Online Appendix for: Who Pays Sin Taxes? (Conlon, Rao, Wang)

A. Comparison of Nielsen Panelists to Government Surveys

A.1. Alcoholic Beverages

An important question is whether the NielsenIQ Consumer Panelist purchase data capture similar patterns to government survey data on alcohol and tobacco consumption. In Table A1 we compare the deciles of weekly alcohol purchases according to our Nielsen Consumer Panelist data and consumption data from the NIAAA's National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). The NESARC deciles are based on the widely-cited tabulations of Cook (2007) with a key adjustment. Because aggregate alcohol consumption in the NESARC survey is roughly half of apparent consumption as measured by shipments reported by manufacturers and tracked by the Alcohol and Tobacco Tax and Trade Bureau of the U.S. Department of the Treasury (TTB), Cook (2007) inflates NESARC survey consumption by roughly a factor of two. We adjust these tabulations, deflating them so that we can compare direct survey data from these two sources. Specifically, we deflate the deciles reported in Cook (2007) by a factor of 1.97. These adjusted deciles are reported in Table A1 under NESARC.

NESARC surveys individuals and the data describes weekly drinks at the individual level. Nielsen on the other hand is panel of households. We provide two measures of weekly purchases from the Nielsen data: drinks per adult in the household and per household. The adjusted NESARC deciles are quite similar to the deciles of the NielsenIQ data. Even at the 80th and 90th percentiles the NESARC data falls between the household level and per adult level measures using the NielsenIQ data. The two main rationales for these descrepancies are: (a) the NESARC survey includes "on premise" consumption (bars and restaurants) while our data look at purchases for "off premise" consumption only; (b) we don't know who within a household consumes the drinks. This means that our per household calculations lie somewhat above the NESARC data, and our per adult calculations lie somewhat below (as this assumes equal consumption within the household). We break out purchases per adult (rather than per household) as measured in "standard drinks" per week in Figure A1. The figure reports the quantiles of "drinks per adult per week". This suggests that alcohol purchase (by ethanol units) are increasing in income, but still dominated by a small number of very heavy drinkers.

While Table A1 compares weekly purchases from the Nielsen panelist data to weekly consumption described by the NESARC survey, Table A2 compares annual average alcoholic beverage volume and liters of ethanol per household by beverage category in the Nielsen data to NIAAA data on apparent consumption, that is alcohol sales, which come from TTB data and shipments reported by manufacturers. It is well-known that aggregate consumption totals from survey responses do not match the quantity of alcohol sold in the U.S.; survey responses generally account for only half of the alcohol sold (Cook, 2007). We sum total annual beer, wine and spirits consumption in the NIAAA apparent consumption data and

	NESARC	Nielsen per adult	Nielsen Households
10%	0	0	0
20%	0	0	0
30%	0	0	0
40%	0.01	0.05	0.10
50%	0.07	0.16	0.31
60%	0.32	0.36	0.69
70%	1.10	0.76	1.42
80%	3.17	1.67	3.11
90%	7.76	4.62	8.47
max	37.49	78.68	154.56

Table A1: Alcoholic Drinks Per Week

Note: The table above reports the average number of drinks per adult aged 18 years or older per week and the average number of drinks per household per week. The number of drinks is calculated according to `https://pubs.niaaa.nih.gov/publications/practitioner/PocketGuide/pocket_guide2.htm` where a standard drink is any drink that contains about 14 grams of pure alcohol (about 0.6 fluid ounces or 1.2 tablespoons). The first column, NIAAA, is based on `https://www.washingtonpost.com/news/wonk/wp/2014/09/25/think-you-drink-a-lot-this-chart-will-tell-you/` but we divided numbers by 1.97 to recover original data. The second and third column are averages from the Nielsen data at the individual adult and household level where we use the same formula to determine standard drinks.

	Nielsen Volume	Nielsen Ethanol	NIAAA Volume	NIAAA Ethanol	Ethanol Discrepancy(%)	On-Premise (%)
Beer	19.68	0.89	188.09	8.46	89.53	23.00
Wine	6.92	0.89	26.98	3.48	74.35	18.50
Spirits	3.29	1.35	16.95	6.97	80.63	21.20
Total	29.89	3.13	232.02	18.91	83.45	

Table A2: Volume and Ethanol Consumption per Household, Nielsen versus NIAAA Data

Note: All units are in liters. To convert volume liters to ethanol liters we use ABV values of 0.045 for beer, 0.129 for wine, and 0.411 for spirits. On-premise shares reported in the final column are from Adams Media Inc. (2019). NIAAA data are from https://pubs.niaaa.nih.gov/publications/surveillance115/CONS18.htm.19 and Census data can be found at `https://fred.stlouisfed.org/series/TTLHH25`.

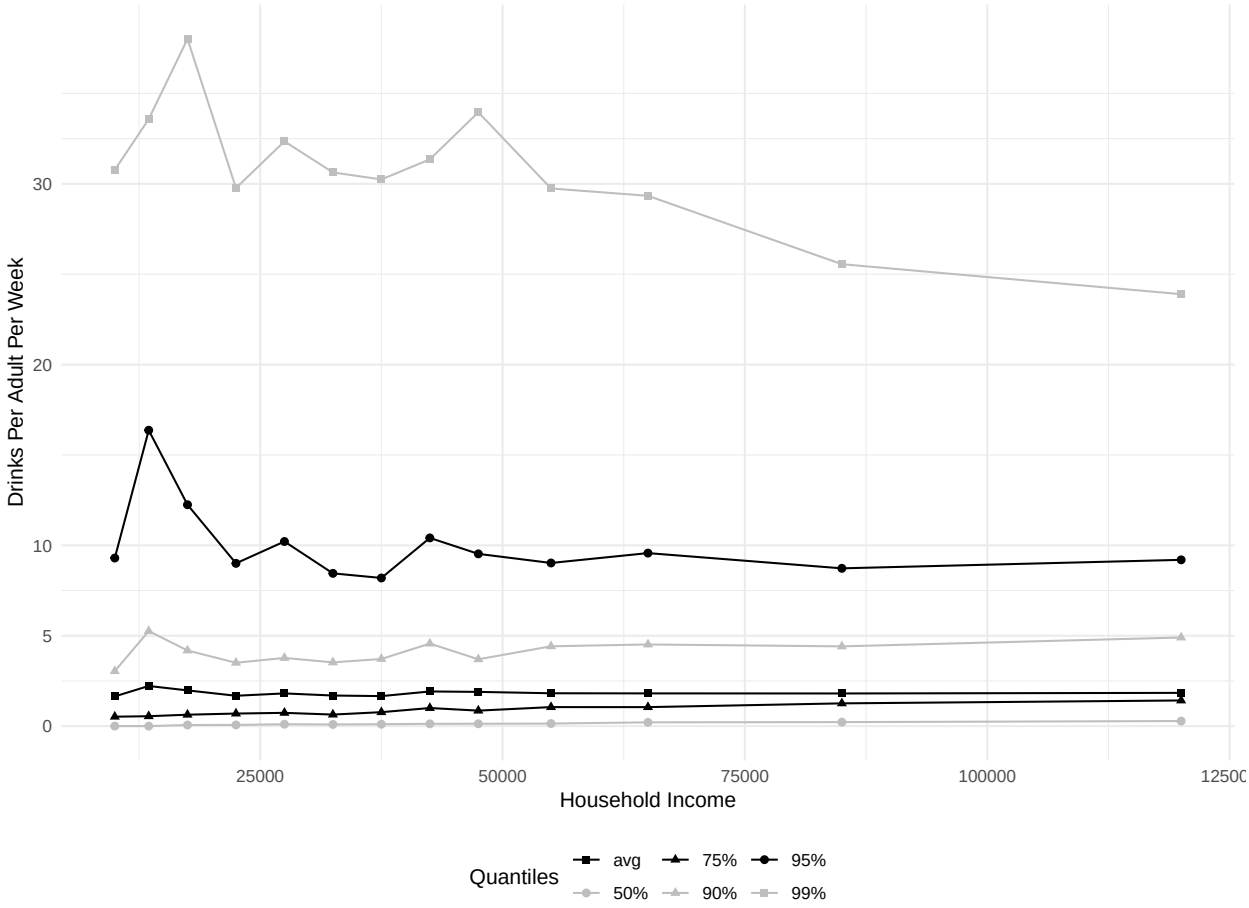

Figure A1: Alcoholic Drinks Per Adult Per Week

Source: NielsenIQ Panelist Data and authors' calculations.
These are computed on a per-household basis and then divided (equally) by household size. We convert liters of ethanol to "standard drinks" (17.7mL of ethanol).

