\theta_i = 2 \Rightarrow p_i^s = 2, \zeta_i = h$. Then:

$$DWL = 0.5[e(1;1) - e(1;1)] + 0.5[e(2;2) - e(1;2) - 1] = 0$$

Proposition 4. *Assume a continuously differentiable and strictly increasing aggregate supply function Q^{supply}, as well as continuously differentiable compensated demand functions h_i and subjective price functions p_i^s $\forall i$. Subjective price functions change one-for-one with sticker prices, so that:*

$$p_i^s(\bar{p}, \tau) = \bar{p} + p_i^s(0, \tau) \ \forall \bar{p} \ \forall \tau \ \forall i$$

Subjective prices also agree with sticker prices when there is no tax:

$$p_i^s(\bar{p}, 0) = \bar{p} \quad \forall \bar{p} \ \forall i$$

We implicitly define the pre-tax sticker price $\bar{p}^{old}$ by:[45]

$$Q^{supply}(\bar{p}^{old}) = \sum_i h_i(\bar{p}^{old}, \nu_i)$$

[45] $\nu_i \equiv u_i(\boldsymbol{d}_i(\boldsymbol{p}, W_i)) \ \forall i$

and the new sticker price $\bar{p}^{new}$ after the imposition of the tax τ when agents are compensated by:

$$Q^{supply}(\bar{p}^{new}) = \sum_i h_i((p_i^s(\bar{p}^{new}, \tau)), \nu_i)$$

Defining deadweight loss by:[46]

$$DWL \equiv \sum_i \Delta CS_i + \int_{\bar{p}^{new}}^{\bar{p}^{old}} Q^{supply}(p)dp - \tau \sum_i q_i^c$$

where

$$\Delta CS_i = (\bar{p}^{new} + \tau - p_i^s(\bar{p}^{new}, \tau))q_i^c + \int_{\bar{p}^{old}}^{p_i^s(\bar{p}^{new},\tau)} h_i(p, \nu_i)dp \ \forall i$$

$$q_i^c \equiv h_i(p_i^s(\bar{p}^{new}, \tau), \nu_i) \ \forall i$$

then aggregate deadweight loss has second order approximation around $\tau = 0$:

$$DWL \approx -\frac{1}{2}[\sum_i m_i \frac{\partial h_i}{\partial p} - \frac{(\sum_i m_i \frac{\partial h_i}{\partial p})^2}{\sum_i \frac{\partial h_i}{\partial p} - \frac{\partial Q^{supply}}{\partial p}}]\tau^2$$

Proof.

$$DWL = \sum_i \int_{\bar{p}^{old}}^{p_i^s(\bar{p}^{new},\tau)} h_i(p, \nu_i)dp + \int_{\bar{p}^{new}}^{\bar{p}^{old}} Q^{supply}(p)dp + \sum_i (\bar{p}^{new} - p_i^s(\bar{p}^{new}, \tau))q_i^c$$

Note that $\bar{p}^{new}$ is a function of τ. One can easily confirm that $\bar{p}^{new}|_{\tau=0} = \bar{p}^{old}$, so that deadweight loss is zero when $\tau = 0$. We can find $\frac{\partial \bar{p}^{new}}{\partial \tau}$ from the Inverse Function Theorem:[47]

$$\frac{\partial Q^{supply}}{\partial p}\frac{\partial \bar{p}^{new}}{\partial \tau} = \sum_i \frac{\partial h_i}{\partial p}[\frac{\partial p_i^s}{\partial \bar{p}^{new}}\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}] = \sum_i \frac{\partial h_i}{\partial p}[\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]$$

$$\frac{\partial \bar{p}^{new}}{\partial \tau} = \frac{\sum_i \frac{\partial h_i}{\partial p}\frac{\partial p_i^s}{\partial \tau}}{\frac{\partial Q^{supply}}{\partial p} - \sum_i \frac{\partial h_i}{\partial p}}$$

We can then take the first derivative of deadweight loss with respect to the tax:

$$\begin{aligned}
\frac{\partial DWL}{\partial \tau} &= \sum_i [\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]h_i - \frac{\partial \bar{p}^{new}}{\partial \tau} Q^{supply}(\bar{p}^{new}) \\
&- \sum_i \left[\frac{\partial p_i^s}{\partial \tau}h_i + (p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new})\frac{\partial h_i}{\partial p}[\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]\right] \\
&= \frac{\partial \bar{p}^{new}}{\partial \tau}\sum_i h_i - \frac{\partial \bar{p}^{new}}{\partial \tau} Q^{supply}(\bar{p}^{new}) - \sum_i \left[(p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new})\frac{\partial h_i}{\partial p}[\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]\right] \\
&= -\sum_i \left[(p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new})\frac{\partial h_i}{\partial p}[\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]\right]
\end{aligned}$$

[46] Note that $\bar{p}^{new} \leq \bar{p}^{old} \ \forall \tau \geq 0$ from the Law of Compensated Demand and the fact that supply is strictly increasing in price.

[47] This claim also uses the fact that aggregate supply is strictly increasing while aggregate compensated demand is weakly decreasing, so that there is always a unique value for $\bar{p}^{new}$.

