Economics Bulletin

Volume 44, Issue 3

Can patterns of household purchases predict the outcome of US presidential elections?

Sabina Crowe
Northeastern University London

Michael Gmeiner
London School of Economics

Sebastian Ille
The American University in Cairo

Abstract

We use NielsenIQ US retail scanner data to show that changes in sales patterns can be used to predict US presidential election results at the county level. Using a probit model, we regress 2016 election results against sales of various products six months prior to the election. We employ the results and the sales data for 2020 to forecast presidential election results in the same year. Comparison to actual election outcomes shows that our work correctly predicts election results in 86.47% of cases across 2,602 US counties. We further study how changes in the consumption of certain goods influences voter turnout as well as Democrat and Republican votes.

The results of this paper are based on the researchers' own analyses calculated based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researchers and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. We thank Northeastern University for supporting this research via a Tier 1 grant.
Citation: Sabina Crowe and Michael Gmeiner and Sebastian Ille, (2024) "Can patterns of household purchases predict the outcome of US presidential elections?", *Economics Bulletin*, Volume 44, Issue 3, pages 1181-1187
Contact: Sabina Crowe - sabina.crowe@nulondon.ac.uk, Michael Gmeiner - m.w.gmeiner@lse.ac.uk, Sebastian Ille - sebastian.ille@aucegypt.edu.
Submitted: August 21, 2024. **Published:** September 30, 2024.

1. Introduction and Literature Review

Various polls and studies to date highlight the differences between liberals and conservatives in terms of their food preferences, pet ownership choices, and propensity towards purchasing particular brands. Silver and Kessel (2021) show democrats are more health focused than their Republican counterparts and are more inclined to dedicate time to their hobbies, to pets, and to being in nature or outdoors. Republican states tend to show higher pet ownership (Coren, 2019) and meat consumption (Hodson and Earle, 2018; Radke, 2020), lower plant-based meat consumption (Yule and Cummings, 2023), a more negative perception of cats (Ivanski et al., 2021), are less fashion oriented (Dante, 2017; Brooke, 2016) and likely to exercise (Kannan and Veazie, 2018), and have a preference for pickup trucks (Howard, 2020).

The question of whether these correlations can be used to predict presidential elections has not been explored. Given the hypothesis that the sales data from shops, and thus consumption pattern of households prior to an election, is indicative of their voting behavior, we link sales data from 35,000 US retail stores to election results at the county level in 2016. We use results from a probit model using 2016 data and sales data from 50,000 retail stores to predict election outcomes for 2020 at an accuracy of 86.47% across 2,602 US counties. Consequently, sales data can serve as a strong indicator of voting behavior six months prior to an election. We further ascertain whether earlier findings regarding the differences in consumption choices between liberal and conservative voters are corroborated by the purchasing behavior of US households in 2020.

2. Methodology and Data

Data on presidential elections by county is sourced from the MIT Election Data and Science Lab (MIT, 2022). The retail sales dataset comprises weekly pricing and volume information from over 90 retail chains representative of the entire US encompassing 35,000 grocery and other stores in 2016 to 50,000 in 2020 and is provided by NielsenIQ. These data represent over half of the US grocery and drug store sales volume and more than 30% of mass merchandiser sales, and cover retail outlets across every market in the United States, excluding Alaska and Hawaii.

Weekly product information is available for 2.6-4.5 million unique products spanning food, non-food grocery items, health and beauty aids, as well as selected general merchandise, and is grouped into 1,100 product categories, further aggregated into 125 product groups and 10 departments. We identify 19 categories of items shown in table I that have been demonstrated to be susceptible to political preferences in the literature. After calculating the monthly sales of each of the 19 categories for each of the 2,569 counties in 2016 and 2,602 counties in 2020, we study the correlation between sales patterns and presidential election outcomes at the county level in 2016 using a probit model containing all or some of these categories. The dependent variable is a binary variable that takes the value 1 if the Democratic candidate received more votes than the Republican candidate. Given that US presidential elections take place in November, we use expenditures from January to May. Since elections at the county level are won by Democrats or Republicans, the omission of other political parties does not impact our analysis.

We use the coefficients from the 2016 logistic regression model with the best fit, and the sales data from January to May 2020 to predict the probability that the Democratic candidate would win the election in a county in 2020 and assign a win if the probability is above 0.5 and a loss if

the probability is lower. Finally, we use OLS regressions to determine correlation between the 19 item categories and voter turnout. This allows us to better understand the differences between our results and prior literature.