divide by the number of U.S. households in 2018 according to the U.S. Census[1]. As NIAAA data are reported in gallons, we convert these values into liters. Similarly, we sum total annual purchases of beer, wine and spirits in the Nielsen data and scale by the number of households. We convert these volumes into ethanol liters per household using standardized alcohol by volume measures (ABV) consistent with the NIAAA data: 0.045 for beer, 0.129 for wine and 0.411 for spirits.

As we would expect, average household alcoholic beverage consumption from the NIAAA data exceeds our tabulations of average household purchases from NielsenIQ in terms of both volume and ethanol. This is especially true for beer, where average ethanol purchased per household according to the NiselsenIQ data is 89.48% lower than apparent consumption

[1]We use 127,586,000 households in all of our calculations. `https://fred.stlouisfed.org/series/TTLHH`

reported in the NIAAA data. This discrepancy is in part explained by the inclusion of on-premise consumption in the NIAAA data. Nielsen data tracks retail purchases and excludes on-premise purchases. As the last column of Table A2 shows, however, industry reports suggest that on-premise sales account for less than a quarter of alcoholic beverage sales (by volume) in any category.

A.2. Cigarettes

Finally, we compare cigarette purchases as recorded in the Nielsen data to consumption data from the Current Population Survey Tobacco Use Supplement (TUS). The TUS surveys individuals rather than households as Nielsen does. Of the 142,577 records in the TUS, 108 are invalid responses and 137,964 or 96% report not smoking at all. In the Nielsen data, 88% of households make zero annual cigarette purchases. We restrict our attention to the top few percentiles due to the small number of purchasers in both datasets. Table A3 reports individual daily cigarette consumption from the TUS data, as well as average cigarette purchases per day from Nielsen at the per adult and per household level.

In part because on-premise sales are less common for cigarettes, purchases and consumption track each other closely. Because we can only measure household purchases in the Nielsen data and not all adults in a household may smoke, our cigarettes per adult per day measure understates true daily consumption by smoking adults. As such the TUS averages generally lie between the Nielsen per adult and per household averages.

	TUS	Nielsen (Per Adult)	Nielsen (Per Household)
95%	0	0.767	1.425
96%	0	1.458	2.740
97%	4	2.822	5.205
98%	10	5.317	9.260
99%	20	10.339	16.969
max	40	52.0	197.26

Table A3: Cigarette Consumption Per Capita, Nielsen versus CPS

Note: This table compares average daily cigarette consumption according to the Current Population Survey Tobacco Use Supplement (TUS) `https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires-data#2018` to average daily Nielsen purchases per adult and household. The TUS surveys individuals while Nielsen tracks the purchases of households. The table presents the number of cigarettes smoked per day where one pack of cigarettes contains 20 cigarettes. We report only the top few percentiles of consumption because the vast majority of respondents in both dataset do not purchase cigarettes. The 2018 contains 142,577 records of which 108 respondents do not provide any valid answer and 137,964 people (about 96%) report not smoking at all. In the Nielsen data 88% households made zero (annual) cigarette purchases.

B. Tax Information

B.1. Sin Tax Rates

In the United States, alcoholic beverages and tobacco are taxed by the federal government as well as most states. Different rates typically apply to beer, wine and spirits. The difference in rates often means that spirits are taxed at a higher rate per unit of ethanol, relative to

wine and beer. For example, the federal government's tax rates for beer wine and spirits of \$0.15, \$0.28 and \$3.57 per liter translate into tax rates of \$3.33, \$2.17 and \$8.93 per liter of ethanol, respectively. Table B1 lists the federal and state tax rates for beer, wine and spirits in 2018 in terms of dollars per liter in the left panel and per liter of ethanol in the next panel. This list includes control states that often do not levy a specific sin tax on alcohol but charge a markup on alcoholic beverages which are exclusively sold in state-run stores. There is substantially heterogeneity in alcohol tax rates across states and also within states in the rates they apply to different beverages categories. For example, tax rates on beer range from \$0.01 to \$0.34 per liter. As on the federal level, state taxes per unit of ethanol are very different across categories even within a state. In most cases tax per unit ethanol is highest for spirits and lowest for beer, but there are exceptions. In Tennessee, for example, beer bears the highest state tax at \$7.57 per ethanol liter while spirits enjoy a relatively low tax \$2.91 per liter. Perhaps unsurprisingly, Tennessee is known for whiskey production.

Cigarettes are also subject to federal taxes as well as additional state taxes in all states. These taxes range from \$0.17 per pack (of 20 cigarettes) in Missouri to \$4.35 per pack in Connecticut and New York.

The final panel of Table B1 reports average sin taxes per household in each state. These averages reflect both state consumption patterns and state tax rates on alcohol and tobacco. Average sin taxes per household range considerably across states, particularly because some control states do not impose explicit sin taxes but instead apply substantial markups at state-run monopoly stores.

It is worth noting that alcohol excise tax rates in the United Kingdom, which have the been the subject of a number of recent studies, are much higher than in the United States. We provide a comparison of tax rates in the UK to those in New York and California in Table B2. The tax rates on beer are are roughly comparable, but rates on wine are around $10\times$ as large in the UK and rates on distilled spirits are around $3-4\times$ as large in the UK. Tobacco taxes in the UK are around \$2.00 per pack higher than in New York.

B.2. SSB Tax Nielsen Modules

We have to determine which products would be subjected to a potential tax on sugar-sweetened beverages. As far as we can tell, all of the implemented SSB taxes apply to total volume (rather than sugar content). Thus a 20oz bottle of moderatly sweetened ice tea is taxed at the same rate as 20oz bottle of full-sugar soda. Jurisdictions differ in which products they exempt. For example Philadelphia and Washington, DC both include diet soda in the tax, whereas other cities (Berkeley, Oakland, San Francisco, and Seattle) do not.[2]

We try as best we can to match the Berkeley, CA or Seattle, WA definitions. We include non-diet soda, sports and energy drinks, and all "juice drinks" that are not 100% juice. This means we treat the following `product_module_code`'s as being subjected to our hypothetical SSB tax: 1030, 1041,1042, and 1484, but exclude any brand name that includes the terms

[2]See `https://www.taxpolicycenter.org/briefing-book/how-do-state-and-local-soda-taxes-work` for more details.

'diet' or 'zero'. Such excluded brand code include 541289, 541308, 688343, 620855, 620862, 754017, 721725, 754017, and 688073. Untaxed SSB modules are 1553 and any module in product group 507 except for 1030, 1041,1042, and 1484 unless the brand name includes 'diet' or 'zero'.

