Tax design in the alcohol market

Rachel Griffith, Martin O'Connell and Kate Smith*

November 5, 2018

Abstract

Alcohol consumption generates negative externalities that are non-linear in the total amount of alcohol consumed. If tastes for products are heterogeneous and correlated with marginal externalities, then varying tax rates on different products can lead to welfare gains. We study this problem in an optimal tax framework and empirically for the UK market. We find that heavy drinkers have systematically different patterns of alcohol demands and welfare gains from optimally varying rates are higher the more concentrated externalities are amongst heavy drinkers.

Keywords: externality, corrective taxes, alcohol
JEL classification: D12, D62, H21, H23
Acknowledgements: The authors 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), RES-544-28-0001, and the Open Research Area, ES/N011562/1, 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

Alcohol consumption is associated with costs to society from anti-social behaviour, crime and public costs of policing and health care. These externalities are non-linear in alcohol consumption, with a small number of heavy drinkers creating the majority of the costs. Governments attempt to reduce problematic alcohol consumption through restricting availability (Seim and Waldfogel (2013) provide a recent analysis) and with policies that aim to increase prices. In this paper we study the design of alcohol taxes.

Our contribution in this paper is to apply the insights from Diamond (1973) to a differentiated product setting to show how the optimal design of alcohol taxes depends on the correlation of consumers' product level demands with the marginal externality their alcohol consumption creates, and to empirically quantify the possible gains from reforms that exploit these correlations. We consider a setting in which ethanol (or pure alcohol) creates an externality and in which the government's motivation for setting alcohol taxes is solely to correct for these externalities. If consumers' tastes for different alcohol products and their price responsiveness are correlated with the marginal externalities that their alcohol consumption creates, then it is optimal to levy different per ethanol tax rates across products.

We study the empirical importance of this using detailed micro data on the UK market for alcohol. We estimate consumer demand for alcohol products and solve for optimal tax rates, showing how the welfare gains from varying tax rates across different types of alcohol depend on the concentration of alcohol externalities among heavy drinkers. A social planner chooses tax rates to maximise the sum of consumer surplus net of external costs of consumption. When the planner is constrained to set a single tax rate applied to all ethanol, the optimal tax is a weighted average of the marginal externalities from alcohol consumption (Diamond (1973)). However, many governments vary tax rates across different types of alcohol; optimally differentiating rates allows the planner to improve on the Diamond prescription if consumers' demands for different alcohol products are correlated with the marginal externality associated with their ethanol consumption. The planner can target the most socially harmful drinking by taxing more heavily the ethanol in products that are both disproportionally consumed by problem drinkers and for which an increase in price leads to a relatively strong reduction in their total ethanol consumption. In lieu of consumer specific taxes, the optimal multi rate system uses correlation in preferences and marginal externalities to "tag" and target consumption that is likely to have high marginal external costs (Akerlof (1978)).

The optimal tax rates depend on the full set of own and cross price elasticities across products and on how they vary with the marginal externalities that consumers create. In our empirical analysis we estimate a model of demand and use the estimates to solve for the optimal rates. We use a discrete choice framework that embeds the decision over whether to buy alcohol, what product to buy and in what size and that captures preference heterogeneity that is correlated with total ethanol consumption (an important driver of the size of externalities). This allows us to capture patterns of substitution between differentiated alcohol products and how demands for different types of alcohol are correlated with the size of externalities an individual generates. A number of papers apply continuous choice demand methods to learn about the responsiveness of alcohol demand to changes in price, 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)). However, this misses potentially important patterns of substitution between, for instance, different beer products, or different types of spirits.

We estimate demand using longitudinal data on the alcohol purchases of a panel of British households. These data contain repeated observations per household, well measured prices and product information for disaggregate products. Consistently heavy drinkers (i.e. those with high total ethanol demands) systematically purchase a different mix of products than lighter drinkers; on average, they buy stronger and cheaper varieties of alcoholic beverages. We find they are much more willing to switch between different alcohol products in response to price changes, and are less willing to switch away from alcohol altogether than lighter drinkers. We also find that there is important substitution between different segments of the market (i.e. from beer to wine, or wine to spirits); neglecting this cross-segment switching would lead to mis-estimation of the optimal rates.

Alcohol markets are a natural setting in which to study optimal corrective taxes. The social costs of alcohol consumption are of concern across the developed world (World Health Organization (2014)).[1] There is considerable evidence that these externalities are non-linear 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 binge drinking accounts for roughly three quarters of the cost of excessive alcohol use (Centers for Disease Control and Prevention (2016)). Despite the evidence that external costs

[1] Negative consumption externalities associated with alcohol include: public healthcare costs, violent behaviour (e.g. Luca et al. (2015)), drink driving (e.g. Ruhm (1996), Jackson and Owens (2011), Hansen (2015)) and negative impacts on prenatally exposed children (Nilsson (2017)).

are convex in alcohol consumption (and hence, at the margin, heavy drinkers tend to create much larger externalities than lighter drinkers), there is uncertainty about the degree of this convexity. We calibrate the mapping from alcohol demands into external costs as a weakly convex function of households' total ethanol demand and we show how our empirical optimal tax results vary across different degrees of convexity of this relationship.

We consider two alternative tax systems. The first is a single tax rate levied on ethanol (the consumption of which maps directly into externalities). The second is a multi rate system in which the planner can vary the tax rate levied on ethanol across a set of alcohol types (based on market segment – spirits, wine, beer or cider – and alcohol strength). If externalities are linear in ethanol and the same across people, then a single ethanol tax rate can achieve the first best; there are no welfare gains from moving to a multi rate system. However, the more convex is the externality function, the larger are welfare gains from being able to set different tax rates across different forms of ethanol. This is because the higher the degree of convexity, the larger the share of externalities are generated by the heaviest drinkers. This enables the planner to target the multi rate tax system more specifically on lowering the ethanol intake of this narrow set of households. By levying a relatively high tax rate on strong spirits the planner is able to target a larger share of the alcohol purchases of heavy than light drinkers, and is able to encourage them to switch to less strong alcohol products, hence lowering their level of ethanol consumption. The size of welfare gains from this additional flexibility depend on how concentrated externalities are among the heaviest drinkers – if, for instance, the 18% of households that purchase the most ethanol account for 95% of the external costs of drinking, the welfare gain from optimally setting different rates is around £400 million.

In practice, alcohol tax systems differ markedly from the optimal prescription. In the US, alcohol is taxed per litre of volume rather than by ethanol content. This means that, within beer, spirits and wine, the effective tax rate per unit of ethanol is declining in the alcoholic strength of the product. In most European countries alcohol is, at least in part, also taxed per litre of volume. The UK system entails taxing wine and cider per litre of volume, while beer and spirits are taxed per unit of ethanol. Our results suggest that if the UK rationalized its current system by moving to an optimally set single ethanol tax there would be substantial welfare gains (of the order of £1.2 billion), with further gains from adopting optimally differentiated rates.

Taxation is clearly not the only instrument available to government to deal with the social costs of excess alcohol consumption. Laws that restrict underage drinking, prohibit drink driving, or provide help to alcoholics, are also well motivated policies. These complement a well-designed system of alcohol taxes. Most governments already tax alcohol over and above general sales tax applied to all products. One of the leading reasons for this is to help tackle social harms from drinking, and it is therefore informative to study the optimal design of such a system.

Our work complements a set of recent papers (Miravete et al. (2018, 2017), Conlon and Rao (2015)) that study how price regulation in two US states affects alcohol pricing and welfare. These papers are principally concerned with the role that government intervention plays in raising revenue and how this interacts with market power of firms. In contrast to this work, we focus on the role of alcohol taxes to corrective for externalities and we assume that pass-through of the tax to consumer prices is complete. Miravete et al. (2017) show that a government that has more flexible instruments (effectively product specific ad valorem taxes) can more effectively meet its goal of raising revenue compared with when it sets a single tax rate. We show a similar result holds with respect to correcting for alcohol related externalities. It has long been highlighted that in absence of externalities (and non-linear income taxes) the optimal set of commodity taxes set to raise a target amount of revenue implies rate differentiation across products (Ramsey (1927)). Our results suggest that in the absence of a revenue raising constraint, in a setting in which taxes are solely targeted at lowering the external costs of ethanol consumption, heterogeneity in marginal externalities provides an alternative rationale for rate differentiation.

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. In Section 4 we outline the empirical demand model and present our demand estimates. We use these along with our optimal tax framework to compute optimal tax rates, which we present, along with welfare results, in Section 5. A final section summarises and concludes. Additional details are provided in the Online Appendix.

2 Corrective tax design

2.1 Model set-up

Let $i \in \{1, ..., N\}$ index consumers; each consumer has income y_i. Let $j \in \{1, \ldots, J\}$ index alcohol products, available at post-tax prices $\mathbf{p} = (p_1, \ldots, p_J)'$. Each product contains a vector of characteristics, $\mathbf{x}_j$. One characteristic (i.e. an element of $\mathbf{x}_j$) is the amount of ethanol (pure alcohol) the product contains, denoted z_j. We denote the matrix of all product characteristics $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_J)'$ and the vector of ethanol contents, $\mathbf{z} = (z_1, \ldots, z_J)'$.

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

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

where α_i is the marginal utility of income and $v_i(\mathbf{p}, \mathbf{x})$ is the indirect utility that arises from the alcohol demands for consumer i. We denote the consumer's demand for product j by $q_{ij} = f_{ij}(\mathbf{p}, \mathbf{x})$ and the consumer's vector of demands by $\mathbf{q}_i = (q_{i1}, \ldots, q_{iJ})'$. Quasi-linear utility means that 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 assume that the external cost associated with an individual's alcohol consumption is given by $\phi_i(Z_i)$, where $Z_i = \sum_j z_j q_{ij}$ denotes individual i's total ethanol demand from all the products in the market. 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. The total external cost from all consumers in the market is then:

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

Consumers ignore the externality when making their choices, and the goal of the social planner is to use taxes to induce consumers to internalize the externality, while taking account of the reduction in consumer surplus that arises due to the higher prices, and returning any revenue raised lump-sum to consumers. We consider specific (or unit) taxes levied on ethanol content. Let τ denote the tax rate applied to the ethanol content in product j; the post-tax price of product j is therefore $p_j = \tilde{p}_j + \tau z_j$, where $\tilde{p}_j$ is the product's pre-tax price. Let $\boldsymbol{\tau}$ denote a vector of tax rates levied per unit of ethanol. We write indirect utility, tax revenue and the externality function directly as functions of $\boldsymbol{\tau}$. The social welfare function is:

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

We make three important assumptions about the planner's problem. First, we abstract from issues of market power by assuming that taxes are fully passed through to consumer prices and there is no producer surplus term in the planner's problem. This is consistent with each of the alcohol products being sold by a set of perfectly competitive retailers that drive prices down to marginal costs. The UK supermarket industry is competitive by international standards and alcohol taxes (only one component of firms' marginal costs) make up, on average, 60% of the price of alcohol sold in retailers. We therefore think that, in this context, the assumption is a reasonable approximation of reality. Second, 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. Third, we do not include a revenue raising constraint in the planner's problem. Inclusion of a revenue raising constraint would result in optimal tax rates comprising two components – a corrective component and a Ramsey-style revenue raising component (see Sandmo (1975) and Kopczuk (2003)).

