To Reduce or to Structure: on Mixed Method Complementarity

James Sears, Michigan State Unversity, searsja1@msu.edu

Invited Paper prepared for presentation at the

2024 Agricultural & Applied Economics Association Annual Meeting

New Orleans, LA

July 28-30, 2024

To Reduce or to Structure: on Mixed Method Complementarity

James Sears*

June 15, 2024

Abstract

Researchers seeking to identify a causal treatment effect of interest often gravitate toward reduced-form modeling approaches, while those interested in characterizing the structure of demand gravitate toward structural models of the full market environment. In this paper we demonstrate that, rather than operate as perfect substitutes, reduced-form and structural approaches can play a complementary role in icharacterizing market dynamics and policy implications. These opportunities are particularly ripe in regards to questions of food policy and marketing strategy impacts, as researchers frequently must balance the need to fully characterize demand and potential feedback loops with the desire to interpret estimated objects in a causal fashion. We provide an example of the complementary use of mixed empirical methods: first we utilize an event study framework to estimate the changes in alcoholic beverage and non-alcoholic beer purchasing in response to the adoption of county-level COVID-19 stay-at-home policies. Second, we estimate a structural model of differentiated alcoholic beverage products that can provide novel insight into the substitution at play behind the growth of the non-alcoholic beer market. Taken together, the results from these two empirical approaches provide novel insight into recent dynamics in the alcoholic beverage market and carry important implications for future food and beverage marketing strategies.

*Michigan State University. Email: searsja1@msu.edu.

1 Introduction

Researchers seeking to identify a causal treatment effect of interest often gravitate toward reduced-form modeling approaches, while those interested in characterizing the structure of demand gravitate toward structural models of the full market environment. In this paper I argue that, rather than operate as perfect substitutes, reduced-form and structural approaches can play a complementary role in informing market dynamics and policy implications. These opportunities are particularly ripe in regards to questions of food policy and marketing strategy impacts, as researchers frequently must balance the need to fully characterize demand and potential feedback loops with the desire to interpret estimated objects in a causal fashion.

In this paper we provide an example of the complementary use of mixed empirical methods. First, we utilize an event study framework to estimate the changes in alcoholic beverage and non-alcoholic beer purchasing in response to the adoption of county-level COVID-19 stay-at-home policies. This approach allows us to recover estimates of the average changes in both total sales dollars and sales volume across the range of alcoholic beverage types (beer, wine, and liquor) and document the rise in sales within the non-alcoholic beer category over the same period. Subsequent heterogeneity and mechanism analysis provides further insights into the role of local demographic characteristics and market structure on changes in alcoholic beverage purchasing patterns.

After discussing results of the event study approach, we use the gained insight to inform the design of a structural model of the alcoholic beverage market, modeling both non-alcoholic beer and alcoholic beverages as differentiated products (Berry, Levinsohn, and Pakes 1995) to model the state of substitutability and complementarity between the different product categories at multiple points in time relative to the COVID-19 pandemic. Results from the structural model reveal the extent to which the growth in the non-alcoholic beer segment arose due to substitution away from other alcoholic products (and which products carried the greatest degree of substitution) or as a result of complementarity with alcoholic beverage consumption. These findings - along with the reduced form results - provide critical insight into the market dynamics for alcoholic beverages and the growth of non-alcoholic alternatives.

