Institute for Fiscal Studies

Design of optimal corrective taxes in the alcohol market

IFS Working Paper W17/02

Rachel Griffith

Martin O'Connell

Kate Smith

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Design of optimal corrective taxes in the alcohol market

Rachel Griffith, Martin O'Connell and Kate Smith*

January 30, 2017

Abstract

Alcohol consumption is associated with costs to society due to its impact on crime and health. Tax can lead consumers to internalise these externalities. We study optimal corrective taxation in the alcohol market. We allow for the fact that the externality generating commodity (ethanol) is available in many differentiated products, over which consumers might have heterogeneous preferences, and that there may also be heterogeneity in marginal externalities across consumers. We show that, if there is correlation in preferences and marginal externalities, setting different tax rates across products can improve welfare relative to a single tax rate on ethanol. We estimate a model of demand in the UK alcohol market and numerically solve for the optimal tax rates. Moving to an optimal system that taxes alcohol types at different rates would close half of the welfare gap between the current 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-2015-AdG-694822, the Economic and Social Research Council (ESRC) under the Centre for the Microeconomic Analysis of Public Policy (CPP), grant number RES-544-28-0001, and the British Academy under pf160093. Data supplied by TNS UK Limited. The use of TNS UK Ltd. data in this work does not imply the endorsement of TNS UK Ltd. in relation to the interpretation or analysis of the data. All errors and omissions remained the responsibility of the authors.

*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

Corrective taxes aim to improve welfare when consumption generates externalities. The market for alcoholic beverages is a leading example; alcohol consumption is associated with costs to society from anti-social behaviour, crime and public health care. If the marginal externality that each consumer creates is constant and equal across consumers then a tax equal to the marginal external cost of each unit consumed can fully correct for the externality (Pigou (1920)). However, if marginal externalities are heterogeneous across consumers then a single linear tax can no longer achieve the first best allocation (Diamond (1973)).

Our contribution in this paper is twofold. First, we characterise optimal corrective taxes in differentiated product markets, in which the externality generating commodity (e.g. ethanol) is available in many different products, and consumers are heterogeneous in their tastes for different products, their price responsiveness and their marginal externalities. We show that, if these aspects of heterogeneity are correlated, then varying tax rates across the ethanol in different products can potentially improve welfare relative to a single ethanol tax rate. Secondly, we show that these theoretical results have empirical relevance when applied to the UK market for alcoholic beverages.

We consider the optimal policy of a planner, whose aim is to set taxes to maximise the sum of consumer surplus and tax revenue minus the external costs of consumption. Consumers' demand for ethanol, which is derived from their demand for alcohol products, potentially gives rise to an externality. The marginal externality associated with drinking can be heterogeneous across individuals. This could be because external costs are nonlinear in ethanol consumption, or, for a given level of ethanol there could be heterogeneity in marginal externalities across individuals. Our optimal tax framework nests the classic results in Pigou (1920) and Diamond (1973).[1] In the special case when marginal externalities are constant and equal across consumers, then a single tax rate on ethanol can fully correct for the externality (as in Pigou (1920)). With heterogeneous marginal externalities, the first best cannot be achieved unless the planner can set consumer specific ethanol taxes. If the planner is restricted to set a single rate of tax on ethanol, the optimal policy entails setting the tax rate equal to the weighted average marginal externality of consumers (as in Diamond (1973)).

Varying tax rates across the ethanol in different products can improve welfare relative to a single ethanol tax rate if consumers' preferences are correlated with

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

their marginal externalities of ethanol consumption. The planner is able to target high externality generating consumption by setting relatively high tax rates on ethanol in products that are disproportionately demanded by consumers that generate large marginal externalities. Taxes are more effective when they induce a larger reduction in the demand of consumers that generate high marginal externalities, but their effectiveness is decreasing in the correlation between consumers' willingness to switch to alternative alcohol products and their marginal externalities. Intuitively, if consumers that generate large marginal externalities have steeply sloped demands then taxes are effective at lowering the most social costly consumption. However, the greater the correlation between cross price effects and marginal externalities, the less responsive is the derived ethanol demand of the high externality consumers to tax, making tax less effective.

We apply these ideas to the UK market for alcoholic beverages. We estimate a flexible model of demand, allowing for heterogeneity in consumers' tastes for different products and their price responsiveness. Consumers' demands for different products give rise to their "derived demand for ethanol". We show that there is correlation between consumers tastes for different products (and, hence, their product level demand curves) and their derived demand for ethanol. Consistently heavy drinkers (i.e. those with high derived ethanol demands) systematically purchase a different mix of products than lighter drinkers. They are also much more willing to switch between different alcohol products in response to price changes, but less willing to switch away from alcohol altogether. As a result, they have derived ethanol demands that are substantially less price sensitive than lighter drinkers. We show that optimally set alcohol taxes can result in substantial welfare gains relative to the current tax system. Moving from the UK system to an optimal single ethanol tax rate would result in a welfare gain of around £0.5 billion and would close about 21% of the welfare gap between the UK system and the first best (consumer specific Pigovian taxation). Moving to an optimal system that sets different rates across different alcohol types (allowing, for instance, for different rates on vodka, gin, beer and so on) would do substantially better, closing half of the welfare gap between the UK system and the first best.

Alcohol markets are a natural setting in which to study optimal taxes in the context of differentiated products and heterogeneous externalities. The social costs of alcohol consumption are of concern across the developed world (World Health Organization (2014)). Negative consumption externalities associated with alcohol include public healthcare costs, violent behaviour (e.g. Luca et al. (2015)) and drink driving (e.g. Ruhm (1996), Jackson and Owens (2011), Hansen (2015)). There is considerable evidence that these externalities are nonlinear in ethanol consumption. For example, in the US frequent binge drinkers represent 7% of the population, but drink 45% of the ethanol consumed by adults (US Department of Justice (2005)) and account for around 75% of the cost of excessive alcohol use (Centers for Disease Control and Prevention (2016)). We show that the purchasing patterns of heavy drinkers differ from those of light drinkers. This suggest that optimally set tax rates, that vary across different alcohol types (many countries already tax ethanol differently depending on its form), has the potential to significantly improve welfare.

This paper is related to several strands of the public finance literature.[2] We focus on the ability of alcohol taxes to correct externalities. A number of papers have considered how public goods should be funded, or externalities corrected, when the government's revenue constraint must be satisfied using distortionary taxation (see Bovenberg and Goulder (1996) and Bovenberg and Goulder (2002) for a survey). Kaplow (2012) shows that, under weak separability of leisure from consumption, the income tax system can be adjusted to perfectly off-set any redistributional effects of corrective taxes. Akerlof (1978) noted the potential value of using observable characteristics to identify (or "tag") the potentially needy to improve the efficiency of the benefit system; others have applied this to labour income taxation (e.g. Mankiw and Weinzierl (2010), Weinzierl (2011)). We use a related idea based on the correlation in consumers' preferences for different alcohol products and their marginal externality to "tag" consumption that is likely to have high marginal external costs. Other papers consider the desirability of differential commodity taxation for revenue raising purposes, in the presence of nonlinear income taxation (e.g. Atkinson and Stiglitz (1976), Cremer et al. (2001), Saez (2002), Laroque (2005), Kaplow (2006)). In contrast to these papers, we focus on the potential of differential taxation of products *within* a single commodity to improve welfare by correcting the externalities associated with consumption.

In order to investigate the empirical relevance of this idea we require estimates of consumer demand in the alcohol market. A number of papers apply continuous choice demand methods to alcohol, either treating alcohol as a homogeneous composite commodity (see, inter alia, Baltagi and Griffin (1995), Manning et al. (1995)), or estimating demand over a set of broad alcohol types (e.g. Irvine and Sims (1993), Crawford et al. (1999)). In contrast to these papers we are interested in capturing substitution patterns between differentiated alcohol products and the correlation in demands for different types of alcohol with derived ethanol demand; these are

[2]There is also a literature that embeds environmental taxation into a DSGE framework to study the optimality of corrective taxes (see, for example, Golosov et al. (2014)).

central to our application as they allow for different tax rates across products to potentially improve on a single ethanol tax rate. However, modelling consumer switching across products in the alcohol market is still important even if the social planner is restricted to use only a single ethanol tax rate. This is because a single rate would change relative prices of different alcohol products (because different products contain differing amounts of ethanol), which leads to switching between products and hence changes in derived ethanol demand. We use a discrete choice demand framework, avoiding the econometric problems that arise in continuous choice models from zero purchases of many of the products available in the market.

Our demand estimates enable us to predict consumers' product, and hence derived ethanol demands, at counterfactual tax rates. To solve for optimal tax rates we combine the estimates with a mapping between a consumer's ethanol consumption and the externality they generate. We use government estimates of the aggregate externality along with medical evidence on the shape of the externality function. We show how the optimal tax rates depend on the assumptions that one makes about the size of the aggregate externality and the shape of the externality function. A higher aggregate externality implies that optimal tax rates are all scaled proportionately upwards, while increasing the convexity of the function (implying the marginal externality increases more strongly in ethanol consumption) increases the differential in optimal rates between lightly and heavily tax products. In all cases we find that welfare improves if we move from the current tax system to an optimal single rate and improves further by moving to optimal tax rates that differ across alcohol types.

Recent papers by Miravete et al. (2016) and Conlon and Rao (2015) complement our work and study the supply side of the US market for spirits. These papers consider how government regulations, in part designed to limit alcohol consumption, interact with firm conduct. Miravete et al. (2016) show that 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

[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.

instead be used to raise the price level and would have the advantage of raising tax revenue. In contrast, we model demand in the entire alcohol market and consider the design of optimal tax policy. We do not directly model the supply side, but discuss the robustness of our results to alternative supply side assumptions.

The rest of the paper is structured as follows. In the next section, we discuss the design of corrective taxes in markets with heterogeneous consumers and with many products that potentially generate externalities. We are primarily interested in alcohol taxes, but these results apply in other markets. In Section 3 we present our empirical model of consumer behaviour and estimates based on data on the UK alcohol market. We use these estimates along with our optimal tax framework to compute optimal tax rates. We present these in Section 4 and compare them to the current UK tax system. A final section summarises and concludes. Several appendices provide additional details.

2 Corrective tax design

2.1 Consumer demand for alcohol

Let $i \in \{1, ..., N\}$ index consumers; each consumer has income y_i. Let $j \in \{1, ..., J\}$ index alcohol products, available at post tax prices $\mathbf{p}_i = (p_{i1}, \dots, p_{iJ})'$, each containing z_j ethanol (pure alcohol) and a vector of other characteristics $\mathbf{x}_j$. We consider as a benchmark consumer specific taxes, hence the i subscript on prices.

We assume consumer indirect utility is quasi-linear in the numeraire good and is given by

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

where α_i is the marginal utility of income and $v_i(\mathbf{p}_i, \mathbf{z}, \mathbf{x})$ is the indirect utility that arises from the alcohol demands for consumer i. We denote the consumer's demand for product j by $q_{ij} = f_{ij}(\mathbf{p}_i, \mathbf{z}, \mathbf{x})$ and the consumer's vector of demands by $\mathbf{q}_i = (q_{i1}, \dots, q_{iJ})'$. Quasi-linear utility means alcohol demands do not depend directly on income; however heterogeneity in preferences (including the marginal utility of income) allows for demand functions to vary flexibly across consumers.

We discuss in more detail the empirical specification of demand in Section 3.2.

2.2 External costs

It is widely accepted that alcohol consumption generates costs that are not considered by individuals when making alcohol consumption decisions. These include costs associated with alcohol induced public health care, violent crime, domestic violence, road accidents and future unanticipated health consequences. This leads to excess consumption from a social perspective, and justifies government intervention.

External costs may be heterogeneous across individuals and there may be nonlinearities in the effect of consumption on external costs; we capture the external cost associated with an individual's alcohol consumption by the function $\phi_i(.)$. There is considerable evidence that the external costs of alcohol consumption are increasing in the total consumption of ethanol, so we specify that the argument of $\phi_i(.)$ is derived ethanol demand. Derived ethanol demand is a function of the consumer's demand for individual alcohol products; we denote it by $Z_i = \sum_j z_j q_{ij}$. The total external cost from all consumers in the market is then:

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

An implication of this form of external cost function is that, conditional on total ethanol demand, the marginal externality from drinking a unit of ethanol is the same across different types of alcohol. Individuals with a high marginal externality, might consume different alcoholic beverages, but, 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.

2.3 Social planner's problem

We consider the social planner's problem of choosing alcohol taxes to maximise 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). We consider specific (or unit) taxes levied on ethanol content.

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. Kaplow (2012) shows that by accompanying externality correcting taxes with a distribution-neutral adjustment to the income tax system, we can offset the effects of the corrective taxes across the income distribution. However, it is harder to correct for the impact of the taxes *within* income classes, if consumers with the same income have different preferences (Kaplow (1996)). 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 tax 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:

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

2.4 Characterising tax policy

We begin by showing how the results of Pigou (1920) and Diamond (1973) can be derived as special cases in this 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 optimal policy prescribes rates that vary across products. This gets (weakly) 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. The optimal tax rates that differ across products are defined implicitly by a system of J equations; they depend on the full matrix of own and cross price effects and their correlation with individuals' marginal externalities. To provide intuition for how the correlation between the marginal externality and shape of demand for products affects the optimal tax rates, we derive analytical expressions for the optimal product taxes under some simplifying assumptions.