Since $p_i^s(\bar{p}^{new}, 0) = \bar{p}^{new}$, it follows that

$$\frac{\partial DWL}{\partial \tau}\bigg|_{\tau=0} = 0$$

Obtaining the second derivative would be straightforward if $h_i \in \mathbb{C}^2 \ \forall i$. Instead, we find the second derivative at $\tau = 0$ from the definition:

$$\frac{\partial^2 DWL}{\partial \tau^2}\bigg|_{\tau=0} = \lim_{\tau \to 0} -\frac{\sum_i \left[(p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new})\frac{\partial h_i}{\partial p}[\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]\right]}{\tau}$$

Note that continuity of $\frac{\partial p_i^s}{\partial \tau}$ with respect to τ for all agents implies that $\frac{\partial \bar{p}^{new}}{\partial \tau}$ is continuous. Since $\frac{\partial Q^{supply}}{\partial p}$ and $\frac{\partial h_i}{\partial p}$ $\forall i$ are also continuous:

$$\begin{aligned}
\frac{\partial^2 DWL}{\partial \tau^2}\bigg|_{\tau=0} &= -\sum_i \frac{\partial h_i}{\partial p}\bigg|_{\tau=0}[\frac{\partial \bar{p}^{new}}{\partial \tau}\bigg|_{\tau=0} + \frac{\partial p_i^s}{\partial \tau}\bigg|_{\tau=0}] \lim_{\tau \to 0} \frac{(p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new})}{\tau} \\
&= -\sum_i \frac{\partial h_i}{\partial p}\bigg|_{\tau=0}[\frac{\partial \bar{p}^{new}}{\partial \tau}\bigg|_{\tau=0} + \frac{\partial p_i^s}{\partial \tau}\bigg|_{\tau=0}]\frac{\partial p_i^s}{\partial \tau}\bigg|_{\tau=0}
\end{aligned}$$

Using the fact that $m_i \equiv \frac{\partial p_i^s}{\partial \tau}|_{\tau=0}$, we can note that:

$$\frac{\partial \bar{p}^{new}}{\partial \tau}|_{\tau=0} = \frac{\sum_i m_i \frac{\partial h_i}{\partial p}}{\frac{\partial Q^{supply}}{\partial p} - \sum_i \frac{\partial h_i}{\partial p}}$$

and so:

$$\frac{\partial^2 DWL}{\partial \tau^2}|_{\tau=0} = -[\sum_i m_i^2 \frac{\partial h_i}{\partial p} + \frac{(\sum_i m_i \frac{\partial h_i}{\partial p})^2}{\frac{\partial Q^{supply}}{\partial p} - \sum_i \frac{\partial h_i}{\partial p}}]$$

Now we can find the second order approximation for deadweight loss:

$$DWL \approx DWL|_{\tau=0} + \frac{\partial DWL}{\partial \tau}|_{\tau=0}\tau + \frac{1}{2}\frac{\partial^2 DWL}{\partial \tau^2}|_{\tau=0}\tau^2$$

$$DWL \approx -\frac{1}{2}[\sum_i m_i^2 \frac{\partial h_i}{\partial p} + \frac{(\sum_i m_i \frac{\partial h_i}{\partial p})^2}{\frac{\partial Q^{supply}}{\partial p} - \sum_i \frac{\partial h_i}{\partial p}}]\tau^2$$

□

A.2 Additional results and proofs from section 3

Proof of Lemma 1. Note that there must be values q^{NT} and $q^{NT'}$ such that:

$$(q(p; \theta_i, l), q^{NT}) \sim_{\theta_i} (q(p'; \theta_i, h), q^{NT'})$$

From local non-satiation:

$$p * q(p; \theta_i, l) + p^{NT} * q^{NT} \leq p * q(p'; \theta_i, h) + p^{NT} * q^{NT'}$$
$$p' * q(p'; \theta_i, h) + p^{NT} * q^{NT'} \leq p' * q(p; \theta_i, l) + p^{NT} * q^{NT}$$

Rearranging yields:

$$p * [q(p; \theta_i, l) - q(p'; \theta_i, h)] \leq p^{NT} * [q^{NT'} - q^{NT}] \leq p' * [q(p; \theta_i, l) - q(p'; \theta_i, h)]$$

Thus, $p' > p \Rightarrow q(p; \theta_i, l) \geq q(p'; \theta_i, h)$. □

Proof of Proposition 3. From lemma 1 and prices being bound away from zero, we can always find a value of $\hat{p}^s$ such that:

$$\int_{\theta_i} q(\hat{p}^s;\theta_i,l)dF_\theta(\theta_i) \leq \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \leq \int_{\theta_i} q(\hat{p}^s;\theta_i,h)dF_\theta(\theta_i)$$

Pick $\lambda \in [0,1]$ such that:

$$\lambda \int_{\theta_i} q(\hat{p}^s;\theta_i,h)dF_\theta(\theta_i) + (1-\lambda)\int_{\theta_i} q(\hat{p}^s;\theta_i,l)dF_\theta(\theta_i) = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

Define $F'_{\theta,\zeta}$ such that $F'_\theta = F_\theta$ and $\zeta = h$ with probability λ, $\zeta = l$ with probability $1-\lambda$, $\theta \perp \zeta$. Then:

$$\begin{aligned}\int_{\theta_i,\zeta_i} q(\hat{p}^s;\theta_i,\zeta_i)dF_\theta(\theta_i) &= \int_{\theta_i}[\lambda q(\hat{p}^s;\theta_i,h) + (1-\lambda)q(\hat{p}^s;\theta_i,l)]dF_\theta(\theta_i) \\ &= \int_{\theta_i,\zeta_i} q(pl_i^s;\theta_i,\zeta_i)dF'_{p^s,\theta,\zeta}(\theta_i,\zeta_i)\end{aligned}$$

