Alcohol tax design

Rachel Griffith, Martin O'Connell and Kate Smith*

October 24, 2016

Abstract

[PRELIMINARY - WORK IN PROGRESS - PLEASE DO NOT CITE]

We study optimal taxation in the alcohol market. We allow for marginal externalities that vary with total ethanol demand, and therefore across consumers, and we consider consumer demand over the differentiated products in the market. We show that differentiating tax rates across alcohol products can improve on the Diamond (1973) optimal tax rate applied to ethanol. The product level system exploits correlation in consumers' demands for individual products with their overall ethanol demand. By specifying and estimating an empirical model of consumer behaviour in the market we use our optimal tax results to numerically solve for the optimal system. We find that moving to an optimal system that differentiates rates across different alcohol types (allowing, for instance, for different rates on vodka, gin, ale and so on) would close 49% of the welfare gap between the UK system and the first best.

Keywords: externality, corrective taxes, alcohol
JEL classification: D12, D62, H21, H23
Acknowledgements: The authors would like to gratefully acknowledge financial support from the European Research Council (ERC) under ERC-2009-AdG grant agreement number 249529, ERC-2015-AdG grant agreement number 694822 and the Economic and Social Research Council (ESRC) under the Centre for the Microeconomic Analysis of Public Policy (CPP), grant number RES-544-28-0001.

*Griffith is at the Institute for Fiscal Studies and University of Manchester, O'Connell is at the Institute for Fiscal Studies and Smith is at the Institute for Fiscal Studies and University College London. **Correspondence:** rgriffith@ifs.org.uk, martin_o@ifs.org.uk and kate_s@ifs.org.uk.

1 Introduction

The potential for taxes to improve welfare when consumption generates externalities motivates many government interventions, including those designed to alter the consumption of alcoholic drinks. Pigou (1920) showed that in the presence of externalities policy should aim to raise price to equal the marginal external cost of each unit consumed. If the marginal externality is constant and equal across consumers, Pigovian taxation can fully correct for the externality. Diamond (1973) showed that when the marginal externality varies across consumers the optimal policy involves setting the tax rate equal to the demand slope weighted average marginal externality (henceforth, "Diamond taxation"). In this case a single linear tax rate can no longer achieve the first best allocation. We study the problem of corrective tax setting in the alcohol market. While it is the consumption of ethanol (pure alcohol) that generates externalities, ethanol is bundled together in products with other characteristics (e.g. alcohol type – vodka, gin and so on – alcoholic strength) and consumers have preferences over all these characteristics. This generates the possibility that consumer preferences over alcohol products and their overall ethanol demand may be correlated, creating the potential for tax design that can improve on Diamond taxation of ethanol.

Our first contribution in this paper is to extend the literature on optimal taxation in the presence of externalities to markets comprising many differentiated products. We consider a social planner that sets taxes to maximise the sum of consumer surplus and tax revenue, minus the external costs of consumption. Consumers choose between many products in the market – their choices lead to their derived demand for ethanol. The external cost of each consumer's consumption is a convex function of their derived ethanol demand – convexity implies the marginal externality of consumption is increasing in derived ethanol demand. Our model nests the classical results in Pigou (1920) and Diamond (1973).[1] If the marginal external cost is constant and equal across all consumers, or if consumer specific taxation is permitted, the first best allocation can be achieved through Pigovian taxation. If the marginal externalities are heterogeneous and the planner is constrained to set a single ethanol tax rate the optimal policy is Diamond taxation. However, if the planner can differentiate tax rates across alcohols of different types, she is able to improve on Diamond taxation. By setting relatively high tax rates on alcohol products that are disproportionately consumed by heavy drinkers the planner is able to target specifically the consumption of high externality generating consumers. The effectiveness of this strategy is increasing in the correlation in demand slopes and marginal externali-

[1] We abstract from revenue raising concerns, but it is straightforward to incorporate a revenue constraint in the model, as in Sandmo (1975).

ties – taxes are more effective when they induce large demand reductions among the high externality consumers – but decreasing in the correlation between demand cross slopes and marginal externalities – taxes are less effective when high externality consumers are very willing to switch their demands to alternative alcohol products.

Our second contribution is to show that these theoretical results have empirical relevance. We apply the model to the UK market for alcoholic beverages, exploiting longitudinal micro data on alcoholic drinks purchased for consumption at home – which account for 77% of total ethanol consumption in the UK. Using these data, we estimate a discrete choice model that is well suited to capturing substitution patterns across products within the market, as well as substitution out of the market. We identify demand parameters using a control function approach, which allows us to isolate price variation driven by cost shifters. We are careful to flexibly specify the distribution of preferences, enabling us to capture correlations in preferences for product characteristics with total derived ethanol demand. We specify the preference distribution as a mixed-normal distribution, exploiting pre-sample information on long run drinking behaviour, and we compare it to a normal preference distribution specification that is more standard to the literature. The demand estimates capture both how consumers respond to changes in the tax system and generate derived ethanol demand, which is the input to the external cost function.

Our empirical estimates show significant variation in demand behaviour, which is correlated with ethanol demand (and hence marginal externalities). Consistently heavy drinkers tend to have relatively strong preferences for strong spirits (e.g. gin, vodka and whisky) and tend to have relatively large cross price elasticities, indicating that they are more willing than lighter drinkers to switch between alcohol products in response to price changes. Despite this higher willingness to switch between alcohol products, heavy drinkers' strong preferences for alcohol means that their ethanol demand is substantially *less* price responsive than lighter drinking households. We use these demand estimates to recover optimal tax rates. Our results show that optimally set alcohol taxes can result in substantial welfare gains relative to the UK system. Moving from the UK system to an optimal single ethanol tax rate (i.e. Diamond taxation) would result in a welfare gain of £0.5 billion and would close 21% of the welfare gap between the UK system and the first best (consumer specific Pigovian taxation). Moving to an optimal system that differentiates rates across different alcohol types (allowing, for instance, for different rates on vodka, gin, ale and so on) would do substantially better, closing 49% of the welfare gap between the UK system and the first best. The optimal rates are higher than current UK taxes on some alcohols (e.g. the set of strong spirits and cider) but lower on others (e.g. ale, stout and flavoured alcoholic beverages). We also show the strong cross price elasticities of heavy drinkers acts to lower optimal rates, as predicted by the theory.

The principal objective of policy intervention in alcohol markets is to reduce the external costs associated with drinking, including the costs of crime (e.g. drink driving, anti-social behaviour, domestic violence) and the public health care cost of treating alcohol related illnesses. Evidence suggests that the external costs of drinking are highly concentrated in a small proportion of heavy drinkers and the marginal external cost is increasing in consumption (see World Health Organization (2014) and Cnossen (2007) for surveys).[2] Nonlinearity in the external cost of consumption means, in lieu of consumer specific taxes, taxation cannot achieve a first best allocation.

Throughout the paper we assume the externality function is positive and convex – a marginal increase in ethanol consumption has associated with it an incremental cost to society that is increasing in the level of ethanol consumption. This is consistent with the bulk of evidence on the external costs of alcohol consumption. We calibrate the function so that it implies an aggregate external cost consistent with official government estimates (UK Cabinet Office (2003)). While there is considerable evidence that the external costs of consumption are convex, there is less evidence of the degree of convexity. In our central calibration we assume convexity consistent with evidence in Taylor et al. (2010) based on the relationship between the amount of ethanol consumed and the probability of having an accident. However, we show robustness of our results to both the calibrated aggregate externality and the degree of convexity; a larger aggregate externality implies higher tax rates across all alcohol products, a more convex function implies a larger tax differential between the products for which heavy drinkers have relatively strong preferences and those for which they have relatively weak preferences.

Specifying the externality function in this way is a parsimonious and appealing way of capturing the welfare costs of drinking, but it is not without limitations. As the argument in the function we use the weekly derived ethanol demand of a household per adult. This reflects the fact that our data is for households rather than individuals. It also means we implicitly assume that the external cost generated by someone consuming 14 units (equivalent to 140ml) of ethanol on Friday night is the same as a person consuming two units every day of the week. We do show however that weekly ethanol consumption and the propensity to binge drink are very strongly correlated, suggesting that our measure does a reasonable job at capturing the first order external costs of drinking.

[2]We use the phrase "externality" to mean all costs that are not taken into account by the consumer at the time of purchase. This may include costs that are borne by the individual at some point in the future, but of which they are not aware at the time of consumption. For instance, O'Donoghue and Rabin (2000) document evidence that decisions made by young people may be based on poor information about the risks involved. This may also lead to suboptimally high consumption and provide a rationale for government intervention.

The framework that we develop is well suited to other applications in which there are heterogeneous consumption externalities in differentiated product markets. For example, concern about obesity and the excess consumption of sugar has led to growing interest in sugar taxes. In this case it is likely the marginal external costs of consumption are heterogeneous across people. For instance, the is particular concern about the consumption of children. If there is correlation between the preferences for different soda products and the marginal externality of sugar consumption, then application of our model would shed light on the design of sugar taxes that reduce the externality while minimising the reduction in consumer surplus.

A number of papers apply continuous choice demand methods to alcohol. These papers either treat alcohol as a homogenous composite commodity (see, inter alia, Deaton and Muellbauer (1980), Blundell et al. (1993), Baltagi and Griffin (1995), Manning et al. (1995), Banks et al. (1997)) or they estimate demand over a set of broad alcohol types (e.g. Crawford et al. (1999), Irvine and Sims (1993) and Purshouse et al. (2010)). In contrast to these papers, we are interested in capturing substitution patterns between disaggregate alcohol products. We use a discrete choice demand framework, avoiding the econometric problems arising from zero purchases (Lee and Pitt (1986), Pudney (1989)) of many of the products available in the market. Modelling consumer choice *between* alcohol products is important for two reasons. First, taxes often change the *relative* prices of different alcohol products, which leads to substitution between them in a way that affects derived demand for ethanol. Consider, for instance, a shopper that typically purchases a bottle of wine each week; if the price of the bottle of wine rises by £1 and the price of a pack of beer falls by £1 and she substitutes to beer, then her derived demand for ethanol changes (beer has less ethanol than wine), but the average price of the two products has not changed. Capturing this kind of substitution is key for evaluating the impact of policies that change relative prices. The second reason is that it allows for the design of taxes that vary across products, which we show can better target the externality costs of drinking, relative to a single tax rate that is common across all products in the market.

Like us, Miravete et al. (2016) and Conlon and Rao (2015) use discrete choice methods to study demand for alcohol. These papers focus on modelling demand in the spirits segment and consider how government regulations, in part designed to limit alcohol consumption, interact with firm conduct. Miravete et al. (2016) show strategic behaviour among distilleries can partially undo the policy objective of the public monopoly that runs alcohol stores in Pennsylvania – for instance, if the public monopoly increases the mark up it sets on alcohol products with the intention of lowering alcohol consumption by

10%, the strategic response of wholesalers would mean consumption would fall by 7.7%.[3] Conlon and Rao (2015) show that post and hold regulations operating in Connecticut – that require wholesalers to post their prices in advance without discriminating across retailers – result in higher retail prices than would otherwise be the case, and that higher levels of alcohol tax could instead be used to raise the price level and would have the advantage of raising tax revenue. In this paper we focus on modelling demand in the entire alcohol market (rather than the spirits segment alone), and we consider the design of optimal policy (rather than the effects of existing policy). However, unlike these papers, we abstract from the supply side of the market. In particular, we consider a social planner that sets taxes to maximise consumer surplus plus tax revenue minus external costs – this captures the realistic situation in which the government is solely concerned with correcting for the external costs of consumption, without reference to any potentially positive mark ups associated with imperfect competition. Due to the large number of products in the alcohol market we also, for UPCs with similar product attributes, aggregate over brands and we assume full pass-through of (specific) taxes to consumer prices. This is an assumption that is unlikely to hold perfectly in reality. For instance Miravete et al. (2016) find pass-through of a regulated downstream mark up (analogous to an ad valorem tax) is less than full, although theoretical results (e.g. Anderson et al. (2001)) point to higher pass-through of specific than ad valorem taxes. We leave as an important avenue for future research combining modelling optimal alcohol taxation with the supply side of the market.

The rest of the paper is structured as follows. In the next section, we set out the general framework for deriving tax rates to correct non-linear consumption externalities in differentiated product markets. In Section 3 we describe the UK market for alcohol, along with our purchase data. We show that consumers that consistently purchase large amounts of ethanol systematically buy different alcohol products from lighter drinks. Section 4 focuses on the details of our empirical model. It details the specifics of our demand model, how we incorporate the external costs of drinking, identification of model parameters and our demand estimates. In Section 5 we present our results on the optimal design of alcohol taxation. Section 6 shows our broad results are robust to the specifics of the calibration of the externality function. A final section summarises and concludes.

[3]Seim and Waldfogel (2013) show the number of stores operated by the public monopoly in Pennsylvania can better be rationalised by a profit maximising motive than welfare maximisation.

2 Corrective tax design

2.1 Model set-up

Consumer demand

Let $i = 1, \ldots, N$ index consumers, $j = 1, \ldots, J$ index alcohol products in the market. Each consumer has y_i of disposable exogenously given income. He decides which, if any, alcohol product to purchase and correspondingly how much income to allocate to a separable bundle of all other consumption goods (the outside good). $j = 0$ denotes the choice to purchase no alcohol products and allocate all income to other consumption. Let $\mathbf{p}_i = (p_{i1}, \ldots, p_{iJ})'$ denote the vector of (post-tax) prices the consumer faces for alcohol products and normalise the price of the outside good to 1. As a benchmark we will consider consumer specific taxes – it is for this reason we write prices with an i index.

Consumer preferences are defined over characteristics of products, both observed (Gorman (1980), Lancaster (1971)) and an unobservable characteristic (Berry (1994), Berry et al. (1995)). These characteristics include the ethanol content of the product, denoted z_j; let $\mathbf{x}_j$ denote a vector of all other product characteristics, which are described in more detail in Section 4.1. The utility the consumer obtains from selecting a product is a function of his valuation of these characteristics and an idiosyncratic utility shock. We assume the consumer selects the option that provides him with the highest utility. We argue in Section 3 that this discrete choice demand framework is well suited to capturing demand in the alcohol market. For instance it rationalises zero purchases and, due to the mapping of preferences into attribute space, does not suffer from the curse of dimensionality of continuous choice demand models.

Integrating out the shocks to utility, the demand framework yields an expected indirect utility function for each consumer, $V_i(y_i, \mathbf{p}_i, \mathbf{z}, \mathbf{x})$ where $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_J)'$ and $\mathbf{z} = (z_1, ..., z_J)'$, and a set of conditional choice probabilities, or demand functions, that describe the probability a consumer selects a given product as a function of product attributes (including prices) and his preferences. Denote the demand function for product j by $q_{ij} = f_{ij}(y_i, \mathbf{p}_i, \mathbf{z}, \mathbf{x})$ and let $\mathbf{q}_i = (q_{i1}, \ldots, q_{iJ})'$ denote the vector of demands for consumer i.

We assume that utility is quasi-linear in the numeraire good, which implies no income effects for demand for alcohol products (and therefore $\mathbf{q}_i = f_i(\mathbf{p}_i, \mathbf{z}, \mathbf{x})$). In our empirical application we allow for rich heterogeneity in the marginal utility of income, allowing demands to vary flexibly across consumers. Therefore the restriction of quasi-linear utility only rules out that price changes induced by taxes induce empirically important income effects. Given our focus on the design of corrective taxes in the alcohol market and given that alcohol spending is a modest share of total income, this assumption is natural. Expected indirect utility can therefore be written:

$$V_i(y_i, \mathbf{p}_i, \mathbf{z}, \mathbf{x}) = \alpha_i y_i + v_i(\mathbf{p}_i, \mathbf{z}, \mathbf{x}), \tag{2.1}$$

where α_i is the marginal utility of income and $v_i(\mathbf{p}_i, \mathbf{z}, \mathbf{x})$ is the expected utility that arises from the consumers' alcohol demands $f_i(\mathbf{p}_i, \mathbf{z}, \mathbf{x})$.

External costs

Alcohol consumption generates a number of costs that are, in general, not considered by individuals when making alcohol consumption decisions; this leads to excess consumption from a social perspective, and justifies government intervention.[4] We use the term external costs to include all costs that are not taken into account by the individual when making their purchasing decision. These include costs that fall on others (e.g. victims of drink driving and alcohol-fuelled violence and the cost burden on publicly funded services such as the police and health service), and those that might fall on the individual in the future (e.g. increased risk of disease), of which they do not take account when making their purchase decision.

We assume that the external costs are a convex function, $\phi_i(\cdot)$ of derived ethanol demand. Derived ethanol demand, Q_i, is a function of the consumer's demand for all the different products in the market and is given by $Q_i = \sum_j z_j q_{ij}$. The total external cost from all consumers in the market is:

$$\Phi = \sum_i \phi_i (Q_i) . \tag{2.2}$$

There is a large amount of evidence which suggests externalities are convex in alcohol consumption. For example, there is evidence of a threshold effect with some diseases: at low levels of alcohol consumption, the risk of disease is not elevated, but this risk increases sharply above a certain point (see Lönnroth et al. (2008) for evidence on tuberculosis,

[4]The World Health Organization (2014) estimate that 5.9% of global deaths and 5.1% of the global burden of disease and injury (measured in disability adjusted life years) is attributable to alcohol. Some diseases in particular are strongly linked to alcohol consumption: in the UK, roughly 70% of cases of liver cirrhosis are attributable to alcohol. Figures from the British Crime Survey (2002) suggest that almost half of violent crime is alcohol-related (Prime Minister's Strategy Unit (2004)). In the UK and Ireland, around a third of domestic violence occurs when the perpetrator is under the influence of alcohol (Mirrlees-Black (1999), Watson and Parsons (2005)). Wells et al. (2005) find that excess alcohol consumption increases the risk of getting into fights after drinking. In the UK, the alcohol-attributable fraction of road traffic deaths is 16.6% for men and 6.7% for women (World Health Organization (2014)).

and Rehm et al. (2010) for evidence on liver cirrhosis).[5] Taylor et al. (2010) find that higher levels of alcohol consumption create an exponential increase in risk of accidents: the odds of injury at 140g of ethanol (roughly 8 pints of beer) are almost 18 times greater than the odds of injury at 14g of ethanol (around 0.5l of beer). In the UK, around 60% of male prisoners and almost 40% of female prisoners are hazardous drinkers, compared with around 30% of male and at most 10% of female general hospital patients (Singleton et al. (2003)).