2.2 Characterising tax policy

If the planner can set consumer specific tax rates, then the first best can be achieved by setting $\tau_i^* = \phi_i'(Z_i(\tau_i^*))$ for each consumer i. This is simply the Pigouvian result that the optimal consumer specific rate is set equal to the consumer's marginal consumption externality at that tax rate. However, in practice, setting consumer specific rates is infeasible for governments.

We consider optimal tax rates that are constrained to be the same across consumers. Let $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_K)'$ denote a set of tax rates. Tax rate τ_k applies to the set of products $\mathcal{K}_k$ (we refer to this as set k). $K \leq J$; $K = 1$ corresponds to a single rate ethanol tax, $K = J$ corresponds to tax rates that vary across all products and $K < J$ captures intermediate cases; most tax systems levy different tax rates on spirits, wine, beer etc. Let $Z_{ik} = \sum_{j \in \mathcal{K}_k} q_{ij}(\boldsymbol{\tau}) z_j$ denote consumer i's ethanol demand from the products belonging to set k and $\frac{\partial Z_{ik}}{\partial \tau_l} = \sum_{j \in \mathcal{K}_k} \frac{\partial q_{ij}}{\partial \tau_l} z_j$ denote the derivative of ethanol demand from set k with respect to a change in the tax rate that applies to products in set l. Tax revenue in this case is given by $R(\boldsymbol{\tau}) = \sum_k \left(\tau_k \sum_i Z_{ik} \right)$.

Taking the derivative of the planner's problem (equation 2.3) with respect to tax rate τ_l and applying Roy's identity yields:

$$\frac{\partial W}{\partial \tau_l} = \sum_i \sum_k (\tau_k - \phi_i') \frac{\partial Z_{ik}}{\partial \tau_l}, \tag{2.4}$$

where $\phi_i' \equiv \phi_i'(Z_i)$. The optimal set of tax rates $\boldsymbol{\tau}^*$ are implicitly defined by setting the first order conditions to zero (equation 2.4 for $l = 1, ..., K$). In general, $\boldsymbol{\tau}^*$

depends on the full set of substitution patterns between the different sets of products and their correlation with the marginal externalities.

When $K = 1$ we recover the optimal tax policy derived in Diamond (1973):

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

where $\bar{Z}' = \frac{1}{N}\sum_i \frac{dZ_i}{d\tau}$ is the average own tax slope of ethanol demand, $\bar{\phi}' = \frac{1}{N}\sum_i \phi_i'$ is the average marginal externality across consumers, and $\text{cov}(\phi_i', Z_i')$ denotes the covariance in the slope of ethanol demand and marginal externalities across consumers. The more strongly correlated are marginal externalities and the tax slope of ethanol demands, the more effective is the tax at correcting for the external costs of consumption and the higher is the optimal rate.

In general, when externalities vary across consumers, setting rates that vary across sets of products improves welfare, relative to a single tax rate. Specifically, this is the case if demand for different types of alcohol are correlated with the marginal externalities that an individual's alcohol consumption creates (i.e. as long as it is not the case that $\text{cov}(\phi_i', \frac{\partial Z_{ik}}{\partial \tau_l}) = 0 \ \forall \ (k, l)$). There are three obvious cases when these covariances are zero: (i) there is no heterogeneity in externalities, so $\phi_i' = \phi'$; (ii) there is no heterogeneity in demands, so $Z_{ik} = \bar{Z}_k \ \forall \ k$; or (iii) the heterogeneity in externalities and demands are uncorrelated. Under (i) all tax rates are set equal to the marginal externality, $\tau_k^* = \phi'$, and the first best is achieved; under (ii) and (iii) all tax rates are set equal to the average marginal externality, $\tau_k^* = \bar{\phi}' \equiv \frac{1}{N}\sum_i \phi_i'$, but the first best is not achieved.

When there is correlated heterogeneity in marginal externalities and demands, the optimal tax rate on a group of alcohol products is increasing in how popular the products are with individuals that generate large marginal externalities and it is increasing in how strongly those consumers reduce their ethanol demand in response to an increase in the tax rate. We show in Section 5 that the welfare gain due to moving from a single tax rate to tax rates that vary across different alcohol types depends crucially on the degree of heterogeneity in demand, externalities and their relationship.

3 Data

To estimate consumer demand in the alcohol market, we use data from the Kantar Worldpanel, which contain rich product information, repeated observations for each household, and accurately measured 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 transaction prices, product size, alcohol type and strength.[2] This type of data are becoming increasingly widely used in research (for example, see Aguiar and Hurst (2007) and Dubois et al. (2014)). For a detailed description of the data, see Griffith and O'Connell (2009) and Leicester and Oldfield (2009); Griffith et al. (2013) contains information on the alcohol segment of the data.

Our data have two substantial advantages over other data sources, such as cross sectional expenditure surveys (e.g. the Living Costs and Food Survey (LCFS) and the Consumer Expenditure Survey (CEX)) and intake diaries (e.g. the Health Survey for England (HSE) and National Health and Nutrition Survey (NHANES)). First, our data track households for a long period of time, meaning we can measure households' long run average alcohol purchases. Second, our data contain detailed information on purchases of alcohol products, including transaction prices and alcohol contents. A drawback of our data is that they do not include purchases of on-trade alcohol (those made in restaurants and bars). Our data covers the 77% of purchases of alcohol that are made off-trade in supermarkets and liquor stores (calculated using the LCFS). In Online Appendix A we show that the distribution of alcohol purchases from our data matches well with other data sources. We also show that the patterns of alcohol purchases are similar for both the off and on-trade parts of the UK market.

3.1 Households

We have a sample of 18,713 households that is representative of the British population; Table 3.1 shows that demographics of the sample of households in the Kantar Worldpanel are similar to the UK's nationally representative consumer expenditure survey, the Living Costs and Food Survey. We observe households for a minimum of 20 weeks in 2011, and for around 40 weeks per year, on average.

We estimate demand for alcohol products, so we use information on the 11,634 households in our sample that we observe buying alcohol in 2010 and 2011. We use the 2011 data to estimate demand for alcohol products and the 2010 data to group households based on how much alcohol they bought in this pre-sample period. Table 3.1 shows that this sample is similar to the full Kantar Worldpanel along key demographics.

[2]Strength is measured as percentage of alcohol-by-volume (ABV). This is defined as the number of millilitres of pure ethanol present in 100ml of solution at 20°C.

Conventions for measuring ethanol volume vary across countries. The US uses "standard drinks"; a standard drink contains 17.7 ml of ethanol. The UK, and many other European countries, use "units"; a unit contains 10 ml of ethanol. For each household we calculate the number of standard drinks that they purchase per adult household member in each week that we observe them in 2010. We take the average for each household across weeks to construct the household's average ethanol purchases in 2010. We observe each household for an average of 40 weeks in 2010, which means we measure whether households are consistently heavy drinkers.

Table 3.1: *Sample descriptive statistics*

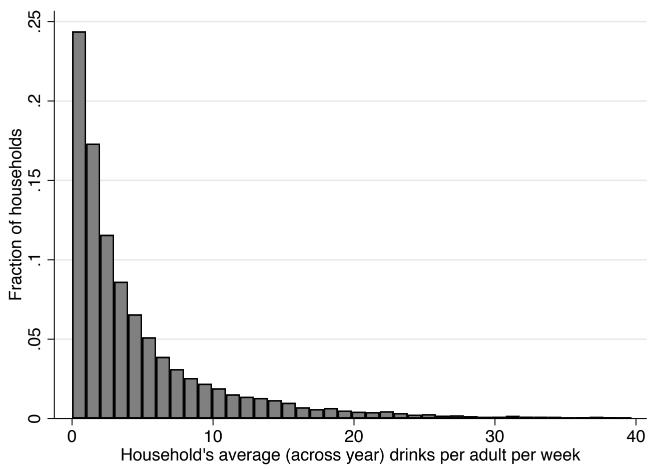
(1)	(2) LCFS	(3) Kantar Worldpanel	(2)
	All households	All households	Alcohol purchasers only
Number of households	5,691	18,713	11,634
Mean age of household's adult members	50.79	50.90	51.87
	[50.36, 51.22]	[50.68, 51.11]	[51.60, 52.14]
Number of household members	2.36	2.58	2.62
	[2.33, 2.39]	[2.56, 2.60]	[2.59, 2.64]
SES: Highly skilled	0.19	0.20	0.22
	[0.18, 0.20]	[0.20, 0.21]	[0.21, 0.23]
SES: Semi-skilled	0.53	0.57	0.58
	[0.51, 0.55]	[0.56, 0.58]	[0.57, 0.59]
SES: Unskilled	0.28	0.23	0.20
	[0.27, 0.29]	[0.22, 0.24]	[0.19, 0.20]
Region: North	0.34	0.35	0.35
	[0.33, 0.36]	[0.34, 0.35]	[0.34, 0.36]
Region: Central	0.34	0.33	0.33
	[0.33, 0.35]	[0.32, 0.33]	[0.32, 0.34]
Region: South	0.32	0.33	0.32
	[0.30, 0.33]	[0.32, 0.33]	[0.31, 0.33]

Notes: The first row shows the number of households in the Living Costs and Food Survey in 2011 (column (2)), the Kantar Worldpanel in 2011 (column (3)), and the Kantar Worldpanel in 2011, conditional on observing households buying alcohol in both 2010 and 2011. The remaining rows show the mean of each variable listed in column (1) for each of the samples. The SES and region variables are dummy variables. 95% confidence intervals are shown below each cell.

