When Do Quality Claims Increase Sales? The Moderating Role of Objective Quality in the Beverage Market

Alina Khindanova

Fall 2025

Abstract

This paper examines the effect of quality claims on beverage sales in U.S. convenience stores and how this effect differs based on the product's objective quality. Using 2024 consumer transaction data from PDI Technologies merged with objective product quality measures from food score datasets, I estimate the impact of quality claims on sales. I construct a fixed-effect model with state and week fixed effects, where identification comes from cross-sectional product differences. The initial results show that quality claims have positive effect for objectively lower quality beverages, while having a negative effect for higher quality products. To further validate these findings and explore heterogeneity in treatment effects, I implement a Double Machine Learning framework to estimate conditional average treatment effects as a function of objective product quality.

1 Introduction

Modern industrial food manufacturing involves both food additives, which are intentionally added to enhance sensory or functional properties, and food contaminants, which are unintentionally introduced through processing, packaging, or environmental exposure. Their presence varies across products and manufacturers, and some products are defined as ultra-processed - a category characterized by industrial formulations that rely largely on refined ingredients and additives rather than whole foods. The recent studies and meta-analyses show that synthetic food additives and contaminants are harmful for human health, especially for children (Trasande et al. 2018, Warner 2024, McCarthy 2019), and consumption of UPFs is related to a variety of gut diseases (Whelan et al. 2024), as well as obesity, diabetes, mental illnesses, cardiovascular diseases, and overall mortality (Levy et al. 2024, Pagliai et al. 2021).

Despite their negative health effects and low nutritional value, synthetic additives remain attractive to food producers as a means of preserving and marketing products, since natural alternatives are typically less effective and more costly (Olas et al. 2021). The consumption of products containing food additives increased by 20% between 2001 and 2019 (Dunford et al. 2023), and synthetic dyes was found in 19% of products in the food and beverage market (Dunford et al. 2025).

The US Food and Drug Administration (FDA), the governmental agency responsible for regulating approximately 80% of the U.S. food supply, permits food additive manufacturers to conduct their own safety assessments under the Generally Recognized as Safe (GRAS) framework. As a result, roughly 10,000 food additives are permitted in the U.S. market—a substantially higher number than in the European Union, where only 411 additives are authorized (Harrison, 2024). Consequently, responsibility for avoiding potentially harmful food additives largely falls on consumers. Despite growing interest in healthier and "cleaner" food options (Asioli et al. 2017), consumers often find it difficult to assess product quality based on ingredient lists alone. Instead, purchasing decisions are primarily driven by other product attributes, such as price, labeling, and marketing cues (Zhao et al. 2021).

In this paper, I am trying to understand how specific ingredient quality claims impact sales in the beverage market. As quality claims I consider the presence of words "premium",

”natural”, “100%”, and ”pure” in the name of the product as they signal the quality of the product and the absence of synthetic ingredients. I use UPC-level transaction data on year 2024 from U.S. independent and small convenience store chains taken from the PDI Technologies database (`https://www.deweydata.io/data-partners/pdi-technologies`) and merge it with food quality scores information from the Food Score Database on the nonprofit Environmental Working Group website (`https://www.ewg.org/foodscores/`). I consider the beverage category in my analysis as it is one of the largest categories to contribute to the artificial food additives sales (Dunford et al. 2025) and is both usually consumed by children and present in convenience stores. To understand the relationship between quality claims and sales and the moderating effect of objective quality, I construct the fixed effects model with week and state fixed effects. The results show that quality claims in the name of the product have positive effect for products with objectively lower quality, and negative effect for products with higher quality. In addition, I support my findings with a causal understanding of the effect using double machine learning to estimate conditional average treatment effect. While many recommendations are concentrated on how to improve food quality on the policy level through changes in regulations (e.g., Whelan et al. 2024), I believe my analysis will be beneficial for marketing specialists and policymakers to better understand the effect of quality signaling on sales.

The structure of the paper is as follows. First, I provide a literature review on past studies that looked at the quality information asymmetry in the market, especially in food markets, as well as studies on the consumer purchasing behavior in the food markets, and explain how my study adds to this literature. Then, I delve deeper to describe my sources of data and provide the summary statistics. In the following sections, I describe the empirical strategy and results, and the last section concludes my analysis.