Consumer specific taxation

Suppose the planner can set different tax rates for each consumer. Let τ_i denote the tax rate for consumer i and $\boldsymbol{\tau} = (\tau_1, ..., \tau_N)'$. 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 $R(\boldsymbol{\tau}) = \sum_i \tau_i \sum_j z_j q_{ij}(\tau_i)$. 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:

$$\tau_i^* = \phi_i', \tag{2.4}$$

where $\phi_i' \equiv \phi_i'(Z_i)$ is the slope of the marginal externality function. The optimal consumer specific tax rate is set to equal the consumer's marginal consumption externality 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 that 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: $\tau'(.) = \phi'(.)$.

Single ethanol tax rate

Now suppose the planner is unable to set consumer specific taxes, and must instead choose one common tax rate to apply to all products. 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, $R(\tau) = \tau \sum_i \sum_j z_j q_{ij}(\tau)$. Denote the slope of demand for derived ethanol as $Z_i' \equiv \frac{\partial Z_i}{\partial \tau} = \sum_j z_j \frac{\partial q_{ij}}{\partial \tau}$. 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^*:

$$\tau^* = \frac{\sum_i \phi_i' |Z_i'|}{\sum_i |Z_i'|} = \overline{\phi}' + \frac{\text{cov}(\phi_i', |Z_i'|)}{|\overline{Z}'|}, \tag{2.5}$$

where $\overline{\phi}' \equiv \frac{1}{N} \sum_i \phi_i'$ and $\overline{Z}' \equiv \frac{1}{N} \sum_i Z_i'$ are the average marginal externality and ethanol demand slopes. The first expression for τ^* 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 expression is an alternative formulation, 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 rate. The intuition is that the more consumers that generate high levels of externalities at the margin have ethanol demands that are particularly sensitive to tax changes, the more effective is the tax at reducing the most socially costly consumption.

Products level tax rates

Now suppose the planner can vary tax rates across products 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 $R(\boldsymbol{\tau}) = \sum_i \sum_j \tau_j z_j q_{ij}(\boldsymbol{\tau})$. Taking the first order condition for τ_j yields:

$$\sum_i \sum_k (\tau_k - \phi'_i)\, z_k \frac{\partial q_{ik}}{\partial \tau_j} = 0 \tag{2.6}$$

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 $\phi'_i(Z_i) = \phi'$. 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.

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 product j as $Z_{ij} = z_j q_{ij}(\tau_j)$ (and the slope of this demand $Z'_{ij} = z_j \frac{\partial q_{ij}}{\partial \tau_j}$); we can then write the optimal tax on product j as:

$$\tau^*_j = \frac{\sum_i \phi'_i |Z'_{ij}|}{\sum_i |Z'_{ij}|} = \overline{\phi}' + \frac{\text{cov}(\phi'_i, |Z'_{ij}|)}{|\overline{Z}'_j|} \tag{2.7}$$

In this case optimal taxation reduces to the single rate optimal tax formula applied to each product. Products with a strong positive covariance between demand slopes and marginal externalities will have higher optimal taxes. If the correlation between demand slopes and the marginal externality is zero, then setting different rates across products cannot improve on the optimal single rate. 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 product j as $w_{ij} = \frac{Z_{ij}}{\sum_{i'} Z_{i'j}}$ and its own price elasticity for product 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 product 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 product j. All else equal, if a product has relatively high demand among consumers that generate a high marginal 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 marginal externalities are most willing to switch away in response to a price rise.

Third, consider the case in which there are two alcohol products 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.1. All else equal, increasing the covariance between the marginal externality and absolute value of the own slope of demand for product i increases the optimal tax on product i and the optimal tax on product j but by less than for product i. As above, if consumers that generate large marginal 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 products i and j decreases the optimal rate on both products, and decreases by more the optimal tax rate on the product with the largest own slope of demand. The greater the correlation between the cross price effects for the two products and the marginal externality, conditional on the correlation between the price sensitivity of own demands and marginal externalities, 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 products 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 Application to the UK alcohol market

We apply the optimal tax policies derived in Section 2 in an empirical setting by considering the UK alcohol market.

To compute the optimal taxes we require (i) flexible estimates of consumer demand in the alcohol market, and, in particular, estimates of how individuals' demand for different products correlates with the marginal externality associated with their consumption, and (ii) the arguments and form of the externality function.

3.1 Data

Our data are from the Kantar Worldpanel – a panel of households selected to be representative of the British population. These data have the significant advantage that they contain rich product information, repeated observations for each household, and accurate prices. Each participating household uses a hand held scanner to record all grocery products, at the UPC level, that are purchased and brought into the home. The data include details of exact transaction prices and product size, alcohol type and strength.[4] The data cover all alcohol purchases made in grocery stores and liquor stores (together known as "off-trade"), which constitute 77% of the alcohol market according to the Living Costs and Food Survey (see Appendix B.1 for a description of this data). The data exclude those purchases made in pubs and restaurants (together known as "on-trade"). In Appendix B.2 we show that purchase patterns off-trade are similar to those on-trade. In particular, we show that for all ethanol purchases and on-trade only ethanol purchases, households that buy relatively large quantities of ethanol get a relatively high share of the their ethanol from spirits and a relatively low share from beer and from cider. For a more detailed description of the data, see Griffith and O'Connell (2009) and Leicester and Oldfield (2009); Griffith et al. (2013) contains more information on the alcohol segment of the data.

We use a sample of 10,289 households, which we observe repeatedly throughout calendar year 2011. Each household is in the data for a minimum of 30 weeks over

[4]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 flavoured alcoholic beverages (FABs). 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. In a small number of cases, we are unable to find the ABV of a product so we use the average ABV for drinks of that type, provided by the UK's Office for National Statistics (Goddard (2007)).

the course of the year and is observed purchasing alcohol at least 3 times.[5] We also observe alcohol purchases made by each household in calendar year 2010. We estimate our model using data from 2011, so we refer to information from 2010 as pre sample information. We compute how much ethanol each household purchased per adult (aged 18 or over) per week on average in the pre sample period. 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.[6]

Table 3.1: *Sample of households*

Household group:	% of Households	% of total ethanol	Avg. weekly ethanol units
Less than 7 units	62.5	23.2	2.9
7-14 units	17.9	20.4	9.2
14-21 units	7.9	15.3	15.6
21-35 units	7.2	19.8	23.9
More than 35 units	4.5	21.2	47.1
Total	100.0	100.0	8.6

Notes: For each household we calculate the average number of units purchased per adult per week. The numbers shown in column 3 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 3.1 we show the fraction of households in our sample that belong to each group. The table shows that the distribution of households across pre sample purchase groups is highly skewed. Most households purchase a moderate number of units of ethanol, however, 7.2% of households consistently purchased between 21-35 units of alcohol per adult per week 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 – households that consistently purchased more than 21 units of alcohol in the pre sample period comprise 42% of all 2011 ethanol purchases. The final column 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. Appendix B.3 contains a detailed

[5] 4,477 households are observed for a minimum of 30 weeks 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.

[6] One unit of ethanol contains 10 millilitres (8 grams) of ethanol. In the US a standard drink measure of spirits is 1.5 fl oz. (or 44 millilitres). For a spirit that has strength 40%, a standard drink contains 17.7 millilitres (or 1.77 units) of ethanol.

breakdown of the proportion of ethanol that the household groups get from different alcohol types; households that buy consistently buy high quantities of ethanol get disproportionately more from high strength spirits, and less from cider and beer.

3.2 Estimating consumer demand

Our objective is to estimate demand in the alcohol market in a flexible way, allowing us to uncover how the shape of households' product level demands is correlated with marginal externalities associated with their consumption.

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. When households purchase alcohol they typically choose one or a small number of options. On 37% of purchase occasions, the household purchases more than one (typically two or three) alcohol products. 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)). We therefore model demand as a discrete choice of which – if any – alcohol option to select. A discrete choice demand framework 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. On occasions on which a household purchases more than one product, we treat this behaviour as the household making multiple separate purchase decisions.

We estimate two variants of the demand model. First, we estimate demand using the standard form of preference distributions used in the literature. Several papers demonstrate that using random coefficients in discrete choice models to capture preference heterogeneity allows the demand system to flexibly capture switching patterns across products (e.g. see Berry et al. (1995) for market level data and Berry et al. (2004) and Train (2003) for consumer level data). Second, we exploit the panel structure of the data to estimate a variant that permits additional flexibility in the distribution of consumer preferences, and its correlation with derived ethanol demand (and hence, the marginal externality of alcohol consumption).

We index households by i and products by j. $j = 0$ denotes the option of purchasing no alcohol, $j = 1, ..., J$ indexes different alcohol products. In this application it is also important to capture the decision households make over quantity. We embed this in the discrete choice model by considering a purchase choice of product and discrete sizes, indexed by s. We thus model the decision over purchasing a product-size, (j, s), with the option to purchase no alcohol denoted $(0, 0)$. We use t to index weeks.

Household preferences are defined over characteristics of products, both observed (Gorman (1980), Lancaster (1971)) and unobserved (Berry (1994), Berry et al. (1995)). We assume that the utility household i obtains from selecting option (j, s) in period t is given by:

$$u_{ijst} = \nu(p_{jst}, z_{js}, \mathbf{x}_{jst}; \theta_i) + \epsilon_{ijst}, \tag{3.1}$$

where p_{jst} is the price of option (j, s) in week t, z_{js} is the ethanol content, $\mathbf{x}_{jst}$ 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 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 $u_{i00t} = \epsilon_{i00t}$.

Households select the option (j, s) that provides them with the highest utility. 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:

$$q_{ijst} = \frac{\exp(\nu(p_{jst}, z_{js}, \mathbf{x}_{jst}; \theta_i))}{1 + \sum_{j'>0,s'>0} \exp(\nu(p_{j's't}, z_{j's'}, \mathbf{x}_{j's't}; \theta_i))} \tag{3.2}$$

and expected utility is given by:

$$v_{it}(\mathbf{p}_{jt}, \mathbf{z}_{jst}, \mathbf{x}_{jst}) = \ln \sum_{j>0,s>0} \exp\{\nu(p_{jst}, z_{js}, \mathbf{x}_{jst}; \theta_i)\} + C \tag{3.3}$$

where C is a constant of integration that differences out when comparisons are made across two different tax regimes. Equations 3.2 and 3.3 give the expressions for q_{ij} and v_i used in Section 2.

We assume the function ν takes the form:

$$\nu(p_{jst}, z_{js}, \mathbf{x}_{jst}; \theta_i) = \alpha_i p_{jst} + \beta_i w_j + \sum_{m=1}^{4} 1[j \in \mathcal{M}_m] \cdot (\gamma_{i,1m} z_{js} + \gamma_{i,2m} z_{js}^2) + \xi_{ijt}. \tag{3.4}$$

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 $m = 1, ..., 4$. $\mathcal{M}_m$ denotes the set of options that belong to segment m. This allows for the possibility that households might value larger or smaller quantities of ethanol 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).

We allow for heterogeneous preferences over observable product attributes and over the unobserved product attributes. The parameters $(\alpha_i, \beta_i, \boldsymbol{\gamma}_i)$ are the household specific weight that the household places on product price, strength and ethanol content when making their purchase decision.

The unobserved product characteristic includes a time invariant component that varies across households, and a time varying component that varies over alcohol-types, indexed k_j,[7] so that

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

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 preferences over the unobserved product attributes. We make two alternative assumptions about the distribution of these – 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. The objective of estimation is to recover the parameters governing these distributions.

The standard approach in the literature is to assume that the distribution of Ψ_i and $\boldsymbol{\eta}_i$ is a multivariate normal distribution. We do this allowing for the possibility that households have correlated tastes for price, alcohol strength, and ethanol content. We refer to this as the "normal preference distribution" specification.

The detailed nature of our data enables us to relax the distributional assumptions commonly made in discrete choice models. 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 purchase histories. In particular, instead of modelling the preference distributions as normal, we model them as a mixture of conditional normal distributions, conditioning on the five groups based on households' pre sample ethanol purchases shown in Table 3.1. If pre sample behaviour is informative about preferences and within sample behaviour then this enables us to more flexibly capture the prefer-

[7] There are 19 alcohol-types, which are (slightly) more aggregate than the products, and are shown in Table B.1. For instance, gin is one alcohol type comprising the alcohol products store brand gin and branded gin.

ence distribution. Conversely, if pre sample behaviour is not informative, the model collapses to the standard normal preference distribution case. We refer to this as the "mixed-normal preference distribution" specification. Further details on this are provided in Appendix C.1.

We have a representative sample of 10,289 households, which we divide into five groups based on their pre sample derived ethanol demand, these are shown in Table 3.1. For the normal preference distribution specification, we randomly draw 1000 households and 25 purchase observations for each household. For the mixed-normal preference distribution specification, we randomly draw 450 households from each group and 25 purchase observations for each household. When computing market level demand and elasticities, we re-weight to take account of this sampling. We estimate demand using maximum simulated likelihood. 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 3.2. This follows from our assumption that the ϵ_{ijst} are i.i.d. type I extreme value. To construct the likelihood function we 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:

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

No closed form for the integral exists, so we use simulation methods.