State	Tax Rate (per L)				Tax Rate (per Ethanol L)			
	Beer	Wine	Spirits	Cigarette	Beer	Wine	Spirits	Tax/HH
FED	0.15	0.28	2.85	1.01	3.40	2.19	7.13	61.95
AL	0.28	0.45	0.72*	0.68	6.16	3.48	1.80*	63.91
AZ	0.04	0.22	0.79	2.00	0.94	1.72	1.98	75.13
AR	0.06	0.20	0.66	1.15	1.41	1.54	1.65	72.42
CA	0.05	0.05	0.87	2.87	1.17	0.41	2.18	50.24
CO	0.02	0.08	0.60	0.84	0.47	0.65	1.51	56.97
CT	0.06	0.19	1.43	4.35	1.41	1.47	3.57	85.23
DE	0.04	0.43	1.19	2.10	0.94	3.34	2.97	81.63
FL	0.13	0.59	1.72	1.34	2.82	4.61	4.29	78.33
GA	0.27	0.62	1.22	0.37	5.93	4.81	3.05	56.16
ID	0.04	0.12	0.95*	0.57	0.88	0.92	2.38*	55.59
IL	0.06	0.37	2.26	1.98	1.36	2.85	5.65	72.16
IN	0.03	0.12	0.71	1.00	0.68	0.96	1.77	60.94
IA	0.05	0.46	0.89*	1.36	1.12	3.58	2.22*	69.52
KS	0.05	0.08	0.66	1.29	1.06	0.61	1.65	38.97
KY	0.02	0.13	0.51	1.10	0.47	1.02	1.27	55.71
LA	0.11	0.20	0.80	1.08	2.37	1.55	2.00	56.21
ME	0.09	0.16	1.43*	2.00	2.05	1.23	3.57*	82.48
MD	0.02	0.11	0.40	2.00	0.53	0.82	0.99	58.93
MA	0.03	0.15	1.07	3.51	0.62	1.13	2.67	61.56
MI	0.05	0.13	0.89*	2.00	1.17	1.04	2.22*	59.54
MN	0.04	0.08	1.33	3.04	0.86	0.61	3.32	55.68
MS	0.11	0.09	0.72*	0.68	2.51	0.72	1.80*	47.30
MO	0.02	0.11	0.53	0.17	0.35	0.86	1.32	54.62
MT	0.04	0.28	0.95*	1.70	0.80	2.17	2.38*	70.06
NE	0.08	0.25	0.99	0.64	1.82	1.95	2.48	65.37
NV	0.04	0.18	0.95	1.80	0.94	1.43	2.38	80.12
NH	0.08	0.16*	1.43*	1.78	1.76	1.23*	3.57*	92.57
NJ	0.03	0.23	1.45	2.70	0.70	1.79	3.63	51.36
NM	0.11	0.45	1.60	1.66	2.41	3.48	4.00	93.62
NY	0.04	0.08	1.70	4.35	0.82	0.61	4.25	54.96
NC	0.16	0.26	0.72*	0.45	3.62	2.04	1.80*	56.04
ND	0.04	0.13	0.66	0.44	0.94	1.02	1.65	53.68
OH	0.05	0.08	0.89*	1.60	1.06	0.66	2.22*	62.77
OK	0.10	0.19	1.47	2.03	2.33	1.47	3.67	71.37
OR	0.02	0.18	0.95*	1.33	0.48	1.37	2.38*	55.21
PA	0.02	0.16*	1.43*	2.60	0.46	1.23*	3.57*	72.06
RI	0.03	0.37	1.43	4.25	0.62	2.87	3.57	66.10
SC	0.20	0.29	0.72	0.57	4.51	2.21	1.80	63.01
SD	0.07	0.25	1.04	1.53	1.58	1.90	2.60	61.41
TN	0.34	0.32	1.16	0.62	7.57	2.48	2.91	60.63
TX	0.05	0.05	0.63	1.41	1.12	0.42	1.59	57.71
UT	0.11	0.22*	0.95*	1.70	2.39	1.72*	2.38*	31.40
VT	0.07	0.15	1.43*	3.08	1.56	1.13	3.57*	94.37
VA	0.07	0.40	0.72*	0.30	1.51	3.09	1.80*	48.71
WA	0.07	0.23	3.77	3.02	1.53	1.78	9.42	76.22
WV	0.05	0.26	0.72*	1.20	1.04	2.05	1.80*	65.64
WI	0.02	0.07	0.89	2.52	0.38	0.51	2.22	83.29
WY	0.01	0.22*	0.95*	0.60	0.12	1.72*	2.38*	51.29
DC	0.02	0.08	0.40	2.50	0.53	0.61	0.99	66.59

Table B1: Federal and State Tax Rates on Sin Goods

Source: Tax Policy Center. cigarettes: `https://www.taxpolicycenter.org/statistics/state-cigarette-tax-rates`,alcohol:`https://www.taxpolicycenter.org/statistics/state-alcohol-excise-tax-rates`.
The table reports federal and state alcohol and cigarette taxes. Alcohol taxes are generally applied in terms of dollars per gallon; we convert these figures into dollars per liter terms. Cigarette taxes are reported in terms of dollars per 20-cigarette pack. Tax rates are from Tax Policy Center. For control states (where the alcohol tax rate is "n.a."), we used the regional median where the four regions are based on Census definitions

Jurisdiction	Beer	Wine	Spirits	Cigarette
CA+Federal	0.20	0.33	3.72	2.87
NY+Federal	0.19	0.36	4.55	4.35
UK	0.23	3.60	13.96	6.39

Table B2: Comparison of US (combined state and local) and UK Sin Taxes

Source: `https://www.gov.uk/tax-on-shopping/alcohol-tobacco` and authors' calculations. Tobacco in the UK also includes a 16.5% *ad valorem* tax (not included).

C. Sample Selection and Outliers

In our analysis we drop a total of 23 households because their reported per capita (per adult) purchases are higher than seems plausible for an adult to regularly consume. We drop households whose per adult purchases exceed either 10 standard drinks or 3 pack of cigarettes per day (over the course of an entire year). One standard drink contains 20mL of ethanol, meaning that 10 drinks per day totals 73L of ethanol per year. Of the 61,384 households in the sample, per capita purchases exceed these thresholds in 23 cases. These observations are excluded from the analysis. Summary stats for those households are provided in Appendix C.

Panel A: Alcohol Outliers - 20

	Beer	Spirits	Wine	Ethanol	Cigarette	Household Size	Adult	SSB	Income
Mean	1394.58	93.42	187.78	121.52	88.88	1.20	1.20	73.58	34849.50
Std	2466.86	121.36	311.52	94.79	167.22	0.41	0.41	90.10	20211.45
Min	0.00	0.00	0.00	74.74	0.00	1.00	1.00	0.00	2499.50
25%	6.39	0.00	0.56	79.39	0.00	1.00	1.00	12.04	24999.50
50%	166.51	20.25	26.63	95.77	0.00	1.00	1.00	38.31	32499.50
75%	1812.94	190.81	193.56	114.00	62.90	1.00	1.00	107.01	42499.50
Max	10830.81	407.15	925.00	505.94	461.00	2.00	2.00	330.69	84999.50

Panel B: Tobacco Outliers - 3

	Beer	Spirits	Wine	Ethanol	Cigarette	Household Size	Adult	SSB	Income
Mean	0.71	1.67	0.96	0.42	1275.00	1.00	1.00	90.58	34499.50
Std	1.23	2.89	1.66	0.39	47.70	0.00	0.00	76.83	27672.19
Min	0.00	0.00	0.00	0.00	1245.00	1.00	1.00	3.55	10999.50
25%	0.00	0.00	0.00	0.23	1247.50	1.00	1.00	61.39	19249.50
50%	0.00	0.00	0.00	0.47	1250.00	1.00	1.00	119.24	27499.50
75%	1.06	2.50	1.44	0.62	1290.00	1.00	1.00	134.10	46249.50
Max	2.13	5.00	2.88	0.78	1330.00	1.00	1.00	148.96	64999.50

Table C1: Distribution of Outliers

Note: The table above describes the 19 households that we drop from our analysis due to their implausibly high per capita purchases. We define an outlier as households whose per capita consumption exceeds 10 standard drinks or 3 packs cigarettes per day. On average, one drink contains 20 ml ethanol, and 10 drinks per day sums to 73L per year. Only 23 households of the 61,384 households in our sample exceed these purchase thresholds and are removed from the main sample.