□

Proof of Theorem 2. From lemma 2 and rationalizability of the data:

$$\begin{aligned}&\int_{p_i^s,\theta_i}[dwl(p^b(p_i^s);\theta_i) + (p_i^s-\bar{p})q(p^b(p_i^s);\theta_i,\zeta_i)]dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\ &\qquad \geq \int_{p_i^s,\theta_i,\zeta_i}[dwl(p_i^s;\theta_i) + (p_i^s-\bar{p})q(p_i^s;\theta_i,\zeta_i)]dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\end{aligned}$$

$$\begin{aligned}&\int_{p_i^s,\theta_i}[dwl(p^b(p_i^s);\theta_i) + p_i^s q(p^b(p_i^s);\theta_i,\zeta_i)]dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\ &\qquad \geq \int_{p_i^s,\theta_i,\zeta_i}[dwl(p_i^s;\theta_i) + p_i^s q(p_i^s;\theta_i,\zeta_i)]dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\end{aligned}$$

$$\begin{aligned}&\int_{p_i^s,\theta_i}[dwl(p^b(p_i^s);\theta_i,\zeta_i) + (p_i^s-\tilde{p}^s)q(p^b(p_i^s);\theta_i,\zeta_i)]dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\ &\qquad \geq \int_{p_i^s,\theta_i,\zeta_i}[dwl(p_i^s;\theta_i) + (p_i^s-\tilde{p}^s)q(p_i^s;\theta_i,\zeta_i)]dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\end{aligned}$$

Rearranging yields:

$$\begin{aligned}\int_{p_i^s,\theta_i,\zeta_i} dwl(p^b(p_i^s);\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) &\geq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\ -\int_{p_i^s,\theta_i,\zeta_i}(p_i^s-\tilde{p}^s)q(p^b(p_i^s);\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) &+ \int_{p_i^s,\theta_i,\zeta_i}(p_i^s-\tilde{p}^s)q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\end{aligned}$$

We can show from lemma 1 and $p_i^s \in [\bar{p},\bar{p}+\bar{m}\tau] = \mathcal{P}\ \forall i$ that the term on the second line is non-negative. Formally for any $p_i^s \in (\bar{p},\bar{p}+\bar{m}\tau),\theta_i,\zeta_i,\zeta_i'$:

$$p_i^s > \tilde{p}^s \Rightarrow p^b(p_i^s) > \tilde{p}^s \Rightarrow q(p^b(p_i^s);\theta_i,\zeta_i') \leq q(p_i^s;\theta_i,\zeta_i)$$

$$p_i^s \leq \tilde{p}^s \Rightarrow p^b(p_i^s) < \tilde{p}^s \Rightarrow q(p^b(p_i^s); \theta_i, \zeta_i') \geq q(p_i^s; \theta_i, \zeta_i)$$

Either way:

$$(p_i^s - \tilde{p}^s) q(p^b(p_i^s); \theta_i, \zeta_i') \leq (p_i^s - \tilde{p}^s) q(p_i^s; \theta_i, \zeta_i)$$

Thus:

$$\begin{aligned}
\int\limits_{p_i^s, \theta_i, \zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i) &= \int\limits_{p_i^s \in int(\mathcal{P}), \theta_i, \zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i) \\
&+ \int\limits_{p_i^s \in \{\bar{p}, \bar{p}+\bar{m}\tau\}, \theta_i, \zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i) \\
&\geq \int\limits_{p_i^s \in int(\mathcal{P}), \theta_i, \zeta_i} (p_i^s - \tilde{p}^s) q(p^b(p_i^s); \theta_i, \zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i) \\
&+ \int\limits_{p_i^s \in \{\bar{p}, \bar{p}+\bar{m}\tau\}, \theta_i, \zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s; \theta_i, \zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i) \\
&= \int\limits_{p_i^s, \theta_i, \zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s; \theta_i, \zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i)
\end{aligned}$$

$$\int\limits_{p_i^s, \theta_i, \zeta_i} dwl(p^b(p_i^s); \theta_i, \zeta_i) dF_{p^s,\theta}(p_i^s, \theta_i) \geq \int\limits_{p_i^s, \theta_i, \zeta_i} dwl(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta}(p_i^s, \theta_i, \zeta_i)$$

□

We now proceed to proving theorem 3, i.e. that deadweight loss is maximized given the available data and distribution F_θ^* by having a cutoff value Δ such that the ratio of $dwl(\bar{p} + \bar{m}\tau; \theta_i, h)$ to $q(\bar{p} + \bar{m}\tau; \theta_i, l) - q(\bar{p} + \bar{m}\tau; \theta_i, h)$ is greater (less than) Δ for those we assign a perceived price of $\bar{p} + \bar{m}\tau$ (respectively $\bar{p}$). However, this reasoning overlooks the choice of the conditional distribution of ζ. Towards determining this distribution, we note that deadweight loss is bounded by the product of the reduction in demand and $\bar{m}\tau$.

