Beverage purchases from stores in Mexico under the excise tax on sugar-sweetened beverages: observational study

M. Arantxa Colchero, PhD; acolchero@insp.mx
Barry M. Popkin, PhD; popkin@unc.edu
Juan A. Rivera, PhD; jrivera@insp.mx
Shu Wen Ng, PhD; shuwen@unc.edu

SUPPLEMENTAL MATERIALS

Supplemental Table 1. Beverage categories and levels for Consumer Packaged Goods beverage products purchased by Mexican households

Level 1*	Level 2	Level 3**
Taxed beverages	Sodas taxed	Sodas taxed
	Other taxed beverages (e.g., flavored water or sweetened juice)	Flavored water taxed
		Sweetened juices taxed
Untaxed beverages	Carbonated drinks untaxed (e.g., diet sodas and sparkling water)	Carbonated drinks untaxed
	Still plain water untaxed	Still plain water untaxed
	Other untaxed beverages (e.g., unsweetened dairy beverages, 100% fruit juices, flavored water without caloric sugars, beer)	Dairy without added sugar untaxed
		Flavored water untaxed
		Juices untaxed
		Beer untaxed
		Other untaxed

*In this study we only present purchases and prices for levels 1 and 2.
**Level 3 beverage categories are most similar to the 2012 Encuesta Nacional de Salud y Nutrición (ENSANUT) categories.

Technical Appendix. Difference-in-Difference (DinD) Fixed Effects Models and Predicted Outcomes

Since the Mexican SSB tax was implemented nationally, it is not possible to construct a true experimental design to study the association between the tax and purchases. Therefore we applied a pre-post quasi-experimental approach using difference-in-difference (DinD) analyses along with fixed effects models (1, 2). Fixed effects models have a number of advantages, the key being that they account for the non-time-varying unobserved characteristics of households (e.g., preference for certain types of beverages). The model adjusts for the preexisting downward trend of purchases of taxed beverages observed since 2012 and for macroeconomic variables that can affect household purchases. We wanted to determine whether there were significant changes in the trends in beverage purchases during the posttax period compared to the pretax period after controlling for household composition and contextual factors. We constructed a counterfactual for what the purchases in the posttax period would have looked like in the absence of the tax and compared the observed posttax purchases to this counterfactual, holding all other factors constant.

The distribution of beverage purchases per capita were skewed and not normally distributed, so we used the logarithm (log) of beverage purchases as outcomes. The continuous explanatory variables were more normally distributed and did not require any transformations. The model specification is:

$$\log(BEV_{hsmy}) = \beta_T T_{hmy} + \beta_M M_{my} + \beta_{TM}\left(T_{hmy} * M_{my}\right) + \delta Q_{qy} + \vartheta SES_{hsy} + \gamma H_{hsy} + \varphi C_{sy} + \alpha_{hs} + \mu_{hsmy}$$

The outcome is the log of the average volume of beverage *BEV* purchased per capita per day by household *h* living in state *s* during month-year *my*. *T* denotes the posttax period, *M* denotes the month-year linear time trend (a continuous measure from 1 to 36), *Q* denotes quarters to account for seasonality in purchases, *SES* denotes socio-economic status , *H* denotes the vector of year-specific household characteristics, *C* denotes contextual measures (state-month level unemployment rate and state-quarter level consumer price index adjusted minimum salary), α denotes the unobserved time-invariant characteristics of each household, and μ denotes the time-varying error. β_{TM} denotes the difference between the change in the log average per capita per day volume of *BEV* purchased during the posttax period compared to the pretax period. β_M denotes the pretax trend in the purchase of *BEV*, and the posttax trend in the purchase of *BEV* will be ($\beta_M + \beta_{TM}$).