D. Details for Calculations in Paper

This section contains alternative version of tables and figures in the main text, as well as calculations referenced in the text.

In Table D1, we calculate the average tax paid for different categories of sin goods by household income. We also report the ratio of average sin tax paid by the top income bin ($> \$100K$) to the average sin tax paid in the bottom income bin ($< \$25K$). We see that beer taxes are pretty evenly distributed across income groups. The highest income group purchase fewer SSBs and would only pay 77% as much as most other income groups (including the very poorest). This is driven mostly by the much higher purchases of diet sodas by the

highest income groups. Taxes on cigarettes are very regressive with the poorest households paying almost 3× as much as the richest. However, taxes on wine and distilled spirits appear to be strongly progressive with the richest households paying $1.87-2\times$ as much in tax as the poorest. It is important to note that this is not driven by more expensive wine and spirits purchases, as the taxes apply to volume, not revenue.

Income Bins	Beer Tax	Wine Tax	Spirits Tax	Cigarette Tax	SSB Tax	Existing Sin Taxes
*Ratio**	1.03	2.40	1.87	0.38	0.79	0.65
<24,999	4.94	2.07	8.99	60.33	37.91	76.32
25,000 - 44,999	5.27	2.53	11.13	48.31	40.01	67.23
45,000-69,999	5.50	3.12	12.59	38.75	38.54	59.96
70,000-99,999	5.58	3.81	15.08	30.43	35.52	54.90
> 100,000	5.11	4.97	16.85	22.83	30.01	49.77

Table D1: Average Tax Burden by Category and Income Level

Source: NielsenIQ Panelist data and authors' calculations.
All units are dollars per household per year.
*Ratio** divides the tax burden for households whose income >$100k by the burden for those < $25k.

We want to be careful about interpreting Table D1 as the definitive information regarding the progressivity or regressivity, which is why it is not the focus of our analysis but provided here for comparison. In Figure D1 we show why average taxes paid by income are not necessarily an ideal comparison. We plot the average annual total sin taxes paid by households against the 13 income levels provided by Nielsen.[3] If we compare the average sin taxes paid for each income level, sin taxes look highly regressive and the correlation coefficient is $\rho = -0.765$. However, if we plot the log of sin taxes paid (plus one dollar) we find that the correlation is strongly positive $\rho = 0.857$.[4] The problem is the extreme heteroskedasticity where the standard deviation of sin taxes paid is more than $210 for lower income households and only $152 for the highest income households, this and the extreme skewness of the distribution explain the discrepancy. The $\log(x)$ transform implies that going from $100 (90th percentile) to $1000 (99th percentile) in sin tax spending is a change of 2.3 log points, which is the same as the median household going from one dollar of sin tax spending to $10 (around 60th percentile).

Table D2 reports key percentiles of the tax burden distribution by race and ethnicity for existing sin taxes as well existing sin taxes and a penny-per-ounce tax on SSBs. As panels A and C show, the burden of existing sin taxes falls more heavily on white households and non-Hispanic households. Taxing SSBs increases taxes on all groups, but also alters the relative tax burdens of racial and ethnic groups. When SSBs are taxed Black households at or below

[3]We consolidate all income bins below $12,000, so that we have 13 bins instead of 16. These only constitute 6% of the population.

[4]The $\log(x+1)$ vs. $\log(x)$ transformation is not driving the result. Even after dropping the zeros, or using $arcsinh(x)$ the result is similar.

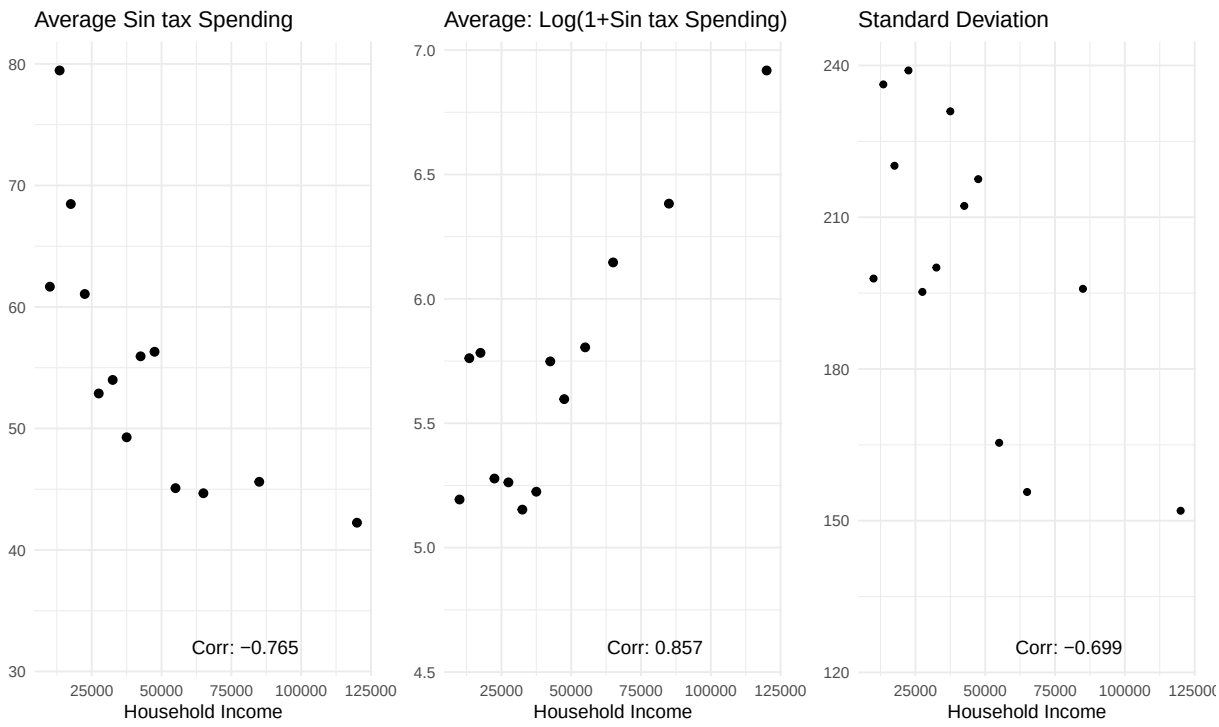

Figure D1: Sin Taxes Paid by Income (Levels vs. Logs)

Source: NielsenIQ Panelist data and authors' calculations.
Income is reported in 13 bins with all incomes below $10,000 consolidated into the first bin.

the median pay higher sin taxes than their white counterparts (Panel B). A similar pattern holds in panel C, which compares the taxes of Hispanic and non-Hispanic households when SSB taxes are included. At higher points in the tax distribution white households pay more sin taxes than Black (and non-Hispanic households pay more than Hispanic households), but the gap is consistently narrower than when just existing sin taxes are considered.

In the text we report the summary results of the regression:

$$\log(1 + \text{Sin Tax}_i) = \beta x_i + \gamma(k_i) + \lambda_{s_i} + \varepsilon_i$$

Where x_i are household demographics and λ_{s_i} are state-fixed effects. The full set of estimates is below in Table D3. The differences are the cluster fixed effects $\gamma(k_i)$ (in columns 2 and 4) and whether the hypothetical SSB tax is included in the overall sin tax burden (columns 3 and 4). The main takeaway is that the explanatory power of demographic variables (Within R^2) is very weak. The explanatory value of the state fixed effects ($R^2 = 0.0359$) is also very weak, especially considering that the statutory tax rates are fully explained by state fixed effects. On the other hand including the $k = 8$ grouped fixed effects for our clusters increases the overall $R^2 = 0.80$ in column 2.