3.3 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 in which choice sets contain 7000 options is infeasible. We therefore define the options in households' choice sets by aggregating the 7000 UPCs into 79 product-size pairs. 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. For 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.[8]

We group the UPCs into 40 alcohol products. We discretise the quantity choice of households by defining a set of sizes for each product.[9] An option is therefore a product-size pairing. There are 79 alcohol options in total (plus the outside option) – these are listed in Appendix B.4. For each option we construct a price index defined over the prices of UPCs that constitute the option. The weights in each price index are held fixed, so movements in the price index reflect only movements in the (average weekly) underlying prices of the UPCs and not changes in the composition of UPCs chosen by households.

3.4 Identification of demand parameters

We use longitudinal micro level data, so we observe each household making repeated choices. This type of 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-product effects) across time. It is this variation that allows us to estimate the parameters governing the preference distribution.[10]

The key identification issue is whether we identify the causal effect of a change in price on demand. 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.

[8]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.

[9]We use these discrete categories to compute the quantity of ethanol purchased per adult per week and are able to closely replicate the observed average quantity for each household group (based on pre sample ethanol demand).

[10]Formally, Berry and Haile (2010) and Fox and Gandhi (2016) establish conditions for non-parametric 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.

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 (which in the mixed-normal preference distribution specification, we also allow to vary across groups of consumers based on their pre sample ethanol purchases). These time effects absorb seasonality in demand and spikes in demand due to advertising campaigns. Prices in the UK grocery market are set nationally, there is limited scope for firms to set geographically varying prices in response to local demand shocks.[11] We expect these rich time effects to absorb the majority of aggregate shocks to alcohol demand that are 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)). We use a set of cost shifters as instruments; these include producer prices for beer and for cider, the sterling-euro and sterling-dollar exchange rates, alcohol duty rates, and wage indices in the retail and the food manufacturing sectors. 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 that we identify the effect of price on demand using differential (across time and products) variation in prices that is driven only by cost shocks. 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 that the rich time effects in demand are indeed doing a good job of absorbing shocks to demand.

3.5 Demand estimates

In Table 3.2 we report the coefficient estimates for our demand model. The first column of the table shows estimates for the "normal preference distribution" specification of the model. The remaining columns present the estimates for the "mixed-normal preference distribution" specification. Panel A shows estimated parameters of the random coefficient distributions over the observable product characteristics

[11]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)).

and Panel B summarises estimates of the parameters governing preferences over unobserved product attributes.

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. The mean of the strength coefficient is not separately identified from the product effects so we normalise it to zero. The variance parameters in column 1 indicate that 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 3.2 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).[12] We restrict the covariance matrix of the unobserved product effects to be diagonal and for 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. 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 the fact that we normalise the no purchase outside option utility to zero. Higher numbers in absolute terms indicate stronger preferences for the set of products in that segment. The variance parameters vary across alcohol segments, indicating that the intensity of within- versus between-alcohol segment substitution varies across segments.

[12] We do not show estimates of the alcohol type-time effects; they are jointly highly statistically significant. They are available from the authors upon request.

Table 3.2: *Estimated preference parameters*

	(1)	(2)	(3)	(4)	(5)	(6)
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 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.

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 basis of pre sample behaviour (see Table 3.1). This allows the correlation in preference parameters over observable attributes to vary by the 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 lighter drinkers to heavier drinkers, the estimated mean product effects in all segment increases – heavy drinkers, on average, have higher valuations of unobserved products attributes for all alcohol products relative to the no purchase outside option than lighter drinkers.

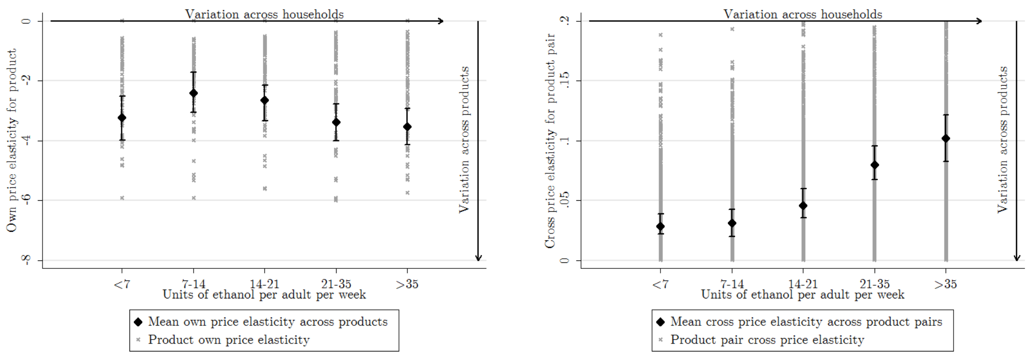
In the next section and in Section 4, we use the mixed-normal preference distribution specification, because of its additional flexibility. In Appendix C.1, we show that the optimal taxes computed under the more restrictive normal preference distribution specification have a similar pattern to those computed with the mixed-normal distribution specification.

3.6 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. The model generates an 80×79 matrix of elasticities for every household group. In Figure 3.1 we summarise this information.

Figure 3.1: *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 five household groups. The black marks are averages across these product level elasticities. The bars are 95% confidence intervals.

Panel (a) shows the own price elasticities and panel (b) shows the cross price elasticities. The vertical variation in the graphs is across products and the horizontal variation is across the five household groups, defined based on pre sample ethanol demand. The grey dots represent the own price elasticities for each product (in panel (a)) and the cross-price elasticities between pairs of products (in panel (b)). The black dots represent the mean elasticity for the particular group of households; the bars are 95% confidence intervals.[13] 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 drinkers 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).

In Table 3.3 we focus on the spirits (including fortified wine) segment of the market and summarise mean own and cross price elasticities for each of the household groups. For elasticities for all segments, as well as confidence intervals, see Appendix C.2. Spirits products are disproportionately purchased by heavy drinkers and, as the table shows, spirits elasticities vary significantly across the 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 table also highlights 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.

[13] We calculate confidence intervals by first obtaining the variance-covariance matrix for the parameter vector estimates using standard asymptotic results. We then take 100 draws of the parameter vector from the joint normal asymptotic distribution of the parameters and, for each draw, compute the statistic of interest, using the resulting distribution across draws to compute Monte Carlo confidence intervals (which need not be symmetric around the statistic estimates).

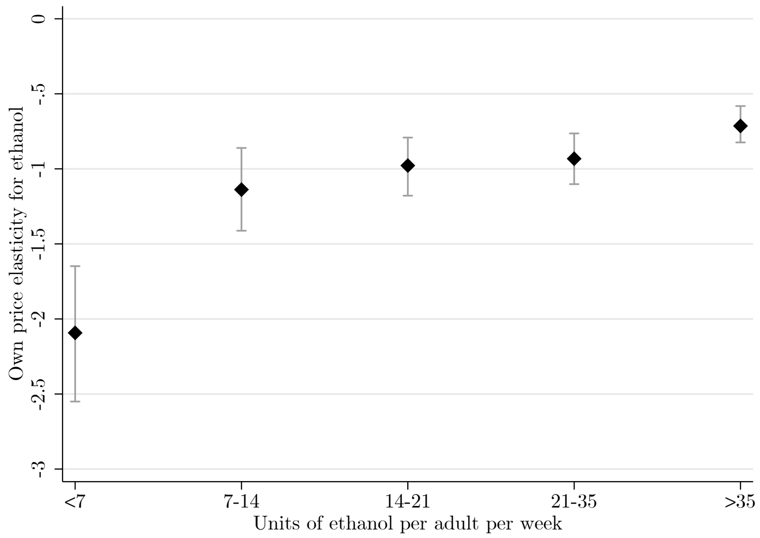
Table 3.3: *Average own and cross price elasticities for spirits products*

	(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 each group of households defined on pre sample ethanol purchases. 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 Appendix C.2.

The price sensitivity of a household's derived ethanol demand to an increase in the price of all alcohol products depends on all their own and cross price elasticities across the alcohol options. We simulate the slope of ethanol demands by uniformly marginally 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 the substitution across all products and sizes that underpin changes in total ethanol demand.

Figure 3.2: *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 does show the mean elasticities across months and households in each household group. The bars show 95% confidence intervals.

Figure 3.2 plots the own elasticity of demand for ethanol for each household group. It shows that, in response to a uniform 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 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.

3.7 Externality function

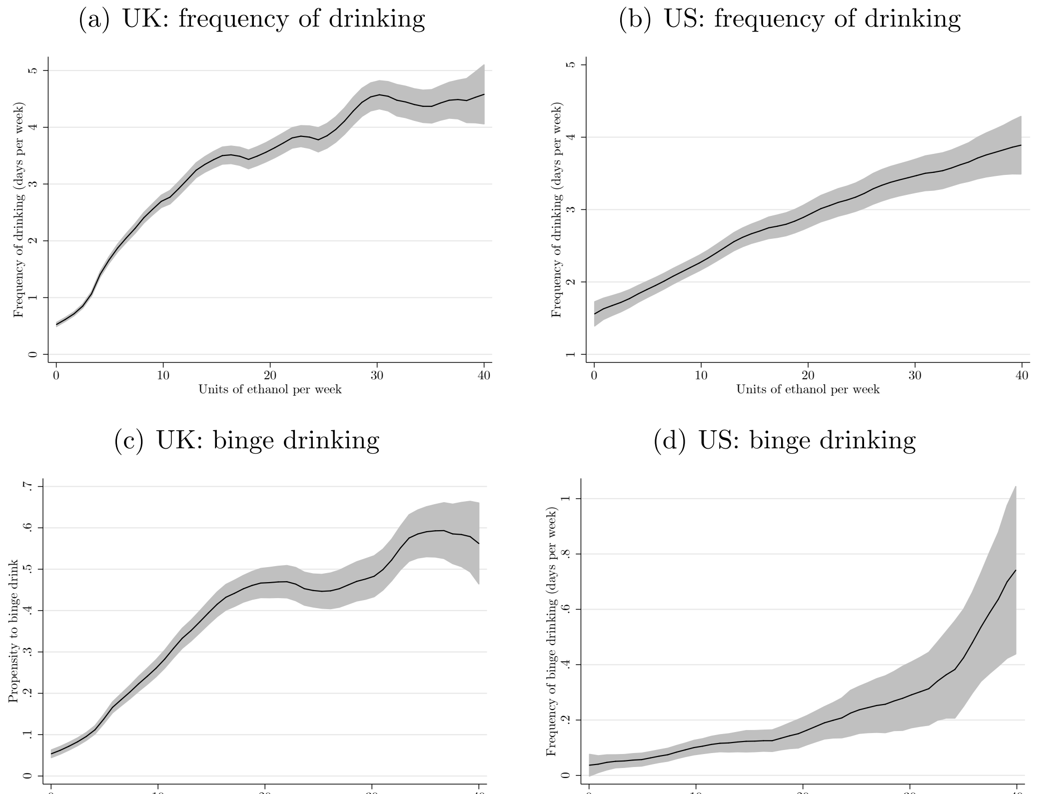
The principal objective of policy intervention in alcohol markets is to reduce the external costs associated with drinking, including the costs of crime (e.g. drink driving, anti-social behaviour, domestic violence) and the public health care costs of treating alcohol related illnesses. The World Health 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. The US Department of Justice (2005) estimate that more than 75,000 deaths are attributable to alcohol consumption each year and the total economic costs associated with alcohol problems total more than $184 billion annually.

In Section 2 we specified the consumer level externality function, ϕ_i, as a function of derived ethanol demand Z_i. We use data on alcohol purchases made by households, so in the empirical implementation we convert total ethanol demand into ethanol demand per adult (person aged 18 or over). We also place some additional structure on the externality function. We assume that ϕ_i is an increasing convex function. This means that a marginal increase in ethanol consumption has associated with it an incremental cost to society that is increasing in the level of ethanol consumption.

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). In Figure 3.3 we show the relationship between total ethanol per week and both the frequency of drinking and the propensity to binge drink. Panels (a) and (b) show that in both the UK and US people that report consuming higher amounts of ethanol also report drinking more days per week. Panels (c) and (d) show that in both countries there is a positive relationship between consumers' total ethanol and whether they reported binge drinking in the previous week.

A large body of evidence suggests that the external costs of drinking are highly concentrated in a small proportion of heavy drinkers (see Cnossen (2007) for surveys). Relatedly, there is a considerable amount of evidence that marginal externalities from alcohol are an increasing function of ethanol consumed. For example, there is evidence of a threshold effect with some diseases: at low levels of ethanol 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, and Rehm et al. (2010) for evidence on liver cirrhosis).

Figure 3.3: *Ethanol consumption, binge and high frequency drinking*



(a) UK: frequency of drinking

(b) US: frequency of drinking

(c) UK: binge drinking

(d) US: binge drinking

Notes: Panels (a) and (c) are drawn using data from the alcohol questionnaire component of the HSE. Panels (b) and (d) are drawn using data from the alcohol questionnaire and food and drink diary components of NHANES. For a detailed description of the data and variable construction, see Appendix B.1.