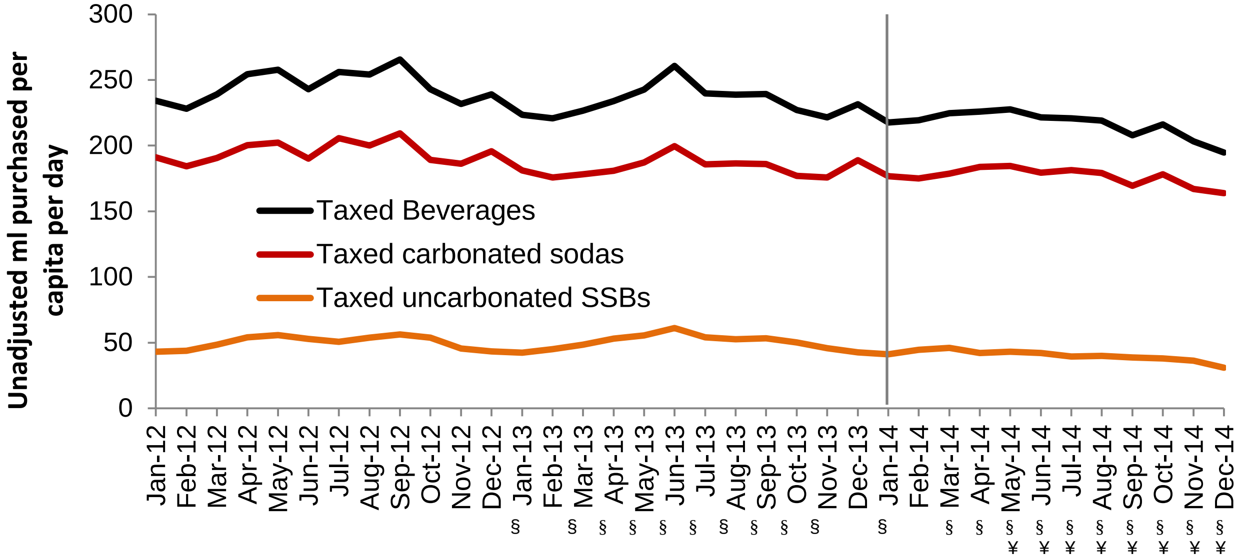
To allow for interpretability in these coefficients, we back-transformed the logged outcomes by calculating and applying Duan smearing factors (3). Specifically, Duan smearing ensures that in the presence of nonzero variances in the volume purchased, the back-transformed predicted outcome is not downward biased (3). This also allowed us to compare in absolute and relative terms the estimated posttax volume purchased in January through December 2014 to the estimated counterfactual posttax volume assuming a pretax trend. We considered presenting predicted values that also detrended seasonality by setting all quarters to the same quarter, but these seasonal trends are interesting and more accurately reflect the changing demand for beverages over the course of the year. We also corrected the standard errors by clustering the analyses at the household level. We conducted all analyses with Stata 13 (4).

For beverage categories where ≥10% of the household quarter observations did not report purchases (taxed sodas and carbonated drinks, other taxed SSBs, and untaxed still plain water), we applied time-varying inverse probability weights to the fixed effects model using -areg, absorb- in Stata (4). We estimated the inverse probability weights from longitudinal (random effects) probit models to address the potential selection bias associated with the probability of purchasing (5). In the case of untaxed carbonated drinks (e.g., diet sodas and sparkling water), because only 27% of the household month observations reported purchases, we used a longitudinal probit model to estimate the probability of purchasing any untaxed carbonated drinks, adjusting for demographic and household composition measures, contextual factors, and region.