This paper makes several contributions to the literature. First, it contributes to the literature on consumer for alcoholic beverages, providing novel insight into the changes to consumption patterns induced by COVID-19 stay-at-home policies. Anecdotal evidence suggests considerable growth in the sale of non-alcoholic beer (Dickson 2021; Bandoim 2020). This growth was confirmed by Nielsen, who reported a 44% year-over-year increase in non-alcoholic beer sales for the week to May 9, 2020 (Hancock 2020). Currently limited evidence exists providing a direct link between COVID-19 conditions and changes in household's alcoholic beverage consumption or the mechanisms underpinning these changes. An online survey by Pollard, Tucker, and Green (2020) found that the frequency of alcohol consumption increased by 16% relative to individuals' 2019 baseline consumption, with another survey confirming that 40% of individuals reported changes to their alcoholic beverage consumption (Wittenberg et al. 2022). Results from a repeated cross-section indicated that, while 60% of respondents reported increasing drinking during COVID-19 (with those experiencing COVID stress reporting increased volume and frequency of consumption), 13% reported drinking less than pre-pandemic levels (Grossman, Benjamin-Neelon, and Sonnenchein 2020). The reduced-form results of this paper provide the first causal estimates of aggregate changes to alcohol consumption as a result of stay-at-home mandates, with results shedding light on the roles that both overall stress and access issues played in informing consumption behavior among affected households. Finally, both the reduced form and structural model results document the rise of the non-alcoholic beer consumption and provide insight into whether this grown arose through substitution away from or complementary consumption with traditional alcoholic beverages.

2 Data

We employ the NielsenIQ retail scanner data obtained through the James M. Kilts Center for Marketing at the University of Chicago Booth School of Business as our primary data source (henceforth Nielsen-Kilts data). The Nielsen-Kilts data provide information on sales prices and volumes at the universal product code (UPC) by week level for 30,000 - 50,000 participating retail stores from roughly 90 different retail chains each year. In addition, a subset of stores report marketing information including their use of features and displays. These data cover every major market within the United States; stores can be linked geographically to a particular Designated Market Area (DMA) and county. As we focus on identifying patterns in alcoholic beverage purchasing and substitution behavior in response to the early stages of COVID-19, we obtain data files for 2019 and 2020.

We combine several Nielsen-Kilts data files to form our main analysis dataset. First, we download the Movement files for the alcoholic beverage department for 20119 and 2020. The Movement files provide information on the number of units of each UPC sold in a given week along with the sales price, and price multiplier (i.e. 3 for $5). In addition, a unique store code is provided identifying the specific retail location. Binary indicators for whether the product was featured or part of a marketing display that week are included for the subset of stores reporting this information. Next, we merge in the Products file containing UPC-level information including text descriptions of the UPC and brand, details on the product size, quantity within the product's packaging (i.e. single bottle or 6-pack), as well as hierarchical categorization variables (i.e. product module, group, and department). Finally, we use the retail store file to gain info on the store's parent chain, the type of retail channel (i.e. Convenience, Drug, Food, Mass Merchandiser, or Liquor store), and it's geographic location (state, county, and DMA).

While the Nielsen-Kilts data represent the most comprehensive source of panel information on food and beverage sales in the United States, there are several important considerations for the study of alcoholic beverage sales. First, laws regarding alcohol sales vary across states: some state allow for beer, wine, and liquor sales in grocery stores while others allow only beer sales within grocery stores, while still others allow for no alcohol sales within these stores. Second, the Nielsen-Kilts data cover a relatively low share of liquor stores, with the majority of liquor sales observed from food, club, or mass merchandiser stores. As a result, our results our only directly applicable to consumers and retailers in states where in-grocery alcohol sales are permitted. Third, the Nielsen-Kilts data do not measure on-premise alcohol sales such as those occurring at restaurants or bars. In the present context this presents fewer concerns related to measurement error, as access to bars and restaurants was heavily constrained during the majority of the study period. This again limits the interpretation of our results; while the observed sales patterns may reflect direct changes in consumption patterns, we may alternatively be observing fixed consumption levels and pure substitution away from and back to on-premise consumption as a function of local public health ordinances.

In order to focus on alcoholic beverage substitution patterns, we restrict the sample in several ways. We first limit the sample to products classified as beer, wine, or liquor[1]. Next, we further limit the sample to the 18 states that permit the sale of beer, wine, and liquor within grocery stores[2]. This process limits the sample to 18,973 stores in 2019 and 2020. Next, we construct aggregate measures of total dollar sales and total volume by the following product categories: non-alcoholic beer and malt beverages, light beer, other beer, non-alcoholic wine, other wine, and liquor. Finally, we merge in county-level demographic characteristics obtained from the 5-Year American Community Survey and information on the date of county's stay-at-home mandate adoption from Sears et al. (2024).