The marginal externality of ethanol consumption may vary across people for two reasons. First, some people may drink more, and therefore, if the externality function is convex, then they will have higher marginal externalities of consumption. This is captured by a convex externality function that is common across people. Second, conditional on a given level of ethanol consumption, some people may generate higher social costs. For example, there is some evidence that low socioeconomic status (SES) households are more likely to generate external costs than higher SES households, conditional on the same drinking patterns (Grittner et al. (2012)). On the other hand, differences in the marginal externality generated across gender are primarily due to differences in the amount drunk. The World Health Organization (2014) argue that men are responsible for more of the external costs because they drink more, and that, *"when the number of health and social consequences is considered for a given level of alcohol use or drinking pattern, sex differences for social outcomes reduce significantly"*.

We parameterise the externality function as an exponential function with parameters, (ϕ_{0i}, ϕ_{1i}):

$$\phi(Z_{it}) = \phi_{0i}(\exp(\phi_{1i}Z_{it}) - 1). \quad (3.5)$$

Subtracting one from the term in the brackets ensures that the external cost of zero ethanol demand is zero. ϕ_{1i} controls the convexity, measured as the ratio of second to first derivatives, and given ϕ_{1i}, ϕ_{0i} governs the aggregate external cost.

Although there is considerable evidence that the external costs of consumption are increasing and convex, there is debate over the size of aggregate external costs and even less evidence on the degree of convexity. In our central calibration we set 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 that the direct tangible social costs are £7.25 billion (in 2011 prices).[14] Direct tangible costs of drinking include costs of alcohol-related disease and the costs of dealing with alcohol-related crime. In our central calibration, we calibrate the convexity of the function to be in line with the estimate of an almost 18 times increase in the probability of an accident after consumer 140g rather than 14g of ethanol found in Taylor et al. (2010).

For our central calibration we assume that (ϕ_{0i}, ϕ_{1i}) are common across households. However in an alternative calibration, we also allow for the marginal exter-

[14]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 the fact that we are using data on alcohol purchases excluding those made in restaurants and pubs (off-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%).

nality to vary by SES, conditional on derived ethanol demand. In line with the estimates in Grittner et al. (2012), we calibrate the function such that ϕ_{0i} for the high SES households is 70% of ϕ_{0i} for the low SES households.

As well as the central calibration and its SES variant, we show results for four alternative calibrations, which are described in Table 3.4. We consider the implications of higher or lower aggregate external costs and if the externality function is more or less convex.

Table 3.4: *Calibration specifications*

	Aggregate external cost (£billion)	Ratio of external costs of heaviest to lightest drinkers	Calibrated parameters (ϕ_{0i}, ϕ_{1i})
Central	7.25	20	$(1.2980, 0.0615)$
Alternative specifications			
By SES	7.25	20	High SES: $(1.1650, 0.0615)$ Low SES: $(1.6580, 0.0615)$
Low aggregate cost	6.00	20	$(1.0740, 0.0615)$
High aggregate cost	8.50	20	$(1.5220, 0.0615)$
Low convexity	7.25	10	$(3.1730, 0.0435)$
High convexity	7.25	30	$(0.8177, 0.0695)$

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.

3.8 Current systems for taxing alcohol

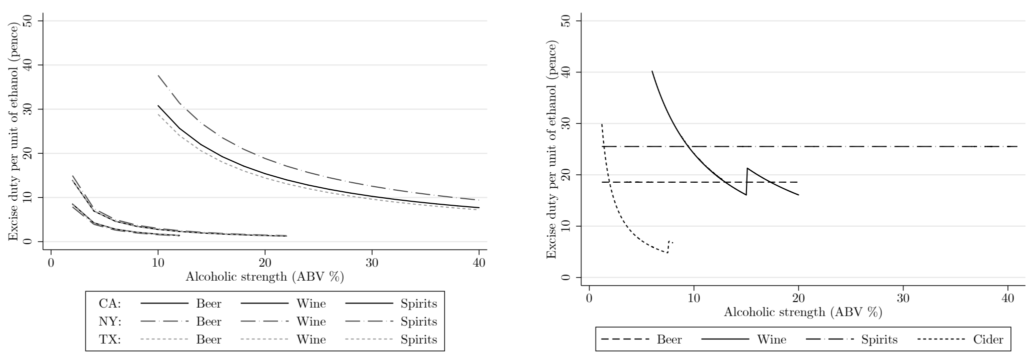
The majority of developed countries tax alcohol; most commonly via specific (unit) excise taxes. Typically alcohol is also subject to broad based sales taxes or value added tax (VAT). Chetty et al. (2009) find that taxes included in posted prices (i.e. excise taxes) reduce alcohol consumption by more than increases in taxes applied at the register (i.e. sales taxes). In the US, alcohol taxes are levied per litre of product (volumetric tax), with different rates for beer, wine and spirits. There are federal alcohol excise taxes in addition to rates that are set by the different states. Volumetric taxes such as these lead to higher strength alcoholic beverages being taxed less heavily per unit of ethanol. To illustrate, Figure 3.4(a) shows the excise taxes (expressed per unit of ethanol) for California, New York and Texas.[15]

[15] 17 states are “control” states, in which the retail sale of (at least one type of) alcohol is controlled by a state run monopoly; there is upstream competition between distillers and alcohol manufacturers. There are also 24 states in which “post and hold” restrictions are in place (again, for at least one type of alcohol); this limits how quickly firms can alter their prices. We abstract

In European countries spirits and beer are taxed per unit of ethanol, while wine and cider are taxed volumetrically, with banded rates (for different strength products) introducing kinks into the tax schedule. To illustrate, Figure 3.4(b) shows how the excise tax in the UK, 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.

Figure 3.4: *Alcohol taxes*

(a) US (selected states) (b) UK



Notes: In the UK, fortified wines with ABV below 22% are taxed as wines and those with ABV above 22% are taxes are spirits and FABs are taxed as spirits. The tax rates in the US include both federal and state alcohol excise taxes.

In the UK the post tax price faced by consumers is $p_{jst} = (1 + t^{VAT})(\tilde{p}_{jst} + t_j^{EX})$ where $\tilde{p}_{jst}$ is the pre tax price, 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. In demand estimation, for each option, we use the post tax price faced by consumers, p_{jst}. In the optimal tax problem we use the pre tax price, $\tilde{p}_{jst}$; the post tax price is $p_{jst} = \tilde{p}_{jst} + \tau_j^* z_{js}$, where τ_j^* denotes the optimal tax rate on product j.

We do not explicitly model the decision over which rate of VAT to set. With tax rates that vary across products we can divide 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, and obtain the system of optimal excise taxes given the prevailing VAT system. We present optimal taxes prior to adjusting for VAT.

from supply side issues in our optimal tax analysis, but we discuss the robustness of our results to various supply side environments in Section 4.3.

4 Optimal alcohol taxes

We use the empirical framework outlined in Section 3 to solve for the tax solutions to the planner's problem set out in Section 2. We solve for both the first best Pigovian consumer specific taxes and for optimal tax rates when rates are constrained to be common across consumers.

4.1 Optimal tax rates

We solve for the optimal rates in four alternative tax regimes:

1. **Consumer specific tax rates**: Each consumer faces its own personalised rate implicitly defined by equation 2.4. This implies Pigovian taxation for each consumer and the first best allocation.

2. **Single ethanol tax rate**: The rate is implicitly defined by equation 2.5.

3. **Product level tax rates**: Tax rates are allowed to vary across products. In the case of product specific taxes the optimal rates are implicitly defined by the set of equations 2.6. We consider two forms of product level taxation:

(a) **Segment taxes**: rates are allowed to be differ across the four segments of the alcohol market (beer, wine, spirits, cider), but are constrained to be the same across products within each segment.

(b) **Type taxes**: rates are allowed to be differ across a set of 18 alcohol types. These types are more aggregate than the products included in our demand system. They impose a common rate across products within each type.

The segment taxes are the most similar to what governments currently do. In both Europe and the US, governments set different tax rates across the four segments, although in most cases these taxes are volumetric rather than levied directly on ethanol. The most flexible tax rates that differ across products that we consider are the type taxes. These allow for different rates on, for instance, gin, vodka etc., but they do not differentiate between branded and store brand products. They also do not set different tax rates for different sizes of the same product. We do not consider a system of more disaggregate product tax rates because we expect this system is infeasible for governments to implement (see Gillitzer et al. (2015) on the difficulties of tax policies that distinguish between very disaggregate products). This choice does not reflect any technical difficulty with solving for more disaggregate tax rates – doing so leads to similar welfare predictions. It is important to note however, that a key feature of our demand framework is that it captures substitution across sizes and products in response to any simulated form of tax, whether this be a single rate or rates that vary across segments and types.

Table 4.1: *Tax rate solutions*

(1)		(2)		(3)		(4)	
Current		Optimal					
				Product level tax rates			
UK tax rates		Single rate		Segment		Type	
Ale	27.7	Ethanol	35.9	Beer	28.3	Ale	22.7
Lager	26.1		·	(inc. lager and ale)	·	Lager	28.7
Stout	29.3		·		·	Stout	23.1
Red wine	26.7		·	Wine	30.4	Red wine	30.6
White wine	31.0		·		·	White wine	29.8
Rose wine	31.9		·		·	Rose wine	24.3
Brandy	32.5		·	Spirits	42.6	Brandy	37.4
Gin	31.7		·	(inc. fortified wine)	·	Gin	42.4
Rum	32.0		·		·	Rum	38.5
Vodka	31.3		·		·	Vodka	44.4
Whisky	32.4		·		·	Whisky	43.2
Liqueurs	35.4		·		·	Liqueurs	19.9
Port	24.4		·		·	Port	16.8
Sherry	24.3		·		·	Sherry	20.4
Vermouth	23.6		·		·	Vermouth	20.8
Other fort. wine	22.2		·		·	Other fort. wine	22.2
Cider	13.8		·	Cider	25.4	Cider	25.2
FABs	39.6		·	(inc. FABs)	·	FABs	16.5

Notes: Column (1) shows the average UK tax rate for each alcohol type (inclusive of VAT and excise duty). Columns (2)-(4) show the tax rates (expressed in pence per unit of ethanol) that maximise consumer welfare (equation (2.3)). Column (2) shows the optimal single tax rate. Columns (3) and (4) show the optimal product level tax rates, at the segment and type level, respectively. The dots represent the tax rate shown in the row above. 95% confidence intervals for the optimal tax rates are shown in Table C.3 in Appendix C.3.

In Table 4.1 we show the optimal rates for the single rate, segment and type taxes (we discuss the consumer specific taxes below), along with the average existing UK tax rates for each alcohol type (column (1)).[16] The optimal single ethanol tax rate is equal to the average marginal externality plus an adjustment for the covariance of marginal externalities and the absolute slope of overall ethanol demands. We compute that the optimal single rate is 35.9p per unit of ethanol. At this rate the average marginal externality, $\bar{\phi}'$, is 43.9p per unit. The fact that the optimal tax rate is lower than the average marginal externality 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 change in the tax rate (see

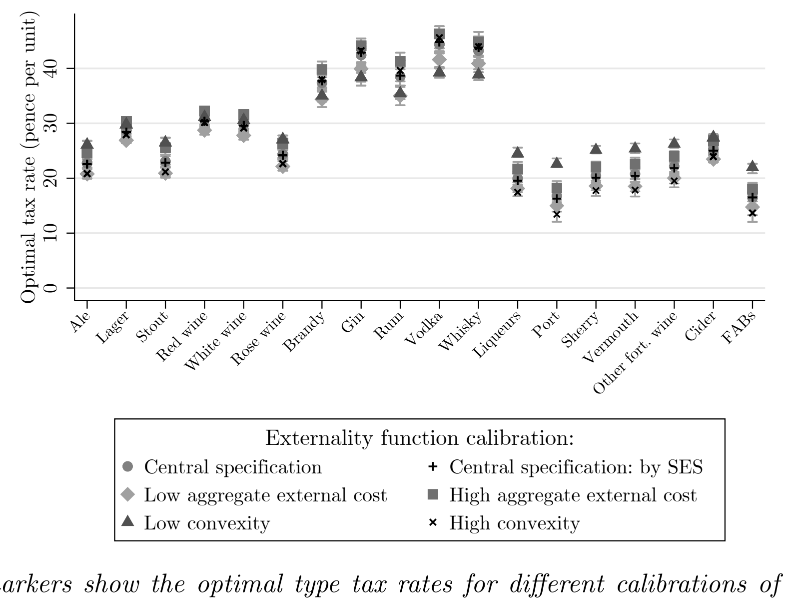
[16]Numbers include both alcohol excise taxes and VAT.

Figure 3.2). This lowers the effectiveness of the tax relative to there being no correlation between derived ethanol demands and the marginal externality, which leads to a lower optimal rate.

In column (3) we show optimal tax rates when rates are allowed to differ 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 Appendix B.5). 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 (4) we show the optimal tax rates that differ across alcohol types. There is considerable variation of the tax rates within segment. For instance, within the spirits segment the tax rate ranges from below 20p per unit of ethanol for port to over 40p per unit for vodka and tax rates are higher on the set of strong spirits than on liqueurs or fortified wines. Other segments also show a significant degree of variation in optimal tax rates. This suggests going from the optimal system of segment taxes to optimal type taxes is likely to result in a considerable improvement in welfare. Under the UK system, high strength spirits and cider are under taxed, relative to the optimal type system, while beer, wine and FABs are all over taxed.