3. Findings and Discussion

Model (1) in table I shows that greater sales of books and magazines, tobacco and accessories, seafood, fresh produce, juice, and outdoor products are associated with a higher probability of that county voting for Democrats. Dairy and eggs are associated with a higher probability of a Democrat election victory in a county, but estimates are statistically insignificant.

Contrary to previous literature, greater spending on alcoholic beverages, coffee and tea, men's toiletries, health related products, and hobbies are correlated with a more likely win of a Republican candidate. The other predictors are not significant in the probit regression, hence we cannot confirm or refute findings from previous literature regarding the correlation between political preferences and consumption.

Model (2) contains all independent variables from model (1) that are significant at the 5% level in addition to dairy and meat (which previous literature identified as relevant). Most of the coefficients that were significant in model (1) continue to be significant and maintain the same sign as in the original model. The probit model (3) only uses the variables that had significant coefficients at the 5% level in model (1) while model (4) only retains the variables that had significant coefficients in model (3). Model 5 includes only those categories that were significant at the 1% level in all other models. It has the best fit in terms of the Bayesian Information Criterion (BIC) and results in the greatest predictive power when applied to 2020 data. All variables maintain significance and sign. Other combinations of regressors have lower fit and predictive power when applied to 2020 data.

Figure 1 shows the results of prediction when applying the estimated coefficients of model (5) to the 2020 retail data and comparing predictions to actual election results. Counties in green are predicted correctly, in red incorrectly, and in white are those counties for which no NielsenIQ sales data are available and which are sparsely populated (541 out of $3,143$). Our model is able to accurately predict the result of the 2020 US presidential election in 86.47% of the 2,602 counties.

A number of the signs of the probit coefficients do not agree with previous findings on political preferences and consumption. Our model produced negative coefficients for the purchases of health-related products, men's toiletries, and coffee and tea. The difference may be explained by a discrepancy between party identification and voters' actions caused by abstention. We therefore study the relationship between consumption and overall votes, votes for Republicans, and votes for Democrats.

Results of the three OLS regressions in table II shed some light on the reason for our different results. For example, consumption of coffee and tea seems to increase abstention of democratic voters and increase the willingness of Republicans to vote for their party at significant levels. Similar results hold for fresh produce sales, which increase overall abstention at the same time. Increases in sales of men's toiletries reduce the willingness of democrats to vote for their party. However, larger consumption of health-related products is associated with more voting across both parties and cannot explain the difference between our findings and earlier results.

Table I: Probit model results for 2016 presidential elections assessing the odds of a Democrat victory at the county level

	Democrat Win				
Model	(1)	(2)	(3)	(4)	(5)
Health-Related Products	-3.04**	-0.94	-1.18		
	(1.36)	(0.95)	(0.72)		
Pet Products	-5.49				
	(5.01)				
Tobacco and Accessories	6.45***	5.13***	5.53***	5.84***	5.86***
	(1.24)	(1.08)	(1.12)	(1.01)	(1.02)
Alcoholic Beverages	-14.14***	-14.24***	-14.99***	-12.13***	-12.64***
	(3.76)	(3.45)	(3.51)	(2.68)	(2.76)
Books and Magazines	43.24***	46.79***	47.78***	44.06***	45.91***
	(9.03)	(7.63)	(7.58)	(7.23)	(7.29)
Hobbies	-2.91**	-0.38	-0.80		
	(1.39)	(1.47)	(1.17)		
Dairy	0.46	0.16			
	(7.33)	(7.06)			
Meat	0.23	-0.99			
	(1.33)	(0.86)			
Seafood	2.32***	1.47**	1.13*		
	(0.74)	(0.61)	(0.59)		
Eggs	136.46				
	(89.29)				
Fresh produce	21.69**	18.96**	17.82**	16.15**	
	(10.06)	(8.22)	(7.22)	(7.45)	
Coffee and Tea	-6.83***	-6.33***	-5.97***	-6.12***	-4.82***
	(1.63)	(1.54)	(1.53)	(1.54)	(1.50)
Automotive	-0.44				
	(0.32)				
Baby Products	-0.09				
	(0.15)				
Kitchen Gadgets	7.72				
	(30.20)				
Juice	42.72***	38.92***	41.48***	32.77***	33.37***
	(13.35)	(11.94)	(9.29)	(5.61)	(5.67)
Prepared Foods and Snacks	-1.25				
	(1.00)				
Men's Toiletries	-4.94***	-4.81***	-4.43***	-2.51***	-2.22***
	(1.29)	(1.17)	(1.11)	(0.51)	(0.51)
Outdoor Products	17.74***	16.90***	15.45***	13.97***	13.68***
	(3.93)	(3.62)	(3.38)	(3.06)	(3.10)
N	2,569	2,569	2,569	2,569	2,569
BIC	2024	1987	1973	1958	1957