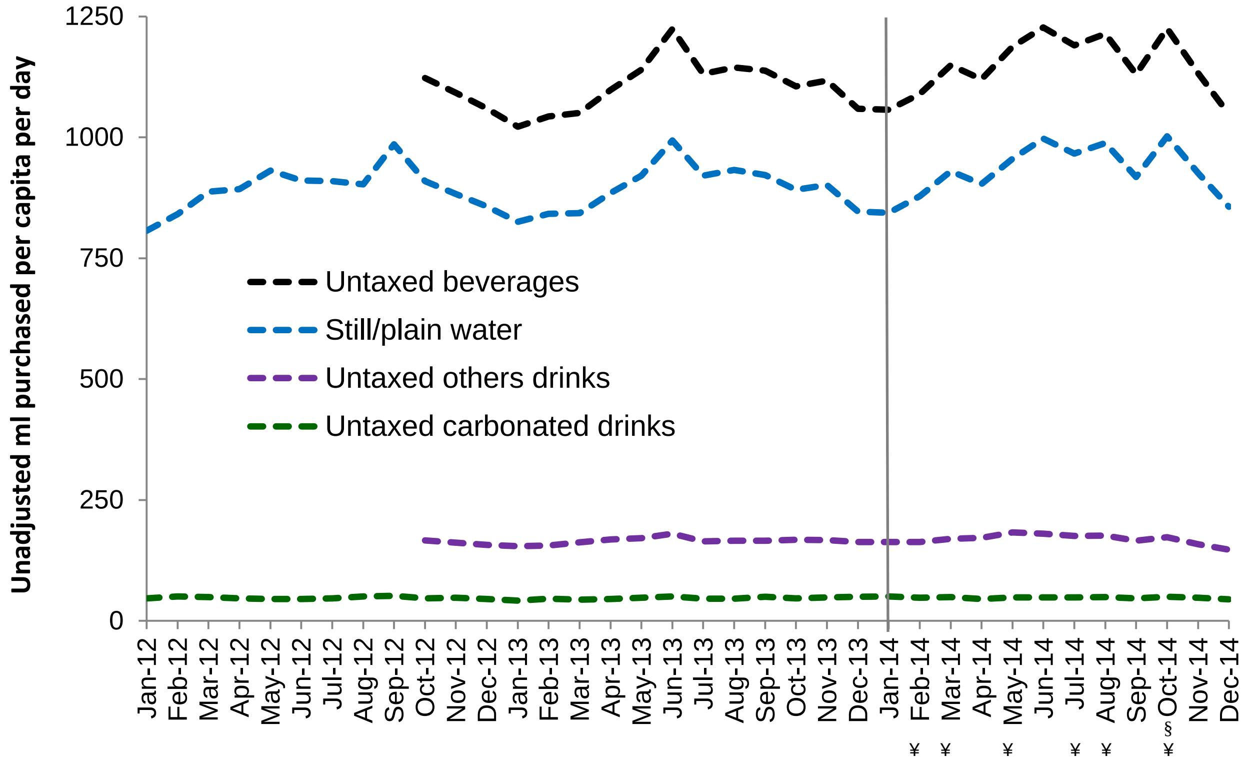
For the models stratified by SES, we used the same modeling approach with the exception of removing household SES from the models and ran three separate models for each outcome for each for the SES subsamples. We based the three SES categories (low, middle, and high) on a six-category measure that the Nielsen Company derived from annually updated questions on household asset ownership (e.g., number of half and full bathrooms in the home, number of bedrooms in the home, number of vehicles owned) and the education of the head of the household.

Supplemental Figure 1. Monthly unadjusted purchases (ml/capita/day) of taxed and untaxed beverages

A. Taxed Beverages



B. Untaxed beverages



[§]Statistically significant difference from the same month in 2012 at p <0.01; [¥] statistically significant difference from the same month in 2013 at p <0.01. Incomplete data for dairy beverages in Jan-Sept 2012.

Source: Authors' own analyses and calculations based on data from Nielsen through its Mexico Consumer Panel Service (CPS) for the food and beverage categories for January 2012 – December 2014.