Figure 4.1: *Comparison of optimal tax rates under different externality function calibrations*



Notes: The markers show the optimal type tax rates for different calibrations of the externality function. The parameters used to calibrate the different externality functions are presented in Table 3.4. 95% confidence intervals are drawn around each marker.

Figure 4.1 shows the optimal type tax rates under different assumptions about the shape of the externality function. We consider five alternatives: (i) a function that varies by the SES of households, (ii) a function that implies a lower aggregate external cost, (iii) a function that implies a higher aggregate external cost, (iv) a less convex function, and (iv) a more convex function.

Allowing for heterogeneity in the externality function across SES groups has little impact on optimal tax rates. This is because there are not large differences in the composition of SES across the household groups (defined bases on pre sample ethanol demand). We could potentially allow for the within household group preference heterogeneity to be correlated with SES and this might lead to different optimal tax rates (relative to the central calibration; we leave this for future work).

Varying the level of aggregate external cost to 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.

Effective average ethanol tax rates across households

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 that the marginal externality is increasing in ethanol content, the optimal tax regime prescribes relatively high taxes on households that purchase relatively large amounts of ethanol.

In practice it is difficult for governments to set consumer specific tax rates. However, by allowing variation in common tax rates across products, the planner can improve on single ethanol rate taxation, getting close to the first best. To illustrate this we compute the "effective average tax rate" (EATR) faced by households under the various tax policies. These describe the tax rate per unit of ethanol a household pays, given their demands under the various tax policies.

Table 4.2 shows these for the UK tax system, for the optimal single rate, segment and type taxes, and for the consumer specific taxes. In each case it shows the average

EATR for each household group. In the case of the single ethanol tax rate the EATR is the same across all households and is equal to the optimal rate of 35.9p. For the consumer specific taxes, each household's EATR is simply the consumer specific tax rate they face. The average consumer specific rate for the heaviest drinking group is 2.3 times the average rate for the lightest drinking group. Under the optimal segment tax rates the average EATR for the heaviest group is 1.1 times that of the lightest group, and under the optimal type tax rates the average EATR for the heaviest group is 1.15 times that of the lightest group. The differences in EATRs across household groups under the taxes rates that differ across products are considerably less than under the first best consumer specific taxes, but are closer to the first best than under either a single ethanol tax rate or the current UK system.

Table 4.2: *Effective average ethanol tax rates under different tax policies*

	Effective average ethanol tax rate				
			Product level tax rates		Consumer
Household group:	UK taxes	Single ethanol rate	Segment	Type	specific taxes
Less than 7 units	27.4	35.9	31.4	27.0	17.1
		[34.8, 36.6]	[30.6, 32.2]	[25.9, 28.2]	[16.2, 18.0]
7-14 units	27.2	35.9	31.8	28.9	24.0
		[34.8, 36.6]	[31.1, 32.5]	[28.2, 29.7]	[23.4, 25.2]
14-21 units	27.0	35.9	32.3	29.6	29.0
		[34.8, 36.6]	[31.6, 33.0]	[28.9, 30.3]	[28.2, 30.4]
21-35 units	27.2	35.9	32.7	29.9	33.4
		[34.8, 36.6]	[32.0, 33.3]	[29.2, 30.6]	[32.5, 36.4]
More than 35 units	27.2	35.9	34.1	31.1	39.2
		[34.8, 36.6]	[33.4, 34.6]	[30.5, 31.8]	[37.2, 45.0]
All households	27.2	35.9	32.3	29.0	22.9
		[34.8, 36.6]	[31.6, 32.9]	[28.4, 29.7]	[22.6, 23.8]

Notes: For each policy we calculate the effective average ethanol tax rate (EATR) by taking the average tax rate across products, using ethanol share weights for each group of households. For the single ethanol tax rate the EATR is equal to the optimal rate and is invariant across households. For the consumer specific taxes a household's EATR equals the single tax rate it faces.

The average EATR across all households falls as the tax regime available to the planner becomes increasingly flexible (i.e. moving from a single rate to consumer specific taxes). The reason for this is that when the planner only has one instrument (i.e. a single tax rate) it has to set a comparatively high rate in order to induce lower ethanol consumption among the very heavy drinkers. This is worth doing, even at the expense of the loss in consumer surplus for the light drinkers, because these small number of households generate considerable externality costs. With more flexible tax policy instruments the planner is able to design tax policies that are increasingly better targeted at the high marginal externality households. This allows for a reduction in the average tax rate across all products. We discuss this in more detail in the context of the impact of the policies on welfare in Section 4.2.

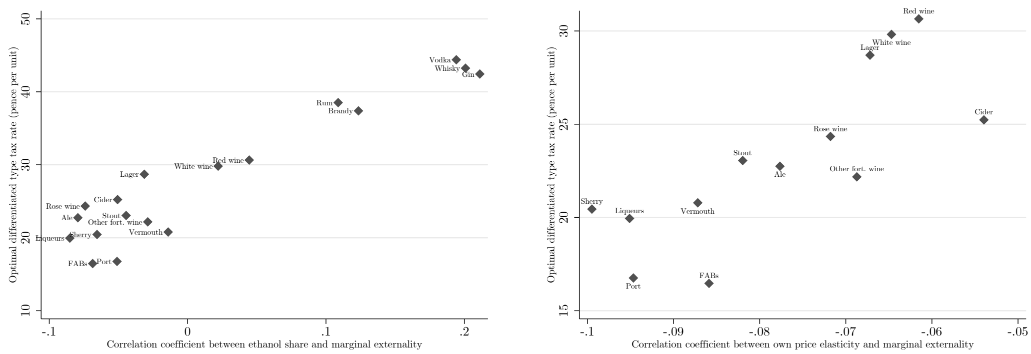
Understanding the variation in the optimal tax rates

In Section 2.4 we make some simplifying assumptions that allow us to gain intuition over the mechanisms leading to variation in optimal tax rates across products. There are three driving factors: (i) correlation of the marginal externality and demand shares, (ii) correlation of the marginal externality and own slopes of demand, and (iii) correlation of the marginal externality and cross price effects. Here, we investigate these patterns empirically. We focus on the optimal tax rates that differ across types.

For each alcohol type we compute the correlation coefficient across households in the share of ethanol from that type and the marginal externality. Panel (a) of Figure 4.2 plots the relationship between these correlation coefficients and the optimal type tax rates. The set of high strength spirits[17] are disproportionately purchased by heavy drinkers and have high optimal rates. For these products the high propensity of heavier drinkers to purchase them is the main driver of the high tax rates. The fact that heavy drinkers drink relatively more spirits is not unique to the UK – we show in Appendix B.5 that this pattern holds in the US too.

Figure 4.2: *Optimal tax rate and correlation between marginal externality and demands*

(a) Ethanol shares (b) Own price elasticities



Notes: In the left hand panel, the x-axis shows the correlation coefficient between the share of ethanol from each type and the marginal externality for each household. In the right hand panel, the x-axis shows the correlation coefficient between the own price elasticity for each type and the marginal externality for each household. In both graphs the y-axis shows the optimal type tax rate.

However, for the other products the relationship between the correlation in demand shares and marginal externalities with optimal taxes is less strong. Much

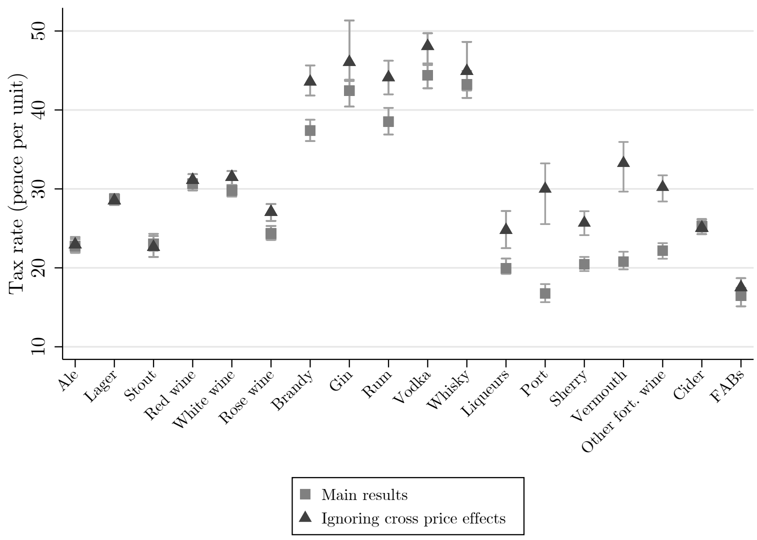
[17] Brandy, Gin, Rum, Vodka and Whisky

of this remaining variation is accounted for by the correlation in alcohol type own price elasticities of demand and marginal externalities. Panel (b) shows that these correlations vary with optimal taxes. Alcohol types with a positive correlation between own price elasticities and marginal externalities tend to attract higher rates. The stronger this correlation, the more effective the tax is at reducing the high externality generating consumption.

In Section 2.4 we show that a high covariance between the marginal externality and the cross price effects acts to lower optimal tax rates. The intuition behind this is that, all else equal, if households that generate high levels of marginal externalities are more willing to switch between alcohol products than to switch out of the market in response to price increases, the less effective the taxes are at lowering the external costs of consumption, while not creating larger losses in consumer surplus.

To quantify the importance of this channel we recompute 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 an application of the single rate formula (equation 2.7) to each alcohol type. 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 4.3: *Comparison of optimal tax rates with and without cross price effects*



Notes: The square markers shows the optimal taxes when the planner takes cross price effects into account; the triangular markers show the tax rates when the planner ignores cross price effects. 95% confidence intervals are drawn around each marker.

In Figure 4.3 we show the optimal rates when the planner ignores cross price effects along with the optimal rates when the planner takes account of cross price effects. Ignoring cross price effects leads to higher than optimal rates for all but three types – on average, the tax rates "overshoot" by around 7%. The reason for this is that heavy drinkers tend to be more willing to switch between alternative alcohol products than light drinkers, meaning when the planner ignores this effect taxes appear more effective at lowering the ethanol demand of high drinkers, leading it to set higher tax rates.

4.2 Welfare

For each of the different tax systems, we calculate the impact on consumer surplus, tax revenue and total external costs of moving from the UK system to that system. We present the results in Table 4.3. 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 billion per year lower than under the UK system, while tax revenue is £0.31 billion higher. However these welfare gains are partially off-set by the fact that consumer surplus falls by £1.85 billion. Overall, moving from the UK tax system to the optimal single rate ethanol tax leads to an increase of £0.46 billion in the sum of consumer surplus and tax revenue net of externalities.

The gain in consumer surplus plus tax revenue net of externalities of moving from the UK tax system to the segment taxes is £0.83 billion – 1.8 times larger than the move from the UK system to the optimal single ethanol rate. 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 these taxes raising marginally less revenue than the UK system).

Moving from the UK system to the optimal type taxes would result in a total gain of £1.05 billion – 2.3 times the gain of moving to the optimal single rate. Relative to the optimal segment taxes, the type taxes result in a slightly higher external cost and raise less revenue. However, this is more than made up for by the fact that this policy results in a much smaller reduction in consumer surplus (compared with the UK tax system). Relative to the more restricted single rate and segment taxes, the optimal type taxes are better able to target the consumption of households with high marginal consumption externalities, while limiting the reduction in consumer surplus.

Table 4.3 also shows the gain that would be achieved from a move from the UK system to the Pigovian first best of optimal consumer specific taxes. The improvement in the sum of consumer surplus and tax revenue net of externalities 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 type tax rates would therefore close nearly half of the gap in welfare between the UK system and the first best.

The welfare changes for the alternative externality function calibrations are shown in Appendix C.4. 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. The less convex the function, the closer the optimal type tax rates gets to the first best. Although there are differences in the precise magnitudes, the key results hold: the UK system performs worst, an optimally set single ethanol tax rate improves on this somewhat, but the optimal type tax rates do better, roughly halving the difference in welfare between the UK taxes and the first best.

