Journal Pre-proof

Banning volume discounts to curb excessive consumption: A cautionary tale

Farasat A.S. Bokhari, Paul W. Dobson, Marcello Morciano, Marc Suhrcke

PII: S0014-2921(23)00109-5
DOI: https://doi.org/10.1016/j.euroecorev.2023.104480
Reference: EER 104480

To appear in: *European Economic Review*

Received date : 2 December 2021
Revised date : 16 April 2023
Accepted date : 7 May 2023

Please cite this article as: F.A.S. Bokhari, P.W. Dobson, M. Morciano et al., Banning volume discounts to curb excessive consumption: A cautionary tale. *European Economic Review* (2023), doi: https://doi.org/10.1016/j.euroecorev.2023.104480.

Banning volume discounts to curb excessive consumption: A cautionary tale

Farasat A.S. Bokhari [a], Paul W. Dobson [b], Marcello Morciano [c], and Marc Suhrcke [d,e]

[a] *School of Economics, Centre for Competition Policy, University of East Anglia*

[b] *Norwich Business School, Centre for Competition Policy, University of East Anglia*

[c] *Marco Biagi Department of Economics, Università degli Studi di Modena e Reggio Emilia, and Research Centre for the Analysis of Public Policies*

[d] *Centre for Health Economics, University of York*

[e] *Luxembourg Institute of Socio-economic Research, Luxembourg*

April 16, 2023

Abstract

Volume discounts encourage consumers to buy more. Banning such discounts should then lead to consumers buying less. This is the thinking behind banning multiple-unit discounts, including multibuy price promotions, to curb excessive harmful consumption of alcohol and high-fat, -sugar, and -salt (HFSS) foods. However, our analysis questions the validity of this thinking, which ignores the possible restraining effect of volume discounts. We find that such a ban for retailing alcohol in Scotland increased rather than reduced sales. Retailers switched to using more straight (single-unit) discounts, which encouraged high-consumption households to increase their shopping frequency and buy more.

Key words: Volume discounts, excessive consumption, multibuy, alcohol

JEL Classification: C54, D04, H23, I12, I18, L81

1. Introduction

Volume discounts are a ubiquitous means to encourage consumers to buy more. They are attractive to firms as a smart means of indirect (second-degree) price discrimination, where customers segment by self-selection when choosing from a menu of price-quantity combinations or a nonlinear pricing schedule to obtain more for less cost per unit (Adams and Yellen, 1976, Wilson, 1993). They can also be efficiency-enhancing when they boost overall sales, help firms achieve scale economies, and intensify competition (Armstrong, 2016). However, governments might not always want consumers to buy more and instead prefer them to buy less when this helps avoid harmful excessive consumption. Such concerns arise with alcohol and unhealthy foods, where moderate consumption may not pose a major health risk, but excessive consumption can seriously harm individuals and impose costs on society. In these circumstances, would banning volume discounts help curb consumption or backfire by driving firms to find alternative means to sell perhaps even more, albeit at lower margins?

This paper considers this question by examining business responses and changes in consumer behavior in the wake of a ban on volume discounts for purchasing alcohol in multiple units. The World Health Organization advocates such bans (WHO, 2010). Countries with volume discount bans on alcohol sales include Canada, Sweden, Finland, Switzerland, Ireland, and Scotland – with the latter being the focus of this paper. Likewise, there are calls to ban similar volume offers for high-fat, -sugar, and -salt (HFSS) foods and drinks, which are viewed as drivers of excess purchases that help fuel obesity (DHSC, 2021). Accordingly, understanding the consequences of banning volume discounts is an important policy consideration in seeing how well-targeted and effective is such an intervention in these critical public health contexts.

In examining the consequences of such a ban, we exploit a (quasi-) natural experiment in Great Britain (GB). Scotland used its devolved powers to introduce a ban on volume discounts, in the form of multiple-unit discounts known as multibuys, for retail alcohol sales starting in October 2011, while the UK government proposed but did not enact the same policy for England and Wales, having applied different impact assessment approaches (Scottish Parliament, 2010, Home Office, 2012). Yet, all three countries share the same British tax system and other policies, along with similar population characteristics and alcohol demand and supply patterns. We use tobit models within a difference-in-differences (DD) framework to compare weekly household-level purchases before and after the introduction of the ban for households in Scotland with those in England and Wales. We distinguish between low-, medium-, and high-consumption households, and examine the impact of the ban for each household type on total alcohol sales as well as by segments of the market separately (spirits, beers, and wines).

We find that the ban had the greatest impact on the primary targets of the ban, i.e. beer and wine sales, where multiple-unit discounts were common and especially on purchases made by heavy consumption households, who made extensive use of these discounts. However, the effects were the exact opposite of those intended. Sales of beers and wines increased, driven primarily by heavy consumption households buying more, not less. With the top third of consumption households accounting for almost three-quarters of alcohol purchases, we find that their extra purchases increased the overall amount of alcohol sold.

We investigate the factors behind this quantity increase and the apparent policy failure. Two key insights emerge. First, retailers responded by replacing multiple-unit discounts, especially in the form of multibuy price promotions (like "buy 6 save 25%" deals), with more straight (single-unit) price reductions. Second, while the ban had limited or no impact on the purchasing patterns of low or moderate purchasers, heavy purchasers responded by increasing the number of shopping trips per week, with the more frequent buying resulting in higher overall quantity.

The former finding is perhaps not surprising when retailers can draw on alternative forms of price promotions to counter and circumvent the effects of a volume discount ban, but the latter finding is somewhat puzzling. If multiple-unit discounts were serving as effective second-degree price discrimination and the ban released the incentive compatibility constraints, we might have expected those consumers not previously bulk buying (i.e. predominantly low- and moderate-consumption households) to buy more as retailers increased the use of straight discounts. However, we find little change for these households, but instead the increased quantity is due to high-consumption households.

Even so, the intriguing aspect is the increased shopping frequency, since the replacement offers would have allowed high-consumption households to continue buying in a single store visit, thereby avoiding additional shopping costs and hassle from more frequent shopping trips. Budgeting to spread out expenses might be a reason, but we find no relationship based on income differences, and this alone does not explain the additional purchases. Instead, it appears that non-linear pricing was restraining heavy purchasers, committed to buying in bulk to obtain the volume discounts, but once the ban came into effect, then straight discounts removed that constraint, opening up the temptation to make additional store visits for further purchases.

Two behavioral economics explanations fit this pattern. First, additional visits and purchases on discounted prices may provide additional transaction utility and segregate perceived gains (Thaler, 1985). Second, multiple unit discounts may facilitate commitment to buying only in bulk as a self-control device to space out purchases and to ration consumption, whereas straight discounts allow for top-up purchases anytime, hence undermining that commitment (Thaler and Shefrin, 1981, Hoch and Loewenstein, 1991). The commitment aspect is consistent with heavy drinkers being prone to time-inconsistent behavior and waiting impulsivity (Mayhew et al., 2020). It is also in line with how sales restrictions, such as limiting store opening hours, may work as a self-control commitment device, helping to reduce store visits and total alcohol purchases (Hinnosaar, 2016), while also helping to curb binge drinking leading to hospitalization (Marcus and Siedler, 2015).

Our findings run counter to two early analyses of the aggregate effect of the ban, where Robinson et al. (2014) found a small sales decrease, while Nakamura et al. (2014) found no overall effect. However, both studies potentially suffer from aggregation biases, do not account for price endogeneity, and have several econometric limitations that we overcome in this paper.

Unlike our household data, Robinson et al. (2014) use Nielsen sales data aggregated at the week-country level. They employ interrupted time series analysis to compare sales in Scotland with England and Wales. Thus, they ignore the heterogeneity in policy responses across different types of households. They also interpolate for several other covariates that are not available at weekly levels, such as weekly population estimates but using differences between UK values and Scotland (yet the sales data is for GB not UK). In contrast, while Nakamura et al. (2014) keep observations at the household level, they ignore the variation in sales over time and collapse all weekly household observations into two single observations: pre-ban total sales and post-ban total sales per household. We overcome both these problems and use weekly observations for each household, allowing zero purchases within any given week in our non-linear DD framework. Similarly, while Nakamura et al. (2014) ignore changes in prices by the retailers and have a potential omitted variable bias in their estimates, Robinson et al. (2014) include prices but treat them as exogenous. In contrast, we do not omit the prices and use control functions to account for endogeneity. Thus, we believe that a careful reanalysis of the policy is warranted, which is what we provide with our new estimates in this paper.

This paper contributes to the literature and ongoing policy debates about alcohol affordability and the wider challenge of curbing excessive consumption of alcohol and unhealthy foods. Evaluating policies affecting alcohol affordability predominantly focuses on taxation effects, finding significant variability in pass-through rates (Nelson and Moran, 2019, Hindriks and Serse, 2019), with no guarantee that the tax will be fully passed on for cheap alcohol products where excessive consumption is most concerning (Ally et al., 2014, Wilson et al., 2021). Existing evidence also underlines the need for potentially complex rate setting for optimal tax design (Griffith, O'Connell and Smith, 2019). Alternatively, regulated state control of prices or imposing minimum prices to prevent discounting provides a more assured way of maintaining high prices, but at the same time, it remains highly contentious as an overtly interventionist policy that curtails competition and potentially promotes inefficiency (Miravete, Seim and Thurk, 2018, 2020, Calcott, 2019, Griffith, O'Connell and Smith, 2022, Conlon and Rao, 2019, 2020). Instead, this paper considers a ban on volume discounts as a less restrictive partial price regulation in an otherwise openly competitive market that may be politically more palatable than the complexity or regressivity of targeted sin taxes or the inflexibility and inflationary nature of imposing minimum prices.

2. Background

Drinking habits have shifted in Britain over the past two decades, with a sharp decline in alcohol consumed at licensed premises but a corresponding sharp increase in alcohol purchased through retailers for consumption at home. Consequently, alcohol policy concerns have shifted from on-premises consumption, e.g., via limiting opening hours to curb binge drinking, to the affordability of cheap alcohol combined with retail price promotions spurring hazardous or harmful consumption at home. The Scottish government has been in the vanguard of policy initiatives to tackle the increase in alcohol consumption at home by seeking to impose a minimum unit price on alcohol and banning volume discounts (Scottish Parliament, 2010). The same policy measures were proposed by the UK government for application in England and Wales, but ultimately not pursued after a lengthy consultation process (Home Office, 2013).

Political and legal challenges prevented the Scottish government from introducing minimum unit pricing until 2018 (Woodhouse, 2020). However, Scotland proceeded in 2011 to introduce a ban on volume discounts for off-trade retailers selling alcohol (which includes supermarkets, off-licenses, and convenience stores selling alcohol for consumption off the premises) as part of the

Alcohol etc. (Scotland) Act 2010.[1] The ban relates to multiple-unit discounts, specifically banning quantity discounts for multipacks and multibuys, where units are purchased as a collection rather than separately. Note that the ban does not cover different product sizes, such as requiring the unit price of a 70cl bottle of spirits not to be less than a 35cl bottle of the same brand. Furthermore, while multipacks are common for beer, they are less relevant to other alcohol categories, and even for beer, the effect of the ban is muted because retailers rarely sell individual cans or bottles of beer of the same unit size that also go into multipacks. Instead, the main target of the ban is on multibuys, which feature extensively in sales of beer, cider, wine, and flavored alcoholic beverages (FABs), but less so for spirits.

Both types of multiple-unit discounts operate as mixed bundling but differ in that multipacks are units physically packaged together, whereas multibuys exist as virtual packages, with the discount applied on individual items bought together. Multipacks are long-established for bulk buying consumer-packaged goods, while multibuys have grown in prominence as price promotions, typically framed as 'buy/get' ('X + N free') offers, like 'buy one get one free' and '3 for 2', or deals that state a fixed price ('X for $Y') or saving on multiple units ('buy X and save Y%').

[1]The Act also provided for other supporting measures, including restricting the location of drinks promotions to within a single area of the store, the requirement of an age verification policy, powers to introduce a social responsibility levy on license holders, and a requirement for Health Boards to be notified of the license applications of premises in their geographical area.