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table II: OLS regression results on voter turnout in the 2020 US presidential elections

	Total Votes	Democrat Votes	Republican Votes
Health-Related Products	8.48***	6.18***	2.30**
	(1.96)	(1.46)	(1.09)
Pet Products	-1.68	7.90	-9.58
	(10.25)	(9.03)	(6.94)
Tobacco and Accessories	0.38	0.81	-0.43
	(0.79)	(1.09)	(0.80)
Alcoholic Beverages	6.27	5.04	1.23
	(4.49)	(3.93)	(3.21)
Books and Magazines	-1.16	14.59	-15.75
	(25.57)	(19.80)	(20.40)
Hobbies	9.84*	6.12	3.71
	(5.40)	(3.93)	(3.46)
Dairy	83.11***	77.45***	5.66
	(18.64)	(17.79)	(14.08)
Meat	-3.80	-3.18	-0.62
	(2.60)	(2.08)	(1.42)
Seafood	-4.68***	-0.89	-3.79***
	(0.99)	(0.86)	(0.67)
Eggs	60.81	46.43	14.39
	(107.74)	(99.64)	(98.55)
Fresh Produce	-17.27***	-12.06***	-5.21**
	(5.12)	(3.97)	(2.29)
Coffee and Tea	2.20	-3.34***	5.54***
	(1.39)	(1.16)	(0.87)
Automotive	1.38**	-0.11	1.48***
	(0.69)	(0.52)	(0.43)
Baby Products	0.81*	0.45	0.37
	(0.43)	(0.30)	(0.26)
Kitchen Gadgets	-5.95	-12.55	6.60
	(109.29)	(85.93)	(78.25)
Juice	-1.65	-32.37	30.71
	(37.92)	(31.38)	(25.64)
Prepared Foods and Snacks	5.32***	0.58	4.74***
	(2.06)	(1.78)	(1.44)
Men's Toiletries	-1.78	-2.57**	0.79
	(1.19)	(1.04)	(0.82)
Outdoor Products	5.49	9.56***	-4.08**
	(3.41)	(3.12)	(1.83)
N	2,602	2,602	2,602
R-squared	0.980	0.976	0.939