Lemma 5. *If $p_i^s \in [\bar{p}, \bar{p} + \bar{m}\tau]$, then $dwl(p_i^s; \theta_i, \zeta_i) \leq [q(\bar{p}; \theta_i, \zeta_i) - q(p_i^s; \theta_i, \zeta_i)]\bar{m}\tau \ \forall \theta_i, \zeta_i$.*

Proof. Using lemma 2:

$$0 = dwl(\bar{p}; \theta_i, \zeta_i) \geq dwl(p_i^s; \theta_i, \zeta_i) - (p_i^s - \bar{p})[q(\bar{p}; \theta_i, \zeta_i) - q(p_i^s; \theta_i, \zeta_i)]$$

$$dwl(p_i^s; \theta_i, \zeta_i) \leq [q(\bar{p}; \theta_i, \zeta_i) - q(p_i^s; \theta_i, \zeta_i)](p_i^s - \bar{p}) \leq [q(\bar{p}; \theta_i, \zeta_i) - q(p_i^s; \theta_i, \zeta_i)]\bar{m}\tau$$

□

Proof of Theorem 3. The outline of the proof is as follows. First, we use lemma 5 to show that the maximal deadweight loss consistent with aggregate demand and F_θ^* comes from a data-generating process in which agents perceiving the price $\bar{p} + \bar{m}\tau$ choose the lowest quantity consistent with preference maximization, whereas the other agents choose the largest such quantity. Then, we show that distributions satisfying such a property yield deadweight loss no larger than the proposed distribution, which exists.

First, consider an arbitrary distribution $F_{p^s,\theta,\zeta}$ (yielding well-defined aggregate demand and deadweight loss) such that $F_\theta = F_\theta^*$ and:

$$F_{p^s} = \begin{cases} 0 & p_i^s < \bar{p} \\ F_{p^s}(\bar{p}) & p_i^s \in [\bar{p}, \bar{p} + \bar{m}\tau) \\ 1 & p_i^s \geq \bar{p} + \bar{m}\tau \end{cases}$$

In words, the above expression says that the support of p^s is contained in $\{\bar{p}, \bar{p} + \bar{m}\tau\}$. By theorem 2, the maximal value of deadweight loss consistent with aggregate demand and F_θ^* must satisfy this property. Consider some value $\rho \in [0, 1]$ such that:

$$\begin{aligned} &\rho \int_{\theta_i} q(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i) + [1 - \rho] \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i) \\ &= \int_{p_i^s, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i) \end{aligned} \tag{10}$$

Such a value of ρ must exist by the Intermediate Value Theorem, since by the definition of l & h and the CLD as expressed in lemma 1:

$$\begin{aligned} \int_{\theta_i} q(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i) &\leq \int_{p_i^s, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i) \\ &\leq \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i) \end{aligned}$$

In words, we are constructing an alternative distribution that rationalizes aggregate demand such that $p^s = \bar{p} + \bar{m}\tau$ & $\zeta = l$ with probability ρ, and otherwise $p^s = \bar{p}$ & $\zeta = h$. We now show that this alternate distribution yields at least as much deadweight loss, thus showing that the maximal value of deadweight loss consistent with aggregate demand and F_θ^* must arise from a distribution in which almost surely $(p^s, \zeta) = (\bar{p}, h)$ or $(p^s, \zeta) = (\bar{p} + \bar{m}\tau, l)$.

From the definition of deadweight loss:

$$\begin{aligned} &\int_{\theta_i, \zeta_i} \bar{m}\tau [q(\bar{p}; \theta_i, \zeta_i) - q(\bar{p}; \theta_i, l)] dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i) \\ &\qquad = \int_{\theta_i, \zeta_i} [dwl(\bar{p} + \bar{m}\tau; \theta_i, l) - dwl(\bar{p}; \theta_i, \zeta_i)] dF_{\theta,\zeta|p^s \neq \bar{p}} \end{aligned}$$

From here, the definition of l, and using the fact that $dwl(\bar{p}; \theta_i, \zeta_i) = 0\ \forall \theta_i, \zeta_i$, we have that $\rho \geq 1 - F_{p^s}(\bar{p})$ implies that:

$$\rho \int_{\theta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i) + (1-\rho) \int_{\theta_i} dwl(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i)$$

$$\begin{aligned}
&= \rho \int_{\theta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i) \\
&\geq [1 - F_{p^s}(\bar{p})] \int_{\theta_i, \zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i) \\
&= [1 - F_{p^s}(\bar{p})] \int_{\theta_i, \zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i) + F_{p^s}(\bar{p}) \int_{\theta_i, \zeta_i} dwl(\bar{p}; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s = \bar{p}}(\theta_i, \zeta_i) \\
&= \int_{\theta_i, \zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta}(\theta_i, \zeta_i)
\end{aligned}$$

Where the inequality follows from the fact that $\rho \geq 1 - F_{p^s}(\bar{p})$ by assumption, and the fact that $dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i)$ and the definition of l. This shows that whenever $\rho \geq 1 - F_{p^s}(\bar{p})$, the proposed alternative distribution yields at least as much deadweight loss. Now suppose instead $\rho < 1 - F_{p^s}(\bar{p})$. From lemma 5:

$$\int_{\theta_i, \zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$
$$\geq \bar{m}\tau \int_{\theta_i, \zeta_i} [q(\bar{p}; \theta_i, h) - q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i)] \bar{m}\tau dF_{\theta_i,\zeta_i|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

In addition, we find it convenient to rewrite the aggregate demand-rationalizing equation as:

$$\int_{p_i^s, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i)$$
$$= [1 - F_{p^s}(\bar{p})] \int_{\theta_i, \zeta_i} q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta|p^s \neq \bar{p}}(\theta_i, \zeta_i) + F_{p^s}(\bar{p}) \int_{\theta_i, \zeta_i} q(\bar{p}; \theta_i, \zeta_i) dF_{\theta|p^s = \bar{p}}(\theta_i, \zeta_i)$$

