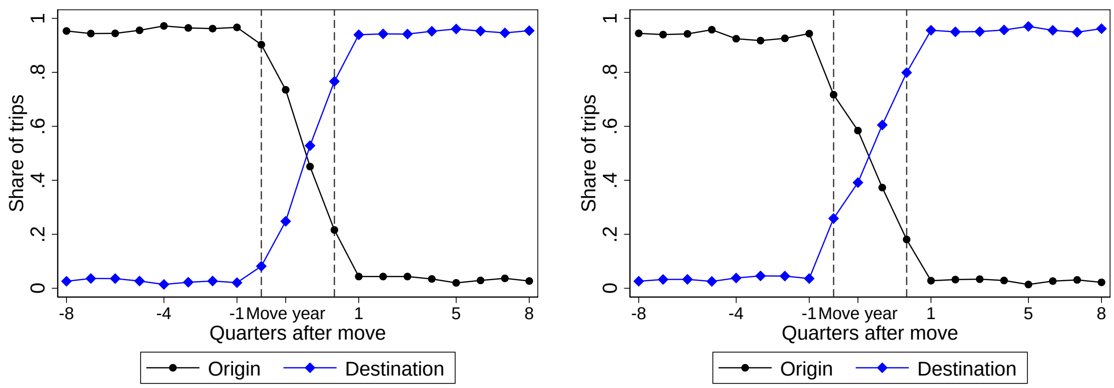
A Online appendix: Additional figures and tables

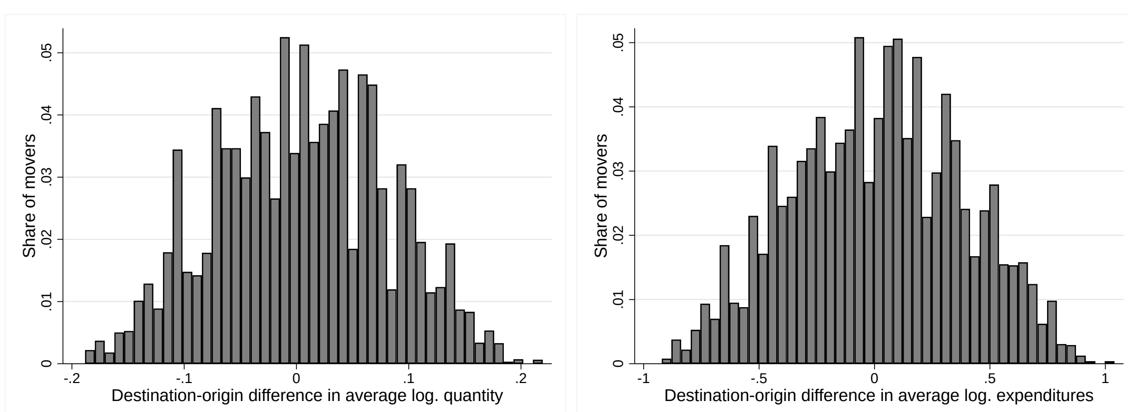


(a) Balanced panel

(b) Unbalanced panel

Figure A.1: Share of shopping trips to the destination versus origin state of residence

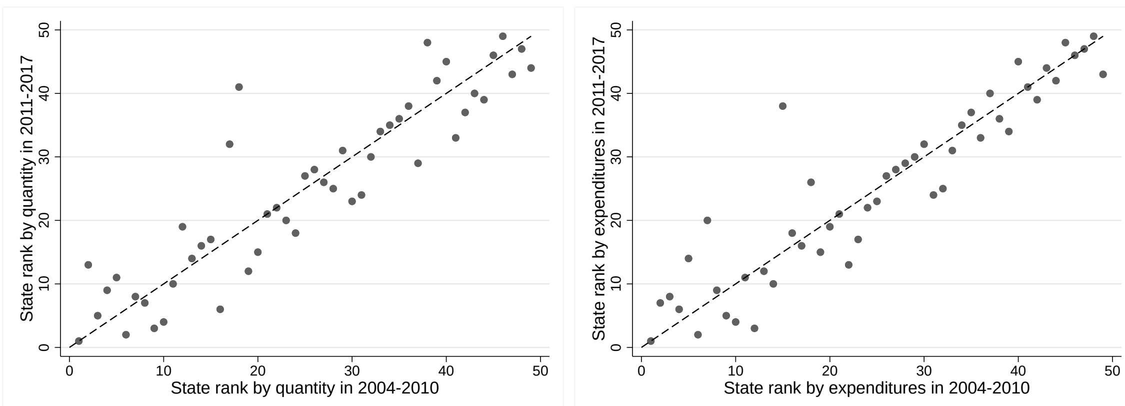
Notes: The shares are calculated using all trips to stores in NielsenIQ Retail Scanner Data. The shares do not necessarily sum to one because there is a very small share of shopping trips in states other than the origin or destination. The sample includes all movers that move across state lines, limiting the sample to those that move only once. On figure A.1a, the sample is further restricted to those who are observed continuously 2 years before and after the move year, while on figure A.1b, the sample includes the movers who are not observed continuously around the move.



(a) Quantity

(b) Expenditures

Figure A.2: Distribution of destination-origin difference in the logarithm of alcohol purchases

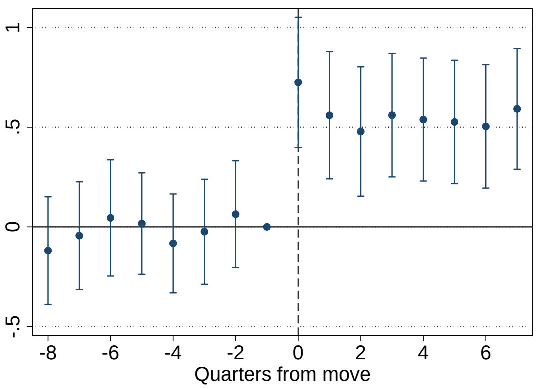


(a) Quantity

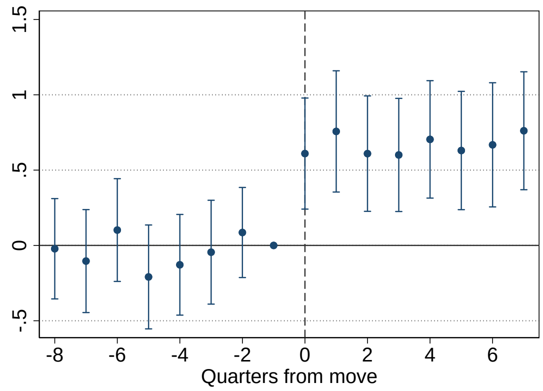
(b) Expenditures

Figure A.3: State ranked by average alcohol purchases before and after 2010

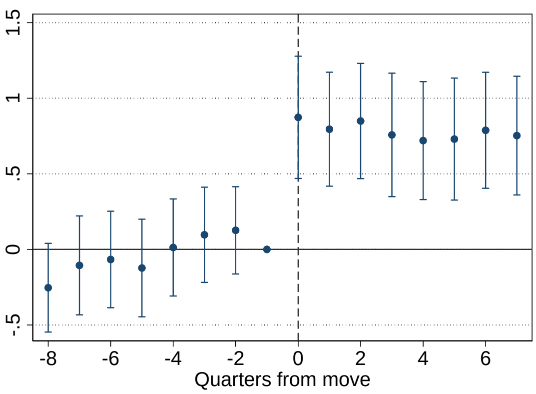
Notes: States ranked by average alcohol purchases of nonmovers in 2004-2010 vs 2011-2017. The dashed line is a 45-degree line.



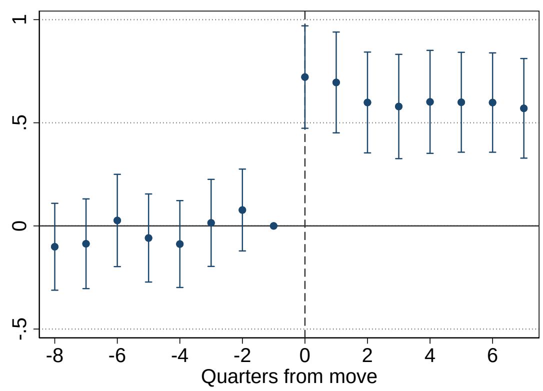
(a) Quantity of beer



(b) Quantity of liquor



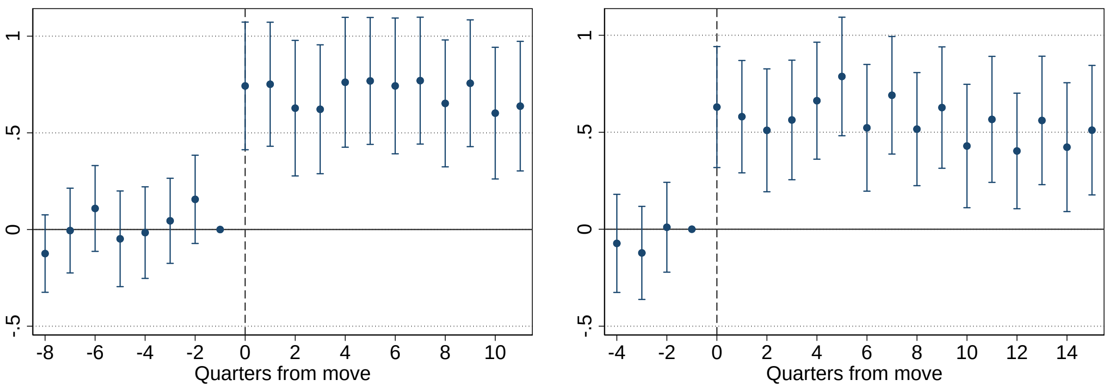
(c) Quantity of wine



(d) Alcohol expenditures

Figure A.4: Event study of alcohol purchases

Notes: Each figure presents the coefficients $\theta_{r(i,t)}$ (point estimates and 95% confidence intervals) estimated from equation (1). The coefficient for the last time period before the move is normalized to 0. The dependent variable is the logarithm of quantity of beer, liquor, or wine purchased (figures A.4a–A.4c), or alcohol expenditures (figure A.4d). Each regression includes quarter-year dummies and household fixed effects. Sample includes all movers that move across state lines, limiting the sample to those that move only once and who are observed continuously 2 years before and after the move; the year of the move itself is excluded. Standard errors are clustered at the household level.