Panel A: Existing Sin Taxes on Alcohol and Tobacco						
	tax 25%	tax 50%	tax 75%	tax 90%	tax 95%	mean
White	0.00	3.36	30.43	152.84	365.49	68.57
Black	0.00	2.10	18.43	92.47	213.97	42.13
Asian	0.00	1.19	8.68	38.35	86.41	21.40
Other	0.00	3.33	23.10	91.60	222.90	48.12
Panel B: Existing Sin Taxes + SSB Taxes						
	tax 25%	tax 50%	tax 75%	tax 90%	tax 95%	mean
White	11.52	35.35	92.19	226.69	434.93	104.33
Black	16.06	39.36	86.09	178.85	292.66	83.22
Asian	5.92	17.68	45.80	98.51	154.60	45.00
Other	14.58	38.22	86.35	178.43	310.71	87.74
Panel C: Existing Sin Taxes on Alcohol and Tobacco						
	tax 25%	tax 50%	tax 75%	tax 90%	tax 95%	mean
Hispanic	0.22	3.56	21.53	82.94	192.50	42.51
Nonhispanic	0.00	2.97	27.76	142.03	340.73	64.66
Panel D: Existing Sin Taxes + SSB Taxes						
	tax 25%	tax 50%	tax 75%	tax 90%	tax 95%	mean
Hispanic	14.87	37.34	83.14	166.18	275.86	80.87
Nonhispanic	11.52	34.92	89.76	216.94	411.08	100.65

Table D2: Sin Tax Paid by Race and Ethnicity

The obvious conclusion is that purchase patterns rather than demographics or geography are the primary source of heterogeneity. There are several demographic bins that are significant, though it is important to remember that significance is relative to our baseline household (White, Non-Hispanic, College Graduate, No Children, \$45,000-\$69,000). Also note that the income coefficients are monotonically increasing in nearly all specifications suggesting that sin taxes are progressive. This is an artifact of the $\log(\cdot)$ transform and the large amount of heteroskedasticity we discuss in Figure D1.

In Table D4 we compute the sin tax burden as a share of income for each household and then compute the appropriately weighted quantiles or mean for each of our $k = 8$ clusters. We report the weighted mean in Table 1 in the main text of the paper. The two panels differ by whether or not the hypothetical SSB taxes are included in the overall sin tax burden.

Since we observe income in bins, we use the midpoint of each bin as the income for every household within that income bin. For the lowest bin we use a value of \$10,000 and for the highest income households we use a value of \$120,000. (These are meant to approximate the median of these bins after fitting a lognormal to the overall income distribution).

As one might expect, the ratio of sin taxes to income is highly skewed both because purchases of sin taxes are highly concentrated among a small number of households and because household income is also quite skewed. Other than *Smokers, Everything* and *Heavy Drinkers* most households pay a negligible amount of income in sin taxes. Even households

in the top 5% of the other clusters rarely pay more than 0.25% of income in sin taxes. The *Everything* and *Smokers* pay a much larger share (often more than 1-2%) both because they are poorer than average and because they face much larger tax burdens – particularly since cigarette taxes are such a large share of the overall sin tax burden.

In Table D5 we report the conditional distribution of demographics given that a household belongs to a particular cluster. For example, a household assigned to our *Everything* cluster has a 5% chance of having received a postgraduate degree. In the main text Table 2, we report the ratio of 5% divided by the overall rate of postgraduates in the data (15%) for 0.33. The ratios are likely to facilitate quicker comparisons and Table 2 could be constructed from Table D5 and vice versa, so this is purely for convenience.

As an alternative to the relative risk we calculate in Table 2, we also fit a multinomial logit regression where we predict cluster assignment as a function of demographics. This has advantages and disadvantages. The main disadvantage is that we need to specify both a baseline set of consumer demographics, and a baseline cluster. This makes the results a little harder to interpret. The main advantage over the results we report in Table 2 is that it better handles the fact that many demographics are highly correlated (such as education and income) so those effects are more moderated.

In Table D7 we provide the bootstrapped confidence intervals for the relative risk in Table 2. The bootstrap procedure is quite straightforward:

1. For each household, assign it to a cluster following (1) from the main text.
2. Re-sample $N = 61{,}332$ households with replacement.
3. Compute $Pr(h \in Demog | h \in Cluster)$ as in Table D5
4. Compute $Pr(h \in Demog)$ and the ratio.
5. Repeat (2)-(4) 500 times and report the $\alpha = 0.025$ and $1 - \alpha$ quantiles.
6. Highlight cells in Table 2 if the confidence interval is strictly above 1.1 or strictly below 0.9.

	Log(Sin Tax)		Log(Sin Tax+SSB Tax)	
	(1)	(2)	(3)	(4)
Income: < 24,999	-0.156	-0.055	-0.097	-0.032
	(0.045)	(0.029)	(0.027)	(0.036)
Income: 25,000-44,999	-0.129	-0.043	-0.055	-0.014
	(0.034)	(0.017)	(0.026)	(0.022)
Income: 70,000-99,999	0.110	0.019	0.036	-0.015
	(0.035)	(0.023)	(0.029)	(0.011)
Income: > 100,000	0.254	0.026	0.109	-0.019
	(0.029)	(0.034)	(0.019)	(0.025)
Race: Asian	-0.541	-0.111	-0.519	-0.170
	(0.033)	(0.033)	(0.035)	(0.022)
Race: Black	-0.210	-0.042	0.112	0.032
	(0.047)	(0.040)	(0.036)	(0.022)
Race: Other	-0.041	0.015	0.029	0.006
	(0.035)	(0.005)	(0.037)	(0.013)
Hispanic: Yes	-0.081	-0.052	-0.008	-0.040
	(0.041)	(0.040)	(0.041)	(0.021)
Children: Yes	-0.180	-0.060	0.294	0.132
	(0.033)	(0.036)	(0.023)	(0.053)
Edu: High School or Less	0.132	-0.006	0.362	0.141
	(0.034)	(0.024)	(0.021)	(0.052)
Edu: Some College	0.165	0.029	0.246	0.074
	(0.036)	(0.011)	(0.023)	(0.021)
Edu: Post College Grad	-0.132	-0.020	-0.272	-0.109
	(0.032)	(0.025)	(0.020)	(0.025)
Age: Under 35	-0.001	-0.010	-0.027	-0.041
	(0.043)	(0.100)	(0.031)	(0.087)
Age: 35-44	0.062	-0.010	0.131	0.023
	(0.046)	(0.085)	(0.029)	(0.067)
Age: 45-54	0.179	0.009	0.277	0.082
	(0.040)	(0.062)	(0.030)	(0.058)
Age: 55-64	0.203	0.037	0.231	0.069
	(0.035)	(0.028)	(0.026)	(0.022)
Standard-Errors	State	State & Cluster Assignment	State	State & Cluster Assignment
Observations	61,332	61,332	61,332	61,332
R^2	0.0359	0.8064	0.0532	0.6716
Within R^2	0.0171	0.0048	0.0399	0.0167
State FE	✓	✓	✓	✓
Cluster Assignment FE		✓		✓

Table D3: Sin Tax Burden with and without Grouped Fixed Effects

Note: The omitted base demographics are for a household that is Race: White, Education: Graduated College, Hispanic: No, Age: over 65, Children at Home: No; Income: 45,000-69,999.
Observations are weighted by Nielsen projection factors.