2.1. **Scotland vs. England and Wales – Institutional details.** There are strong similarities across the constituent countries of GB for demographic profiles, income levels, culture, education levels, and economic conditions. Alcohol demand and supply are also very similar. Accordingly, we might expect responses to an alcohol consumption policy to be very similar between these countries. Hence, they may provide a good comparative setting for drawing inferences when there are policy differences. Indeed, this is how the respective governments tend to evaluate their policy measures (Giles and Richardson, 2020, Woodhouse, 2020).

There is also a large body of academic research evaluating alcohol policy measures based explicitly on comparisons for Scotland with England and Wales. For example, Nakamura et al. (2014), Robinson et al. (2014, 2018), and this paper make such comparisons in evaluating the ban on multiple-unit discounts, while O'Donnell et al. (2019), Xhurxhi (2020), Anderson et al. (2021), Robinson et al. (2021), Griffith, O'Connell and Smith (2022) and Vandoros and Kawachi (2022) use these cross-country comparisons for evaluating minimum unit pricing, and similarly Cooper, Gehrsitz and McIntyre (2020) and Francesconi and James (2021) for evaluating the impact of changed drink-driving limits, and Green, Heywood and Navarro (2014) on liberalizing opening hours.

Nevertheless, there are important institutional and market features to consider. First, licensing distinguishes between "on-trade" sales for on-premises consumption within public hospitality venues and "off-trade" sales through retailers for off-premises consumption. The two segments serve as broad substitutes but appear distinct in the character of their demand and have exhibited opposing sales trends over time. On-trade prices are also on average more than three times higher than corresponding off-trade prices (Giles and Richardson, 2020). Given this considerable price gap and the different sales trends, with off-trade accounting for two-thirds of consumption, we would not expect the modest price effect from banning multiple-unit discounts applied to off-trade sales to materially impact on-trade sales. This is in line with the UK competition authorities viewing the on-trade and off-trade as constituting separate markets (Griffith, O'Connell and Smith, 2022, Appendix A.3). This is also supported by the annual sales trends reported by Giles and Richardson (2020), showing that on-trade sales in Scotland have continued in lockstep decline with on-trade sales in England and Wales – not perceptively influenced by the off-trade price regulation measures introduced over the past decade.

Second, in terms of consumption patterns, the volume of pure alcohol sold per adult in Scotland has historically been higher than in England and Wales. However, the volume gap has narrowed over the past twenty years (Giles and Richardson, 2020). There are differences in the composition of sales – more spirits but less beer are consumed per adult in Scotland compared to England and Wales – but both the long-term trends and weekly sales patterns are very similar (Giles, Robinson and Beeston, 2019). Furthermore, while alcohol prices (per unit of alcohol sold) have been increasing over time, especially in the on-trade, the levels and trends are remarkably similar between Scotland and England & Wales. The close price comparability reflects similar economic conditions across the three countries, the same tax rates, and the dominance of chain operators using national uniform pricing, especially in the off-trade.

3. Data and Methods

3.1. **Sample and Variables.** Our main data are drawn from the Kantar WorldPanel database, which contains repeated information on purchases from grocery stores by a representative sample of households from Scotland, England, and Wales. Each participating household uses a handheld scanner to record take-home purchases. For each product purchased in a given transaction, the data include the quantity purchased and transaction prices, together with information on the type of promotion (if any), the identity of the store/chain where it was purchased, and the date of purchase. For each product, we also know its exact identity (via a unique product number) and manufacturer information along with physical characteristics such as type of package, number of units in the pack, size, and strength of each unit (e.g., Carlsberg lager beer, 4 cans pack, 500 ml with 5% ABV), and selected nutrient values associated with each unit (calories, sugar, proteins, carbohydrates, fat, saturated fat, fibers, sodium, and an overall British Food Standard Agency

(FSA) nutrient profiling score).[2] With each transaction, we have a household ID, which is linked to a companion dataset on household socio-demographics that includes household size, social and economic status, and main adult shopper information on age, education, and ethnic status. Importantly, the geographic location of the household is also available at 4-digit postcode level (e.g., NR31).

We measured the aggregate volume of ethanol purchased by a household per week by multiplying the aggregate volume of alcohol purchased by its strength (ABV), divided by 1,000. The advantage of this approach is that it standardizes for differences in strength across products. Moreover, it is equivalent to the 'units of alcohol' (10ml of pure ethanol) measure used in the UK and in EU countries for measuring ethanol volume. Units of alcohol purchased per week were further divided by the number of adults in the household and log-transformed to account for the skewness of the data. Thus, we have measures of units per adult per week for all alcoholic products combined (S00 - All) and by four alcohol segments: Spirits and Fortified Wines (S01 - Spirits for short: mean ABV 30.15%); Ales, Lagers, and Ciders (S02 - Beers for short: mean ABV 4.83%); Wines and Sparkling Wines (S03 - Wines for short: mean ABV

[2] Strength is measured as a percentage of alcohol-by-volume (ABV, the number of milliliters of pure ethanol present in 100 ml of solution at 20 degree Celsius). For products with missing ABV, we performed online searches to impute their values. We do not observe any values for fat, saturated fat, or fibers for spirits.

12.11%) and Flavored Alcoholic Beverages (S04 FABs for short: mean ABV 5.68%).

For each transaction, we observe the list price of a given item (pack/bottle etc.), the associated promotion code (if any), and the total amount paid after promotion for the bundle or singleton of alcohol purchased. Thus, we compute the price per unit of alcohol as the total expenditure paid after promotions, divided by the total units of alcohol purchased. We also use the information on the list prices to compute the associated discount per unit of ethanol as the difference between the list and transaction prices of the bundle. For each bundle, we also compute the values of other characteristics (including calories, sugar, etc., as mentioned earlier) as the share weighted average of individual items in the bundle.[3] Similarly, we compute overall (S00) and segment-specific (S01-S04) prices per unit and discounts. In a given week, a household might not purchase any alcoholic product, so the quantity variable is zero and prices are missing. However, rather than discard the observation, we assign a weekly price that corresponds to the average weekly price paid by other households for that segment in the same household group and region of the UK (for 15 regions: 1 for Wales, 9 for England, and 5 for Scotland). The same holds true for the discount and other product characteristics listed above.

[3] For instance, if a household purchases four beers and a bottle of wine, we compute the share of expenditure on each item and then use these weights to compute the mean value of calories per unit of alcohol.

In our sample, observations are over 83 weeks spanning from Jan/1/2011 to Jul/31/2012, and include only those households that made any purchase of alcohol during this period. The multibuy ban started on Oct/1/2011, which corresponds to week 41. We focused on only those households that were continuously enrolled during this period. We further restricted the analysis to households that purchased at least the equivalent of 5.5 British pints of typical beer (4.5% ABV) or more over weeks 2-12 inclusive (where we omitted the first week of January as the new year is celebrated more in Scotland). This is equivalent to 2 pints of beer per adult per month.[4] This restriction discards households that purchase alcohol sporadically and are not of concern from a policy viewpoint. (Over the 11-week period, excluded households purchased 3.61% of total alcohol.) We also discard households that lived within 35km of the Scottish-English border so as not to contaminate the analysis by those who can easily engage in cross-border purchases.

We grouped the remaining households into country-specific tertiles (HH-type = low, medium, or high) of per-adult alcohol purchase over the first 11-week period. The final consumption pattern and the number of households per group and country are given in Table 1. This 11-week period was used only for classifying households in tertiles and was omitted from the main analysis.

[4]Thus we required that the total household purchase per adult over weeks 2-12 be more than 14.113 units of alcohol. A British pint is 568ml and a typical beer is 4.5% ABV, which is 25.56ml of ethanol. A unit of alcohol is 10ml of ethanol, and so one typical pint of beer is 2.556 units of alcohol.

TABLE 1. Household consumption patterns (weeks 2-12)

HH-type	England & Wales	Scotland	Total
	Households (#)		
Low	2,565	220	2,785
Medium	2,568	225	2,793
High	2,568	230	2,798
Total	7,701	675	8,376
	Consumption (%)		
Low	7.59	6.82	7.53
Medium	18.75	18.50	18.73
High	73.65	74.68	73.74

Percentage based on total (per adult) purchase.

By construction, each household group has 1/3 of the total observations per country. During the first 11 weeks, the HH-type=low were responsible for 7.5% of all alcohol purchases, while the HH-type=high purchased about 73.7% of the total. These patterns are somewhat similar in England and Wales (EW) vs. Scotland. This skewed pattern, where a third of the households are responsible for almost three quarters of all purchases, is consistent over the entire observational period used in the main analysis (weeks 13-83).

3.2. **Empirical Specification.** We used the difference-in-differences (DD) methodology and compared household-level alcohol purchase patterns before and after the introduction of the ban in Scotland to those in England and Wales. We did so in the context of a panel setting where we observed each household for 71 weeks (weeks 13-83 inclusive) and where the household may or may not purchase any alcoholic product during a given week.

Let y^*_{it} be the latent variable that represents the (log of) quantity purchased per adult in household i in week t (henceforth, we use the terms consumption, purchase, and quantity interchangeably and assume no stockpiling) and is given by

$$y^*_{it} = \beta_1 S_i + \beta_2 B_{it} + \beta_3 (S_i B_{it}) + \boldsymbol{x}'_{4it}\boldsymbol{\beta}_4 + \boldsymbol{x}'_{5it}\boldsymbol{\beta}_5 + \boldsymbol{\tau}'_{it}\boldsymbol{\beta}_6 + \epsilon_{it}. \tag{1}$$

In the equation above, S_i is an indicator variable set to one if the household is located in Scotland and zero otherwise (i.e, when a household is from England or Wales). Similarly, B_{it} is also an indicator variable set to one in the post-ban period (week 41 onwards). We assume that ϵ_{it} is a mean zero standard normal error term (while allowing for observations to be correlated over time for a given household), and so the latent variable has the same distribution as ϵ_{it}. We observe the latent variable, if the value is greater than zero, so $y_{it} = \max\{0, y^*_{it}\}$. Accordingly, we estimated random effects tobit models, where the dependent variable was the log of quantity purchased per adult in a household (to avoid taking log of zeros, we added one to the quantity before logging).[5]

[5] In linear DD models, an identifying assumption is that the time effect is constant across groups and the group effect is constant across time. In turn, the treatment effect is constant across the treated population and allows for constructing a counterfactual. In contrast, in non-linear models such as the tobit, the treatment effect is not constant across the treated population, and hence identification is not straightforward (Athey and Imbens, 2006). Furthermore, as mentioned by Ai and Norton (2003), even if the interaction term coefficient (which is the term of primary interest) is zero, the cross difference/derivative term is generally nonzero. However, as shown by Puhani

The vectors $\boldsymbol{x}'_{4it}$ and $\boldsymbol{x}'_{5it}$ are additional covariates assumed to influence purchase decisions. The vector $\boldsymbol{x}'_{4it}$ includes the group to which the household belongs (HH-type=low, medium or high) and household socio-demographics (number of children in the house, social and economic status codes, age, age square, level of education and race of the main shopper in the house). The vector $\boldsymbol{x}'_{5it}$, included in some specifications, consists of log of price, information on product characteristics, and the discount variable described earlier.

The vector $\boldsymbol{\tau}'_{it}$ is a set of time-related indicator dummies coded differently across specifications. In the baseline model, it consists of dummies for each 4-week period (a pseudo-month) to account for non-linear trends in purchase patterns. In other specifications, we removed it completely (so it becomes a pure DD model) or replaced it with a set of recurring seasonal monthly dummies (January, February, etc., equal to one when the observation is from that calendar month for either year, and zero otherwise) or seasonal monthly dummies by country. Results from these alternative specifications are discussed later in section 4.3.2.

(2012), in nonlinear but strictly monotonic functions, the interaction term is not equal to a simple cross-difference but rather a difference between cross-differences. Specifically, the interaction term is equal to the cross difference of the conditional expectation of the observed outcome minus the cross difference of the conditional expectation of the potential outcome without treatment (i.e., the counterfactual). Importantly, the treatment effect is equal to the difference in the cross-differences, and hence the sign of the treatment effect in non-linear monotone increasing DD models is equal to the sign of the coefficient of the interaction term.