3 Empirical Framework

In order to demonstrate the complementary roles of reduced-form and structural modeling approaches, we employ two primary analytical approaches: an event study design and a structural model for alcoholic beverage demand (Berry, Levinsohn, and Pakes 1995).

1. Whenever a reported UPC's product module, brand code, pack size, or container size changes, NeilsenIQ flags the updated version as a new UPC version. To limit analysis to uniquely-identified UPCs, we remove all instances of ambiguous UPCs due to ambiguous attribute changes during the reporting period. This filtering drops 998 versions of 447 UPC with pack-size disparities (e.g. switches from a single bottle to a 12 pack). An additional 1,591 versions of 719 UPCs are removed due to unit-size changes (e.g. 12 to 16 ounce). Of the remaining 173,681 UPC versions, 2,294 for 1,060 unique UPCs change product module (e.g. "Beer" to "Ale") and 7,571 versions of 3,719 UPCs exhibit slight variations in brand name. As these underlying products are unchanged, we retain the latest version of these UPCs. This process results in 163,788 unique alcoholic beverage UPCs.

2. We retain data for California, Washington, Nevada, Arizona, New Mexico, South Dakota, Nebraska, Iowa, Missouri, Louisiana, Wisconsin, Illinois, Michigan, Ohio, West Virginia, Maryland, Massachusetts, and Maine.

3.1 Event Study

To directly model the dynamic nature of alcoholic beverage purchasing in response to the spread of COVID–19 and subsequent implementation of county-level stay-at-home (SAH) mandates, we estimate changes in alcoholic beverage sales using a traditional event study model of the form:

$$Y_{jcsw} = \alpha + \sum_{k=\underline{k}}^{\bar{k}} \beta_k \cdot \text{Weeks Since SAH}^k_{csw} + \mathbf{P}_{csw} + \eta_{cs} + \delta_w + \varepsilon_{csw} \quad (1)$$

where Y_{jcsw} is the sales outcome of interest for product category j during the week ending on date w in county c located in state s, while P_{csw} a vector of controls for other COVID-19 policies (state of emergency declarations, school closures, and reopenings). η_{cs} and δ_w are county and week fixed effects, respectively, and ε_{csw} is an idiosyncratic error term capturing remaining unobservable determinants of weekly alcoholic beverage sales in a county and product category. Our inclusion of η_{cs} helps account for time-invariant county characteristics likely to inform alcoholic beverage purchasing behavior (i.e. area, population density, local regulations) while δ_w absorbs seasonal differences in consumption patterns.

Turning to the evaluation of stay-at-home (SAH) mandates, the vector "Days Since SAH^k_{csw}" is comprised of indicator variables defined as equal to one when k weeks have elapsed since the week that a county's stay-at-home mandate first came into effect (Days Since $\text{SAH}^k_{csw} = 1$ when $w = SAH_{cs} + k$) and are zero otherwise. Values of k greater than zero reflect dates following adoption of a stay-at-home mandate in adopting counties, while $k < 0$ corresponds to pre-treatment dates and $k = 0$ indicates the first week that a county's mandate was in effect[3]

We normalize all event-time effects to the week prior to mandate adoption ($k = -1$) and bin endpoints to ensure identification of dynamic treatment effects separate from

3. We define the first week as the first week in which residents of a county were required to stay at home for at least one full workday (either by a county-level mandate or a state-level ordinance, depending on whichever is earlier). For example, if a county's mandate was implemented at 10am on Monday, March 23, we set SAH_{cs} equal to the week ending March 29. If a mandate were instead adopted at 4pm on Sunday, March 29, we assign SAH_{cs} equal to the following week ending on Sunday, April 5. If a county did not adopt its own stay-at-home mandate, we replace the county date with the date of stay-at-home adoption for the state where the county is located. Counties that never adopted mandates and are located in non-adopting states serve as pure control units.

time trends (Schmidheiny and Siegloch 2019). In this way the coefficient β_k yields an estimate of the average sales response k weeks relative to mandate adoption. To control for comparable dynamic responses to other county-level policies, we include vectors of event-time coefficients for each of the three alternate policies.