Panel A: Existing Sin Taxes on Alcohol and Tobacco

	Everything	Smokers	Heavy Drinkers	Moderate Spirits	Mostly Wine	Moderate Beer	SSB only	Nothing
min	0.018	0.005	0.009	0.003	0.001	0.001	0.000	0.000
ratio 25%	0.278	0.128	0.068	0.023	0.013	0.008	0.000	0.000
ratio 50%	0.730	0.444	0.140	0.045	0.028	0.017	0.000	0.000
ratio 75%	1.860	1.483	0.332	0.100	0.065	0.046	0.003	0.005
ratio 90%	4.138	3.752	0.740	0.238	0.155	0.125	0.010	0.018
ratio 95%	6.756	6.159	1.225	0.436	0.279	0.240	0.019	0.039
mean	1.882	1.541	0.347	0.123	0.082	0.063	0.005	0.011
max	155.155	114.796	39.463	19.633	11.259	22.706	1.113	4.412

Panel B: Existing Sin Taxes Plus New Taxes on SSBs

	Everything	Smokers	Heavy Drinkers	Moderate Spirits	Mostly Wine	Moderate Beer	SSB only	Nothing
min	0.022	0.007	0.013	0.006	0.002	0.004	0.001	0.000
ratio 25%	0.374	0.251	0.094	0.047	0.025	0.034	0.022	0.001
ratio 50%	0.895	0.680	0.186	0.096	0.050	0.075	0.054	0.004
ratio 75%	2.109	1.862	0.411	0.212	0.112	0.179	0.140	0.011
ratio 90%	4.542	4.402	0.878	0.462	0.245	0.406	0.341	0.027
ratio 95%	7.282	7.021	1.429	0.788	0.417	0.687	0.594	0.051
mean	2.099	1.861	0.426	0.236	0.127	0.192	0.164	0.015
max	155.155	124.276	39.825	55.554	12.777	25.266	59.279	4.412

Panel A1: Existing Sin Taxes with potential $1 increase in tobacco tax

	Everything	Smokers	Heavy Drinkers	Moderate Spirits	Mostly Wine	Moderate Beer	SSB only	Nothing
min	0.022	0.009	0.009	0.003	0.001	0.001	0.000	0.000
ratio 25%	0.357	0.188	0.068	0.023	0.014	0.008	0.000	0.000
ratio 50%	0.967	0.656	0.140	0.045	0.028	0.018	0.000	0.000
ratio 75%	2.514	2.180	0.334	0.101	0.066	0.047	0.003	0.005
ratio 90%	5.613	5.539	0.743	0.241	0.158	0.130	0.011	0.019
ratio 95%	9.134	8.892	1.231	0.443	0.287	0.251	0.022	0.041
mean	2.533	2.239	0.348	0.124	0.084	0.065	0.005	0.012
max	200.404	185.370	39.463	19.633	13.678	22.706	1.393	6.173

Panel B1: Existing Sin Taxes Plus New Taxes on SSBs with potential $1 increase in tobacco tax

	Everything	Smokers	Heavy Drinkers	Moderate Spirits	Mostly Wine	Moderate Beer	SSB only	Nothing
min	0.026	0.012	0.013	0.006	0.002	0.004	0.001	0.000
ratio 25%	0.459	0.322	0.094	0.048	0.025	0.034	0.022	0.001
ratio 50%	1.131	0.896	0.186	0.096	0.050	0.076	0.055	0.004
ratio 75%	2.764	2.567	0.413	0.214	0.113	0.181	0.141	0.011
ratio 90%	6.022	6.119	0.882	0.465	0.248	0.411	0.342	0.028
ratio 95%	9.738	9.723	1.431	0.791	0.421	0.693	0.595	0.053
mean	2.750	2.560	0.427	0.237	0.129	0.194	0.165	0.016
max	200.404	194.850	39.825	55.554	14.977	25.266	59.279	6.173

Table D4: Distribution of Sin Tax to Income Ratios (%)

	Everything	Smokers	Heavy Drinkers	Mostly Wine	Moderate Spirits	Moderate Beer	SSB only	Nothing
Race: White (74.9%)	0.78	0.83	0.80	0.69	0.80	0.76	0.72	0.78
Race: Black (12.5%)	0.12	0.10	0.09	0.16	0.10	0.10	0.16	0.08
Race: Asian (4.4%)	0.03	0.02	0.03	0.04	0.04	0.03	0.04	0.07
Race: Other (8.2%)	0.07	0.05	0.08	0.11	0.07	0.10	0.09	0.07
Hispanic: No (86.8%)	0.89	0.92	0.87	0.85	0.87	0.81	0.86	0.90
Hispanic: Yes (13.2%)	0.11	0.08	0.13	0.15	0.13	0.19	0.14	0.10
Children: Yes (31.3%)	0.22	0.30	0.26	0.34	0.27	0.37	0.37	0.18
Children: No (68.7%)	0.78	0.70	0.74	0.66	0.73	0.63	0.63	0.82
Age: $<$ 35 (12.9%)	0.08	0.08	0.10	0.14	0.13	0.14	0.13	0.14
Age: 35 to 44 (18.0%)	0.13	0.17	0.15	0.19	0.17	0.21	0.20	0.14
Age: 45 to 54 (21.8%)	0.23	0.25	0.23	0.26	0.19	0.24	0.22	0.17
Age: 55 to 64 (22.7%)	0.33	0.28	0.25	0.22	0.22	0.22	0.21	0.23
Age: $>$ 65 (24.6%)	0.23	0.22	0.26	0.19	0.30	0.20	0.23	0.31
Income: $<$ 24,999 (20.4%)	0.28	0.38	0.09	0.15	0.11	0.15	0.22	0.23
Income: 25,000 - 44,999 (17.7%)	0.21	0.21	0.12	0.16	0.13	0.16	0.19	0.18
Income: 45,000-69,999 (18.2%)	0.18	0.18	0.17	0.18	0.17	0.19	0.19	0.18
Income: 70,000-99,999 (15.5%)	0.14	0.09	0.19	0.17	0.18	0.19	0.15	0.14
Income: $>$ 100,000 (28.1%)	0.18	0.14	0.43	0.34	0.41	0.31	0.25	0.27
Edu: High School or less (27.4%)	0.36	0.45	0.21	0.23	0.18	0.28	0.29	0.24
Edu: Some College (31.4%)	0.41	0.36	0.31	0.35	0.26	0.32	0.32	0.27
Edu: Graduated College (26.3%)	0.18	0.14	0.30	0.28	0.32	0.28	0.25	0.28
Edu: Post College Grad (14.9%)	0.05	0.04	0.18	0.14	0.24	0.13	0.13	0.20

Table D5: Share of households with certain demographics by cluster: $Pr(h \in Demog | h \in Cluster)$

	cluster = Everything		**cluster = Smokers**		**cluster = Heavy Drinkers**		**cluster = Moderate Spirits**	
	coef	std err	coef	std err	coef	std err	coef	std err
Intercept	-2.4002	0.086	-1.8103	0.066	-0.6951	0.051	-1.1415	0.054
Edu: High School or Less	0.8314	0.082	0.9944	0.061	0.1193	0.059	0.2703	0.059
Edu: Post College Grad	-0.8504	0.108	-0.7952	0.082	-0.4063	0.048	-0.4951	0.051
Edu: Some College	0.6750	0.072	0.6262	0.056	0.1495	0.047	0.2596	0.047
Race: Black	0.4988	0.093	0.2452	0.076	0.0971	0.068	0.7027	0.059
Race: Asian	-0.9562	0.206	-0.9968	0.153	-1.2282	0.112	-0.6916	0.095
Race: Other	0.3199	0.133	-0.0704	0.113	-0.0023	0.093	0.2332	0.087
Hispanic: Yes	-0.0212	0.130	-0.1186	0.103	0.1342	0.081	0.1986	0.077
Age: 35 - 44	0.3184	0.111	0.3551	0.080	-0.2591	0.068	0.2891	0.066
Age: 45 - 54	0.6845	0.085	0.6626	0.064	-0.1025	0.056	0.4795	0.056
Age: 55 - 64	0.5662	0.072	0.5006	0.055	-0.0474	0.045	0.2353	0.049
Age: Under 35	-0.1903	0.155	-0.3062	0.113	-0.4749	0.086	0.1844	0.078
Children: Yes	0.2122	0.086	0.7563	0.062	0.2949	0.056	0.6250	0.051
Income: $<$ 24,999	-0.3971	0.088	0.0491	0.063	-1.2542	0.078	-0.7966	0.067
Income: 25,000 - 44,999	-0.3815	0.080	-0.2252	0.060	-0.6441	0.058	-0.5645	0.056
Income: 70,000 - 99,999	-0.2857	0.086	-0.6282	0.072	0.1508	0.053	-0.0484	0.054
Income: $>$ 100,000	-0.4351	0.099	-0.7032	0.081	0.4627	0.054	0.1373	0.055