Equation 1 is estimated for all alcohol segments combined ($\ln Y_{s00}$) and then separately by segments: spirits ($\ln Y_{s01}$), beers and ciders ($\ln Y_{s02}$) and wines ($\ln Y_{s03}$). The segment analysis would assess the presence of heterogeneous effects of the ban by alcohol type.[6] In the latter three segment-specific estimations, the vector $\boldsymbol{x'_{5it}}$ includes prices of all four segments ($\ln p_{s01}, \ln p_{s02}, \ln p_{s03}$, and $\ln p_{s04}$) rather than just the price of the own segment, thereby allowing for substitutive or complementary effects.

3.3. **Endogeneity.** If we omit prices from the tobit specifications (included in the vector $\boldsymbol{x_{5it}}$), the total effect of the volume discount ban can be identified via the coefficient β_3 in a DD specification. However, retailers (or manufacturers) may change other alcohol promotion policies in Scotland in response to the ban on multiple-unit discounts, which in turn affects the price of a purchased bundle or multipack of alcohol. Since consumers would react to this change in the final price of the bundle, the total effect of the ban would consist of the direct effect of the ban, plus the indirect effect via the changed prices. Thus, for the identification of the effect of the direct ban, prices must be included in the regression. One way to think about this is to disentangle the total effect of the policy into its direct and indirect components, as is done in a causal mediation analysis (Keele, Tingley and Yamamoto, 2015, Albert and Nelson, 2011).

[6]As noted earlier, while FABs are expected to be affected by the ban, we omitted the estimation of the FABs segment (Y_{s04}) due to very few sales in the observed period.

However, consumers can react to the policy change or the associated price changes of individual items, and adjust the contents of the bundle of alcohol they purchase by substituting cheaper items or those with different product characteristics. In effect then, because the consumers choose the contents of a bundle of alcohol, the price may be endogenous, i.e., correlated with the error term, due to unobserved bundle characteristics that are correlated with price and the quantity purchased. Since we observe many of these bundle characteristics, we added the vector $\boldsymbol{x_{5it}}$, which includes price and other bundle characteristics (including discount) to the specification. In turn, this should attenuate the problem of correlation between price and the error term.

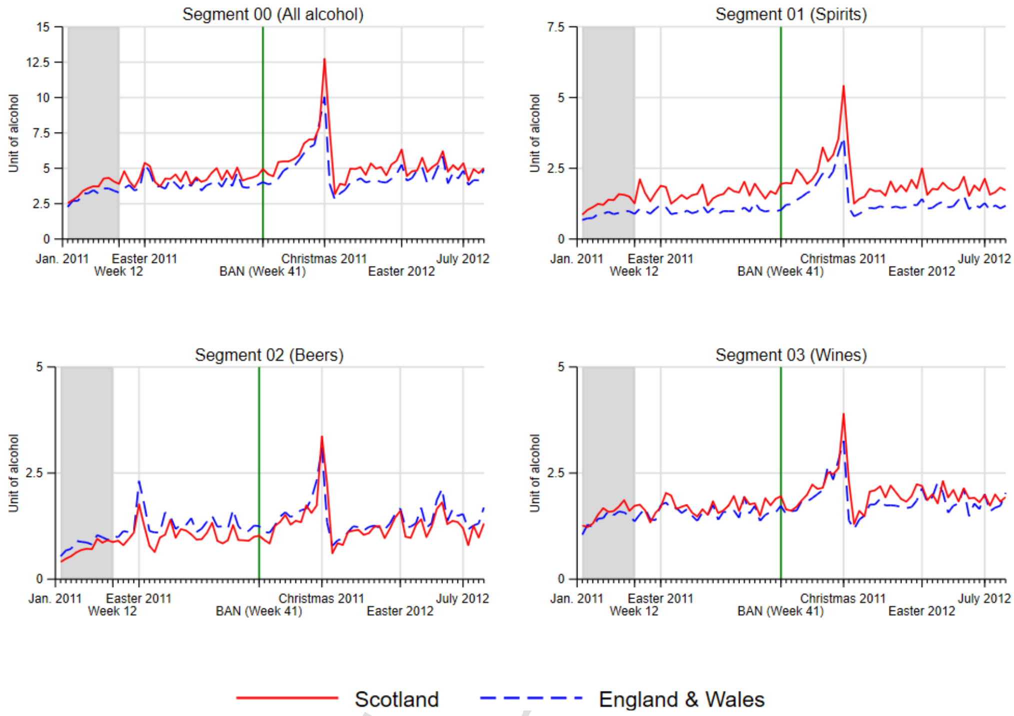
Nonetheless, we cannot rule out the possibility that other omitted demand-side variables in the error term are not correlated with prices. For instance, display location within a store may influence the choice of items in the bundle and may be correlated with price. Such (to the econometrician) unobservable additional bundle characteristics would cause a bias in the estimated coefficients. To account for these, we included control functions for prices, using instrumental variables that we expect will affect retail prices but not directly the demand for alcohol. Specifically, following Griffith, O'Connell and Smith (2019), our instruments include monthly factory price indexes for beer, cider/fruit wines, and an overall index for all such beverages. We also use weekly exchange rates between sterling and US dollar, and between sterling and euro, as these will affect the prices of imported alcohol and import duties paid on them. One reason for the regional price variation in the UK is the overall coverage by the main grocery stores. Thus, we include the market share of chain grocery stores and others by region as additional instruments. Finally, we also use weekly diesel prices, as they would be cost shifters for retailers, and interact them with shares of grocery stores by region. Variations in these instruments and their construction are described in detail in appendix A.1.

The control functions are constructed as residuals from first-stage regressions of log prices on all the exogenous variables plus the instruments listed above. To obtain standard errors that account for first-stage regression, we used block bootstrap by household and included the first- and second-stage (random effects) tobit within a draw by household with replacement (100 replications).

4. Results

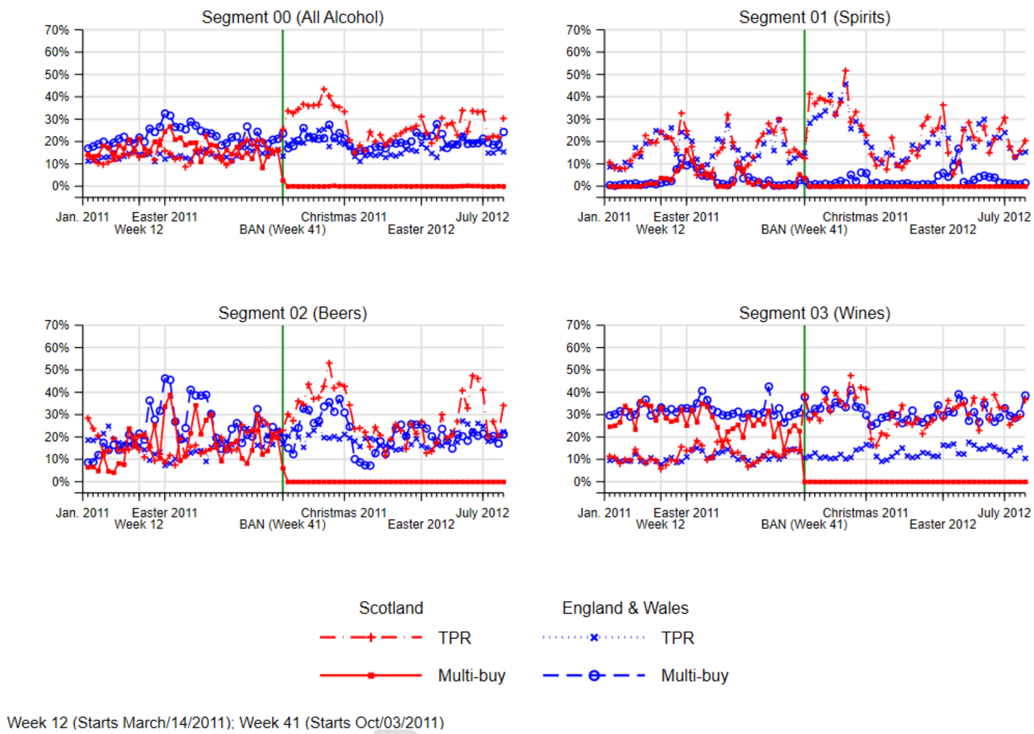
4.1. **Descriptive Analysis.** Figure 1 shows average consumption per adult by segment and country over the study period, including weeks 2-12 (shown in gray) used for classifying households by HH-type. The vertical line marked as 'BAN (Week 41)' corresponds to Monday, Oct/3/2011 (the multiple-unit discount ban came into effect on Oct/1/2011). While alcohol consumption seems to be increasing as Christmas/New Year approaches, there is no discernable difference in aggregate consumption patterns before and after the implementation of the policy across Scotland vs. England and Wales.

FIGURE 1. Weekly purchase per adult



In contrast, Figure 2 shows a very clear drop in multibuy promotions in Scotland after the ban was imposed. The figure shows average household alcohol expenditure by promotion type – in the form of either a temporary price reduction (TPR), which is a straight discount, or a multibuy discount – as a percentage of total alcohol household expenditure by country and week. The top left panel reports trends for all alcoholic products combined (S00). In week 41, the percentage of expenditure on multibuy promotions dropped to near zero in Scotland, while that on TPR jumped up from about 10% to around 30%. In contrast, there were no similar changes in the shares computed for England and Wales.

FIGURE 2. Expenditures (as %) by promotion type



A further breakdown by segments shows that this change in promotion types was mostly in beer and wine segments (the lower two panels, S02 and S03). In contrast, spirits (S01) were minimally affected by the ban. This is because multibuy promotions are not a typical promotion type for spirits; they were rarely used even before the ban came into effect and no discernible changes were found in the post-ban period.

4.2. **Price Regressions.** Figure 2 for expenses by promotion type suggests that effective prices may have changed. To check this, we tested whether consumers in Scotland chose bundles with different prices after the ban was introduced. Using household-level weekly data, we estimated random effects linear regressions of log price on $S_i, B_{it}, S_i \times B_{it}$, household characteristics in $\boldsymbol{x}_{4it}$ as well as weekly dummies $\boldsymbol{\tau}'$ and other product characteristics listed in $\boldsymbol{x}_{5it}$. Selected regression coefficients are reported in Table 2. Except for spirits (S01), we find a small but statistically significant reduction in prices for bundles selected by consumers in Scotland after the ban, which is most evident for beers (S02) and wine segments (S03).

TABLE 2. Reduced form regressions for (ln) prices

	(1) $\ln p_{s00}$ (All)	(2) $\ln p_{s01}$ (Spirits)	(3) $\ln p_{s02}$ (Beers)	(4) $\ln p_{s03}$ (Wines)	(5) $\ln p_{s04}$ (FABS)
B: PostBan	0.031***	0.103***	0.053***	0.013***	0.019***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)
S: Scotland	0.007*	0.019***	0.003	-0.001	-0.096***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
$S \times B$:	-0.014***	-0.003	-0.016***	-0.023***	0.067***
Scotland×PostBan	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)

All regressions include product characteristics, household characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively, and standard errors are clustered by household.

4.3. **Tobit estimates.** The tobit specifications were estimated for all segments and we added in price and bundle characteristics ($\boldsymbol{x}_{5it}$) sequentially, followed by control functions to further account for endogeneity. This allows us to measure the total effect of the ban, as well as the effect net of any price changes. We further estimated the models by household type. We discuss here only the left-censored marginal effects of select variables of interest, i.e., $\beta_k \Phi(\cdot)$

where Φ is the CDF for the normal distribution. The marginal effects were computed with S, B and $S \times B$ equal to one, and over the 4 weeks immediately following the ban, and other variables at their sample mean values.[7]

4.3.1. *Overall analysis - S00.* Table 3 reports marginals for the overall alcohol segment (S00). Panel A includes the time trend described earlier. Following the ban, there is evidence of an overall increase in alcohol purchases. Without controlling for prices, discounts, and product attributes (Column 1), alcohol purchases went up by 8.5% in the post-ban period. This result holds when controlling for prices, discount, and product attributes (Column 2). In this same column, the marginal effect for log price is -0.811 and for the discount is 0.454 (implying that own-price elasticity is -0.811% and a 10p increase in discount changes the quantity purchased by 4.54%).