2 Literature Review

My study of the effect of labeling on sales adds contribution to two bodies of literature. First, I add to the large body of literature that studies the consumer behavior in the presence of asymmetric information. There are papers that try to understand the effect of signaling in these markets and advertising, as well as the moderating effect of objective quality (Kopalle et al. 2018). I add to this analysis by studying the effect of quality claims on sales and the moderating effect of objective quality for the beverage market.

Second, I add to the general discussion of how food product labeling affects consumer perception and purchase intention. Most previous studies have concentrated on the effect of nutritional labeling, such as the impact of disclosing saturated fats and sugar on sales. Some papers studied the effect of voluntary nutritional labeling in the US before mandatory nutrition labeling reform (Mojduszka and Caswell 2000), and there are many papers on the reform itself and its effect on population health; as well as the effect of labeling policies in other countries (Araya et al. 2022). The effect of nutrition label is higher for "healthy" categories, e.g. in Chile effect is significant for cereals but not significant for cookies and chocolates, meaning that pre-existing knowledge matters a lot (Maesen et al. 2022, Barahona et al. 2023). While mandatory nutrition labeling is on the back of the package, there are many additional studies that are trying to figure out how (and which) additional front-of-package label may impact consumer purchasing decision (figuring out best practices), welfare and health (Crosbie et al. 2023, Lim et al. 2020, Oswald et al. 2022). There is also a meta-study that confirms that nutrition labeling is a good way to improve healthy choices of consumers (Roberto et al. 2021).

Most studies that examined the influence of labeling that signals quality of ingredients, rather than nutritional value, are concentrated on the organic label and organic products. Organic food is becoming more popular among consumers, and there is research on why people buy those products (Pant et al. 2024, Sahelices-Pinto et al. 2021, Hughner et al. 2007). Overall, consumers tend to perceive organic products as higher quality and assume they are healthier (Massey et al. 2018, Gundala and Singh 2021).

There are only a small number of papers that considered the effect of labeling that does not have established requirements, for example, the experiment by Skubisz (2017), where they found that putting the words "natural" makes people believe that the product is healthier to consume. Additional analysis on how words "natural" affects consumer perceptions in the juice market (Musicus et al. 2025) found that natural claims (such as "all natural") increase purchase intention and several other related characteristics. Additional study found that added credibility did not make producers meet labeling criteria, but instead made them drop the label (Ippolito 2003). In my analysis, I look at the similar problem, but instead of running an experiment, I conduct a field study, where I study how quality claims affects sales.

3 Data

For my analysis, I merge two sources of data: consumer transaction data and food quality scores data.

Consumer Transaction Data. The main dataset is taken from the PDI Technologies consumer transactions data. The dataset, titled "Consumer Transactions Daily Aggregation," was accessed from the Dewey portal. The data covers transactions from convenience stores across the United States. I took data from the year 2024, and only for packaged beverages (non-alcoholic). I merged the main dataset with the additional dataset on the location of the store, which will allow me to control for state in my analysis. The total number of stores is 34,163. Then, I grouped data by product identifier (GTIN), state, and week, to get the weekly sales for each product (GTIN) in each state.

Food Quality Scores Data. I scraped data from the Beverage section of the Environmental Working Group Food Scores database (`https://www.ewg.org/foodscores/`). There are such beverages as coffee, frozen juices, fruit and vegetable juices and drinks, iced

