Recovering the Policy Maker's Unobservable Tax Preferences: Theory and Applications*

Seungjin Han† Josip Lesica‡

November 17, 2023

Abstract

This paper proposes methods that allow us to identify the policymaker's unobserved preferences given the choices of different tax rates on differentiated products. We apply the method to understand the excise tax policy adopted by the Canadian province of Ontario on beer products, which not only violates the inverse elasticity rule, but is inconsistent with revenue maximization. Using unique retail sales data on beer products at the SKU level and neighborhood socio-economic characteristics from Ontario, we identify the government's social welfare weights over the total surplus and net externalities associated with the production and consumption of local craft beer vs. large brewers' beer. According to our estimates, the government believes that micro beer brings more positive externality by \$1.41 per liter on average, which justifies the reduced excise tax on local micro beer. Our methods can be applied to various policy problems such as a government's choice of different excise tax rates for differentiated tobacco products based on its nicotine content if they were to introduce it, a government's decision on merger, etc. *JEL classifications*: D12, D62, H21, H23.

*We would like to thank Benjamin Dachis, Youngki Shin, Mike Veall and Ian Irvine for their comments and suggestions. We thank seminar participants at the 56th annual conferences of the Canadian Economics Association, Journées Louis-André Gérard-Varet 2022 and Canadian Public Economics Group 2022 conference for their comments and suggestions. Han gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada. All views expressed herein are solely those of the authors and do not reflect policy opinions or positions of the Statistics Canada in any way.

†Department of Economics, McMaster University. Email: hansj@mcmaster.ca

‡Economic Analysis Division, Statistics Canada. Email: josip.lesica@gmail.com.

Contents

1 Introduction

It is important to citizens and interest groups to understand the social welfare functions employed by governments when trying to influence them to establish policies and on making decisions on whether they agree with Decision Maker (DM)'s policy choice or they want to change the DM through an election. This paper demonstrates how reduced-form methods can be used for estimating unobservable parameters in the social welfare function using a very unique dataset on beer sales in the Canadian province of Ontario from 2005-2015.

Most governments impose a tax on alcoholic drinks and some apply different excise tax rates on different categories of products. A notable example is the two-tier excise taxes imposed on beer products differentiating large and small (i.e. micro) producers in many countries in Europe and North America. Several EU countries have implemented reduced rates for small and independent breweries as opposed to large ones. In the United States, the 2017 Tax Cuts and Jobs Act provided a reduction in federal excise taxes for all brewers, with particularly lower rates for small domestic brewers.[1]

Ontario adopted a policy of reduced excise tax rates for small, independent (craft) breweries. However, because we find that manufacturers' beer is more elastic with respect to its own price than micro beer is, this violates the inverse elasticity rule. Based on the inverse elasticity rule, it would make sense to tax micro beer higher. The observed tax rates are also inconsistent with the revenue maximizing ones in that the observed tax rates are consistently higher than the revenue maximizing ones.[2]

To understand such a taxation policy that cannot be explained by the inverse elasticity rule and revenue maximization, it is important to identify the DM's social welfare function. We assume that the DM (e.g., government, legislature, etc.) has an underlying preference over the total surplus and the net externality of the production and consumption of each type of beer, represented by the social welfare function (SWF) as the weighted sum of the total surplus and the negative value of the equilibrium quantity of each type of beer. These weights are the DM's private information and she maximizes her objective function by setting the two differentiated beer excises tax rates. We are interested in not just why taxes are different across beer products, but why they are set differently by the government.

To our best knowledge, our paper is the first to offer a clear articulation of how to

[1] For example, the rates are \$3.50 per barrel on the first 60,000 barrels for domestic brewers producing fewer than two million barrels annually as opposed to \$18 per barrel rate for producing over six million barrels. The U.S. Congress made these rates permanent at the end of 2020.

[2] See Section 3.2 for details.

identify the DM's social welfare function. Our methodology is sufficiently general to be extended or applied to various policy problems. Furthermore, our methodology would inform governments on the consequences of what has been done in the past with respect to implicit weights on different stakeholder interests, which could then inform future policy decisions. We believe that this is a really neat demonstration of what can be done with relatively simple methods, which most policy makers are candidly, unaware of, and would be an invaluable addition to their toolkit.

1.1 Methods

The method proposed in our theory section guides us how to retrieve unobservable weights in the SWF step by step. In our model, producers sell their differentiated products in markets. The equilibrium price for a product correctly reflects the excise tax imposed on the producer's beer product and the sum of the producer's marginal cost and the mark-up. It is also equal to the consumer's maximum willingness to pay, which can be estimated by the (inverse) demand function. Since the (retail) prices and excise taxes are observed in practice, we can derive the sum of the producer's marginal cost and the mark-up. Replacing the price with the corresponding excise rate and the sum of the marginal cost and the mark-up for each product yields the system of equations. Solving this system yields the equilibrium quantity of each type of product that depends not only on its own excise tax, but also on the excise tax on the other type of products. Foreseeing the impact of excise taxes on the equilibrium quantities, the DM sets the optimal excise tax rate for each type of beer products that maximizes the social welfare.

Each optimal excise tax rate is a function of parameters from the two demand functions, the sum of the marginal cost and the mark-up, and the weights in the DM's social welfare function. With the optimal excise tax rates assumed to be the ones observed in practice, the only unknowns in the solutions for optimal excise taxes are the weights in the DM's social welfare function that consist of the weight sum of consumer surplus, producer surplus, tax revenue and net externalities from beer products. However, the identifiability of the weights in the DM's social welfare function crucially depends on the functional form of the two demand functions.

The purpose of our studies is not to contribute to the literature on alcohol product demand estimation. Rather, the focus is to illustrate how to retrieve unobservable parameters (weights) in the social welfare function using observed behavior of consumers and the policy maker. For this reason, we parsimoniously choose linear demand functions, which allows for solving a system of linear equations for the weights in the DM's social welfare function.

1.2 Application to excise taxation in the beer market

Importantly, we do not impose a sign restriction on the weight of the equilibrium quantity of each beer in the DM's SWF as it reflects the DM's evaluation of the net externality of the production and consumption of each type of beer. This leaves the possibility that the DM thinks a type of beer may induce a positive net externality. For example, craft beer may generate the same negative externality on consumers as the big manufacturers beer products do (e.g., health cost, drunk driving, lost productivity, etc.). However, there could be positive societal effects of craft beer producers (i.e. 'externalities') associated with the supply side such as (i) the protection of the nascent local micro brewing industry, given that small business owners are the main suppliers of micro beer; (ii) the positive impact on the local economy such as local employment and supply of ingredients for beer manufacturing, expansion of local tourism; (iii) improved environmental effects. We are able to examine which of these two externalities the DM thinks is greater by identifying the weight on the net externality. We apply our theory to the excise beer tax policy in the Canadian province of Ontario that enacted dual beer excise tax rates per liter of beer produced in 2010, one for micro brewers and another for big producers.

We start with a simple estimation of linear demand function coefficients for micro and non-micro beer using a unique data set obtained under a research contract from the Liquor Control Board of Ontario (LCBO): weekly prices and quantities for all beer products sold in Ontario during 2005-2015 period. LCBO is a government corporation that retails and distributes alcoholic beverages throughout the Canadian province of Ontario. It maintains a quasi-monopoly on the retail sale of beer, with The Beer Store being the only other retailer during the span of our data. Furthermore, we link each store's location to its respective area's household socio-economic characteristics from a custom tabulation for households' units and average income created by Statistics Canada using the Canadian personal income tax returns, the T1 Family File. The estimation of the demand functions indicates that craft beer and big beer products are substitutes.

Next, using the estimated demand parameters we derive the elasitictes of each type of beer products. Non-micro beer is getting slightly more elastic with respect to its own price over the years, whereas micro beer is getting more inelastic. In later years, manufacturers' beer is more elastic with respect to its own price than micro beer is. Based on the inverse elasticity rule, this would tell us the government should tax micro beer higher in later years and hence the inverse elasticity rule is violated.

With the estimated demand parameters, we also derive the annual revenue maximizing excise tax rates during 2010 - 2015 that the DM would choose if they were to care only about tax revenue (See Table 6). Our estimated revenue maximizing excise tax rates follow the same pattern as the actually observed rates: excise tax rate for micro beer is consistently lower than that for non-micro beer in every year. However, it is notable that the revenue maximizing excise tax rates for each type of beer are consistently lower than the corresponding observed excise tax rate in every year.

Therefore we explore the gap between the observed micro beer and non-micro beer tax rates by identifying the social welfare weights. Table 7 shows that the SWF weights are very stable during 2010 - 2015. The total surplus maintains the weight in the range of 0.5906 to 0.5570 during this period, which indicates the government places a slightly more weight on the total surplus than externalities. The weight on the negative value of the equilibrium quantity of micro beer ranges from -0.1976 to -0.1860 and it is increasing at a very low rate, whereas the weight on the negative value of the equilibrium quantity of non-micro beer ranges from 0.6009 to 0.6290 and it is increasing at a very low rate. Importantly, the negative weight on the negative value of the equilibrium quantity of micro beer implies that the DM believes that the *positive* effects on the supply side of micro beer outweighs the *negative* externality on the demand side of micro beer. On the other hand, the positive weight on the negative value of equilibrium quantity of non-micro beer implies that the DM believes the negative externality outweighs the positive effects.

Using identified weights we can derive the marginal rate of substitution (MRS) for each type of beer, which specifies the amount by which total surplus would decrease for a change in the net externality due to a marginal decrease in the equilibrium quantity of each type of beer products, holding social welfare constant. During 2010-2015, the MRS of micro beer is in the range of -0.3243 to -0.3404 and it steadily increased, except for 2010-2011 and 2011-2012. On the other hand, the MRS of non-micro beer is in the range of 1.0580 to 1.1293 and it also steadily increased.

A negative MRS of micro beer implies that in order to hold a constant social welfare the DM should increase the total surplus by \$0.33 on average during 2010-2015 following a decrease in the positive net externality due to one litre decrease in the equilibrium quantity of micro beer. On the other hand, a positive MRS of non-micro beer means that, in order to keep social welfare constant, total surplus would have to be reduced by \$1.08 on average following a decrease in the negative net externality due to a one litre decrease in the equilibrium quantity of non-micro beer.

In other words, in terms of the DM's point of view, social welfare *increases* by \$0.33 on average due to the positive net externality associated with a 1L increase in the equilibrium quantity of micro beer, whereas it *decreases* by \$1.08 on average due to the negative net externality associated with a 1L increase in the equilibrium quantity of non-micro beer. If we assume that both types of beer create the same amount of the negative externality per litre, this implies that the DM believes that mirco beer creates more positive externality by \$1.41 per liter.

Why is there a high discrepancy of MRSs between two types of beers with the opposite signs? In other words, what is the source of the DM's high subjective cost of reducing the equilibrium quantity of micro beer through taxation? During 2010-2015, the average excise tax on micro beer (23.1 cents per litre) is only a third of that on non-micro beer (73.1 cents per litre). However, the straight comparison of elasticities does not provide the whole story because the two markets differ by size; the market for micro beer is only 9% of the market for non-micro beer on average during 2010-2015.