We base the externality function on the consumer's weekly derived ethanol demand, which is a function of their conditional choice probabilities for alcohol products. This means if a consumer is entertaining visitors in a particular week and so has unusually high alcohol demand (which is captured by idiosyncratic shocks to utility) this will not feed into them generating an unusually high level of externality. Given that we want to capture the fact that externalities are increasing in an individual's consumption, this is sensible. A possible concern with modelling weekly demand for alcohol is whether we miss some externalities created by a consumer engaging in uncharacteristic binge drinking. In Section 4.2, we show that binge drinking and having high average ethanol purchases over a longer period are highly correlated, while Wells et al. (2005) find that the correlation between binge drinking ("heavy episodic drinking") and the likelihood of fighting after drinking became insignificant once drinking frequency and drinking volume are taken into account.

Note that we assume, conditional on total ethanol demand, that the marginal externality from drinking a unit of ethanol (10ml, or equivalently 8g, of ethanol) is the same across different types of alcohol. This of course does not rule out the possibility that people whose ethanol consumption has a high marginal externality (i.e. those that drink a lot), might consume different alcoholic beverages. Indeed, this is precisely the variation that will allow for the design of a system of alcohol taxes that improves on taxation of ethanol. Rather the assumption is that, conditional on a given level of derived ethanol demand, consuming an extra unit of ethanol in the form of beer or in the form of spirits has the same marginal external cost.

Consumers ignore the externality when making their choices, and the goal of the social planner is to use taxes to induce consumers to internalise the externality, while minimising the reduction in consumer surplus that arises due to the higher prices.

[5] Some evidence even suggest modest positive health benefits at low levels of drinking (Peele and Brodsky (2000), Corrao et al. (2000), National Institute on Alcohol Abuse and Alcoholism (2000))

Social planner's problem

We consider the social planner's problem of choosing alcohol taxes to maximise total consumer welfare. We consider specific taxes levied on ethanol content. Total consumer welfare consists of the sum of consumers' indirect utilities (given by equation 2.1) plus revenue raised from tax, R, minus the total external costs of consumption (given by equation 2.2).[6] We make two important assumptions about the planner's problem. First, we write the objective function in money metric form. This means we abstract from any questions of redistribution, focusing exclusively on the design of taxes to correct externalities. Typically alcohol taxes are applied as excise taxes and are in addition to broad based sales or value added taxes. The justification for alcohol taxes over and above broad sales or value added taxes is as a means to correct suboptimally high consumption. Second, the objective function is based on consumer (and not producer) surplus. The planner takes pre-tax prices as given and makes no attempt to correct for the existence of any mark-ups associated with imperfect competition.

Let $\boldsymbol{\tau}$ denote a vector of taxes rates levied per unit of ethanol. We assume tax changes are passed directly to consumer prices and that non-price product characteristics do not change as a result of the tax; we therefore write indirect utility, tax revenue and the externality function as functions of $\boldsymbol{\tau}$. The consumer welfare function is:

$$W(\boldsymbol{\tau}) = \sum_i \left[y_i + \frac{v_i(\boldsymbol{\tau})}{\alpha_i} \right] + R(\boldsymbol{\tau}) - \Phi(\boldsymbol{\tau}). \tag{2.3}$$

2.2 Characterising tax policy

We begin by showing how the results of Pigou (1920) and Diamond (1973) can be derived as special cases in our model. Our main result is to show that, in general, if the planner is constrained to set linear tax rates that are the same across consumers, then the second best policy prescribes rates that vary across products. In general, this gets closer to the Pigouvian first best than the Diamond prescription of an optimal ethanol tax rate that is constant across products and that is equal to the demand slope weighted average marginal externality of ethanol consumption. We also derive the optimal product taxes under simplifying assumptions to provide intuition for how correlation between the marginal externality and shape of demand for products affects the optimal tax rates.

[6] We do not specify how tax revenue is used, but we do assume a dollar of tax revenue is valued the same as an additional dollar of income by consumers. A standard assumption consistent with this is that tax revenue is redistributed lump sum to consumers. An alternative assumption is that tax revenue is used to fund a public good that has total value to consumers equal to public good expenditure.

Pigouvian taxation

Suppose the planner can set separate tax rates for each consumer. Let τ_i denote the tax rate for consumer i and $\boldsymbol{\tau} = (\tau_1, ..., \tau_N)'$. Let $Q_i(\tau_i) = \sum_j z_j q_{ij}(\tau_i)$ denote the derived ethanol demand of consumer i when facing tax rate τ_i. The taxes are levied per unit of ethanol in the product; in this case post-tax prices are given by $p_{ij} = \tilde{p}_j + \tau_i z_j$, where $\tilde{p}_j$ denotes the pre-tax price, and tax revenue is $R(\boldsymbol{\tau}) = \sum_i t_i \sum_j z_j q_{ij}(\tau_i)$. Taking the first order condition for τ_i and applying Roy's identity, $q_{ij} = -\frac{1}{\alpha_i}\frac{\partial v_i}{\partial p_{ij}}$, yields the familiar Pigouvian tax result:

$$\tau_i^* = \phi_i'\left(Q_i(\tau_i^*)\right). \tag{2.4}$$

The optimal consumer specific tax rate is set to equal the consumer's marginal consumption externality evaluated at that tax rate. If consumer specific taxes are feasible it is possible to fully correct for the consumption externality and hence achieve the first best.

Note that an implication of this is, if the externality function is the same across consumers, $\phi_i(\cdot) = \phi(\cdot)$, then the first best outcome could also be achieved by a single tax schedule that is nonlinear in the derived ethanol demand. The schedule that achieves this is such that the marginal tax rate is equal to the marginal externality: $\tau'(.) = \phi'(.)$.

Diamond taxation

Now suppose the planner is unable to set consumer specific taxes, and must instead choose one tax rate to apply to all products that is common across consumers. Specifically, consider a tax that leads to post-tax prices: $p_j = \tilde{p}_j + \tau z_j$. Revenue in this case is given by, $R(\tau) = \tau \sum_i \sum_j z_j q_{ij}(\tau)$ and derived ethanol demand is $Q_i(\tau) = \sum_j z_j q_{ij}(\tau)$. Denote the slope of demand for Q_i as $Q_i' \equiv \frac{\partial Q_i}{\partial \tau} = \sum_j z_j \frac{\partial q_{ij}(\tau)}{\partial \tau}$ and the slope of the externality function as $\phi_i' \equiv \phi_i'(Q_i)$. Taking the first order condition with respect to τ and applying Roy's identity yields the following expression for the optimal ethanol tax rate, τ^*:

$$\begin{aligned}\tau^* =& \frac{\sum_i \phi_i' |Q_i'|}{\sum_i |Q_i'|} \\ =& \overline{\phi}' + \frac{\text{cov}(\phi_i', |Q_i'|)}{|\overline{Q}'|},\end{aligned} \tag{2.5}$$

where $\overline{\phi}' \equiv \frac{1}{N}\sum_i \phi_i'$ and $\overline{Q}' \equiv \frac{1}{N}\sum_i Q_i'$ are the average marginal externality and ethanol demand slopes and we use $|.|$ to denote the absolute value. The first line is Diamond's (1973) formulation that the optimal tax rate equals the weighted marginal externality, where the weights are the slopes of demand. The second line is an alternative expression, which states that the optimal commodity tax is equal to the (unweighted) average marginal externality plus an adjustment based on the covariance of the marginal externality and (absolute value of) the slope of derived ethanol demands – all else equal, the more positively correlated are ethanol demand slopes and marginal externalities, the higher the optimal ethanol tax. The intuition is that the more consumers that generate high levels of externalities have ethanol demands that are particularly sensitive to tax changes, the more effective is the tax at lowering particularly costly consumption.

Second best product taxation

Now suppose the planner can set a separate tax rate for each product in the market. The planner chooses the vector of taxes $\boldsymbol{\tau} = (\tau_1, ..., \tau_J)'$, with post-tax prices given by $p_j = \tilde{p}_j + \tau_j z_j$. Revenue in this case is given by $R(\boldsymbol{\tau}) = \sum_i \sum_j \tau_j z_j q_{ij}(\boldsymbol{\tau})$ and derived ethanol demand is $Q_i(\boldsymbol{\tau}) = \sum_j z_j q_{ij}(\boldsymbol{\tau})$. Taking the first order condition for τ_j yields:

$$\sum_i \sum_k (\tau_k - \phi_i'(Q_i(\boldsymbol{\tau}))) z_k \frac{\partial q_{ik}}{\partial \tau_j} = 0 \tag{2.6}$$

The set of conditions across products $j = 1, ..., J$ implicitly define the set of optimal product taxes. In general, the optimal product taxes will depend on the full matrix of own and cross price effects and their correlation with the marginal externality.

To obtain some intuition for this condition consider three special cases. First, suppose the marginal externality is constant (and therefore independent of consumption), so $\phi_i' = \phi'$. In this case the optimal tax rate on each product is the same and equal to the marginal externality, $t_j^* = \phi'$ for all j, and we have Pigouvian taxation and the first best outcome.[7]

Second, suppose that cross price effects between alcohol products are zero, so $\frac{\partial q_{ij}}{\partial \tau_k} = 0$ for all $k \neq j$. Denote the ethanol demand of consumer i from good j as $Q_{ij} = z_j q_{ij}(\tau_j)$ (and the slope of this demand $Q_{ij}' = z_j \frac{\partial q_{ij}}{\partial \tau_j}$) and his marginal externality $\phi_i' \equiv \phi_i'(Q_i(\boldsymbol{\tau}))$; we can then write the optimal tax on good j as:

$$\begin{aligned} \tau_j^* =& \frac{\sum_i \phi_i' |Q_{ij}'|}{\sum_i |Q_{ij}'|} \\ =& \overline{\phi}' + \frac{\text{cov}(\phi_i', |Q_{ij}'|)}{|\overline{Q}_j'|} \end{aligned} \tag{2.7}$$

In this case optimal taxation reduces to Diamond taxation but on a product by product basis. Products with a strong positive covariance between demand slopes and externalities tend to have higher optimal taxes. If the correlation between demand slopes and the

[7]Note, due to differences in ethanol contents across products, the amount of tax levied on each product, equal to $\phi' z_j$, will vary.

marginal externality is zero, then differentiating rates across products cannot improve on Diamond taxation. Notice, though, that even in the case of no cross price effects, the optimal tax rates for products are linked through their common effect on the externality function.

A useful alternative way to write condition (2.7) is in terms of consumers' contribution to total product demand and the own price demand elasticity. If we define consumer i's contribution to total ethanol demand from good j as $w_{ij} = \frac{Q_{ij}}{\sum_i Q_{ij}}$ and his own price elasticity for good j as $\varepsilon_{ij} = \frac{\partial q_{ij}}{\partial p_j}\frac{p_j}{q_{ij}}$, we can re-write condition 2.7 as:

$$\tau_j = \widehat{\phi}'_j + \frac{\widehat{\text{cov}}(\phi'_i, |\varepsilon_{ij}|)}{|\widehat{\varepsilon}_j|} \tag{2.8}$$

where

$$\widehat{\phi}'_j = \sum_i w_{ij}\phi'_i, \quad \widehat{\varepsilon}_j = \sum_i w_{ij}\varepsilon_{ij}, \quad \widehat{\text{cov}}(\phi'_i, \varepsilon_{ij}) = \sum_i w_{ij}(\phi'_i - \widehat{\phi}'_j)(\varepsilon_{ij} - \widehat{\varepsilon}_j)$$

This formulation shows that, in the absence of cross price effects, the optimal choice of tax rate for good j depends on two interpretable terms. The first is the weighted average marginal externality, where consumers are weighted by their contribution to total demand for good j. All else equal, if a good has relatively high demand among consumers who generate a high externality it should attract a higher tax rate. The second term is the weighted covariance of consumer level marginal externalities and own price elasticities, scaled by the average own price elasticity. All else equal, the more positive is the correlation between the marginal externality and the (absolute value of) the elasticity of demand, the higher should be the tax rate. Taxes should therefore be relatively high on products for which consumers generating high externalities are most willing to switch away in response to a price rise.

Third, consider the case in which there are two alcohol goods that are substitutable (and hence have a positive cross price effect). In this case the optimal product tax rates can be expressed as (implicit) functions of average demand slopes, cross slopes and the correlations of demand slopes and cross slopes with marginal externalities – see Appendix A. All else equal, increasing the covariance between the marginal externality and absolute value of the own slope of demand for good i increases the optimal tax on good i and the optimal tax on good j but by less than for good i. As above, if consumers that generate large externalities have steeply sloped demands, tax is more effective at lowering their demands and at the optimum the rate is higher. On the other hand, increasing the covariance between the marginal externality and the cross slope of demand between goods i and j decreases the optimal rate on both goods, and decreases by more the optimal tax on the good with the largest own slope of demand. Given the correlation between the price sensitivity of own demands and externalities, the greater is the correlation between cross price effects between the two goods and the externalitiy, the less responsive is the derived ethanol demand of the high externality consumers to tax. This makes tax less effective, acting to lower optimal rates.

In our empirical application we allow for cross price effects across all goods in the alcohol market and we allow demands to vary flexibly across consumers with different marginal externalities. We use our model to solve conditions (2.6) in order to recover the optimal product taxes.

3 UK alcohol market

Our application is to the UK off-trade market for alcoholic beverages. In this section we describe our main data source, some key features of the market that influence how we model demand and some broad purchase patterns that provide some insight into what products in the market are likely to attract relatively high tax rates. In the following section (Section 4) we provide details of our empirical model.

Our data contain comprehensive information on purchases of alcohol off-trade. Off-trade alcohol purchases consist of purchases of alcohol products made in stores (including supermarkets, corner stores and off-licenses[8]). Off-trade alcohol purchases constitute 77% of ethanol purchased in alcoholic drinks in the UK.[9] We show that the patterns of alcohol purchases, and crucially how it varies with total ethanol demand, is similar for off- and on-trade alcohol.

Our data are from the Kantar Worldpanel – a panel of households selected to be representative of the UK (excluding Northern Ireland) population. Each participating household uses a hand held scanner to record all products, at the UPC level, purchased from grocery stores and off-licenses that are brought into the home. The data include details of exact transaction prices and product size, alcohol type and strength.[10] For a more detailed description of the data, see Griffith and O'Connell (2009) and Leicester and

[8]The term used in the UK to describe stores solely dedicated to the sale of alcohol.

[9]Calculated using the *Living Costs and Food Survey (2011).*

[10]Strength is measured as percentage alcohol by volume (ABV). This is defined as the number of millilitres of pure ethanol present in 100ml of solution at 20°C. ABV is recorded for beer, cider and alcopops. For wine and spirits ABV information is only partly recorded in the data. We use information from retailer and manufacturer websites to fill in missing ABV values. Where we are unable to find the ABV of a product we apply the Office for National Statistics standard assumed alcohol content for drinks of that type, see Goddard (2007).

Oldfield (2009); Griffith et al. (2013) contains more information on the alcohol segment of the data.

We use a sample of 10,289 households, who we observe repeatedly throughout calendar year 2011. Each household is in the data for a minimum of 30 weeks over the course of the year and are observed purchasing alcohol at least 3 times.[11] We also observe alcohol purchases the households made in calendar year 2010. As we estimate our model using data from 2011, we refer to information from 2010 as pre-sample information.

3.1 Purchase patterns

For each household in our sample we compute how much ethanol per adult (aged 18 or over) per week they purchase on average in the pre-sample period. We measure ethanol in units – one unit of ethanol is defined as 10 millilitres (8 grams) of ethanol. In the UK, a standard one drink measure of a strong spirit (e.g. vodka) contains 1 unit of ethanol. We categorise households based on this measure – defining groups based on those that purchased fewer than 7, 7–14, 14–21, 21–35, and above 35 units per adult per week on average. This grouping splits households into sets of light to heavy drinkers. The UK National Health Service (see NICE (2010)) classifies men that regularly drink 21–28 units of ethanol per week, and women that regularly drink 14–21 units per week as "increasing-risk drinkers". Those that regularly drink in excess of these amounts are classified as "higher-risk drinkers".

In Table 3.1 we show the fraction of households in our sample that belong to each group, along with some broad measures of alcohol purchases made by households in each group. The table shows the distribution of households across pre-sample purchase groups is highly skewed – most households tend to purchase a moderate number of units of ethanol, however, 7.2% of households consistently purchased between 21-35 units of alcohol and 4.5% of households consistently purchased in excess of 35 units per adult per week. While they account for a relatively small fraction of households, the consistently heavy purchasers account for a much higher fraction of alcohol – in households that consistently purchased more than 21 units of alcohol in the pre-sample period comprise 42% of all 2011 ethanol purchases. Column 3 shows that households' pre-sample ethanol purchases and ethanol purchased in sample (in 2011) is very strongly related; households that are low, moderate and heavy purchasers of alcohol in the pre-sample period continue to be low, moderate and heavy in the sample period. As we make clear in Section 4.1, in demand estimation we use the pre-sample purchase groups as a conditioning variable

[11] 4,477 households are observed for a minimum of 30 and either do not purchase alcohol or do so only once or twice. These households account for less than 1% of total alcohol purchased and therefore have no impact on our empirical implementation of optimal corrective taxes.

for the preference distribution. Incorporating pre-sample information that is informative in explaining behaviour allows us to incorporate additional flexibility into the demand model.