	cluster = Mostly Wine		**cluster = Moderate Beer**		**cluster = SSB only**	
	coef	std err	coef	std err	coef	std err
Intercept	-0.4361	0.048	-0.8420	0.049	0.6703	0.034
Edu: High School or Less	-0.2491	0.061	0.5403	0.051	0.4373	0.036
Edu: Post College Grad	-0.1335	0.044	-0.5765	0.047	-0.3508	0.031
Edu: Some College	-0.0756	0.046	0.2792	0.043	0.2252	0.031
Race: Black	0.2077	0.063	0.2606	0.060	0.7081	0.042
Race: Asian	-0.9303	0.093	-0.8447	0.089	-0.4679	0.053
Race: Other	-0.0358	0.089	-0.0020	0.081	0.1751	0.059
Hispanic: Yes	0.2465	0.076	0.5328	0.067	0.1179	0.053
Age: 35 - 44	-0.3206	0.063	0.2543	0.060	0.1258	0.042
Age: 45 - 54	-0.3113	0.054	0.3333	0.052	0.2574	0.036
Age: 55 - 64	-0.2832	0.044	0.2048	0.044	0.1126	0.030
Age: Under 35	-0.4764	0.077	0.1047	0.071	-0.0928	0.049
Children: Yes	0.3685	0.053	0.7894	0.047	0.9368	0.035
Income: $<$ 24,999	-0.9614	0.070	-0.9177	0.062	-0.3890	0.039
Income: 25,000 - 44,999	-0.4958	0.054	-0.5729	0.051	-0.2883	0.035
Income: 70,000 - 99,999	0.0789	0.051	0.0169	0.049	-0.1624	0.036
Income: $>$ 100,000	0.2744	0.052	0.0728	0.051	-0.2022	0.038

Table D6: Multinomial Logit Regression

	Everything	Smokers	Heavy Drinkers	Moderate Spirits	Mostly Wine	Moderate Beer	SSB only	Nothing
Race: White (74.9%)	(1.01, 1.08)	(1.08, 1.13)	(1.04, 1.09)	(0.9, 0.95)	(1.04, 1.09)	(0.99, 1.04)	(0.95, 0.96)	(1.03, 1.06)
Race: Black (12.5%)	(0.77, 1.14)	(0.72, 0.96)	(0.65, 0.84)	(1.15, 1.39)	(0.7, 0.9)	(0.75, 0.92)	(1.2, 1.27)	(0.57, 0.67)
Race: Asian (4.4%)	(0.3, 0.84)	(0.25, 0.55)	(0.5, 0.85)	(0.71, 1.09)	(0.64, 1.0)	(0.58, 0.87)	(0.92, 1.04)	(1.54, 1.81)
Race: Other (8.2%)	(0.65, 1.14)	(0.52, 0.79)	(0.8, 1.13)	(1.16, 1.51)	(0.69, 0.97)	(1.13, 1.42)	(1.0, 1.1)	(0.72, 0.89)
Hispanic: No (86.8%)	(1.0, 1.06)	(1.04, 1.07)	(0.99, 1.03)	(0.96, 1.0)	(0.98, 1.01)	(0.91, 0.95)	(0.99, 1.0)	(1.03, 1.05)
Hispanic: Yes (13.2%)	(0.63, 1.03)	(0.51, 0.75)	(0.82, 1.07)	(1.01, 1.26)	(0.91, 1.15)	(1.35, 1.58)	(1.0, 1.07)	(0.7, 0.83)
Children: Yes (31.3%)	(0.6, 0.79)	(0.87, 1.02)	(0.77, 0.89)	(1.03, 1.16)	(0.8, 0.92)	(1.14, 1.26)	(1.17, 1.2)	(0.55, 0.61)
Children: No (68.7%)	(1.09, 1.18)	(0.99, 1.06)	(1.05, 1.11)	(0.93, 0.99)	(1.04, 1.09)	(0.88, 0.94)	(0.91, 0.92)	(1.17, 1.21)
Age: $<$ 35 (12.9%)	(0.45, 0.8)	(0.52, 0.76)	(0.67, 0.9)	(0.97, 1.21)	(0.87, 1.1)	(0.99, 1.22)	(1.0, 1.07)	(1.01, 1.16)
Age: 35 to 44 (18.0%)	(0.6, 0.88)	(0.82, 1.03)	(0.76, 0.94)	(0.97, 1.15)	(0.84, 1.01)	(1.07, 1.23)	(1.08, 1.13)	(0.74, 0.83)
Age: 45 to 54 (21.8%)	(0.93, 1.21)	(1.06, 1.25)	(0.97, 1.13)	(1.09, 1.26)	(0.81, 0.96)	(1.01, 1.14)	(1.0, 1.05)	(0.76, 0.84)
Age: 55 to 64 (22.7%)	(1.3, 1.6)	(1.13, 1.31)	(1.04, 1.19)	(0.91, 1.05)	(0.89, 1.02)	(0.9, 1.02)	(0.92, 0.97)	(0.98, 1.06)
Age: $>$ 65 (24.6%)	(0.81, 1.03)	(0.83, 0.98)	(1.01, 1.14)	(0.72, 0.85)	(1.14, 1.27)	(0.75, 0.86)	(0.92, 0.96)	(1.23, 1.32)
Income: $<$ 24,999 (20.4%)	(1.21, 1.53)	(1.76, 1.99)	(0.36, 0.48)	(0.67, 0.82)	(0.47, 0.59)	(0.68, 0.82)	(1.04, 1.09)	(1.09, 1.2)
Income: 25,000 - 44,999 (17.7%)	(1.06, 1.35)	(1.09, 1.3)	(0.63, 0.78)	(0.8, 0.97)	(0.68, 0.82)	(0.85, 0.99)	(1.06, 1.11)	(0.95, 1.05)
Income: 45,000-69,999 (18.2%)	(0.87, 1.14)	(0.87, 1.06)	(0.85, 1.01)	(0.9, 1.07)	(0.86, 1.01)	(0.99, 1.13)	(1.01, 1.06)	(0.91, 1.01)
Income: 70,000-99,999 (15.5%)	(0.78, 1.08)	(0.5, 0.66)	(1.13, 1.32)	(1.03, 1.21)	(1.07, 1.26)	(1.11, 1.28)	(0.94, 1.0)	(0.86, 0.97)
Income: $>$ 100,000 (28.1%)	(0.53, 0.75)	(0.44, 0.57)	(1.45, 1.6)	(1.14, 1.28)	(1.39, 1.53)	(1.03, 1.15)	(0.87, 0.91)	(0.93, 1.01)
Edu: High School or less (27.4%)	(1.17, 1.43)	(1.55, 1.72)	(0.69, 0.82)	(0.76, 0.9)	(0.6, 0.72)	(0.94, 1.06)	(1.04, 1.09)	(0.86, 0.94)
Edu: Some College (31.4%)	(1.21, 1.43)	(1.08, 1.23)	(0.93, 1.05)	(1.05, 1.17)	(0.78, 0.9)	(0.95, 1.06)	(1.01, 1.04)	(0.83, 0.9)
Edu: Graduated College (26.3%)	(0.6, 0.8)	(0.49, 0.61)	(1.08, 1.22)	(1.0, 1.13)	(1.14, 1.28)	(1.0, 1.11)	(0.94, 0.98)	(1.03, 1.12)
Edu: Post College Grad (14.9%)	(0.24, 0.41)	(0.24, 0.36)	(1.09, 1.3)	(0.86, 1.05)	(1.49, 1.7)	(0.8, 0.96)	(0.87, 0.93)	(1.28, 1.4)

Table D7: Relative Risk - Bootstrap Result

E. Alternative Specifications

In Table E1, and Table E2 we consider robustness to our choice of cluster assignments in the main text $k = 7$ and $k = 9$ (instead of $k = 8$). We did not perform "hierarchical clustering" where cluster assignments are nested inside one another by construction. Instead, we perform standard k-means with a different choice of k. However, we can see that with $k = 7$ we lose the ability to separate the *Moderate Spirits* drinkers and with $k = 9$ we further separate an additional group of non-smoking *Heavy Beer* drinkers (mostly taken from the *Heavy Drinkers* and *Mostly Beer* drinkers).