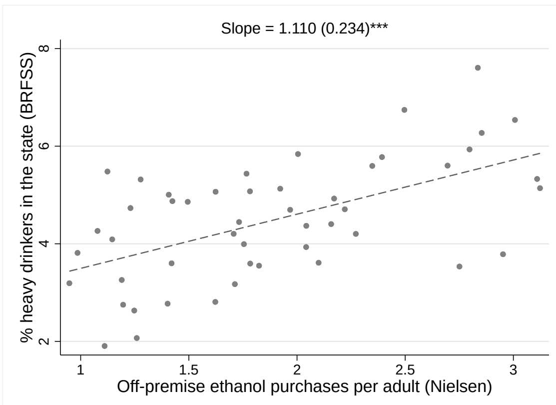


(a) 2 years before, 3 years after

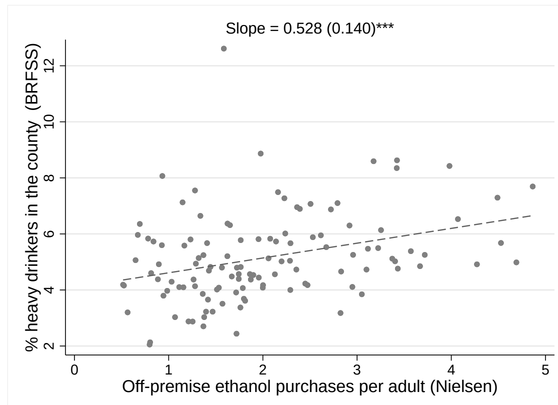
(b) 1 year before, 4 years after

Figure A.5: Event study of alcohol purchases

Notes: Each figure presents the coefficients $\theta_{r(i,t)}$ (point estimates and 95% confidence intervals) estimated from equation (1). The coefficient for the last time period before the move is normalized to 0. The dependent variable is the logarithm of quantity of total ethanol purchased. Each regression includes quarter-year dummies and household fixed effects. Sample includes all movers that move across state lines, limiting the sample to those that move only once and who are observed continuously: on figure A.5a, 2 years before and 3 years after the move year and on figure A.5b, 1 year before and 4 years after the move year. The year of the move itself is excluded. Quarters more than 2 years before or 3 years after the move year are included in the estimation but not shown on the figure A.5a. Similarly, quarters more than 1 year before or 4 years after the move year are included in the estimation but not shown on the figure A.5b. On figure A.5a, the sample includes 1,045 households and 40,896 observations. On figure A.5b, the sample includes 1,089 households and 42,420 observations. Standard errors are clustered at the household level.



(a) State level



(b) County level

Figure A.6: Off-premise alcohol purchases (Nielsen) vs the share of heavy drinkers (BRFSS) at state and county level

Notes: Each dot represents either a state (figure A.6a) or a county (figure A.6b). Figure A.6b includes only counties with at least 100 respondents in both BRFSS and Nielsen's dataset. Y-axis reflects the share of heavy drinkers (defined by BRFSS as "adult men having more than two drinks per day and adult women having more than one drink per day") in a given locality in 2007 BRFSS dataset. X-axis displays the yearly off-premise alcohol purchases per adult in a given county or state calculated using Nielsen sample weights (and information on the number of adults) from the NielsenIQ Consumer Panel of 2007. We choose 2007 since that was the first year with a larger number of households (approximately 60,000) in the NielsenIQ Consumer Panel. The line of the best fit is obtained from OLS regressions. Its slope coefficient and standard error (in parentheses) are reported on the graph. The coefficients are significant at 1% level on both figures.

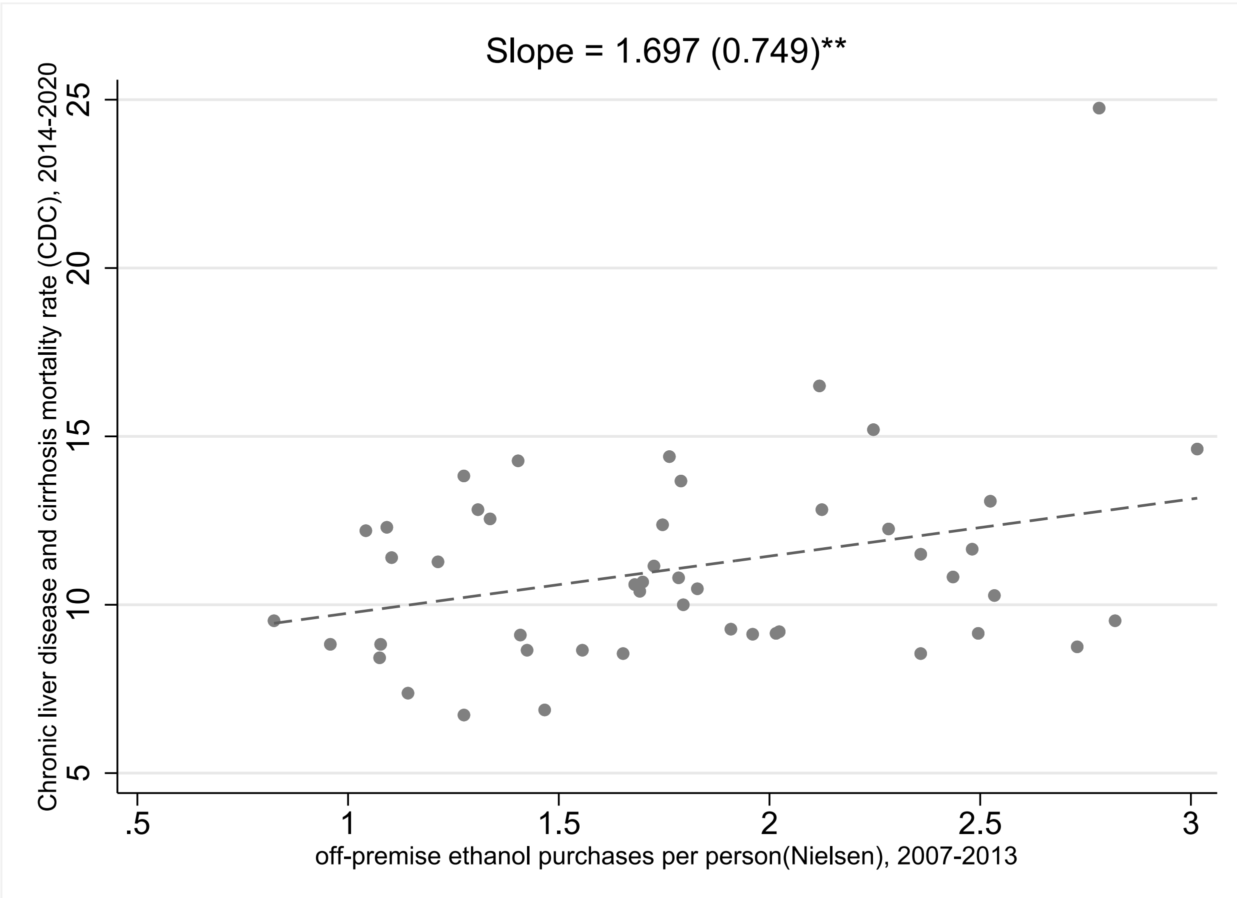


Figure A.7: Off-premise alcohol purchases (Nielsen) vs cirrhosis and chronic liver disease mortality rate at state level (CDC)

Notes: Each dot represents a state. X-axis reflects the average annual off-premise alcohol purchases per adult in a given state calculated using Nielsen sample weights (and information on the number of adults) from the NielsenIQ Consumer Panel of 2007-2014. Y-axis displays the average annual age-adjusted mortality rate per 100,000 population for cirrhosis and chronic liver disease at state level between 2014 and 2020 from the CDC website. We choose 2007 since that was the first year with a larger number of households (approximately 60,000) in the NielsenIQ Consumer Panel. The mortality rate data is only available for 2005, 2014-2020 from the CDC. We lag the outcome by many years since chronic liver disease and cirrhosis are caused by years of excessive drinking. The line of the best fit is obtained from OLS regressions. Its slope coefficient and standard error (in parentheses) are reported on the graph. The coefficients are significant at 5% level.