Column 4 shows that households that consistently purchase a large amount of ethanol tend to purchase a higher weekly volume of alcohol products (i.e. more litres of alcohol products). This means the higher amount of total ethanol purchased by the groups of heavier drinkers is not driven entirely by heavy drinkers purchasing stronger drinks. However, column 5 shows it is, to some extent, driven by heavy drinkers tending to purchase stronger drinks. We use alcohol by volume (ABV) to measure alcohol strength – defined as millilitres of ethanol per 100 millilitres of volume at $20°C$. The average ABV of alcohol products purchased by households in the more than 35 units group is 13.8%, while it is 10.1% for the group of lightest drinkers. The final column of Table 3.1 shows the average price paid per unit of ethanol by households in each purchase group showing that, on average, heavier drinkers purchase alcohol products that are cheaper in per unit of ethanol terms. For instance the heaviest drinkers on average pay 41 pence per unit of ethanol, just 80% of the average per unit price paid by the lightest drinking group.

Table 3.1: *Variation in average attributes of purchases by group*

	(1)	(2)	(3)	(4)	(5)	(6)
	Percentage of		Average weekly		Average	
Household group:	Households	Total ethanol	Ethanol units	Litres	ABV	Price per unit (p)
Less than 7 units	62.5	23.2	2.9	0.3	10.1	51.0
7-14 units	17.9	20.4	9.2	1.0	11.4	46.6
14-21 units	7.9	15.3	15.6	1.7	11.7	44.3
21-35 units	7.2	19.8	23.9	2.4	12.3	42.8
More than 35 units	4.5	21.2	47.1	4.2	13.8	41.0
Total	100.0	100.0	8.6	0.9	10.8	48.6

Notes: For each household we calculate the average number of units and litres purchased per adult per week and the (quantity-weighted) average price paid per unit of ethanol and average strength (ABV). The numbers shown in the table are the average across all households within each household group. Groups are based on average amount of alcohol purchased per adult per week in the year preceding the period of time on which we estimate the model.

In Table 3.2 we show how the share of total units of ethanol accounted for by different types of alcohol products varies across the pre-sample purchase groups. The table shows that heavier drinking groups get a much lower fraction of their alcohol units from the beer (including lager and ale) segment than lighter drinking groups (e.g. 13.5% on average for the greater than 35 units group and 25.2% for lower than 7 unit group). Conversely, heavy drinking groups get a much higher fraction of their total ethanol units from the spirits (including fortified wine) segment than lighter drinking groups (39.5% for the heaviest drinking groups versus 27.3% for the lightest drinking group). The wine segment and cider (including flavour alcoholic beverages (FABs)) segments have a non-monotonic relationship with the drinking groups; the lightest and heaviest drinking groups both get lower shares of their ethanol from wine than other groups, but more from cider and FABs. Within the broad alcohol segments there is also strong relationships between purchase patterns and long run ethanol consumption. For instance, within the spirits segment, heavier drinking households get more of their ethanol from each of brandy, gin, vodka and whisky than lighter drinking households, but less from liqueurs and fortified wines.

Table 3.2: *Variation in share of ethanol from different alcohol types by group*

	Household group				
	<7	7-14	14-21	21-35	>35
Beer (inc. lager and ale)	*25.2*	*21.7*	*21.7*	*20.2*	*13.5*
Ale	6.9	5.9	5.5	4.4	3.0
Lager	17.1	14.9	15.6	15.2	10.3
Stout	1.3	0.9	0.7	0.5	0.3
Wine	*38.3*	*43.1*	*40.9*	*41.5*	*38.7*
Red wine	15.3	20.0	19.3	18.1	16.3
White wine	16.3	17.9	18.1	19.9	19.7
Rose wine	6.7	5.2	3.6	3.5	2.7
Spirits (inc. fortified wine)	*27.3*	*27.7*	*29.1*	*32.2*	*39.5*
Brandy	2.1	2.4	2.6	2.8	3.0
Gin	2.0	2.9	2.9	4.3	5.1
Rum	2.2	2.6	2.1	1.9	2.2
Vodka	4.0	5.1	5.4	5.5	9.1
Whisky	7.4	8.3	10.3	11.4	15.0
Liqueurs	5.5	3.0	2.4	1.9	1.2
Port	1.1	0.6	0.5	0.5	0.7
Sherry	1.5	1.4	1.3	1.8	0.7
Vermouth	0.5	0.6	0.7	0.9	1.0
Other fort. wine	1.0	0.8	1.0	1.4	1.5
Cider (inc. FABs)	*9.1*	*7.4*	*8.2*	*6.1*	*8.2*
Cider	8.1	6.9	7.8	5.8	8.0
Pre-mixed spirits	0.2	0.1	0.1	0.1	0.1
Alcopops	0.8	0.4	0.3	0.2	0.2

Notes: For each group of households we calculate the total number of units of ethanol purchased in 2011 and the number purchased from each type of alcohol. The numbers in the table are the percentage of total units from each type.

Taken together, Tables 3.1 and 3.2 show rich correlations between purchase patterns and a measure of long run average ethanol demand. This points to the possibility that optimally set taxes rates that vary across disaggregate alcohol types may significantly improve on Diamond taxation of ethanol considered as single commodity. In Section 2 we showed that, all else equal, optimal tax rates are likely to be higher on products purchased disproportionately by consumers that generate large marginal externalities. This points towards relatively high tax rates on strong spirits products. However, optimal taxes also depend on the pattern of own and cross price elasticities across products and how this varies across consumers based on the marginal externality they creates. In Section 4.1 we describe a demand model that allows us to flexibly capture substitution patterns across products and how they vary across different consumers based on their total ethanol demand and hence will allow us to compute optimal tax rates.

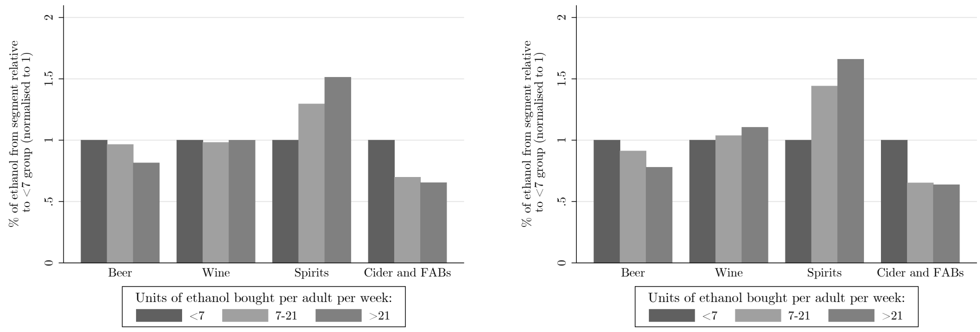
On-trade alcohol

Our data contain very detailed information on purchases of alcohol products off-trade, but they do not contain information on alcohol purchases on-trade. The Living Cost and Food Survey (LCFS) contains information on alcohol purchased both on- and off-trade. These data are based on two week household diaries. Unlike the Kantar data, they do not contain repeated observations for the same households over time, they do not contain product level information and they do not contain transaction prices or any measure of alcohol strength. Nevertheless, we can use these data to get an idea of whether purchase patterns are similar between off-trade alone and on- and off-trade alcohol together.

To do this we impute the strength of the alcohol categories collected in the EFS. For instance, for the category beer we use 4% ABV – the average from the Kantar data. Based on this, in 2011, we compute that 77% of units of ethanol purchased was done so off-trade. We also categorise households into those that purchased fewer than 7, 7-21 and above 21 units of ethanol per adult per week over their two week diary period. In Figure 3.1 we compare the share of their ethanol each group of households get from the main alcohol segments – beer, wine, spirits and cider. We normalise the share for each segment of the group purchasing less than 7 units to 1, and show the share for other groups relative to the group of lightest drinkers. Panel (a) shows numbers for off-trade only and (b) shows numbers based on both on- and off-trade. The figure shows that the pattern of households with relatively large ethanol demands getting a relatively low share of their units from beer and a relatively high share from spirits holds for both off-trade alone and on- and off-trade together. This suggests our focus on off-trade purchases is unlikely to result in a substantially different pattern of optimal taxes across products than would result if we estimated demand including the 23% of ethanol purchased on-trade.

Figure 3.1: *Alcohol purchases: on- and off-trade*

(a) Off-trade only (b) On- and off-trade



Notes: For each group of households we calculate the share of ethanol units that are bought from products in each segment of the market. The bars show the ethanol share for each type for each group of households relative to the group of households that purchase fewer than 7 *units of ethanol per adult per week (normalised to 1). The left hand panel shows off-trade purchases only and the right panel shows on- and off-trade purchases, calculated using the Living Costs and Food Survey 2011.*

3.2 Alcohol products

There are over 7000 distinct alcohol UPCs (barcodes), and almost 3000 alcohol brands, recorded as being purchased in our data. Estimating a demand system where choice sets contain 7000 options is infeasible. A common approach in the industrial organisation literature is to take the top few brands as representative of the market and estimate demand using purchases of these products. However, even if we selected the top 50 brands (which we would wish to subdivide into separate sizes), we would capture less than half of the alcohol market, and only 25% of the cider segment. There are significant differences in the price and alcohol strength of brands that are and are not selected when using this criteria e.g. on average, beer brands that are included based on this criteria are 20% cheaper than those not included.

We therefore define the options in households' choice sets by aggregating the 7000 UPCs into 79 product-size pairs. In doing this we group together UPCs that have similar product attributes, including alcohol strength, alcohol type, branding (branded versus store brand) and pack type and that also have similar movements in price. As an example, the spirits option that contributes the most to the aggregate number of ethanol units purchased is "Whisky; branded c. 1.4l". Five UPCs make up 74% of all expenditure on this option. These UPCs all have 40% ABV and have average prices per unit ethanol in the range of 36-40p, with similar movements over time.

We group all the UPCs into 40 alcohol products. For each product we define a set of sizes based on the total quantity that households buy on a given purchase occasion. An option is therefore a product-size pairing. There are 79 alcohol options in total (plus the outside option) – these are listed in the Appendix in Table B.1. For each option we construct a price index defined over the prices of UPCs which constitute the option. The index weights are held fixed, so movements in the price index reflects only movements in the (average weekly) underlying prices of the UPCs and not changes in the composition of UPCs chosen by households. Table B.1 shows the means of the price index for each option – we use this in demand estimation. It also shows the total number of ethanol units in each option and the alcoholic strength of each product.

We model the alcohol purchase a household makes on a "purchase occasion". We define a purchase occasion as a week in which the household is recorded buying groceries. Alcohol is purchased on 53.4% of purchase occasions. On the remaining purchase occasions households choose the "outside option" of no alcohol.

Our demand framework assumes that households make a series of discrete decisions over which alcohol products to buy. Households typically choose one or a small number of options – on 37% of purchase occasions, households purchase more than one (typically two or three) alcohol products. The large number of zero demands at this granular level means a discrete choice model is a natural way to model consumer demand.[12] On occasions on which a household purchases more than one product, we treat this behaviour as the household making multiple separate purchase decisions. Heuristically we can think of the household making a separate purchase for Friday night and Saturday night, or for two separate members of the household.

3.3 UK tax system

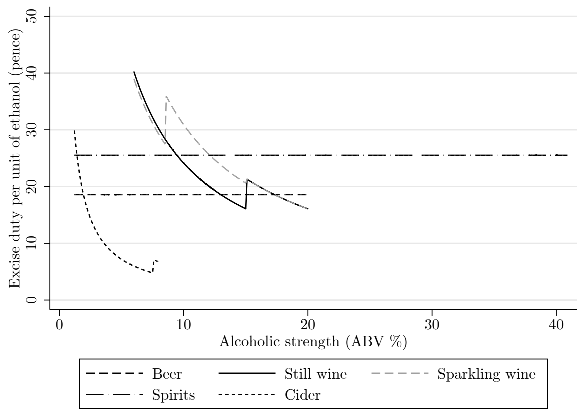
In the UK alcohol is subject to two taxes – an alcohol excise tax, and value added tax (VAT). VAT is a broad based tax which applies to the majority of goods. Alcohol excise taxes are differentiated across broad alcohol segments, with beer, wine, spirits and cider attracting different tax rates. The tax base for beer and spirits is product ethanol content, while the tax base for wine and cider is the volume of the product. The relationship between the post-tax price faced by consumers, p_j and pre-tax price, $\tilde{p}_j$, is:

$$p_j = (1 + t^{VAT})(\tilde{p}_j + t_j^{EX}),$$

[12] A large number of corner solutions (or zero demands) creates econometric problems for the estimation of continuous choice demand models (see Lee and Pitt (1986), Pudney (1989)).

were t^{VAT} is the VAT tax rate, which was 20% in 2011 and t_j^{EX} is the excise tax that applies to product j. For beer and spirits the excise tax is levied on ethanol – i.e. for spirits product j the excise tax is $t_j^{EX} = t^{SPIRIT} z_j$. For wine and cider the excise tax is levied on product volume – i.e. for wine product j the excise tax is $t_j^{EX} = t^{WINE} volume_j$. In Figure 3.2 we show how the excise tax, expressed per unit of ethanol, varies across alcohol segments and with alcohol strength. For wine and cider tax per unit is declining in strength. The kinks in the schedules correspond to different tax bands based on alcohol strength. We show in our optimal tax calculations that consumer welfare can be significantly improved by moving from the current system to tax rates that are set to maximise consumer welfare net of the external costs of drinking (equation 2.3). This is true not only for the system of optimal product taxes, but also if the planner is constrained to set optimal taxes which are common across products within alcohol segment or a single rate applied to all ethanol.

Figure 3.2: *UK system of alcohol excise taxes*



Notes: Fortified wines with ABV below 22% are taxed as wines and those with ABV above 22% are taxes are spirits. FABs are taxed as spirits.

In demand estimation, for each option, we use the post-tax price faced by consumers – i.e. p_j. As explained above we construct this as a price index based on the transaction prices of UPCs that comprise the option. In the optimal tax problem we use the option price after removing VAT and the excise tax as the pre-tax price – i.e. $\tilde{p}_j$. Using demand estimates and the problem outlined in Section 2 we then solve for the set of optimal alcohol tax rates. So, for instance, in the case of optimal product taxes, τ_j^*, the post tax price is

$$p_j = \tilde{p}_j + \tau_j^* z_j.$$

Note, we do not explicitly model the decision over which rate of VAT to set – this will depend on demand across all goods in the economy as well as on labour supply decisions. However, by dividing the set of optimal taxes $\boldsymbol{\tau}^* = (\tau_1^*, ..., \tau_J^*)'$ by the VAT rate – effectively reducing all optimal taxes by the common factor 1.2 – we can obtain the system of optimal excise taxes given the prevailing VAT system. This optimal excise tax system takes account of the fact that a £1 increase in excise taxes translate into a $(1 + t^{VAT})$ price change. As the difference between the optimal system of alcohol taxes and the optimal system of alcohol excise taxes (given a particular VAT rate) is only to shift all taxes by a common factor, in our results we present results on the former.

4 Empirical model

In this section we describe our empirical model, which we use to solve the optimal tax problem outlined in Section 2. We begin by describing our empirical model of demand, the externality function and identification. We then present estimates of demand parameters and price elasticities.

4.1 Empirical model of consumer demand

As in Section 2 we use i to index consumers (i.e. households[13]) and j to index products. $j = 0$ denotes the option of purchasing no alcohol, $j = 1, ..., J$ index different alcohol products. At this point it is useful to introduce two additional subscripts. We model the purchase decision households make in each week of a particular calendar year (2011) – we use t to index weeks. Each product is available in a number of different sizes. We model the decision households make over both which product and which size to select. We use s to index sizes and denote an option in the households' choice set as (j, s); the option to purchase no alcohol is denoted $(0, 0)$.

We assume the utility consumer i obtains from selecting option (j, s) in period t is given by:

$$u_{ijst} = \nu(p_{jst}, z_{js}, \mathbf{x}_{jst}; \theta_i) + \epsilon_{ijst}, \tag{4.1}$$

where p_{jst} be the price of option (j, s) in week t, z_{js} is the ethanol content of the option (j, s), $\mathbf{x}_{jst}$ is a vector of other option characteristics (including a time-varying unobserved attribute), and θ_i is a vector of household level preference parameters. ϵ_{ijst} is an idiosyncratic shock to utility; we assume ϵ_{ijst} is distributed i.i.d. type I extreme value. We normalise the price and product attributes of the choice not to purchase alcohol to zero, meaning the utility from purchasing no alcohol is given by $u_{i00t} = \epsilon_{i00t}$.

[13] We have data on purchases made by households. We assume a unitary model household decision making.

Integrating across the demand shocks, $\boldsymbol{\epsilon}_{it}$, yields conditional choice probabilities, which describe the probability that household i selects option (j, s) in week t, conditional on prices, product attributes and preferences. At the household level the conditional choice probability for option $j > 0, s > 0$ takes the closed-form:

$$q_{ijst} = \frac{\exp(\nu(p_{jst}, z_{js}, \mathbf{x}_{jst}; \theta_i))}{1 + \sum_{j'>0,s'>0} \exp(\nu(p_{j's't}, z_{j's'}, \mathbf{x}_{j's't}; \theta_i))} \tag{4.2}$$

and expected utility is given by:

$$v_{it}(\mathbf{p}_{jt}, \mathbf{z}_{jst}, \mathbf{x}_{jst}) = \ln \sum_{j>0,s>0} \exp\{\nu(p_{jst}, z_{js}, \mathbf{x}_{jst}; \theta_i)\} + C \tag{4.3}$$

where C is a constant of integration which differences out when comparisons are made across two different policy regimes. Equations (4.2) and (4.3) give the expressions for q_{ij} and v_i used in Section 2.