In column (3), we add the control function. Doing so attenuates the overall effect of the ban as the marginal effect changes from 8.6% to 7.7% and the price elasticity increases in magnitude to -1.384. The next three columns report the analysis in column (3) by household type (HH-type: low, medium, and high). The impact of the ban is not present in low-consumption households, i.e., the marginal on the interaction term is not statistically significant, but

[7]For a non-interactive variable x_k, the left censored marginal is given by $\partial E(\ln y|X_i)/\partial x_k = \beta_k \Phi(\cdot)$ and we provide it here for the interactive term as well. Results from truncated marginal, i.e., $\partial E(\ln y|X_i, y_i > 0)/\partial x_k$ are similar and omitted in the interest of space. The full set of all regression coefficients is available in the online Appendix C.

Table 3. Per capita alcohol purchase – Marginals ($\beta_k\Phi(\cdot)$) for segment S00 (All Alcohol)

	(1)	(2)	(3)	(4)	(5)	(6)
Sample Households	All	All	All	Low	Medium	High
Panel A						
$S \times B$: Scotland×PostBan	0.085***	0.086***	0.077***	-0.004	0.074***	0.195***
	(0.018)	(0.018)	(0.018)	(0.023)	(0.030)	(0.035)
$\ln p_{s00}$: ln price ethanol		-0.811***	-1.384***	-1.091***	-1.317***	-1.793***
		(0.036)	(0.073)	(0.120)	(0.122)	(0.128)
d_{s00}: segment discount		0.454***	0.442***	0.163***	0.426***	0.706***
		(0.032)	(0.032)	(0.049)	(0.054)	(0.063)
Panel B						
	No time trends (pure DD – $\boldsymbol{\tau}$ excluded)					
$S \times B$: Scotland×PostBan	0.085***	0.087***	0.079***	-0.002	0.077***	0.191***
	(0.019)	(0.018)	(0.018)	(0.023)	(0.031)	(0.036)
Sample	594,696	594,694	594,694	197,735	198,302	198,657
Households	8,376	8,376	8,376	2,785	2,793	2,798
	First-difference (Scotland Only)					
B	0.097***	0.037	-0.028	-0.056	-0.131***	0.198***
(B = 1 if post-ban)	(0.038)	(0.036)	(0.036)	(0.049)	(0.064)	(0.081)
Sample	47,925	47,923	47,923	15,620	15,974	16,329
Households	675	675	675	230	225	220
Prices	✗	✓	✓	✓	✓	✓
Control functions	na	✗	✓	✓	✓	✓

All regressions include household characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Columns (1,2,3) additionally contain dummy variables for the type of household (low, medium, and high), while columns (4,5,6) provide sub-analysis by HH-type. Column (1) does not contain prices, discount, or observable product characteristics. Column (2) adds prices, discount, and observable product characteristics. Column (3) adds in control variables as residuals from first-stage regressions where price is regressed on exogenous variables and additional excluded instruments. Columns (4,5,6) are similar to (3) but restrict that sample by household type. Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively, and are based on bootstrapped standard errors. The first-difference results in the bottom panel include data only from Scotland, which has fewer observations, and none of the regressions include a dummy for Scotland or the interaction term $S \times B$.

in medium- and high-consumption households it is positive, significant, and progressively increases in magnitude (7.4% and 19.5% in columns 5 and 6). The price sensitivity also increases as we move from low- to high-consumption households.

Columns (3) onwards rely on the use of instruments described earlier. While the exogeneity condition of the instruments is credible (they are cost shifters and should not directly affect demand except through prices), their relevance in determining prices is largely an empirical issue. Table A-2 provides first-stage F-tests for joint significance of the excluded instruments (weak instruments tests). For the overall alcohol segment (segment S00), the F-value is 44.2, while for sub-analysis by HH-type, the values are 32.6, 20.3, and 14.0, indicating that in all cases we can reject the null hypothesis of no relationship between price and the excluded instruments.[8]

Panel B of Table 3 provides the marginal effects for the dummy for post-ban when we re-estimate the tobit using no time trends and if we use data from Scotland only, i.e., first-difference estimates. In the former case, the marginals are quite similar to the case with time trends. In the latter case, the first-difference estimates show the extent to which our results are driven by the

[8]The regression coefficients are given in the online Appendix C . The first-stage residuals from these regressions were added as control functions in the second-stage tobit models and were all statistically significant. Table A-2 also provides first-stage F-tests for segment-level analysis to follow. In each case, the F-test values are large and hence reject a null hypothesis of there being no relationship between our instruments and prices.

trend in the control group. Column (1) shows that post-ban there was an increase in alcohol consumption of about 9.7% in Scotland, while column (2) indicates that this could mostly be explained by changes in prices. However, results in columns (4)-(6) show that first-difference results by household type are not similar to those from the DD analysis. For instance, column (5) indicates that consumption declined for households with medium-level consumption. By contrast, column (6) shows that in households with high consumption, the DD results are not driven just by the control group.

4.3.2. *Country-specific seasonality (S00).* Seasonality in alcohol consumption is a well-documented phenomenon, as can also be seen in Figure 1. While we have included a non-linear time trend, it is possible that our specification does not correctly identify the effect of the ban, if there is a seasonal effect by calendar months that differs by country. For instance, Christmas week and January are included in only the post-ban period, and if the spike in consumption during this period is higher for Scotland than in the control countries, then our estimates would overestimate the net impact of the ban.

To address such concerns, we re-estimated the models with alternative specifications. First, in $\boldsymbol{\tau}'$ we replaced the time trend with seasonality via dummies for calendar months (i.e., dummies for January, February, etc., equal to one when the observation is from that calendar month for either year, and zero otherwise). Second, we interacted these with the country dummy to allow for country-specific seasonal effects. Third, we re-estimated the model with country-specific seasonal effects on the observations for months that are available in both the pre- and post-treatment periods, so observations from months March-July of both years only, and still no observations for weeks 1-12 (recalling that in the initial analysis, we used data up to week 12 to classify households). Fourth, we re-estimated similarly to the previous case but with retained observations from January-July of both years (to allow the January effect to be different in both countries). Finally, to isolate any Christmas effect, we re-estimated the models using observations for just two months around Easter (March/April) each year and without any week dummies. The results for these five cases (summarized in Table A-3) do not differ much from the baseline case, and hence we retain our original specification for the rest of the analysis.

4.4. **Segment Analysis - S01, S02 and S03.** Table 4 summarizes the marginal effects (i.e., $\beta_k\Phi(\cdot)$) for each segment. The top panel of the table shows minimal negative net impact of the ban on spirits for all households combined. It becomes significant at the 5% level only for HH-type=medium, and their overall purchased quantity declined by about 4.6%. This result is in line with the observation that the ban was not a binding constraint, as this type of promotion was hardly used for spirits (see Figure 2). However, there were other supporting measures in the legislation (see footnote (1)) that could have had a marginal effect, such as restricting the in-store display area. For instance, if alcohol can no longer be displayed at the checkout counter, this would not act as a potential reminder to medium-level drinkers about a purchase, while the low and heavy drinkers would not be tempted or reminded by it anyway.

Our results show a net decrease among medium-level consumers, but neither the high nor low HH-type were affected. We also find that price sensitivity increases by HH-type, and adding control functions does increase the magnitude of the price coefficient. Also, the cross-price effects become positive (and often significant) after adding the control functions.

TABLE 4. Per capita alcohol purchase – marginals ($\beta_k\Phi(\cdot)$) by segment

	(1)	(2)	(3)	(4)	(5)	(6)
Sample Households	All	All	All	Low	Medium	High
			Segment S01 (Spirits)			
$S \times B$: Scotland×PostBan	-0.005	-0.010	-0.016	-0.016	-0.046**	-0.002
	(0.012)	(0.012)	(0.014)	(0.012)	(0.022)	(0.028)
$\ln p_{s01}$: ln price sprits		-0.208***	-1.958***	-0.230**	-1.726***	-3.228***
		(0.030)	(0.186)	(0.100)	(0.279)	(0.325)
$\ln p_{s02}$: ln price beers		-0.031***	0.225***	0.071*	0.324***	0.263***
		(0.011)	(0.036)	(0.041)	(0.076)	(0.078)
$\ln p_{s03}$: ln price wines		-0.027***	0.230***	0.100	0.256***	0.340***
		(0.012)	(0.051)	(0.064)	(0.095)	(0.105)
$\ln p_{s04}$: ln price FABs		-0.003	0.126***	0.020	0.049	0.274***
		(0.006)	(0.033)	(0.032)	(0.040)	(0.091)

All regressions include household characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Columns (1,2,3) additionally contain dummy variables for the type of household (low, medium, and high), while columns (4,5,6) provide sub-analysis by HH-type. Column (1) does not contain prices, discount, or observable product characteristics. Column (2) adds prices, discount, and observable product characteristics. Column (3) adds control variables as residuals from first-stage regressions where price is regressed on exogenous variables and additional excluded instruments. Columns (4,5,6) are similar to (3) but restrict that sample by household type. Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively, and are based on bootstrapped standard errors.

TABLE 4. Per capita alcohol purchase – marginals ($\beta_k \Phi(\cdot)$) by segment

	(1)	(2)	(3)	(4)	(5)	(6)
Sample Households	All	All	All	Low	Medium	High
d_{s01}: segment discount		0.571***	-0.055	0.032	-0.185	0.886***
		(0.085)	(0.122)	(0.089)	(0.206)	(0.198)
			Segment S02 (Beers)			
$S \times B$: Scotland×PostBan	0.063***	0.052***	0.034***	-0.011	0.054***	0.068***
	(0.011)	(0.010)	(0.009)	(0.012)	(0.019)	(0.024)
$\ln p_{s01}$: ln price sprits		0.013**	0.169***	0.079***	0.182***	0.276***
		(0.008)	(0.022)	(0.023)	(0.037)	(0.052)
$\ln p_{s02}$: ln price beers		-0.364***	-1.138***	-0.635***	-0.984***	-1.824***
		(0.023)	(0.064)	(0.090)	(0.112)	(0.159)
$\ln p_{s03}$: ln price wines		-0.054***	0.138***	0.126***	0.110*	0.130**
		(0.007)	(0.029)	(0.042)	(0.067)	(0.060)
$\ln p_{s04}$: ln price FABs		0.001	0.093***	0.063***	0.054**	0.176***
		(0.004)	(0.019)	(0.020)	(0.024)	(0.053)
d_{s02}: segment discount		0.110***	0.091***	0.022	0.132***	0.107*
		(0.021)	(0.023)	(0.031)	(0.042)	(0.056)
			Segment S03 (Wines)			
$S \times B$: Scotland×PostBan	0.033***	0.027***	0.005	-0.019	-0.008	0.042
	(0.013)	(0.013)	(0.012)	(0.014)	(0.022)	(0.028)
$\ln p_{s01}$: ln price sprits		0.018**	0.130***	0.040	0.225***	0.224***
		(0.010)	(0.026)	(0.025)	(0.055)	(0.069)
$\ln p_{s02}$: ln price beers		-0.025***	0.231***	0.122***	0.235***	0.274***
		(0.011)	(0.040)	(0.038)	(0.057)	(0.081)
$\ln p_{s03}$: ln price wines		-0.437***	-1.324***	-0.594***	-1.345***	-2.104***
		(0.027)	(0.094)	(0.106)	(0.146)	(0.207)
$\ln p_{s04}$: ln price FABs		-0.000	0.067***	0.016	0.089***	0.067
		(0.005)	(0.023)	(0.032)	(0.030)	(0.074)

All regressions include household characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Columns (1,2,3) additionally contain dummy variables for the type of household (low, medium, and high), while columns (4,5,6) provide sub-analysis by HH-type. Column (1) does not contain prices, discount, or observable product characteristics. Column (2) adds prices, discount, and observable product characteristics. Column (3) adds control variables as residuals from first-stage regressions where price is regressed on exogenous variables and additional excluded instruments. Columns (4,5,6) are similar to (3) but restrict that sample by household type. Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively, and are based on bootstrapped standard errors.