Table A.1: Summary statistics

	Ever-movers (1)	Non-movers (2)
	Demographic characteristics	
Household income	69414.9	63820.3
Household size	2.4	2.7
Average age of household heads	50.8	49.9
College	0.597	0.522
Male household head employed	0.749	0.765
Female household head employed	0.588	0.614
Race: white non-Hispanic	0.810	0.785
Married	0.642	0.646
	First observed residence	
Northeast	18.2	16.9
Midwest	22.8	25.5
South	36.5	38.0
West	22.5	19.6
	Alcohol purchases	
Quantity of total pure alcohol	0.64	0.49
Quantity of beer	3.69	3.37
Quantity of liquor	0.73	0.53
Quantity of wine	1.54	1.05
Alcohol expenditures	27.50	20.45
Purchasing alcohol	0.74	0.66
Number of households	3267	172827

Notes: All characteristics are measured during the first year in the sample. In the survey, income is reported in intervals; to calculate *Household income* in dollars, midpoints of the intervals are used. Income and expenditures are deflated to 2015 dollars using the consumer price index for urban consumers. For households with two heads, *College* indicates whether at least one of them has attended college. Alcohol purchases are measured per adult per quarter. Quantity of alcohol is measured in liters and expenditures in dollars. *Purchasing alcohol* is an indicator variable for whether the household has purchased any alcohol during the current and past 3 quarters. Ever-movers include all movers that move across state lines, limiting the sample to those that move only once. Non-movers includes all households that don't ever move across state lines. For all demographic characteristics (except marital status) and alcohol purchases, the differences between movers and non-movers are statistically significantly different from zero at 1 percent significance level according to the Wilcoxon rank-sum test.

Table A.2: Migration patterns (percentages)

Origin region	Destination region			
	Northeast	Midwest	South	West
Northeast	4.4	1.3	10.9	1.7
Midwest	1.0	6.9	10.9	4.0
South	3.1	6.3	22.0	5.0
West	1.3	3.0	7.1	11.1

Notes: The sample of movers.

Table A.3: Comparison of movers to states where alcohol consumption is lower versus higher than in the state of origin

	Movers to states with alcohol purchases		t-test	Wilcoxon test
	lower	higher	p-value	p-value
	(1)	(2)	(3)	(4)
	Demographic characteristics			
Household income	70316.9	68398.3	0.003	0.005
Household size	2.4	2.5	0.007	0.000
Average age of household heads	51.0	50.5	0.053	0.006
College	0.610	0.582	0.001	0.001
Male household head employed	0.740	0.760	0.016	0.016
Female household head employed	0.588	0.589	0.943	0.943
Race: white non-Hispanic	0.815	0.805	0.180	0.180
Married	0.639	0.646	0.412	0.412
	First observed residence			
Northeast	25.2	10.4		
Midwest	21.8	23.8		
South	36.8	36.1		
West	16.1	29.8		
	Alcohol purchases			
Quantity of total pure alcohol	0.58	0.71	0.000	0.000
Quantity of beer	3.24	4.20	0.000	0.000
Quantity of liquor	0.69	0.77	0.106	0.000
Quantity of wine	1.37	1.74	0.001	0.004
Alcohol expenditures	26.18	28.99	0.026	0.000
Purchasing alcohol	0.72	0.75	0.092	0.092
Number of households	1731	1536		

Notes: All demographic characteristics and alcohol purchases are measured during the first year in the sample. In the survey, income is reported in intervals; to calculate *Household income* in dollars, midpoints of the intervals are used. Income and expenditures are deflated to 2015 dollars using the consumer price index for urban consumers. For households with two heads, *College* indicates whether at least one of them has attended college. Alcohol purchases are measured per adult per quarter. Quantity of alcohol is measured in liters and expenditures in dollars. Column 3 presents the p-value of the t-test and column 4 of the Wilcoxon rank-sum test for whether the difference between the movers to states with higher versus lower alcohol consumption is significantly different from zero. The sample includes all movers that move across state lines, limiting the sample to those that move only once.

Table A.4: Robustness: Change in alcohol purchases after move, alternative samples

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
Panel A: Sample: continuously 2 years before and after					
$\Delta\cdot$ After move	0.669***	0.616***	0.650***	0.886***	0.687***
	(0.106)	(0.114)	(0.129)	(0.130)	(0.077)
Households	1379	1379	1379	1379	1379
Observations	50964	50964	50964	50964	50964
Panel B: Sample: balanced panel 2 years around the move					
$\Delta\cdot$ After move	0.651***	0.553***	0.708***	0.811***	0.632***
	(0.101)	(0.104)	(0.116)	(0.134)	(0.075)
Households	1379	1379	1379	1379	1379
Observations	22064	22064	22064	22064	22064

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. The sample includes all movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In panel A, the sample is further restricted to those who are observed continuously from 2 years before the move to 2 years after the move; quarters more than 2 years before or after the move are also included in the estimation. In panel B, the sample is restricted to the balanced panel of households observed 2 years before and after the move. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.5: Robustness: Change in alcohol purchases after move, households with constant demographic characteristics

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
Panel A: Sample: constant demographics 3 years around the move					
$\Delta\cdot$ After move	0.655***	0.572***	0.430**	1.016***	0.663***
	(0.156)	(0.148)	(0.200)	(0.209)	(0.109)
Households	656	656	656	656	656
Observations	8392	8392	8392	8392	8392
Panel B: Sample: constant demographics 5 years around the move					
$\Delta\cdot$ After move	0.723***	0.474**	0.223	1.329***	0.823***
	(0.221)	(0.239)	(0.249)	(0.331)	(0.163)
Households	271	271	271	271	271
Observations	3424	3424	3424	3424	3424

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. The sample includes all movers that move across state lines, limiting the sample to those that move only once. The sample includes time periods within two years before and after the year of the move. The year of the move is excluded from the sample. The sample is restricted to those whose employment status, marital status, household size, the number of members aged 21 and above, income, and occupation stay constant three years around the move (from the year before to the year after the move year) on panel A; or five years around the move (from two years before to two years after the move year) on panel B. As in the dataset, income is reported in 16 intervals, we consider income to be constant if it remains in the same income interval. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.6: Robustness: Change in alcohol purchases after move, alternative samples

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
	Panel A: Sample: includes the year of move				
$\Delta\cdot$ After move	0.698***	0.678***	0.608***	0.798***	0.671***
	(0.077)	(0.081)	(0.088)	(0.085)	(0.053)
Households	3267	3267	3267	3267	3267
Observations	110928	110928	110928	110928	110928
	Panel B: Sample: includes non-movers				
$\Delta\cdot$ After move	0.719***	0.707***	0.610***	0.830***	0.704***
	(0.083)	(0.087)	(0.097)	(0.094)	(0.058)
Households	176094	176094	176094	176094	176094
Observations	3110488	3110488	3110488	3110488	3110488
	Panel C: Sample: Excludes Utah				
$\Delta\cdot$ After move	0.723***	0.703***	0.621***	0.844***	0.716***
	(0.085)	(0.089)	(0.099)	(0.097)	(0.060)
Households	3206	3206	3206	3206	3206
Observations	96348	96348	96348	96348	96348

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. In panel A, the sample includes all movers that move across state lines, limiting the sample to those that move only once. In panel B, the sample adds non-movers. In panel C, the sample excludes movers to and from Utah. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.7: Robustness: Change in alcohol purchases after move, sample excludes large purchases

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
$\Delta\cdot$ After move	0.785***	0.719***	0.749***	0.914***	0.763***
	(0.083)	(0.086)	(0.094)	(0.088)	(0.058)
Households	3267	3267	3267	3267	3267
Observations	97860	97860	97860	97860	97860

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. The sample includes all movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. The sample excludes days when the household made a large alcohol purchase: purchasing either more than a six-pack of beer, more than a typical large bottle of liquor (1.75 liters), or more than four bottles of wine. These thresholds don't imply that smaller quantities are consumed by a single household in one sitting. Instead, the thresholds reflect alcohol purchasing and storage behavior, and allow price sensitive households to obtain quantity discounts. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.8: Robustness: Change in alcohol purchases after move. Aggregate purchases calculated based on two years prior the move year in the state of origin and two years after the move year in the state of destination

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
$\Delta\cdot$ After move	0.706***	0.661***	0.618***	0.815***	0.693***
	(0.080)	(0.083)	(0.091)	(0.092)	(0.057)
Households	3267	3267	3267	3267	3267
Observations	97860	97860	97860	97860	97860

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. The average purchases are calculated based on two years prior the move year in the state of origin and two years after the move year in the state of destination. The sample includes all movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.9: Robustness: Change in alcohol purchases after move, functional form

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
Panel A: Inverse Hyperbolic Sine transformation of purchases					
$\Delta\cdot$ After move	0.710***	0.695***	0.612***	0.819***	0.697***
	(0.084)	(0.085)	(0.097)	(0.091)	(0.056)
Households	3267	3267	3267	3267	3267
Observations	97860	97860	97860	97860	97860
Panel B: Percentile rank of purchases					
$\Delta\cdot$ After move	0.710***	0.699***	0.627***	0.775***	0.704***
	(0.051)	(0.070)	(0.078)	(0.059)	(0.051)
Households	3267	3267	3267	3267	3267
Observations	97860	97860	97860	97860	97860
Panel C: Actual value of purchases					
$\Delta\cdot$ After move	0.641***	0.527***	0.519***	0.743***	0.540***
	(0.122)	(0.154)	(0.157)	(0.197)	(0.127)
Households	3267	3267	3267	3267	3267
Observations	97860	97860	97860	97860	97860