Figure 3.1 plots the distribution of average drinks per adult per week across households (based on 2010). We refer to this as the distribution of ethanol purchases. We use this to group households into five quintiles, with each quintile accounting for 20% of all drinks purchased. We show the share of households in each quintile and summary statistics for drinks per adult per week in Table 3.2. 64% of households are in the first, or bottom, quintile: that is the lightest 64% of alcohol consumers account for 20% of all drinks bought. The fifth, or top, quintile accounts for 20% of drinks, but contains only 3% of households. We use these quintiles as conditioning variables in our demand estimation. In Online Appendix A we show the distribution of average drinks per adult per week implied by the Living Cost and Food Survey, the main UK cross sectional expenditure survey (similar to the Consumer Expenditure Survey in the US) is very similar.

High average drinks per adult per week over a sustained period of time may be due to consumers drinking large amounts regularly or engaging in less regular very high consumption (binge drinking). Both types of drinking behaviour can lead to externalities, although the nature of these externalities may differ. In Online Appendix A we show that in both the UK and the US people who report consuming more ethanol also report drinking more days per week and are more likely to have reported binge drinking in the previous week.

Figure 3.1: *Distribution of drinkers*



Notes: Distribution drawn across 11,634 households over 2010 (pre-sample).

Table 3.2: *Quintiles of drinking distribution*

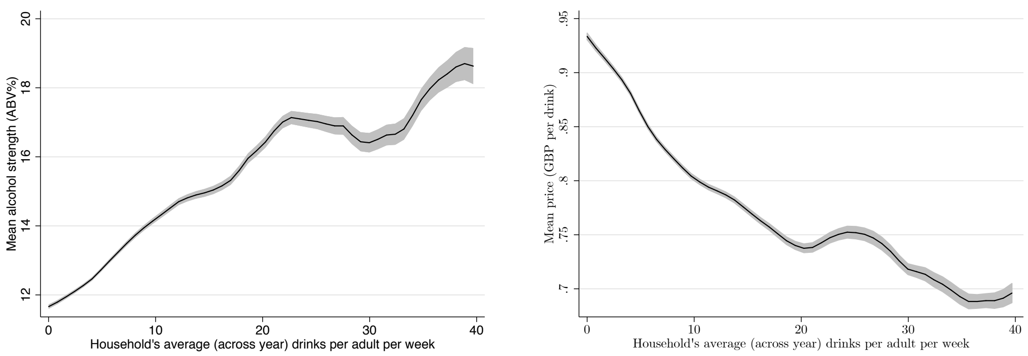
	Quintile of drinkers				
	1	2	3	4	5
% total ethanol consumption	20%	20%	20%	20%	20%
% households	64%	17%	10%	6%	3%
Standard drinks per adult per week:					
mean	1.7	6.1	10.9	17.9	35.1
min	0.0	4.4	8.5	14.1	23.7

Notes: We split the distribution of drinkers into five quintiles based on total drinks purchased (per adult) in 2010. Table shows the percentage of households in each quintile and the mean, minimum drinks per adult per week across households for each group in 2010. The minimum for each group defines the cut-off between the groups. A standard drink is 17.7ml of ethanol.

Alcohol purchase patterns vary substantially across the distribution of drinkers. Figure 3.2 plots the relationship across households between mean average alcoholic strength (panel (a)) and price of products purchased (panel (b)) with the average number of drinks purchased per adult per week. Heavier drinkers tend to purchase stronger types of alcohol, on average. This is because the heaviest drinking households both buy proportionately more spirits, and less beer, than lighter drinkers, and buy stronger products within these broad categories. The heaviest drinkers also buy products that are cheaper in per-drink terms. This suggests that a tax system that increases the relative prices of strong and cheap products may successfully target the consumption of heavy drinkers. Whether this is indeed the case depends on: (i) how strongly different households (e.g. light versus heavy drinkers) switch away from the products in response to a tax rise; (ii) how strongly and to what alternative alcohols they switch; and (iii) what fraction of drinking externalities are accounted for by the set of heavy drinkers.

Figure 3.2: *Average alcoholic strength and price of products purchased, across distribution of drinkers*

(a) Alcoholic strength (b) Price



Notes: For each household-week in 2011 we calculate the average alcoholic strength and price per drink. The lines plot an estimated local polynomial (bandwidth = 2) between these variables and the average number of drinks purchased per adult per week, measured in the pre-sample period, for each household. The grey lines are 95% confidence intervals.

3.2 Products

Over 7000 alcohol UPCs (barcodes) and 3000 brands are recorded as being purchased in our data. We aggregate UPCs into 32 "products" to focus on the margins of substitution that are most relevant to our application, see Table 3.3. It is important that we capture heterogeneity in the shape of demand for sets of UPCs that are impacted similarly by alcohol tax changes and it is also important that we capture how changes in taxes and hence prices affect the total quantity of alcohol that households purchase. We are therefore careful not to aggregate over UPCs that have different alcohol strengths or UPCs likely to be subject to different tax treatment and, as much as possible, only aggregate across UPCs that are of a similar alcohol type, quality and price.

Table 3.3: *Product definition and characteristics*

(1) Product definition	(2) Top brand and within-product share (%)	(3) No. brands	(4) Mean ABV	(5) Market share (%)	(6) No. sizes
Beer					
Premium beer; ABV $<$ 5%	Newcastle Brown Ale (6.1)	386	4.4	1.8	3
Premium beer; ABV $\geq$ 5%	Old Speckled Hen (16.5)	238	5.5	2.1	3
Mid-range bottled beer	Budweiser Lager (19.6)	94	4.7	4.6	3
Mid-range canned beer; ABV $<$ 4.5%	Carlsberg Lager (28.8)	17	3.9	5.8	3
Mid-range canned beer; ABV $\geq$ 4.5%	Stella Artois Lager (72.0)	15	5.0	2.7	3
Budget beer	John Smiths Bitter (23.6)	72	4.2	3.2	3
Wine					
Red wine	Tesco Wine (6.2)	439	12.6	18.4	4
White wine	Tesco Red Wine (7.8)	327	12.1	17.1	4
Rose wine	Echo Falls Wine (8.6)	67	11.5	4.2	2
Sparkling wine	Lambrini Sparkling Wine (8.4)	125	9.2	3.1	2
Champagne	Lanson Champagne (12.7)	42	11.8	0.8	1
Port	Dows Port (22.0)	23	19.8	0.7	1
Sherry	Harveys Bristol Cream (18.7)	25	16.8	1.2	1
Vermouth	Martini Extra Dry (11.8)	33	15.0	0.6	1
Other fortified wines	Tesco Fortified Wine (21.8)	37	14.6	0.9	1
Spirits					
Premium gin	Gordons Gin (59.6)	21	38.3	1.6	2
Budget gin	Tesco Gin (22.3)	15	38.3	1.3	2
Premium vodka	Smirnoff Red Vodka (39.0)	54	37.6	3.1	2
Budget vodka	Tesco Vodka (31.4)	17	37.5	1.8	2
Premium whiskey	Jack Daniels Bourbon/Rye (19.6)	80	40.5	2.1	2
Budget whiskey	Bells Scotch Whiskey (18.7)	56	40.0	8.1	2
Liqueurs; ABV <30%	Baileys (25.9)	203	18.4	3.1	2
Liqueurs; ABV $\geq$30%	Southern Comfort (27.2)	41	37.0	0.8	2
Brandy	Tesco Brandy (22.1)	55	37.3	2.4	2
Rum	Bacardi White Rum (29.1)	58	37.1	2.0	2
Pre-mixed spirits	Gordons Gin+Tonic (14.7)	43	6.1	0.2	1
Alcopops	Smirnoff Ice Vodka Mix (17.3)	147	4.8	0.8	1
Cider					
Apple cider, <5% ABV	Magners Original Cider (26.9)	52	4.4	1.6	3
Apple cider, 5-6% ABV	Strongbow Cider (63.1)	49	5.3	2.0	3
Apple cider, >6% ABV	Scrumpy Jack Cider (18.7)	71	7.0	0.8	2
Pear cider	Bulmers Pear Cider (24.2)	33	4.9	0.7	2
Fruit cider	Jacques Fruit Cider (21.4)	48	4.4	0.5	2

Notes: Column (1) shows the product definition. Column (2) lists the brand that constitutes the largest share of spending within each product; its within-product expenditure share is shown in parentheses. Column (3) lists the number of brands within each product. Column (4) shows the mean alcoholic strength (ABV) of each product. Column (5) shows the share of the alcohol market accounted for by each product. Column (6) shows the number of bins used to divide the quantity distribution.

Table 3.3 shows that for many of our 32 products, one brand constitutes the majority of the spending on that product. However, other products consist of many smaller brands, for example, wine and premium bottled beer. We consider tax systems that vary tax rates across different types and strengths of alcohol. The most important consumer substitution resulting from changes in tax systems such as these is between different alcohol types and strengths. This switching is well captured by our 32 products. If we were interested in alcohol tax systems that set different rates for, say, Spitfire Kentish Ale and Badger Golden Glory – two different brands of premium beers, each with 4.5% ABV – then it would be important to model substitution within low strength premium beers.

To model the quantity of a product that households choose, we discretize the within product quantity distribution into a set of sizes for each product. We allow the number of sizes to vary across products depending on how dispersed the quantity distribution is – in Table 3.3 we list the number of size categories we use for each product. For example, for red wine, which constitutes around 18% of total alcohol spending, we define 4 categories – 1 bottle, 2 bottles, 3 bottles and more than 3 bottles. In total there are 69 product-sizes. In Online Appendix A we plot the distributions of quantity for each product, we also show that the distribution of drinks purchased per household-week computed using our discrete size categories very closely matches the distribution observed directly in the data.

3.3 Prices

For each product-size we compute a price index that we use in our model. The index captures price movements of the underlying UPCs that comprise the product. We compute a weighted average of the UPC prices using weights that are fixed over time. Let b index UPC (or barcode), j index product, s index size, r index region, t index time, and f index retailer. The barcode b is sold at price ρ_{bft} in retailer f at time t. In the UK the main retailers set national prices. Let $\mathcal{B}_{js}$ denote the set of barcodes that belong to product j in size s. The region r, time t price index for product j in size s is:

$$p_{jsrt} = \sum_{b \in \mathcal{B}_{is}, f} w_{bfr} \rho_{bft}, \qquad w_{bfr} = \frac{N_{bfr}}{\sum_{b' \in \mathcal{B}_{is}, f'} N_{b'f'r}} \tag{3.1}$$

where N_{bfr} denotes the number of purchases of barcode b from retailer f in region r across the entire time period. The regional dimension to the weights captures geographical variation in retailer coverage. Note, we also allow the weights to vary across the five drinking quintiles, capturing the possibility that the popularity of UPCs within product-sizes varies across these quintiles – however, we omit a household quintile index for notational simplicity.