Table 4.3: *Welfare impact of tax changes*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(4) $(2) + (3) - (1)$
UK taxes	7.25	7.16	–	–
Optimal:				
Single ethanol rate	-2.00	0.31	-1.85	0.46
	[-2.33, -1.67]	[0.14, 0.48]	[-2.03, -1.65]	[0.35, 0.56]
% difference	*-27.6*	*4.3*	–	–
Segment tax rates	-2.24	-0.13	-1.28	0.83
	[-2.54, -1.89]	[-0.25, 0.03]	[-1.42, -1.10]	[0.69, 0.98]
% difference	*-30.9*	*-1.8*	–	–
Type tax rates	-2.15	-0.47	-0.63	1.05
	[-2.44, -1.79]	[-0.58, -0.34]	[-0.77, -0.47]	[0.89, 1.20]
% difference	*-29.7*	*-6.6*	–	–
First best:				
Consumer specific taxes	-1.38	0.57	0.19	2.14
	[-1.74, -0.86]	[0.27, 0.85]	[0.00, 0.31]	[1.70, 2.39]
% difference	*-19.0*	*8.0*	–	–

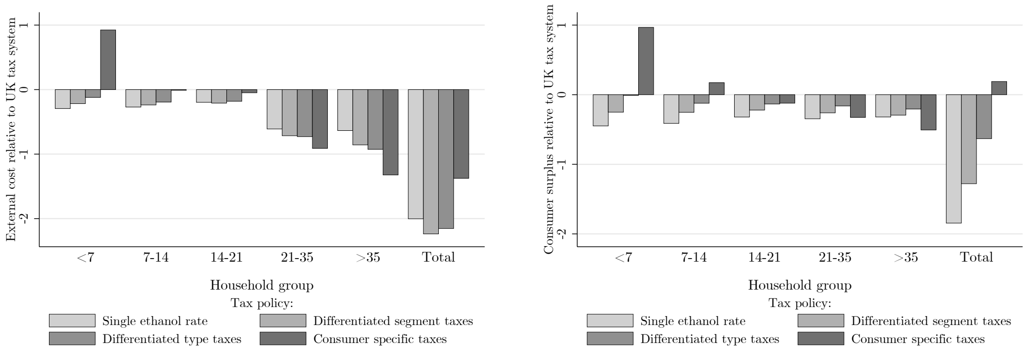
Notes: The first row shows the external cost and tax revenue under the UK tax system for our central calibration of the externality function (Table 3.4). The rows below show the difference relative to the UK system for each tax policy. Column (1) shows the external cost, column (2) the tax revenue, column (3) the change in consumer surplus relative to the UK system, and column (4) the change in the sum of consumer surplus and tax revenue minus the external cost. All numbers are expressed in £billion per year. Numbers in italic are the percentage differences relative to the UK system. 95% confidence intervals for the differences relative to the UK tax system are shown in square brackets.

Under the first best consumer specific taxes, both the external costs of alcohol consumption and consumer surplus are higher than under the optimal single rate or either of the product level tax rate systems. In Figure 4.4 we show how the changes in external costs (panel (a)) and consumers surplus (panel (b)), relative to the UK tax system, vary across the five household groups.

Relative to the UK system, the consumer specific taxes result in an increase in external costs and consumer surplus of the lightest drinking group and substantial reductions for heavier drinkers. The current UK tax system effectively over-taxes light drinkers. Consumer specific taxes allow the planner to reduce taxes on this group, while increasing taxes on heavier drinking groups (thereby both lowering the externality from these groups, as well as their consumer surplus). The single ethanol tax rate does not provide the planner with scope to differentially target light and heavy drinkers through the tax system – relative to the UK system it induces lower external costs and consumer surplus across the total ethanol distribution. Setting tax rates that differ across products provide some scope for better targeting the consumption of heavy drinkers. The optimal type tax rates result in much larger falls in external costs among heavy drinkers, leaving light drinkers only slightly worse off relative to the UK system.

Figure 4.4: *External costs and consumer surplus across households*

(a) External costs (b) Consumer surplus



Notes: The left hand figure shows the change in external costs generated by each group of households under each policy, relative to the UK tax system. The right hand figure shows the change in consumer surplus experienced by each group of households under each policy, relative to the UK tax system. Numbers are in £ billion per year.

4.3 Robustness

Supply responses

The tax solutions we compute assume full pass-through of tax to consumer prices. If we knew pass-through rates of taxes to all product prices in the alcohol industry, then we could adjust tax rates to ensure post tax equilibrium prices are the optimal ones we compute. For instance, suppose pass-through of taxes in the alcohol market was 75%. To get to the optimal post tax price we would simply need to scale our computed tax rates by the factor 4/3. If pass-through is differential across products, the scaling factor required would also vary across products. Our results therefore can complement those that study tax pass-through in alcohol markets.

In principle we could extend our empirical framework to formally model tax pass-through. While it may be possible to devise some plausible nesting structure, which would make estimation of the 100s of brand level elasticities and costs feasible, implementation would be challenging due to the scale of the problem. In this paper our focus is on tax setting in the alcohol market as a whole, so we have decided to defer formal incorporation of the supply side into our framework to future research.

However, we can use our demand estimates to provide an indication of whether there is likely to be differential tax pass-through across products in the market. A number of papers (including Anderson et al. (2001) and Weyl and Fabinger (2013)) have highlighted that the curvature of product demand is a key component of pass-through. In a monopoly setting with constant marginal cost, if demand is log-convex/log-concave tax will be over/under-shifted to consumer prices. In more complex settings with price competition and many sellers the determinants of pass-through become more complex, but the curvature of product demands remains a key determinant (Weyl and Fabinger (2013)). We find no relationship between the second derivatives of the log of product demands and the optimal taxes we compute (see Appendix C.5). This suggests there is nothing about the shape of consumer demands that points towards differential tax pass-through being systematically linked to optimal tax rates.

Demand dynamics

We allow for individual level heterogeneity in preferences and therefore statistical dependence in households' purchases, through time, through the random coefficients $(\Psi_i, \boldsymbol{\eta}_i)$. However, we do not model state dependence arising, conditional on preference heterogeneity, from the effect of past purchases on current behaviour. Current choice may depend on past choices due to high frequency habit 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 can show some reduced form evidence that suggests that these forms of dynamics are not likely to be of first order importance 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 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 Appendix C.6).

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. 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 Appendix C.6).

5 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) many products potentially create consumption externalities. We consider the alcohol market, in which the externality arises from ethanol, but where this is bundled together in products with other characteristics over which the consumer has preferences. 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.

We show that setting different tax rates across products can improve on a single tax rate applied to all ethanol. The reason is that varying rates across the ethanol in different products 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 consumption that generates high externalities. We show that while moving from the UK tax system to an optimal single ethanol tax rate would close 21% of the welfare gap between the UK system and the first best, moving to an optimal system that varies 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 firms, however focusing on a narrower segment of the market could provide a means for doing this.

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, there 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.

References

Akerlof, G. A. (1978). The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning. *American Economic Review 68*(1), 8–19.

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

Atkinson, A. and J. Stiglitz (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics 6*(1-2), 55–75.

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., 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.

Berry, S. T. and P. A. Haile (2010). Nonparametric Identification of Multinomal Choice Demand Models with Heterogeneous Consumers. *Cowles Foundation Discussion Paper* (1718), 1–41.

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

Bovenberg, A. L. and L. H. Goulder (1996). Optimal Environmental Taxation in the Presence of Other Taxes: General- Equilibrium Analyses. *American Economic Review 86*(4), 985–1000.

Bovenberg, A. L. and L. H. Goulder (2002). Environmental taxation and regulation. *Handbook of Public Economics 3*, 1471–1545.

Centers for Disease Control and Prevention (2016). CDC - Data and Maps - Alcohol. http://www.cdc.gov/alcohol/data-stats.htm.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and Taxation: Theory and Evidence. *American Economic Review 99*(4), 1145–1177.

Cnossen, S. (2007). Alcohol taxation and regulation in the European Union. *International Tax and Public Finance 14*(6), 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*.

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

Cremer, H., P. Pestieau, and J.-C. Rochet (2001). Direct Versus Indirect Taxation: The Design of the Tax Structure Revisited. *International Economic Review 42*(3), 781–800.

Diamond, P. A. (1973). Consumption externalities and corrective imperfect pricing. *The Bell Journal of Economics 4*(2), 526–538.

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

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

Gillitzer, C., H. J. Kleven, and J. Slemrod (2015). A Characteristics Approach to Optimal Taxation: Line Drawing and Tax-Driven Product Innovation. *Scandinavian Journal of Economics*.

Goddard, E. (2007). *Estimating Alcohol Consumption From Survey Data: Updated Method of Converting Volumes to Units*. National Statistics Methodological Series.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014). Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica 82*(1), 41–88.

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

Griffith, R., A. Leicester, and M. O'Connell (2013). Price-based measures to reduce alcohol consumption. IFS Briefing Note 138.

Griffith, R. and M. O'Connell (2009). The use of scanner data for research into nutrition. *Fiscal Studies 30*(3-4), 339–365.

Grittner, U., S. Kuntsche, K. Graham, and K. Bloomfield (2012, September). Social Inequalities and Gender Differences in the Experience of Alcohol-Related Problems. *Alcohol and Alcoholism 47*(5), 597–605.

Hansen, B. (2015). Punishment and Deterrence: Evidence from Drunk Driving. *American Economic Review 105*(4), 1581–1617.

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 Environmental Taxation. *Journal of Public Economics 52*(1), 101–119.

Jackson, C. K. and E. G. Owens (2011). One for the road: Public transportation, alcohol consumption, and intoxicated driving. *Journal of Public Economics 95*, 106–121.

Kaplow, L. (1996). The Optimal Supply of Public Goods and the Distortionary Cost of Taxation. *National Tax Journal 49*(4), 513–533.

Kaplow, L. (2006). On the undesirability of commodity taxation even when income taxation is not optimal. *Journal of Public Economics 90*(6–7), 1235–1250.

Kaplow, L. (2012). Optimal Control of Externalities in the Presence of Externalities. *International Economic Review 53*(2), 487–509.

Lancaster, K. (1971). *Consumer Demand: A New Approach*. Columbia University Press New York.

Laroque, G. R. (2005). Indirect taxation is superfluous under separability and taste homogeneity: A simple proof. *Economics Letters 87*(1), 141–144.

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). Using scanner technology to collect expenditure data. *Fiscal Studies 30*, 309–37.

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).

Luca, D. L., E. Owens, and G. Sharma (2015). Can Alcohol Prohibition Reduce Violence Against Women? *American Economic Review 105*(5), 625–629.

Mankiw, N. G. and M. Weinzierl (2010). The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution. *American Economic Journal: Economic Policy 2*(1), 155–176.

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. Technical report.

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

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.

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

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.

Ruhm, C. J. (1996). Alcohol policies and highway vehicle fatalities. *Journal of Health Economics 15*(4), 435–454.

Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics 83*(2), 217–230.

Sandmo, A. (1975). Optimal Taxation in the Presence of Externalities. *Swedish Journal of Economics 77*(1), 86–98.

Seim, B. 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), 831–862.

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.

Train, K. E. (2003). *Discrete Choice Methods with Simulation.* Cambridge University Press.

UK Cabinet Office (2003). Alcohol misuse: How much does it cost? Technical report.

US Department of Justice (2005). *Drinking in America: Myths, Realities, and Prevention Policy.* Report by Office of Juvenile Justice and Delinquency Prevention Published: www.udetc.org/documents/Drinking_in_America.pdf, accessed 13 April 2014.

Weinzierl, M. (2011). The Surprising Power of Age-Dependent Taxes. *Review of Economic Studies 78*(4), 1490–1518.

Weyl, E. G. and M. Fabinger (2013). Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition. *Journal of Political Economy 121*(3), 528–583.

World Health Organization (2014). Global Status Report on Alcohol and Health. Technical report, Geneva: World Health Organization.

Appendix

Design of optimal corrective taxes in the alcohol market

Rachel Griffith, Martin O'Connell and Kate Smith

January 30, 2017

A Additional theoretical results

A.1 Optimal tax formulae with cross-price effects

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

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

$$\Rightarrow \tau_1 \sum_i Z'_{i1j} - \sum_i \phi'_i Z'_{i1j} + \tau_2 \sum_i Z'_{i2j} - \sum_i \phi'_i Z'_{i2j} = 0$$

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

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

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

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

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

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

where can re-write the condition as:

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

and express the difference in optimal taxes as:

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

B Additional details on the data

B.1 Data sets

Here we describe three of the data sets we use in addition to the Kantar Worldpanel, which is described in Section 3.1.

National Health and Examination Survey (NHANES)

The National Health and Examination Survey (NHANES) combines interviews and physical examinations to assess the health and nutritional status of adults and children in the United States. We use data on 15,699 individuals over the age of 21 from the 2007 – 2011 surveys. We use two components from the survey.

The first is the diary component. Individuals record all foods and beverages consumed during the 24-hour period of the interview (midnight to midnight). Individuals are interviewed twice: the first dietary recall interview is collected in-person, and the second interview is collected by telephone 3 to 10 days later. To construct the variable measured on the x-axis of Figures 3(b) and 3(d) we average all ethanol consumed over the two separate diary days, and convert to units (1 unit = 8g ethanol).

The second component we use is the alcohol questionnaire, which focuses on lifetime and current use (past 12 months). We use the answers to two questions to draw Figures 3.3(b) and 3.3(d). Figure 3.3(b) uses questions *ALQ120Q* ("How often did you drink alcohol over the past 12 months?") and *ALQ120U* (unit of measure for question *ALQ120Q*) to construct the average per week drinking frequency. Figure 3.3(d) uses questions *ALQ141Q* ("On how many days over the past

12 months did you consume 4 or 5 alcoholic beverages?") and *ALQ141U* (unit of measure for question ALQ141Q) to construct the average number of days per week on which the individual engaged in binge drinking. Figures 3.3(b) and 3.3(d) fits local polynomial regressions between these variables and the ethanol consumption variable constructed from the diary data for the subset of individuals who record consuming non-zero quantities of ethanol in the diary (3234 individuals).