Notes: In panel A, the outcome variable is the Inverse Hyperbolic Sine transformation of alcohol purchases that equals $\ln\left[x+\sqrt{x^2+1}\right]$, where x is the quantity or expenditures of alcohol purchases. In panel B, the outcome variable is the percentile rank of alcohol purchases that equals $(Rank_{it}-1)/(N_t-1)$, where $Rank_{it}$ is household i's rank in the national distribution of alcohol purchases calculated for a given time period t, and N_t is the number of households in that time period. In panel C, the outcome variable is the quantity or expenditure of alcohol purchases. In each panel, Δ_i is the difference in average outcome variable between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes all movers that move across state lines, limiting the sample to those that move only once; the year of the move is excluded from the sample. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.10: Robustness: Change in alcohol purchases after move, including additional variables

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
Panel A: Sample of movers, includes demographic char.					
$\Delta\cdot$ After move	0.689***	0.679***	0.595***	0.806***	0.686***
	(0.082)	(0.087)	(0.097)	(0.093)	(0.058)
Demogr. Char.	Yes	Yes	Yes	Yes	Yes
Households	3267	3267	3267	3267	3267
Observations	97860	97860	97860	97860	97860
Panel B: Sample of movers and nonmovers, includes quarter since move FE					
$\Delta\cdot$ After move	0.698***	0.673***	0.610***	0.805***	0.687***
	(0.082)	(0.087)	(0.096)	(0.093)	(0.058)
Quarter since move FE	Yes	Yes	Yes	Yes	Yes
Households	176094	176094	176094	176094	176094
Observations	3110488	3110488	3110488	3110488	3110488

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. The sample includes all movers that move across state lines, limiting the sample to those that move only once; the year of the move is excluded from the sample. In panel B, the sample also includes non-movers. Each regression includes quarter-year dummies and household fixed effects. In panel A, regressions also include the logarithm of income, household size, number of adults aged 21+, marital status, employment status, and an indicator for children aged 0-5. In panel B, regressions include event time fixed effects. In the panel B, we include non-movers, because if we only use movers, event time fixed effects and time period fixed effects would be perfectly collinear. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.11: Robustness: Change in alcohol purchases after move; clustering at origin-destination states pair

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
Clustering at origin-destination states pair					
$\Delta \cdot$ After move	0.708***	0.689***	0.608***	0.821***	0.695***
	(0.083)	(0.091)	(0.101)	(0.105)	(0.064)
Households	3267	3267	3267	3267	3267
Observations	97860	97860	97860	97860	97860

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes all movers that move across state lines, limiting the sample to those that move only once; the year of the move is excluded from the sample. Standard errors are clustered at the origin-destination states pair (1,012 clusters). *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.12: Robustness: Change in alcohol purchases after move, extensive margin, aggregate and asymmetric effects. Difference-in-differences estimates. Extensive margin calculated based on half a year.

	(1)	(2)	(3)	(4)
Extensive margin: whether bought alcohol in half-a-year				
		Purchasing		
	Alcohol	Beer	Liquor	Wine
Panel A: Average effects				
$\Delta \cdot$ After move	0.619***	0.706***	0.592***	0.664***
	(0.050)	(0.071)	(0.078)	(0.055)
Households	3267	3267	3267	3267
Observations	93944	93944	93944	93944
Panel B: Asymmetric effects				
$\Delta \cdot 1[\Delta > 0] \cdot$ After move	0.841***	0.945***	0.902***	0.938***
	(0.078)	(0.108)	(0.122)	(0.090)
$\Delta \cdot 1[\Delta < 0] \cdot$ After move	0.386***	0.421***	0.317***	0.389***
	(0.075)	(0.110)	(0.120)	(0.080)
Wald test, coef. equality, p-value	0.000	0.002	0.002	0.000
Households	3267	3267	3267	3267
Observations	93944	93944	93944	93944

Notes: Each column-panel combination presents estimates from a separate regression. Dependent variable is an indicator for purchasing beer (column 1), liquor (column 2) or wine (column 3); it is calculated as a rolling average over two quarters. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in the share of households purchasing any beer (column 1), liquor (column 2) or wine (column 3) between the destination and origin state. $1[\Delta_i > 0]$ is an indicator for Δ_i being strictly positive, that is, a move to a state with a larger share of households purchasing that category of alcohol products, and $1[\Delta_i < 0]$ indicates a move to a state with smaller share of households purchasing that category of alcohol products. Each regression includes quarter-year fixed effects and household fixed effects. Sample includes all movers that move across state lines, limiting the sample to those that move only once; the year of the move is excluded from the sample. In panel B, to test whether the effect is asymmetric, p-value of the Wald test for the equality of the two coefficients is included. Standard errors (in parentheses) are clustered at the household level. *** Indicates significance at 1 percent level.

Table A.13: Robustness: Change in alcohol purchases after move; geographic area is defined as state; move inside versus across census regions

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
Panel A: move across census regions					
$\Delta\cdot$ After move	0.766***	0.806***	0.613***	0.900***	0.754***
	(0.105)	(0.107)	(0.121)	(0.127)	(0.076)
Households	1815	1815	1815	1815	1815
Observations	54292	54292	54292	54292	54292
Panel B: move inside census region					
$\Delta\cdot$ After move	0.622***	0.500***	0.598***	0.714***	0.616***
	(0.134)	(0.148)	(0.159)	(0.138)	(0.089)
Households	1452	1452	1452	1452	1452
Observations	43568	43568	43568	43568	43568

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes all movers that move across state lines, limiting the sample to those that move only once; the year of the move is excluded from the sample. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.14: Robustness: Change in alcohol purchases after move; geographic area is defined as county; move inside versus across states

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
	Panel A: move across states				
$\Delta\cdot$ After move	0.646***	0.461***	0.625***	0.995***	0.746***
	(0.177)	(0.173)	(0.183)	(0.187)	(0.131)
Households	543	543	543	543	543
Observations	16336	16336	16336	16336	16336
	Panel B: move inside state				
$\Delta\cdot$ After move	0.275	0.055	0.308	0.103	0.298*
	(0.168)	(0.167)	(0.208)	(0.220)	(0.159)
Households	947	947	947	947	947
Observations	29900	29900	29900	29900	29900

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin county of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across county borders, limiting the sample to those that move only once, and move between counties where at any time period there are at least 50 households in the dataset; the year of the move is excluded from the sample. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.15: Robustness: Change in alcohol purchases after move; geographic area is defined as 3-digit zip code; move inside versus across states

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
	Panel A: move across states				
$\Delta\cdot$ After move	0.422***	0.649***	0.362*	0.655***	0.596***
	(0.152)	(0.158)	(0.188)	(0.149)	(0.106)
Households	701	701	701	701	701
Observations	21032	21032	21032	21032	21032
	Panel B: move inside state				
$\Delta\cdot$ After move	0.183	-0.062	0.337	0.142	0.251**
	(0.125)	(0.138)	(0.213)	(0.135)	(0.109)
Households	1262	1262	1262	1262	1262
Observations	39568	39568	39568	39568	39568

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin 3-digit zip code of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across 3-digit zip codes, limiting the sample to those that move only once, and move between 3 digit zip codes where at any time period there are at least 50 households in the dataset; the year of the move is excluded from the sample. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.16: Heterogeneity: Change in alcohol purchases after move; separating by retail conditions: liquor control states

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
Panel A: Moves between non-liquor-control states					
$\Delta \cdot$ After move	0.665***	0.624***	0.676***	0.907***	0.717***
	(0.107)	(0.117)	(0.122)	(0.127)	(0.079)
Households	1738	1738	1738	1738	1738
Observations	52060	52060	52060	52060	52060
Panel B: Moves between liquor control states					
$\Delta \cdot$ After move	0.899**	1.017***	0.535	0.584**	0.829***
	(0.423)	(0.352)	(0.411)	(0.292)	(0.244)
Households	198	198	198	198	198
Observations	5496	5496	5496	5496	5496
Panel C: Moves: from CS to not CS; or from not CS to CS					
$\Delta \cdot$ After move	0.770***	0.765***	0.503***	0.750***	0.648***
	(0.136)	(0.138)	(0.168)	(0.151)	(0.090)
Households	1331	1331	1331	1331	1331
Observations	40304	40304	40304	40304	40304

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In panel A, the sample is further restricted to moves between states that are not liquor control states; in panel B, to moves between states that are liquor control states; in panel C, to the remaining moves (either from liquor control state to non-control state or from non-control state to control state). Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.17: Heterogeneity: Change in alcohol purchases after move; by retail conditions: whether wine available in the majority of grocery stores

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
	Panel A: Moves between states wine available				
$\Delta \cdot$ After move	0.504***	0.386**	0.732***	0.511***	0.401***
	(0.147)	(0.183)	(0.152)	(0.162)	(0.109)
Households	1795	1795	1795	1795	1795
Observations	53660	53660	53660	53660	53660
	Panel B: Moves to/from wine not available				
$\Delta \cdot$ After move	0.807***	0.772***	0.490***	0.993***	0.799***
	(0.100)	(0.099)	(0.120)	(0.115)	(0.068)
Households	1472	1472	1472	1472	1472
Observations	44200	44200	44200	44200	44200

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In panel A, the sample is further restricted to moves between states where wine is available in more than 80% grocery stores (calculated from the NielsenIQ Retail Scanner Data as described in Online Appendix D); in panel B, to moves to/from states where wine is not available in more than 80% of the grocery stores. There is not enough moves between the states where wine is not available in the grocery stores to estimate analogous regressions for those. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.18: Heterogeneity: Change in alcohol purchases after move; by retail conditions: whether beer available in the majority of grocery stores