In Online Appendix A we report average prices and plot the price series for each of the 69 product-size pairs. There is considerable differential time series variation in price across products. We discuss how this price variation allows us to identify the effect of price changes on demand in Section 4.2.

4 Demand model

We specify a model of consumer demand in the alcohol market. The model embeds the decision of whether or not buy alcohol, what product to choose and in what quantity. It also incorporates heterogeneity in preferences, allowing for the possibility that the shape of product level demands are correlated with where households are in the distribution of ethanol purchases (and hence what level of externality their drinking is likely to create).[3]

4.1 Empirical demand specification

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 54% of purchase occasions. On the remaining purchase occasions households choose the "outside option" of no alcohol. We model the decision over whether to buy alcohol and which option to choose as a discrete choice. A discrete choice demand framework naturally accommodates 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 17% of household-week observations, a household purchases more than one (typically two or three) alcohol products. We treat this behaviour as the household undertaking multiple separate purchase occasions. In total we have data on 632,810 purchase occasions.

We index households by i and products by j. $j = 0$ denotes the option of purchasing no alcohol, $j = 1, \ldots, J$ indexes different alcohol products. Products are available to the consumer in discrete sizes, indexed by s. We model the decision over which product-size, (j, s), to select, with the option to purchase no alcohol denoted $(0, 0)$. We use t to index time (i.e. weeks).[4]

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 that household i obtains from selecting option (j, s) in period t is given by:

$$u_{ijst} = \nu(p_{jsrt}, \mathbf{x}_{jst}; \boldsymbol{\theta}_i) + \epsilon_{ijst}, \tag{4.1}$$

[3]Although we condition the entire preference distribution on pre-sample ethanol consumption, we do not explicitly model state dependence. In Online Appendix B we provide some reduced form evidence that once preference heterogeneity is accounted for state dependence in demand appears not to be of first order importance.

[4]For households that purchase multiple (i.e. two or three) different alcohol products in a weeks, we have multiple observations per week.

where p_{jsrt} is the price of option (j, s) in period t and region r, $\mathbf{x}_{jst}$ is a vector of option characteristics (including a time-varying unobserved attribute), and $\boldsymbol{\theta}_i$ 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 utility from purchasing no alcohol so that $u_{i00t} = \epsilon_{i00t}$.

Households select the option (j, s) that provides them with the highest utility. Integrating across the demand shocks, $\boldsymbol{\epsilon}_{it} = (\epsilon_{i00t}, ..., \epsilon_{iJSt})'$, 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:

$$q_{ijst} = \frac{\exp(\nu(p_{jsrt}, \mathbf{x}_{jst}; \boldsymbol{\theta}_i))}{1 + \sum_{j'>0,s'>0} \exp(\nu(p_{j's'rt}, \mathbf{x}_{j's't}; \boldsymbol{\theta}_i))} \tag{4.2}$$

and expected utility is given by:

$$v_{it}(\mathbf{p}_{rt}, \mathbf{x}_t) = \ln \sum_{j>0,s>0} \exp\{\nu(p_{jsrt}, \mathbf{x}_{jst}; \boldsymbol{\theta}_i)\} + C \tag{4.3}$$

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

We use $d = 1, \ldots, D$ to index the household quintiles defined in Section 3.1. Each quintile comprises 20% of total drinks purchased (the first quintile contains the lightest drinkers, the fifth contains the heaviest). We index the four segments of the alcohol market (cider, beer, wine and spirits), $m = 1, \ldots, 4$. We assume that the payoff function ν for household i belonging to drinking quintile $d(i)$ and for option (j, s) in alcohol segment $m(j)$ takes the form:

$$\nu_{it}(.) = \alpha_i p_{jsrt} + \beta_i w_j + (\gamma_{i0} + \gamma_{d(i)m(j)1}) z_{js} + \gamma_{d(i)m(j)2} z_{js}^2 + \xi_{d(i)j} + \eta_{im(j)} + \chi_{d(i)jt}. \tag{4.4}$$

w_j denotes the alcohol strength of the product and z_{js} denotes the amount of pure alcohol (ethanol) in the option.[5] We allow z to affect the utility from option (j, s) through a quadratic function. We model the first order term as having a household specific component (γ_{i0}), which captures heterogeneity in preferences for total ethanol across households, and a component that is household group specific and varies across alcohol segments ($\gamma_{d(i)m(j)1}$), capturing the fact that preferences over ethanol may vary depending on ethanol type. We also model the second order

[5] The alcohol strength of a product is the amount of ethanol it contains per litre of product i.e. $w_j = z_{js}/L_{js}$, where L_{js} is the size in litres of product-size (j, s). It is only necessary to include two of these three variables to capture both preferences over product size and alcohol strength.

term as household group specific and variable across alcohol segments ($\gamma_{d(i)m(j)2}$). The quadratic term allows for there to be diminishing marginal utility in product size and for this to vary across the four segments of the market.

We also include $\xi_{d(i)j}$, which are time-invariant product effects that capture consumers' preferences over unobserved product attributes that are fixed over time. We allow the product effects to vary across the five household quintiles, capturing the possibility that preferences over unobserved product attributes are correlated with how heavily a household consumes alcohol. $\eta_{im(j)}$ are household specific preferences over alcohol segments. This captures the possibility that some households have preferences for certain types of alcohol – e.g. one household may have a stronger preference for beer than another. The $\eta_{im(j)}$ effects allow 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. In addition, we include $\chi_{d(i)jt}$, which captures variation in preferences for unobserved alcohol attributes over time (due, for instance, to the effects of advertising or seasonal demand patterns).[6]

We model the household specific preferences, which are over price (α_i), alcohol strength (β_i), ethanol content (γ_{i0}) and segments ($\eta_{i1}, \ldots, \eta_{i4}$) as random coefficients. In each case we specify the distribution to be normal conditional on household quintile d. Therefore, the overall random coefficient distribution is a mixture of normal distributions. Conditional on d we also allow for correlation between preferences over price, strength and ethanol content. Note that the means (conditional on d) of the strength and alcohol segment random coefficients are collinear with the product effects, therefore we normalise these means to zero.

4.2 Identification of demand parameters

We use longitudinal micro data; for each household in our sample we observe many repeated choices. The vector of product prices that households face varies cross sectionally across regions and over time. How households adjust their behaviour in response to these changes aids identification of the preference parameters. A number of papers (e.g. Berry and Haile (2010), Berry et al. (2004)) have highlighted the powerful identifying role that micro data (compared with more commonly used market level data) plays in pinning down parameters in choice models.[7]

[6] Note, we write these as varying across products here, but in practice, for reasons of parsimony, we constrain some of the time varying effects to be the same for similar products – for example for branded and store brand gin.

[7] Berry and Haile (2010) and Fox and Gandhi (2016) establish conditions for nonparametric identification of random coefficients in random utility discrete choice models by placing restrictions

We aim to exploit price variation that is driven by factors such as input prices and alcohol tax rates that are determinants of marginal cost. A possible concern is that some of the variation in price that we use reflects demand shocks, due to firms altering their prices in response to fluctuations in demand, and hence prices are correlated with changes in demand that are not controlled for and are collected in the term ϵ_{ijst}.

To limit the possibility of this contaminating our estimates we control for a detailed set of fixed effects – including product effects, time varying product effects and region effects. We also include a control function based on the instruments: tax changes, producer prices, exchange rates and regional transport costs. Our exclusion restriction is that *conditional on all the controls and fixed effects in demand* these instruments affect demand through their impact on price and are independent of residual demand variation in the error term. The threat to identification is that this restriction does not hold and changes in the instruments are correlated with ϵ_{ijst}. We think that is unlikely to be the case for the following reasons.

We include a rich set of unobserved characteristics that control for a number of possible sources of price endogeneity arising from demand side price drivers. The vector of product effects controls for unobserved quality differences across products, which are likely to be correlated with price, and the product-time effects control for seasonality in demand and spikes in demand due to advertising campaigns. In addition, the practice of UK supermarkets of pricing products nationally limits the scope for geographical variation in prices driven by local demand shocks.[8]

Nevertheless, we cannot rule out the possibility that there may by some residual omitted demand side variables that are correlated with prices. We therefore include a control function for price that isolates price variation driven by a set of instruments that we expect to shift firm costs, but not to directly impact on demand (see Blundell and Powell (2004) and, for multinomial discrete choice models, Petrin and Train (2010)).

Our instrument set includes alcohol duty rates; Table B.3 in the Online Appendix describes changes in the duty applied to different types of alcohol over our estimation period. It also includes producer prices for beer and cider, which are likely to be drivers of the consumer price of beer and cider options. The price

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.

[8]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)).

indices are factory-gate prices for beer and cider and are produced by the UK Office for National Statistics. Producer prices are also likely to vary seasonally due to demand fluctuations. It is therefore important that in demand we control for time effects, which addresses this concern.[9] Also included are the sterling-euro and sterling-dollar exchange rates, which affect the price of imported alcohol.

The main reason for regional variation in prices in the UK is differences in the geographical coverage of food retailers (which set national prices). We control for differences in regional coverage of retailers by including regional effects interacted with a dummy for the outside option. This controls for the possibility that retailers with systematically high or low prices may choose to locate in regions that have consumers with either systematically strong or weak preferences for alcohol. We use as an instrument the price of oil interacted with region dummies to capture regional variation in transport costs, and thus differential time series variation in regional costs. In Online Appendix B we describe in more detail the variation in the instruments.

The F-statistic for a test of the (ir)relevance of the instruments is 36.4, which means we strongly reject the hypothesis of no relationship between price and the instruments.[10] In demand estimation we control for the predicted residuals of the first stage regression. The residuals enter positively and statistically significantly (see Table 4.1) and the price coefficients become more negative when the control function is included. This indicates that the omission of the control function would lead to a (modest) bias towards zero of the price coefficients.

4.3 Estimation

We estimate preferences conditional on the quintile to which households belong, meaning we can estimate the model separately quintile-by-quintile. We estimate demand using maximum simulated likelihood. Conditional on preference draws, the probability a household selects a given option on a given purchase occasion is given by the closed form of equation (4.2). This follows from our assumption that the idiosyncratic utility shocks, ϵ_{ijst}, are i.i.d. type I extreme value. To construct the likelihood function we integrate across the random coefficient distribution. Let $(1, \ldots, T_i)$ 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

[9] We also estimate the model dropping producer prices from the instrument set. This results in somewhat weaker instruments, but does not have a material impact on the estimated demand coefficients.