We assume the function ν takes the form:

$$\nu(p_{jst}, z_{js}, \mathbf{x}_{jst}; \theta_i) = \alpha_i p_{jst} + \beta_i w_j + \sum_{m=1}^{4} 1[j \in \mathcal{M}_m] \cdot (\gamma_{i,1m} z_{js} + \gamma_{i,2m} z_{js}^2) + \xi_{ijt}. \tag{4.4}$$

We allow total ethanol content to affect the utility from option (j, s) through a quadratic function with parameters that we allow to vary across the four segments of the alcohol market, beer, wine, spirits and cider – indexed $m = 1, ..., 4$ where $\mathcal{M}_m$ denotes the set of options that belong to segment m. This allows for the possibility that households might value larger or smaller quantities of alcohol differently, depending on what type of alcohol they are buying. We also allow the product's alcohol strength, w_j, and a household specific time varying unobserved product attribute, ξ_{ijt}, to affect the utility from option (j, s). The parameters $(\alpha_i, \beta_i, \boldsymbol{\gamma}_i)$ are all household specific and capture the weight the household places on product price, strength and ethanol content when making their purchase decision. We now discuss how we model the preference distribution over observable attributes, $(\alpha_i, \beta_i, \boldsymbol{\gamma}_i)$, and over the unobservable product attribute ξ_{ijt}.

Preference heterogeneity

We allow for heterogeneous preferences over observable product attributes and over the unobserved product attributes. For the unobserved product attribute, we split it into a time invariant component that varies across households, and a time varying component that varies over alcohol-types. The alcohol-types are defined to be (slightly) more aggregate than the products and we use k_j to denote the alcohol-type product j belongs to.

The 19 alcohol types are shown in Table 3.2.[14] In particular, we assume:

$$\xi_{ijt} = \eta_{ij} + \zeta_{k_j t}$$

Let $\Psi_i = (\alpha_i, \beta_i, \boldsymbol{\gamma}_i)$ denote the household specific preferences over observable product attributes and $\boldsymbol{\eta}_i = (\eta_{i1}, ..., \eta_{iJ})'$ denote the household specific preference over the unobserved product attributes. To close the demand model we need to specify a distribution over each of these. We make two alternative assumptions – the first is similar to the standard distributional assumption made in the discrete choice demand literature, the second is a more flexible assumption that exploits the household specific histories of pre-sample purchases that we observe.

A: Normal preference distribution

As a means of comparing our full demand model to that based on a demand specification more standard to the literature, we estimate a version of our demand model assuming preference distributions are normal. In particular, we assume $\Psi_i \sim \mathcal{N}(\mu, \Omega)$, where we allow Ω, the variance-covariance matrix, to be unrestricted. This allows for the possibility that households have correlated tastes for price, alcohol strength and ethanol content and turns out to be empirically important. For the unobserved product effects we assume $\boldsymbol{\eta}_i \sim \mathcal{N}(\bar{\eta}, \Sigma)$. As there are many products it is necessary to place some restrictions on the covariance matrix Σ. We restrict it to be diagonal and we restrict products within each of the four segments of the market – beer, wine, spirits and cider – to have common variance components. This allows for the possibility that households' willingness to substitute between products in each of these segments differs from their willingness to switch between products in different segments. A substantial literature (including Berry et al. (1995), Nevo (2001)) has shown that making similar distributional assumptions for random coefficients in logit choice models results in the models being able to flexibility capture substitution patterns across products.

B: Mixed-normal preference distribution

The very detailed nature of our data enables us to relax the distributional assumptions commonly made in discrete choice models. Given our application, the most important dimension in which to do this is to allow for flexible correlation in substitution patterns across consumers with different derived ethanol demands (and hence marginal externalities). We do this in a tractable and easily implementable way by exploiting pre-sample

[14] For instance, gin is one alcohol type comprising the alcohol products store brand gin and branded gin, with each product being available in 0.7l and 1l sizes.

purchase histories. In particular, instead of modelling the preference distributions as normal we model them as a mixture of conditional normal distributions, conditioning on households' pre-sample ethanol purchases. If pre-sample behaviour is informative about preferences and in sample behaviour this will enable us to much more flexibly capture the preference distribution. Conversely, if pre-sample behaviour is not informative, the model will collapse to the standard normal preference distribution case.

Specifically, we divide households into $d = 1, \ldots, D$ household groups based on the average amount of ethanol per adult per week purchased over the previous year. These groups are shown in Table 3.1. Let $\mathcal{D}_d$ denote the set of households in group d. In our full model we assume $\Psi_i | i \in \mathcal{D}_d \sim \mathcal{N}(\mu_d, \Omega_d)$ and $\boldsymbol{\eta}_i | i \in \mathcal{D}_d \sim \mathcal{N}(\bar{\eta}_d, \Sigma_d)$. The mean and covariance parameters are then specific to the group of households, d. In this specification we also allow for the possibility that the time varying component of the unobserved effects vary across groups d. This mixture of normals specification is a tractable way of incorporating more information into the model that may be informative about preference distributions and hence allows us to model them more flexibly than is standard. As we show, this added flexibility affects our precise quantitative results on optimal alcohol taxes.

We estimate demand using maximum simulated likelihood based on a sample of 56,250 purchase occasions.[15] Conditional on the draws from the random coefficient distributions, Ψ_i and $\boldsymbol{\eta}_i$, the probability a household selects a given option in a given week takes the closed form of equation 4.2. This follows from our assumption that the ϵ_{ijst} are i.i.d. type I extreme value. To construct the likelihood function we have to integrate across the random coefficient distribution. Let $(1, ..., T_i)$ denote the stream of sampled purchase occasions on which we see decisions of household i and let (j_t^*, s_t^*) denote the option the household chooses on purchase occasion t. The contribution household i makes to the likelihood function is then:

$$l_{it} = \ln \int \prod_{t=(1,...,T_i)} q_{ij_t^* s_t^* t} dF(\Psi_i) dF(\boldsymbol{\eta}_i)$$

As no closed form for the integral exists, we use simulation methods.

4.2 Externality function

In Section 2 we assume that the externality function is a convex function of the consumer's derived ethanol demand. Here we impose the additional assumption that the function is

[15] We randomly draw 450 households from each group d and 25 purchase occasions per household.

the same across households i.e. $\phi_i(\cdot) = \phi(\cdot)$; this means that differences in the external cost generated by different households is due exclusively to differences in their derived ethanol demand. We parameterise the function as an exponential function with two parameters, (ϕ_0, ϕ_1), which allows us to separately calibrate the aggregate external cost and the convexity of the function:

$$\phi(Q_{it}) = \phi_0(\exp(\phi_1 Q_{it}) - 1), \tag{4.5}$$

where $Q_{it} = \sum_{j,s} z_{js} q_{ijst}$ is derived ethanol demand. Subtracting one from the term in the brackets ensures that the external cost of zero ethanol demand is zero. Convexity of the function means that higher levels of ethanol are associated with a higher marginal external cost.

The externality function takes as its argument the total derived ethanol demand of the consumer. As we use data on the household purchases, we convert total ethanol demand into demand per adult (person aged 18 or over). This avoids large households being considered as generating high externalities even when their ethanol demand per adult is not particularly high.

High ethanol demand may be due to consumers drinking large amounts regularly or engaging in less regular very high consumption (binge drinking). Both types of drinking behaviour are likely to lead to externalities, although the nature of these externalities may differ somewhat (e.g. both types of consumption are likely to be bad for health, but binge drinking is arguably more likely to lead to criminal behaviour). We use data from the *Health Survey for England (HSE)*, in which over 6000 adults are asked to record their total alcohol consumption over a one week period in a diary. We categorise individuals into five groups: those that consumed fewer than 7, 7–14, 14–21, 21–35, and above 35 units per adult per week on average.[16]

In Table 4.1 we show that high alcohol consumption is highly correlated with both frequency of consumption and binge drinking episodes. For each group, we calculate the proportion of individuals who report binge drinking in their surveyed week (based on their heaviest drinking day during that week). The Office for National Statistics defines binge drinking as consuming more than 8 units a day for a man, and more than 6 units a day for a woman. The table also shows the average number of days on which they drank in the week. There is a strong correlation between average alcohol consumption

[16] *HSE* contains details of consumption from on- and off-trade alcohol. The purchase data we use to estimate the model contains information on the latter (which constitute 77% of total ethanol units purchased, on average). We use the *Living Costs and Food Survey (2011)* to calculate the share of ethanol units off-trade, conditional on total ethanol purchases. For each individual in the HSE data, we use this to predict the number of ethanol units they consume off-trade and group individuals by this in Table 4.1.

and the propensity to binge drink: over 70% of people who consume more than 35 units per week engaged in binge drinking on at least one occasion, compared to less than 10% of individuals who consume fewer than 7 units per week. People who drink more than 35 units on average per week also drink much more frequently: on 5.3 days, on average, in comparison to on 1.1 days for those that consume less than 7 units per week.

Table 4.1: *Drinking patterns across individuals*

Average alcohol consumption (weekly units)	% of individuals	% that binge drink	Days per week spent drinking
Less than 7 units	61.9	9.7	1.1
7-14 units	16.4	38.3	3.3
14-21 units	7.7	46.2	3.7
21-35 units	7.1	51.3	4.4
More than 35 units	6.9	73.4	5.3

Notes: For each individual we calculate the number of ethanol units consumed from purchases brought off-trade, and we the individuals into the groups shown in the table . Column (1) shows the percentage of individuals in each group. Column (2) shows the percentage of individuals that engaged in binge drinking in the seven days preceding the survey (defined as more than 8 units a day for a man, and more than 6 units a day for a woman). Column (3) shows the average number of days on which alcohol was consumed for individuals in each group.

4.2.1 Calibration

We calibrate the two parameters of the externality function, (ϕ_0, ϕ_1). ϕ_1 controls the convexity, measured as the ratio of second to first derivatives, and given ϕ_1, ϕ_0 controls the aggregate external cost.

We calibrate the aggregate external cost based on a study by the UK Cabinet Office (2003). Using this study Cnossen (2007) categorises estimates of the various costs associated with alcohol misuse in the UK. The report estimates the direct tangible social costs are £7.25 billion (in 2011 prices).[17] We use this to calibrate the aggregate external costs of alcohol consumption. Direct tangible costs of drinking include costs of alcohol-related disease and the costs of dealing with alcohol-related crime. Cnossen (2007) also contains estimates of the cost of lost output and intangible costs (for instance, emotional cost to victims of crime). We do not include estimates of the former as a large portion of this cost is likely to be borne by the drinker himself and we do not include the latter as the

[17] The estimate reported in the paper was £7.5 billion in 2001 prices; we uprate this to 2011 prices using the Retail Price Index and scale to account for off and on trade purchases. We assume that the share of external costs generated by off-trade alcohol consumption is proportional to the number of units consumed off-trade (77%).

costs are highly uncertain. Nonetheless, it is likely £7.25 billion is a lower bound on the total external costs of drinking.

While there is considerable evidence that the marginal externality of alcohol consumption is increasing (and hence the external cost function is convex), there is not much evidence on the degree of convexity. In our central calibration we set $\phi_1 = 0.0615$. This implies that someone that consumes 140g of ethanol has a marginal external cost of over 17 times that of someone that consumer 14g of ethanol. This is in line with the study by Taylor et al. (2010) who find the same increase in the odds of having an accident when comparing consumption of 140g of ethanol with 14g.

We test the robustness of our results to four alternative calibrations, which are described in Table 4.2. We consider the implications of higher or lower aggregate external costs and if the externality function is more or less convex. We show in Section 6 that although the calibration affects the solution to the optimal tax problems, our qualitative welfare results hold.

Table 4.2: *Calibration specifications*

	Aggregate external cost (£billion)	Ratio of external costs of heaviest to lightest drinkers	Calibrated parameters (ϕ_0, ϕ_1)
Central	7.25	20	$(1.2980, 0.0615)$
Alternative specifications			
High aggregate cost	8.50	20	$(1.5220, 0.0615)$
Low aggregate cost	6.00	20	$(1.0740, 0.0615)$
High convexity	7.25	30	$(0.8177, 0.0695)$
Low convexity	7.25	10	$(3.1730, 0.0435)$

Notes: The aggregate external cost is calculated as the sum of external costs over all households given their ethanol demand at UK prices in 2011. The ratio of the external costs of heaviest to lightest drinkers is calculated as the total external cost of all households that buy more than 35 units of ethanol per adult per week over the the total external cost of all households that buy less than 7 units of ethanol per adult per week.

4.3 Identification

There are three principal identification issues. This first is whether we identify the causal effect of a change in price on demand. The second is how we identify the distribution of household preferences. The third is whether dynamics in demand might bias our estimates of the preference parameters.

Price endogeneity

A classical potential problem in demand estimation is whether, conditional on the other variables in the model, price is correlated with the "demand shocks". In our case that would entail a correlation between alcohol prices and the ϵ_{ijst} shocks to utility conditional on the observable product attributes, unobserved product effects and time varying-alcohol type effects included in the model. Such a relationship would result in inconsistent demand estimates.

Our strategy for avoiding this problem is twofold. First, we include in the model a full set of product dummies and, for the set of alcohol types (e.g. gin, vodka, whisky etc.), we include quarterly varying time effects, (in the full specification, allowing for heterogeneity in their effect on demand across the D household groups). These time effects absorb seasonality in demand and spikes in demand due to advertising campaigns. As price-setting in the UK grocery market is done nationally, there is limited scope for firms to set geographically varying prices in response to local demand shocks.[18] We expect these rich time effects to absorb the majority of aggregate shocks to alcohol demand possibly correlated with price.

Second, we include a control function for price (see Blundell and Powell (2004) and, for multinomial discrete choice models, Petrin and Train (2010)). As instruments we use a set of cost shifters that include producer prices for beer and for cider, the pound-euro and pound-dollar exchange rates, alcohol duty rates, and wage indices in the retail and the food and catering sector. We estimate a first stage regression of price on the instruments (interacted with option effects) and the other variables included in the utility model. The F-stat for a test of the (ir)relevance of the instruments is 19.7, leading us to strongly reject the hypothesis of no relationship between price and the instruments. In demand estimation we control for the predicted residuals of the first stage regression. This means we use only price variation that is conditionally correlated with movements in cost shifters. The coefficient on the control function in demand estimation is statistically significant, indicating evidence of some price endogeneity in the absence of instrumenting. However, our qualitative predictions are similar with and without including the control function, indicating the rich time effects in demand are indeed doing a good job of absorbing shocks to demand.

[18] The large UK supermarkets, which make up over three quarters of the grocery market, agreed to implement a national pricing policy following the Competition Commission's investigation into supermarket behaviour (Competition Commission (2000)).

Preference distribution

A second identification issue is how we identify the preference distribution (which we model with random coefficients). We use data that are at the micro level and that are longitudinal so that we observe each household making repeated choices. Micro data has been shown to be particularly useful in identifying and estimating substitution patterns (see Berry and Haile (2010), Berry et al. (2004)). We observe households facing different choice situations (e.g. different price vectors and time-brand effects) plus there is within household variation in choice situations across time. It is this variation that allows us to estimate the parameters governing the preference distribution.

In our full model – with the mixed-normal preference distribution – we both include random coefficients and a measure of pre-sample average ethanol demand. It is important that we take into account dependence of the random coefficients on the measure of pre-sample demand. For instance, a household with a strong taste draw for vodka is likely to have purchased more alcohol in the past. We do this by modelling the distribution of household preferences *conditional* on which pre-sample purchase group the household belongs to.[19]

Formally, Berry and Haile (2010) and Fox and Gandhi (2016) establish conditions for nonparametric identification of random coefficients in random utility discrete choice models by placing restrictions on the covariate supports. Fox et al. (2012) show that the identification conditions are weaker in the case where ϵ_{ijst} shocks are distributed type I extreme value, and that even with cross sectional data the model is always identified if utilities are a function of linear indices with continuously distributed covariates.

Dynamics

A third identification issue is whether there are dynamics in demand that we fail to take into account. While allow for individual level heterogeneity in preferences and therefore statistical dependence in households' purchases, through time, through the random coefficient $(\Psi_i, \boldsymbol{\eta}_i)$. However, we do not model state dependence arising, conditional on heterogeneity, from the effect of recent purchases on current behaviour. Current choice may depend on recent past choices due to high frequency habits formation. It may also arise if households stockpile during sales periods (Hendel and Nevo (2006a)). We cannot categorically rule out these forms of state dependence but we present here some reduced

[19] The unconditional distribution of preferences is a mixture of these conditional distributions. If instead we directly integrated across the unconditional preference distribution, we would implicitly be assuming independence of the random coefficients and all the other variables in the model, which, given we include a measure of pre-sample purchases would be unreasonable.

form evidence that suggests these forms of short run dynamics are not very strong once we take account of household level preference heterogeneity.

We test for evidence of habit formation by running two regressions. The dependent variable in the first regression is a dummy equal to one if a household purchases alcohol in a given week and the dependent variable in the second regression is, conditional on purchasing, how may units of ethanol the household purchased. We regress these variables on the number of units of ethanol the household purchased in each of the past eight weeks, plus week dummies. We estimate each regression both omitting household fixed effects and including them. When we omit the fixed effects, there is a moderate relationship between recent past behaviour and current behaviour – for instance, purchasing 1 unit more alcohol per adult two weeks previously is associated with an increase in the probability of purchasing alcohol of 0.2 percentage points and conditional on buying, is associated with purchasing 0.13 (or 0.6%) more units. However, once we include household fixed effects, these numbers fall to just 0.02 percentage points and 0.01 units (for full results see Table C.1 in the Appendix).