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
	Panel A: Moves between states beer available				
$\Delta\cdot$ After move	0.660***	0.582***	0.730***	0.696***	0.597***
	(0.105)	(0.135)	(0.113)	(0.112)	(0.075)
Households	2392	2392	2392	2392	2392
Observations	71324	71324	71324	71324	71324
	Panel B: Moves to/from beer not available				
$\Delta\cdot$ After move	0.766***	0.765***	0.222	1.092***	0.831***
	(0.132)	(0.111)	(0.180)	(0.170)	(0.092)
Households	875	875	875	875	875
Observations	26536	26536	26536	26536	26536

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In panel A, the sample is further restricted to moves between states where beer is available in more than 80% of the grocery stores (calculated from the NielsenIQ Retail Scanner Data as described in Online Appendix D); in panel B, to moves to/from states where beer is not available in more than 80% of the grocery stores. There is not enough moves between the states where beer is not available in the grocery stores to estimate analogous regressions for those. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.19: Heterogeneity: Change in alcohol purchases after move; heavy drinkers and moderate drinkers

	Quantity				Expend.
	Total	Beer	Liquor	Wine	Total
	(1)	(2)	(3)	(4)	(5)
	Panel A. Sample: Heavy drinkers				
$\Delta\cdot$ After move	0.740**	0.797**	0.522	1.236***	0.539***
	(0.354)	(0.346)	(0.455)	(0.395)	(0.171)
Households	521	521	521	521	521
Observations	15880	15880	15880	15880	15880
	Panel B. Sample: Moderate drinkers				
$\Delta\cdot$ After move	0.714***	0.702***	0.656***	0.779***	0.750***
	(0.079)	(0.093)	(0.091)	(0.094)	(0.067)
Households	2201	2201	2201	2201	2201
Observations	67716	67716	67716	67716	67716

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In panel A, the sample is further restricted to movers whose before move alcohol purchases where above the 80th percentile (heavy drinkers); in panel B, to movers who purchased alcohol before the move but who purchases were below the 80th percentile (moderate drinkers). Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.20: Heterogeneity: Change in alcohol purchases after move; by education, income, and age

	College		Income > median		Age		
	Yes	No	Yes	No	<45	45-60	61+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A. Average effects						
$\Delta\cdot$ After move	0.746***	0.657***	0.736***	0.679***	0.553***	0.726***	0.770***
	(0.103)	(0.135)	(0.121)	(0.112)	(0.140)	(0.126)	(0.158)
Households	1935	1332	1615	1652	1010	1329	928
Observations	57808	40052	49508	48352	24492	42536	30832
	Panel B. Asymmetric effects						
$\Delta \cdot 1[\Delta > 0]\cdot$ After move	0.846***	0.870***	1.054***	0.659***	0.473**	0.974***	0.949***
	(0.149)	(0.218)	(0.193)	(0.159)	(0.224)	(0.198)	(0.218)
$\Delta \cdot 1[\Delta < 0]\cdot$ After move	0.624***	0.417**	0.341*	0.701***	0.635***	0.440**	0.532*
	(0.170)	(0.204)	(0.178)	(0.186)	(0.228)	(0.187)	(0.272)
Wald test, coef. equality, p-value	0.358	0.165	0.013	0.872	0.649	0.070	0.263
Households	1935	1332	1615	1652	1010	1329	928
Observations	57808	40052	49508	48352	24492	42536	30832

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In each regression the sample is further restricted by the described characteristic. All characteristics are measured in the last year before the move year. College indicates whether either the male or female household head has college education. Age indicates the age of the youngest household head. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.21: Heterogeneity: Change in alcohol purchases after move; by changes in income and occupation and employment

	Income			Occupation/employment	
	Increase	Decrease	Constant	Change	Constant
	(1)	(2)	(3)	(4)	(5)
	Panel A. Average effects				
$\Delta\cdot$ After move	0.714***	0.689***	0.715***	0.504***	0.917***
	(0.166)	(0.147)	(0.123)	(0.109)	(0.123)
Households	939	957	1371	1624	1643
Observations	26832	29500	41528	48360	49500
	Panel B. Asymmetric effects				
$\Delta \cdot 1[\Delta > 0]\cdot$ After move	0.598**	0.932***	0.933***	0.531***	1.185***
	(0.236)	(0.245)	(0.181)	(0.171)	(0.183)
$\Delta \cdot 1[\Delta < 0]\cdot$ After move	0.837***	0.415*	0.434**	0.472***	0.601***
	(0.300)	(0.212)	(0.183)	(0.173)	(0.194)
Wald test, coef. equality, p-value	0.572	0.145	0.066	0.824	0.040
Households	939	957	1371	1624	1643
Observations	26832	29500	41528	48360	49500

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In each regression the sample is further restricted by the described characteristic. The comparison is between the first year after the move year and the latest year before the move year. As in the dataset, income is reported in 16 intervals, in columns 1–3, we consider income to be constant if it remains in the same income interval, and increasing or decreasing otherwise. In columns 4–5, the sample is split by whether at least one of the household head changed occupation or employment status. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.22: Heterogeneity: Change in alcohol purchases after move; by the number of adults in the household, and changes in the number of adults

	Number of adults		
	1	> 1	Change
	(1)	(2)	(3)
	Panel A. Average effects		
$\Delta\cdot$ After move	0.666***	0.768***	0.590***
	(0.193)	(0.105)	(0.171)
Households	690	1896	681
Observations	22324	55076	20460
	Panel B. Asymmetric effects		
$\Delta \cdot 1[\Delta > 0]\cdot$ After move	0.679***	0.982***	0.734***
	(0.253)	(0.174)	(0.261)
$\Delta \cdot 1[\Delta < 0]\cdot$ After move	0.650**	0.512***	0.436*
	(0.325)	(0.160)	(0.256)
Wald test, coef. equality, p-value	0.946	0.074	0.446
Households	690	1896	681
Observations	22324	55076	20460

Notes: Dependent variable is the logarithm of alcohol purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of alcohol purchases between the destination and origin state of the mover. Each regression includes quarter-year dummies and household fixed effects. The sample includes movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In each regression the sample is further restricted by the number of adults aged 21 or above, or the change in the number of adults. In columns 1–2, the number of adults remains constant, and in column 3, it changes. The comparison is between the first year after the move year and the latest year before the move year. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.23: Change in food purchases after move. Dependent variable: the logarithm of food purchases

	Calories (1)	Expenditures (2)
$\Delta \cdot$ After move	0.138	0.136
	(0.098)	(0.093)
Households	3267	3267
Observations	97860	97860

Notes: Dependent variable is the logarithm of food of purchases, either calories (column 1) or expenditures (column 2). Food purchases include all grocery purchases, except breakfast foods and random weight purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of the calories or expenditures of purchases between the destination and origin state of the mover. The sample includes all movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.24: Change in cigarette purchases after move. Dependent variable: the logarithm of cigarettes quantity

	Cigarettes quantity (1)
$\Delta \cdot$ After move	0.293***
	(0.109)
Households	3267
Observations	97860

Notes: Dependent variable is the logarithm of cigarettes quantity. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of cigarettes quantity between the destination and origin state of the mover. The sample includes all movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.25: Change in purchases after move: beer brands, soda, and fruits and vegetables

	Quantity		
	(1)	(2)	(3)
		Panel A: Beer	
	Bud Light	Miller Light	Coors Light
$\Delta\cdot$ After move	0.377***	0.625***	0.667***
	(0.113)	(0.129)	(0.202)
Households	3267	3267	3267
Observations	97860	97860	97860
		Panel B: Beer	
	Budweiser	Natural Light	Busch Light
$\Delta\cdot$ After move	0.531**	0.484**	0.569**
	(0.207)	(0.210)	(0.284)
Households	3267	3267	3267
Observations	97860	97860	97860
		Panel C: Other products	
	Regular soda	Low-calorie soda	Fruits & vegetables
$\Delta\cdot$ After move	0.088	-0.129	0.136
	(0.098)	(0.114)	(0.105)
Households	3267	3267	3267
Observations	97860	97860	97860

Notes: Dependent variable is the logarithm of quantity of purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of the quantity of purchases between the destination and origin state of the mover. The sample includes all movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. In panel A and B, the dependent variable is one of the six most sold brands of beer. In panel C, dependent variable is either regular soda, low-calorie soda, fruits and vegetables. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

Table A.26: Change in food purchases after move as a response to move to high or low alcohol consumption region. Dependent variable: the logarithm of food purchases

	Calories (1)	Expenditures (2)
Δ log. alcohol quantity · After move	-0.045	
	(0.154)	
Δ log. alcohol expend. · After move		0.019
		(0.027)
Households	3267	3267
Observations	97860	97860

Notes: Dependent variable is the logarithm of food of purchases, either calories (column 1) or expenditures (column 2). Food purchases include all grocery purchases, except breakfast foods and random weight purchases. $\Delta_i = \bar{y}_{D,i} - \bar{y}_{O,i}$ is the difference in average logarithm of total alcohol quantity or expenditures between the destination and origin state of the mover. The sample includes all movers that move across state lines, limiting the sample to those that move only once, the year of the move is excluded from the sample. Each regression includes quarter-year dummies and household fixed effects. Standard errors are clustered at the household level. *** Indicates significance at the 1 percent level, ** at a 5 percent level, * at a 10 percent level.