[10] In Online Appendix B we provide the F-statistics for the joint significance of subsets of the instruments.

purchase occasion t. The contribution that household i makes to the likelihood function is then:

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

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

4.4 Demand estimates and elasticities

In Table 4.1 we report the coefficient estimates for our demand model.[11] The means of the price preference distribution for each household quintile are all negative and statistically significant. The variances of the preferences for price, strength and ethanol for each household quintile all indicate statistically significant within quintile preference heterogeneity The covariance parameters show that, within each quintile, more price sensitive consumers typically have relatively strong preferences over quantity of ethanol and alcohol strength, with the exception of the heaviest drinkers. Panel B presents estimates of the average of the mean product effects within each alcohol segment (relative to the utility from the outside option). The light drinking households in the bottom quintile of the ethanol purchase distribution have the lowest mean product effects for each segment on average. However, the segment specific variance parameters indicate that, as with the observable product attributes, there is a high degree of within quintile preference heterogeneity.

Table 4.2 summarises the key patterns in the own and cross price elasticities; the full matrices of elasticities for each quintile are available in the Online Appendix. There is some variation in the mean own price elasticity across quintile, with the top quintile, on average, having the most price elastic product level demand. However, the variation in the mean cross price elasticity across the household quintiles is much more striking. The mean cross price elasticity of households in the heaviest drinking quintile is over 3.5 times as high as the mean for the lightest drinking bottom quintile. The heaviest 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). A consequence of this is that when we simulate the overall price elasticity of demand for ethanol (i.e. what is the % change in demand that follows a 1% price increase in all alcohol) households in the top quintile are much less price sensitive; their own price elasticity is -0.96 compared with -2.26 for the bottom quintile.

[11] To estimate the model we randomly sample 500 households from each household group (with the exception of group 5, for which we use all households). For each drawn household we use 50 randomly sampled purchase occasion (or all of their purchase occasions if this is less than 50). We conduct all post estimation analysis on the full sample.

Table 4.1: *Estimated preference parameters*

Drinking quintile:	1	2	3	4	5
Panel A: Preferences for observable product characteristics					
Means					
Price	-0.336	-0.260	-0.294	-0.341	-0.388
	(0.032)	(0.021)	(0.018)	(0.018)	(0.019)
Beer*Quantity of ethanol	0.139	0.167	0.163	0.193	0.229
	(0.018)	(0.012)	(0.010)	(0.010)	(0.011)
Wine*Quantity of ethanol	0.010	0.021	0.084	0.127	0.155
	(0.020)	(0.013)	(0.011)	(0.011)	(0.013)
Spirits*Quantity of ethanol	0.309	0.247	0.271	0.325	0.430
	(0.036)	(0.023)	(0.020)	(0.020)	(0.021)
Cider*Quantity of ethanol	0.069	0.090	0.120	0.182	0.188
	(0.019)	(0.012)	(0.011)	(0.011)	(0.011)
Beer*Quantity of ethanol^2	-0.002	-0.002	-0.001	-0.002	-0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wine*Quantity of ethanol^2	0.001	0.001	0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Spirits*Quantity of ethanol^2	-0.004	-0.003	-0.003	-0.003	-0.004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cider*Quantity of ethanol^2	-0.001	-0.001	-0.002	-0.002	-0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Variances ×100					
Price	1.826	1.443	1.842	1.060	0.777
	(0.289)	(0.152)	(0.127)	(0.080)	(0.096)
Quantity of ethanol	0.620	0.276	0.387	0.175	0.337
	(0.069)	(0.023)	(0.022)	(0.010)	(0.019)
Strength	0.194	0.281	0.307	0.451	0.522
	(0.024)	(0.016)	(0.015)	(0.020)	(0.029)
Covariances ×100					
Price*Quantity of ethanol	-0.838	-0.441	-0.646	-0.269	-0.414
	(0.130)	(0.057)	(0.046)	(0.024)	(0.040)
Price*Alcohol strength	-0.283	-0.223	-0.463	-0.219	0.345
	(0.041)	(0.033)	(0.030)	(0.018)	(0.022)
Quantity of ethanol*Alcohol strength	-0.025	-0.055	0.011	-0.096	-0.233
	(0.028)	(0.013)	(0.016)	(0.009)	(0.013)
Panel B: Preferences for unobserved product characteristics					
Mean product effects for each segment					
Beer	-5.652	-4.400	-3.293	-3.349	-2.959
	(0.152)	(0.129)	(0.139)	(0.127)	(0.185)
Wine	-3.660	-2.559	-2.232	-2.024	-1.929
	(0.181)	(0.141)	(0.157)	(0.136)	(0.201)
Spirits	-8.861	-6.645	-5.711	-6.050	-6.651
	(0.357)	(0.265)	(0.251)	(0.246)	(0.295)
Cider	-5.237	-4.106	-3.544	-3.945	-3.226
	(0.173)	(0.136)	(0.131)	(0.130)	(0.165)
Variances					
Beer	1.903	2.423	2.039	1.944	2.189
	(0.155)	(0.133)	(0.098)	(0.097)	(0.126)
Wine	1.542	1.217	1.935	1.424	1.597
	(0.111)	(0.065)	(0.094)	(0.076)	(0.103)
Spirits	0.800	0.407	1.077	0.681	0.880
	(0.098)	(0.046)	(0.092)	(0.085)	(0.081)
Cider	3.060	5.577	3.105	4.252	4.074
	(0.272)	(0.337)	(0.189)	(0.256)	(0.268)
Control function	0.264	0.174	0.282	0.303	0.337
	(0.040)	(0.024)	(0.019)	(0.018)	(0.020)
Product effects	Yes	Yes	Yes	Yes	Yes
Type-time effects	Yes	Yes	Yes	Yes	Yes
Outside option-region effects	Yes	Yes	Yes	Yes	Yes
Number of households	500	500	500	500	351
Number of purchase occasions	21,638	22,820	23,616	23,958	16,959

Notes: Drinking quintiles are defined in Table 3.2. 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.

Table 4.2: *Mean own and cross price elasticities*

Drinking quintile:	1	2	3	4	5
Mean option level own price elasticities					
All alcohol	-4.01	-2.91	-3.59	-4.07	-4.69
Beer	-4.17	-2.78	-3.00	-3.70	-4.87
Wine	-3.31	-2.45	-3.28	-3.72	-4.58
Spirits	-5.71	-4.12	-5.00	-5.09	-5.04
Cider	-2.70	-1.79	-2.11	-2.23	-3.14
Mean option level cross price elasticities					
All alcohol	0.045	0.038	0.077	0.101	0.169
Total alcohol own price elasticities					
All alcohol	-2.26	-1.28	-1.10	-1.21	-0.96

Notes: Drinking quintiles are defined in Table 3.2. The top panel shows average option level elasticities own price elasticities; the middle panel shows average option level elasticities cross price elasticities. The bottom panel shows the total alcohol own price elasiticities; what is the % change in ethanol demand when all alcohol options price rise by 1%.

In Table 4.3 we show the cross-segment diversion ratios by quintile of the drinking distribution. For example, the first panel shows which segments consumers shift toward if the price of strong beer increases; each column sums to 100. For the lowest quintile of drinkers 48.9% of the substitution away from strong beer is towards the outside option, 22.3% is toward weak beer, 18.3% toward wine, 6% toward spirits and 4.5% toward cider. These patterns of substitution are heterogeneous across alcohol segments and across drinking quintiles. The quintiles of lighter drinkers substitute more heavily toward the outside option, but more so in some segments than others. For the highest quintile, when the price of high strength spirits increases, almost 50% of the demand that switches away goes to wine, while this is only 20% for the lowest quintile. This table illustrates the importance of modelling demand for all alcohol segments jointly, rather than considering the segments independently.

Table 4.3: *Cross segment diversion ratios*

Drinking quintile::	1	2	3	4	5
Beer (>5% ABV)					
Beer (≤5% ABV)	22.3	39.2	36.4	33.6	39.5
Wine	18.3	15.2	21.6	29.6	24.0
Spirits	6.0	8.1	16.3	13.9	13.5
Cider	4.5	6.4	7.0	6.1	10.5
Outside option	48.9	31.1	18.6	16.8	12.4
Beer (≤5% ABV)					
Beer (>5% ABV)	8.2	16.8	19.8	14.5	19.5
Wine	23.4	20.4	27.4	38.3	32.6
Spirits	7.1	10.6	20.1	17.3	17.9
Cider	5.6	8.8	9.1	7.8	13.8
Outside option	55.7	43.4	23.6	22.1	16.2
Wine (>14% ABV)					
Beer	8.8	10.4	12.5	14.4	14.2
Wine (≤14% ABV)	37.7	36.5	47.8	43.0	46.2
Spirits	6.8	14.3	17.7	22.2	23.9
Cider	3.3	6.3	6.4	6.3	6.8
Outside option	43.4	32.6	15.5	14.1	8.9
Wine (≤14% ABV)					
Beer	14.9	16.2	21.8	26.9	27.6
Wine (>14% ABV)	5.6	5.1	11.3	9.3	10.4
Spirits	8.4	17.4	25.3	28.5	33.2
Cider	4.6	10.3	12.4	11.2	13.6
Outside option	66.5	51.0	29.2	24.2	15.2
Spirits (>20% ABV)					
Beer	11.5	14.0	22.1	17.8	22.2
Wine	20.1	32.7	37.9	46.5	49.9
Spirits (≤20% ABV)	4.2	5.0	5.4	4.8	6.8
Cider	5.1	9.9	12.3	8.2	10.2
Outside option	59.1	38.4	22.3	22.6	11.0
Spirits (≤20% ABV)					
Beer	8.4	14.7	17.4	16.3	16.6
Wine	17.1	25.1	24.9	32.8	25.6
Spirits (>20% ABV)	8.6	16.3	28.1	26.8	35.1
Cider	5.2	6.8	8.5	5.6	8.6
Outside option	60.6	37.1	21.1	18.6	14.1
Cider (>5% ABV)					
Beer	13.5	14.4	14.5	14.1	24.6
Wine	15.9	21.1	26.5	29.5	27.4
Spirits	8.0	10.4	19.5	15.2	12.9
Cider (≤5% ABV)	9.9	28.6	22.4	22.4	20.7
Outside option	52.7	25.6	17.0	18.9	14.4
Cider (≤5% ABV)					
Beer	10.2	14.1	13.8	14.6	17.1
Wine	15.0	21.0	22.8	29.2	21.4
Spirits	7.0	11.0	16.3	13.3	14.2
Cider (>5% ABV)	7.9	26.4	28.1	25.7	33.9
Outside option	60.0	27.5	19.1	17.3	13.5

Notes: Drinking quintiles are defined in Table 3.2. The rows show the % demand that goes to each segment when the price of the alcohol type indicated at the top of each panel increases. Within each panel, the rows sum to 100.