TABLE 4. Per capita alcohol purchase – marginals ($\beta_k\Phi(\cdot)$) by segment

	(1)	(2)	(3)	(4)	(5)	(6)
Sample Households	All	All	All	Low	Medium	High
d_{s03}: segment discount		0.196***	0.113***	0.073**	0.097**	0.175***
		(0.022)	(0.022)	(0.036)	(0.039)	(0.046)
Sample	594,696	584,067	584,067	193,673	193,088	197,306
Households	8,376	8,376	8,376	2,785	2,793	2,798
Prices	✗	✓	✓	✓	✓	✓
Control functions	na	✗	✓	✓	✓	✓

All regressions include household characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Columns (1,2,3) additionally contain dummy variables for the type of household (low, medium, and high), while columns (4,5,6) provide sub-analysis by HH-type. Column (1) does not contain prices, discount, or observable product characteristics. Column (2) adds prices, discount, and observable product characteristics. Column (3) adds control variables as residuals from first-stage regressions where price is regressed on exogenous variables and additional excluded instruments. Columns (4,5,6) are similar to (3) but restrict that sample by household type. Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively, and are based on bootstrapped standard errors.

The second part of Table 4 provides marginals for the beers and ciders segment. The estimated net impact of the ban is about 6.3% in column (1), but is almost halved (3.4%) in column (3), when we control for prices. While the low HH-type group was largely unaffected by the ban, an increase of 5.4% and of 6.8% in the purchased quantity was estimated in medium and high HH-types.

Finally, the third part of Table 4 provides marginals for the wine segments. While the net effect of the ban is positive and significant (3.3% increase), this seems to be driven mostly by the associated price changes. The net effect of the ban is eliminated, when we add in the prices and control functions (see column (3)). As before, adding the control functions results in a negative and significantly larger magnitude in own-price coefficients, while the price coefficients of other alcohol segments become almost all positive and significant. Furthermore, as columns (4) and (6) reveal, while the effect of the ban is not statistically significant for either group after accounting for endogeneity of prices, the sign is negative for low HH-types and positive for high HH-types. Results not shown here – by HH-type and without adding in prices – also have negative and positive marginal effects for low and high types, respectively, but were statistically significant in those cases. In turn, it shows that the net effect of the ban worked through prices but in the opposite direction for the low and high HH-types.

4.5. **Expenditures.** Since prices decreased while quantities increased, the per capita household spending can increase or decrease. To assess the net impact on expenditures, we used a DD design similar to the quantity regressions, which control for household characteristics and time trends. The marginal effect for the interaction term for expenditures shows an increase of about 6.5% for all households, none for low, 6.1% for medium, and 16% for high-consumption households (see Table A-4 in the appendix).

4.6. **Visits per week.** The foregoing analysis indicates that overall alcohol quantity increased even after accounting for the price changes. This is the exact opposite of the expected policy outcome, raising the question of why consumers responded this way.

One possibility is that shopping patterns changed. The presence of volume discounts may simultaneously provide both a financial inducement and a constraint to bulk in buy. With the ban in place, though, and replacement of multibuys with straight discounts, the constraint is removed even if the incentive and ability to buy on discount still exist, allowing for both buying in bulk and for extra incremental purchases. The implication is that consumers could then spread out their purchases over time, rather than focus their buying on single large shopping trips. Moreover, beyond any budgeting benefit that this may afford, there may be psychological drivers to increase shopping trips, either to gain additional transaction utility and segregate perceived gains from buying on straight discounts, in line with Thaler (1985), or because the absence of multibuys made it harder to commit to spaced-out shopping trips and avoid top-up shopping when consumers exhibit time-inconsistent preferences and face self-control problems, in line with Thaler and Shefrin (1981) and Hoch and Loewenstein (1991).

To illustrate the latter possibility, consider a consumer who makes a fixed number of visits per week to grocery stores to purchase alcohol for the entire week. Prior to the ban, she takes into account the non-linear prices and buys four packs of her favorite alcohol where the fourth unit is at a lower price per unit. If she runs out of alcohol before the end of the week, she waits until the next week to purchase a similar total amount, rather than buy a fifth unit at a higher marginal price. However, after the multiple-unit discount ban is imposed, so all units are sold at the same uniform price, the marginal price of the fifth unit of alcohol is the same as that of the earlier four units. In this case, she might be tempted to make an additional visit to the store during the same week to purchase the extra unit of alcohol, and perhaps even more. Thus, the removal of the multibuy constraint may take the brake off store visits, similarly to how deregulating store opening hours might remove the commitment device curbing visit frequency and spending (Hinnosaar, 2016), and in turn binge drinking (Marcus and Siedler, 2015), especially in view of heavy drinkers being prone to time-inconsistent behavior and waiting impulsivity (Mayhew et al., 2020).

Whether it is about better budgeting, segregating perceived transaction utility gains, or weakened ability to commit to limiting store visits, one might expect that higher consumption households with a greater desire for additional alcohol may be more susceptible to increasing the number of store visits after the ban. To test this hypothesis, we computed the total number of alcohol purchase visits per week for each household and used it as the outcome variable in our DD design. Specifically, using the count of the number of shop visits per week as the dependent variable, we estimated the random coefficients poisson models with over-dispersion (i.e., negative binomial models to allow the variance of the dependent variable to be larger than its mean).[9] The regressions control

[9]Since we allowed for clustering, over-dispersion can be rejected in favor of a simple poisson estimate. In models without clustering, over-dispersion is not rejected and hence the negative

for household characteristics as before, but now we also control for the size of the household, as larger households may shop more often. Table 5 shows the results, revealing that the mean number of visits increased for medium- and, especially, for high-consumption households in line with the hypothesis of increased shopping frequency.

TABLE 5. Poisson regression (visits per week)

Sample Households	(1) All	(2) Low	(3) Medium	(4) High
Visits (mean)	0.588	0.313	0.491	0.959
Visits (variance)	(0.676)	(0.310)	(0.459)	(1.034)
S: Scotland	0.052*	0.014	0.062	0.098*
	(0.031)	(0.063)	(0.049)	(0.051)
B: PostBan	0.035***	0.058**	0.019	0.036***
	(0.009)	(0.023)	(0.017)	(0.013)
$S \times B$: Scotland-PostBan	0.078***	-0.025	0.073***	0.112***
	(0.017)	(0.044)	(0.028)	(0.023)
alpha (log of)	-0.906	-0.723	-0.943	-1.063
(dispersion parameter)	(1.343)	(2.209)	(2.364)	(2.327)
Sample	594,696	197,735	198,303	198,658
Households	8,376	2,758	2,793	2,798

All regressions include household characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Column (1) includes observations from all households while columns (2,3,4) restrict by HH-type (low, medium, and high). Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively. All standard errors are clustered by household.

We further investigated whether this pattern was driven by income differences, so essentially if it is about budgeting. Note, though, that the regressions reported above already control for household characteristics including income.

binomial is preferred in that case. In fact, the mean and variance of visits per week are not too different (given in the table) and hence the poisson model may be appropriate.

Nevertheless, we performed two further tests using additional count models by sub-samples of income groups. If the pattern was simply about the desire to smooth spending and better budgeting, then we might expect – upon restricting the low-drinking households in column (2) to just the lowest-income households in our sample (annual income less than £20K) – that the non-significant -0.025 coefficient on the interaction term may become positive and significant. Conversely, if we restrict the highest-drinking households in column (4) to the relatively well-off households (annual income above £40K), then the significant and positive coefficient of 0.112 on the interaction term may become negative and significant, as this group would not be budget-constrained and may want to have fewer visits due to higher opportunity costs. Neither of these tests turned out to be true, and the results across different income sub-samples stayed qualitatively similar to those reported above in Table 5. These additional results are available upon request.[10]

Similarly, we also computed the average amount of alcohol purchased per trip, rather than per week, for each household both before and after the ban (using just two values for each household computed from all the trips in the pre- and post-ban periods). In a similar DD design as above, the interaction

[10]For the household group classified as low, they are not statistically different from zero for all income levels. Similarly, for households classified as high, the coefficients on the interaction term are positive and significant for all income levels (0-20K, 20-40K, 40-60K) except for the top income group of 60K+ p.a., where it is still positive but not statistically significant due to small sample size.

term did not show a decrease in the amount of alcohol purchased per trip in Scotland after the ban relative to England and Wales for the low- or high-consumption household groups. It did show a slight decrease for those in the medium level of consumption at the 10% significance level. Taken together, the additional analyses point to increased shopping frequency and purchases after the ban being more than simply due to budgeting, lending credence to the aforementioned behavioral arguments as well as leaving open the possibility of other unmodeled factors.

Table 6. Pre-ban parallel trends test (p-values)

Sample Households	(1) All	(2) Low	(3) Medium	(4) High
Panel A				
Segment S00 (All Alcohol)	0.011	0.316	0.820	0.103
Segment S01 (Spirits)	0.069	0.289	0.943	0.108
Segment S02 (Beers)	0.666	0.397	0.588	0.302
Segment S03 (Wines)	0.068	0.594	0.205	0.092
Panel B				
Segment S00 (All Alcohol)	0.043	0.219	0.981	0.227
Segment S01 (Spirits)	0.633	0.239	0.953	0.930
Segment S02 (Beers)	0.983	0.425	0.724	0.335
Segment S03 (Wines)	0.224	0.545	0.496	0.239

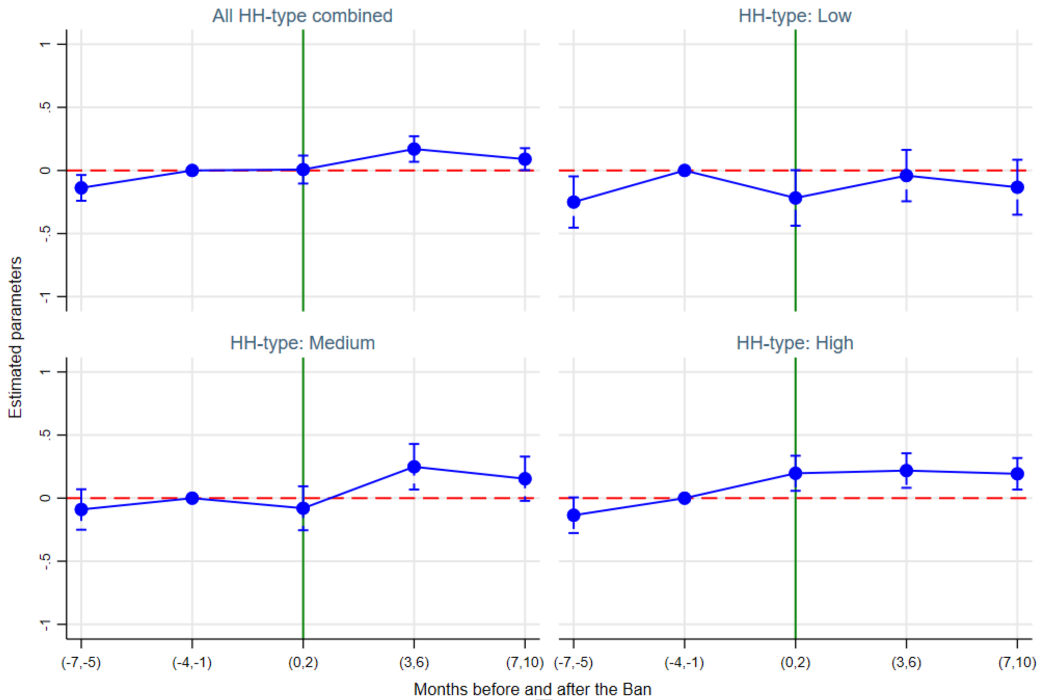
All regressions include household characteristics, product characteristics, prices, residuals from the first stage regressions, a dummy for each 4-week period (a pseudo-month) and a dummy for Scotland (1 for Scotland, 0 for England and Wales). Column 1 is for all households and contains a dummy for the household type while columns (2,3,4) provide sub-analysis by HH-type. All regressions also include interaction terms of country dummy with the 4-week monthly dummies and the sample is restricted to the pre-ban period. The reported p-values are for the joint F-test with a null that the interaction terms are zero. Panel B excludes observations from (pseudo-) month five.