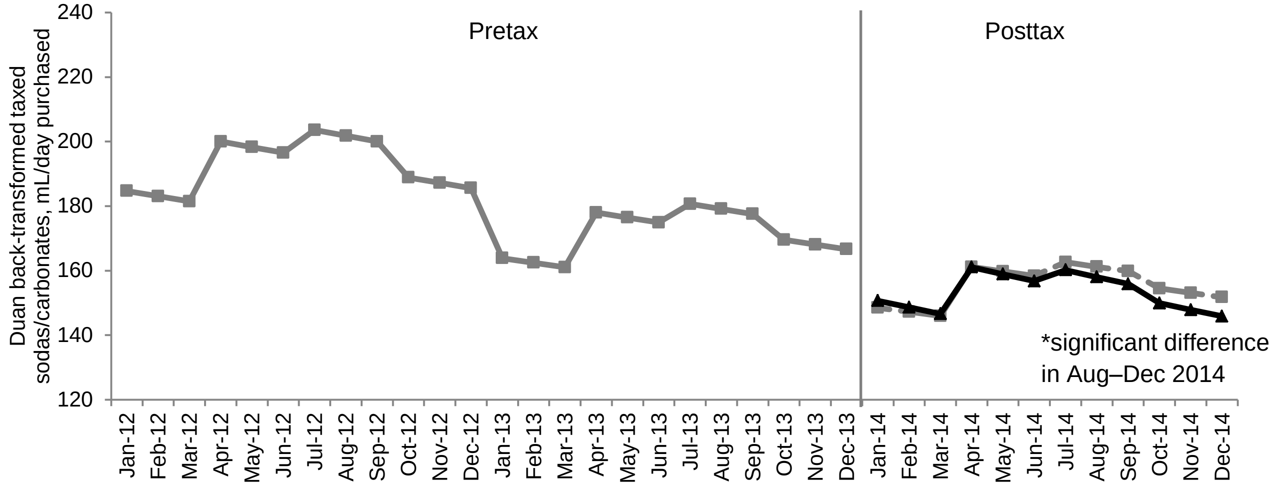
Supplemental Table 2. Coefficient estimates from DinD model results, β (P value)

Beverage outcome	Pretax trend			DinD in trends			Posttax dummy		
	β_M	P		β_{TM}	P		β_T	P	
log(volume purchased taxed beverages)[a]	-0.007	(0.000)	**	-0.015	(0.000)	**	0.254	(0.000)	
log(volume purchased taxed carbonated drinks)[a, b]	-0.009	(0.000)	**	-0.005	(0.001)	**	0.131	(0.005)	*
log(volume purchased taxed noncarbonated drinks)[a,b]	-0.003	(0.000)	**	-0.028	(0.000)	**	0.583	(0.000)	**
log(volume purchased untaxed beverages)[a, d]	-0.004	(0.001)	**	-0.006	(0.000)	**	0.258	(0.000)	**
log(volume purchased untaxed water)[a, b]	0.003	(0.000)	**	-0.011	(0.000)	**	0.383	(0.000)	**
log(volume purchased untaxed other)[a, d]	-0.004	(0.000)	**	-0.011	(0.000)	**	0.327	(0.000)	**
Pr(any untaxed carbonated drinks)[c]	-0.003	(0.002)	*	-0.004	(0.116)		0.115	(0.143)	