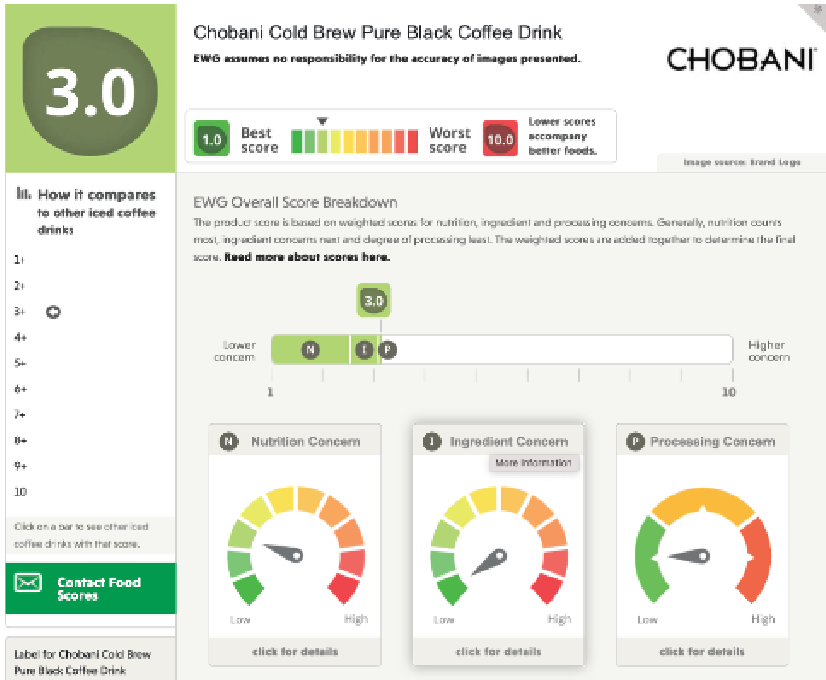


Figure 1: Food Scores Example

teas, powdered drinks, and sodas. I collected information on the name of the product, its GTIN, as well as total score, nutrition, ingredient, and processing scores, as shown in the picture below.

The main variable of interest is the Ingredient Concern score, as it captures the quality of the product. It captures "the likely presence of key contaminants, pesticides, hormones, and antibiotics and possible health implications of food additives" (`https://www.ewg.org/foodscores/content/methodology-ingredients/`). It ranges from 1 to 10, where 10 is the best possible quality.

After getting data on both datasets, I merged them based on GTIN. As the food score database does not cover all the possible beverages in the market, some elements did not merge – from more than 5 million initial observations, I left with 713,367 final observations. Though additional analysis is needed to overcome the attrition, this attrition mostly happened because EWG food scores data do not capture some weight but capture another, or because it captures single or pack but not both, though the product itself is the same. Additionally, I currently constructed price variable as the ratio between total revenue and total sales for a particular product within a week, working on getting regular price data in the future.

Table 1: Summary Statistics

Variable	Mean	SD	Min	Max
Total Sales	276.26	1277.55	1.00	53744.00
Total Revenue	740.13	3270.20	0.01	125470.04
Nutrition Score	5.20	1.44	1.00	9.00
Ingredient Score	3.76	1.16	0.00	8.00
Processing Score	1.99	0.45	1.00	3.00
Price	3.20	1.95	0.01	114.49

4 Empirical Strategy

I consider three different variations for the objective ingredient quality variable. Besides considering the initial 10 point objective quality measure (INGR), I constructed two binary variables based on score distribution: first is equal to 1 if score is >=5 (INGR_high1); second is equal to 1 if score is >=4 (INGR_high2).

Additionally, I constructed a variable that captures whether there is a presence of word signaling quality in the name of the product. It equals 1 if the product name has words "Premium", "Natural", "100%", or "Pure". (For an additional robustness checks I constructed other binary variables based on the subset of key words: {"Premium", "Natural", Health", "Healthy", "100%", "Real", "Pure"} and {"Premium", "Natural", "Health", "Healthy", "Organic", "100%", "Real", "Pure"} and {"Health", "Healthy", "Organic"}).

I consider the following model:

$$\begin{aligned}\ln(\text{Sales}_{ist}) = \beta_0 + \beta_1\,\text{QualityFlag}_i + \beta_2\,\text{IngredientScore}_i \\ +\beta_3\,(\text{QualityFlag}_i \times \text{IngredientScore}_i) + \beta_4\,\text{Price}_{ist} + \beta_5\,\text{Group}_i \\ +\beta_6\,\text{NutritionScore}_i + \gamma_t + \delta_s + \varepsilon_{ist}\end{aligned} \tag{1}$$

where QualityFlag is the dummy variable of quality claims in the beverage product name (I consider four different specifications based on the cutoff value described above); IngredientScore is the dummy variable for the objective quality of the product (I use two different specifications based on the cutoff value described above); Price is the proxy of price for the product; Group is the category of the product (coffee, juice, soda etc.); Nutrition Score is defined in the EWG website as the healthfulness of the product based on the presence of sodium, sugar, etc.; gamma and delta is week and state fixed effects respectively.