5 Optimal alcohol taxes

In this section we combine our estimates of households' alcohol demands with the optimal tax framework from Section 2 to calculate optimal alcohol taxes.

5.1 Externality function

In Section 2 we specify the argument of the consumer level externality function, $\phi_i(.)$, to be total ethanol demand Z_i. We do not observe individual household members' alcohol consumption. Our data have details of alcohol purchases made by households, so 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 (weakly) convex function and that its shape does not vary across households i.e. $\phi_i(.) = \phi(.)$ for all i; hence, differences in marginal externalities across people are driven by differences in their level of ethanol demand.[12]

We parametrise the externality function as quadratic with parameters, (ϕ_0, ϕ_1):

$$\phi(Z_{it}) = \phi_0 Z_{it} + \phi_1 Z_{it}^2 \tag{5.1}$$

(ϕ_0, ϕ_1) jointly determine the aggregate external cost and degree of convexity of the function. We calibrate the externality function to match the aggregate external cost estimate 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).[13] There are many potential costs associated with excess alcohol consumption, including health costs, the impact of crime, lost productivity etc. We use the estimates from Cnossen (2007) of the 'direct' costs, which are primarily the costs of alcohol related crimes (e.g. administering the criminal justice system, victim support) and the health costs of alcohol related diseases. The

[12]It is generally accepted that men generate more externalities from drinking than women. The World Health Organization argues this is principally a consequence of men drinking more rather than creating more externality for a given level of consumption ; *"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"*. (World Health Organization (2014)).

[13]The estimate reported in the paper was £7.5 billion in 2001 prices; we uprate this to 2011 prices using the Retail Price Index (RPI) 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%).

costs of drink driving are included, but constitute only 1% of the total external costs of drinking that we use for calibration.

There is a large body of evidence that suggests the external costs of drinking are highly concentrated among a small number of heavy drinkers (see Cnossen (2007) for a survey). Relatedly, there is a considerable amount of evidence that externalities from alcohol consumption are convexly increasing in ethanol consumed (and hence the marginal externality associated with an additional drink is increasing in number of drinks 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). A recent, large global study found that though there is no 'safe' or 'beneficial' levels of alcohol consumption, the health risks from low levels of alcohol consumption are small (GBD 2016 Alcohol Collaborators (2018)). Although there is considerable evidence that the external costs of alcohol consumption are convex, there is little evidence on the precise degree of convexity of the relationship. Therefore we remain agnostic about this and show how the optimal tax rates change as we vary the degree of convexity in the relationship from zero (i.e. constant marginal externality) to a high degree of convexity.

Specifically, we define households in the top three quintiles of the ethanol distribution as "heavy drinkers", which are households that buy more than government recommended levels. On average they buy more than 8.5 standard drinkers per adult per week. In the UK, the recommendations are to drink no more than 8 standard drinks,[14] irrespective of gender. The US government recommends that women consume no more than 1 drink per day, and men no more than 2.[15] This group of heavy drinkers purchase 60% of total ethanol, but make up only 19% of households. We refer to the remaining 81% of households that purchase less than government recommended levels as "light drinkers"; together they buy 40% of ethanol.

If the marginal external costs of drinking are constant, then the heavy drinkers would generate 60% of the external costs (as they buy 60% of the ethanol). As the convexity of the externality function increases, the share of costs generated by the heavy drinkers increases. We calibrate (ϕ_0, ϕ_1) to eight specifications, in which heavy drinkers generate: 60%, 65%, 70%, 75%, 80%, 85%, 90% and 95% of the external costs of drinking. The assumption that the heavy drinkers generate 95% of the externality implies that the marginal externality of somebody who drinks 40

[14]The UK government recommends not more than 14 units per week, which is equivalent to 8 standard drinks.

[15]https://www.cdc.gov/alcohol/fact-sheets/moderate-drinking.htm

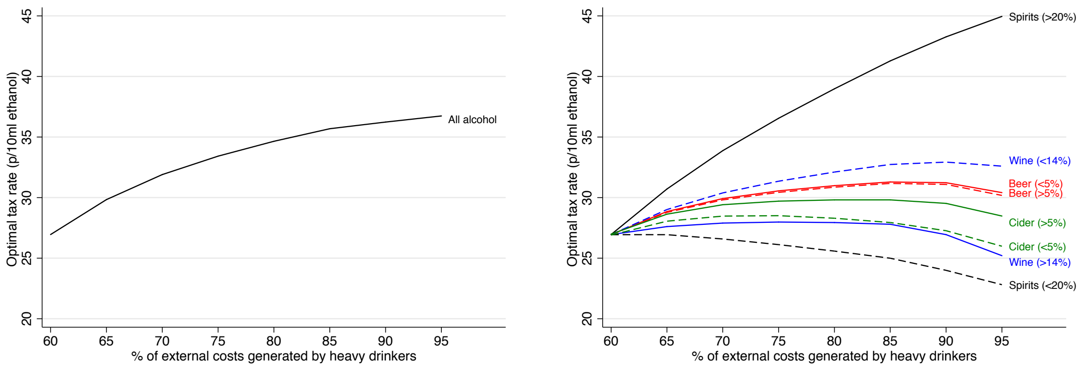
standard drinks a week is 8 times as large as the marginal externality of somebody who drinks 8 standard drinks a week. We plot the externality functions for the 60%, 65%, 80% and 95% specifications in Online Appendix B.[16] The Centers for Disease Control and Prevention (2016) estimate that binge drinkers, who constitute 17% of the population, are responsible for 77% of the costs of excessive alcohol use.

5.2 Optimal tax rates

We consider two optimal tax systems. First, we consider a situation in which the planner is constrained to set a single tax rate levied per unit of ethanol, common across all products. Second, we consider the case in which the planner can set multiple tax rates, varying them across eight alcohol types; beer (>5% ABV), beer (≤5% ABV), wine (>14% ABV), wine (≤14% ABV), spirits (>20% ABV), spirits (≤20% ABV), cider (>5% ABV) and cider (≤5% ABV). In this multi rate case, we choose to let rates vary across alcohol types in a way that is similar to existing tax systems in many countries; allowing for tax rates to vary over more disaggregate alcohols (e.g. a different vodka and gin rate) simply magnifies our conclusions about the welfare gains of rate differentiation.

Figure 5.1: *Comparison of optimal tax rates under different externality function calibrations*

(a) Optimal single rate　　　　(b) Optimal multi rate



Notes: The figures show the optimal tax rates under various calibrations of the convexity of the externality function, shown on the horizontal axis. Light drinkers are those that on average buy below government recommended levels, and heavy drinkers are those that buy on average above recommended levels. The vertical axis show the optimal tax rate (p/10ml ethanol). The top panel shows the optimal single tax rate applied to all alcohol products and the bottom panel shows the optimal multi rate system applied to the 8 different alcohol types.

In Figure 5.1 we show how the optimal tax rates vary with how convex the externality function is. When heavy drinkers account for 60% of the external costs,

[16] Note, we place a lower bound on the externality function of zero. For the most convex specifications this binds at low levels of ethanol consumption.

the marginal externality is constant. The higher is the fraction of externalities accounted for heavy drinkers, the more convex the externality function becomes. Panel (a) shows the optimal single rate and panel (b) shows the optimal multi rate system. We express the rates per 10ml of ethanol. The optimal single rate is 27p/10ml ethanol when marginal drinking externalities are constant, and increases as we vary the externality function to make it increasingly convex (rising to 37p/10ml ethanol when 95% of externalities arise from the set of heavy drinkers). The reason the optimal rate is higher the more convex is the function is that heavy drinkers reduce their ethanol more *in levels* (though not in percent terms) than light drinkers do as a consequence of an increase in alcohol prices. Therefore, the greater the share of aggregate externalities that are accounted for by the heavy drinkers, the more effective is tax at lowering the social costs of alcohol consumption and therefore the higher is the optimal rate.

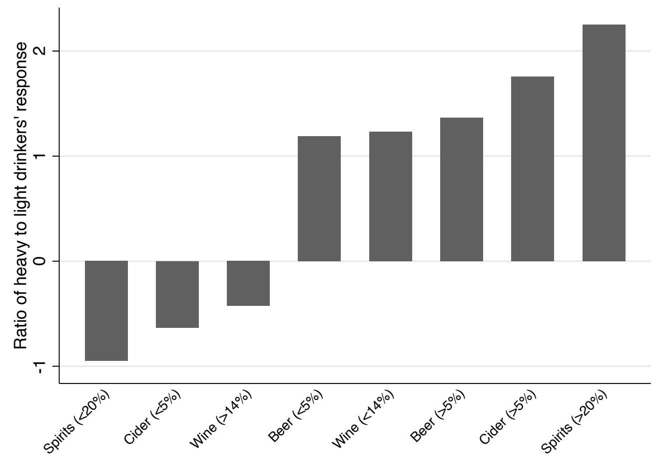
In the case of the multi rate system, all optimal rates are 27p/10ml ethanol when marginal externalities are constant. In this case there is no gain from rate differentiation and the optimal single and multi rate systems coincide. However, as the externality function becomes increasingly convex the optimal rates in the multi rate system diverge. High strength spirits (those with ABV>20%) attract the highest tax rate (over 40p/10ml ethanol when the heavy drinkers account for over 80% of externalities). Table wine (wine with ABV≤14%) attracts the next highest tax rate followed by beer and strong cider. The lowest rates apply to fortified wine (wine with ABV>14%), weak cider and weak spirits.

What drives the variation in tax rates?

The variation in optimal rates is driven by the correlation between how strongly a tax rate induces households to switch away from ethanol and their marginal externalities. If taxing an alcohol type induces heavy drinkers to switch strongly away from that alcohol type (without switching too strongly to alternative sources of ethanol), while leaving the decisions of light drinkers relatively unchanged, then that particular tax is effective at discouraging the most socially costly forms of drinking.