And so, using equation 10 and rearranging terms,

$$\rho \int_{\theta_i, \zeta_i} [q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) - q(\bar{p} + \bar{m}\tau; \theta_i, l)] dF_{\theta,\zeta}(\theta_i, \zeta_i)$$
$$\begin{aligned}
&= (1-\rho) \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i) - [1 - F_{p^s}(\bar{p}) - \rho] \int_{\theta_i, \zeta_i} q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i) \\
&\quad - F_{p^s}(\bar{p}) \int_{\theta_i, \zeta_i} q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)
\end{aligned}$$

Using lemma 5 and plugging in yields:

$$\rho \int_{\theta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,l)dF_{\theta|p^s\neq\bar{p}}(\theta_i) + (1-\rho)\int_{\theta_i} dwl(\bar{p};\theta_i,h)dF_{\theta|p^s=\bar{p}}(\theta_i)$$

$$\begin{aligned}
&= \rho \int_{\theta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,l)dF_{\theta|p^s\neq\bar{p}}(\theta_i) \\
&= \rho \int_{\theta_i,\zeta_i} \big[dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i) - q(\bar{p}+\bar{m}\tau;\theta_i,l) + q(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)\big] dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) \\
&= \rho \int_{\theta_i,\zeta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) \\
&+ \rho \int_{\theta_i,\zeta_i} [q(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]dF_{\theta,\zeta}(\theta_i,\zeta_i) \\
&= \rho \int_{\theta_i,\zeta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) + (1-\rho)\int_{\theta_i} q(\bar{p};\theta_i,h)dF_{\theta|p^s=\bar{p}}(\theta_i) \\
&- [1-F_{p^s}(\bar{p})-\rho]\int_{\theta_i,\zeta_i} q(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) \\
&- F_{p^s}(\bar{p})\int_{\theta_i,\zeta_i} q(\bar{p};\theta_i,\zeta_i)dF_{\theta,\zeta|p^s=\bar{p}}(\theta_i,\zeta_i) \\
&= \rho \int_{\theta_i,\zeta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) + F_{p^s}(\bar{p})\int_{\theta_i} q(\bar{p};\theta_i,h)dF_{\theta|p^s=\bar{p}}(\theta_i) \\
&+ [1-F_{p^s}(\bar{p})-\rho]\int_{\theta_i,\zeta_i} \big[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)\big]dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) \\
&- F_{p^s}(\bar{p})\int_{\theta_i,\zeta_i} q(\bar{p};\theta_i,\zeta_i)dF_{\theta,\zeta|p^s=\bar{p}}(\theta_i,\zeta_i) \\
&\geq \rho \int_{\theta_i,\zeta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) + F_{p^s}(\bar{p})\int_{\theta_i} q(\bar{p};\theta_i,h)dF_{\theta|p^s=\bar{p}}(\theta_i) \\
&+ [1-F_{p^s}(\bar{p})-\rho]\bar{m}\tau\int_{\theta_i,\zeta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) \\
&- F_{p^s}(\bar{p})\int_{\theta_i,\zeta_i} q(\bar{p};\theta_i,\zeta_i)dF_{\theta,\zeta|p^s=\bar{p}}(\theta_i,\zeta_i) \\
&\geq \rho \int_{\theta_i,\zeta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) \\
&+ [1-F_{p^s}(\bar{p})-\rho]\bar{m}\tau\int_{\theta_i,\zeta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) \\
&= [1-F_{p^s}(\bar{p})]\int_{\theta_i,\zeta_i} dwl(\bar{p}+\bar{m}\tau;\theta_i,\zeta_i)dF_{\theta,\zeta|p^s\neq\bar{p}}(\theta_i,\zeta_i) + F_{p^s}(\bar{p})\int_{\theta_i,\zeta_i} dwl(\bar{p};\theta_i,\zeta_i)dF_{\theta,\zeta|p^s=\bar{p}}(\theta_i,\zeta_i)
\end{aligned}$$

Thus, we know that the maximal deadweight loss consistent with aggregate demand and F_θ^* is generated by a distribution in which with probability one either $(p^s,\zeta)=(\bar{p},h)$ or $(p^s,\zeta)=(\bar{p}+\bar{m}\tau,l)$. We refer to distributions of this sort as *binary* distributions.

Now, we show that the proposed distribution maximizes deadweight loss among all binary distributions, and thus among all distributions, that rationalize aggregate demand such that $F_\theta = F_\theta^*$. Towards that end, we first show that the proposed distribution exists. Note by lemma 5 and the CLD as in lemma 1:

$$\int\limits_{\theta_i} \tilde{q}_{\bar{m}\tau,1}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i) \leq \int\limits_{p^s_i,\theta_i,\zeta_i} q(p^s_i;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p^s_i,\theta_i,\zeta_i) \leq \int\limits_{\theta_i} \tilde{q}_{0,0}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i)$$

In words, aggregate demand is contained between when all agents perceive a high price & have type h and when all agents perceive a low price & have type l. Furthermore, one can confirm that for any $\Delta, \Delta', \gamma, \gamma'$ such that $0 \leq \Delta < \Delta' \leq \bar{m}\tau$ and $0 \leq \gamma < \gamma' \leq 1$:

$$\int_{\theta_i} \tilde{q}_{\Delta,\gamma'}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i) \leq \int_{\theta_i} \tilde{q}_{\Delta,\gamma}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i)$$

$$\int_{\theta_i} \tilde{q}_{\Delta',\gamma'}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i) \geq \int_{\theta_i} \tilde{q}_{\Delta,\gamma}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i)$$