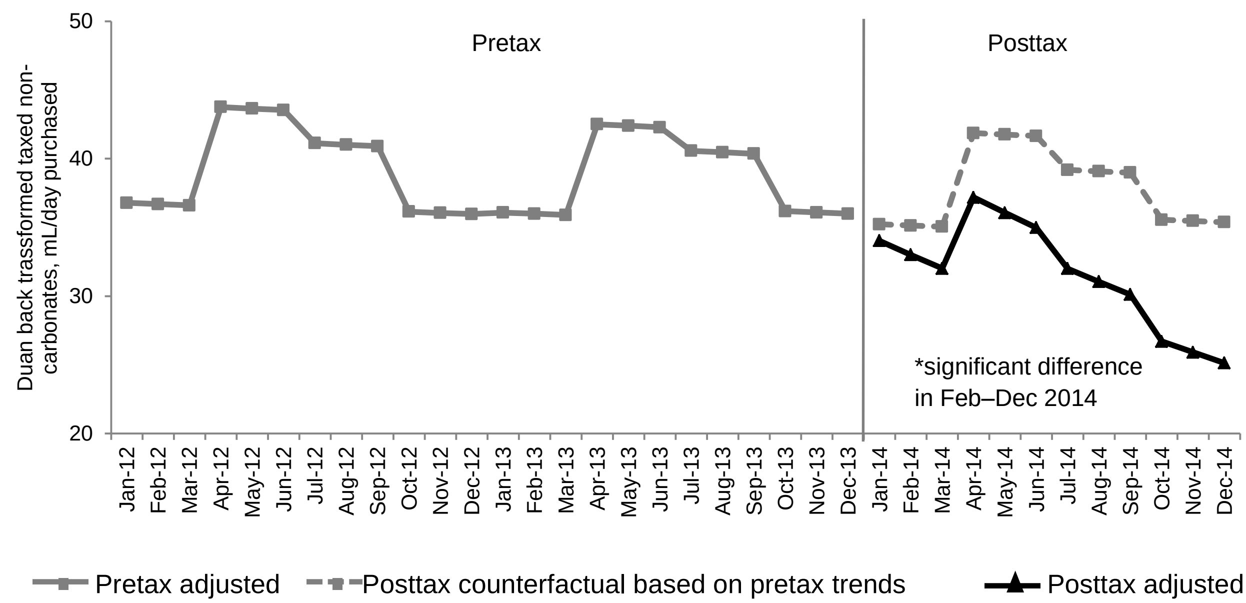
[a] Fixed effects model that uses the log(*BEV* volume) = f(mthyr, posttax, posttax*mthyr, quarter, contextual measures, household composition, household SES) clustered by household. Unless otherwise noted, 36 months of data, n = 205,112 observations from 6,253 households.
[b] Due to >10% nonpurchasing household month observations, the model also accounts for time-varying inverse probability weight for probability of purchasing said beverage in given month with fixed effects in Stata using -areg, absorb-.
[c] Random effects model of the probability of purchasing untaxed carbonated drinks.
[d] Limited to October 2012–December 2014 (27 months of data only); n = 153,387 observations from 6,239 households.
* Statistically significant at p <0.01; ** statistically significant at p <0.001.
Source: Authors' own analyses and calculations based on data from Nielsen through its Mexico Consumer Panel Service (CPS) for the food and beverage categories for January 2012 – December 2014. Copyright © 2015, The Nielsen Company. Nielsen is not responsible for and had no role in preparing the results reported herein.

Supplemental Figure 2. Monthly predicted purchases of taxed sodas and carbonated drinks and taxed noncarbonated SSBs comparing the counterfactual to posttax

A. Taxed sodas/carbonated drinks



B. Taxed noncarbonated SSBs



* Statistically significant at p <0.01. Predictions do not adjust for quarter in order to show seasonal trends in beverage purchases. Back-transformation of predicted log(*BEV* volume) from DinD fixed effects models used Duan smearing factors to handle potential heteroskedasticity.

Source: Authors' own analyses and calculations based on data from Nielsen through its Mexico Consumer Panel Service (CPS) for the food and beverage categories for January 2012 – December 2014. Copyright © 2015, The Nielsen Company. Nielsen is not responsible for and had no role in preparing the results reported herein.