4.7. **Parallel trends.** The graphical inspection of Figure 1 suggests that treated and control countries had similar trends in the period prior to the ban (weeks 13-39) and that the common trends is a reasonable assumption. This is consistent with other studies reporting common trends between these countries covering the same period with different alcohol purchase data (e.g., Robinson et al. (2014) with Nielsen data and Rehm et al. (2022) with Alcovision data) as well as studies covering later periods (Griffith, O'Connell and Smith (2022), Anderson, Kokole and Jané-Llopis (2022), inter alia). Furthermore, there is no reason to assume that apart from the Alcohol etc. (Scotland) Act 2010, there has been any major change over our study period to policy or to industry structure and demand that would have had a differential impact on Scotland compared to England & Wales. Even so, there is the possibility of other factors or events affecting consumption patterns differently at specific points in time.

We formally tested for parallel trends using data only from the pre-ban period and creating interaction terms between the country dummy and the 4-week monthly dummies. We then tested for the joint significance of these interaction terms and report the p-values in panel A of Table 6. In four out of 16 cases, the p-values are below 0.10. We further investigated these and generally found deviations from parallel trends around week 20 (weeks 19-21 and 25). We repeated the tests by dropping observations from month 5 (which includes week 20) and report the results in panel B of the table. In all but one case the p-values are now above 0.10 (the exception is S00-All). Finally, we further checked if the results in the main analysis reported in Table 3 and Table 4 hold up when we drop observations from month 5. The results, summarized in Appendix A.4, are very similar to the main analysis. While not perfect, these results, alongside the qualitative evidence cited above, and the evidence from event-study style analysis in the next section, suggest that the DD design method is suitable for identification in this context.

4.8. **Persistent effects.** It is possible that the effect of the ban diminishes over time. To check this, we re-estimated all the models but replaced the interaction term $S \times B$ with $\sum_j S \times B \times \rho_j$ where ρ_j is a 1/0 dummy equal to one if the observation is from the j-th period. We have 18 months of data with eight before and inclusive of the month of the policy change, and ten after. We grouped these into five periods consisting of either three or four months. These are (-7,-5), (-4,-1), (0,2), (3,6), and (7,10), where the reference period is (-4,-1), which is the period just before the policy change. Figure 3 displays the estimated coefficients for overall alcohol purchasing (S00) by household type (and similar figures for spirits (S01), beers and ciders (S02), and wines (S03) are in Appendix A.5). Generally, we observe that for all types of alcohol combined, and particularly for beers and ciders, the net impact of the ban on consumption is positive and well above the zero line for most of the post-ban time window, notably for the high-consumption group (bottom right panels).

FIGURE 3. Interaction terms $S \times B$ over time (Segment S00)



Plot of coefficients for $\sum_j S \times B \times \rho_j$ where ρ_j is a 1/0 dummy equal to one if the observation is from the j-th period. Error bars are 95% confidence intervals.

4.9. **Robustness and treatment heterogeneity.** Our main analysis has household-weekly level data and estimates tobit models. The results were robust to some sample section criteria, such as the exclusion of households from Wales or the inclusion of households living within 35 km of the Scottish-English border. They were also robust to using linear instead of tobit models with estimates of the coefficients on the interaction $S \times B$ summarized in Table B-1 (in the online Appendix B). We also re-estimated the models for S00 segment with regional dummies to check if their inclusion affects the conclusions (recalling there are 15 regions in our data and five of them are in Scotland). It did not, and the results are given in Table B-2 in the online appendix.

We further checked if the aggregation over time or households changes the results. In terms of using household-monthly observations, with some exceptions, the analysis at the monthly level generally confirms previous findings in terms of the sign and significance of the $S \times B$ parameter, albeit the marginal effects were found to be slightly higher in magnitude and significant at the 10% level for some consumer groups in the spirits (S01) and wines (S03) segments. In a separate analysis, we aggregated the data at the regional level and used linear models without control functions. The signs of the $S \times B$ parameter are in line with those reported in the main analysis, the net effects were found to be statistically significant only for the S00 and S02 analyses (approximately 5% and 16%, respectively).

Finally, women on average drink much less than men in the UK, and hence, perhaps they react differently to policy changes.[11] Thus, we re-estimated the tobit models by splitting the sample by the gender of the main shopper, and then by further restricting it to single-person households (so the gender of the shopper is the same as the person most likely consuming the alcohol). Results are given in Table B-3 and show that the marginal effects for households

[11] The recommended lower-risk consumption limit is 30% less for women than men, and on average women drink considerably less than men in the UK. See NICE (2010, 7-Glossary), Scottish Government (2020, Table 3.1) for Scotland and NHS Digital (2022, Table 1) for England.

with female shoppers are lower than for households with male shoppers, i.e., 0.073 vs. 0.108, and the contrast becomes greater for high-volume purchasers (though, the differences are not statistically significant based on overlapping confidence intervals).[12] When we restrict the sample further to single-person households by gender, the marginal effect for medium- or high-volume households with females becomes larger than when we did not impose the single-person restriction.[13] For further details on results described in this section please see the online Appendix B.

5. Summary and Conclusions

Volume discounts are designed to encourage consumers to buy more. In principle, banning these discounts should lead to consumers buying less. Our results do not support this finding for the ban on multiple-unit discounts on alcohol sales introduced across Scotland in October 2011, as retailers switched to using

[12]This trend is reversed in households classified as medium-level purchasers. One possible explanation is because our low, medium, and high classification are not gender specific, and since women drink less than men (on average), female-shopper households falling in the medium group are more akin to the high-volume households among male shoppers.

[13]The sample for single-person by gender and household type is very small and we are reluctant to push this analysis too far, particularly for single-male-occupant households. For instance, in Scotland, there were only 13 households with single-male occupants classified as low, and we were not able to compute marginal effects for this group. Even for other groups, the failure rate for convergence of the tobit models under bootstrap sample draws was very high.

more single-unit discounts and high-consumption households increased their shopping frequency to buy more.

We are mindful that our analysis only considers a relatively short period following the introduction of the ban, and perhaps market outcomes changed in subsequent years. Even so, for policymakers, our findings suggest that the effectiveness of such a ban can depend on two critical factors: (1) strategic responses by firms, when they have leeway to switch profitably to using other forms of discounts to counter any loss of sales; and (2) behavioral reactions of consumers, particularly heavy users, who may use non-linear pricing as a commitment mechanism to curb their total consumption.

The original intention of the Scottish government was to initiate the ban on volume discounts alongside introducing minimum unit pricing (MUP), which was delayed until 2018 but early indications point to reduced purchases by high-consumption households since then (Griffith, O'Connell and Smith, 2022). In this context, the policy combination could be effective when MUP limits the ability of retailers to offer deep straight discounts or deep volume discounts for large size containers and multipacks. On this basis, a volume discount ban may work well in tandem with other measures affecting alcohol affordability, if not so well on its own.[14]

[14] Separately, we note that the UK government has delayed introducing a ban on multibuys in England for HFSS foods and drinks (DHSC, 2022).

Acknowledgements. We wish to thank the editorial team and reviewers for their suggestions and guidance. In addition, we thank Rachel Griffith, Josh Kraindler, Eugenio Miravete, Katja Seim, and seminar participants at the Center for Competition Policy and the Health Economics Study Group UK (2019) for helpful comments and feedback. MS is a former member of Behaviour and Health Research Unit (BHRU) at the University of Cambridge which owns the data used in this paper. We are very grateful to its director, Professor Theresa Marteau, for allowing us to use the data and to the authors of Nakamura et al. (2014) for making the data available to us. Neither the funders nor the data providers bear any responsibility for the analyses or interpretations presented here.

References

Adams, William James, and Janet L. Yellen. 1976. "Commodity bundling and the burden of monopoly." *Quarterly Journal of Economics*, 90(3): 475–498.

Ai, Churong, and Edward C. Norton. 2003. "Interaction terms in logit and probit models." *Economics Letters*, 80(1): 123–129.

Albert, Jeffrey M., and Suchitra Nelson. 2011. "Generalized Causal Mediation Analysis." *Biometrics*, 67(3): 1028–1038.

Ally, Abdallah K., Yang Meng, Ratula Chakraborty, Paul W. Dobson, Jonathan S. Seaton, John Holmes, Colin Angus, Yelan Guo, Daniel Hill-McManus, Alan Brennan, and Petra S. Meier. 2014. "Alcohol tax pass-through across the product and price range: do retailers treat cheap alcohol differently?" *Addiction*, 109(12): 1994–2002.

Anderson, Peter, Amy O'Donnell, Eileen Kaner, Eva Jané-Llopis, Jakob Manthey, and Jürgen Rehm. 2021. "Impact of minimum unit pricing on alcohol purchases in Scotland and Wales: controlled interrupted time series analyses." *The Lancet Public Health*, 6(8): e557–e565.

Anderson, Peter, Daša Kokole, and Eva Jané-Llopis. 2022. "Impact of minimum unit pricing on shifting purchases from higher to lower strength beers in Scotland: Controlled interrupted time series analyses, 2015–2020." *Drug and Alcohol Review*, 41(3): 646–656.

Armstrong, Mark. 2016. "Nonlinear pricing." *Annual Review of Economics*, 8(1): 583–614.

Athey, Susan, and Guido W. Imbens. 2006. "Identification and inference in nonlinear difference-in-differences models." *Econometrica*, 74(2): 431–497.

Calcott, Paul. 2019. "Minimum unit prices for alcohol." *Journal of Health Economics*, 66: 18–26.

Conlon, Christopher T., and Nirupama L. Rao. 2019. "The price of liquor is too damn high: alcohol taxation and market structure." New York University NYU Wagner Research Paper 2610118.

Conlon, Christopher T., and Nirupama L. Rao. 2020. "Discrete prices and the incidence and efficiency of excise taxes." *American Economic Journal: Economic Policy*, 12(4): 111–43.

Cooper, Benjamin, Markus Gehrsitz, and Stuart G. McIntyre. 2020. "Drink, death, and driving: Do blood alcohol content limit reductions improve road safety?" *Health Economics*, 29(7): 841–847.

DHSC. 2021. "Consultation outcome: Restricting volume promotions for high fat, sugar, and salt (HFSS) products." Department of Health and Social Care, United Kingdom Consultation Outcome (updated July 19, 2021). Reference No: RPC-DHSC-4333(3).

DHSC. 2022. "Government delays restrictions on multibuy deals and advertising on TV and online." *Department of Health and Social Care, UK Government [Press Release]*.

Francesconi, Marco, and Jonathan James. 2021. "None for the road? Stricter drink driving laws and road accidents." *Journal of Health Economics*, 79(102487): 1–23.

Giles, Lucie, and Elizabeth Richardson. 2020. "Monitoring and Evaluating Scotland's Alcohol Strategy (MESAS): Monitoring Report 2020." Public Health Scotland PHS MESAS Monitoring Report, Edinburgh.

Giles, Lucie, Mark Robinson, and Clare Beeston. 2019. "Minimum Unit Pricing (MUP) Evaluation. Sales-based consumption: a descriptive analysis of one year post-MUP off-trade alcohol sales data." NHS Health Scotland PHS MESAS Monitoring Report, Edinburgh.

Green, Colin P., John S. Heywood, and Maria Navarro. 2014. "Did liberalising bar hours decrease traffic accidents?" *Journal of Health Economics*, 35: 189–198.

Griffith, Rachel, Martin O'Connell, and Kate Smith. 2019. "Tax design in the alcohol market." *Journal of Public Economics*, 172: 20–35.