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

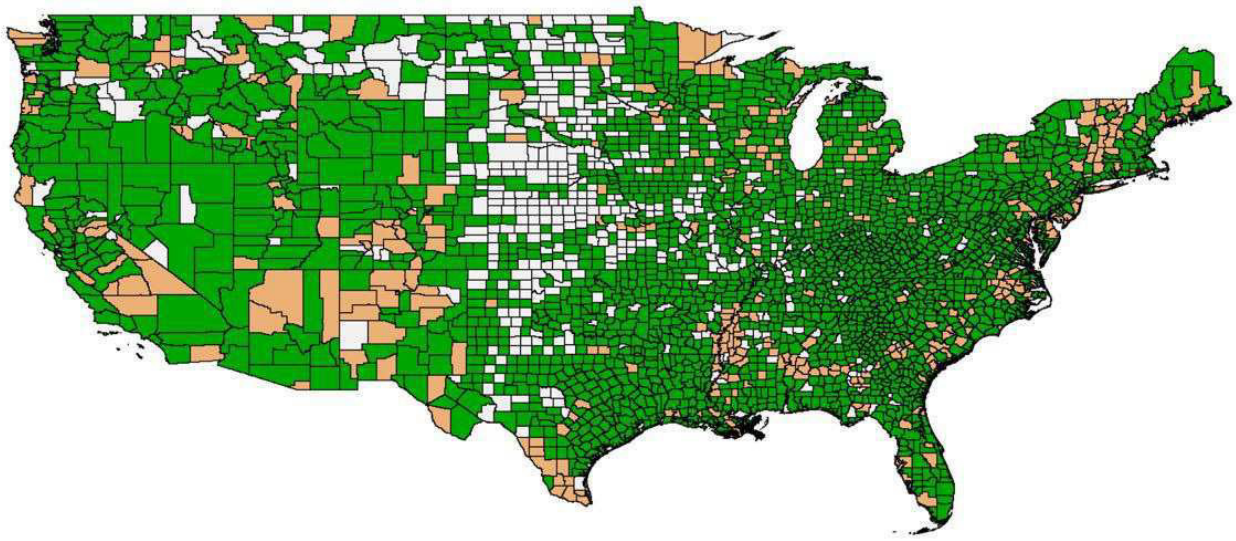

Figure 1: Model predictions of 2020 presidential election outcomes at the county level.
Legend: Green = Correct; Red = Incorrect; White = Missing Data

4. Conclusion

The contributions of the paper are twofold. First, we link sales data to political preferences and provide a model that predicts election outcomes at the US county level in 2020 with 86.47% accuracy.[1] Secondly, we evaluate the relationship between voter turnout and sales. We show that increased sales of some products might lead to abstention. While this can explain some differences with the literature, we are unable to confirm earlier results for some product categories. This study relies on an expansive dataset and offers a more nuanced analysis compared to existing literature. Some limitations remain. Data was unavailable for 541 counties and data limitations did not allow us to include other potentially relevant categories, such as plant-based foods.

5. Disclaimer

The results of this paper are based on the researchers' own analyses calculated based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researchers and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

[1] There may be regional variation in both consumption patterns and voting. If so, prediction accuracy of our model may only be a result of capturing regional trends rather than consumption directly predicting voting patterns. To address this, we included state fixed effects. Prediction accuracy was 82%.

6. Acknowledgements

We thank Northeastern University for supporting this research via a Tier 1 grant.

7. Disclosure of interest

There are no conflicts of interest to declare.

References

Brooke, E. (2016). "Do Democrats and Republicans Actually Shop Differently?". Available at: https://www.racked.com/2016/11/5/13522724/democrats-republicans-shop-stats.

Coren, S. (2019). "Dog Ownership Predicts Voting Behavior—Cats Do Not". Available at: https://www.psychologytoday.com/gb/blog/canine-corner/201908/dog-ownership-predicts-voting-behavior-cats-do-not.

Dante, C. (2017). "Liberal or Conservative? Where You Shop Reveals How You Vote". Available at: https://www.nbcnews.com/storyline/business-of-the-holidays/liberal-or-conservative-where-you-shop-reveals-how-you-vote-n831536.

Hodson, G. and M. Earle (2018). "Conservatism Predicts Lapses from Vegetarian/Vegan Diets to Meat Consumption (Through Lower Social Justice Concerns and Social Support)". *Appetite 120*, 75–81.

Howard, B. (2020). "Vehicles And Voting: What Your Car Might Say About How You'll Vote". Available at: https://www.forbes.com/wheels/news/what-your-car-might-say-about-how-you-vote/.

Ivanski, C., R. F. Lo, and R. A. Mar (2021). "Pets and Politics: Do Liberals and Conservatives Differ in Their Preferences for Cats Versus Dogs?". *Collabra: Psychology* (1), 28391.

Kannan, V. D. and P. J. Veazie (2018). "Political Orientation, Political Environment, and Health Behaviors in the United States". *Preventive Medicine 114*, 95–101.

MIT (2022). "County Presidential Election Returns 2000-2020". Available at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VOQCHQ.

Radke, A. (2020). "A Unique Poll Shows Correlation Between Food & Voting". Available at: https://www.beefmagazine.com/farm-life/unique-poll-shows-correlation-between-food-voting.

Silver, L. and P. V. Kessel (2021). "Both Republicans and Democrats Prioritize Family, but They Differ Over Other Sources of Meaning in Life". Available at: https://www.pewresearch.org/short-reads/2021/11/22/both-republicans-and-democrats-prioritize-family-but-they-differ-over-other-sources-of-meaning-in-life/.

Yule, J. A. and K. H. Cummings (2023). "Conservative Consumer Disinterest in Plant-based Meat: A problem of Message Incongruence". *Appetite 187*, 106574.