Thus, we can pick Δ such that:

$$\int\limits_{\theta_i} \tilde{q}_{\Delta,1}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i) \leq \int\limits_{p^s_i,\theta_i,\zeta_i} q(p^s_i;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p^s_i,\theta_i,\zeta_i) \leq \int\limits_{\theta_i} \tilde{q}_{\Delta,0}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i)$$

If both sides hold with equality, we can define γ arbitrarily. Otherwise, we define γ so that the market clears:

$$\gamma \equiv \frac{\int_{p^s_i,\theta_i,\zeta_i} q(p^s_i;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p^s_i,\theta_i,\zeta_i) - \int_{\theta_i} \tilde{q}_{\Delta,0}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i)}{\int_{\theta_i} \tilde{q}_{\Delta,1}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i) - \int_{\theta_i} \tilde{q}_{\Delta,0}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i)}$$

We now have the values Δ and γ such that the market clears. Suppressing Δ and γ subscripts from $\tilde{q}$, we can say that:

$$\int_{\theta_i} \tilde{q}(\theta_i) dF^*_{p^s,\theta}(p^s_i,\theta_i) = \int_{p^s_i,\theta_i,\zeta_i} q(p^s_i;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p^s_i,\theta_i,\zeta_i)$$

Finally, to show that the proposed distribution maximizes deadweight loss, consider arbitrary binary distribution $F_{p^s,\theta,\zeta}$ that rationalizes aggregate demand. Defining $\mathbb{P}_F(p^s \neq \bar{p}|\theta_i) \equiv 1 - F_{p^s|\theta=\theta_i}(\bar{p} + \bar{m}\tau)$ as the probability that $(p^s, \zeta) = (\bar{p} + \bar{m}\tau, l)$ conditional on θ_i, rationalizing aggregate demand with $F_\theta = F^*_\theta$ means that:

$$\int_{\theta_i} \left[\mathbb{P}_F(p^s \neq \bar{p}|\theta_i)\ q(\bar{p} + \bar{m}\tau; \theta_i, l) + F_{p^s|\theta=\theta_i}(\bar{p})\ q(\bar{p}; \theta_i, h)\right] dF^*_\theta(\theta_i)$$
$$= \int_{p^s_i,\theta_i,\zeta_i} q(p^s_i;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p^s_i,\theta_i,\zeta_i)$$

We can now write the difference in generated values of aggregate deadweight loss as:

$$\int_{\theta_i} \big[\frac{\tilde{q}(\theta_i) - q(\bar{p};\theta_i,h)}{q(\bar{p}+\bar{m}\tau;\theta_i,l) - q(\bar{p};\theta_i,h)} - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)\big] dwl(\bar{p}+\bar{m}\tau;\theta_i,l) dF^*_\theta(\theta_i)$$

$$= \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) > \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} [1 - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)] dwl(\bar{p}+\bar{m}\tau;\theta_i) dF^*_\theta(\theta_i)$$

$$+ \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) = \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} [\gamma - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)] dwl(\bar{p}+\bar{m}\tau;\theta_i) dF^*_\theta(\theta_i)$$

$$- \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) < \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} \mathbb{P}_F(p^s \neq \bar{p}|\theta_i) dwl(\bar{p}+\bar{m}\tau;\theta_i) dF^*_\theta(\theta_i)$$

$$\geq \Delta \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) > \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} [1 - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)]\ [q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)] dF^*_\theta(\theta_i)$$

$$+ \Delta \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) = \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} [\gamma - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)]\ [q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)] dF^*_\theta(\theta_i)$$

$$- \Delta \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) < \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} \mathbb{P}_F(p^s \neq \bar{p}|\theta_i) [q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)] dF^*_\theta(\theta_i)$$

We complete the proof by showing the right-hand side of the last inequality is zero. Since both distributions rationalize the same aggregate demand:

$$\int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) > \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} q(\bar{p}+\bar{m}\tau;\theta_i,l) dF^*_\theta(\theta_i)$$

$$+ \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) = \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} [\gamma q(\bar{p}+\bar{m}\tau;\theta_i,l) + (1-\gamma) q(\bar{p};\theta_i,h)] dF^*_\theta(\theta_i)$$

$$+ \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) < \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} q(\bar{p};\theta_i,h) dF^*_\theta(\theta_i)$$

$$= \int_{\theta_i} \big[\mathbb{P}_F(p^s \neq p^s|\theta_i)\ [q(\bar{p}+\bar{m}\tau;\theta_i,l) - q(\bar{p};\theta_i,h)] + q(\bar{p};\theta_i,h)\big] dF^*_\theta(\theta_i)$$

Subtracting both sides from $\int_{\theta_i} q(\bar{p};\theta_i,h) dF^*_\theta(\theta_i)$ yields:

$$\int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) > \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} [q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)] dF^*_\theta(\theta_i)$$

$$+ \int_{\theta_i : dwl(\bar{p}+\bar{m}\tau) = \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]} \gamma [q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)] dF^*_\theta(\theta_i)$$

$$= \int_{\theta_i} \mathbb{P}_F(p^s \neq p^s|\theta_i)\ [q(\bar{p}+\bar{m}\tau;\theta_i,l) - q(\bar{p};\theta_i,h)] dF^*_\theta(\theta_i)$$

Finally, subtracting the right-hand side from the left-hand size and multiplying by zero yields the desired result. Thus:

$$\int_{\theta_i} \big[\frac{\tilde{q}(\theta_i) - q(\bar{p};\theta_i,h)}{q(\bar{p}+\bar{m}\tau;\theta_i,l) - q(\bar{p};\theta_i,h)} - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)\big] dwl(\bar{p}+\bar{m}\tau;\theta_i,l) dF^*_\theta(\theta_i) = 0$$