Supplemental Table 3. Coefficient estimates from SES stratified DinD models

Lowest SES[1]	Pretax trend		DinD in trends		Posttax dummy	
	β_M	P	β_{TM}	P	β_T	P
log(volume purchased taxed beverages)[a]	-0.004	0.075	-0.017	0.000**	0.374	0.000**
log(volume purchased taxed carbonated drinks)[a, b]	-0.005	0.001**	-0.009	0.006**	0.183	0.061
log(volume purchased taxed noncarbonated drinks)[a, b]	0.001	0.788	-0.035	0.000**	0.784	0.000**
log(volume purchased untaxed beverages)[a, d]	-0.003	0.203	-0.005	0.193	0.186	0.064
log(volume purchased untaxed water)[a, b]	0.010	0.000**	-0.012	0.004**	0.310	0.018*
log(volume purchased untaxed other)[a, d]	-0.012	0.000**	-0.008	0.080	0.277	0.020*
Pr (any untaxed carbonated drinks)[c]	-0.002	0.490	-0.008	0.214	0.177	0.375
Middle SES[2]	**Pretax trend**		**DinD in trends**		**Posttax dummy**	
	β_M	P	β_{TM}	P	β_T	P
log(volume purchased taxed beverages)[a]	-0.005	0.000**	-0.015	0.000**	0.369	0.000**
log(volume purchased taxed carbonated drinks)[a, b]	-0.008	0.000**	-0.010	0.000**	0.303	0.000**
log(volume purchased taxed noncarbonated drinks)[a, b]	0.002	0.088	-0.032	0.000**	0.670	0.000**
log(volume purchased untaxed beverages)[a, d]	-0.004	0.011*	-0.010	0.000**	0.420	0.000**
log(volume purchased untaxed water)[a, b]	0.003	0.005*	-0.017	0.000**	0.577	0.000**
log(volume purchased untaxed other)[a, d]	-0.002	0.209	-0.016	0.000**	0.481	0.000**
Pr (any untaxed carbonated drinks)[c]	-0.002	0.322	-0.003	0.455	0.096	0.402
Highest SES[3]	**Pretax trend**		**DinD in trends**		**Posttax dummy**	
	β_M	P	β_{TM}	P	β_T	P
log(volume purchased taxed beverages)[a]	-0.011	0.000**	-0.003	0.415	-0.012	0.892
log(volume purchased taxed carbonated drinks)[a, b]	-0.011	0.000**	0.005	0.080	-0.168	0.067
log(volume purchased taxed noncarbonated drinks)[a, b]	-0.008	0.000**	-0.017	0.000**	0.301	0.003**
log(volume purchased untaxed beverages)[a, d]	-0.003	0.048	0.000	0.852	0.040	0.517
log(volume purchased untaxed water)[a, b]	-0.001	0.535	-0.003	0.468	0.120	0.265
log(volume purchased untaxed other)[a, c]	-0.004	0.026	-0.005	0.109	0.130	0.121
Pr (any untaxed carbonated drinks)[c]	-0.005	0.015	-0.004	0.358	0.099	0.448

[1] 36 months: 37,123 observations from 1,421 households; 27 months: 28,661 observations from 1,416 households.
[2] 36 months: 104,905 observations from 3,794 households; 27 months: 76,989 observations from 3,790 households.
[3] 36 months: 63,084 observations from 2,126 households; 27 months: 47,737 observations from 2,121 households.
[a] Fixed effects model that uses the log(*BEV* volume)= f(mthyr, posttax, posttax*mthyr, quarter, contextual measures, household composition) clustered by household.
[b] Due to >10% nonpurchasing household month observations, the model also accounts for time-varying inverse probability weight for probability of purchasing said beverage in given month with fixed effects in Stata using -areg, absorb-.
[c] Random effects model of the probability of purchasing untaxed carbonated drinks.
[d] Limited to October 2012–December 2014 (27 months of data only), n = 153,387 observations from 6,239 households.
* Statistically significant at p <0.01; ** significant at p <0.001.
Source: Authors’ own analyses and calculations based on data from Nielsen through its Mexico Consumer Panel Service (CPS) for the food and beverage categories for January 2012 – December 2014.

Supplemental Table 4. Differences between the counterfactual and posttax predictions in monthly purchases of beverages in 2014 from SES stratified DinD models