Griffith, Rachel, Martin O'Connell, and Kate Smith. 2022. "Price floors and externality correction." *The Economic Journal*, 132(646): 2273–2289.

Hindriks, Jean, and Valerio Serse. 2019. "Heterogeneity in the tax pass-through to spirit retail prices: evidence from Belgium." *Journal of Public Economics*, 176: 142–160.

Hinnosaar, Marit. 2016. "Time inconsistency and alcohol sales restrictions."

European Economic Review, 87: 108–131.

Hoch, Stephen J., and George F. Loewenstein. 1991. "Time-inconsistent preferences and consumer self-control." *Journal of Consumer Research*, 17(4): 492–507.

Home Office. 2012. "Alcohol Multi-buy promotions." Home Office, United Kingdom Impact Assessment.

Home Office. 2013. "Next steps following the consultation on delivering the Government's alcohol strategy." Home Office, United Kingdom Consultation Report.

Keele, Luke, Dustin Tingley, and Teppei Yamamoto. 2015. "Identifying mechanisms behind policy interventions via causal mediation analysis." *Journal of Policy Analysis and Management*, 34(4): 937–963.

Marcus, Jan, and Thomas Siedler. 2015. "Reducing binge drinking? The effect of a ban on late-night off-premise alcohol sales on alcohol-related hospital stays in Germany." *Journal of Public Economics*, 123: 55–77.

Mayhew, Matthew J., James M. Byrne, Jane H. Powell, and Tim Meynen. 2020. "Are hazardous drinkers more impulsive than light drinkers? A comprehensive assessment in young adults." *Alcohol*, 84: 9–20.

Miravete, Eugenio J., Katja Seim, and Jeff Thurk. 2018. "Market power and the Laffer curve." *Econometrica*, 85(5): 1651–1687.

Miravete, Eugenio J., Katja Seim, and Jeff Thurk. 2020. "One markup to rule them all: taxation by liquor pricing regulation." *American Economic Journal: Microeconomics*, 12(1): 1–41.

Nakamura, Ryota, Marc Suhrcke, Rachel Pechey, Marcello Morciano, Martin Roland, and Theresa M. Marteau. 2014. "Impact on alcohol purchasing of a ban on multi-buy promotions: a quasi-experimental evaluation comparing Scotland with England and Wales." *Addiction*, 109(4): 558–567.

Nelson, Jon P., and John R. Moran. 2019. "Effects of alcohol taxation on prices: a systematic review and meta-analysis of pass-through rates." *The B.E. Journal of Economic Analysis and Policy.*, 20(1): 20190134.

NHS Digital. 2022. "Health Survey for England 2011 to 2019: Alcohol Additional Analyses."

NICE. 2010. "Alcohol-use disorders: prevention." National Institute for Health and Care Excellence (NICE).

O'Donnell, Amy, Peter Anderson, Eva Jané-Llopis, Jakob Manthey, Eileen Kaner, and Jürgen Rehm. 2019. "Immediate impact of minimum unit pricing on alcohol purchases in Scotland: controlled interrupted time series analysis for 2015-18." *British Medical Journal*, 366(15274): 1–9.

Puhani, Patrick A. 2012. "The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models." *Economic Letters*, 115(1): 85–87.

Rehm, Jürgen, Amy O'Donnell, Eileen Kaner, , Eva Jané-Llopis, Jakob Manthey, and Peter Anderson. 2022. "Differential impact of minimum unit pricing on alcohol consumption between Scottish men and women: controlled interrupted time series analysis." *BMJ Open*, 12(7): e054161.

Robinson, Mark, Claudia Geue, James Lewsey, Daniel Mackay, Gerry McCartney, Esther Curnock, and Clare Beeston. 2014. "Evaluating the impact of the alcohol act on off-trade alcohol sales: a natural experiment in Scotland." *Addiction*, 109(12): 2035–2043.

Robinson, Mark, Daniel Mackay, Lucie Giles, Jim Lewsey, Elizabeth Richardson, and Clare Beeston. 2021. "Evaluating the impact of minimum unit pricing (MUP) on off-trade alcohol sales in Scotland: an interrupted time–series study." *Addiction*, 16(10): 2697–2707.

Robinson, Mark, Janet Bouttell, James Lewsey, Daniel Mackay, Gerry McCartney, and Clare Beeston. 2018. "The short-term impact of the alcohol act on alcohol-related deaths and hospital admissions in Scotland: a natural experiment." *Addiction*, 113(3): 429–439.

Scottish Government. 2020. "The Scottish Health Survey." 2018 Edition (amended February 2020) ed.

Scottish Parliament. 2010. "Alcohol (Minimum Pricing)(Scotland) Bill." The Scottish Parliament SPICe Briefing 10-13, Edinburgh.

Thaler, Richard. 1985. "Mental accounting and consumer choice." *Marketing Science*, 4(3): 199–214.

Thaler, Richard H., and H. M. Shefrin. 1981. "An economic theory of self-control." *Journal of Political Economy*, 89(2): 392–406.

Vandoros, Sotiris, and Ichiro Kawachi. 2022. "Minimum alcohol pricing and motor vehicle collisions in Scotland." *American Journal of Epidemiology*, 191(5): 867–873.

WHO. 2010. "Global strategy to reduce the harmful use of alcohol." World Health Organization, Geneva, Switzerland. Last accessed on May 17, 2021.

Wilson, Luke B. Robert Pryce, Colin Angus, Rosemary Hiscock, Alan Brennan, and Duncan Gillespie. 2021. "The effect of alcohol tax changes on retail prices: how do on-trade alcohol retailers pass through tax changes to consumers?" *European Journal of Health Econonomics*, 22(3): 381–392.

Wilson, Robert B. 1993. *Nonlinear Pricing.* New York:Oxford University Press.

Woodhouse, John. 2020. "Alcohol: minimum pricing." UK Parliament, House of Commons Library.

Xhurxhi, Irena Palamani. 2020. "The early impact of Scotland's minimum unit pricing policy on alcohol prices and sales." *Health Economics*, 29(12): 1637–1656.

Appendix A. (Additional Results)

A.1. **Instruments and first-stage regression stats.** Our instrumenting strategy follows very closely that of Griffith, O'Connell and Smith (2019). To generate exogenous shocks to price, we used several variables that influence costs but are not likely to directly influence demand for alcohol. Some of these variables generate variation over time, while others give geographic variation. To this end, we used exchange rates for EUR and USD, which vary over time and may affect the prices differently for products that are imported vs. home-brewed. Similarly, we also used factory gate prices (indexes) for beer, cider and fruit wines, and for overall alcoholic beverages, as recorded by the Office of National Statistics. Weekly diesel prices were also used and were interacted with shares of stores by geographic coverage (see Figure A-1). To compute the latter, we used alcohol purchase data from the first 12 weeks and aggregated it up to store and regional level to compute shares by store type (seven type of stores) for each of the 15 regions separately (see Table A-1).

Figure A-1. Variation in price instruments over time

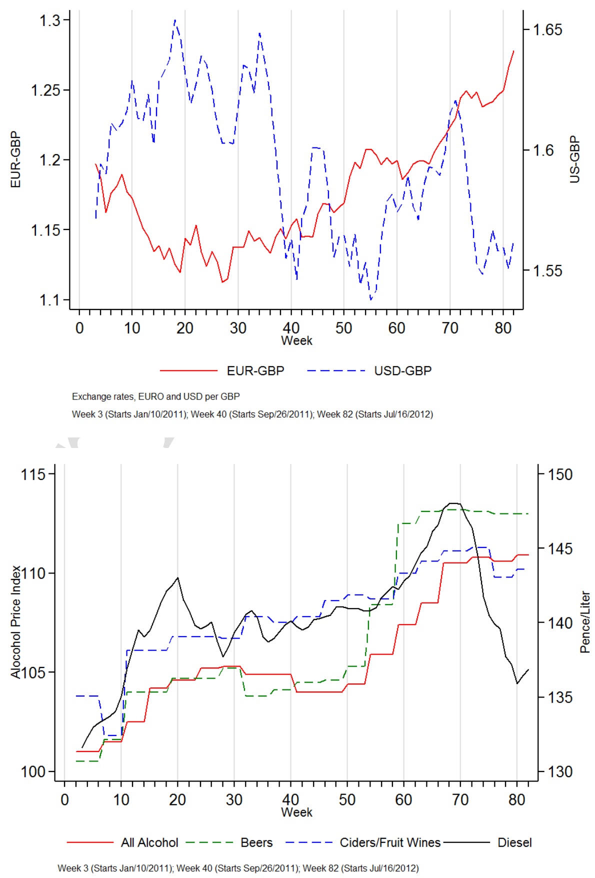


TABLE A-1. Shares of alcohol sales by stores per region

Region	Tesco	Sainsbury's	Asda	Morrisons	Discounter	Upmarket	Other
North East	0.22	0.13	0.27	0.14	0.07	0.03	0.14
North West	0.27	0.13	0.28	0.12	0.06	0.02	0.12
Yorkshire & The Humber	0.23	0.13	0.20	0.21	0.05	0.02	0.17
East Midlands	0.30	0.15	0.17	0.15	0.07	0.02	0.13
West Midlands	0.25	0.19	0.20	0.14	0.06	0.03	0.13
East of England	0.42	0.19	0.13	0.09	0.05	0.03	0.09
London	0.31	0.30	0.14	0.06	0.03	0.07	0.09
South East	0.36	0.25	0.15	0.07	0.05	0.05	0.07
South West	0.36	0.18	0.16	0.11	0.06	0.04	0.08
North Eastern Scotland	0.43	0.12	0.24	0.06	0.08	0.02	0.06
Highlands and Islands	0.42	0.00	0.26	0.09	0.06	0.01	0.17
Eastern Scotland	0.35	0.11	0.25	0.13	0.05	0.03	0.08
West Central Scotland	0.28	0.10	0.28	0.16	0.07	0.02	0.09
Southern Scotland	0.28	0.06	0.23	0.20	0.08	0.03	0.12
Wales	0.31	0.09	0.26	0.11	0.11	0.02	0.10

Shares based on weeks 1-12 purchases (Jan/1-Mar/20, 2011). Discounters are Aldi and Lidl, upmarket is Waitrose and Marks & Spencer, and others are independent stores.

Thus, our first-stage instruments consisted of exchange rates, alcohol price index, ex-factory prices for beer, and for cider/fruit wines, diesel prices, shares of store types by region, and the interactions of store shares by region with diesel prices. Second-stage equations are estimated separately for each segment, and by household type and all households combined. Each of these contains four different price variables and slightly different exogenous variables in second-stage equations. For example, for the beers segment (S02), the four endogenous variables are prices of spirits, beers, wines and FABs, and there are four such regressions by household type, hence there are a total of 16 first-stage regressions for this segment. In total, control variables were constructed from 52 separate first-stage regressions. Table A-2 provides F-tests from first-stage regression of (log) prices on all exogenous variables in the segment and for the household type, where the test is the restriction test of excluded instruments (i.e., a weak instruments test). In all cases, the test statistic is reasonably high and above the rule-of-thumb value of 10.

Table A-2. First Stage F-Test for Excluded Instruments

Sample Households		(1) All	(2) Low	(3) Medium	(4) High
Segment S00: All Drinks Combined					
ln price	Overall	44.2	32.6	20.3	14.0
Segment S01: Spirits and Fortified Wines					
ln price	Sprits	237.5	479.8	245.1	48.7
	Beers	66.6	53.5	52.4	21.3
	Wines	140.1	139.7	57.4	34.4
	FABS	609.4	5237.3	4132.3	4022.3
Segment S02: Beers and Ales					
ln price	Sprits	232.5	416.1	245.9	47.8
	Beers	67.0	54.0	53.3	21.3
	Wines	136.4	133.8	54.2	35.0
	FABS	607.7	3092.2	1948.7	2463.1
Segment S03: Wines and Bubblies					
ln price	Sprits	234.8	448.3	252.3	47.4
	Beers	66.1	53.0	53.1	21.2
	Wines	131.8	129.7	54.8	32.3
	FABS	606.7	3717.0	4683.2	4515.0

ln price regressed on instruments and exogenous variables. Regressions are by alcohol segment and by household type.