A correct measure is one that divides the elasticity by its own excise tax rate. This measures the percentage decrease in the equilibrium quantity of each beer type caused by a one-dollar increase in its own tax. Over the 2010-2015 period, a one-dollar increase in the micro beer tax decreases its equilibrium quantity by 14.5%, whereas the same increase in the non-micro beer tax decreases its equilibrium quantity by 4.1% only. Even though the price of micro beer is higher than that of non-micro beer, a one-dollar increase in tax on micro beer decreases its equilibrium quantity *3.6 times more* in terms of the percentage point than the same amount of tax increase on non-micro beer shrinks its equilibrium quantity.

Our calculation shows that during 2010-215, on average, a 7.9 cent increase in the micro beer tax reduces its own equilibrium quantity by one percent, whereas a 24.2 cent increase in the non-micro beer tax is needed to reduce its own equilibrium quantity by one percent. Therefore, the same percentage decrease in the equilibrium quantity of micro beer would be observed only when the excise tax on micro beer is 32.6% of the excise tax on non-micro beer. This magnitude is aligned with the observed taxes: During 2010-2015, the average excise tax on micro beer is 32% of the one on non-micro beer (23.1 cents vs. 72.3 cents).

Given that the positive externality associated with the production of micro beer outweighs the negative externality associated with its consumption, it is not optimal for the government to impose the same amount of tax on each type. The same amount of tax increase shrinks the market for micro beer disproportionately more, 3.6 times more. This may be one of the reasons why the DM imposed a lower tax on micro beer.

1.3 Related literature

Our theory is related to the inverse optimum problem in Ahmad and Stern [1984]. It essentially shows how to calculate non-negative weights on all the households in the DM's social welfare function. If no such welfare weights exist, then a marginal Pareto improvement is possible in the economy. This approach does not consider how to deal with externalities that are not explicitly considered in the model. Our approach incorporates the possibilities of externalities associated with the consumption and production of goods in that the DM's social welfare function is the weighted sum of the total surplus and the net externalities from different products. According to the criterion for no Pareto improvement from Ahmad and Stern [1984], the weight on the total surplus must be non-negative, which is confirmed in our application to excise taxation in the beer market.

Our application is related to the literature on the design and incidence of alcohol taxes. In recent work Griffith et al. [2019] study the optimal corrective tax design in the alcohol market, where the government's only motivation is to correct the negative externality associated with the consumption of different alcohol products (beer, wine, liquor). Their normative result is that the optimal corrective taxation involves differentiating tax rates across alcohol products according to varying levels of ethanol. This policy consideration has received particular attention from the public health perspective.[3]

Consistent with Griffith et al. [2019] we set up a social welfare function consisting of total surplus and the externalities of alcohol consumption. However, in our approach, we assume that the DM's social welfare function is not observable because the social welfare is a weighted sum of the total surplus and the *net* externalities of different products, where the weights are assumed to be private information. We take the revealed preference approach to identity these weights that are consistent with the observed tax rates. Further, while Griffith et al. [2019] fix the DM's social welfare function to study the optimal corrective tax design that deals only with the negative externality, we allow for the presence of both positive and negative effects associated with the consumption and production of alcohol

[3]For example, the possibility of differentially taxing tobacco products based on its nicotine content to maximize incentives for tobacco users to switch from the most harmful products to the least harmful ones are discussed in Chaloupka et al. [2019, 2015].

products. The identified weights reflect the DM's valuation of the *net* externality associated with both the production and consumption of each type of beer product.

The revenue and redistributive concerns in our paper are in line with Miravete et al. [2018, 2020] and Conlon and Rao [2020]. They do not consider beer excise taxation, but are particular concerned with the relationship between excise taxes and consumer welfare in the distilled spirits market. We find that the revenue maximizing excise taxes are below the set ones, which is consistent with Miravete et al. [2018].

Another reason why it is important to study the setting of differential beer taxes is the pass-through of taxes to retail prices. In a perfectly competitive market taxes are fully passed through, but the literature often finds that pass-through rates exceed unity. See Kenkel [2005] and Shrestha and Markowitz [2016] for the U.S. evidence. Pass-through of excise taxes has also received attention in the context of multi-product markets, which is more akin to the beer market studied in this paper. Hamilton [2009] theoretically examines the pass-through of excise taxes for multi-product firms while Friberg and Romahn [2018] analyzes how pass-through depends on the number of products that firms control in the Swedish beer market.

1.4 Other applications

While the identification of social welfare weights is based on two excise taxes on two differentiated products, it can be extended to an arbitrary number of different products. It would be very interesting and relevant to identify social welfare weights based on excise taxes on all types of alcohol products.

Differentiated excise taxation has received particular attention from the public health perspective. For example, the possibility of differentially taxing tobacco products based on its nicotine content to maximize incentives for tobacco users to switch from the most harmful products to the least harmful ones are discussed in Chaloupka et al. [2019, 2015]. One can use our method to analyze the government's choice of optimal excise tax rates on differentiated tabacco products if they were to introduce it for differentiated tabacco products.

Total surplus or its variations are commonly used as a measure for social welfare. For example, competition policy authorities around the world use variations of total surplus in recommending or stopping mergers, which include implicit weights on consumer and producer surplus. In Canada, there is often heated debate on whether the Competition

Bureau of Canada needed to put equal weights on consumer and producer surplus to determine the welfare effect of a merger or could choose to assign different weight (e.g., the Superior Propane and ICG vs. Comissioner of Compeition case in Canada).[4] Given the DM's decision on a merger case, we can derive the lower bound of the weight on producer surplus when the merger is approved.

2 Theory

It is important to citizens and interest groups to understand the DM's social welfare function when trying to influence the DM to set up the policy on excise tax in their favor. It is also important when they make a decision on whether they agree with DM's policy choice or want to change the DM through an election. In this section we show how to identify a DM's parameters in the social welfare function.

The DM may choose multiple excise taxes, one for each good when goods are able to be differentiated, such as alcoholic beverages. The notable examples are the two-tier excise taxes imposed on beer products in many jurisdictions in Europe and North America. The DM's preferences over the total surplus and the externality associated with the production and consumption of beer products can be captured by her social welfare function.

The DM may choose multiple excise taxes, one for each good when goods are able to be differentiated, such as alcoholic beverages. The notable examples are the two-tier excise taxes imposed on beer products in many jurisdictions in Europe and North America. The DM's preferences over the total surplus and the externality associated with the production and consumption of beer products can be captured by her social welfare function.

There are two types of beer products in the market: beer produced by multiple local microbrewers and beer produced by multiple big manufacturers. Producers sell their differentiated products in markets. Let p_1 and p_2 denote the unit prices of products 1 (local craft beer) and 2 (large manufacturers' beer) respectively. Let t_1 and t_2 denote the excise taxes imposed on product 1 and product 2 respectively. Let c_1 and c_2 denote the marginal costs of product 1 and product 2 respectively. Let m_1 and m_2 denote the industry mark-up of product 1 and product 2 respectively. Let $b_i = c_i + m_i$ be the sum of the marginal cost and the mark-up.

[4] Canada has a small open economy with most industries located in very close proximity to much larger competitors in the U.S. This underlies Canada's traditional tolerance for higher concentration levels in many leading industries.

In terms of market structure, the dominant firm competitive fringe model may explain well the markets observed in practice. When the dominant firm adopts dynamic limit pricing, they foresee that the size of the fringe is expanded as entry continues. As the expansion of the size of the fringe causes a flatter residual demand curve, the dominant firm's price will fall until it is close to the limit price (i.e., the marginal cost). Alternatively, we can assume that produces sell their products in oligopoly markets operating *a la* Bertrand (i.e., price competition). In either market structure, the equilibrium mark-up would not be sufficiently high to make the producer's profit significantly larger than the opportunity cost.[5]

The equilibrium quantities are determined at the point where consumers' marginal willingness to pay for each product is equal to the price chosen by producers. Consumers' maximum willingness to pay is characterized by the *inverse demand function*. Let q_1 and q_2 denote the (market) quantities of products 1 and 2 respectively. The inverse demand functions for product i $(i = 1, 2)$ is given by

$$p_i = f_i(q_1, q_2, \psi_i), \tag{1}$$

where ψ_i is a vector of parameters. In equilibrium, the prices are $p_1 = b_1 + t_1$ and $p_2 = b_2 + t_2$. Substituting $b_i + t_i$ for p_i $(i = 1, 2)$ in (1) yields

$$b_1 + t_1 = f_1(q_1, q_2, \psi_1), \tag{2}$$

$$b_2 + t_2 = f_2(q_1, q_2, \psi_2). \tag{3}$$

Solving (2) and (3) for q_1 and q_2 yields the equilibrium quantities for good 1 and good 2: $q_1^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2)$ and $q_2^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2)$.

The DM cares about not only the total surplus but also the (net) externality associated with the production and consumption of each type of products. For example, a micro beer product can create two types of externalities. First, a negative externality associated with overconsumption such as health costs, impact on crime, lost productivity, etc. On the other hand, it can also create a positive 'externality' on the supply side such as growth of small businesses in local micro brewing industry, positive impact on the local supply of ingredients, employment, tourism, environmental benefits of small scale farming and sustainable water use, etc.

[5]The anecdotal evidence is that the craft beer industry in Canada is highly competitive, with the number of them doubling in the five years before the pandemic, and very low profitability. This would suggest a higher degree of competition in the craft beer industry. See Kirby and Lundy [2022] for example.

Although these societal benefits of successful craft beer producers are not an externality in the traditional sense, we interpret them as general equilibrium effects not fully captured in our partial equilibrium model. How researchers model the externalities in the social welfare function matters for the result, and we chose to not focus only on the traditional negative effects of alcohol consumption because a government might have reasons to promote the local beer industry.

If the DM only cares about the tax revenue, she will choose excise taxes that maximize

$$R(t_1, t_2, b_1, b_2, \psi_1, \psi_2) = t_1 q_1^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2) + t_2 q_2^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2).$$

Let t_1^* and t_2^* be the revenue maximizing excise taxes. However, the DM may consider together (i) the total surplus and (ii) the net externalities associated with the production and consumption of products that are not internalized in the market. The DM's social welfare function is then given by

$$S = \omega_0 TS(t_1, t_2, b_1, b_2, \psi_1, \psi_2) - \omega_1 q_1^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2) - \omega_2 q_2^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2),$$

where $TS(t_1, t_2, b_1, b_2, \psi_1, \psi_2)$ denote the total surplus and $\omega_0 + \omega_1 + \omega_2 = 1$. The weights $(\omega_0, \omega_1, \omega_2)$ are the DM's private information and not observable by the public. We assume that $\omega_0 > 0$. However we do not impose the sign of ω_1 and ω_2. If $\omega_i > 0$. for $i \in \{1, 2\}$, the DM thinks that the net externalities of product i is negative. If it is negative, the DM thinks that the net externalities of product i is positive. Note that we can standardize the social welfare function in dollar terms by simply dividing the weights by ω_0, and interpreting the net-externality weights as $\frac{\omega_1}{\omega_0}$ and $\frac{\omega_2}{\omega_0}$. Because total surplus is measured in dollar terms, $\frac{\omega_i}{\omega_0}$ measures the reduction in the social welfare in dollar terms due to the (net) externalities created by the consumption of one unit of product i. In fact, $\frac{\omega_i}{\omega_0}$ is the DM's marginal rate of substitution of the net externalities created by one unit of product i for total surplus.