5 Results

The output in table 2 shows the effect of main variable of interest on log sales for different moderators. I excluded some control variables from the table to make the output more concise. Here, I also used clustered standard errors based on GTIN.

As we can see, the effect is objective ingredient quality is positive given all other characteristics the same. In contrast, the main effect of quality claims is small and statistically insignificant in the binary-quality specifications and only marginally significant when ingredient quality is modeled continuously. The interaction between quality claims and objective quality is negative and statistically significant in all models. This indicates that the effectiveness of quality claims diminishes as objective quality increases.

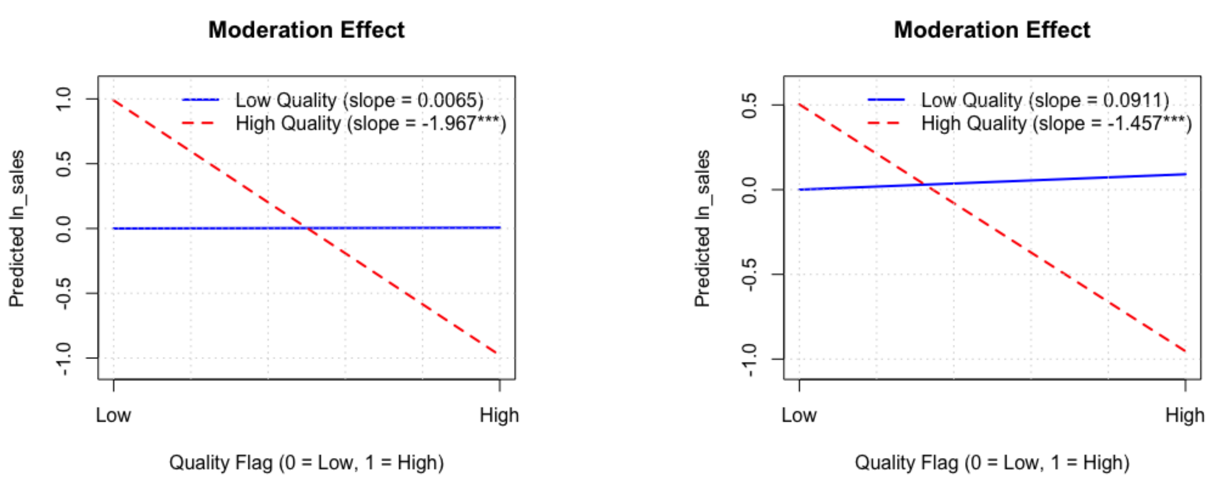
The effect of price is negative and significant in all three specifications, and the effect of nutritional quality is positive and significant at 5% level.

Unfortunately, using OLS, we can see that the value of treatment and moderation effects really depends on the the way how treatment variable is measured. It could be dew to the fact that the values of the INGR variable are highly concentrated between values 3 and 5, with some observations having scores 1 or 10, which can bias OLS estimates. Therefore, additional analysis is needed to overcome this issue.

Table 2: The effect of quality claims on sales. Dependent variable: log(Sales)

Variable of interest:	INGR_high1	INGR_high2	INGR
"premium/natural/100%/pure"	0.006	0.091	1.154*
	(0.255)	(0.276)	(0.595)
INGR_high1	0.988***		
	(0.203)		
"premium/natural/100%/pure" × INGR_high1	−1.973***		
	(0.559)		
INGR_high2		0.503**	
		(0.195)	
"premium/natural/100%/pure" × INGR_high2		−1.548***	
		(0.429)	
INGR			0.327***
			(0.111)
"premium/natural/100%/pure" × INGR			−0.423***
			(0.163)
Num.Obs.	713 319	713 319	713 319
R2	0.143	0.114	0.119
R2 Adj.	0.142	0.114	0.119

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001



(a) Moderator: INGR_high1

(b) Moderator: INGR_high2

Figure 2: Simple slopes of quality claims by product quality.