To illustrate empirically what drives the optimal tax results we compute the change in the quantity of ethanol demanded resulting from a 1% increase in the price of each of the eight alcohol types, doing this separately for the set of heavy and light drinkers. In Figure 5.2 we show the ratio of the change for the heavy drinkers to the change for the light drinkers. An increase in the price of strong spirits reduces the ethanol demand of the heavy drinkers by over twice as much as it reduces ethanol demand for the light drinkers. For beer, table wine and strong cider, a marginal increase in the price of products belonging to each of these alcohol types stimulates larger reductions in total ethanol from heavy drinkers compared with the light group by a factor of between 1.2 and 1.75. This is why the optimal tax rate on strong spirits is greater than those on beer, table wine and strong cider. For weak spirits, fortified wine and weak cider, a marginal increase in the price of products in each of these groups actually increases the ethanol demand of the set of heavy drinkers (though it lowers demand among the lighter drinkers) – which is why the bars for these alcohol types in Figure 5.2 are negative and why they are the alcohols with the lowest optimal rates. Raising the tax rate on these types of alcohols encourages the heavy drinkers to switch to alternative stronger alcohols, leading the optimal rates on these alcohol types to be relatively low.

Figure 5.2: *Response of heavy and light drinkers to increases in the price of different alcohol types*



Notes: For heavy and light drinkers we calculate the change in total ethanol demand as a result of a 1% increase in the price of the alcohol type shown on the horizontal axis. The bars show the ratio of the change for the heavy drinkers to the light drinkers. Light drinkers are those that on average buy below government recommended levels, and heavy drinkers are those that buy on average above recommended levels.

Differences in the impact on ethanol demand of a change in tax rate for a given alcohol type across light and heavy drinkers may be due to: (i) differences in the level of ethanol obtained from that alcohol type; (ii) the strength of switching away from it (i.e. the alcohol type own price effect); or (iii) differences in the propensity to switch to alternative alcohol types. In Table 5.1 we compare each of these between the set of light and heavy drinkers.

Columns (1) and (3) show the number of standard drinks per adult per week that the group of light and heavy drinkers get from each alcohol type; column (3) shows the ratio. Heavy drinkers tend to purchase more of each alcohol type. However, by the far the biggest discrepancy is for strong spirits – heavy drinkers get, on average, over 4 times as much ethanol from this source as the light drinkers. In columns (4) and (5) we show the own price effects for the different alcohol types (i.e. the % change in ethanol demanded from the type following a 1% increase in the price of all products of that type) for the light and heavy drinkers and in column (6) we show the ratio between the two groups. Strong spirits is one of only two alcohol types that the heavy drinkers have a lower (in absolute terms) elasticity compared with lighter drinkers. However, the much larger level of strong spirits demanded by the heavy drinkers means that they switch more strongly in level terms away from this alcohol type in response to a price rise. Finally in columns (7)–(9) we show the impact of changes in price for each alcohol type on overall ethanol demand, taking into account substitution towards other alcohols. Columns (7), for light, and (8), for heavy drinkers, show the percentage change in total ethanol that results from a 1% increase in the price of each alcohol type; column (9) shows the ratio. For all alcohol types the heavy drinkers switch away from ethanol less in percent terms than the light drinkers. This is because heavy drinkers are much more inclined than lighter drinkers to respond by switching to alternative alcohol types.

Table 5.1: *Demand responses by alcohol types*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				% reduction in ethanol demand following 1% increase in price of type, accounting for:					
	Number of standard drinks			Own price effect			Total effect		
	Light	Heavy	Ratio	Light	Heavy	Ratio	Light	Heavy	Ratio
Spirits (<20%)	0.15	0.18	1.1	-2.83	-3.08	1.1	-0.04	0.02	-0.4
Cider (<5%)	0.18	0.32	1.8	-1.66	-2.03	1.2	-0.02	0.01	-0.2
Wine (>14%)	0.16	0.37	2.3	-2.41	-2.54	1.1	-0.04	0.01	-0.2
Beer (<5%)	0.97	2.02	2.1	-2.92	-2.78	0.9	-0.42	-0.19	0.5
Wine (<14%)	1.65	3.43	2.1	-2.16	-2.55	1.2	-0.48	-0.23	0.5
Beer (>5%)	0.23	0.59	2.5	-3.17	-3.36	1.1	-0.08	-0.04	0.5
Cider (>5%)	0.25	0.85	3.4	-2.31	-1.94	0.8	-0.08	-0.05	0.7
Spirits (>20%)	0.91	3.89	4.3	-4.27	-2.97	0.7	-0.68	-0.59	0.9

Notes: Columns (1) and (2) show the number of standard drinks from each alcohol type for light and heavy drinkers respectively. Column (3) shows the ratio of column (2) to (1). We simulate a 1% increase in the price of each alcohol type and calculate the % change in ethanol demanded from that type (shown in columns (4) and (5)) and the % change in total ethanol demanded from all types (shown in columns (7) and (8)). Column (6) shows the ratio of column (5) to (4), and column (9) shows the ratio of column (8) to (7). Light drinkers are those that on average buy below government recommended levels, and heavy drinkers are those that buy on average above recommended levels.

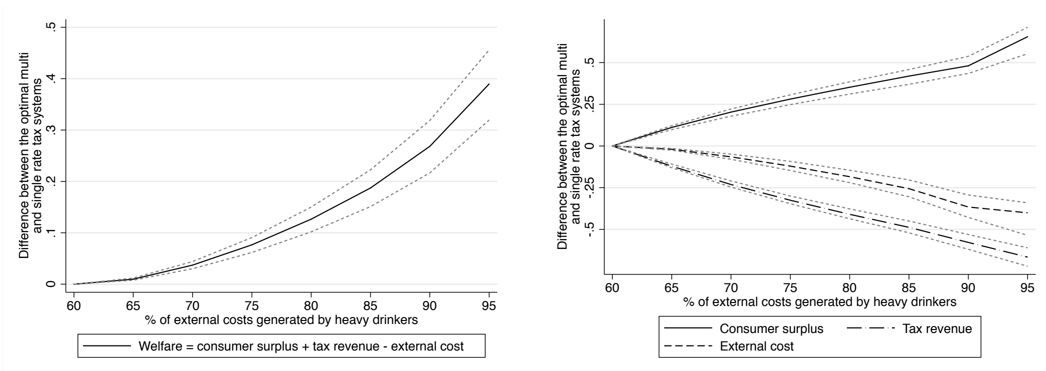
Hence, the relatively high tax rate on strong spirits is driven by three factors: (i) heavy drinkers get a large share of their ethanol from these products; (ii) they are reasonably price sensitive with respect to these products (though not as much as light drinkers); (iii) although they tend to switch to alternative alcohol types to a much greater extent than lighter drinkers, alternatives to strong spirits tend to contain much less alcohol. Therefore, taxing strong spirits at a relatively high tax rate is an effective way to reduce the ethanol purchased by the heavy drinking group without imposing large costs on lighter drinkers.

5.3 Welfare

In Figure 5.3 we show how social welfare under the optimal multi rate system differs from welfare under the optimal single rate system, and how this difference varies with the degree of convexity of the externality function. Panel (a) shows the impact on total social welfare (the sum of consumer surplus and tax revenue minus external costs) and panel (b) shows the impact on the constituent parts – consumer surplus, tax revenue and the external costs of drinking. When 60% of externalities are generated by the set of heavy drinkers (who, recall, consume 60% of ethanol) there is no difference in the optimal multi and single rate systems.

Figure 5.3: *Comparison of welfare under optimal single and multi rate tax systems*

(a) Total welfare (b) Welfare components



Notes: Differences are measured in £billion per year. Light drinkers are those that on average buy below government recommended levels, and heavy drinkers are those that buy on average above recommended levels. 95% confidence intervals are shown in grey.

The welfare gain from moving from a single to multi rate system becomes increasingly large as we increase the degree of convexity of the externality function. If the heavy drinkers generate 95% of the external costs, the optimal multi rate system increases welfare by over £350 million, relative to an optimally set single rate system. To put this number into context, the UK Cabinet Office estimate that around £370 million is incurred by alcohol related accident and emergency visits (Cnossen (2007)).[17]

The welfare gain is driven both by lower external costs and higher consumer surplus. The flexibility afforded by the optimal multi rate system enables the planner to target the consumption of the heaviest drinkers by raising the tax rate on strong spirits and lowering the tax rates on other forms of alcohol below the the optimal single rate. This means the multi rate system is able to more effectively reduce ethanol consumption among the heaviest drinkers, while actually reducing the average tax rate on alcohol relative to the optimal single rate. As a consequence, on aggregate, consumer surplus rises even as the total external costs from drinking are brought down. However, as a result, the optimal multi rate system raises less tax revenue than the single rate system and this revenue loss offsets, to some extent, the welfare gains arising from higher consumer surplus and lower externalities.

We assume that the planner maximises a utilitarian social welfare function, and do not take stand on the incidence of the external costs of drinking on heavy or light drinkers. This means that we can only state that social welfare is higher under both the optimal systems vis-a-vis a no tax world, not whether heavy or light drinkers are better or worse off as a result of the optimal tax systems. However, we can analyse how the consumer surplus for different groups varies under the optimal tax systems.

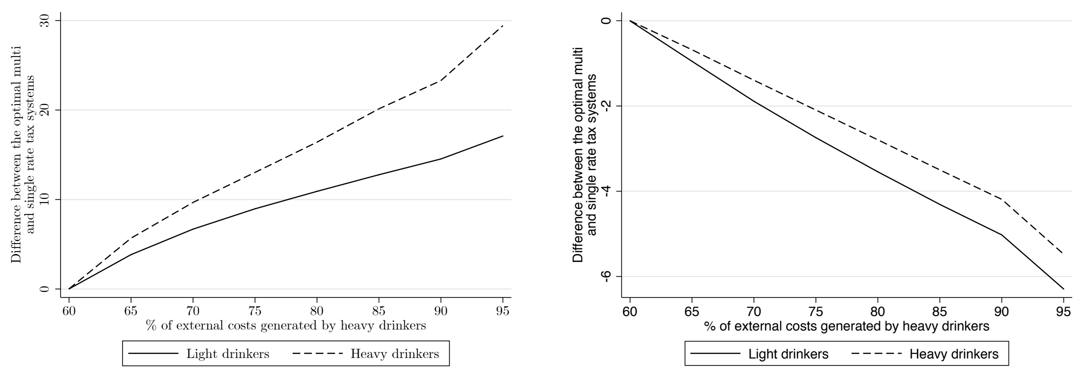
The gains in consumer surplus associated with having an optimal multi rather than single rate system vary across households. In Figure 5.4 we show (panel (a)) the average consumer surplus changes for heavy and light drinkers (light drinkers are those that on average buy below government recommended levels, and heavy drinkers are those that buy on average above recommended levels) and how this varies with the convexity of the externality function. For any strictly convex function both groups have, on average, higher consumer surplus under the multi rate system. Panel (b) shows that the reason for this is that for both light and heavy drinkers, the average tax rate they face for their alcohol is lower under the multi than single rate system. Although the light drinkers see the largest reduction in their average tax rate, their consumer welfare gain in £ terms is lower than for the heavy drinkers. This is because heavy drinkers purchase more ethanol than light drinkers – as a fraction of alcohol expenditure their consumer surplus gain is smaller than for the light drinkers.