We also assess evidence for omitted state dependence arising from consumers stockpiling during sale periods; if short-run price reductions, such as a sale, leads to an increase in alcohol purchases, which are then stored rather than immediately consumed, this would lead us to over-estimate the own price elasticities of demand (see e.g. Hendel and Nevo (2006a)). To test for such an effect we follow one of the suggestions in Hendel and Nevo (2006b). We assume that each household has a constant consumption rate (equal to their weekly average number of alcohol units purchased) and use this along with their purchases to compute an inventory for each household at the beginning of each week.[20] We then regress (i) the probability of purchase in a week and (ii) the number of units purchased (conditional on purchasing a positive amount) on this constructed inventory variable, week effects (which control for price changes, promotions, advertising etc.) and household fixed effects. Hendel and Nevo (2006b) argue that if stockpiling is present, then a high inventory is likely to lead to a lower probability of purchase or lower quantity purchased conditional on purchasing. In contrast, we find a very weak positive relationship between the inventory variable and both the probability and quantity of alcohol purchased (see Table C.2 in the Appendix).

4.4 Coefficient estimates

In Table 4.3 we report the coefficient estimates for the demand model we outline in Section 4.1. The first column of the table shows estimates for the "normal preference

[20]Since we include household fixed effects we can normalise the initial inventory to zero.

distribution" specification of the model. The remaining columns presents the estimates for the "mixed-normal preference distribution". Panel A shows estimated parameters of the random coefficient distributions over the observable product characteristics and Panel B summarises estimates of the parameters governing preferences over unobserved product attributes.[21]

In the case of the normal preference distribution specification, we assume preferences over the observed attributes, price, strength and ethanol, are normally distributed with an unrestricted covariance matrix. As the mean of the strength coefficient is not separately identified from the product effects we normalise it to zero. The variance parameters in column 1 indicate dispersion in preferences over price, strength and ethanol content is statistically significant. The covariance parameters show that more price sensitive consumers tend to have stronger preferences for total ethanol content and strength, but, in this specification, the covariance between preferences over ethanol content and alcohol strength is zero.

In Panel B of Table 4.3 we present estimates of the unobserved product effects. Rather than present estimates of the means of all the product effects, we present estimates of the average of the mean product effects within each alcohol segment (relative to the utility from the outside option). So, for instance, for the normal preference distribution specification, the average of the product effects for beer products is -1.127 and the estimated variance of the beer random coefficient distribution is 2.548. For each segment the mean of the products effects is negative – this reflects that we normalise the no purchase outside option utility to zero. Higher numbers in absolute terms indicate stronger preferences. The variance parameters vary across alcohol segments, indicating that the intensity of within- versus between-alcohol segment substitution varies across segments.

Columns 2-6 show the coefficient estimates for the mixed-normal preference distribution specification. Here all the parameters vary across the five household groups defined on the bases of pre-sample behaviour. This allows the correlation in preference parameters over observable attributes to vary by household groups. It also allows for richer and more complex correlations in preferences across all households than the more standard normal preference distribution specification. Moving from light drinking groups to heavier drinkers, the estimated mean product effects in all segment increases – heavy drinkers, on average, have higher valuations of unobserved products attributes for beer, wine, spirits and cider products relative to the no purchase outside option than lighter drinkers.

[21] Rather than show estimates of all J product effects, we show the means within each segment of the market. We also do not show estimates of the alcohol type-time effects. These are available from the authors upon request.

Table 4.3: *Estimated preference parameters*

Distribution of household preferences:	*Normal*	*Mixed normal*				
Household group:	·	< 7	7-14	14-21	21-35	> 35
Panel A: Preferences for observable product characteristics						
Means						
Price	-0.249	-0.327	-0.258	-0.254	-0.273	-0.283
	(0.019)	(0.039)	(0.028)	(0.025)	(0.023)	(0.024)
Beer*Total ethanol content	0.249	0.271	0.268	0.229	0.232	0.238
	(0.011)	(0.022)	(0.016)	(0.014)	(0.014)	(0.014)
Wine*Total ethanol content	0.022	0.030	0.036	0.047	0.064	0.107
	(0.012)	(0.025)	(0.017)	(0.015)	(0.014)	(0.013)
Spirits*Total ethanol content	0.146	0.336	0.144	0.089	0.049	0.064
	(0.028)	(0.061)	(0.057)	(0.041)	(0.047)	(0.039)
Cider*Total ethanol content	0.172	0.224	0.181	0.183	0.208	0.187
	(0.015)	(0.029)	(0.022)	(0.020)	(0.022)	(0.020)
Beer*Total ethanol content^2	-0.309	-0.339	-0.337	-0.221	-0.201	-0.191
	(0.015)	(0.030)	(0.021)	(0.017)	(0.017)	(0.018)
Wine*Total ethanol content^2	0.098	0.056	0.070	0.107	0.121	0.057
	(0.019)	(0.046)	(0.027)	(0.021)	(0.020)	(0.017)
Spirits*Total ethanol content^2	-0.139	-0.415	-0.108	0.008	0.091	0.095
	(0.038)	(0.085)	(0.080)	(0.056)	(0.063)	(0.051)
Cider*Total ethanol content^2	-0.281	-0.486	-0.269	-0.263	-0.267	-0.169
	(0.034)	(0.076)	(0.052)	(0.046)	(0.057)	(0.040)
Variances						
Price	0.060	0.043	0.047	0.068	0.061	0.053
	(0.005)	(0.009)	(0.006)	(0.007)	(0.006)	(0.004)
Total ethanol content	0.018	0.010	0.006	0.009	0.012	0.009
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Strength	0.233	0.312	0.490	0.387	0.332	0.374
	(0.022)	(0.037)	(0.041)	(0.030)	(0.022)	(0.030)
Covariances						
Price*Total ethanol content	-0.031	-0.018	-0.014	-0.023	-0.026	-0.021
	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
Price*Alcohol strength	-0.009	-0.013	-0.058	-0.050	0.020	0.012
	(0.006)	(0.011)	(0.009)	(0.010)	(0.006)	(0.005)
Total ethanol content*Alcohol strength	0.001	-0.016	-0.005	-0.003	-0.018	-0.005
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.002)
Panel B: Preferences for unobserved product characteristics						
Mean product effects for each segment						
Beer	-1.127	-1.349	-1.130	-0.970	-0.865	-0.786
	(0.021)	(0.037)	(0.030)	(0.030)	(0.029)	(0.030)
Wine	-5.792	-6.545	-5.477	-5.068	-4.274	-4.074
	(0.079)	(0.137)	(0.113)	(0.114)	(0.105)	(0.111)
Spirits	-4.945	-6.586	-4.303	-3.746	-3.116	-2.782
	(0.164)	(0.314)	(0.326)	(0.232)	(0.286)	(0.239)
Cider	-5.756	-8.449	-5.010	-4.022	-2.402	-2.002
	(0.362)	(0.703)	(0.755)	(0.527)	(0.655)	(0.556)
Variances						
Beer	2.548	2.303	2.109	2.895	2.292	1.805
	(0.141)	(0.199)	(0.209)	(0.234)	(0.188)	(0.144)
Wine	1.892	1.817	1.505	2.341	2.494	1.525
	(0.111)	(0.172)	(0.128)	(0.199)	(0.181)	(0.119)
Spirits	1.110	1.016	0.431	2.121	1.007	2.191
	(0.128)	(0.264)	(0.087)	(0.294)	(0.119)	(0.209)
Cider	3.521	1.766	3.688	3.301	2.582	3.069
	(0.254)	(0.226)	(0.322)	(0.323)	(0.242)	(0.274)
Product effects	Yes			Yes		
Type-time effects	Yes			Yes		
Control function	Yes			Yes		
Number of households	1000			2250		
Number of purchase occasions	25000			56250		

Notes: Column (1) contains the parameter estimates for the model in which the distribution of random coefficients across all households is modelled as normal. Columns (2)-(6) contain the parameter estimates for the model in which the distribution of preferences is a mixture of conditional normal distributions, conditioning on on households' pre-sample ethanol purchases. Panel A shows estimated parameters for the distribution of preferences over observable product characteristics, Panel B shows estimated parameters for the distribution of preferences over unobserved product characteristics. Standard errors are reported below the coefficients.

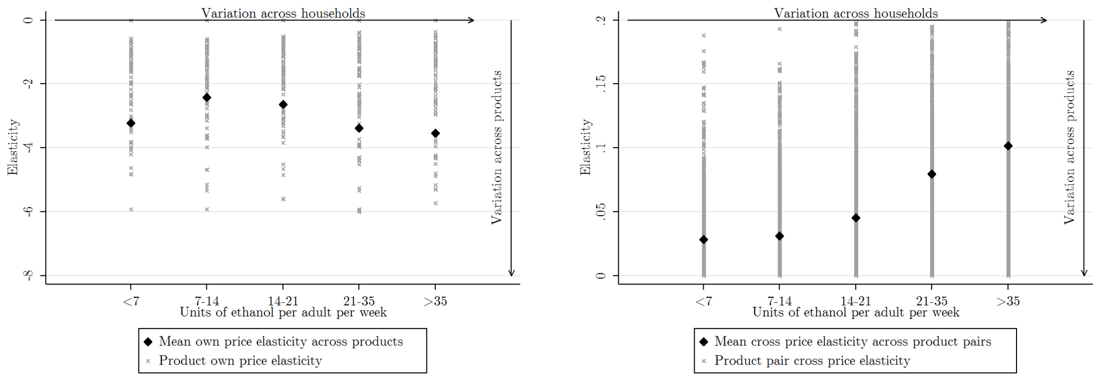
Given its additional flexibility, we present elasticity results in the next section and main optimal tax results, using the mixed-normal preference distribution specification. However, in Section 5.3, we compare the optimal taxes computed under the more restrictive normal preference distribution specification with those from the mixed-normal specification.

4.5 Price elasticities

The demand model estimates generate a set of own and cross price elasticities that capture how households switch between all the options (product-sizes) in the market, as well as towards the no purchase outside option, in response to marginal price changes. There is a 80×79 matrix of elasticities for every consumer (i.e. every draw from the preference distribution). In Figure 4.1 we summarise this information. For each of the five conditional distributions that comprise the overall preference distribution, we compute option level own and cross price elasticities (integrating across the conditional preference distribution). In the left hand panel of the figure we show the own price elasticities and in the right hand panel we show the cross price elasticities. The vertical variation in the graphs is across products and the horizontal variation is across the household groups. The graph highlights that variation in elasticities across products is substantial. It also shows some variation in the mean own price elasticity across groups – the lightest and the two heaviest groups of drinkers tend to have more elastic demand than those in the middle groups. However, the variation in the mean cross price elasticity across the household groups is much more striking. The mean cross price elasticity of households in the heaviest drinking group is over 3.5 times as high as the mean in the lightest group. Heavy drinks are much more likely to respond to an increase in a product's price by switching to alternative products (rather than out of the market).

Figure 4.1: *Summary of own and cross price elasticities*

(a) Own price elasticities (b) Cross price elasticities



Notes: The grey markers represent product level elasticities, computed separately for households in each of the 5 purchase level groups. The black marks are averages across these product level elasticities.

In Table 4.4 we focus on the spirits (including fortified wine) segment of the market and summarise mean (across options) own and cross price elasticities for each of the household groups.[22] Spirits products are disproportionately purchased by heavy drinkers and, as the table shows, spirits elasticities vary significantly across the household groups.

The first column shows the average own price elasticity for spirits options, the second column shows the average cross price elasticity between spirits options and the remaining columns show the average cross price elasticity between spirits options and options in each of the other three segments on the market. The table highlights that households are considerably more willing to switch from one spirit option to another, than they are from a spirit to a non-spirit option. For instance, the average cross price elasticity between spirits options is between four (for group 21-35 units) and six (for group 7-14 units) times the average cross price elasticity between spirits and ciders. The second main observation based on the table is that while the heaviest drinkers have higher cross price elasticities in general, the pattern is particularly strong for cross price elasticities between spirits options; for instance the average within spirits cross price elasticity for the heaviest drinking group is five times the value for the lightest drinking group.

[22] For elasticities for all segments see Table D.1 in the Appendix.

Table 4.4: *Average own and cross price elasticities for spirits products*

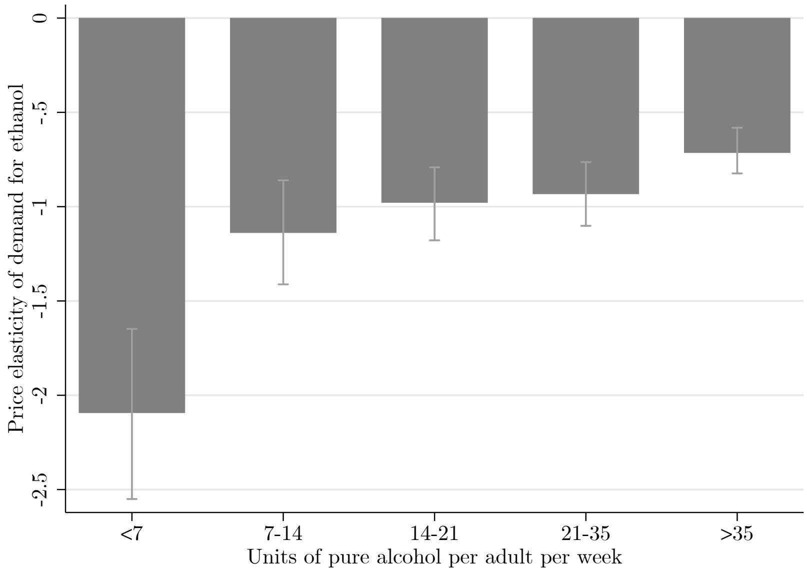
	(1)	(2)	(3)	(4)	(5)
	Mean	Mean cross price elasticity			
Household group	own price	Spirits	Beer	Wine	Cider
Less than 7 units	-4.094	0.039	0.016	0.019	0.007
7-14 units	-3.641	0.061	0.015	0.034	0.010
14-21 units	-3.721	0.104	0.024	0.033	0.020
21-35 units	-3.579	0.105	0.048	0.062	0.027
More than 35 units	-3.905	0.192	0.031	0.075	0.039

Notes: Each row shows the estimated elasticities for a different group of households. Column (1) shows the mean own price elasticity for spirits options. Columns (2)-(5) show the average cross price elasticity of spirits options with respect to a price change of an option in the alcohol segment indicated in the first row. The elasticities are a weighted averages of the option level elasticities where the weights are the options share of total units demanded. 95% confidence intervals are shown in Table D.1.

The price sensitivity of a household's overall ethanol demand to an increase in the price of all alcohol products will depend on all of their own and cross price elasticities across the alcohol options. We simulate the slope of ethanol demands by marginally uniformly increasing the price of all alcohol options. A number of papers have directly estimated the slope of ethanol demand (e.g. Banks et al. (1997), Baltagi and Griffin (1995), Manning et al. (1995)). These papers treat alcohol as a single commodity, implicitly assuming the composition of alcohol products does not alter as the price of all alcohol is raised. In contrast our model captures all the substitutions across all products and sizes that underlie changes in total ethanol demand.

Figure 4.2 plots the own elasticity of demand for ethanol for each household group. It shows that, in response to an increase in the price of all alcohol options, lighter drinking households reduce their ethanol demand, in percentage terms, by significantly more than heavier drinking households. The price elasticity of demand for ethanol for households that purchase fewer than 7 units of ethanol per adult per week in the pre-sample period is -2.09, compared to -0.71 for households that purchase more than 35 units per adult per week. This is driven by the fact that households that typically purchase relatively large quantities of ethanol are much more willing to switch between alcohol products and are also less likely to switch towards the outside option of not buying alcohol at all. This results in their ethanol demand being less elastic than more moderate drinkers. This pattern of response drives a negative correlation between the marginal externality of consumption and price sensitivity of ethanol demand, which as described in the following section, results in the optimal single rate ethanol tax being lower than the (unweighted) average marginal externality.

Figure 4.2: *Price elasticity of demand for all alcohol, by household group*



Notes: The elasticity is the percentage change in total alcohol units demanded following a 1% increase in the price of all options. The bars show the mean elasticities across months, for each group of households. 95% confidence intervals are shown.

5 Optimal alcohol taxes

In this section we use the empirical framework described in Section 4 to solve for the tax solutions to the planner's problem set out in Section 2. We describe the system of optimal alcohol taxes and how they compare with the current UK tax system. We also show the impact that modelling the preference distribution as a mixture of conditional normals has relative to the more standard normal distribution specification on optimal taxes, and we show the effect that household willingness to switch between alcohol products has on the solution.

5.1 Tax solutions

The first best Pigovian solution, which completely corrects for the externality, involves setting household specific tax rates per unit of ethanol. For each household, the rate is equal to the household's marginal consumption externality (evaluated at that rate) – see equation 2.4. Given the marginal externality is increasing in ethanol content, the policy prescribes relatively high taxes on households that purchase relatively large amounts of ethanol. We summarises these rates in Table E.1 of the Appendix.

In practice, it is hard for governments to implement consumer specific taxes. In Table 5.1 we show a set of optimal systems of alcohol taxation when tax rates are constrained to be common across consumers. In column 1 we report the optimal tax rate when there is a single rate applied to all products. We refer to this as Diamond taxation – the optimal tax rate is implicitly defined by equation 2.5. In this case the optimal alcohol tax is equal to the average marginal externality plus an adjustment for the covariance of marginal externalities and the absolute slope of overall ethanol demands. Given our empirical model, we compute the optimal rate is 35.9p per unit of ethanol. At this rate the (unweighted) average marginal externality, $\bar{\phi}$, is 43.9p per unit. The fact the optimal tax is below this level reflects the fact that households with high ethanol demands (and therefore high marginal externalities) tend to have ethanol demands that are less sensitive to a marginal changes in the tax rate (see Figure 4.2). This lowers the effectiveness of the tax relative to when there is no correlation, leading to a lower optimal rate.