	Low SES		Middle SES		High SES	
Taxed beverages	Absolute difference (ml/day)	% of counterfactual	Absolute difference (ml/day)	% of counterfactual	Absolute difference (ml/day)	% of counterfactual
Jan. 2014	-0.69	-0.4%	4.41	2.4%	-8.40**	-4.2%
Feb. 2014	-3.99*	-2.0%	1.61	0.9%	-8.78**	-4.5%
Mar. 2014	-7.20**	-3.7%	-1.11**	-0.6%	-9.15**	-4.7%
Apr. 2014	-11.72**	-5.3%	-4.31**	-2.1%	-10.54**	-5.0%
May 2014	-15.19**	-6.9%	-7.27**	-3.5%	-10.92**	-5.2%
June 2014	-18.57**	-8.5%	-10.15**	-5.0%	-11.30**	-5.4%
July 2014	-21.43**	-10.0%	-13.31**	-6.4%	-11.73**	-5.7%
Aug. 2014	-24.59**	-11.6%	-16.12**	-7.8%	-12.09**	-5.9%
Sept. 2014	-27.66**	-13.1%	-18.86**	-9.1%	-12.44**	-6.1%
Oct. 2014	-29.03**	-14.5%	-20.66**	-10.5%	-12.32**	-6.4%
Nov. 2014	-31.83**	-16.0%	-23.13**	-11.8%	-12.63**	-6.6%
Dec. 2014	-34.54**	-17.4%	-25.55**	-13.1%	-12.94**	-6.8%
Average over 2014	-18.87**	-9.1%	-11.20**	-5.6%	-11.10**	-5.5%

	Low SES		Middle SES		High SES	
Untaxed beverages‡	Absolute difference (ml/day)	% of counterfactual	Absolute difference (ml/day)	% of counterfactual	Absolute difference (ml/day)	% of counterfactual
Jan. 2014	37.59**	5.0%	97.58**	12.1%	16.73**	1.8%
Feb. 2014	33.80**	4.5%	87.87**	10.9%	16.29**	1.7%
Mar. 2014	30.04**	4.0%	78.33**	9.7%	15.84**	1.7%
Apr. 2014	31.84**	3.6%	85.13**	8.6%	17.70**	1.6%
May 2014	27.39**	3.1%	73.76**	7.5%	17.20**	1.6%
June 2014	23.00**	2.6%	62.58**	6.4%	16.70**	1.6%
July 2014	18.29**	2.1%	52.46**	5.3%	15.77**	1.5%
Aug. 2014	14.08**	1.6%	41.48**	4.2%	15.28**	1.5%
Sept. 2014	9.91**	1.1%	30.69**	3.1%	14.81**	1.4%
Oct. 2014	5.26**	0.7%	18.66**	2.0%	13.40**	1.4%
Nov. 2014	1.56	0.2%	8.99**	1.0%	12.96**	1.3%
Dec. 2014	-2.10	-0.3%	-0.52	-0.1%	12.52**	1.3%
Average over 2014	19.22**	2.4%	53.08**	5.9%	15.43**	1.5%

‡ Analysis only uses data from October 2012 onward due to incomplete dairy data from January 2012 to September 2012.
* Statistically significant at $p < 0.01$; ** statistically significant at $p < 0.001$. Predictions do not adjust for quarter in order to show seasonal trends in beverage purchases. Back-transformation of predicted log(*BEV* volume) from DinD fixed effects models used Duan smearing factors to handle potential heteroskedasticity.
Source: Authors' own analyses and calculations based on data from Nielsen through its Mexico Consumer Panel Service (CPS) for the food and beverage categories for January 2012 – December 2014. Nielsen is not responsible for and had no role in preparing the results reported herein.

SUPPLEMENTAL MATERIALS REFERENCES

1. Athey S, Imbens GW. Identification and Inference in Nonlinear Difference-in-Differences Models. Econometrica. 2006;74(2):431-97.

2. Donald SG, Lang K. Inference with Difference-in-Differences and Other Panel Data. Review of Economics and Statistics. 2007 2007/05/01;89(2):221-33.

3. Duan N. Smearing Estimate: A Nonparametric Retransformation Method. Journal of the American Statistical Association. 1983;78(383):605-10.

4. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.; 2014.

5. Cole SR, Hernán MA. Constructing Inverse Probability Weights for Marginal Structural Models. American Journal of Epidemiology. 2008 September 15, 2008;168(6):656-64.