[17]The number reported in Cnossen (2007) is €447 million. We convert this to pounds and uprate to 2011 prices using the RPI.

In Online Appendix B, we show how other aspects of alcohol purchasing change under the optimal multi versus single rate system. Heavy drinkers buy fewer standard drinks under the optimal multi rate system, and the difference in purchases relative to the single rate system is broadly constant in the convexity of the externality function. Conversely, light drinkers buy more standard drinks under the optimal multi rate system compared with the single rate system. Heavy drinkers, conditional on buying alcohol, switch much more strongly away from high ABV alcohol under the multi rate system than under the single rate system. Light and heavy drinkers, conditional on purchasing alcohol, actually buy larger packs, and also switch less strongly to the outside option under the optimal multi rate system.

Figure 5.4: *Welfare and alcohol purchases under optimal single and multi rate tax systems for heavy and light drinkers*

(a) Consumer surplus (b) Average tax rate faced



Notes: Difference in consumer surplus measured in £ per household per year. Difference in tax rate is p/10ml ethanol. Light drinkers are those that on average buy below government recommended levels, and heavy drinkers are those that buy on average above recommended levels.

The flexibility of the multi rate tax system (relative to a single ethanol tax rate) creates welfare gains through achieving higher average alcohol taxes for heavy relative to light drinkers, thus focusing on lowering the consumption of high ethanol products by heavy drinkers while leaving relatively less affected the consumption of light drinkers. The optimal single tax rate on ethanol prescribes a relatively high tax on all alcohol, while the multi rate system can focus much more on reducing spirits consumption among the very heaviest drinkers. This lowers the social costs of drinking while also achieving consumer surplus gains for the majority of households.

Comparison to first best

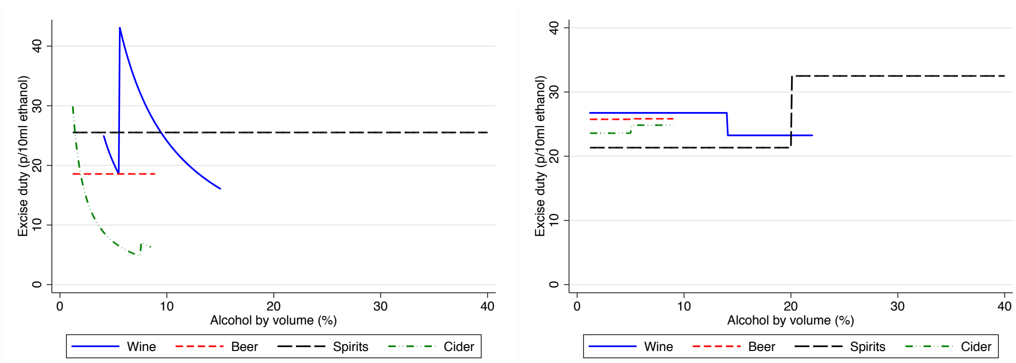
An important feature of the optimal tax systems depicted in Figure 5.1 is that the planner is restricted to set tax rates that are common across households. If the planner could set household specific ethanol tax rates, it could achieve the first best by setting these such that, for each household, they equal the marginal externality associated with that household's ethanol consumption (at that tax rate). In practice, achieving perfectly discriminating ethanol tax rates is (currently) beyond the bounds of feasibility, but it represent a useful benchmark, and allows us to measure the loss in welfare that results from the restriction that tax rates are common across households. For the calibration in which 80% of the external costs of drinking are generated by the set of heavy drinkers, the optimal set of household specific tax rates imply an increase in welfare of £1.23 billion over the optimal multi rate system. Moving from the optimal single rate system to the optimal multi rate system closes 10% of the gap between the single rate system and first best.

Comparison to the UK system

Our primary focus is on how exploiting correlation between heterogeneous marginal externalities and demands can lead to significant welfare gains from varying tax rates across different forms of ethanol. However, we can also use our framework to assess how close the UK tax system gets to an optimal tax system. In the UK there are different tax rates levied on beer, spirits, wine and cider with some variation in rates across different ABV contents. However, both wine and cider are taxed per litre of product (rather than per amount of ethanol), and the system is not coherently targeted at consumers that generate large marginal externalities. In Online Appendix B we provide more details on alcohol taxes in the US and EU. Alcohol in the US is taxed per litre of product, and not per unit of ethanol. This means that the schedules have a similar downward slope, with (within alcohol type) stronger alcohol taxed more lightly per unit of ethanol.

Figure 5.5: *Comparison of the current UK system with the optimal multi rate system*

(a) Current UK system (b) Optimal UK system



Notes: The optimal rates are shown for the calibration in which 80% of the external costs of drinking are generated by heavy drinkers. Heavy drinkers are those that buy on average above recommended levels. The UK tax rates are those in place in 2011.

Figure 5.5 compares the optimal tax rates for the calibration in which 80% of the external costs of drinking are generated by heavy drinkers, and the UK tax system in 2011. In the UK, in addition to alcohol excise duty, there is a broad based Value Added Tax (VAT) levied at the rate of 20%. To make our optimal taxes comparable to the UK duty component, in Figure 5.5 we divide them by 1.2.[18] Panel (a) shows the current UK tax system, and panel (b) shows the optimal multi rate system. It is clear that the current system is far from optimal: significant welfare gains could be achieved from: (i) levying taxes on ethanol rather than on volume, (ii) increasing the tax rate on cider, (iii) reducing the tax rate on spirits below 20% ABV, and increasing the rate on spirits above 20% ABV.

Table 5.2: *Comparison of welfare under the optimal rates with UK system*

	Difference with UK system for optimal:	
£bill per year unless stated	Single rate	Multi rate
Consumer surplus	-1.55	-1.20
of light drinkers (total)	-0.63	-0.40
of light drinkers (£per household)	*-32.38*	*-20.94*
of heavy drinkers (total)	-0.92	-0.79
of heavy drinkers (£per household)	*-138.23*	*-118.58*
External cost	-2.70	-2.88
Tax revenue	-0.04	-0.45
Social welfare	1.11	1.24

Notes: Differences are measured in £billion per year, unless stated. The numbers are shown for the calibration of the externality function under which the heavy drinkers generate 80% of the external costs of drinking. Heavy drinkers are those that buy on average above recommended levels.

This size of the welfare gains depends both on the aggregate social costs of drinking and how concentrated they are among heavy drinkers. Under the assumption that the aggregate social costs are £7.25 billion (from Cnossen (2007)), if the heaviest 19% of drinkers account for 80% of the externalities the average tax rates under the optimal and UK systems are similar. Thus, the better performance of the optimal system is not driven by the calibration of the aggregate externality to £7.25 billion. The inferior performance of the UK system is partly driven by not levying tax per unit of ethanol: the welfare gain from moving from the UK to a single rate system would be £1.24 billion, see Table 5.2. The value from optimally differentiating rates across alcohol types can be seen from the move from the optimal single

[18] Our optimal tax estimates are for the total tax levied on alcohol. With a VAT tax of 20%, the optimal alcohol duty rates equal our estimates divided by 1.2.

to optimal multi rate system, which results in a further 10% improvement. Adopting either of the optimal systems would involve reductions in consumer surplus and tax revenue, but these would be more than made up for by reducing drinking externalities. Although these precise numbers depend on the level and degree of convexity of the externality function, the fact that simple improvements to the UK tax system could yield substantial welfare gains holds generally.

6 Summary and conclusions

In this paper we consider corrective tax design in markets in which an externality generating commodity is available in many products and consumers are potentially heterogeneous in both the externalities that their consumption creates and in their demands for different products. We focus on the alcohol market. There is much evidence that consumption of ethanol (which is available in many products bundled together with other attributes) is associated with externalities and these externalities are non-linear. We consider varying degrees of non-linearity in this relationship, ranging from marginal externalities that are constant in ethanol consumption to the external costs of alcohol being disproportionally generated by a small group of heavy drinkers.

Our results show that varying tax rates across different forms of alcohol can lead to significant welfare gains relative to an optimally set single ethanol tax rate. Optimally varying rates exploits correlations in households' preferences for product attributes (and hence product level demands) with their total demand for ethanol, enabling the tax system to better target consumption that generates high externalities. Welfare gains are larger the more convex externalities are in ethanol consumption. To implement our optimal tax framework empirically we estimate a demand model using detailed longitudinal micro data on households' alcohol purchases.

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, specific groups may be more prone to generate externalities (including on their future self) – for instance, there are particular concerns surrounding children's sugar consumption. 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

Aguiar, M. and E. Hurst (2007). Life-cycle prices and production. *American Economic Review 97*(5), 1533–1559.

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.

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.

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.

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

Cnossen, S. (2007). Alcohol taxation and regulation in the European Union. *International Tax and Public Finance 14*(6), 699–732.

Competition Commission (2000). *Supermarkets: A Report on the Supply of Groceries from Multiple Stores in the United Kingdom.*

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.

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

Dubois, P., R. Griffith, and A. Nevo (2014). Do Prices and Attributes Explain International Differences in Food Purchases? *American Economic Review 104*(3), 832–867.

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.

GBD 2016 Alcohol Collaborators (2018). Alcohol use and burden for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet 392*(10152), 1015–1035.

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.

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

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.

Kopczuk, W. (2003). A note on optimal taxation in the presence of externalities. *Economics Letters 80*(1), 81–86.

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

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.

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 (2017). One Markup to Rule Them All: Taxation by Liquor Pricing Regulation. Technical report, National Bureau of Economic Research.

Miravete, E. J., K. Seim, and J. Thurk (2018, April). Market Power and the Laffer Curve. *Econometrica.*

Nilsson, J. P. (2017). Alcohol Availability, Prenatal Conditions, and Long-Term Economic Outcomes. *Journal of Political Economy 125*(4), 1149–1207.

Petrin, A. and K. Train (2010). A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research 47*(1), 3–13.

Ramsey, F. P. (1927). A Contribution to the Theory of Taxation. *The Economic Journal 37*(145), 47–61.

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.

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.

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.

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