6 Heterogeneous Quality Claims Effects

To study how objective product quality moderates the effect of quality claims on sales, I estimated Group Average Treatment effects (GATEs) based on double/debiased machine learning framework (Chernozhukov et al. 2018).

As before, I partitioned products based on their objective quality, and defined "high-quality" group as products that have ingredient score greater or equal to 4. As controls, I used available sets of variables such as price, product category, nutrition score, and state. Random forests were used to estimate the probability that a product carries a quality label (the propensity score), and expected sales outcomes for labeled and unlabeled products, conditional on observed variables.

Table 3: Group Average Treatment Effects of Quality Labeling

Group	**GATE (Log Sales)**	**95% Confidence Intervals**
INGR <4	-0.1878	(-0.260; -0.115)
INGR >=4	-0.6396	(-0.676; -0.603)
Difference	-0.4518	(-0.533; -0.370)

Table 3 reports group average treatment effects of ingredient quality claims on log sales. Labeling reduces sales on average in both groups; however, the effect is substantially larger in magnitude for products with higher quality. The difference in GATEs indicates that labeling is significantly more detrimental for objectively higher-quality products.

7 Conclusion

This study can be improved and extended in several directions. First, the data used in the analysis could be extended to include information on the weight of the product, price, as well as which products got discounts.

Second, while this analysis focuses on quality claims inferred from product names,such information may not fully capture the set of quality signals presented to consumers at the point of purchase. Therefore, it would be interesting to consider the additional front-of-package information to the analysis.

Lastly, I am planning to estimate the effect of quality claims separately for each beverage product category. For example, consumers purchasing sodas may be less responsive to quality-related labels than consumers purchasing juices or other beverages perceived as healthier.

8 References

Araya, S., Elberg, A., Noton, C., & Schwartz, D. (2022). Identifying food labeling effects on consumer behavior. Marketing Science, 41(5), 982-1003.

Asioli, D., Aschemann-Witzel, J., Caputo, V., Vecchio, R., Annunziata, A., Næs, T., & Varela, P. (2017). Making sense of the "clean label" trends: A review of consumer food choice behavior and discussion of industry implications. Food research international, 99, 58-71.

Barahona, N., Otero, C., & Otero, S. (2023). Equilibrium effects of food labeling policies. Econometrica, 91(3), 839-868.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters.

Crosbie, E., Gomes, F. S., Olvera, J., Patiño, S. R. G., Hoeper, S., & Carriedo, A. (2023). A policy study on front–of–pack nutrition labeling in the Americas: emerging developments and outcomes. The Lancet Regional Health–Americas, 18.

Dunford, E. K., Galligan, T. M., Taillie, L. S., & Musicus, A. A. (2025). All the Colors of the Rainbow: Synthetic Dyes in US Packaged Foods and Beverages in 2020. Journal of the Academy of Nutrition and Dietetics.

Dunford, E. K., Miles, D. R., & Popkin, B. (2023). Food additives in ultra-processed packaged foods: an examination of US household grocery store purchases. Journal of the Academy of Nutrition and Dietetics, 123(6), 889-901.

Gundala, R. R., & Singh, A. (2021). What motivates consumers to buy organic foods? Results of an empirical study in the United States. Plos one, 16(9), e0257288.

Harrison, R. (2024). How a Legal Loophole Allows Unsafe Ingredients in U.S. Foods. https://www.nyu.edu/about/news-publications/news/2024/august/legal-loophole-unsafe-ingredients.html

Hughner, R. S., McDonagh, P., Prothero, A., Shultz, C. J., & Stanton, J. (2007). Who are organic food consumers? A compilation and review of why people purchase organic food. Journal of Consumer Behaviour: An International Research Review, 6(2-3), 94-110.

Ippolito, P. M. (2003). Asymmetric information in product markets: looking to other sectors for institutional approaches. American Journal of Agricultural Economics, 85(3), 731-736.