In either market structure, the producer surplus net of the opportunity cost is not significant. The consumer surplus $CS(t_1, t_2, b_1, b_2, \psi_1, \psi_2)$ is the sum of two consumer surpluses in two markets

$$CS(t_1, t_2, b_1, b_2, \psi_1, \psi_2) = CS_1(t_1, t_2, b_1, b_2, \psi_1, \psi_2) + CS_2(t_1, t_2, b_1, b_2, \psi_1, \psi_2),$$

where, for $i = 1, 2$ and $j \neq i$

$$CS_i(t_1, t_2, b_1, b_2, \psi_1, \psi_2) = \int_0^{q_i^*(t_1,t_2,\psi_1,\psi_2)} f_i(q_i, q_j^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2), \psi_i) dq_i - (b_i + t_i)\, q_i^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2).$$

Therefore, we have that $TS(t_1, t_2, \psi_1, \psi_2) = CS(t_1, t_2, \psi_1, \psi_2) + R(t_1, t_2, \psi_1, \psi_2)$. Then, the DM's social welfare function becomes

$$S = \omega_0 \left[CS(t_1, t_2, b_1, b_2, \psi_1, \psi_2) + R(t_1, t_2, b_1, b_2, \psi_1, \psi_2) \right] - \omega_1 q_1^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2) - \omega_2 q_2^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2). \quad (4)$$

The DM's optimal taxes maximize her social welfare S given $(\omega_0, \omega_1, \omega_2)$ and (ψ_1, ψ_2). Because $\omega_0 = 1 - \omega_1 - \omega_2$, the optimal taxes depend on $(\omega_1, \omega_2, b_1, b_2, \psi_1, \psi_2)$ and let them denoted by $t_1^\circ(\omega_1, \omega_2, b_1, b_2, \psi_1, \psi_2)$ and $t_2^\circ(\omega_1, \omega_2, b_1, b_2, \psi_1, \psi_2)$. We assume that the observed taxes, t_1 and t_2, are the same as the optimal taxes, $t_1^\circ(\omega_1, \omega_2, b_1, b_2, \psi_1, \psi_2)$ and $t_2^\circ(\omega_1, \omega_2, b_1, b_2, \psi_1, \psi_2)$:

$$t_1 = t_1^\circ(\omega_1, \omega_2, b_1, b_2, \psi_1, \psi_2) \quad (5)$$

$$t_2 = t_2^\circ(\omega_1, \omega_2, b_1, b_2, \psi_1, \psi_2) \quad (6)$$

ψ_1 and ψ_2 are estimated from the demand estimation. b_1 and b_2 are retrievable from $p_1 = b_1 + t_1$ and $p_2 = b_2 + t_2$ because prices and taxes are observable. Therefore, only unknowns in (5) and (6) are ω_1 and ω_2. Given all the other parameters, ω_1 and ω_2 are the solutions for the system of equations (5) and (6).

While one may conduct some comparative statics by imposing reasonable properties in the demand functions f_1 and f_2, the idenfitication of ω_1 and ω_2 requires the restrictions of the demand functions that leads to a closed-form solutions for t_1° and t_2° and for ω_1 and ω_2. This is a challenging task even with commonly used log-linear demand functions.

The primary purpose of our studies is not to contribute to the literature on alcohol products demand estimation. Rather, the focus is to show how to retrieve unobservable parameters (weights) in the social welfare function using observed behavior of consumers and policymaker. To makes this possible, a given simple functional form of the demand functions is necessary to derive closed form solutions. This allows us to analyze the motivation of the observed tax policy given the identified social welfare function.

We choose linear demand functions[6]:

$$f_1(q_1, q_2, \psi_1) = \alpha_1 + \beta_1 q_1 + \gamma_1 q_2, \quad (7)$$

$$f_2(q_1, q_2, \psi_2) = \alpha_2 + \gamma_2 q_1 + \beta_2 q_2, \quad (8)$$

where $\psi_1 = [\alpha_1, \beta_1, \gamma_1]$, $\psi_2 = [\alpha_2, \beta_2, \gamma_2]$, $\alpha_1, \alpha_2 > 0$ and $\beta_1, \beta_2 < 0$. If γ_1 and γ_2 are positive (negative), the two goods are substitutes (complements). Assuming that the consumption pattern in the unobserved part of the market follows a similar pattern to that in the observed data, we think of (7) and (8) as individual demand functions or market demand functions. This makes the empirical analysis tractable in the case where the data on complete market quantities are not observable.

Given the linear demand functions, the equilibrium quanities are

$$q_1^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2) = X + Yt_1 - Zt_2, \quad (9)$$

$$q_2^*(t_1, t_2, b_1, b_2, \psi_1, \psi_2) = x + yt_2 - zt_1, \quad (10)$$

where

$$X = \frac{-\beta_2(\alpha_1 - b_1) + \gamma_1(\alpha_2 - b_2)}{\beta_1\beta_2 - \gamma_1\gamma_2}, \quad x = \frac{-\beta_1(\alpha_2 - b_2) + \gamma_2(\alpha_1 - b_1)}{\beta_1\beta_2 - \gamma_1\gamma_2},$$
$$Y = \frac{\beta_2}{\beta_1\beta_2 - \gamma_1\gamma_2}, \quad y = \frac{\beta_1}{\beta_1\beta_2 - \gamma_1\gamma_2},$$
$$Z = \frac{\gamma_1}{\beta_1\beta_2 - \gamma_1\gamma_2}, \quad z = \frac{\gamma_2}{\beta_1\beta_2 - \gamma_1\gamma_2}.$$

The revenue maximizing taxes are

$$t_1^* = \frac{2\beta_1\beta_2 - \gamma_1\gamma_2 - \gamma_2^2}{4\beta_1\beta_2 - (\gamma_1 + \gamma_2)^2}(\alpha_1 - b_1) + \frac{\beta_1(\gamma_2 - \gamma_1)}{4\beta_1\beta_2 - (\gamma_1 + \gamma_2)^2}(\alpha_2 - b_2), \quad (11)$$

$$t_2^* = \frac{2\beta_1\beta_2 - \gamma_2\gamma_1 - \gamma_1^2}{4\beta_1\beta_2 - (\gamma_1 + \gamma_2)^2}(\alpha_2 - b_2) + \frac{\beta_2(\gamma_1 - \gamma_2)}{4\beta_1\beta_2 - (\gamma_1 + \gamma_2)^2}(\alpha_1 - b_1). \quad (12)$$

Taking the partial derivatives of S with respect to t_1 and t_2 and setting them to zero respectively we obtain first-order conditions. Solving the first-order conditions for t_1 and t_2 simultaneously yields the optimal taxes t_1° and t_2° as follows.

[6] It is interesting to derive the less restrictive properties on the demand function that ensure the identification of ω_1 and ω_2, but it is beyond the scope of this paper and we leave it for future research. For robustness analysis, we compare the elasticities derived from log-linear demand functions with those derived from our linear demand functions and we also do a simple IV estimation with the available instrument. See page 22.

Proposition 1 *For any given* $(\omega_0, \omega_1, \omega_2)$, *the DM's optimal taxes are*

$$t_1^{\circ} = \frac{\left(-D + \frac{\omega_1}{\omega_\circ}Y - \frac{\omega_2}{\omega_\circ}z\right)g + \left(d + \frac{\omega_1}{\omega_\circ}Z - \frac{\omega_2}{\omega_\circ}y\right)G}{Fg - fG}, \tag{13}$$

$$t_2^{\circ} = \frac{F(-d - \frac{\omega_1}{\omega_\circ}Z + \frac{\omega_2}{\omega_\circ}y) + f\left(D - \frac{\omega_1}{\omega_\circ}Y + \frac{\omega_2}{\omega_\circ}z\right)}{Fg - fG}, \tag{14}$$

where

$$\begin{aligned}
D &= \frac{1}{2}\left[\left(\alpha_1 + \gamma_1 x - b_1\right)Y - \left(\gamma_1 z + 1\right)X - \left(\alpha_2 + \gamma_2 X - b_2\right)z + \gamma_2 Y x\right] + X, \\
F &= 1 - z\gamma_1 - z\gamma_2, \\
G &= \frac{1}{2}\left(\gamma_1 yY + \gamma_2 yY + \gamma_1 zZ + \gamma_2 zZ - Z - z\right), \\
d &= \frac{1}{2}\left[-\left(\alpha_1 + \gamma_1 x - b_1\right)Z + \gamma_1 yX + \left(\alpha_2 + \gamma_2 X - b_2\right)y - \left(\gamma_2 Z + 1\right)x\right] + x, \\
f &= \frac{1}{2}\left(\gamma_1 yY + \gamma_2 yY + \gamma_1 zZ + \gamma_2 zZ - Z - z\right), \\
g &= \left(1 - Z\gamma_1 - Z\gamma_2\right)y.
\end{aligned}$$

Given (13) and (14), (5) and (6) become the system of two linear equations as follows:

$$\left(1 + Q\right)\omega_1 + Q\omega_2 = Q, \tag{15}$$

$$P\omega_1 + (1 + P)\omega_2 = P, \tag{16}$$

where

$$\begin{aligned}
Q &= \frac{\left(t_1^{\circ}\left(Fg - fG\right) + Dg - dG\right)\left(Fy + fz\right) + \left(t_2^{\circ}\left(Fg - fG\right) + Fd - fD\right)\left(zg + yG\right)}{(Yg + ZG)(Fy + fz) - (FZ + fY)(zg + yG)}, \\
P &= \frac{\left(Yg + ZG\right)\left(t_2^{\circ}\left(Fg - fG\right) + Fd - fD\right) + \left(Fz + fY\right)\left(t_1^{\circ}\left(Fg - fG\right) + Dg - dG\right)}{(Yg + ZG)(Fy + fz) - (FZ + fY)(zg + yG)}.
\end{aligned}$$

We can identify ω_1 and ω_2 by solving the system of two equations (15) and (16) for ω_1 and ω_2 and they are

$$\omega_1 = \frac{Q(1 + P) - PQ}{\left(1 + Q\right)\left(1 + P\right) - PQ}, \tag{17}$$

$$\omega_2 = \frac{(1 + Q)P - PQ}{\left(1 + Q\right)\left(1 + P\right) - PQ}. \tag{18}$$

First, we can see if and how much observed taxes t_1° and t_2° are higher than the revenue maximizing taxes t_1^* and t_2^*. Second, by identifying ω_i for $i = 1, 2$ we can see whether the net externality associated with the production and consumption of each product is positive or negative. If $\omega_i < 0$ (> 0), the net externality is positive (negative). Third, from the identified weights in the social welfare function we can identify the DM's stand on how to balance the total surplus and the net externalities associated with different types of products. The DM's marginal rate of substitution of the net externality for the total surplus, $\omega_i/\omega_0 = dTR/dq_i^*$, specifies the amount by which total surplus is reduced for a change in the externalities due to a marginal decrease in the equilibrium quantity of product i, holding social welfare constant.

Suppose that $\omega_2 > 0$, which implies that the net externalities of good 2 is negative. If $\omega_1/\omega_0 < \omega_2/\omega_0$ with $\omega_2 > 0$, then an increase in the (positive) externality due to a marginal reduction of equilibrium quantity of good 2 must be accompanied by a relatively more decrease in the total surplus, *holding social welfare constant*, than an change in the externality due to a marginal reduction of equilibrium quantity of good 1 is. This implies that the value of the reduction of equilibrium quantity of good 2 measured in the total surplus is higher than that of good 1. Therefore, the DM would be willing to control the equilibrium quantity of good 2 more aggressively through her tax policy.

