Associations between Mexico's sugar-sweetened beverages tax and soft drinks consumption in adults: an open cohort longitudinal analysis of the Health Workers Cohort Study

Luz María Sánchez-Romero (https://orcid.org/0000-0001-7951-3965); Francisco Canto-Osorio; Romina González-Morales; M. Arantxa Colchero; Shu-Wen Ng; Paula Ramírez-Palacios, Jorge Salmerón; Tonatiuh Barrientos-Gutiérrez.

APPENDIX 1

Table A1- 1. Distribution of The Health Workers Cohort Study sample by exclusion criteria

Exclusion Criteria	Wave 1: 2003	Wave 2: 2010	Wave 3: 2017
N by wave	1579	2141	1214
Less than 19 years old	169	199	98
Pregnant	5	2	0
SSB consumption not reported	18	44	65
Total Kcal intake <600 kcal	5	25	31
SSB daily intake ≥1500ml	12	9	5
Other (no dates of interview)	2	4	5
Observations with missing co-variates data	165	235	50
Total analytical sample	1203	1623	960

The OLCRE model uses level-1 independent variables in their original form (x_{it}). It assumes that the non-time varying level-2 error term (ϑ_i) depends on (x_{it}) via the mean across observations of the same individual $(\overline{x_i})$. In our case, the model equation was fit as follows (1):

$$g(\mu_{it}) = \beta_w x_{it} + (\beta_\beta - \beta_w)\overline{x_i} + \gamma c_i + \vartheta_i, \text{ where}$$

$g(\mu_{it})$ denotes *the ordered log-odds of being in a given soft drink consumption category for the individual i at time t (vs. being in a lower category of consumption);*
$\beta_w x_{it}$ *denotes the estimated within-individual effect* (β_w) *given repeated measures per individual (i) over time (t), for a vector of time-varying covariates;*
β_β denotes the estimated between individual effect;
$(\beta_\beta - \beta_w)$*denotes the comparison of the within- and between- effects (difference effects),*
$\overline{x_i}$ denotes the vector of means of each time varying covariate across observations of the same individual;
γc_i denotes the coefficient of the variable that varies only between individuals (age at baseline/time invariant covariates)

The OLCRE model produces ordered log-odds of the within and difference effects, that we transformed to proportional odds ratios. For this paper, our main focus is on the within-individual effect $(\beta_w x_{it})$ and time-invariant (γc_i) coefficients of the SSBs tax. The difference-effect coefficients $((\beta_\beta - \beta_w)\overline{x_i})$ refer to the product of subtracting the within -effect coefficient from the between effect coefficient. It tells us whether individual unobserved heterogeneity matters and corrects for it but is not a main outcome of interest. Between-effects are not displayed when using this OLCRE STATA command. We recommend that readers use caution in interpreting the difference-effect log-odds coefficients. Due to the characteristics of our variables (panel data), these coefficients do not offer a contextual effect for our analysis and the interpretation can be deceiving. More information about their understanding can be found in Schunck, R. 2017(1) and Bell, A 2019(2).

Table A1-2. Summary of the ordered logistic regression effects of the sugar sweetened beverages tax on soft drink consumption, adjusted by covariates (n=3786) from the Health Workers Cohort Study.

	Within effects (W)		Difference effects (D)†	
Soft Drink Consumption categories	**Odds Ratio**	**95% Confidence Interval**	**Odds Ratio**	**95% Confidence Interval**
Tax	0.385	(0.159 to 0.937)	1.080	(0.302 to 3.856)
Time (years)	1.076	(0.933 to 0.969)	0.860	(0.740 to 0.999)
Age at baseline*time	0.999	(0.998 to 1.000)	1.001	(0.997 to 1.005)
Education				
Elementary school or less	Ref		Ref	
Middle and high school	0.243	(0.066 to 0.896)	5.632	(1.419 to 22.360)
College and higher	0.161	(0.036 to 0.726)	3.689	(0.763 to 17.846)
Income				
Low	Ref		Ref	
Middle	2.027	(1.245 to 3.300)	0.431	(0.236 to 0.785)
High	2.062	(1.141 to 3.726)	0.385	(0.190 to 0.779)
Temperature (Celsius)	1.023	(0.979 to 1.068)	0.951	(0.867 to 1.044)
Gross domestic product	0.999	(0.999 to 1.000)	1.001	(1.000 to 1.001)
Inflation	0.859	(0.722 to 1.022)	1.098	(0.708 to 1.702)
Age at baseline (years)‡	0.951	(0.933 to 0.969)		

Proportional odds ratios from the OLCRE Model 1. †Difference effects estimated as difference of the estimated within and between-individual coefficients. ‡ Variable that only vary between individuals (time-invariant).

Table A1-3. Summary of the ordered logistic regression effects on soft drinks consumption categories including an interaction between tax and income

	Within effects (W)		Difference effects (D)†	
Soft Drink Consumption categories	**Odds Ratio**	**95% Confidence Interval**	**Odds Ratio**	**95% Confidence Interval**
Tax	0.520	(0.203 to 1.332)	0.691	(0.158 to 3.016)
Income				
Low	Ref		Ref	
Middle	2.477	(1.448 to 4.237)	0.352	(0.186 to 0.664)
High	2.503	(1.319 to 4.750)	0.284	(0.128 to 0.630)
Tax*income			1.518	(0.749 to 3.075)
Low	Ref			
Middle	0.629	(0.066 to 0.384)		
High	0.691	(0.133 to 0.427)		

Time (years)	1.073	(0.942 to 1.222)	0.862	(0.742 to 1.001)
Age at baseline*time	0.999	(0.997 to 1.000)	1.001	(0.997 to 1.006)
Education				
Elementary school or less	Ref		Ref	
Middle and high school	0.241	(0.065 to 0.893)	5.673	(1.425 to 22.588)
College and higher	0.161	(0.036 to 0.728)	3.688	(0.759 to 17.910)
Temperature (Celsius)	1.021	(0.978 to 1.066)	0.952	(0.867 to 1.044)
Gross domestic product	0.999	(0.999 to 1.000)	1.001	(1.000 to 1.001)
Inflation	0.870	(0.731to 1.035)	1.087	(0.700 to 1.686)
Age at baseline (years)‡	0.951	(0.932 to 0.969)		

Proportional odds ratios from the OLCRE Model. †Difference effects estimated as difference of the estimated within and between-individual coefficients. ‡ Variable that only vary between individuals (time-invariant).