In words, deadweight loss from the proposed distribution is at least as great as the deadweight loss from any binary distribution that also rationalizes aggregate demand and with the true distribution of preference types. From the first part of the proof, any distribution that rationalized aggregate demand and had the support of perceived prices contained in $\partial d\mathcal{P}$ yielded deadweight loss no greater than what one could obtain with a binary distribution that rationalized aggregate demand with $F_\theta = F_\theta^*$. Theorem 2 noted that any distribution that rationalized aggregate demand with $F_\theta = F_\theta^*$ yielded deadweight loss no greater than that one could obtain with a distribution that had the support of perceived prices contained in $\partial d\mathcal{P}$, rationalized aggregate demand, and had $F_\theta = F_\theta^*$. Therefore, any distribution that rationalizes aggregate demand and with $F_\theta = F_\theta^*$ yields deadweight loss no greater than the proposed distribution. □

Claim: If $m_{it} \perp \beta_{it}$, $\beta_{it} \leq 0$ with probability one, and m_{it} has an exponential distribution, then:

$$\hat{m} \leq 0.5 \Rightarrow \mathbb{E}[dwl_{it}|\tau] \leq -\frac{1}{2}\mathbb{E}[\tilde{\beta}_{it}]\tau^2 \ \forall \tau$$

Proof. Since the variance of an exponentially-distributed random variable is its squared expected value:

$$\mathbb{E}(m_{it}^2) = Var(m_{it}) + (\mathbb{E}[m_{it}])^2 = 2(\mathbb{E}[m_{it}])^2$$

Also, from the independence of salience and sticker price responsiveness:

$$\hat{m} = \frac{\mathbb{E}(\tilde{\beta}_{it})}{\mathbb{E}(\beta_{it})} = \frac{\mathbb{E}(m_{it})\mathbb{E}(\beta_{it})}{\mathbb{E}(\beta_{it})} = \mathbb{E}(m_{it})$$

Combining these results and again using $m_{it} \perp \beta_{it}$ yields:

$$\hat{m} \leq 0.5 \Rightarrow \mathbb{E}[dwl_{it}|\tau] = -\frac{1}{2}\mathbb{E}[m_{it}^2]\mathbb{E}[\beta_{it}]\tau^2 = -\frac{1}{2}(2(\mathbb{E}[m_{it}])^2)\mathbb{E}[\beta_{it}]\tau^2 \leq -\frac{1}{2}\mathbb{E}[m_{it}]\tau^2$$

The inequality uses the fact that $\mathbb{E}[\beta_{it}] \leq 0 \leq \mathbb{E}[m_{it}]$. □

A.3 Additional results and proofs from section 4

Proof of Theorem 4. Pick any $\lambda_1 \neq 0$ and $\lambda_2 \in \mathbb{R}$. Pick a sequence $(\bar{p}_k, \tau_k)_{k=1}^{\infty}$ contained within the support of $(\bar{p}, \tau)$ that blows up in magnitude such that $\frac{\tau_k}{\bar{p}_k} \to \frac{\lambda_2}{\lambda_1}$. Then for each $k \in \mathbb{N}$, since price and consumption are observable, one can identify the distribution of:

$$\frac{q}{\bar{p}_k} = \frac{\alpha + \epsilon}{\bar{p}_k} + \beta + \tilde{\beta}\frac{\tau_k}{\bar{p}_k}$$

Taking the limit as $k \to \infty$ and multiplying both sides by λ_1 means that we can identify:

$$\lambda_1 \lim_{k\to\infty} \frac{q}{\bar{p}_k} = \lambda_1[\lim_{k\to\infty} \frac{\alpha + \epsilon}{\bar{p}_k} + \beta + \tilde{\beta}\frac{\tau_k}{\bar{p}_k}]$$

Since $\bar{p}_k$ blows up in magnitude as $k \to \infty$:

$$\lambda_1 \lim_{k\to\infty} \frac{q}{\bar{p}_k} = \lambda_1\beta + \lambda_2\tilde{\beta}$$

Note that we can alternatively pick a sequence $(\bar{p}_k, \tau_k)_{k=1}^{\infty}$ contained within the support of $(\bar{p}, \tau)$ so that $\bar{p}_k$ remains bounded while τ_k blows up. In this case, we can identify:

$$\lim_{k\to\infty} \frac{q}{\tau_k} = \lim_{k\to\infty} \frac{\alpha + \epsilon}{\tau_k} + \beta\frac{\bar{p}_k}{\tau_k} + \tilde{\beta} = \tilde{\beta}$$

Multiplying both sides by arbitrary $\lambda_2 \in \mathbb{R}$ yields:

$$\lambda_2 \lim_{k \to \infty} \frac{q}{\tau_k} = \lambda_2 \tilde{\beta}$$

Thus, we can identify the distribution of $\lambda_1 \beta + \lambda_2 \tilde{\beta}$ for any $(\lambda_1, \lambda_2) \in \mathbb{R}^2$. By the Cramèr-Wold theorem, we can identify the joint distribution of $(\beta, \tilde{\beta})$, and so can identify $\mathbb{E}[dwl_i] = -\frac{1}{2}\mathbb{E}[\frac{\tilde{\beta}^2}{\beta}]\tau^2$ for any τ. □