3 Analysis of Excise Taxation in the Ontario Beer Market

The empirical analysis consists of two main parts. First, we are interested in evaluating how the excise tax rates observed in practice in Ontario compare with the implication of tax revenue maximizing rates derived above. In order to evaluate that observation, we need to estimate the demand functions specified in equations (7) and (8). Then, using the estimated demand parameters $\alpha_i, \beta_i, \gamma_i$ for $i \in [1, 2]$ we can directly calculate the optimal excise taxes from eqs. (11) and (12) and compare them to the actually set rates.

Second, with the t_1^* and t_2^* we can calculate our social welfare weights of interest ω_1 and ω_2 from equations (17) and (18) in order to identify the decision maker's social welfare function.

3.1 Data and Estimation of Demand Functions

We make use of two unique datasets for our demand estimation: retail beer prices and sales over the 2005-2015 period in Ontario and the matching local households' income and population data from administrative tax record for Ontario. Both data sets are obtained under a research contract and are proprietary.

The beer sales data is the store-level panel data obtained from Liquor Control Board of Ontario (LCBO) under Ontario's freedom of information laws. LCBO is a government corporation that retails and distributes alcoholic beverages throughout the Canadian province of Ontario. LCBO maintains a quasi-monopoly on the retail sale of beer, with The Beer Store being the only other retailer.[7] See Sen [2013] for an economic analysis of The Beer Store in Ontario.[8]

This is a higher frequency data containing weekly observations on the pairs of retail price and quantity sold for each beer product at the Stock Keeping Unit (SKU) level for every LCBO store in Ontario.[9] Given that Ontario changed its beer tax policy on July 1st, 2010 we obtained a symmetric time period of five years before and after the policy change. Therefore, our weekly data includes every week between January 2005 - December 2015.

Besides SKU level product information, our dataset includes information on each LCBO store's GPS location, a full address, and the opening and closing dates. This allows us to link each store's geographic location to its respective area's household socio-economic characteristics. For the area where a store is located we chose the Forward Sortation Area (FSA). The FSA designates a geographical unit based on the first three digits in a Canadian postal code, for both urban and rural areas. Given that we have the full address for each LCBO store we can easily determine its FSA. An FSA captures a wider area than a postal code, which can even be very store specific, thus not allowing us to determine its economic

[7] As of 2016, LCBO authorized some supermarkets to sell beer within their grocery aisles, thereby seemingly weakening the Beer Stores and LCBO's quasi-monopoly. This does not affect our estimates since our data includes years 2005 to 2015. Also, the grocers are not allowed to negotiate their own supply and pricing terms directly with the brewers and the beer prices are set by LCBO. Further, only those grocery stores that are located far enough from the next LCBO location are allowed to sell beer.

[8] The Beer Store holds a monopoly on selling large packs of beer, such as 24-bottle cases, although LCBO can sell these in some locations. LCBO also has more outlets and pays higher (union) wages to its employees. In the Beer Store qualified brewers are free to list their products and set their own selling prices. However, the beer prices are still subject to LCBO price approval that must comply with legislated minimum and uniform pricing requirements. Ontario's uniform pricing regulation precludes price competition between beer retailers, LCBO and the Beer Store, effectively making LCBO the key pricing agent in the beer market.

[9] Stock Keeping Unit product information includes product and producer name, unit size, packaging count, price. It also provides stores with inventory information, although we don't observe those.

characteristics.[10] Then, using the Canadian personal income tax returns, the T1 Family File, we obtained a custom tabulation of family (household) unit counts and average total income for each FSA in Ontario from Statistics Canada.

Each LCBO store sells many different beer products from many producers. Following our theory, we group these into two distinct producers: micro brewers (craft beer) and manufacturers (big beer). Ontario Ministry of Finances maintains a list of beer producers that are subject to a different beer tax rate following the change in tax policy, either a micro brewer or manufacturer rate. This allows us to precisely identify which beer producers are subject to which of the two taxes in our dataset.

Table 1 shows the yearly counts of LCBO stores, retail weeks and FSAs in our data. With population growth, the number of stores and FSAs increased from 2005 to 2015 in Ontario. The last column of Table 1 counts the unique store-week combinations for which both micro brewers and manufacturers beer are sold in each store per year. To note in particular is that, assuming all stores always carry big beer products, as the number of stores carrying craft beer increases the count of unique store-week combinations rises over the years. This gives us an average of 24,461 store-weeks per year and a total of 293,530 store-week observations for the whole period.

	Store	Weeks	FSA	Store-Week ID
2005	450	53	307	17405
2006	484	52	324	19869
2007	512	52	326	21567
2008	555	52	342	23955
2009	587	52	347	26972
2010	596	52	348	28299
2011	610	53	351	28997
2012	629	52	356	29626
2013	634	52	358	31303
2014	643	52	361	32344
2015	653	52	365	33192

Table 1: Counts of LCBO stores, sales weeks, Forward Sortation Areas and unique store-week combinations in Ontario over the years.

Using weekly data allows us to observe changes in prices due to temporary discount sales on a subset of products. However, there is no price variation across stores, which eliminates the possibility of store bargain shopping for consumers. Further, weekly data

[10]This is either due to a lack of information for such a narrow geographic location or because of confidentiality rules by Statistics Canada and Canada Revenue Agency.

reflects the strong seasonal effects in beer sales across each year, such as major holidays during which beer sales surge. As an example, this seasonality is illustrated in Figure 1 for 2015, the last year in our data.

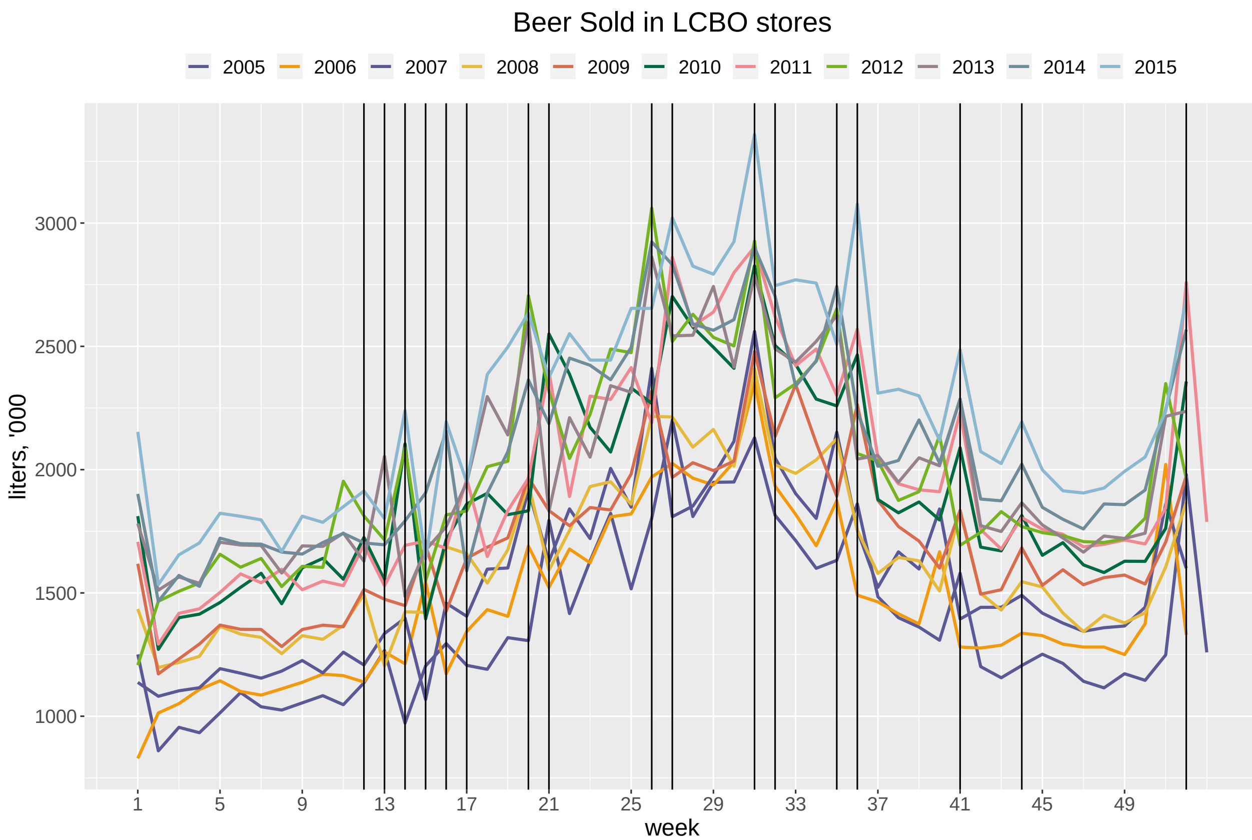

Figure 1: Weekly liters of beer sold in all Ontario LCBO stores from 2005 to 2015. The vertical lines indicate relevant statutory holidays. We use these dates to controls for the holiday weeks in our demand regressions.

Each LCBO store carries hundreds of beer products in different package standards and unit sizes. This requires us to standardize units sold and their price across all unit types in all LCBO stores. Table 2 provides the summary counts of unique products offered (SKU numbers), unique unit sizes in which those products are supplied (measured in mL), number of unique producers in our data. Notice that the number of unique producers and their products steadily increases over the years, primarily because of new micro breweries entering the Ontario market.

The last column in Table 2 indicates the number of unique bottles or cans per pack offered across all the various producers. For example, a producer may sell the same or different beer brands in bottles and/or cans (tall and short), a four-pack or a six-pack or some special promotional packaging with extra units.[11] Because our data comes at the SKU level we have considerable heterogeneity in the types of beer products sold by a brewer at any given store. Therefore, as a standardized measure of quantity sold in a given store-week we calculate a single quantity, total liters sold for each producer's beer. Liter is the natural choice for our setting since the beer excise taxes on micro and non-micro beer are imposed on a per liter basis.

[11] For example, producers often run promotional campaigns where a product might come with extra "free" units, such that a standard 12-pack is sold with bonus three more units, making it a 15-pack.

	Unique SKU	Unique Name	Unique Unit Size	Unique Producers	Bottles per Pack
2005	466	442	42	31	12
2006	467	443	39	33	11
2007	496	466	42	33	14
2008	487	468	40	33	12
2009	496	480	38	37	12
2010	567	548	48	41	15
2011	624	606	53	42	15
2012	678	661	57	47	15
2013	749	732	59	58	14
2014	820	797	63	71	14
2015	906	882	63	93	14

Table 2: Count of unique beer product characteristics

		N	Mean	SD	Min	Pc(25)	Median	Pc(75)	Max
Manufacturers' Beer Sold, L	q_{2ysw}	293,529	2,940.04	2,009.39	1.89	1,417.21	2,576.14	4,000.36	24,635.80
Microbrewers' Beer Sold, L	q_{1ysw}	293,529	209.17	342.25	0.28	17.47	68.6	251.42	4,997.12
Manufacturers' Price per L	p_{2ysw}	293,529	4.55	0.17	3.92	4.44	4.55	4.67	6.30
Microbrewers' Price per L	p_{1ysw}	293,529	6.05	0.43	3.15	5.84	6.10	6.32	9.95
Household Income in FSA, $	i_s	292,988	75,564.3	28,511.9	25,920.0	61,200.0	70,090.00	83,210.0	706,940.00
No. of Households in FSA	N	292,988	15,357.5	9,519.59	90.0	8,380.0	13,900.0	19,700.0	48,940.0
No. of Stores per FSA	S	293,529	3.43	3.72	1	1	2	5	15
No. of Households per Store	n_s	292,988	7,638.71	5,703.03	90.00	3,015.56	6,420.00	10,490.00	44,340.00

Table 3: Summary statistics of main variables of interest used in regression analysis for the 2005-2015 period.