3.2 Structural Demand Model

3.2.1 Structural Model Identification

Here, as in nearly all structural demand models, the endogeneity of prices stands as the primary barrier to identification (Berry, Levinsohn, and Pakes 1995; Berry and Haile 2016). Following Nevo (2001), we construct our instrumental variable set to leverage the correlation between brand's prices across counties while ensuring the independence assumption holds. For each brand sold in a given focus county, we instrument for the current price with the state-level weekly average price (excluding the focus county). Under this framework, identification requires that, after controlling for brand-specific averages and county-level demographics, the unobserved county-specific portions of the product's valuation are independent across counties. As both cities and counties tend to represent relatively large geographic areas, the likelihood of correlation in valuation between either neighboring counties or all counties within a state is minimized. A violation of this assumption would occur if persistent state-level shocks occurred for certain brands: if residents of southern states valued bourbon more than individuals residing in northern states (in a way not otherwise captured by demographics or heterogeneity), then these pervasive preference differences would induce correlation between the state-specific dummy variables and the idiosyncratic error term.

3.2.2 Differentiated Beverage Model

Below we develop a differentiated goods structural model that will allow estimation of own and cross-price elasticities following the approach from the discrete-choice literature (Berry, Levinsohn, and Pakes 1995; Nevo 2001).

Consider the alcoholic beverage market consisting of $m = 1, ..., M$ markets each with $i = 1, ..., I$ consumers choosing consumption of $j = 1, ..., J$ products. Given the structure of our employed data, we define a market as a county-quarter pair.

The indirect utility received by consumer i in market m from product j is

$$u_{ijm} = x_j\beta_i^* - \alpha_i^* p_{jm} + \xi_j + \Delta\xi_{jm} + \epsilon_{ijm} \tag{2}$$

where x_j is a vector of observable characteristics for beverage j, p_{jm} is the price of the beverage in market m, and ϵ_{ijm} is an idiosyncratic error term. The two ξ terms capture willingness-to-pay for each beverage: ξ_j captures the nationwide mean valuation for econometrically-unobserved beverage characteristics, while $\Delta\xi_{jm}$ allows for deviations away from this national average in every county and every quarter. α_i^* and β_i^* are $K+1$ vectors of coefficients allowed to vary at the individual level.

This framework allows us to both directly model preferences for observable beverage characteristics and simultaneously account for the role that unobservable characteristics play in informing consumer choices. We take advantage of the product characteristics provide in the NielsenIQ data to include the alcohol content, organic label, and beverage category membership as observable product characteristics. The unobserved nationwide component ξ_j is accounted for through a vector of product-specific dummy variables, while the county-specific deviations (accounting for actions like local advertising differences) are absorbed in a composite error term under the assumption of commonality across consumers.

We allow the consumer taste parameters to depend on both observed and unobserved demographic characteristics following Nevo (2001):

$$\alpha_i^* = \alpha + \Pi D_i + \sum \nu_i \tag{3}$$

$$\beta_i^* = \beta + \Pi D_i + \sum \nu_i \tag{4}$$

where D_i is a vector of demographic characteristics, $\sum$ a scaling matrix, and $\nu_i \sim N(0, I_{K+1})$. With the addition of an outside good (such that consumers can choose not to purchase an alcoholic beverage) and the assumption that consumers will choose the product that provides them with the highest utility (and that ties do not occur), we can express the market share for beverage j in market m as

$$s_{jm}(x, p_{.m}, \delta_{.m}; \theta) = \int_{A_{jm}} dP^*(\epsilon) dP^*(\nu) dP^*(D) \tag{5}$$

where $P^*()$ represents the population distribution functions, $.m$ the vectors indexed $1m, ..., Jm$, and θ the vector of nonlinear parameters $(\Pi, \sum)$. To avoid assuming substitution occurs in proportion to market share, we follow Nevo (2001) and estimate θ^* as the solution to a nonlinear GMM estimator solved numerically using a contraction mapping approach.