In column 2 we show optimal tax rates when rates are allowed to be differentiated across the four segments of the alcohol market. The tax rate on spirits, 42.6p per unit of ethanol, is highest, followed by wine (30.4p), beer (28.3p) and cider (25.4p). Spirits are disproportionately purchased by heavy drinkers (see Table 3.2). A relatively high tax on spirits is therefore, to some extent, able to specifically target the alcohol consumption of the consumers that generate a relatively high marginal externality.

In column 3 we consider tax at a more disaggregate “type” level. In this case the tax rate is allowed to be differentiated across a set of alcohol types. These types are more aggregate than the products including in our demand system. We impose a common rate across different sizes of products and across products that are similar. For instance, the type tax involves a single tax rate for gin. It does not allow separate tax rates across different sizes of bottles, or between branded and generic/store brand gin. We impose this as it is unlikely the government would choose to differentiate taxes across these dimensions. This choice does not reflect any technical difficulty with solving for more disaggregate tax rates – doing so results in similar welfare predictions. It is important to note however, that a key feature of our demand framework is that it will capture substitution across sizes and products in response to any simulated form of tax, whether this be ethanol, segment or type taxes.

Table 5.1: *Tax rate solutions*

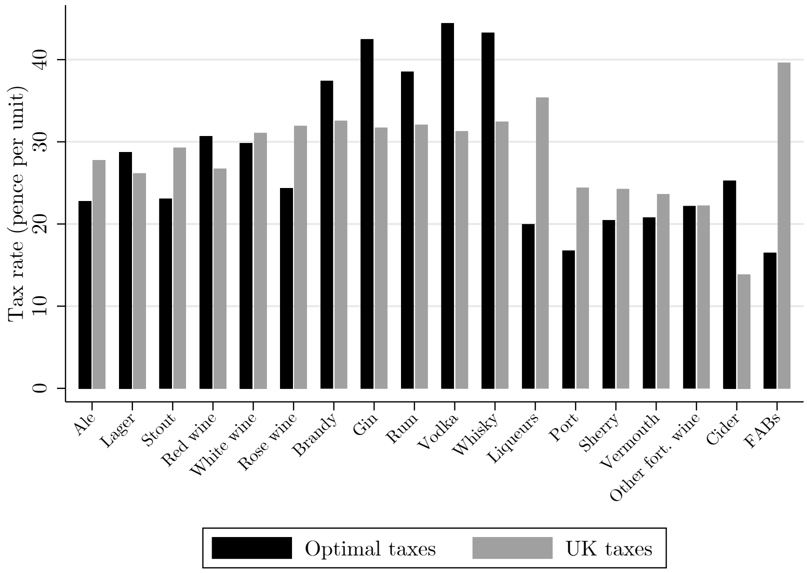
Optimal tax rates at level of:					
(1)		(2)		(3)	
Ethanol		Segment		Type	
Ethanol	35.9	Beer (inc. lager and ale)	28.3	Ale	22.7
	·		·	Lager	28.7
	·		·	Stout	23.1
	·	Wine	30.4	Red wine	30.6
	·		·	White wine	29.8
	·		·	Rose wine	24.3
	·	Spirits (inc. fortified wine)	42.6	Brandy	37.4
	·		·	Gin	42.4
	·		·	Rum	38.5
	·		·	Vodka	44.4
	·		·	Whisky	43.2
	·		·	Liqueurs	19.9
	·		·	Port	16.8
	·		·	Sherry	20.4
	·		·	Vermouth	20.8
	·		·	Other fort. wine	22.2
	·	Cider (inc. FABs)	25.4	Cider	25.2
	·		·	FABs	16.5

Notes: Each column shows the tax rates (expressed in pence per unit of ethanol) that maximise consumer welfare (equation (2.3)). Column (1) shows the optimal ethanol tax. Column (2) shows the optimal segment-level tax rates. Column (3) shows the optimal type-level taxes. The dots represent the tax rate shown in the row above.

The type taxes do involve considerable differentiation within alcohol segments. For instance, within the spirits segment the tax rate ranges from 16.8p per unit for port to 44.4p per unit for vodka and tax rates are higher on the set of strong spirits than on liqueurs or fortified wines. A similar pattern is true in the other segments – i.e. a significant degree of within segment variation in optimal taxes. This suggests going from the optimal system of segment taxes to optimal type taxes is likely to result in a reasonable degree of improvement in welfare.

Figure 5.1 shows how the optimal system of type level taxes compased with UK tax system that was in place in 2011 (including both alcohol excise tax and VAT components). The average tax rate across all alcohol types, weighted by demand, is similar under the optimal tax rates than under the UK system (27.5p compared with 27.8p). More substantively, there are significant differences in relative tax rates across different alcohol types. The system of optimal taxes, compared with the UK system, prescribes higher tax rates for brandy, gin, rum and vodka (all alcohols purchased disproportionately by heavy drinks) and a higher tax on cider (which for political economy reasons is taxed lightly in the UK). It also involved marginally higher taxes on lager and red wine but lower taxes on all other alcohols – especially FABs (which are taxed very highly in per unit terms in the UK as a consequence of having relatively low ABV but being taxed as spirits, which means they attract unusually high VAT per unit).

Figure 5.1: *Comparison of optimal taxes and UK tax rates*



Notes: The black bars shows the optimal type level taxes shown in Table 5.1; the grey bars show the per unit tax applied under the 2011 UK system (including both excise and VAT components).

5.2 Welfare

For each of the different tax policies, we calculate the impact on total consumer welfare of moving to the system from the UK system of taxes. The change in total consumer welfare equals the change in consumer surplus plus the change in tax revenue (which we assume is redistributed to consumers lump-sum) minus the change in the total externality costs. We present the results in Table 5.2. Under the UK system of taxes, the total external cost of consumption is £7.25 billion per year. This follows from our calibration of the externality function to match this figure. UK tax revenues are £7.16 billion per year.

Under the optimal single rate ethanol tax, the external costs of alcohol consumption are £2.00 billion per lower than under the UK system, while tax revenue is £0.31 billion higher. However these gains to welfare are partially off-set by the fact that consumer surplus falls by £1.85 billion. However, in total consumer welfare is £0.46 billion higher under the optimal single rate ethanol tax than under the UK tax system.

The consumer welfare gain of moving from the UK system to the optimal segment taxes is £0.83 billion – 1.8 times larger than the move from the UK system to the optimal ethanol tax. This larger welfare gain is due to a larger reduction in the external costs and a smaller fall in consumer surplus (although this is partially off-set by the segment tax raising marginally less revenue than the UK system).

Moving from the UK system to the optimal system of alcohol type taxes would result in a welfare gain of £1.05 billion – 1.3 times the gain of moving to the optimal segment tax. Relative to the segment taxes, the type taxes results in a slightly higher external cost and raises less revenue. However, this is more than made up for by the fact that they result in a much smaller loss in consumer surplus (when moving from the UK tax system). Relative to the ethanol and segment taxes, the optimal type taxes are better able to target the consumption of households with high marginal consumption externalities, while limiting the loss to consumer surplus.

Table 5.2: *Welfare impact of tax changes*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(3) + (4) − (2) Change in consumer welfare
UK taxes	7.25	7.16	–	–
Ethanol tax	-2.00	0.31	-1.85	0.46
% difference	*-27.6*	*4.3*	–	–
Segment taxes	-2.24	-0.13	-1.28	0.83
% difference	*-30.9*	*-1.8*	–	–
Type taxes	-2.15	-0.48	-0.63	1.05
% difference	*-29.7*	*-6.7*	–	–
Consumer specific taxes	-1.38	0.57	0.19	2.14
% difference	*-19.0*	*8.0*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for our central calibration of the externality function (Table 4.2). The rows below show the difference relative to the UK system for each tax policy. Column (1) shows the external cost, column (2) the tax revenue, column (3) the change in consumer surplus relative to the UK system, and column (4) the overall change in welfare. All numbers are expressed in £billion per year. Numbers in italic are the percentage differences relative to the UK system.

Table 5.2 also shows the welfare gain that would be achieved from a move from the UK system to the Pigovian first best of optimal consumer specific taxes. The gain to total consumer welfare would be £2.14 billion – comprised of a lower external cost, higher tax revenue and an increase in consumer surplus. Moving from the UK system to the optimal system of alcohol type taxes therefore would close nearly half of the gap in welfare between the UK system and the first best.

Distributional consequences

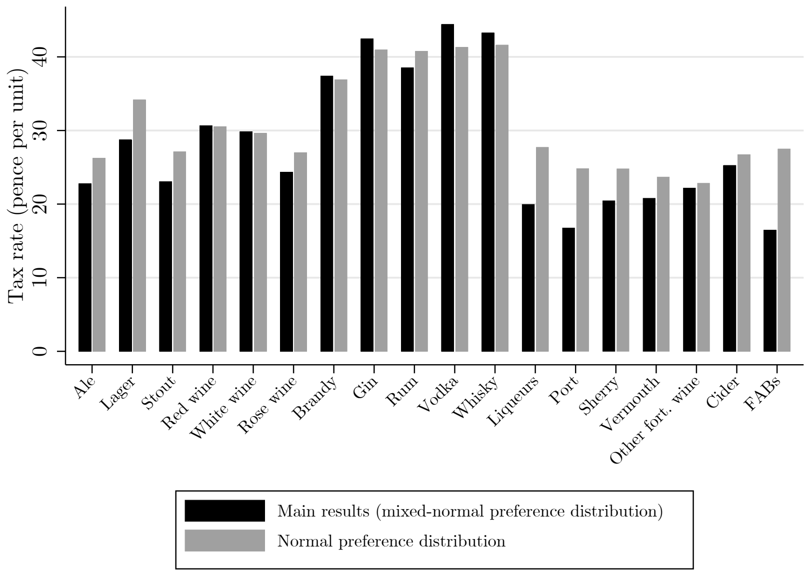
We assume the social planner does not consider questions of redistribution across consumers when setting alcohol taxes. This is sensible, as other parts of the tax and benefit system are better suited to achieving the government's distributional goals. Nevertheless, it is still informative to assess the distributional consequences of moving from the UK to the optimal system, as this will help inform which consumers require compensation. For each household in our data we know the socioeconomic group to which they belong. This measure is based on occupational status of the main shopper in the household and proxies for permanent income. Groups AB are managerial professionals, group C are skilled manual workers and groups DE are unskilled workers and those that rely on the state for their income. We do not explicitly include socioeconomic status in our model, but we can use the fraction of each of the five household groups comprising each socioeconomic category to infer how annual consumer surplus changes (moving from the UK to optimal alcohol type system) vary across socioeconomic status. We find that there is little variation in monetary terms across groups – those from groups AB see a reduction of £22.74 per year, those from group C see a reduction of £24.05 and those from groups DE see a reduction of £23.78. These falls do not account for tax revenue and savings from external costs. Assuming these are evenly spread across consumers, all groups would see an increase in their welfare.

5.3 Impact of preference heterogeneity and cross price effects

In Figure 5.2 we show the impact that modelling the preferences distribution as a mixture of conditional normal distributions (conditioning on past ethanol demand) instead of the standard assumption of normal distributions has on optimal tax results. In particular we show the optimal alcohol type taxes under both specifications of the preference distribution. The black bars are for the mixed normal preference distribution specification – these repeat the rates shown in Table 5.1 and Figure 5.1. The grey bars show the optimal taxes computed using the normal preference distribution specification.[23]

[23]In Appendix Table F.1 we show the set of optimal rates for the normal preference distribution specification.

Figure 5.2: *Comparison of optimal tax rates under different preference heterogeneity specifications*



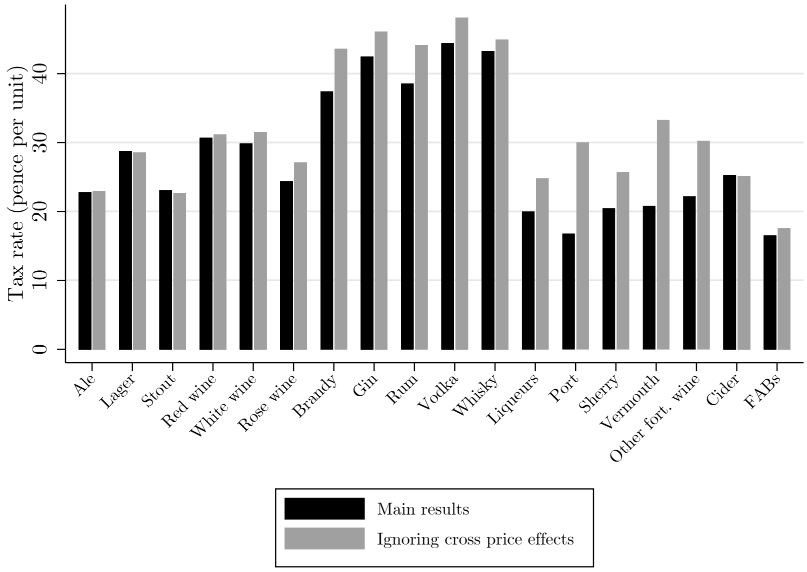
Notes: The black bars shows the optimal taxes under the mixed normal distribution of preferences; the grey bars show the optimal taxes under the normal distribution of preferences.

The pattern of optimal taxes is broadly similar under both specifications – in each case the set of strong spirits (brandy, gin, rum, vodka and whisky) attract relatively high rates. However, the exact tax rates differ – the mixed-normal specification, compared with normal specification, prescribes higher rates on strong spirits and lower rates on most other alcohol types. The reason for this is the more restricted normal specification is less able to fully capture the correlations in derived ethanol demand (and hence marginal externality) with preferences for strong spirits.

In Section 2.2 we argued that, given the average slopes and cross slopes of demand and given the covariance between households' marginal externalities and own slope of demands, increasing the covariance between marginal externalities and cross slopes of demand will lower optimal tax rates. The intuition behind this is that, all else equal, if households that generate high levels of externalities are more willing to switch between alcohol products than out of the market in response to price increases, the less effective the taxes will be at lowering the external costs of consumption, while not creating larger losses in consumer surplus.

We quantify the importance of this channel by recomputing the optimal tax rates that would result if the social planner were to ignore all cross price effects. In this case the tax rates are given by alcohol type-by-alcohol type application of the Diamond formula – see equation 2.7. Comparison of these "naive" optimal rates and the true optimal rates highlights the influence that cross price effects, and in particular how they vary with the marginal externality, has on optimal taxes. Figure 5.3 depicts this comparison. The black bars repeat the optimal rates and the grey bars show the optimal rates of a naive planner that ignores cross price effects. For all spirits and fortified wines, ignoring cross price effects results in lower taxes. The reason for this is that households that generate high marginal externalities tend to have higher cross price elasticities between alcohol products, particularly within the alcohol and fortified wine segment of the market – see Table 4.4.

Figure 5.3: *Comparison of optimal tax rates with and without cross price effects*



Notes: The black bars shows the optimal taxes when the planner takes cross price effects into account; the grey bars show the tax rates when she ignores cross price effects.

6 Robustness

In this section we discuss the robustness of our results to alternative calibration specifications. We consider four alternatives: (i) a function that gives a lower aggregate external cost, (ii) a function that gives a higher aggregate external cost, (iii) a less convex function, and (iv) a more convex function. Details of the calibration of these functions are given in Section 4.2.1. Tables that are the analogs to Table 5.2 are in the Appendix (Tables G.1-G.4).

Figure 6.1 presents the alcohol type tax solutions under the alternative calibration specifications. Varying the level of aggregate external cost at which the externality function is calibrated acts to scale the optimal tax rates up or down, relative to our central specification. If the planner thinks that £8.5 billion is a more realistic estimate of the aggregate external costs of alcohol consumption, then the optimal tax rates are roughly 6.5% higher than under our central calibration of a £7.25 billion aggregate external cost. On the other hand, if the planner is more conservative and thinks that £6 billion is a more sensible estimate, then the tax rates are roughly 7.5% lower.

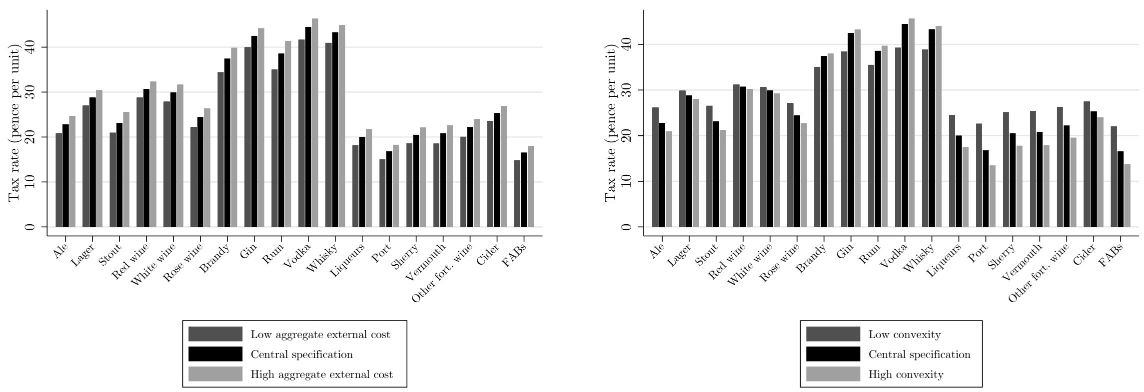
Varying the convexity of the function also changes the optimal tax rates in a way that makes intuitive sense: a more convex externality function leads to higher rates on high strength spirits, but lower rates on other products. In other words, the greater the proportion of external costs generated by the heaviest drinkers, the higher the relative tax rates on the alcohol types they purchase disproportionately.