The increased popularity of craft beer products and their demand has already received some interest on its own in Toro-Gonzalez et al. [2014]. The literature studying beer demand and brewing industry employs several empirical approaches, with the Almost Ideal Demand System of Deaton and Muellbauer [1980] as employed by Hausman et al. [1994] and the BLP model of Berry et al. [1995] as in Miller and Weinberg [2017] being most prominent. Although beer demand regression is a part of our empirical analysis, we do not seek to contribute to this extensive demand estimation literature in this paper nor establish a definitive causal relationship.[12] We leave this goal for future research using our point-of-sale data from the specific alcohol market in Ontario.

Inverse demand functions in equations (7) and (8) are the basis for demand estimation. Regressions estimates are performed on a panel data of all LCBO stores s per week w in year y over the 2005-2015 period. Therefore, the unit of analysis is a store-week. Using a longer time series is important as it allows for more precise estimation of demand coefficients. The coefficients we estimate are demand correlation parameters and the goal is that they have a correct sign in the price-quantity demand relationship. Following (7) and (8), our regressions are done in levels, rather than logs, in order to be able to derive expressions of various welfare metrics. While linear demand functions are restrictive, and a logit-type models would allow for more substitution patters and heterogeneity in demand, we require estimates of demand parameters that are intuitive in order to derive revenue maximizing excise taxes and social welfare weights, the primary aim of the study.

Let q_{1ysw} be the quantity of micro brewers' beer sold in the area where store s is located in week w of year y and p_{1ysw} its price. We define q_{2ysw} and p_{2ysw} analogously for manufacturers' (non-micro brewers) beer. The demand functions estimation is set up as follows:

$$p_{1ysw} = \alpha_1 + \alpha_{1y} + \alpha_{1s} + \alpha_{1h}h(w) + \alpha_{1n}n_s + \alpha_{1i}i_s + \beta_1 q_{1yws} + \gamma_1 q_{2yws} + u_{1ysw}, \quad (19)$$

$$p_{2ysw} = \alpha_2 + \alpha_{2y} + \alpha_{2s} + \alpha_{2h}h(w) + \alpha_{2n}n_s + \alpha_{2i}i_s + \beta_2 q_{2yws} + \gamma_2 q_{1yws} + u_{2ysw}. \quad (20)$$

Quantities sold for micro brewers (q_{1ysw}) and manufacturers (q_{2ysw}) are measured in total liters. With beer sold standardized into a single quantity sold, we also need a single price. The dependent price variable p_{1ysw} is the weighted average price per liter across all micro brewers' beer sold in week w at store s, while p_{2ysw} is the weighted average weekly store

[12] Toro-Gonzalez et al. [2014] provides a useful summary of the beer demand literature, Friberg and Romahn [2018] for beer demand in Sweden. Rojas [2008] studies pricing conduct in the U.S. beer industry using beer excise tax increase as a natural experiment.

price per liter across all manufacturers' beer sold. The weights used are fixed product shares of each year's sales to avoid contaminating prices with quantity responses on a weekly basis. Therefore, for each store-week observation we have a standardized liter price and liters sold for micro brewers and manufacturers' beer. Note that we pick liter as the standardized unit of measurement because the beer taxes are set in terms of cents per liter.

The coefficients β_1 and β_2 indicate own product effects. Consistent with the theory of demand we expect them both to be negative. On the other hand, γ_1 and γ_2 capture product effects which indicate whether manufacturers and micro brewers' beer are substitutes or complements. We conjecture they are substitutes because the nature of the products allows them to be used for the same purpose and because of the anecdotal evidence that demand for big brewers' beer is flat or in decline over years as is being displaced by the rise of craft breweries. Positive coefficients would indicate they are substitutes.

To account for various determinants of demand other than price and eliminate unobserved heterogeneity we include certain controls in our estimation. First, each year sees predictable spikes in demand around holidays, such as Victoria Day, the unofficial start of summer in Canada, Canada Day on July 1^{st}, Labour Day in September, and end of the year holidays in December. We control for these unusually high, but temporary increases in beer demand, which are anticipated by producers with price discounts, by including a time dummy α_{1h} for holiday weeks, where $h(w) = 1$ if a week w contains a holiday.[13] These spikes in demand are visible in Figure 1 for 2015, where the vertical lines indicate the holidays being controlled for in every year.

Second, store fixed effects α_{1s} capture time-invariant heterogeneity across regions and cities of Ontario, while year fixed effects α_{1y} account for the time varying factors that might have common influence on beer demand across the entire province, such as the recession of 2008-09.

Finally, we control for the number of households and average yearly income in each store's area. The average number of households who live in the area where store s is located, n_s, is taken as the consumer base of that particular store, while the average household income i_s captures that area's purchasing power. Table 3 presents summary statistics for the variables used in demand estimation.

As discussed above, we set the Forward Sortation Area (FSA) for the area where a store is located. Not every FSA in Ontario contains an LCBO, but some FSAs may contain a

[13]Holidays accounted for are: Christmas and Boxing Day, New Year Day, Easter, Victoria Day (last Monday preceding May 25), Civic Holiday (Ontario provincial holiday in August), Labour Day, Thanksgiving Day (second Monday in October) and Halloween (All Saints Day).

few. In particular, this is the case for rural areas. If $S > 1$ stores are located in an FSA we divide the total number of households by S to calculate the average number of households for each store in the FSA, n_s, and take the average family income in the FSA for each store in the FSA.

	Dependent variable:					
	Price Micro	Price Manu	Price Micro	Price Manu	Price Micro	Price Manu
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Constant	5.9845***	4.5691***	5.9013***	4.5006***	5.5573***	4.7294***
	(0.0108)	(0.0043)	(0.0222)	(0.0100)	(0.0314)	(0.0112)
Micro L Sold	0.3809***	0.2470***	0.3498***	0.2174***	−0.0299***	0.0038
	(0.0261)	(0.0150)	(0.0171)	(0.0090)	(0.0106)	(0.0037)
Non-Micro L Sold	−0.0060*	−0.0256***	−0.0032	−0.0252***	0.0032*	−0.0083***
	(0.0032)	(0.0014)	(0.0022)	(0.0010)	(0.0018)	(0.0005)
Store FE?	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Year FE?	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Holidays?	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Households?	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	293,529	293,529	292,988	292,988	292,988	292,988
Clusters	663	663	663	663	663	663
Adjusted R^2	0.0886	0.2535	0.0924	0.2936	0.4421	0.8197

Note: Statistical significance: *p<0.1; **p<0.05; ***p<0.01

Table 4: This table shows regression results for demand equations 19 and 20 by week and store during the January 2005 to December 2015 period. Own demand coefficient estimates are labeled by a β, while cross demand coefficients are denoted by a γ in the two equations. Three different specifications are presented differing by set of controls used. Columns 3a and 3b are our preferred specifications, controlling for a full set of fixed effects, holiday weeks and household characteristics, average income and size. We use these coefficient estimates in our revenue maximizing and optimal beer tax calculations. Standard errors are clustered at the store level.

Table 4 presents simultaneous regression results of equations 19 and 20 in the set of three different specifications. The first two columns represent basic OLS results with no controls. The own quantity coefficients do not have negative signs as expected from demand theory. The next columns (2a) and (2b) introduce the holiday week dummies, household income and count controls. However, this does not improve the initial specification as the key parameters β_i and γ_i still do not have the expected demand signs.

The last two columns (3a) and (3b) are our preferred estimation specification, accounting for a full set of store and year effects while keeping the controls from the previous specification. The estimated coefficients on own quantity sold, β_1 and β_2 have the expected negative sign, while the cross effects, γ_1 and γ_2 indicate that microbrewery and manufacturer's beer are, as conjectured, substitutes.

Our implicit assumption is that the weekly supply of beer is relatively inelastic for the following reasons. Unlike other alcohol products, freshly brewed beer is beer in its optimum state which happens immediately at the end of the brewing cycle. Best consumed at this point and typically holds its perfect flavour for around 3 weeks. After that, the flavours quickly deteriorate, leading to the presence of stale flavours and a decrease in the drinkability of the beer. Therefore, it is not desirable to maintain a large volume of the inventory. Furthermore, the brewing process is also time-intensive. For example, it takes anywhere between 4-8 weeks to brew a commercial lager and between 2-3 weeks to brew a commercial ale.[14] Because of those reasons described above, it would be difficult to accommodate a sudden change in the weekly demand observed in the area covered by a particular store. Fortunately, brewers can make their production plans fairly accurately in advance because the amount of beer consumption is very predictable. As Figure 1 shows there is a high demand for beer in the summer and during winter holidays.

However, given the possibility of endogeneity we would prefer to conduct a robustness check of our demand estimation using an instrumental variable. As a quick check, we perform a 2SLS regression of our most preferred specifications, the last two columns in Table 4 using the 52 week lag of beer sold (i.e., the same week one year earlier) as an instrumental variable for its own respective beer type. These results are presented and discussed in Appendix A.1.

[14]See How Long Does It Take to Brew Commercial Beer?

3.2 Observed tax rates, inverse elasticity rule and revenue maximizing tax rates

With the demand functions estimates in hand we have all the relevant parameter values to derive the revenue maximizing excise taxes t_1^* and t_2^* based on expressions (11) and (12). Our preferred regression estimates in columns (3a) and (3b) of Table 4 give us the values for α_i, β_i and γ_i for both beer types. The sum of the marginal cost and industry mark-up for each beer producer, b_1 and b_2, can easily be identified by subtracting the observed beer excise taxes t_1 and t_2 from the respective liter prices, p_1 and p_2. Actually set beer tax rates in Ontario are available from the Ontario Ministry of Finance website since their introduction in 2010.

Year	$\bar{q}_1^*$	$\bar{q}_2^*$	e_{11}	e_{22}	e_{12}	e_{21}
2010	145.930	3136.183	-0.047	-0.028	-0.066	-0.001
2011	190.136	3137.972	-0.039	-0.029	-0.051	-0.001
2012	237.076	3109.595	-0.033	-0.030	-0.042	-0.001
2013	282.872	2998.111	-0.030	-0.031	-0.036	-0.001
2014	355.635	2968.294	-0.025	-0.032	-0.029	-0.001
2015	468.299	3052.996	-0.020	-0.032	-0.022	-0.001
Mean	279.991	3067.192	-0.032	-0.030	-0.041	-0.001

Table 5: Own (e_{ii}) and Cross Beer (e_{ij}) Tax Elasticities. The first two columns indicate the average weekly quantity of type i beer sold per store in each year.

We first examine if the actual tax rates follows the inverse elasticity rule. Table 5 presents the resulting own and cross elasticities on a yearly basis, along with estimated equilibrium quantities. Manufacturers' beer is getting slightly more elastic with respect to its own price over the years, whereas micro beer is getting more inelastic. In later years, manufacturers' beer is more elastic with respect to its own price than micro beer is. Based on the inverse elasticity rule, it would make sense to tax micro beer higher. Therefore, the actual tax rates violates the inverse elasticity rule.