B Application of Linear Model

We apply the linear specification of section 3.3 to data gathered by CLK (2009) on the aggregate consumption of beer in U.S. states between 1970 and 2003. First, we translate their model (in logs) to our linear specification. Second, we estimate the same equation of interest under different sets of controls:

$$y_{st} = \alpha + \beta \tau_{st}^{e} + \tilde{\beta} \tau_{st}^{s} + \gamma X_{st} + \epsilon_{st}$$

For each linear specification, we compute $\hat{m} = \frac{\tilde{\beta}}{\beta}$, which gives us the ratio of upper bound of deadweight loss to lower bound of deadweight loss (assuming that maximal attention, $\bar{m} = 1$). Results are presented in table 3. We also estimate a number of other specifications, again following CLK (2009), presented in table 4. These are meant to address concerns for spurious results – in particular, it could be the case that consumers react differently to the two tax rates as while sales taxes affect a variety of goods, excise taxes on beer affect only beer prices. In particular, the second last column of table 4 show estimates for a regression only for those states that exempt food (a likely substitute of beer) from sales tax, demonstrating that even in this restricted sample beer consumption is quite insensitive to sales tax. Finally, the last column addresses the potential concern that people might be substituting toward other alcoholic beverages when they face a beer tax increase, and not when they face a sales tax increase. As we can see, the share of ethanol people consume in the form of beer is quite insensitive to either tax rate.

We repeat the exercise for Goldin and Homonoff (2013), who have a similar set-up with individual-level, cross-sectional data on cigarette consumption. The ratio of average responsiveness to sales taxes relative to the average responsiveness to excise taxes of about zero, although the estimate quite uncertain (see table 5 for details). Rosen (1976) uses a linear model, so we directly use his estimates from table 1 of his paper. Results are reported in table 6.

	Baseline	Business cycle	Alcohol regulations	Region trends
Δ(excise tax)	-.08 (.03)	-.11 (.03)	-.1 (.03)	-.07 (.03)
Δ(sales tax)	-.05 (0.07)	-.02 (.07)	-.02 (.07)	-.03 (.07)
Δ(population)	-.0002 (0.0002)	-.0002 (.0002)	-.0001 (.0002)	-.0002 (.0002)
Δ(income per cap.)		.0002 (.00006)	.0002 (.00006)	.0002 (.00006)
Δ(unemployment)		-.09 (.03)	-.1 (.03)	-.09 (.03)
Alcohol reg. controls			X	X
Year FE	X	X	X	X
Region FE				X
$\hat{m}$	.67 (.87)	.2 (.62)	.23 (.65)	.46 (.99)
Sample size	1,607	1,487	1,487	1,487

Table 3: Estimating $\hat{m}$ with several sets of controls, following the specifications in Chetty et al. (2009) in the context of a linear model.

	Policy IV for excise tax	3-Year differences	Food exempt	Dep. var.: share of ethanol from beer
Δ(excise tax)	-.12 (.06)	-.24 (.1)	-.1 (.04)	.0003 (.0005)
Δ(sales tax)	-.02 (.07)	-.03 (.07)	-.05 (.07)	.001 (.001)
Δ(population)	-.0001 (0.0002)	-.002 (.0015)	-.00002 (.0002)	-.0000 (.0000)
Δ(income per cap.)	.0001 (.00006)	.0002 (.00007)	.0002 (.00007)	.0000 (.0000)
Δ(unemployment)	-.09 (.03)	-.03 (.03)	-.06 (.03)	-.0001 (.0004)
Alcohol reg. controls Year FE	X X	X X	X X	X X
$\hat{m}$	.17 (.54)	.11 (.3)	.52 (.68)	
Sample size	1,487	1,389	937	1,487

Table 4: Estimating $\hat{m}$ following the strategy of CLK (2009) in the context of a linear model. As in CLK, we use nominal excise tax rate divided by the average price of a case of beer from 1970 to 2003 as an IV for excise tax to eliminate tax-rate variation coming from inflation erosion. Next, we run the same regression in 3-year differences. Next, we run it only for states where food is exempt from sales-tax, to address concerns about whether consumers react differently to changes in the two taxes only because sales taxes apply to a broad set of goods. Finally, the last column addresses the concern that beer taxes may induce substition with other alcoholic products, biasing the coefficient on excise tax relative to the one on sales tax. It shows that beer excise taxes have no discernable effect on the share of ethanol consumed from beer.

	Specification 1	2	3
Excise Tax	-4.85	-4.73	-4.87
	(.92)	(.92)	(.94)
Sales Tax	-.72	.42	.32
	(4.89)	(5.25)	(5.26)
Demographic controls	X	X	X
Econ. conditions controls		X	X
Income trend controls			X
State,year, and month FE	X	X	X
$\hat{m}$	.15	-.09	-.07
	(1.01)	(1.11)	(1.10)
Sample size	274,137	274,137	274,137

Table 5: Estimating $\hat{m}$ based on the intensive response of cigarette consumption to sales taxes (not included in sticker price) and excise taxes (included in the sticker price). The specifications are a linearized version of the specifications in Goldin and Homonoff (2013).

	Hours/Year MTR at zero hours	Hours/Year MTR at full time	Hours/Week MTR at zero hours	Hours/Week MTR at full time
Wage	990.4	1218.1	18.8	21.96
	(74.88)	(106.3)	(1.75)	(2.48)
MTR*Wage	-950.7	-1480.7	-21.1	-26.58
	(269.7)	(324.5)	(6.29)	(7.58)
$\hat{m}$	0.96	1.21	1.12	1.21
	(.217)	(.175)	(.257)	(.227)
Sample size	2,545	2,545	2,545	2,545

Table 6: We report regression results directly from Table 1 of Rosen (1976).