The different specifications of the externality functions lead to differing predictions of the magnitude of welfare losses, relative to the first best, but the qualitative results from our central specification hold. Figure 6.2 presents the social welfare loss (relative to the first best consumer specific taxes) for the UK taxes, ethanol tax, and alcohol type taxes under each calibration specification. The less convex the function, the closer the optimal tax system (when tax rates are constrained to be common across households) gets to the first best. Although there are differences in the precise magnitudes, the key results hold: the UK system performs worst, the ethanol tax improves this somewhat, but the alcohol type taxes do better, roughly halving the difference in welfare between the UK taxes and the first best.

Figure 6.1: *Product tax solutions under different calibration specifications*

(a) Aggregate external cost

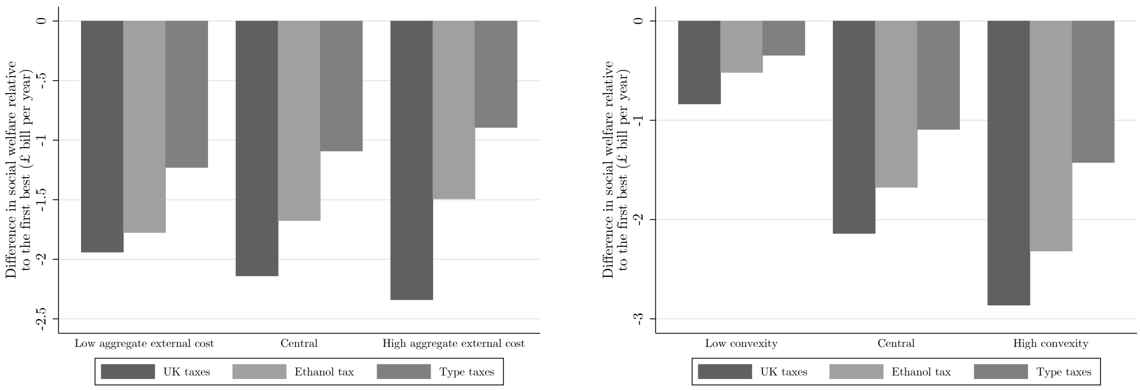
(b) Convexity of the externality function



Notes: Graphs show optimal alcohol type rates. Panel (a) shows how the numbers vary with the calibrated aggregate externality, panel (b) shows how they vary with the calibrated convexity of the externality function.

Figure 6.2: *Social welfare under different calibration specifications*

(a) Aggregate external cost (b) Convexity of the externality function



Notes: The graphs show the welfare difference between UK system, the optimal ethanol tax rate and the optimal alcohol type tax rates with the first best. Panel (a) shows how the numbers vary with the calibrated aggregate externality, panel (b) shows how they vary with the calibrated convexity of the externality function. Numbers are shown in Tables G.1-G.4 in the Appendix.

7 Summary and conclusions

In this paper we consider tax design to correct consumption externalities in markets in which i) marginal externalities vary across consumers and ii) there are many differentiated products. We specifically consider the alcohol market, in which heterogeneity in marginal externalities is driven by nonlinearity in the externality function – the marginal externality of the umpteenth drink is higher than that of the first – and where the externality arises from ethanol, but where this is bundled together in products with other characteristics over which the consumer has preferences. However, the framework can also be applied to other markets. A topical example would be to consider junk food markets, where the externality arises from sugar and where heterogeneity in externalities across consumers is driven by the particular concerns about children and/or obese people.

We show that differentiating tax rates across products can improve on Diamond taxation of ethanol. The reason is this exploits correlations in preferences of all product characteristics (and hence product level demands) with derived demand for ethanol, allowing for the tax system to better target the high externality generating consumption. We show that while moving from the UK tax system to Diamond taxation of ethanol would close 21% of the welfare gap between the UK system and the first best, moving to an optimal system that differentiates rates between alcohol types would close 49% of the gap.

Our focus in this paper has been on the correction of externalities. We have envisaged a social planner that sets taxes to maximise the sum of consumer surplus and tax revenue minus external costs. The social planner therefore does not take account of the existence of positive mark ups arising from competition. We have also assumed complete pass-through of tax to consumer prices. In the UK alcohol market we believe these are defensible abstractions; the UK supermarket segment, by international standards, is very competitive and policy is concerned with tackling excessive consumption. However, an important avenue for future research will be incorporating supply side considerations into the optimal tax framework. Doing so whilst considering the entire alcohol market is unlikely to be profitable, given the larger number of players, however focusing on a narrower segment of the market could provide a means for doing this.

References

Anderson, S. P., A. De Palma, and B. Kreider (2001). Tax incidence in differentiated product oligopoly. *Journal of Public Economics 81*(2), 173–192.

Baltagi, B. H. and J. M. Griffin (1995). A dynamic demand model for liquor: the case for pooling. *Review of Economics and Statistics 77*(3), 545–554.

Banks, J., R. Blundell, and A. Lewbel (1997). Quadratic Engel Curves and Consumer Demand. *Review of Economics and Statistics 79*(4), 527–539.

Berry, S. and P. Haile (2010). Nonparametric identification of multinomial choice demand models with heterogenous consumers. Technical report, Cowles Foundation Discussion Paper 1718.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market. *Journal of Political Economy 112*(1), 68–105.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics 25*(2), 242–262.

Blundell, R., P. Pashardes, and G. Weber (1993). We do we learn about consumer demand patterns from micro data? *American Economic Review 83*(3), 570–597.

Blundell, R. W. and J. L. Powell (2004). Endogeneity in semiparametric binary response models. *The Review of Economic Studies 71*(3), 655–679.

Cnossen, S. (2007). Alcohol taxation and regulation in the European Union. *International Tax and Public Finance 14*, 699?732.

Conlon, C. T. and N. S. Rao (2015). The Price of Liquor is Too Damn High: Alcohol Taxation and Market Structure. *NYU Wagner Research Paper No. 2610118*.

Corrao, G., L. Rubbiati, V. Bagnardi, A. Zambon, and K. Poikolainen (2000). Alcohol and coronary heart disease: a meta-analysis. *Addiction 95*(10), 1505–1523.

Crawford, I., Z. Smith, and S. Tanner (1999). Alcohol taxes, tax revenue and the Single European Market. *Fiscal Studies 20*(3), 287–304.

Deaton, A. and J. Muellbauer (1980). An almost ideal demand system. *The American Economic Review 70*(3), 312–326.

Diamond, P. (1973). Consumption Externalities and Imperfect Corrective Pricing. *The Bell Journal of Economics and Management Science 4*(2), 526–538.

Fox, J. T. and A. Gandhi (2016). Nonparametric identification and estimation of random coefficients in multinomial choice models. *The RAND Journal of Economics 47*(1), 118–139.

Fox, J. T., K. il Kim, S. P. Ryan, and P. Bajari (2012). The random coefficients logit model is identified. *Journal of Econometrics 166*(2), 204–212.

Goddard, E. (2007). Estimating alcohol consumption from survey data: Updated method of converting volumes to units.

Gorman, W. M. (1980). A Possible Procedure for Analysing Quality Differentials in the Egg Market. *The Review of Economic Studies 47*(5), 843–856.

Griffith, R., A. Leicester, and M. O'Connell (2013). Price-based measures to reduce alcohol consumption. Technical report, IFS Briefing Note No. 138.

Griffith, R. and M. O'Connell (2009). The Use of Scanner Data for Research into Nutrition. *Fiscal Studies 30*, 339–365.

Hendel, I. and A. Nevo (2006a). Measuring the implications of sales and consumer inventory behavior. *Econometrica 74*(6), 1637–1673.

Hendel, I. and A. Nevo (2006b). Sales and consumer inventory. *RAND Journal of Economics 37*(3), 543–561.

Irvine, I. J. and W. A. Sims (1993). The welfare effects of alcohol taxation. *Journal of Public Economics 52*(1), 83–100.

Lancaster, K. (1971). *Consumer demand: A new approach.* Columbia University Press New York.

Lee, L.-F. and M. M. Pitt (1986). Microeconometric demand system with binding nonnegativity constraints: the dual approach. *Econometrica 54*(5), 1237–1242.

Leicester, A. and Z. Oldfield (2009). An analysis of consumer panel data. Technical report, IFS Working Papers W09/09.

Lönnroth, K., B. G. Williams, S. Stadlin, E. Jaramillo, and C. Dye (2008). Alcohol use as a risk factor for tuberculosis–a systematic review. *BMC Public Health 8*(1), 1.

Manning, W. G., L. Blumberg, and L. H. Moulton (1995). The demand for alcohol: the differential response to price. *Journal of Health Economics 14*(2), 123–148.

Miravete, E. J., K. Seim, and J. Thurk (2016). Naive Policy Design and Legislative Inertia in the Regulation of Alcohol. Previously circulated as *CEPR DP No. 9641*.

Mirrlees-Black, C. (1999). Domestic violence: findings from a new british crime survey self-completion questionnaire. Technical report, Home Office Research Study 191.

National Institute on Alcohol Abuse and Alcoholism (2000). Tenth special report to the US Congress on alcohol and health. Technical report, Washington: US Government Printing Office.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

NICE (2010). Alcohol-use disorders: Preventing harmful drinking. Technical report, National Institute for Clinical Excellence.

O'Donoghue, T. and M. Rabin (2000). The economics of immediate gratification. *Journal of Behavioral Decision Making 13*(2), 233–250.

Peele, S. and A. Brodsky (2000). Exploring psychological benefits associated with moderate alcohol use: a necessary corrective to assessments of drinking outcomes? *Drug and alcohol dependence 60*(3), 221–247.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research 47*(1), 3–13.

Pigou, A. C. (1920). *The Economics of Welfare*. McMillan&Co., London.

Prime Minister's Strategy Unit (2004). Alcohol Harm Reduction Strategy for England.

Pudney, S. (1989). *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*. http://iserwww.essex.ac.uk/home/spudney.

Purshouse, R. C., P. S. Meier, A. Brennan, K. B. Taylor, and R. Rafia (2010). Estimated effect of alcohol pricing policies on health and health economic outcomes in england: an epidemiological model. *The Lancet 375*(9723), 1355–1364.

Rehm, J., B. Taylor, S. Mohapatra, H. Irving, D. Baliunas, J. Patra, and M. Roerecke (2010). Alcohol as a risk factor for liver cirrhosis: A systematic review and meta-analysis. *Drug and alcohol review 29*(4), 437–445.

Sandmo, A. (1975). Optimal taxation in the presence of externalities. *The Swedish Journal of Economics 77*(1), 86–98.

Seim, K. and J. Waldfogel (2013). Public monopoly and economic efficiency: Evidence from the pennsylvania liquor control board's entry decisions. *American Economic Review 103*(2).

Singleton, N., M. Farrell, and H. Meltzer (2003). Substance misuse among prisoners in england and wales. *International review of psychiatry 15*(1-2), 150–152.

Taylor, B., H. Irving, F. Kanteres, R. Room, G. Borges, C. Cherpitel, T. Greenfield, and J. Rehm (2010). The more you drink, the harder you fall: a systematic review and meta-analysis of how acute alcohol consumption and injury or collision risk increase together. *Drug and alcohol dependence 110*(1), 108–116.

UK Cabinet Office (2003). Alcohol misuse: how much does it cost? http://alcoholresearchuk.org/wp-content/uploads/2014/01/strategy-unit-alcohol-costs-2003.pdf. London: Strategy Unit.

Watson, D. and J. Parsons (2005). Domestic abuse of women and men in Ireland: report on the national study of domestic abuse. Dublin: Stationery Office.

Wells, S., K. Graham, M. Speechley, and J. J Koval (2005). Drinking patterns, drinking contexts and alcohol-related aggression among late adolescent and young adult drinkers. *Addiction 100*(7), 933–944.

World Health Organization (2014). European status report on alcohol and health 2014. Technical report, Geneva: World Health Organization.

Appendix

Alcohol tax design

Rachel Griffith, Martin O'Connell and Kate Smith
October 24, 2016

A Optimal tax formulae with cross-price effects

Consider a two-product market, $j = 1, 2$. Let $Q_{ijk} = z_j \frac{\partial q_{ij}}{\partial t_k}$. Suppose own demands are downward slopping ($Q'_{ijj} < 0$) and the products are substitutes ($Q'_{ijk} > 0$). The first order condition for product j is:

$$\sum_i [(\tau_1 - \phi'_i)\, Q'_{i1j} + (\tau_2 - \phi'_i)\, Q'_{i2j}] = 0$$
$$\Rightarrow \tau_1 \sum_i Q'_{i1j} - \sum_i \phi'_i Q'_{i1j} + \tau_2 \sum_i Q'_{i2j} - \sum_i \phi'_i Q'_{i2j} = 0$$

Denote the average slope of demand $\bar{Q}'_{jk} = \frac{1}{N} \sum_i \bar{Q}'_{ijk}$. For own demands it is convenient to use the absolute slope, $|Q_{ijj}|$, which is positive. The first order condition for τ_2 implies:

$$\tau_2 = \frac{1}{N|\bar{Q}'_{22}|} \left[t_1 N \bar{Q}'_{12} - \sum_i \phi'_i Q'_{i12} + \sum_i \phi'_i |Q'_{i22}| \right]$$

Assuming symmetry of demands $Q'_{i12} = Q'_{i21}$ and substituting the expression for τ_2 into the first order condition for τ_1 yields:

$$\tau_1 = \frac{|\bar{Q}'_{22}|}{N(|\bar{Q}'_{11}||\bar{Q}'_{22}| - \bar{Q}_{12}\bar{Q}_{12})} \left[\sum_i \phi'_i |Q'_{i11}| + \frac{\bar{Q}'_{12}}{|\bar{Q}'_{22}|} \sum_i \phi'_i |Q'_{i22}| - \left(1 + \frac{\bar{Q}'_{12}}{|\bar{Q}'_{22}|}\right) \sum_i \phi'_i Q'_{i12} \right]$$

As long as the own price slope of market demand is steeper than the cross price slope, $|\bar{Q}'_{jj}| > \bar{Q}'_{jk}$, then the pre-multiplying term is greater than zero, $\frac{|\bar{Q}'_{22}|}{N(|\bar{Q}'_{11}||\bar{Q}'_{22}| - \bar{Q}_{12}\bar{Q}_{12})} > 0$. Noting that:

$$\sum_i \phi'_i Q'_{ijk} = N\text{Cov}(\phi'_i, Q'_{ijk}) + N\bar{\phi}'\bar{Q}'{}_{jk},$$

where can re-write the condition as:

$$\tau_1 = \bar{\phi}' + \frac{|\bar{Q}'_{22}|}{|\bar{Q}'_{11}||\bar{Q}'_{22}| - \bar{Q}'_{12}\bar{Q}'_{12}} \left(\text{Cov}(\phi'_i, |Q'_{i11}|) + \frac{\bar{Q}'_{21}}{|\bar{Q}'_{22}|} \text{Cov}(\phi'_i, |Q'_{i22}|) - \left(1 + \frac{\bar{Q}'_{21}}{|\bar{Q}'_{22}|}\right) \text{Cov}(\phi'_i, Q'_{i12}) \right)$$

and express the difference in optimal taxes as:

$$\begin{aligned}\tau_1 - \tau_2 = &\frac{|Q'_{22}| - |Q'_{21}|}{|\bar{Q}'_{11}||\bar{Q}'_{22}| - \bar{Q}'_{12}\bar{Q}'_{12}}\mathrm{Cov}(\phi'_i, |Q'_{i11}|) \\ &- \frac{|Q'_{11}| - |Q'_{21}|}{|\bar{Q}'_{11}||\bar{Q}'_{22}| - \bar{Q}'_{12}\bar{Q}'_{12}}\mathrm{Cov}(\phi'_i, |Q'_{i22}|) \\ &- \frac{|Q'_{22}| - |Q'_{11}|}{|\bar{Q}'_{11}||\bar{Q}'_{22}| - \bar{Q}'_{12}\bar{Q}'_{12}}\mathrm{Cov}(\phi'_i, |Q'_{i12}|)\end{aligned}$$

B Alcohol products and sizes

Table B.1: *Observed product attributes*

	Product (j)	Size (s)	Price in £ (p_{jst})	Alcohol units (z_{js})	Alcohol strength (w_j)
Beer					
(1)	Ale: low strength	c. 500ml	1.97	2.51	3.60
(2)		c. 4x440ml	3.38	6.31	3.60
(3)		c. 12x440ml	11.53	25.73	3.60
(4)	Ale: mid strength, bottles	c. 500ml	3.24	4.69	4.54
(5)		> 1x500ml	6.60	11.86	4.54
(6)	Ale: mid strength, cans	c. 4x500ml	6.71	16.03	4.50
(7)	Ale: high strength	c. 500ml	2.98	4.91	5.67
(8)		> 1x500ml	7.89	16.34	5.67
(9)	Lager: branded, low strength	c. 4x440ml	3.78	7.15	3.91
(10)		c. 12x440ml	9.70	22.37	3.91
(11)		c. 20x440ml	17.46	45.92	3.91
(12)	Lager: branded, mid strength	c. 4x330ml	3.99	6.85	4.65
(13)		c. 12x330ml	11.30	23.53	4.65
(14)	Lager: branded, high strength, bottles	c. 660ml	2.37	3.91	5.11
(15)		c. 4x330ml	3.87	6.94	5.11
(16)		c. 12x275ml	6.00	12.17	5.11
(17)		c. 15x275ml	12.78	31.17	5.11
(18)	Lager: branded, high strength, cans	c. 4x440ml	4.34	10.39	5.47
(19)		c. 10x440ml	12.73	33.11	5.47
(20)	Lager: store brand	c. 4x500ml	5.06	15.91	4.10
(21)	Stout	c. 500ml	2.43	3.13	4.23
(22)		c. 4x440ml	4.47	6.90	4.23
(23)		c. 10x440ml	13.55	25.04	4.23
Wine					
(24)	Red wine: store brand	c. 750ml	5.66	12.38	12.52
(25)		> 1x750ml	12.00	30.21	12.52
(26)	Red wine: branded	c. 750ml	8.24	15.66	12.60
(27)		c. 2x750ml	11.90	23.61	12.60
(28)		> 2x750ml	17.19	38.66	12.60
(29)	White wine: still, store brand	c. 750ml	5.08	10.77	11.91
(30)		> 1x750ml	11.32	27.60	11.91
(31)	White wine: still, branded	c. 750ml	7.21	13.64	12.28
(32)		c. 2x750ml	11.08	21.62	12.28
(33)		> 1x750ml	16.84	37.32	12.28
(34)	White wine: sparkling, store brand	c. 750ml	5.56	8.11	10.45
(35)		> 1x750ml	13.06	20.93	10.45
(36)	White wine: sparkling, branded	c. 750ml	6.86	8.04	9.14
(37)		> 1x750ml	9.16	21.50	9.14
(38)	Rose wine: still, store brand	c. 750ml	4.26	9.44	11.84
(39)		> 1x750ml	10.20	25.25	11.84
(40)	Rose wine: still, branded	c. 750ml	5.05	9.56	11.41
(41)		> 1x750ml	12.20	25.08	11.41
(42)	Rose wine: sparkling, store brand	c. 750ml	6.73	10.48	9.42
(43)	Rose wine: sparkling, branded	c. 750ml	6.17	8.02	10.17
(44)		> 1x750ml	15.53	21.00	10.17

Notes: Price is the average price of the product-size pair across months. p_{jst} is a price index constructed using fixed weights for the UPCs in each option, therefore changes in p_{jst} reflect movements in the prices of the underlying UPCs, not changes in the composition of barcodes bought. Column (2) shows the number of alcohol units (10ml of ethanol) in each option. Column (3) shows the alcoholic strength (ABV) of each option. Both z_{js} and w_j are time-invariant.