B Online appendix: Alternative decomposition of alcohol purchases to individual characteristics and location effects

Using data on both movers and non-movers, we estimate the following equation for household i in state j in period t:

$$y_{ijt} = \alpha_i + \gamma_j + \tau_t + \rho_{it} + \varepsilon_{ijt} \tag{2}$$

where households are indexed by i, states by j, and periods by t, and where α-s are household fixed effects, γ-s are state fixed effects, τ-s are time period fixed effects, and for movers ρ_{it} are relative period effects since the move.

This approach provides an alternative way of measuring the importance of the environment and, as such, it works as a robustness check for the analysis described in the main text.

Table B.1 presents the additive decomposition of the difference between the high and low alcohol consumption areas. It reports the overall difference between the two areas R and R' in average log. consumption $\hat{y}_R - \hat{y}_{R'}$, the difference in estimated state-level location effects $\hat{\gamma}_R - \hat{\gamma}_{R'}$, and the share of the difference due to the location effects $S_{locat}(R, R') = \frac{\hat{\gamma}_R - \hat{\gamma}_{R'}}{\hat{y}_R - \hat{y}_{R'}}$.

While to identify the location effects equation (2) uses the variation from movers, the results of the decomposition need not to be the same as the difference-in-differences estimates in table 2. There are two main differences. First, here, the model is estimated using data also from non-movers, which helps to measure time varying effects. Second, here, the comparision is between two groups instead of all households.

The estimates (table B.1) show that about 60% of the difference between the above- and below- median alcohol consumption states and 70% of the difference between the states in the top and bottom alcohol-consumption deciles is due to the current environment. Overall, the results are similar to those from the event study and difference-in-differences analysis in the main text.

Table B.1: Additive decomposition of log. alcohol purchases to location and household effects

	Quantity			Expenditures		
	Above & below median	Top & bottom 25%	Top & bottom 10%	Above & below median	Top & bottom 25%	Top & bottom 10%
	(1)	(2)	(3)	(4)	(5)	(6)
Difference in average log. purchases:						
Overall	0.109	0.169	0.219	0.522	0.800	0.995
Due to location	0.066	0.102	0.154	0.361	0.438	0.576
Share of difference due to:						
Location	0.607	0.601	0.705	0.692	0.548	0.579
	(0.123)	(0.107)	(0.141)	(0.089)	(0.081)	(0.109)

Notes: Table is based on the estimates from equation (2). Dependent variable is log. quantity (columns 1–3) or log. expenditures (columns 4–6). Each regression includes quarter-year dummies, indicators for quarters since move, state fixed effects, and household fixed effects. Each column defines areas R and R' based on percentiles of average log. consumption. The first row measures the overall difference in average log. consumption $\hat{y}_R - \hat{y}_{R'}$, where $\hat{y}_R$ is calculated in 3 steps: first, taking averages across households in a given state in a given quarter to calculate the state-quarter average; second, calculating the average across time to obtain the state average; third, calculate average across the states in region R. The second row measures the difference due to location effect $\hat{\gamma}_R - \hat{\gamma}_{R'}$, and the third row reports the share of the difference due to location effects. Sample includes movers and non-movers (3,110,488 consumer-quarters). Standard errors (in parentheses) are calculated using a bootstrap with 50 repetitions at the household level.

C Online appendix: Characteristics of high-alcohol-consumption environments

Which characteristics are common to environments with high alcohol consumption? To answer the question, first, using data on both movers and non-movers, we regress alcohol purchases on household, state, and time fixed effects in order to quantify the role of locations in explaining the variation in alcohol purchases. Second, we measure which state-level characteristics are correlated with the estimated state-level location effects. For the first step, we use the estimates of equation (2) from online appendix B. This gives us the estimates of the state-level location effects. Using the estimates we proceed to step two.

Correlates of state-level location effects. We study which state-level characteristics are correlated with the estimated state-level location effects ($\hat{\gamma}$). The state-level characteristics describe alcohol availability, taxes, and prices. Two types of measures are used to capture alcohol availability. First, whether the state is a control state (i.e., has a state monopoly over selling wine or liquor). Even in states where alcohol is not restricted to be sold only in state-owned liquor stores, there could still be dry counties and restrictions on the number of alcohol licenses available for grocery stores. Therefore, we also measure availability as the percentage of grocery stores selling alcohol. Alcohol price includes excise taxes and, as such, captures both taxes and market conditions. Details of the construction of the characteristics are in the Online Appendix D.

Figure C.1 presents the correlates of estimated location effects: bivariate OLS estimates are presented on the left, and post-Lasso multivariate regression estimates are on the right. The bivariate OLS estimates show that states, where more grocery stores are selling stronger alcohol and where alcohol prices are lower, have statistically significantly higher location effects. In the post-Lasso multivariate regression, alcohol price and the share of grocery stores selling wine remain significant at the 90% level.[17] Similar estimates for alcohol expenditures and for beer, wine, and liquor quantities, are presented on figures C.2–C.5.

We note that figures C.1–C.5 do not necessarily describe causal effects and instead could capture endogeneous responses to voters' and consumers' preferences. This could happen both across states and time. For example, during economic downturns, governments tend to increase alcohol taxes (because other revenue sources have decreased), and alcohol consumption also tends to increase during economic downturns. This positive correlation between taxes and consumption does not mean that higher taxes lead to more drinking. In cross-section, in states with high demand, voters may support lower taxes and fewer restrictions on alcohol availability.

[17] For the share of grocery stores selling wine, the p-value equals 0.085, and for alcohol price, $p < 0.001$.

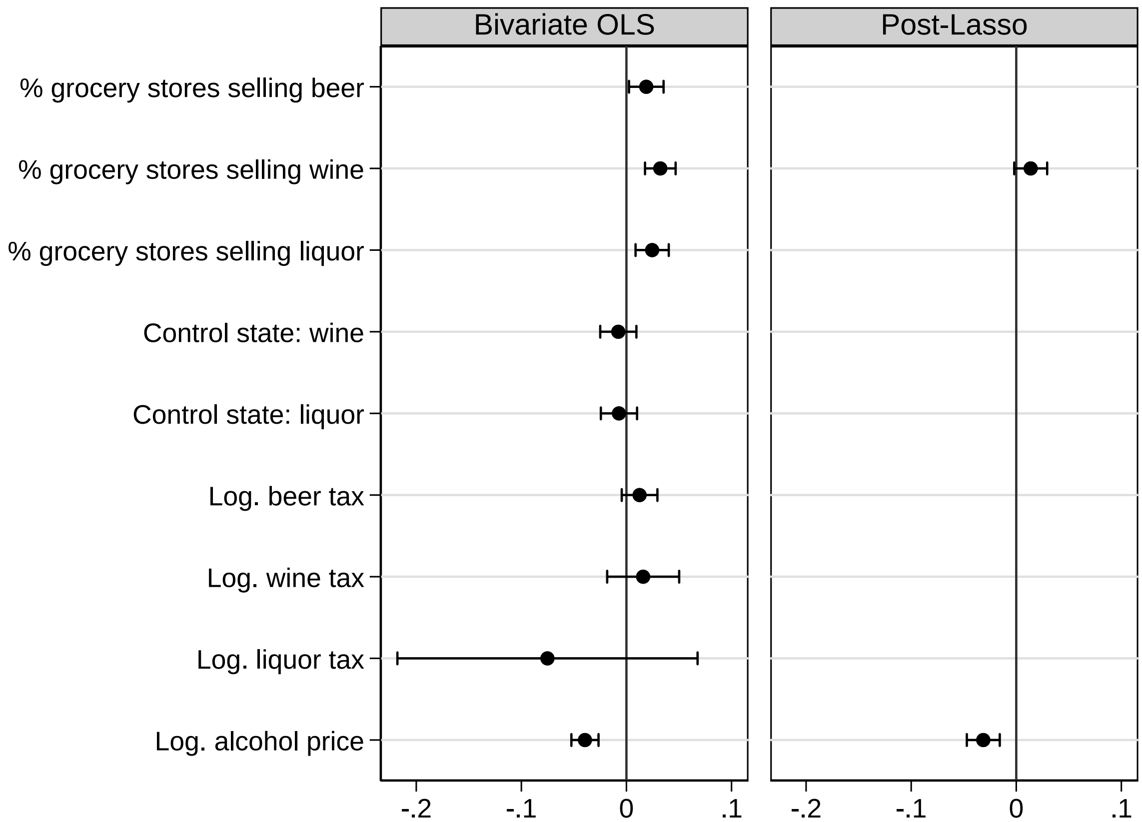


Figure C.1: Correlates of average place effects

Notes: Figure presents point estimates and 95% confidence intervals from bivariate OLS regressions (left) and a post-Lasso multivariate regression (right) of average state level location effects on state-level characteristics. Average state-level location effects ($\hat{\gamma}$) are estimated from equation (2) where dependent variable is the logarithm of alcohol quantity, and the sample includes movers and non-movers (3,110,488 household-quarters). In the OLS, Lasso, and Post-Lasso regressions the unit of observation is a state and the sample includes 49 observations (48 mainland states and Washington, D.C.), except for the excise taxes, where the sample includes only non-control states. All covariates are standardized to mean zero and standard deviation one. Post-Lasso estimates are obtained by first running a Lasso regression on the full set of covariates and then an OLS regression using the set of covariates chosen by Lasso. In the Lasso regression, the penalty level is chosen by the theory-based rule by Belloni et al. (2012), which is also applicable with heteroscedasticity.