Now we compare the values of calculated revenue maximizing beer tax rates to each other and whether they are lower or higher than the ones actually set for Ontario.[15]

Table 6 compares the calculated beer excise taxes (t_1^* and t_2^*) from eq. (11) and eq. (12) with their actually observed counterparts (t_1 and t_2). The patterns of observed taxes over time are consistent with the pattern of the predicted revenue-maximizing tax rates based

[15]Throughout the paper we refer to the observed beer excise taxes set by the government of Ontario as 'actual' and those those estimated as revenue maximizing based on our theory as 'calculated'.

	Calculated				Actual			
Year	t_1^*	t_2^*	$\%\Delta t_1^*$	$\%\Delta t_2^*$	t_1	t_2	$\%\Delta t_1$	$\%\Delta t_2$
2010	0.110	0.337			0.196	0.697		
2011	0.113	0.368	2.588	9.073	0.209	0.709	6.633	1.722
2012	0.122	0.374	8.110	1.950	0.223	0.723	6.699	1.975
2013	0.136	0.383	11.326	2.328	0.240	0.740	7.623	2.351
2014	0.147	0.391	7.787	1.901	0.254	0.754	5.833	1.892
2015	0.160	0.399	8.819	2.114	0.266	0.766	4.724	1.592
Mean	0.131	0.375	7.726	3.473	0.231	0.732	6.302	1.906

Table 6: Calculated revenue maximizing beer taxes for microbrewers (t_1^*) and manufacturers (t_2^*) compared to the actually set beer taxes, t_1 and t_2 respectively, in Ontario starting in 2010.

on the estimated demand functions: the tax on beer made by Ontario micro brewers is consistently lower than that on beer made by large manufacturers. However, we also find that both t_1^* and t_2^* are consistently lower than their respective actually set rates observed over the period 2010-15. Specifically, we calculate t_1^* as consistently slightly less than half of t_1, while t_2^* is slightly more than half of t_2 Figure 2 plots the beer tax rates from table 6 to illustrate these points. The gap between the observed and revenue maximizing tax rates and the scale difference in the gap might be indicative that the government's tax policy is not consistent with just revenue maximization objective.[16]

Therefore, the inverse elasticity rule and the revenue maximization cannot explain the levels and patterns of the observed tax rates. To explore the reasons for changes in observed tax rates and the gap between the observed tax rates for micro beer and non-micro beer, we quantify the weights that the DM attaches in the social welfare function.

3.3 Identification of the Social Welfare Function

The observed taxes are assumed to be set to maximize the social welfare in eq. (4), which describes the government's preferences over total surplus and net externalities created by the production and consumption of beer products.

We can identify the ω weights on consumption of craft beer and big manufacturers beer in the DM's social welfare function from equations (17) and (18) respectively which have all the necessary parameter values: β and γ come from our regressions, and for t_1° and t_2° we

[16]As a robustness check, in appendix A.2 we calculate revenue maximizing tax rates using an alternative method: revenue maximizing rates are determined only in the initial year, 2010, and then increased in future years following the CPI rule. These alternative revenue maximizing tax rates are not substantially different from the ones in Table 6.

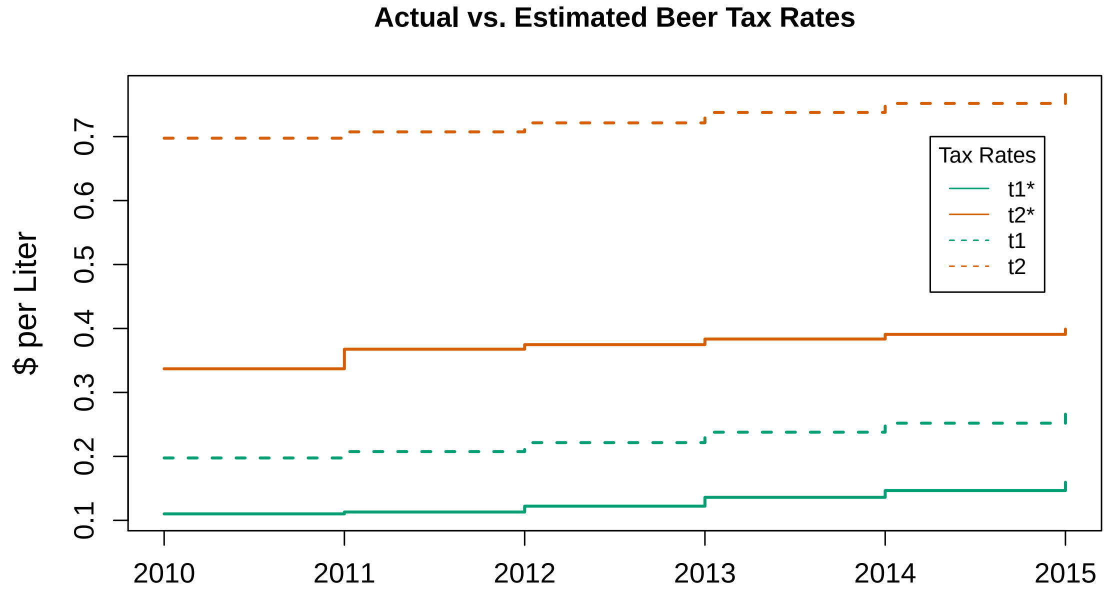

Figure 2: Comparing calculated revenue maximizing beer taxes from eqs. (11) and (12) (solid lines) with the actually set ones (dashed lines). The actually set tax rates are always higher than the respective revenue maximizing ones, but in both cases the rate which applies to microbrewers' beer is lower than the one for manufacturers' beer.

take the actually set beer excise taxes.

Table 7 shows the calculated values of ωs over the years. The sign of ω_1 is negative, whereas the sign of ω_2 is positive during 2010-2015. Therefore, the DM thinks that the net externality of the production and consumption of micro beer is positive, whereas non-micro beer brings the negative net externality. The weights seem very stable over time. The weight on the total surplus, ω_0, is in the range of 0.5906 to 0.5570, meaning the government places slightly more weight on the total surplus. The weight ω_1 ranges from -0.1976 to -0.1860 and it is increasing at a very low annual rate except from 2010 to 2011.[17] The weight ω_2 ranges from 0.6009 to 0.6290 and it is very stable, increasing only at a very low annual rate.

Given the identified weights, we can derive the DM's marginal rate of substitution (MRS) for each type of beer product. The MRS is defined as $\frac{\omega_1}{\omega_0}$ and $\frac{\omega_2}{\omega_0}$ and specifies the amount by which the total surplus is reduced for a change in the net externality due to a marginal decrease in the equilibrium quantity of each type of beer products, holding social welfare

[17] A high decreasing rate of ω_1 from 2010 to 2011 may be related to the fact that the differentiated excise tax policy was introduced in July 2010 so that ω_1 for 2010 was identified with only half-year year.

Year	ω_1	ω_2	ω_0	$\frac{\omega_1}{\omega_0}$	$\frac{\omega_2}{\omega_0}$	$\%\Delta\frac{\omega_1}{\omega_0}$	$\%\Delta\frac{\omega_2}{\omega_0}$
2010	-0.1915	0.6009	0.5906	-0.3243	1.0174		
2011	-0.1976	0.6157	0.5819	-0.3395	1.0580	-4.7080	3.9991
2012	-0.1962	0.6198	0.5764	-0.3404	1.0753	-0.2393	1.6291
2013	-0.1926	0.6236	0.5690	-0.3385	1.0959	0.5505	1.9134
2014	-0.1903	0.6270	0.5633	-0.3378	1.1130	0.1926	1.5662
2015	-0.1860	0.6290	0.5570	-0.3339	1.1293	1.1499	1.4620
Mean	-0.1924	0.6193	0.5730	-0.3357	1.0815	-0.6109	2.1140

Table 7: Social Welfare Function weights and their changes.

constant. During 2010-2015, the MRS of micro beer is in the range of -0.3243 to -0.3404 and it steadily increases, except for 2010-2011 and 2011-2012. On the other hand, the MRS of non-micro beer is in the range of 1.0580 to 1.1293 and it also steadily increases.

Because the MRS of micro beer is negative, total surplus should increase by 33 cents on average during 2010-2015 for a decrease in the positive net externality due to one litre decrease in the equilibrium quantity of micro beer, holding the social welfare constant. On the other hand, because the MRS of non-micro beer is positive, total surplus should decreases by a dollar eight cents ($1.08) on average for a decrease in the negative net externality due to one litre decrease in the equilibrium quantity of non-micro beer.

In other words, from the DM's point of view, social welfare increases by $0.33 on average due to the positive net externality associated with a one-liter increase in the equilibrium quantity of micro beer, whereas it decreases by $1.08 on average due to the negative net externality associated with an one-litre increase in the equilibrium quantity of non-micro beer. If we assume that both types of beer create the same amount of the negative externality per litre, this implies that the DM believes that mirco-beer brings more positive externality by $1.41 per liter. This implies that the DM would be willing to control the equilibrium quantity of non-micro beer more aggressively through her tax policy.

Appendix A.3 shows the results on social welfare function weights based on four additional different specifications of the demand functions: adding (i) week fixed effects, (ii) month fixed effects, (iii) quadratic trends, (iv) quadratic trends and month fixed effects. Social welfare function weights show the same patterns under all four specifications. That is, ω_1 is negative, whereas ω_2 and ω_3 are positive. Therefore, the MRS of micro beer is negative, whereas the MRS of non-micro beer is positive.

However, the additional positive externality that the DM believes micro beer brings is higher with all these four specifications: $1.93 with week fixed effects in the demand function,

\$1.78 with month fixed effects, \$1.43 with quadratic trends, (iv) \$1.94 with quadratic trends and month effects. We chose the current parsimonious specification for the demand function without those fixed effects and quadratic trends because it provides the most conservative estimate of the additional positive externality that the DM believes micro beer brings (\$1.41).

Why is there a high discrepancy and opposite signs of MRSs between two types of beers? Given that micro beer is locally supplied mostly by small business owners[18], the positive externality associated with the production of micro beer would include (i) the protection of small business owners who are the main suppliers in the nascent micro brewing industry, and (ii) the positive impact on the local suppliers of ingredients and employment. The negative value of ω_1 suggests that the government believes the positive externality associated with the production of micro beer outweighs the negative externality associated with the consumption of micro beer. Does this belief justify a gap between excise taxes on micro and non-micro beer products? On average, the excise tax on micro beer (23.1 cents per litre) is only a third of that on non-micro beer (73.1 cents per litre). (See Table 6.)

As alluded in Section 3.2, a straight comparison of elasticities does not tell us the whole story. This is because the two markets differ by size. Table 5 shows that the market for micro beer is only 9% of the marker for non-micro beer on average during 2010 - 2015 period. Then, suppose the government imposes the same dollar tax increase on each type of beer. Each type's equilibrium quantity will then shrink by a different percentage; a decrease in micro beer is larger than in non-micro beer. In that case, a better measure to compare is $\frac{e_{ij}}{t_i}$ for both micro and non-micro beer.