.

Table A1-4. Summary of the regression effects on soft drinks consumption categories including an interaction between tax and education

Soft Drink Consumption categories	**Within effects (W)**		**Difference effects (D)†**	
	Odds Ratio	**95% Confidence Interval**	**Odds Ratio**	**95% Confidence Interval**
Tax	0.696	(0.257 to 1.885)	1.447	(0.122 to 17.211)
Education				
Elementary school or less	Ref		Ref	
Middle and high school	0.280	(0.075 to 1.038)	5.943	(1.480 to 23.855)
College and higher	0.183	(0.041 to 0.827)	3.872	(0.769 to 19.511)
Tax*education			0.900	(0.368 to 2.201)
Elementary school or less	Ref			
Middle and high school	0.437	(0.245 to 0.781)		
College and higher	0.518	(0.289 to 0.931)		
Time (years)	1.078	(0.946 to 1.228)	0.855	(0.736 to 0.993)
Age at baseline*time	0.998	(0.997 to 1.000)	1.001	(0.997 to 1.005)
Income				
Low	Ref		Ref	
Middle	2.063	(1.266 to 3.362)	0.425	(0.233 to 0.776)
High	2.127	(1.176 to 3.848)	0.374	(0.184 to 0.760)
Temperature (Celsius)	1.020	(0.977 to 1.065)	0.953	(0.868 to 1.045)
Gross domestic product	0.999	(0.999 to 1.000)	1.001	(1.000 to 1.001)
Inflation	0.877	(0.736 to 1.044)	1.088	(0.700 to 1.690)
Age at baseline (years)‡	0.953	(0.935 to 0.972)		

Proportional odds ratios from the OLCRE model. †Difference effects estimated as difference of the estimated within and between-individual coefficients. ‡ Variable that only vary between individuals (time-invariant).

References

1. Schunck R. Within- and between-cluster effects in generalized linear mixed models: A discussion of approaches and the xthybrid command. The Stata Journal. 2017;17(1):89-115.
2. Bell A, Fairbrother M, Jones K. Fixed and random effects models: making an informed choice. Quality & Quantity. 2019;53:1051-74.

APPENDIX 2
Sensitivity Analysis

Table A2-1. Matched sample characteristics of the Health Workers Cohort sample for each wave of data collection

Variables	Wave 1 2004	Wave 2 2010	Wave 3 2017
N	650	650	650
Age at baseline	47.5 ±12.2	47.5 ±12.2	47.5 ±12.2
Male (%)	21.5	21.7	21.8
Time in the cohort (years)	0	7.0 ± 1.1	12.4 ± 1.0
Education (%)			
Elementary school or less	15.4	14.8	13.7
Secondary school or High-school	38.3	36.8	36.2
College and higher	46.3	48.5	50.2
Income (%)			
Low	33.3	33.3	33.5
Middle	34.0	34.2	41.7
High	32.7	32.6	24.8
Temperature (ºC)	24.9 ± 1.9	25.5 ± 2.1	23.7 ± 2.5
Annual Gross Domestic Product (GDP)	7697.7 ± 363.3	9435.1 ± 366.0	9267.0 ± 100.0
Inflation (%)	4.1 ±0.6	3.3 ±0.2	4.6 ±0.7

Table A2-2 Summary of the ordered logistic regression effects of the sugar sweetened beverages tax on soft drink consumption, adjusted by covariates using the matched Health Workers Cohort Study sample.

Soft Drink Consumption categories	Within effects (W)		Difference effects (D)†	
	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval
Tax	0.261	(0.085 to 0.801)	NULL*	
Time (years)	1.163	(0.845 to 1.212)	0.873	(0.577 to 1.322)
Age at baseline*time	0.999	(0.998 to 1.001)	0.993	(0.965 to 1.021)
Education				
Elementary school or less	Ref		Ref	
Middle and high school	0.249	(0.056 to 1.114)	5.176	(1.000 to 26.792)
College and higher	0.272	(0.047 to 1.591)	2.241	(0.330 to 15.199)
Income				
Low	Ref		Ref	

Middle	1.636	(0.998 to 2.682)	0.589	(0.275 to 1.262)
High	1.725	(0.911 to 3.266)	0.373	(0.152 to 0.910)
Temperature (Celsius)	1.055	(1.000 to 1.112)	0.863	(0.729 to 1.021)
Gross domestic product	0.999	(0.998 to 1.000)	1.002	(1.001 to 1.003)
Inflation	0.855	(0.698 to 1.046)	1.239	(0.606 to 2.532)
Age at baseline (years) ‡	1.012	(0.845 to 1.212)		

Proportional odds ratios from the OLCRE Model 1 using only individuals with consumption data for the three HWCS waves (matched). †Difference effects estimated as difference of the estimated within and between-individual coefficients. ‡ Variable that only vary between individuals (time-invariant).

*All individuals (n=650) in the matched sample were exposed to the tax 1/3 of their time, thus, this model estimates within-individual effects but the between individual component is null, rendering null results for the difference-effects.