For Figure 3.3(b) we use only the diary component – the alcohol questionnaire does not have any details on the types of alcohol consumed. We use the ethanol measure described above to group individuals into quartile (quartile values: $[7.0, 14.3, 27.9]$). For each group of individuals, we compute the share of their ethanol that comes from beer, wine, spirits, cider/FABs. The figure shows the average ethanol share for each alcohol type and group of households relative to the bottom quartile.

Health Survey for England (HSE)

The Health Survey for England combines interviews and physical examinations to assess the health status of adults and children in the United Kingdom. We use data on 8281 individuals over the age of 18 in the 2011 survey. We use the alcohol questionnaire component of the survey. We use the derived variable *totalwu* – total units of alcohol per week, which is derived from questions about the individuals drinking habits to construct the units of ethanol per week variable used on the x-axis of Figures 3.3(a) and 3.3(c).

Figure 3.3(a) uses the responses to question *d7many* ("How many days in the last 7 have you had a drink?") in a local polynomial regression to estimate the relationship between ethanol consumption and frequency of drinking. We use the responses to question *d7unitwg* ("Number of units drunk on the heaviest day in the past 7 days") to construct a variable indicating the propensity to binge drink. This is equal to 1 if the individual was male (female) and recorded drinking over 8 (6) units on their heaviest drinking day out of the past 7. Figure 3(c) uses a local polynomial regression to estimate the relationship between ethanol consumption and propensity to binge drink.

Living Costs and Food Survey (LCFS)

The Living Costs and Food Survey (LCFS) is a two week diary survey used to measure the spending patterns of UK households. We use data on 3688 households in the 2011 survey that record buying alcohol over the two week survey period. The LCFS records the quantity (in litres) purchased of different alcohol types (both on-and off-trade), but not the ethanol purchased. We use the Kantar Worldpanel to impute the average ABV content for different alcohol types to construct the average weekly ethanol purchased per adult for each household.

B.2 On-trade alcohol

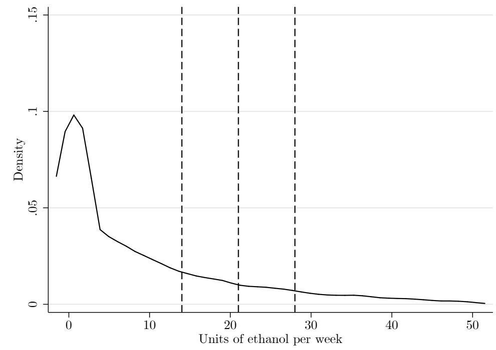
Our data contain comprehensive information on purchases of alcohol off-trade. Off-trade alcohol purchases consist of purchases of alcohol products made in stores (including supermarkets, corner stores and liquor stores). Off-trade alcohol purchases constitute 77% of ethanol purchased in alcoholic drinks in the UK (Living Costs and Food Survey (2011). 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 contain very detailed information on purchases of alcohol products off-trade, but they do not contain information on alcohol purchases on-trade. The Living Costs 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.

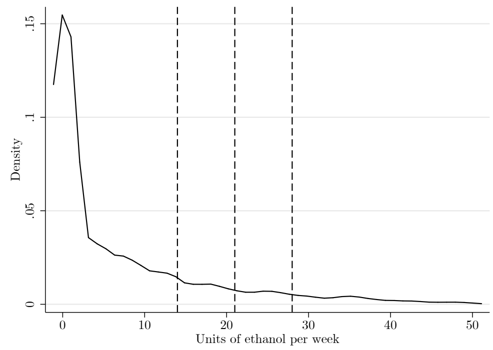
To do this we impute the strength of the alcohol categories collected in the LCFS. For instance, for the category beer we use 4% ABV – the average from the Kantar data. Based on this, in 2011, we compute that 77% of units of ethanol purchased was done so off-trade. Figure B.1 plots the distributions of ethanol purchases for on- and off-trade. Figure B.2 shows the how the share of ethanol from different alcohol segments varies across the total ethanol purchase distribution. 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 B.1: *Distribution of ethanol purchases: on- and off-trade*

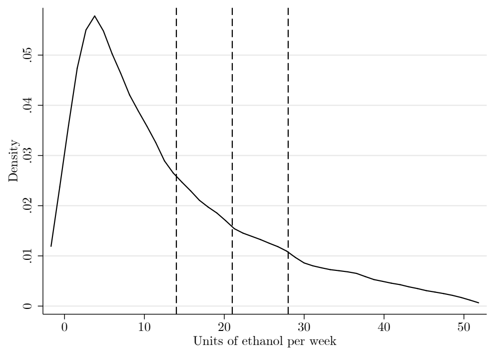
(a) On- and off-trade



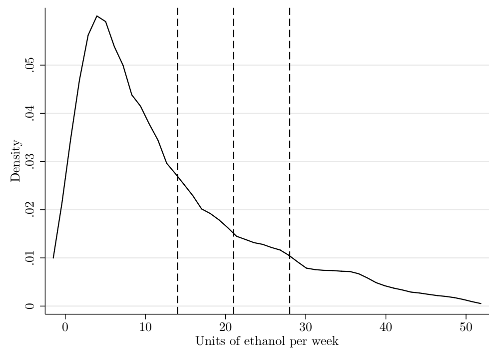
(b) Off-trade only



(c) On- and off-trade: conditional on buying



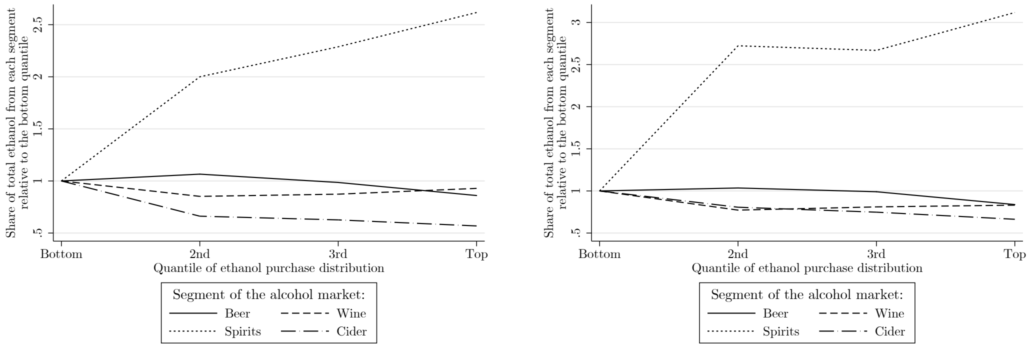
(d) Off-trade only: conditional on buying



Notes: Both panels use data from the Living Costs and Food survey 2011; the left hand panel shows alcohol purchases made on- and off-trade, the right hand panel shows alcohol purchased made off-trade only. We calculate the total amount of ethanol purchased per adult in each household over the two week period; the figures show the distribution of this variable. The top two figures show the unconditional distributions across all households; the bottom two figures show the distributions conditional on purchasing alcohol (at all (left), or off-trade only (right)) during the two week survey period.

Figure B.2: *Alcohol purchases: on- and off-trade*

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



Notes: Both panels use data from the Living Costs and Food survey 2011; the left hand panel shows alcohol purchases made on- and off-trade, the right hand panel shows alcohol purchased made off-trade only. We calculate the total amount of ethanol purchased per adult in each household over the two week period, and divide households into quartiles based in this measure. The y-axis shows the proportion of ethanol that comes from beer, wine, spirits, cider/FABS relative to the bottom quantile.

B.3 Additional data description

Tables and figures in this section are computed using the Kantar Worldpanel.

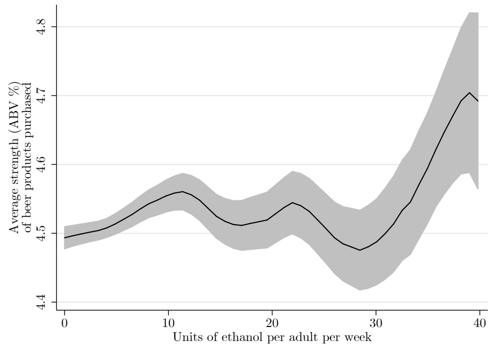
Table B.1: *Variation in share of ethanol from different alcohol types by group*

	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

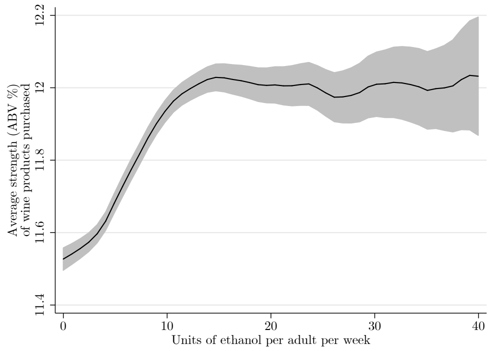
Notes: For each group of households we calculate the total number of units of ethanol purchased in 2011 and the number purchased from each type of alcohol. The numbers in the table are the percentage of total units from each type.

Figure B.3: *Variation in strength of alcohol purchased across the total ethanol distribution*

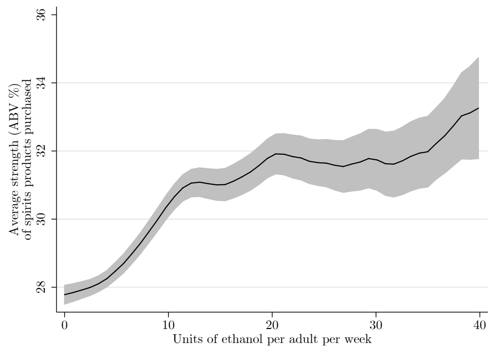
(a) Beer segment



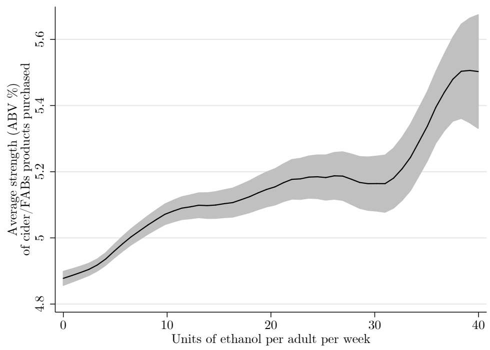
(b) Wine segment



(c) Spirits segment



(d) Cider/FABs segment



Notes: The x-axis shows the average ethanol purchased per adult over a calendar year for each household in the sample. Each panel shows the average strength (measured in ABV (%)) of products belonging to the segment indicated in the caption bought by each household, using ethanol shares as weights. A local polynomial is fitted in each panel.

B.4 Alcohol products and sizes

Table B.2: *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.

B.5 Correlation between heterogeneity in product demands and high ethanol consumption

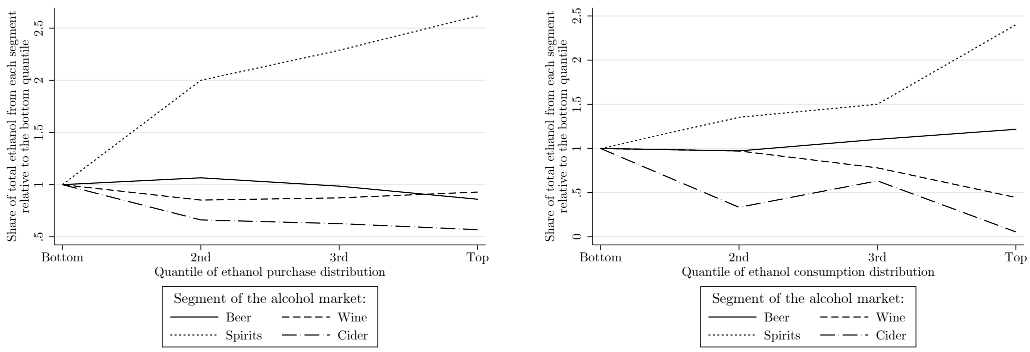
In Figure B.4 we show how the share of ethanol consumers demand from each of the four main segments of the alcohol market varies with their total ethanol demand.[18]. The x-axis splits consumers in the quartiles of the total ethanol distribution. The y-axis shows how the share of units from a particular segment of the market com-

[18]The beer segment includes lager and ales, the spirits segment includes fortified wines and the cider segment includes flavoured alcohol beverages (FABs)

pares with the share demanded by consumers in the bottom quartile of the total ethanol distribution. Panel (a) is for the UK and panel (b) is for the US. In both countries heavier drinkers obtain considerably more of their ethanol from spirits and considerably less from cider than lighter drinkers.[19]

Figure B.4: *Variation in alcohol segment across the total ethanol distribution*

(a) UK (b) US



Notes: Panel (a) uses data from the Living Costs and Food survey and Panel (b) uses data from the food and drink diary component of NHANES. For a detailed description of the data and variable construction, see Appendix B.1.

C Additional details on the empirical analysis

C.1 Preference heterogeneity distribution

We use $\Psi_i = (\alpha_i, \beta_i, \boldsymbol{\gamma}_i)$ to denote the household specific preferences over observable product attributes and $\boldsymbol{\eta}_i = (\eta_{i1}, ..., \eta_{iJ})'$ to 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

To compare 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 $\Psi_i \sim \mathcal{N}(\mu, \Omega)$,

[19] One reason for this larger divergence in shares across segments in the US compared to the UK is that the NHANES data for the US is based on a two day diary while the LCFS data for the UK is based on a two week diary. See Appendix B.1 for further details.