By a simple manipulation of the elasticity definition we obtain the following equivalence: For all $i, j \in \{1, 2\}$

$$\frac{e_{ij}}{t_j} = \frac{x_{ij}}{q_i^*} \tag{21}$$

where $x_{ij} := \frac{\partial q_i^*(t_i, t_j)}{\partial t_j}$. If $j = i$, then $\frac{e_{ii}}{t_i}$ is the percentage decrease in the equilibrium quantity of type-i beer if tax for that type increases by \$1; a semi-elasticity of beer tax. We prefer this semi-elasticity for comparison because the Ontario government adjusts beer taxes in equal dollar terms and the semi-elasticity gives the percentage change in q_i^* when

[18] For example, the anecdotal evidence compiled by the Trillium Network for Advanced Manufacturing indicates that the number of breweries in Ontario increased from less than 100 to more than 320 between 2010 and 2019 and the number of people employed by breweries jumped from 2,220 in 2010 to more than 5,800 in 2019. This craft breweries' employment growth offset declines in employment at the big brewers. See Sweeney [2020].

Year	$\frac{\partial q_1^*}{\partial t_1}/\bar{q}_1^*$	$\frac{\partial q_2^*}{\partial t_2}/\bar{q}_2^*$	$\frac{\partial q_1^*}{\partial t_2}/\bar{q}_1^*$	$\frac{\partial q_2^*}{\partial t_1}/\bar{q}_2^*$	$\left\lvert\frac{q_1^*}{x_{11}}\right\rvert$	$\left\lvert\frac{q_2^*}{x_{22}}\right\rvert$	$\left\lvert\frac{q_1^*}{x_{12}}\right\rvert$	$\left\lvert\frac{q_2^*}{x_{21}}\right\rvert$
2010	-0.241	-0.040	-0.094	-0.005	0.041	0.247	0.106	1.968
2011	-0.185	-0.040	-0.072	-0.005	0.054	0.247	0.139	1.969
2012	-0.148	-0.041	-0.058	-0.005	0.067	0.245	0.173	1.951
2013	-0.124	-0.042	-0.049	-0.005	0.080	0.236	0.206	1.882
2014	-0.099	-0.043	-0.039	-0.005	0.101	0.234	0.259	1.863
2015	-0.075	-0.042	-0.029	-0.005	0.133	0.241	0.341	1.916
Mean	-0.145	-0.041	-0.056	-0.005	0.079	0.241	0.204	1.924

Table 8: Own and Cross Equilibrium Percentage Quantity Decreases in Response to a $1 Tax Increases. Multiply by 100 to obtain percentage in the first four columns.

t_j changes by $1.

The first four columns in Table 8 tell us the percentage decrease in the equilibrium quantity of type-i beer if tax on type-j beer increases by $1 for all $i, j \in \{1, 2\}$. The first two columns show that a one-dollar increase in the tax on micro beer shrinks the equilibrium quantity of micro beer disproportionately a lot more than the same one-dollar increase in the tax on non-micro beer shrinks its own equilibrium quantity: Over the period 2010-2015, a one dollar increase in the tax on micro beer decreases its equilibrium quantity by 14.5%, whereas the same increase in the tax on non-micro beer decreases its equilibrium quantity by 4% only. Therefore, even though the price of micro beer is higher than that of non-micro beer, a one-dollar increase in tax on micro beer decreases its equilibrium quantity *3.6 times more* in percentage point terms than the same amount of tax increase on non-micro beer decreases its equilibrium quantity. This suggests that the same amount of tax on each type of beer would have a more detrimental effects on the positive externality associated with the supply side of the nascent local microbrewing industry that the government thinks is important such as (i) the well-being of small business owners who are the main suppliers of micro beer and (ii) the impact on the local economy that the micro brewing industry. Given that the government believes the positive externality from the production of micro beer outweighs the negative externality of the consumption of it, it is not optimal for the government to impose the same amount of tax on each type of beer because the same amount of tax increase shrinks the market for micro beer disproportionately a lot more. This led them to impose a lower tax on micro beer.

The last four columns show the amount of tax increase on type j in order to shrink the equilibrium quantity of type-i beer by 1%. (i.e., the absolute value of the reciprocal of the first four columns divided by 100): On average, a 7.9 cent increase in tax on micro beer shrinks its own equilibrium quantity by one percent, whereas a 24.2 cent increase in tax on non-micro beer is needed to shrinks its own equilibrium quantity by one percent. Therefore, the same percentage decrease in the equilibrium quantity of micro beer would be observed only when the excise tax on micro beer is 32.6% of the excise tax on non-micro beer. This magnitude is closely aligned with the observed taxes: During 2010-2015, the average excise tax on micro beer is 32% (23.1 cents vs. 72.3 cents).

4 Extensions and Other Applications

While the identification of social welfare weights is based on two excise taxes on two differentiated products, it can be extended to an arbitrary number of different products. It would be very interesting and relevant to identify social welfare weights based on excise taxes on all types of alcohol products.

Suppose that there are N differentiated products. Then, the SWF becomes

$$S = \omega_0 TS(t, b, \psi) - \sum_{n=1}^{N} \omega_n q_n^*(t, b, \psi)$$

with $\omega_0 + \sum_{n=1}^{N} \omega_n = 1$, $t := [t_1, \ldots, t_N]$, and $b := [b_1, \ldots, b_N]$, and $\psi := [\psi_1, \ldots, \psi_N]$.

Once ψ is identified from the demand estimation, one can derive the optimal excise tax rates t° that maximize the value of the SWF. For each product n, let $t_n^\circ(\omega, b, \psi)$ denote the optimal with $\omega := [\omega_1, \ldots, \omega_N]$. Then we can solve the system of N equations that is set up similar to (5) - (6), for N unknowns, $[\omega_1, \ldots, \omega_N]$.

Differentiated excise taxation has received particular attention from the public health perspective. For example, the possibility of differentially taxing tobacco products based on its nicotine content to maximize incentives for tobacco users to switch from the most harmful products to the least harmful ones are discussed in Chaloupka et al. [2019, 2015]. One can use our method to analyze the government's choice of optimal excise tax rates on differentiated tabacco products if they were to introduce it for differentiated tabacco products. In particular, our method can be extended to an arbitrary number of differentiated products.

Total surplus or its variations are commonly used as a measure for social welfare. For example, competitition policy authorities around the world use variations of total surplus in recommending or stopping mergers, which include implicit weights on consumer and producer surplus. In Canada, there is often heated debate on whether the Competition

Bureau of Canada needed to put equal weights on consumer and producer surplus to determine the welfare effect of a merger or could choose to assign different weight (e.g., the Superior Propane and ICG vs. Comissioner of Compeition case in Canada). What can we learn about the DM's preferences from his decision on a merger case?

For simplicty, suppose that the DM faces a decision on merger between two firms that face the same market demand. Let $CS^{\circ}(b^{\circ}, \psi)$ and $PS^{\circ}(b^{\circ}, \psi)$ be consumer surplus and producer surplus prior to merge, where $b^{\circ} := [b_1^{\circ}, b_2^{\circ}]$ with b_i° being the sum of the marginal cost and the industry markup. Let $CS^1(b^1, \psi)$ and $PS^1(b^1, \psi)$ be consumer surplus and producer surplus after merge, where b^1 is the sum of the marginal cost and the markup for the monopolist. Suppose that the DM's SWF is the weighted sum of consumer surplus and producer surplus. The value of social welfare prior to merger is then $S^{\circ} = \omega_0 CS^{\circ}(b^{\circ}, \psi) + \omega_1 PS^{\circ}(b^{\circ}, \psi)$ with $\omega_0 + \omega_1 = 1$. The value of social welfare after merger is $S^1 = \omega_0 CS^1(b^1, \psi) + \omega_1 PS^1(b^1, \psi)$. If the merger is approved, it implies that

$$S^1 \geq S^{\circ} \Leftrightarrow \omega_1 \geq \omega_1^* := \frac{CS^{\circ} - CS^1}{(PS^1 - CS^1) - (PS^{\circ} - CS^{\circ})}$$

It is easy to see the threshold weight ω_1^* on producer surplus is in $(0, 1)$ because $CS^{\circ} > CS^1$, $(PS^1 - CS^1) > (PS^{\circ} - CS^{\circ})$, and $PS^1 > PS^0$. Note that b_i° $(i = 1, 2)$ and b^1 are the equilibrium prices before and after merger and that ψ can be identified from demand estimation. Then, one can derive the threshold weight ω_1^*. Therefore, if the DM approves merger, it reveals that his weight on producer surplus is at least as high as ω_1^*. The higher the value of ω_1^* is, the more business friendly the DM is or the more tolerant for higher concentration level the DM is in the concerned industry. This revelation may affect how interest groups interact with the DM to influence his decisions on future merger cases or industry policies.

5 Concluding Remarks

To our best knowledge, our paper is the first to offer a clear articulation of how to identify the DM's social welfare function. This methodology would inform governments on the consequences of what has been done in the past with respect to implicit weights on different stakeholder interests, which could then inform future policy decisions. We believe that this is a really neat demonstration of what can be done with relatively simple methods, which most policy makers are candidly, unaware of, and would be an invaluable addition to their toolkit.

Applying our methods, we identify the government's social welfare weights over the total surplus and net externalities associated with the production and consumption of micro and non-micro beer with a unique point-of-sale data on beer products at the SKU level and neighborhood socio-economic characteristics from Ontario. This uncovers why the government of Ontario chooses differentiated excise tax rates that contradicts the inverse elasiticity rule, allowing us to interpret the government's stance on how to balance the total surplus and net externalities from the consumption and production of different types of beer products.

The results indicate that government believes the positive externality associated with the production of micro beer outweighs the negative externality associated with its consumption. According to our estimates, the government believes that on average micro beer creates more positive externality by \$1.41 per liter consumed. The implication is that the government believes the cost of reducing the equilibrium quantity of beer through taxation is significantly higher for micro than non-micro beer. The observed gap in excise taxes is aligned with the discrepancy in the percentage decreases in the equilibrium quantities of two types of beer with respect a dollar increase in its own excise tax respectively. Equivalently, the discrepancy in the amounts of excise tax required to decrease the equilibrium quantities of two types of beer by one percent.

Another important aspect of the differentiated excise taxation, which we do not analyze in this paper is the regressiveness of such a policy. In order to analyze if the social welfare function in section 2 is robust to the possibility of regressiveness of differentiated beer taxes, we need to examine whether and how regressive the dual beer excise tax structure is. Specifically, differentiated excise tax rates on beer can be regressive if the share of lower taxed micro beer consumed is increasing in the household income. Using the LCBO data Han and Lesica [2023] establish that the share of micro beer sold in Ontario is increasing in household income and that it increased somewhat more for higher income households over time, and then estimate the degree of regressiveness of a differentiated beer excise tax policy adopted in Ontario.

References

Ahmad, E. and N. Stern (1984): “The Theory of Reform and Indian Indirect Taxes,” *Journal of Public Economics*, 25, 259–298.

Berry, S., J. Levinsohn, and A. Pakes (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63, 841–890.

Chaloupka, F. J., L. M. Powell, and K. E. Warner (2019): “The Use of Excise Taxes to Reduce Tobacco, Alcohol, and Sugary Beverage Consumption,” *Annual Review of Public Health*, 40, 187–201, pMID: 30601721.