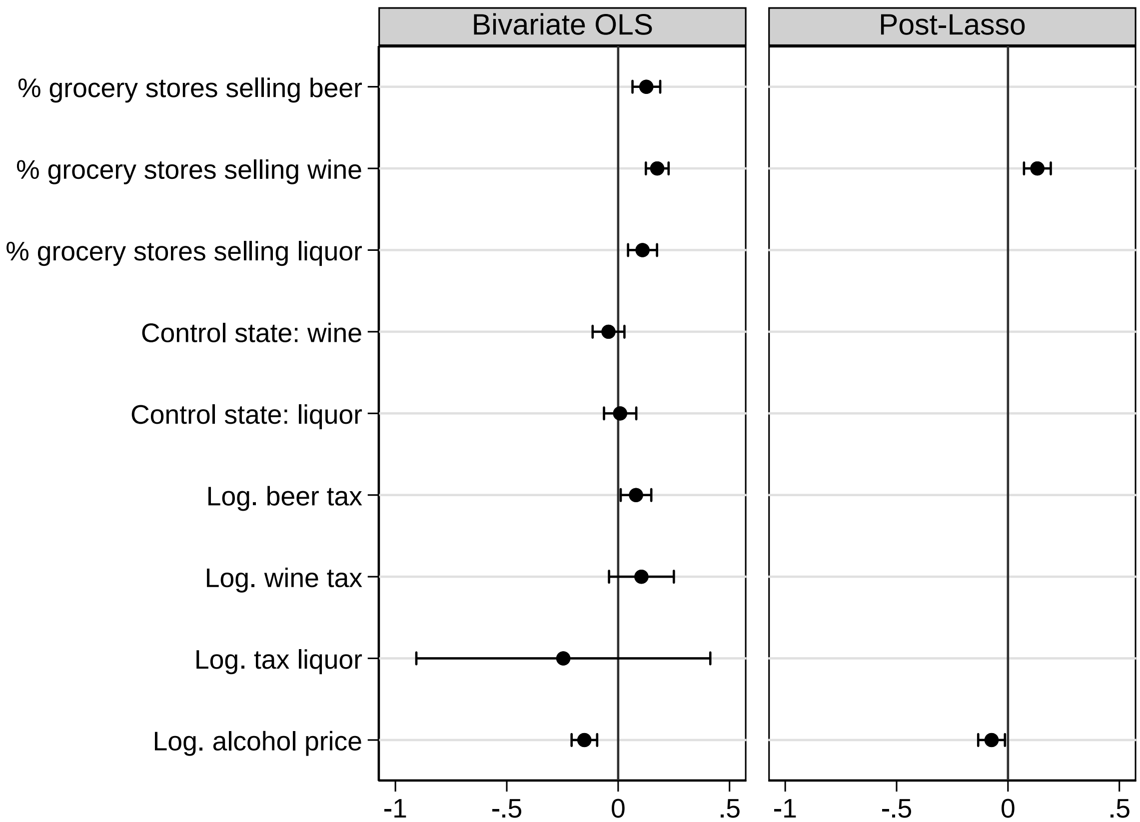


Figure C.2: Correlates of average place effects, log. expenditure of alcohol

Notes: Figure presents point estimates and 95% confidence intervals from bivariate OLS regressions (left) and a post-Lasso multivariate regression (right) of average state-level location effects on state-level characteristics. Average state-level location effects ($\hat{\gamma}$) are estimated from equation (2) where dependent variable is the logarithm of alcohol expenditures, and the sample includes movers and non-movers (3,110,488 household-quarters). In the OLS, Lasso, and Post-Lasso regressions the unit of observation is a state and the sample includes 49 observations (48 mainland states and Washington, D.C.), except for the excise taxes, where the sample includes only non-control states. All covariates are standardized to mean zero and standard deviation one. Post-Lasso estimates are obtained by first running a Lasso regression on the full set of covariates and then an OLS regression using the set of covariates chosen by Lasso. In the Lasso regression, the penalty level is chosen by the theory-based rule by Belloni et al. (2012), which is also applicable with heteroscedacity.

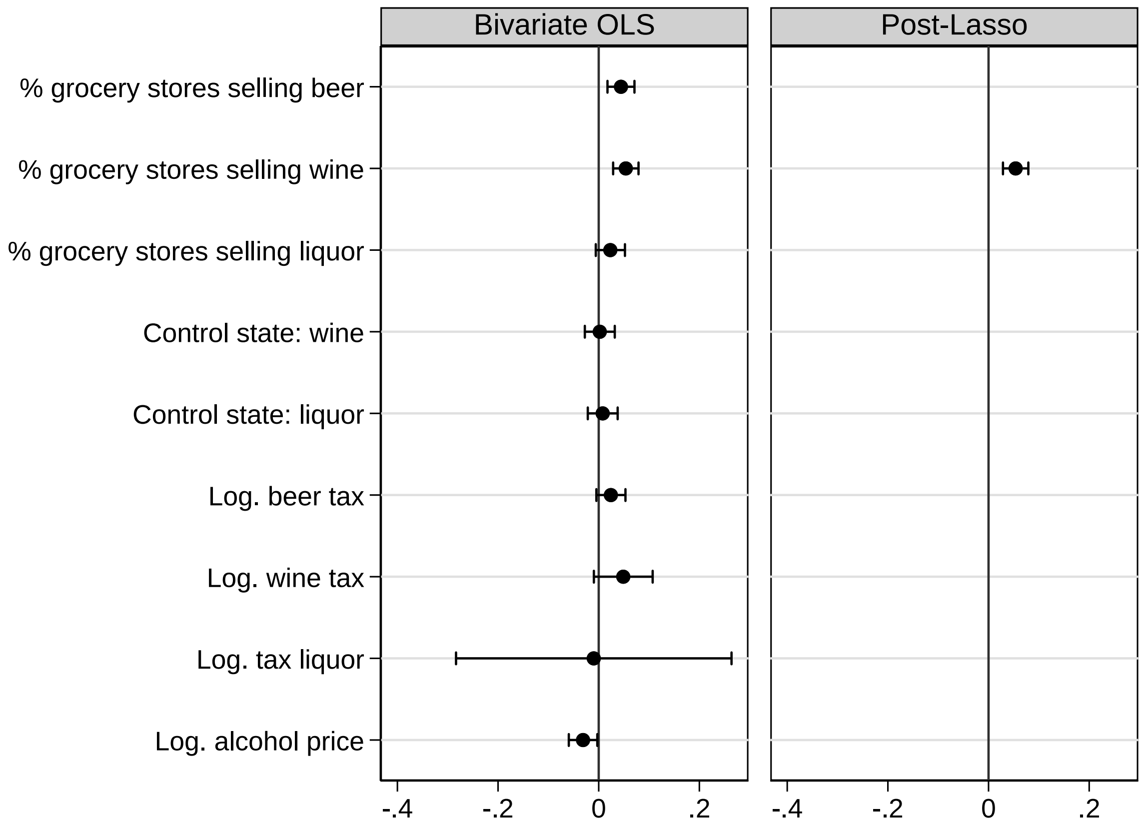


Figure C.3: Correlates of average place effects, log. quantity of beer

Notes: Figure presents point estimates and 95% confidence intervals from bivariate OLS regressions (left) and a post-Lasso multivariate regression (right) of average state-level location effects on state level characteristics. Average state-level location effects ($\hat{\gamma}$) are estimated from equation (2) where dependent variable is the logarithm of beer quantity, and the sample includes movers and non-movers (3,110,488 household-quarters). In the OLS, Lasso, and Post-Lasso regressions the unit of observation is a state and the sample includes 49 observations (48 mainland states and Washington, D.C.), except for the excise taxes, where the sample includes only non-control states. All covariates are standardized to mean zero and standard deviation one. Post-Lasso estimates are obtained by first running a Lasso regression on the full set of covariates and then an OLS regression using the set of covariates chosen by Lasso. In the Lasso regression, the penalty level is chosen by the theory-based rule by Belloni et al. (2012), which is also applicable with heteroscedacity.