4 Conclusion

In this paper we discuss the potential tradeoffs and and benefits of employing mixed empirical methods before providing a concrete example of just such a scenario. We focus on studying changes in alcoholic beverage sales and substitution with non-alcoholic beer alternatives during the COVID-19 pandemic, a prime example of a research setting where the combined use of reduced form and structural methods can provide critical insights to questions in the food policy and marketing contexts. The results of the reduced form event study model provide estimates of the average dynamic consumption responses to adoption of county-level stay-at-home policies, shedding new insight into the degrees to which individuals changed their consumption behavior in response to new sources of stress and disruptions to everyday routines. These results also reveal the degree to which these responses persisted during the remainder of the stay-at-home period and reveal the temporal nature of the growth in the non-alcoholic beer sector.

Following estimation of reduced-form methods, our structural model of differentiated beverage products allows for direct estimation of the substitution patterns underpinning alcoholic beverage choices. The use of a structural approach allows for flexible estimation of own and cross-price elasticities, which then allows for a complete understanding of the degree to which non-alcoholic beer's growth served as a complement to current alcoholic beverage consumption or if it acted as a substitute for traditionally-consumed beverages (and reveal which product types were substituted away from). Taken together, the results of the two approaches provide novel understanding of how consumption behavior responded to major changes in daily habits and access conditions, yielding important insights for beverage brands as they seek to position their products in the current product landscape.

References

Bandoim, Lana. 2020. Nonalcoholic drinks grow in popularity during Coronavirus pandemic. *Forbes,* https://www.forbes.com/sites/lanabandoim/2020/10/22/nonalcoholic-drinks-grow-in-popularity-during-coronavirus-outbreak/?sh=3742915516aa.

Berry, Steven, and Philip Haile. 2016. Identification in differentiated products markets. *Annual Review of Economics* 8:27–52.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* 63 (4): 841–890.

Dickson, Andrew. 2021. How the Coronavirus has boosted non-alcoholic beer. *Business Insider,* https://www.businessinsider.com/how-the-coronavirus-has-boosted-non-alcoholic-beer-2021-2.

Grossman, Elyse R., Sara E. Benjamin-Neelon, and Susan Sonnenchein. 2020. Alcohol consumption during the COVID-19 pandemic: a cross-sectional survey of US adults. *International Journal of Environmental Research and Public Health* 17 (24): 9189.

Hancock, Edith. 2020. Nonalcoholic drinks grow in popularity during Coronavirus pandemic. *The Drinks Business,* https://www.thedrinksbusiness.com/2020/05/how-drinking-habits-have-changed-during-lockdown/.

Nevo, Aviv. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* 69 (2): 307–342.

Pollard, Michael S., Joan S. Tucker, and Harold D. Green. 2020. Changes in adult alcohol use and consequences during the COVID-19 pandemic in the US. *JAMA Network Open* 3 (9): e2022942.

Schmidheiny, Kurt, and Sebastian Siegloch. 2019. *On event study designs and distributed-lag models: equivalence, generalization and practical implications.* CESifo Working Paper Series 7481. CESifo. https://ideas.repec.org/p/ces/ceswps/_7481.html.

Sears, James, Sofia B. Villas-Boas, Molly Sears, Carly Trachtman, Andrew Stevens, and Linlin Fan. 2024. Deniers and compliers: unpacking the heterogeneous effectiveness of u.s. stay-at-home mandates during the covid-19 pandemic. *Working Paper.*

Wittenberg, Eve, Collin Labutte, Benjamin Thornburg, Abraham Gebreselassie, Carolina Barbosa, and Jeremy W. Bray. 2022. Alcohol consumption and health-related quality of life in the US during the COVID-19 pandemic: a US national survey. *Journal of Patient-Reported Outcomes* 6 (1): 106.