Chaloupka, F. J., D. Sweanor, and K. E. Warner (2015): “Differential Taxes for Differential Risks — Toward Reduced Harm from Nicotine-Yielding Products,” *New England Journal of Medicine*, 373, 594–597, pMID: 26267620.

Conlon, C. T. and N. L. Rao (2020): “Discrete Prices and the Incidence and Efficiency of Excise Taxes,” *American Economic Journal: Economic Policy*, 12, 111–143.

Deaton, A. S. and J. Muellbauer (1980): “An Almost Ideal Demand System,” *American Economic Review*, 70, 312–326.

Friberg, R. and A. Romahn (2018): “Pass-Through by Multi-Product Firms,” *International Journal of the Economics of Business*, 25, 265–295.

Griffith, R., M. O’Connell, and K. Smith (2019): “Tax design in the alcohol market,” *Journal of Public Economics*, 172, 20–35.

Hamilton, S. F. (2009): “Excise Taxes with Multiproduct Transactions,” *American Economic Review*, 99, 458–471.

Han, S. and J. Lesica (2023): “Regressiveness of Differentiated Excise Taxation,” *Working Paper*.

Hausman, J., G. Leonard, and J. D. Zona (1994): “Competitive Analysis with Differentiated Products,” *Annals of Economics and Statistics*, 143–157.

Kenkel, D. S. (2005): “Are Alcohol Tax Hikes Fully Passed Through to Prices? Evidence from Alaska,” *American Economic Review*, 95, 273–277.

Kirby, J. and M. Lundy (2022): “Canada’s crowded craft beer industry is tapped out. What brewers say must happen to stay afloat,” *The Globe and Mail*.

Miller, N. H. and M. C. Weinberg (2017): “Understanding the Price Effects of the MillerCoors Joint Venture,” *Econometrica*, 85, 1763–1791.

Miravete, E. J., K. Seim, and J. Thurk (2018): “Market Power and the Laffer Curve,” *Econometrica*, 86, 1651–1687.

——— (2020): "One Markup to Rule Them All: Taxation by Liquor Pricing Regulation," *American Economic Journal: Microeconomics*, 12, 1–41.

Rojas, C. (2008): "Price Competition In U.S. Brewing," *Journal of Industrial Economics*, 56, 1–31.

Sen, A. (2013): "The Beer Store, Monopoly Profits and the Potential for Government Revenue: An Economic Analysis," *University of Waterloo, Technical Report.*

Shrestha, V. and S. Markowitz (2016): "The Pass-Through Of Beer Taxes To Prices: Evidence From State And Federal Tax Changes," *Economic Inquiry*, 54, 1946–1962.

Sweeney, B. A. (2020): "Good Things Brewing: Ontario's Craft Beer Industry, 2010-2019," *Trillium Network.*

Toro-Gonzalez, D., J. J. McCluskey, and R. Mittelhammer (2014): "Beer Snobs Do Exist: Estimation of Beer Demand by Type," *Journal of Agricultural and Resource Economics*, 39, 1–14.

A Online Appendix (Not for Publication)

A.1 IV robustness check

We are estimating inverse demand functions, where we need to instrument for the quantity. What instruments are available, especially at the individual brand level?

Preferably, we would be able to employ the Hausman instruments: quantities or prices in other markets, such as a different city, as instruments for the same product in the same time period in another market. These have been employed in Hausman et al. [1994] and Toro-Gonzalez et al. [2014] for example. Unfortunately, the LCBO pricing is uniform throughout all the stores in Ontario, which rules out using uncorrelated demand shocks across cities as an identification assumption. Alternative instruments include factor prices, such as malt barley, hops or brewers' wages. We tried to obtain those for the brewing industry in Ontario, but could not find a complete source on prices for the required period in our data. Ontario imports 97% of all hops used for brewing. Hops are a variety based industry in which prices are drastically different depending on the variety and location grown.

Table 9 presents the results using the 52 week lag of beer sold (i.e., the same week one year earlier) as an instrumental variable in columns (3a) and (3b) and quarterly lags, every 52, 39, 26, and 13 weeks a multiple instruments in columns (4a) and (4b).

The past year week quantity is highly correlated with the current year quantity sold, while the first-stage F statistics suggest we do not have a weak instrument.

The resulting coefficients are not significantly different from those in columns (3a) and (3b) of Table 4 regardless of the type of instrument used here. However, it is not obvious that a lagged quantity fully satisfies the exclusion restriction requirement. Even if the exclusion restrictions are not met, it still might be of some interest to see these results.

	Dependent variable:			
	Price Micro	Price Manu	Price Micro	Price Manu
	(3a)	(3b)	(4a)	(4b)
Non-Micro L Sold	0.0013		0.0020	
	(-0.0026,0.0051)		(-0.0019,0.0058)	
Micro L Sold		0.0095^{*}		0.0081
		(-0.0013,0.0202)		(-0.0025,0.0187)
IV Micro L Sold	-0.0364^{**}		-0.0458^{***}	
	(-0.0652,-0.0077)		(-0.0769,-0.0147)	
IV Non-Micro L Sold		-0.0079^{***}		-0.0072^{***}
		(-0.0097,-0.0062)		(-0.0088,-0.0055)
Store FE?	*Yes*	*Yes*	*Yes*	*Yes*
Year FE?	*Yes*	*Yes*	*Yes*	*Yes*
Holidays?	*Yes*	*Yes*		
Households?	*Yes*	*Yes*		
1st Stage Wald F	8399.94	1032.34	$4,267.6$	899.3
Sargan test			721.9	$4,871.6$
Observations	258,875	258,875	258,875	258,875
Adjusted R^2	0.4365	0.8258	0.4365	0.8258
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 9: **IV Regression Estimates**. The estimates here are compared to those presented in columns (3a) and (3b) in Table 4. Each regression instruments the quantity sold with a 52 week lag of beer sold, i.e., the same week one year earlier, for its own beer type. The first-stage F statistics suggest we do not have a weak instrument problem. The parenthesis below coefficient estimates show the 95% confidence interval, based on cluster-robust standard errors.

A.2 Alternative revenue maximizing tax rates

An alternative method to calculate revenue maximizing excise tax rates is to suppose that the DM determines tax rates t_1^* and t_2^* from expressions (11) and (12) only for year 2010, the initial year when Ontario set beer taxes. Then we simply follow the government's own tax adjustment rule and increase the rates in the subsequent years by the CPI rule. In other words, this method is a robustness check exercise that answers the question: What would the calculated revenue maximizing tax rates be if we follow the CPI rule in setting actual taxes?

	t_1^*	t_2^*	t_1	t_2	Δt_1	Δt_2	$\hat{t}_1^*$	$\hat{t}_2^*$
2010	0.1102	0.3370	0.1960	0.6970				
2011	0.1130	0.3676	0.2090	0.7090	0.0663	0.0172	0.1175	0.3428
2012	0.1222	0.3747	0.2230	0.7230	0.0670	0.0197	0.1206	0.3748
2013	0.1360	0.3835	0.2400	0.7400	0.0762	0.0235	0.1315	0.3836
2014	0.1466	0.3908	0.2540	0.7540	0.0583	0.0189	0.1439	0.3907
2015	0.1595	0.3990	0.2660	0.7660	0.0472	0.0159	0.1535	0.3970

Table 10: Alternative calculation of revenue maximizing tax rates based only on CPI rule.

Table 10 presents the results of this exercise. The first four columns show our calculated revenue maximizing beer taxes and actual beer taxes from table 6. We want to compare these to revenue maximizing taxes calculated following the CPI rule.

First, in columns 5 and 6 we calculate the change in actual beer taxes from columns 3 and 4, assuming that these were increased according to the government's rule on yearly adjustment to to the excise tax rates. Then, in the last two columns we calculate revenue maximizing tax rates as $t_i^* \times (1 + \Delta t_i)$, for $i = [1, 2]$. We denoted them as $\hat{t}_i^*$ to distinguish them from tax rates in columns 1 and 2. As we can see, these are very close to the initially calculated revenue maximizing tax rates and employing this alternative method would not change any of the subsequent results in a meaningful way.

A.3 Robustness checks on estimated social welfare weights

The following tables show alternative estimates of social welfare function weights (ωs) by adding four additional control variables in demand specification: (i) week fixed effects, (ii) month fixed effects, (iii) quadratic trends, (iv) quadratic trends and month fixed effects.

Year	ω_1	ω_2	ω_0	$\frac{\omega_1}{\omega_0}$	$\frac{\omega_2}{\omega_0}$
2010	-1.2852	1.0183	1.2669	-1.0144	0.8038
2011	-1.3068	1.0366	1.2702	-1.0288	0.8161
2012	-1.3601	1.0793	1.2809	-1.0619	0.8426
2013	-1.4528	1.1446	1.3082	-1.1106	0.8750
2014	-1.4111	1.1311	1.2801	-1.1024	0.8836
2015	-1.4575	1.1693	1.2882	-1.1314	0.9077
Mean	-1.3789	1.0965	1.2824	-1.0749	0.8548

Table 11: Social welfare function weights with week fixed effects in the demand specifications.

Year	ω_1	ω_2	ω_0	$\frac{\omega_1}{\omega_0}$	$\frac{\omega_2}{\omega_0}$
2010	-0.6601	0.7933	0.8668	-0.7615	0.9152
2011	-0.6854	0.8229	0.8625	-0.7947	0.9541
2012	-0.6878	0.8318	0.8560	-0.8035	0.9717
2013	-0.6857	0.8402	0.8455	-0.8110	0.9938
2014	-0.6849	0.8476	0.8373	-0.8180	1.0124
2015	-0.6801	0.8527	0.8275	-0.8219	1.0305
Mean	-0.6807	0.8314	0.8493	-0.8018	0.9796

Table 12: Social welfare function weights with month fixed effects in the demand specifications.

specifications of the demand functions by adding (i) week fixed effects, (ii) month fixed effects, (iii) quadratic trends, (iv) quadratic trends and month fixed

Year	ω_1	ω_2	ω_0	$\frac{\omega_1}{\omega_0}$	$\frac{\omega_2}{\omega_0}$
2010	-0.2962	0.6180	0.6782	-0.4367	0.9113
2011	-0.3053	0.6354	0.6700	-0.4558	0.9484
2012	-0.3037	0.6399	0.6638	-0.4575	0.9639
2013	-0.3000	0.6448	0.6552	-0.4578	0.9841
2014	-0.2977	0.6491	0.6485	-0.4590	1.0010
2015	-0.2935	0.6526	0.6409	-0.4580	1.0182
Mean	-0.2994	0.6400	0.6594	-0.4541	0.9711

Table 13: Social welfare function weights with quadratic quantities in the demand specifications.

Year	ω_1	ω_2	ω_0	$\frac{\omega_1}{\omega_0}$	$\frac{\omega_2}{\omega_0}$
2010	-0.8154	0.8872	0.9282	-0.8784	0.9558
2011	-0.8476	0.9216	0.9261	-0.9153	0.9952
2012	-0.8513	0.9312	0.9201	-0.9252	1.0121
2013	-0.8499	0.9403	0.9096	-0.9344	1.0338
2014	-0.8497	0.9483	0.9015	-0.9426	1.0519
2015	-0.8447	0.9536	0.8911	-0.9479	1.0701
Mean	-0.8431	0.9303	0.9128	-0.9240	1.0198

Table 14: Social welfare function weights with quadratic quantities and month fixed effects in the demand specifications.