	Everything	Heavy Drinker	Moderate Beer	Mostly Wine	Nothing	SSB only	Smoker
Beer (mean)	149.08	77.83	60.08	20.05	2.63	0.71	2.27
Wine (mean)	21.34	29.01	1.64	38.22	0.99	0.80	0.62
Spirits (mean)	18.45	26.67	1.66	1.70	0.45	0.33	0.47
Tobacco (mean)	146.88	0.31	0.35	0.37	0.20	0.11	131.10
SSB (mean)	151.84	95.02	121.98	53.12	3.80	111.94	175.90
Ethanol (mean)	15.42	16.04	3.32	6.34	0.37	0.22	0.31
Beer 50%	46.84	25.55	21.87	4.26	0.00	0.00	0.00
Beer 75%	187.19	78.78	51.10	15.55	0.59	0.00	1.42
Beer 95%	569.72	325.03	240.61	77.71	12.78	4.26	12.78
Wine 50%	2.25	9.62	0.75	17.52	0.00	0.00	0.00
Wine 75%	12.22	27.50	2.62	38.25	0.75	0.75	0.00
Wine 95%	114.57	128.02	6.37	135.75	5.25	4.50	3.75
Spirits 50%	5.31	16.00	0.00	0.75	0.00	0.00	0.00
Spirits 75%	17.59	31.30	2.50	2.63	0.00	0.00	0.00
Spirits 95%	93.02	86.54	7.14	6.57	2.53	2.25	2.97
Tobacco 50%	75.00	0.00	0.00	0.00	0.00	0.00	60.00
Tobacco 75%	205.00	0.00	0.00	0.00	0.00	0.00	178.00
Tobacco 95%	500.00	2.00	2.00	1.00	0.00	0.20	464.95
SSB 50%	87.58	53.26	73.40	26.79	2.84	65.25	101.21
SSB 75%	207.76	126.57	154.69	64.82	6.62	139.45	245.02
SSB 95%	519.69	328.53	393.71	189.63	10.74	360.12	573.56
SSB per Person/Week	1.35	0.80	0.98	0.47	0.05	0.99	1.55
Drinks per Week	16.71	17.38	3.60	6.87	0.40	0.23	0.34
Drinks per Adult	9.98	10.05	2.10	4.26	0.28	0.13	0.19
Effective Ethanol Tax/L	8.19	10.10	7.57	5.56	3.62	3.68	4.24
Total Tax Share	25.82	23.17	4.97	5.36	1.29	1.99	37.40
Alcohol Tax Share	14.60	54.51	11.29	12.31	2.49	4.09	0.71
Tobacco Tax Share	34.02	0.26	0.36	0.27	0.41	0.46	64.22
SSB Tax Share	4.34	8.64	14.92	4.85	0.77	55.85	10.63
Tax Burden/Income (%)	2.28	0.39	0.23	0.11	0.01	0.14	1.87
# Households	1433	5437	6622	5784	11958	27262	2836
Share of Households	2.61	8.30	11.17	8.34	18.47	45.58	5.52

Table E1: 7 clusters

	Everything	Heavy Beer	Heavy Drinker	Moderate Beer	Moderate Spirits	Mostly Wine	Nothing	SSB only	Smoker
Beer (mean)	147.48	180.24	78.70	11.36	8.20	14.94	2.23	0.10	2.41
Wine (mean)	21.56	2.77	45.60	0.99	2.51	38.42	0.98	0.83	0.63
Spirits (mean)	19.88	2.33	29.22	0.25	11.08	1.21	0.34	0.14	0.42
Tobacco (mean)	158.09	0.66	0.44	0.20	0.44	0.37	0.22	0.10	131.81
SSB (mean)	150.88	104.97	87.78	114.91	115.27	52.64	3.02	107.76	176.39
Ethanol (mean)	15.81	9.14	19.28	0.70	3.86	5.99	0.32	0.15	0.30
Beer 50%	46.72	98.66	33.48	8.23	4.12	6.10	0.00	0.00	0.00
Beer 75%	181.48	204.41	84.47	14.91	11.88	17.03	0.00	0.00	1.42
Beer 95%	572.29	585.84	317.10	33.36	31.91	55.36	12.78	0.89	13.84
Wine 50%	3.00	0.88	19.88	0.00	1.50	18.75	0.00	0.00	0.00
Wine 75%	13.13	3.75	47.42	1.50	3.75	38.25	0.75	0.75	0.00
Wine 95%	108.09	10.50	171.65	4.50	9.00	135.00	5.25	4.50	3.75
Spirits 50%	5.88	0.75	17.63	0.00	6.51	0.20	0.00	0.00	0.00
Spirits 75%	18.57	3.00	36.31	0.00	11.40	1.75	0.00	0.00	0.00
Spirits 95%	94.61	9.50	92.75	1.75	33.25	4.72	1.88	1.12	2.62
Tobacco 50%	84.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	60.00
Tobacco 75%	230.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	178.87
Tobacco 95%	532.28	4.00	2.00	1.00	3.00	1.00	0.00	0.00	467.87
SSB 50%	89.44	58.87	45.92	69.36	67.77	26.00	2.00	60.80	102.19
SSB 75%	206.53	139.28	116.58	146.33	146.54	63.57	5.39	132.71	244.94
SSB 95%	521.87	376.07	307.95	373.51	358.78	193.81	8.81	354.67	579.20
SSB per Person/Week	1.36	0.86	0.74	0.94	0.94	0.45	0.04	0.97	1.54
Drinks per Week	17.13	9.91	20.89	0.76	4.18	6.50	0.35	0.16	0.33
Drinks per Adult	10.21	5.96	11.85	0.42	2.59	4.04	0.24	0.09	0.19
Effective Ethanol Tax/L	8.28	5.97	8.85	6.02	15.35	5.10	3.31	2.65	4.16
Total Tax Share	25.36	4.78	18.28	1.10	6.78	4.27	1.00	1.00	37.43
Alcohol Tax Share	14.06	10.94	42.92	2.30	15.63	9.79	1.80	1.88	0.67
Tobacco Tax Share	33.62	0.28	0.27	0.21	0.31	0.23	0.42	0.36	64.30
SSB Tax Share	3.97	5.14	5.54	14.82	8.74	4.25	0.54	46.38	10.62
Tax Burden/Income (%)	2.40	0.27	0.40	0.13	0.31	0.10	0.01	0.14	1.88
# Households	1337	2754	3853	6814	4164	5119	10657	23815	2819
Share of Households	2.40	4.47	5.76	11.78	6.93	7.38	16.46	39.32	5.50

Table E2: 9 clusters

References

Adams Media Inc. (2019) "Liquor Handbook 2019," in *Adams/Jobson's liquor handbook.*, New York, NY :: Adams Media Inc.

Cook, Philip J. (2007) *Paying the Tab: The Costs and Benefits of Alcohol Control*, Princeton, NJ: Princeton University Press.