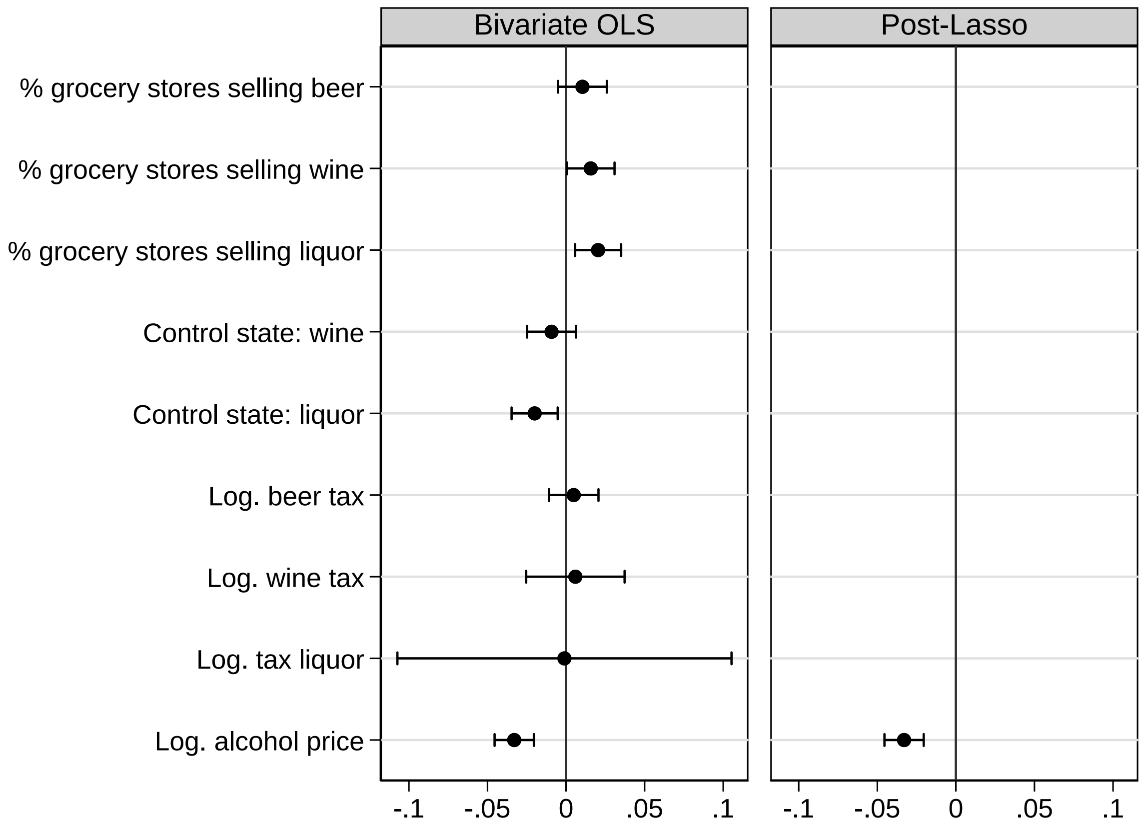


Figure C.4: Correlates of average place effects, log. quantity of liquor

Notes: Figure presents point estimates and 95% confidence intervals from bivariate OLS regressions (left) and a post-Lasso multivariate regression (right) of average state-level location effects on state level characteristics. Average state-level location effects ($\hat{\gamma}$) are estimated from equation (2) where dependent variable is the logarithm of liquor quantity, and the sample includes movers and non-movers (3,110,488 household-quarters). In the OLS, Lasso, and Post-Lasso regressions the unit of observation is a state and the sample includes 49 observations (48 mainland states and Washington, D.C.), except for the excise taxes, where the sample includes only non-control states. All covariates are standardized to mean zero and standard deviation one. Post-Lasso estimates are obtained by first running a Lasso regression on the full set of covariates and then an OLS regression using the set of covariates chosen by Lasso. In the Lasso regression, the penalty level is chosen by the theory-based rule by Belloni et al. (2012), which is also applicable with heteroscedacity.

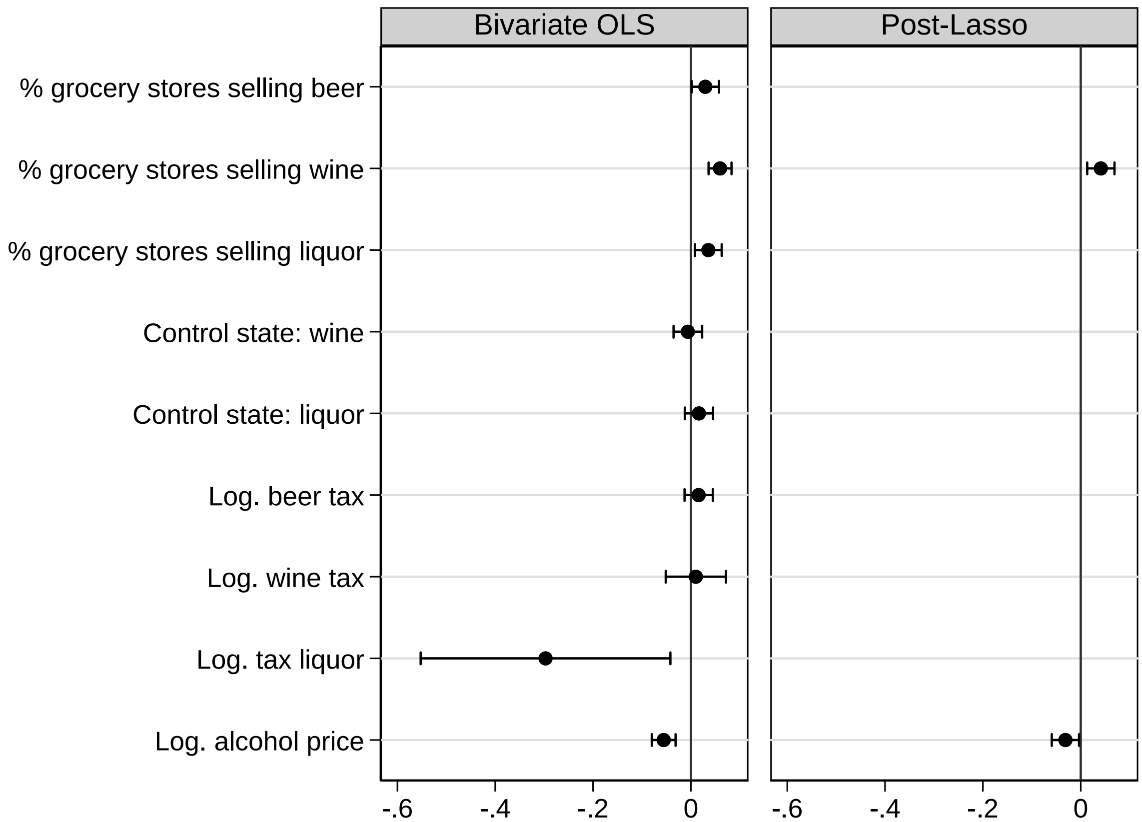


Figure C.5: Correlates of average place effects, log. quantity of wine

Notes: Figure presents point estimates and 95% confidence intervals from bivariate OLS regressions (left) and a post-Lasso multivariate regression (right) of average state-level location effects on state level characteristics. Average state-level location effects ($\hat{\gamma}$) are estimated from equation (2) where dependent variable is the logarithm of wine quantity, and the sample includes movers and non-movers (3,110,488 household-quarters). In the OLS, Lasso, and Post-Lasso regressions the unit of observation is a state and the sample includes 49 observations (48 mainland states and Washington D.C.), except for the excise taxes, where the sample includes only non-control states. All covariates are standardized to mean zero and standard deviation one. Post-Lasso estimates are obtained by first running a Lasso regression on the full set of covariates and then an OLS regression using the set of covariates chosen by Lasso. In the Lasso regression, the penalty level is chosen by the theory-based rule by Belloni et al. (2012), which is also applicable with heteroscedacity.

D Online appendix: Variable definitions

Control state (for wine or liquor) and beer, wine, and liquor state-level excise taxes. Data about control states and state-level alcohol excise tax rates is obtained from The Urban-Brookings Tax Policy Center.[18] The dataset includes yearly data from 2004–2017. It includes indicators of whether the state is a wine control state and liquor control state. It also includes excise tax rates per gallon of beer, wine, and liquor. It does not include tax rates on wine and beer for control states (because the control states set the alcohol prices). We deflate tax rates to 2015 dollars using the consumer price index for urban consumers. For all the variables, we calculate averages across all years in 2004-2017. In the OLS regressions on figure C.1, dependent variables are logarithms of the average (across years) alcohol taxes plus one cent.

Percent of grocery stores selling beer, wine, or liquor. To calculate these measures, we use data from NielsenIQ Retail Scanner Data from 2006-2017.[19] We start with a sample of all food (grocery) stores that in a given calendar year sold at least one unit of commonly bought grocery products: milk (dairy refrigerated milk) or bread. For each store and year, we measure whether the store sold at least a unit of a given alcohol product (beer, wine, or liquor). Then for each year and alcohol product, we calculate the share of stores selling it. Finally, for each product, we take the simple average of the shares across years. We define beer products as all beer products excluding near beer. We exclude from wine products non-alcoholic wine and wine-flavored refreshments. Liquor products consist of brandy, gin, rum, tequila, vodka, and whiskey, excluding all products that can have lower alcohol content than regular liquor (for example, ready-made cocktails).

Alcohol price. To generate the state-level aggregate alcohol prices, we use NielsenIQ Consumer Panel from 2004-2017. We calculate prices in the following steps. First, we convert all the prices into comparable units (price per unit of ethanol). Second, we calculate the weighted average price of ethanol separately for each product group (beer, liquor, and wine) for each year and state. It is calculated as the weighted average of brand and product size pair prices in a given product group, year, and state. The weight of a brand and product size pair

[18]The Urban-Brookings Tax Policy Center. 2020. "State Alcohol Excise Taxes". (`https://www.taxpolicycenter.org/statistics/state-alcohol-excise-taxes`, accessed July 8, 2020.)

[19]While we would like to use the same time period as in our main dataset, years 2004–2005 are not available in NielsenIQ Retail Scanner Data. This is unlikely to be a problem because over time there are not many changes in alcohol regulation that could affect alcohol availability.

is equal to its share in volume in that product group in a given year. To allow market shares to evolve, the weights vary across years. Each year weights are the same across states, to avoid that aggregate price is artificially cheaper in states where consumers purchase cheaper brands. We don't use price data on local brands, to be able to compare prices of the same products across states. Third, we calculate weighted average price of ethanol for each state and year, averaging across the product groups (beer, liquor, and wine). The weight of a product group is equal to its share in volume (including all brands, local and non-local) in a given year. Again, weights are constant across states and vary over years. Fourth, we deflate all the prices to 2015 dollars using the consumer price index for urban consumers and take simple average over years. We have also calculated aggregate prices in alternative ways, and it has led to similar estimates.