Alcohol products and sizes, cont.

	Product (j)	Size (s)	Price in £ (p_{jst})	Alcohol units (z_{js})	Alcohol strength (w_j)
Spirits					
(45)	Brandy	c. 700ml	10.75	24.26	37.28
(46)		c. 1.4l	17.71	40.93	37.28
(47)	Gin; store brand	c. 700ml	8.74	24.63	38.38
(48)		c. 1.4l	15.29	43.96	38.38
(49)	Gin; branded	c. 700ml	11.52	26.33	38.23
(50)		c. 1.4l	18.44	44.10	38.23
(51)	Rum	c. 700ml	10.73	25.50	37.15
(52)		c. 1.4l	17.20	42.77	37.15
(53)	Vodka; store brand	c. 700ml	8.08	22.42	37.55
(54)		c. 1.4l	15.95	44.35	37.55
(55)	Vodka; branded	c. 700ml	10.38	25.79	37.63
(56)		c. 1.4l	16.35	43.05	37.63
(57)	Whisky; store brand	c. 700ml	10.61	25.87	40.00
(58)		c. 1.4l	17.89	45.64	40.00
(59)	Whisky; branded	c. 700ml	14.97	28.42	40.11
(60)		c. 1.4l	17.17	41.93	40.11
(61)	Liqueurs	c. 700ml	10.55	16.70	21.50
(62)		c. 1.4l	15.68	25.70	21.50
(63)	Port	c. 750ml	8.61	17.26	19.82
(64)	Sherry	c. 750ml	7.51	18.86	16.74
(65)	Vermouth	c. 1.4l	6.65	18.04	14.94
(66)	Other fort. wine	c. 1l	6.22	17.88	14.61
Cider and flavoured alcoholic beverages (FABs)					
(67)	Dry cider, low strength	c. 1l	2.47	3.95	4.36
(68)		c. 4l	6.32	18.09	4.36
(69)	Dry cider, high strength, store brand	c. 2l	2.28	9.95	5.82
(70)		c. 5l	5.36	27.42	5.82
(71)	Dry cider, high strength, branded	c. 500ml	3.05	6.61	5.99
(72)		c. 2l	3.84	11.51	5.99
(73)		c. 12x440ml	10.01	34.80	5.99
(74)	Pear cider	c. 568ml	2.36	4.70	5.01
(75)		c. 3l	6.77	18.72	5.01
(76)	Fruit cider	c. 1l	4.63	6.00	4.47
(77)	Pre-mixed spirit	c. 750ml	4.13	4.54	6.16
(78)	Alcopops	c. 700ml	3.66	4.32	4.90
(79)		c. 2x700ml	8.27	10.03	4.90

Notes: Price is the average price of the product-size pair across months. p_{jst} *is a price index constructed using fixed weights for the UPCs in each option, therefore changes in* p_{jst} *reflect movements in the prices of the underlying UPCs, not changes in the composition of barcodes bought. Column (2) shows the number of alcohol units (10ml of ethanol) in each option. Column (3) shows the alcoholic strength (ABV) of each option. Both* z_{js} *and* w_j *are time-invariant.*

C Short-run persistence and stockpiling

Table C.1: *Dependence of current purchase decisions on past alcohol purchases*

	(1) Purchased alcohol?	(2) Purchased alcohol?	(3) Quantity	(4) Quantity
Number of units purchased per adult per week:				
1 week before	0.0016	-0.0005	0.0942	-0.0150
	(0.0001)	(0.0001)	(0.0028)	(0.0027)
2 weeks before	0.0024	0.0002	0.1238	0.0113
	(0.0001)	(0.0001)	(0.0028)	(0.0027)
3 weeks before	0.0022	0.0001	0.1079	0.0013
	(0.0001)	(0.0001)	(0.0028)	(0.0027)
4 weeks before	0.0023	0.0001	0.1132	0.0103
	(0.0001)	(0.0001)	(0.0029)	(0.0027)
5 weeks before	0.0021	-0.0000	0.1017	0.0008
	(0.0001)	(0.0001)	(0.0029)	(0.0028)
6 weeks before	0.0019	-0.0002	0.0953	-0.0039
	(0.0001)	(0.0001)	(0.0029)	(0.0028)
7 weeks before	0.0019	-0.0002	0.1020	0.0014
	(0.0001)	(0.0001)	(0.0029)	(0.0028)
8 weeks before	0.0021	-0.0001	0.1074	0.0069
	(0.0001)	(0.0001)	(0.0029)	(0.0028)
Mean of dependent variable	0.3833	0.3833	19.7637	19.7637
Time effects?	Yes	Yes	Yes	Yes
Household fixed effects?	No	Yes	No	Yes

Notes: The dependent variable in columns (1) and (2) is a dummy equal to one if the household purchase alcohol in that week. The dependent variable in columns (3) and (4) is the number of units purchased per adult in that week, conditional on making a non-zero purchase. The table shows the estimated coefficients on the number of units purchased per adult in the preceding one, two, three, etc. weeks. Standard errors are shown in parentheses. Week effects are included, and household fixed effects are include in columns (2) and (3).

Table C.2: *Dependence of current purchase decisions on inventory*

	(1) Purchase alcohol?	(2) Quantity
Inventory	0.0015	0.0897
	(0.0000)	(0.0010)
Mean of dependent variable	0.3833	19.7637
Time effects?	Yes	Yes
Household fixed effects?	Yes	Yes

Notes: The dependent variable in column (1) is a dummy equal to one if the household purchase alcohol in that week. The dependent variable in columns (2) is the number of units purchased per adult in that week, conditional on making a non-zero purchase. The table shows the estimated coefficients on a variable for the household's alcohol inventory. This is calculated by assuming that the household has a fixed level of consumption (equal to its mean purchases over the year) and an initial inventory of zero. Standard errors are shown in parentheses. Week effects and household fixed effects are included in both regressions.

D Elasticities

Table D.1: *Average own and cross price elasticities within and between alcohol segments, by household group*

	Mean	Mean cross price elasticity			
<7 units	own price	Beer	Wine	Spirits	Cider
Beer and lager	-3.413	0.087	0.017	0.010	0.008
	[-4.156, -2.665]	[0.063, 0.122]	[0.012, 0.023]	[0.007, 0.015]	[0.006, 0.015]
Wine	-2.740	0.018	0.044	0.011	0.007
	[-3.342, -2.071]	[0.013, 0.027]	[0.032, 0.058]	[0.008, 0.017]	[0.005, 0.010]
Spirits	-4.094	0.016	0.019	0.039	0.007
	[-5.089, -3.111]	[0.012, 0.026]	[0.014, 0.025]	[0.028, 0.061]	[0.005, 0.011]
Cider and FABs	-2.135	0.030	0.020	0.014	0.028
	[-2.749, -1.651]	[0.021, 0.052]	[0.015, 0.026]	[0.010, 0.020]	[0.020, 0.046]
7-14 units					
Beer and lager	-2.010	0.050	0.027	0.012	0.011
	[-2.609, -1.332]	[0.030, 0.076]	[0.014, 0.039]	[0.009, 0.017]	[0.008, 0.016]
Wine	-2.000	0.015	0.052	0.015	0.009
	[-2.593, -1.284]	[0.008, 0.023]	[0.027, 0.074]	[0.010, 0.020]	[0.006, 0.012]
Spirits	-3.641	0.015	0.034	0.061	0.010
	[-4.488, -2.774]	[0.010, 0.021]	[0.024, 0.043]	[0.045, 0.084]	[0.007, 0.012]
Cider and FABs	-1.815	0.024	0.036	0.017	0.058
	[-2.169, -1.396]	[0.017, 0.037]	[0.026, 0.047]	[0.013, 0.022]	[0.044, 0.075]
14-21 units					
Beer and lager	-2.456	0.095	0.038	0.020	0.026
	[-3.084, -1.975]	[0.073, 0.121]	[0.026, 0.051]	[0.015, 0.028]	[0.020, 0.032]
Wine	-1.964	0.025	0.067	0.018	0.015
	[-2.622, -1.440]	[0.018, 0.034]	[0.041, 0.097]	[0.013, 0.025]	[0.012, 0.019]
Spirits	-3.721	0.024	0.033	0.104	0.020
	[-4.536, -2.879]	[0.018, 0.032]	[0.024, 0.043]	[0.073, 0.138]	[0.014, 0.026]
Cider and FABs	-2.309	0.063	0.055	0.036	0.131
	[-2.683, -1.988]	[0.051, 0.091]	[0.043, 0.068]	[0.027, 0.048]	[0.106, 0.156]
21-35 units					
Beer and lager	-3.649	0.167	0.080	0.034	0.038
	[-4.266, -3.086]	[0.143, 0.198]	[0.066, 0.098]	[0.028, 0.043]	[0.030, 0.049]
Wine	-3.121	0.054	0.155	0.031	0.024
	[-3.745, -2.535]	[0.044, 0.067]	[0.126, 0.186]	[0.024, 0.039]	[0.020, 0.030]
Spirits	-3.579	0.048	0.062	0.105	0.027
	[-4.323, -2.921]	[0.041, 0.059]	[0.050, 0.076]	[0.083, 0.131]	[0.022, 0.033]
Cider and FABs	-2.853	0.120	0.107	0.055	0.167
	[-3.221, -2.460]	[0.099, 0.149]	[0.090, 0.124]	[0.048, 0.066]	[0.136, 0.200]
> 35 units					
Beer and lager	-3.463	0.114	0.105	0.055	0.063
	[-3.972, -2.866]	[0.090, 0.136]	[0.084, 0.125]	[0.044, 0.066]	[0.052, 0.077]
Wine	-3.295	0.043	0.164	0.057	0.045
	[-3.922, -2.673]	[0.034, 0.053]	[0.130, 0.201]	[0.044, 0.069]	[0.038, 0.055]
Spirits	-3.905	0.031	0.075	0.192	0.039
	[-4.600, -3.226]	[0.024, 0.038]	[0.060, 0.092]	[0.152, 0.230]	[0.031, 0.048]
Cider and FABs	-2.934	0.073	0.122	0.080	0.235
	[-3.266, -2.554]	[0.059, 0.089]	[0.103, 0.139]	[0.067, 0.091]	[0.195, 0.281]

Notes: Each panel shows the estimated elasticities for a different group of households. The second column shows the average own price elasticity for options within each alcohol segment. Columns (3)-(6) show the average cross price elasticity of options in the alcohol segment indicated in the first column with respect to a price change of an option in the alcohol segment indicated in the first row. The elasticities are a weighted averages of the option level elasticities where the weights are the options share of total units demanded. 95% confidence intervals in square brackets below each number.

E First best

Table E.1: *Consumer specific tax rates*

	Tax per unit of ethanol	
Household group:	Mean	Std. dev
Less than 7 units	17.08	4.19
7-14 units	23.97	10.29
14-21 units	28.99	11.99
21-35 units	33.41	15.50
More than 35 units	39.15	28.44
All households	22.94	13.60

Notes: Numbers show mean and standard deviation of optimal consumer specific (first best) tax rates within each household purchase level group and across all households.

F Preference heterogeneity distribution

Table F.1: *Tax rate solutions: normally distributed preference heterogeneity*

Optimal tax rates at level of:					
(1)		(2)		(3)	
Type		Segment		Single rate	
Ethanol	38.7	Beer and lager	33.9	Ale	26.2
	·		·	Lager	34.1
	·		·	Stout	27.1
	·	Wine	30.3	Red wine	30.5
	·		·	White wine	29.6
	·		·	Rose wine	27.0
	·	Spirits	41.4	Brandy	36.9
	·		·	Gin	40.9
	·		·	Rum	40.8
	·		·	Vodka	41.3
	·		·	Whisky	41.6
	·		·	Liqueurs	27.7
	·		·	Port	24.8
	·		·	Sherry	24.8
	·		·	Vermouth	23.7
	·		·	Other fort. wine	22.8
	·	Cider and FABs	27.2	Cider	26.7
	·		·	FABs	27.5

Notes: Each column shows the tax rates (expressed in pence per unit of ethanol) that maximise consumer welfare (equation (2.3)). Column (1) shows the optimal type-level taxes. Column (2) shows the optimal segment-level tax rates. The final column shows the optimal commodity tax rate for alcohol. The dots represent the tax rate shown in the row above.

G Robustness

Table G.1: *Welfare impact of tax changes: low aggregate external cost*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(3) + (4) − (2) Change in social welfare
UK taxes	6.00	7.16	–	–
Commodity tax	-1.19	0.32	-1.34	0.17
% difference	*-19.8*	*4.5*	–	–
Segment taxes	-1.43	-0.13	-0.81	0.49
% difference	*-23.9*	*-1.9*	–	–
Type taxes	-1.42	-0.56	-0.15	0.71
% difference	*-23.6*	*-7.8*	–	–
Consumer specific taxes	-0.63	0.35	0.96	1.94
% difference	*-10.4*	*4.9*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for our central calibration of the externality function (Table 4.2). The rows below show the difference relative to the UK system for each tax policy. Column (1) shows the external cost, column (2) the tax revenue, column (3) the change in consumer surplus relative to the UK system, and column (4) the overall change in welfare. All numbers are expressed in £billion per year. Numbers in italic are the percentage differences relative to the UK system.

Table G.2: *Welfare impact of tax changes: high aggregate external cost*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(3) + (4) − (2) Change in social welfare
UK taxes	8.50	7.16	–	–
Commodity tax	-2.83	0.29	-2.28	0.85
% difference	*-33.3*	*4.0*	–	–
Segment taxes	-3.01	-0.11	-1.66	1.25
% difference	*-35.4*	*-1.5*	–	–
Type taxes	-2.87	-0.40	-1.03	1.45
% difference	*-33.8*	*-5.6*	–	–
Consumer specific taxes	-2.11	0.81	-0.58	2.34
% difference	*-24.8*	*11.3*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for our central calibration of the externality function (Table 4.2). The rows below show the difference relative to the UK system for each tax policy. Column (1) shows the external cost, column (2) the tax revenue, column (3) the change in consumer surplus relative to the UK system, and column (4) the overall change in welfare. All numbers are expressed in £billion per year. Numbers in italic are the percentage differences relative to the UK system.

Table G.3: *Welfare impact of tax changes: less convex function*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(3) + (4) − (2) Change in social welfare
UK taxes	7.25	7.16	–	–
Commodity tax	-1.43	0.32	-1.43	0.32
% difference	*-19.7*	*4.5*	–	–
Segment taxes	-1.51	0.03	-1.12	0.42
% difference	*-20.8*	*0.5*	–	–
Type taxes	-1.49	-0.19	-0.82	0.49
% difference	*-20.6*	*-2.6*	–	–
Consumer specific taxes	-0.81	0.85	-0.83	0.84
% difference	*-11.2*	*11.9*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for our central calibration of the externality function (Table 4.2). The rows below show the difference relative to the UK system for each tax policy. Column (1) shows the external cost, column (2) the tax revenue, column (3) the change in consumer surplus relative to the UK system, and column (4) the overall change in welfare. All numbers are expressed in £billion per year. Numbers in italic are the percentage differences relative to the UK system.

Table G.4: *Welfare impact of tax changes: more convex function*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(3) + (4) − (2) Change in social welfare
UK taxes	7.25	7.16	–	–
Commodity tax	-2.28	0.30	-2.04	0.54
% difference	*-31.5*	*4.2*	–	–
Segment taxes	-2.58	-0.18	-1.27	1.13
% difference	*-35.6*	*-2.4*	–	–
Type taxes	-2.48	-0.63	-0.41	1.44
% difference	*-34.2*	*-8.7*	–	–
Consumer specific taxes	-1.69	0.40	0.78	2.86
% difference	*-23.3*	*5.6*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for our central calibration of the externality function (Table 4.2). The rows below show the difference relative to the UK system for each tax policy. Column (1) shows the external cost, column (2) the tax revenue, column (3) the change in consumer surplus relative to the UK system, and column (4) the overall change in welfare. All numbers are expressed in £billion per year. Numbers in italic are the percentage differences relative to the UK system.