A.2. **Seasonality, further results.** This appendix provides results when changing the specification for seasonality as described in the main text. (1) Replace 4-week pseudo months with a set of dummies corresponding to the true calendar months; (2) Interact the new calendar dummies with the country dummy; (3) Same as the previous case but restrict observations from months March-July of both years only, and no observations for weeks 1-12; (4) Similar to the previous case but now retained observations from January-July of both years. The results for these four cases for the interaction term are summarized in Table A-3 in the four rows labeled 'w/ seasonality #' for all the previous six specifications for segment S00. The table reports only the interaction coefficient β_3 scaled by $\Phi(\cdot)$ (i.e., the marginal effect) for these 24 different tobit models. For ease of comparison with our initial results, the first row shows the interaction terms from the original specifications reported in Table 3.

Table A-3. alcohol purchase – marginals $\beta_3\Phi(\cdot)/(s.e.)$ with seasonality

	(1)	(2)	(3)	(4)	(5)	(6)
Sample Households	All	All	All	Low	Medium	High
$S \times B$: Scotland×PostBan	0.085***	0.086***	0.077***	-0.004	0.074***	0.195***
(original with time trends)	(0.018)	(0.018)	(0.018)	(0.023)	(0.030)	(0.035)
w/ seasonality 1	0.085***	0.085***	0.076***	-0.005	0.072***	0.198***
	(0.018)	(0.018)	(0.018)	(0.023)	(0.030)	(0.035)
w/ seasonality 2	0.092***	0.100***	0.098***	0.035	0.074***	0.224***
	(0.023)	(0.021)	(0.021)	(0.029)	(0.032)	(0.048)
w/ seasonality 3	0.092***	0.103***	0.103***	0.037	0.084***	0.226***
	(0.023)	(0.022)	(0.022)	(0.028)	(0.038)	(0.049)
w/ seasonality 4	0.100***	0.106***	0.095***	0.001	0.109***	0.202***
	(0.020)	(0.019)	(0.019)	(0.024)	(0.036)	(0.042)
March/April only	0.059*	0.070**	0.070**	-0.007	0.081*	0.161**
(no time dummies)	(0.033)	(0.032)	(0.033)	(0.037)	(0.048)	(0.069)

All regressions include household characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Columns (1,2,3) additionally contain dummy variables for the type of household (low, medium, and high) while columns (4,5,6) provide sub-analysis by HH-type. Column (1) does not contain prices, discount, or observable product characteristics. Column (2) adds prices, discount, and observable product characteristics. Column (3) adds in control variables as residuals from first-stage regressions where price is regressed on exogenous variables and additional excluded instruments. Columns (4,5,6) are similar to (3) but restrict that sample by household type. Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively, and are based on bootstrapped standard errors.

A.3. **Total expenditures, further results.** This appendix provides the marginal effects when estimating the tobit model for the log of expenditure per household adult. The regressions include a dummy for country, for post-ban period, their interaction, time trends, controls for household characteristics and product characteristics. The results reported in the paper are given below.

TABLE A-4. Expenditure (Interaction term $\times\Phi(\cdot)/(s.e.)$

	(1)	(2)	(3)	(4)
Sample Households	All	Low	Medium	High
$S \times B$: Scotland×PostBan	0.065***	-0.001	0.061***	0.160***
	(0.010)	(0.012)	(0.017)	(0.022)
Sample	594,694	197,735	198,303	198,658
Households	8,376	2,785	2,793	2,798

All regressions include household characteristics, exogenous product characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Column (1) additionally contain dummy variables for type of household (low, medium and high) while columns (2,3,4) provide sub-analysis by HH-type. Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively.

A.4. **Excluding month five data, further results.** In the section on parallel tests, in panel B of Table 6, we provide p-values for the joint significance test of the interaction terms between the country dummy and the 4-week monthly dummies after excluding data for month five. Those tests show that after excluding data for this period, the p-values are above 0.10 in 23 out of 24 cases. However, the marginal effects reported in Table 3 and Table 4 do not exclude these data. Thus, Table A-5 below shows the comparable marginal effects for the interaction terms ($\beta_k\Phi(\cdot)$) when month five is also excluded from the main analysis, and they are similar to the ones reported in the text.

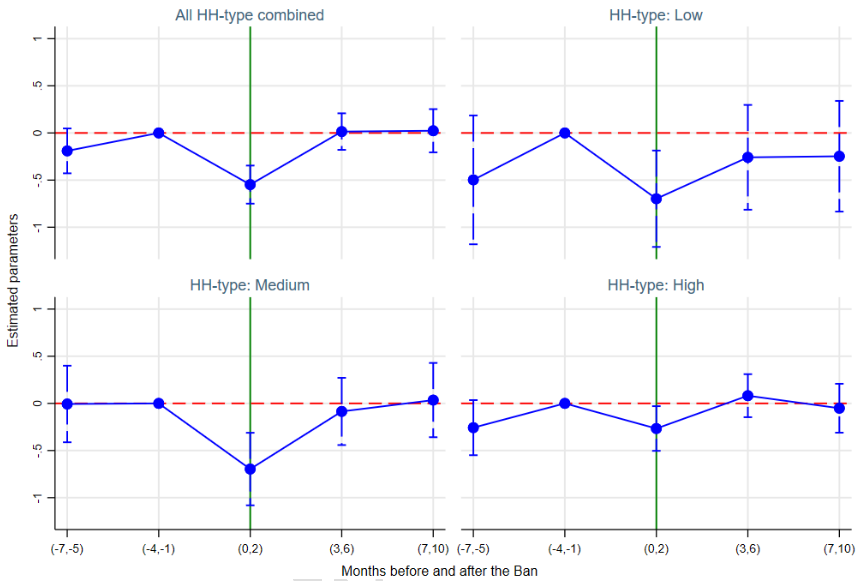
TABLE A-5. Marginals $\beta_3\Phi(\cdot)/(s.e.)$ — w/out month five data

Sample Households	(1) All	(2) All	(3) All	(4) Low	(5) Medium	(6) High
Segment S00 (All Alcohol)	0.074***	0.075***	0.067***	-0.007	0.066***	0.177***
	(0.019)	(0.018)	(0.018)	(0.022)	(0.031)	(0.036)
Segment S01 (Spirits)	-0.008	-0.014	-0.025	-0.021*	-0.052**	-0.021
	(0.013)	(0.013)	(0.015)	(0.012)	(0.022)	(0.028)
Segment S02 (Beers)	0.053***	0.045***	0.030***	-0.017	0.056***	0.063**
	(0.011)	(0.010)	(0.009)	(0.011)	(0.020)	(0.025)
Segment S03 (Wines)	0.029**	0.022*	-0.001	-0.017	-0.018	0.031
	(0.013)	(0.013)	(0.013)	(0.015)	(0.022)	(0.029)

All regressions include household characteristics, a dummy for Scotland, a dummy for pre-post ban, their interaction, and a dummy for each 4-week period. Columns (1,2,3) additionally contain dummy variables for the type of household (low, medium, and high), while columns (4,5,6) provide sub-analysis by HH-type. Column (1) does not contain prices, discount, or observable product characteristics. Column (2) adds prices, discount, and observable product characteristics. Column (3) adds control variables as residuals from first-stage regressions where price is regressed on exogenous variables and additional excluded instruments. Columns (4,5,6) are similar to (3) but restrict that sample by household type. Superscripts ***,**,* indicate significance at 1%, 5%, and 10% respectively, and are based on bootstrapped standard errors. All regressions exclude observations from (pseudo-) month five.

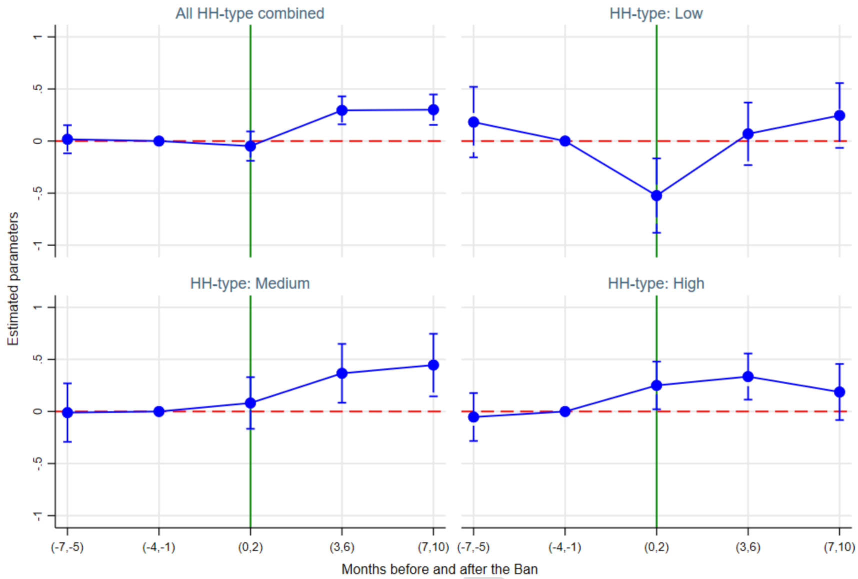
A.5. **Persistence effects – further results.** This appendix provides additional graphs for the interaction terms $S \times B$ with $\sum_j S \times B \times \rho_j$ where ρ_j is a 1/0 dummy equal to one if the observation is from the j-th period, where the period is a quarter and ranges from two quarters before to two quarters after. The main text in the paper provides the interaction terms for the all alcohol segment (S00) (Figure 3). This appendix provides similar graphs for the remaining three segments: spirits (S01), beer and ciders (S02) and wines (S03).

FIGURE A-2. Interaction terms $S \times B$ over time (Segment S01)



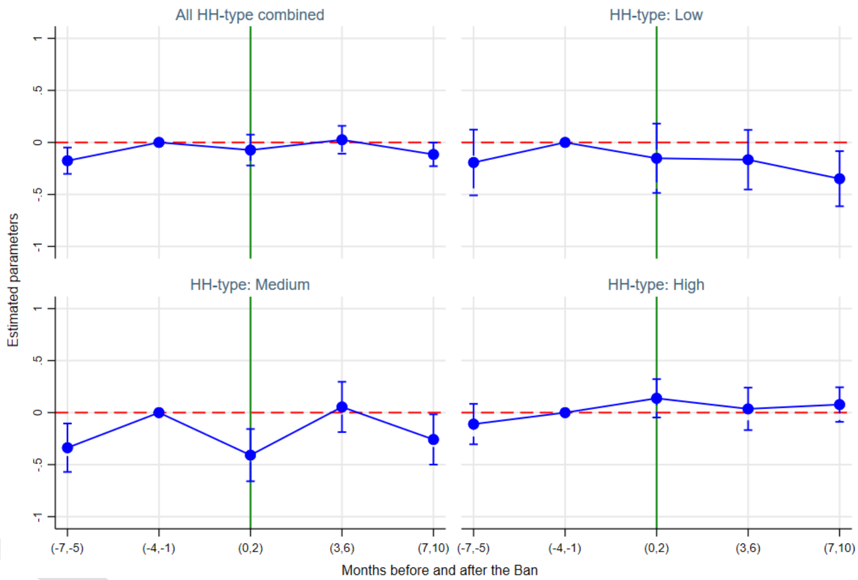
Plot of coefficients for $\sum_j S \times B \times \rho_j$ where ρ_j is a 1/0 dummy equal to one if the observation is from the j-th period. Error bars are 95% confidence intervals.

Figure A-3. Interaction terms $S \times B$ over time (Segment S02)



Plot of coefficients for $\sum_j S \times B \times \rho_j$ where ρ_j is a 1/0 dummy equal to one if the observation is from the j-th period. Error bars are 95% confidence intervals.

Figure A-4. Interaction terms $S \times B$ over time (Segment S03)



Plot of coefficients for $\sum_j S \times B \times \rho_j$ where ρ_j is a 1/0 dummy equal to one if the observation is from the j-th period. Error bars are 95% confidence intervals.