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 $\boldsymbol{\eta}_i \sim \mathcal{N}(\bar{\eta}, \Sigma)$. There are many products so 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. We do this in a tractable and easily implementable way by exploiting pre sample 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.

Specifically, we divide households into $d = 1, \dots, D$ 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 of the paper. 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 components 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.

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. 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.

In Table C.1 we show the optimal tax rates computed with the normal preference distribution (the analogue of Table 4.1 in the paper). In Figure C.1 we show a comparison of the optimal type taxes when modelling the preference 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. The darker markers are for the mixed normal preference distribution specification – these repeat the rates shown in Table 4.1. The lighter markers show the optimal type tax rates computed using the normal preference distribution specification.

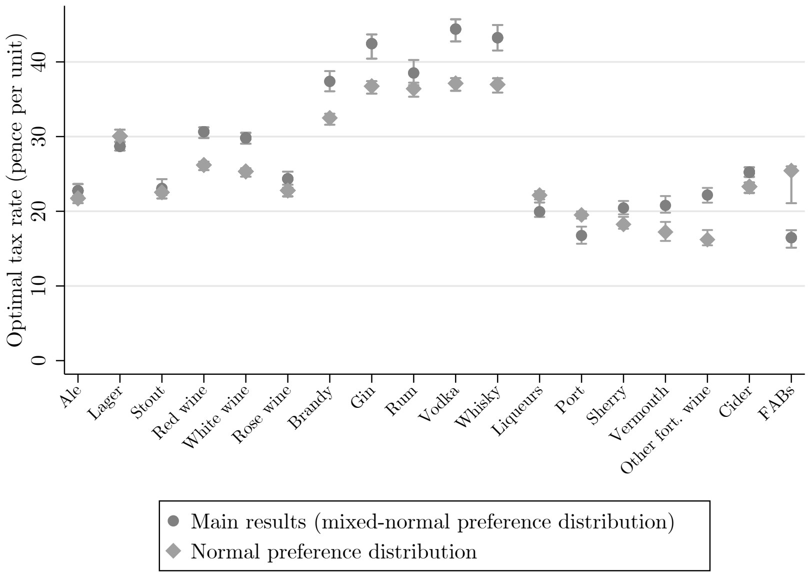
The pattern of optimal taxes is broadly similar under both specifications – in each case the set of strong spirits (brandy, gin, rum, vodka and whisky) attract relatively high rates. However, the exact tax rates differ – the mixed-normal specification, compared with normal specification, prescribes higher 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.

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

(1)		(2)		(3)	
		Product level tax rates			
Single rate		Segment		Type	
Ethanol	35.3 [34.4, 36.3]	Beer (inc. lager and ale)	30.1 [29.3, 31.2]	Ale	21.7 [21.1, 22.1]
	·		·	Lager	30.1 [29.3, 30.9]
	·		·	Stout	22.5 [21.7, 23.0]
	·	Wine	26.2 [25.6, 26.8]	Red wine	26.2 [25.5, 26.6]
	·		·	White wine	25.3 [24.6, 25.7]
	·		·	Rose wine	22.8 [22.0, 23.2]
	·	Spirits (inc. fortified wine)	38.0 [36.9, 39.2]	Brandy	32.5 [31.6, 33.1]
	·		·	Gin	36.7 [35.7, 37.4]
	·		·	Rum	36.4 [35.3, 37.2]
	·		·	Vodka	37.1 [36.1, 37.8]
	·		·	Whisky	37.0 [35.9, 37.8]
	·		·	Liqueurs	22.2 [21.6, 22.7]
	·		·	Port	19.5 [19.0, 20.0]
	·		·	Sherry	18.2 [17.7, 19.2]
	·		·	Vermouth	17.2 [16.0, 18.6]
	·		·	Other fort. wine	16.2 [15.4, 17.5]
	·	Cider (inc. FABs)	24.0 [23.4, 24.7]	Cider	23.3 [22.5, 23.8]
	·		·	FABs	25.4 [21.1, 26.0]

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 single tax rate. Columns (2) and (3) show the optimal product level tax rates, at the segment and type level, respectively. The dots represent the tax rate shown in the row above. 95% confidence intervals are shown below each tax rate.

Figure C.1: *Comparison of optimal tax rates under different preference heterogeneity specifications*



Notes: The darker markers shows the optimal taxes under the mixed normal distribution of preferences; the lighter markers show the optimal taxes under the normal distribution of preferences.

C.2 Elasticities

Table C.2: *Average own and cross price elasticities within and between alcohol segments, by household group*

<7 units	Mean own price	Mean cross price elasticity 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.

C.3 Additional optimal tax results

Table C.3: *Tax rate solutions*

(1)		(2)		(3)	
		Differentiated taxation			
Single rate		Segment		Type	
Ethanol	35.9 [34.8, 36.6]	Beer (inc. lager and ale)	28.3 [27.7, 29.0]	Ale	22.7 [21.9, 23.6]
	·		·	Lager	28.7 [28.1, 29.3]
	·		·	Stout	23.1 [22.2, 24.3]
	·	Wine	30.4 [29.6, 31.0]	Red wine	30.6 [29.8, 31.2]
	·		·	White wine	29.8 [29.0, 30.5]
	·		·	Rose wine	24.3 [23.6, 25.3]
	·	Spirits (inc. fortified wine)	42.6 [41.1, 43.9]	Brandy	37.4 [36.1, 38.8]
	·		·	Gin	42.4 [40.4, 43.7]
	·		·	Rum	38.5 [36.9, 40.3]
	·		·	Vodka	44.4 [42.7, 45.7]
	·		·	Whisky	43.2 [41.5, 44.9]
	·		·	Liqueurs	19.9 [19.2, 21.2]
	·		·	Port	16.8 [15.6, 17.9]
	·		·	Sherry	20.4 [19.6, 21.4]
	·		·	Vermouth	20.8 [19.8, 22.0]
	·		·	Other fort. wine	22.2 [21.2, 23.1]
	·	Cider (inc. FABs)	25.4 [24.8, 26.1]	Cider	25.2 [24.6, 25.9]
	·		·	FABs	16.5 [15.1, 17.5]

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 single tax rate. Columns (2) and (3) show the optimal product level tax rates, at the segment and type level, respectively. The dots represent the tax rate shown in the row above. 95% confidence intervals are shown below each tax rate.

C.4 Welfare and parameterisation of externality function

Table C.4: *Welfare impact of tax changes: low aggregate external cost*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(4) (2) + (3) − (1)
UK taxes	6.00	7.16	–	–
Optimal:				
Single ethanol rate	-1.19	0.32	-1.34	0.17
	[-1.44, -0.94]	[0.22, 0.45]	[-1.54, -1.17]	[0.10, 0.24]
% difference	*-19.8*	*4.5*	–	–
Segment tax rates	-1.43	-0.13	-0.81	0.49
	[-1.67, -1.16]	[-0.22, -0.01]	[-0.96, -0.64]	[0.40, 0.58]
% difference	*-23.9*	*-1.9*	–	–
Type tax rates	-1.42	-0.56	-0.15	0.71
	[-1.64, -1.13]	[-0.63, -0.45]	[-0.29, 0.02]	[0.60, 0.82]
% difference	*-23.6*	*-7.8*	–	–
First best:				
Consumer specific taxes	-0.63	0.35	0.96	1.94
	[-0.95, -0.21]	[0.02, 0.63]	[0.80, 1.04]	[1.61, 2.18]
% difference	*-10.4*	*4.9*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for our low aggregate external cost calibration of the externality function (Table 3.4). 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. 95% confidence intervals for the differences relative to the UK tax system are shown in square brackets.

Table C.5: *Welfare impact of tax changes: high aggregate external cost*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(4) (2) + (3) − (1)
UK taxes	8.50	7.16	–	–
Optimal:				
Single ethanol rate	-2.83	0.29	-2.28	0.85
	[-3.22, -2.42]	[0.08, 0.51]	[-2.45, -2.06]	[0.67, 1.01]
% difference	*-33.3*	*4.0*	–	–
Segment tax rates	-3.01	-0.11	-1.66	1.25
	[-3.37, -2.59]	[-0.28, 0.08]	[-1.79, -1.47]	[1.05, 1.45]
% difference	*-35.4*	*-1.5*	–	–
Type tax rates	-2.87	-0.40	-1.03	1.45
	[-3.22, -2.44]	[-0.53, -0.24]	[-1.18, -0.86]	[1.23, 1.65]
% difference	*-33.8*	*-5.6*	–	–
First best:				
Consumer specific taxes	-2.11	0.81	-0.58	2.34
	[-2.56, -1.60]	[0.51, 1.00]	[-0.69, -0.38]	[1.93, 2.71]
% difference	*-24.8*	*11.3*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for our high aggregate external cost calibration of the externality function (Table 3.4). 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. 95% confidence intervals for the differences relative to the UK tax system are shown in square brackets.

Table C.6: *Welfare impact of tax changes: less convex function*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(4) (2) + (3) − (1)
UK taxes	7.25	7.16	–	–
Optimal:				
Single ethanol rate	-1.43	0.32	-1.43	0.32
	[-1.64, -1.24]	[0.20, 0.46]	[-1.59, -1.30]	[0.27, 0.36]
% difference	*-19.7*	*4.5*	–	–
Segment tax rates	-1.51	0.03	-1.12	0.42
	[-1.71, -1.31]	[-0.07, 0.16]	[-1.26, -0.99]	[0.37, 0.47]
% difference	*-20.8*	*0.5*	–	–
Type tax rates	-1.49	-0.19	-0.82	0.49
	[-1.70, -1.29]	[-0.28, -0.07]	[-0.95, -0.69]	[0.43, 0.55]
% difference	*-20.6*	*-2.6*	–	–
First best:				
Consumer specific taxes	-0.81	0.85	-0.83	0.84
	[-1.09, -0.56]	[0.59, 1.11]	[-1.06, -0.65]	[0.64, 1.01]
% difference	*-11.2*	*11.9*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for less convex calibration of the externality function (Table 3.4). 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. 95% confidence intervals for the differences relative to the UK tax system are shown in square brackets.

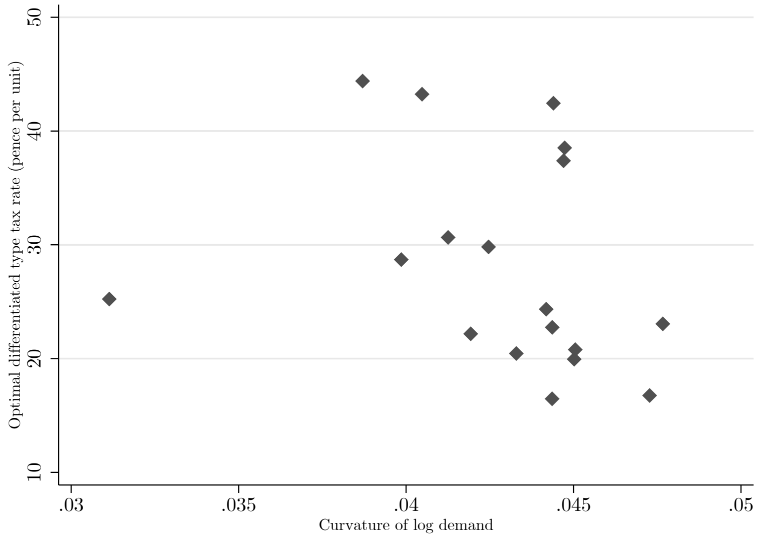
Table C.7: *Welfare impact of tax changes: more convex function*

£billion per year	(1) External cost	(2) Tax revenue	(3) Change in consumer surplus	(4) (2) + (3) − (1)
UK taxes	7.25	7.16	–	–
Optimal:				
Single ethanol rate	-2.28	0.30	-2.04	0.54
	[-2.67, -1.87]	[0.12, 0.49]	[-2.22, -1.79]	[0.39, 0.69]
% difference	*-31.5*	*4.2*	–	–
Segment tax rates	-2.58	-0.18	-1.27	1.13
	[-2.92, -2.17]	[-0.31, -0.01]	[-1.42, -1.08]	[0.92, 1.33]
% difference	*-35.6*	*-2.4*	–	–
Type tax rates	-2.48	-0.63	-0.41	1.44
	[-2.82, -2.05]	[-0.72, -0.49]	[-0.56, -0.24]	[1.21, 1.65]
% difference	*-34.2*	*-8.7*	–	–
First best:				
Consumer specific taxes	-1.69	0.40	0.78	2.86
	[-2.16, -1.10]	[0.03, 0.61]	[0.70, 0.95]	[2.39, 3.30]
% difference	*-23.3*	*5.6*	–	–

Notes: The first row shows the external cost and tax revenue under the 2011 UK tax system for our more convex calibration of the externality function (Table 3.4). 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. 95% confidence intervals for the differences relative to the UK tax system are shown in square brackets.

C.5 Demand curvature and optimal tax rates

Figure C.2: *Correlation between optimal type tax rates and the second derivative of log demand*



Notes: We compute the second derivative of log demand at UK prices, and take a demand share weighted average across options within each alcohol type. The figure plots this average against the optimal type tax rate.

C.6 Short-run persistence and stockpiling

Table C.8: *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 (4).

Table C.9: *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.