Levy, R. B., Barata, M. F., Leite, M. A., & Andrade, G. C. (2024). How and why ultra-processed foods harm human health. Proceedings of the Nutrition Society, 83(1), 1-8.

Lim, J. H., Rishika, R., Janakiraman, R., & Kannan, P. K. (2020). Competitive effects of front-of-package nutrition labeling adoption on nutritional quality: evidence from facts up front–style labels. Journal of Marketing, 84(6), 3-21.

Maesen, S., Lamey, L., Ter Braak, A., & Jansen, L. (2022). Going healthy: how product characteristics influence the sales impact of front-of-pack health symbols. Journal of the Academy of Marketing Science, 50(1), 108-130.

Massey, M., O'Cass, A., & Otahal, P. (2018). A meta-analytic study of the factors driving the purchase of organic food. Appetite, 125, 418-427.

McCarthy, C. (2019). Common food additives and chemicals harmful to children. https://www.health.harvard.edu/blog/common-food-additives-and-chemicals-harmful-to-children-2018072414326

Mojduszka, E. M., & Caswell, J. A. (2000). A test of nutritional quality signaling in food markets prior to implementation of mandatory labeling. American Journal of Agricultural Economics, 82(2), 298-309.

Musicus, A. A., Jensen, M. L., Mita, C., Winters, D. R., Roberts, M. T., Mancini, S., ... & Krieger, J. W. (2025). The relationship between fruit drink front-of-package claims, fruit imagery, and ingredient disclosures and consumer perceptions, intentions, and behavior: A systematic review. Journal of the Academy of Nutrition and Dietetics.

Olas, B., Białecki, J., Urbańska, K., & Bryś, M. (2021). The effects of natural and synthetic blue dyes on human health: A review of current knowledge and therapeutic perspectives. Advances in Nutrition, 12(6), 2301-2311.

Oswald, C., Adhikari, K., & Mohan, A. (2022). Effect of front-of-package labels on consumer product evaluation and preferences. Current research in food science, 5, 131-140.

Pagliai, G., Dinu, M., Madarena, M. P., Bonaccio, M., Iacoviello, L., & Sofi, F. (2021). Consumption of ultra-processed foods and health status: a systematic review and meta-analysis. British journal of nutrition, 125(3), 308-318.

Pant, S. C., Saxena, R., Pant, D. K., & Singh, R. (2024). Organic food consumption: A bibliometric–content analysis. Journal of Consumer Behaviour, 23(5), 2730-2750.

PDI. (2022). Transaction Items [Dataset]. Dewey Data. `https://doi.org/10.82551/2EZQ-RR86`

Roberto, C. A., Ng, S. W., Ganderats-Fuentes, M., Hammond, D., Barquera, S., Jauregui, A., & Taillie, L. S. (2021). The influence of front-of-package nutrition labeling on consumer behavior and product reformulation. Annual review of nutrition, 41(1), 529-550.

Sahelices-Pinto, C., Lanero-Carrizo, A., & Vázquez-Burguete, J. L. (2021). Self-determination, clean conscience, or social pressure? Underlying motivations for organic food consumption among young millennials. Journal of Consumer Behaviour, 20(2), 449-459.

Skubisz, C. (2017). Naturally good: Front-of-package claims as message cues. Appetite, 108, 506-511.

Trasande, L., Shaffer, R. M., Sathyanarayana, S., Council on Environmental Health, Lowry, J. A., Ahdoot, S., ... & Woolf, A. D. (2018). Food additives and child health. Pediatrics, 142(2), e20181410.

Warner, J. O. (2024). Artificial food additives: hazardous to long-term health?. Archives of disease in childhood, 109(11), 882-885.

Whelan, K., Bancil, A. S., Lindsay, J. O., & Chassaing, B. (2024). Ultra-processed foods and food additives in gut health and disease. Nature Reviews Gastroenterology & Hepatology, 21(6), 406-427.

Zhao, H., Yao, X., Liu, Z., & Yang, Q. (2021). Impact of pricing and product information on consumer buying behavior with customer satisfaction in a mediating role. Frontiers in psychology, 12, 720151.