Title
Economic Effects of Plant-Based Milk on Prices and Quantities of Cow’s Milk, with Milk Policy Impacts

Author
Lee, Sangwon

Publication Date
2024

Peer reviewed|Thesis/dissertation

Economic Effects of Plant-Based Milk on Prices and Quantities of Cow's Milk, with Milk Policy Impacts

By

SANGWON LEE
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Agricultural and Resource Economics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Danial A. Sumner, Chair

Richard J. Sexton

Julian M. Alston

Committee in Charge

2024

Abstract

Plant-based milk alternatives have gained significant attention in recent years, transitioning from a niche market to the mainstream. This dissertation studies the economic impacts of the introduction and wide availability of plant-based milk alternatives on prices and quantities of cow's milk products and the implications for price policy for farm milk.

Chapter 2 examines the implications of the introduction of refrigerated almond milk for the price and quantity of fluid cow's milk and soymilk products sold in retail stores. The entry of almond milk—accounting for about 70% of the quantity share in the plant-based milk market in 2020—into refrigerated shelves was a key to the substantial growth of plant-based milk products in the 2010s. This chapter uses the staggered rollout of refrigerated almond milk across retail stores, mostly between 2008 and 2010, to assess its impact. Empirical results, using recently developed econometric methods, show that soymilk experienced a short-run 6% quantity fall, and organic cow's milk and lactose-free cow's milk saw 3% declines, whereas conventional cow's milk demand remained largely unaffected. These quantity effects align with expectations that products that are expected to be more substitutable for almond milk experienced greater reductions in quantity. The estimated effects indicate that the annual per capita quantity of cow's milk and soymilk decreased by 0.055 to 0.086 gallons. During the 2008–2010 period, the annual per capita quantity of refrigerated almond milk increased by 0.152 gallons, suggesting refrigerated almond milk expanded the overall milk market rather than merely cannibalizing demand for other milk types. Price declines were less than 1% for all product categories.

Chapter 3 explores the extent to which recent declines in consumption of retail fluid cow's milk products are attributable to the availability of plant-based milk. Chapter 3

presents estimates of discrete-choice demands for many cow's milk and plant-based milk products. The econometric estimation uses household purchase data and matches it with store-level scanner data to represent the households' choice sets. The supply side is modeled as an oligopolistic market where processors follow Bertrand-Nash price competition. The estimated demand parameters generate own- and cross-price elasticities of demand for plant-based milk and conventional, lactose-free, and organic cow's milk. With all these models and estimates, the chapter uses counterfactual simulations to find that the removal of all plant-based milk products from the choice set causes a 23% increase in the retail quantity of organic cow's milk, 16% for lactose-free cow's milk and 11% for conventional cow's milk. Over the 2006 to 2020 period, the availability of plant-based milk products accounted for 38% of the historical decline in U.S. cow's milk consumption.

Chapter 4 explores the implications of the farm milk price policy. By using the demand parameters and the marginal costs of processing cow's milk and plant-based milk estimated in Chapter 3, this chapter conducts two milk price policy simulations: (1) an increase in the price of farm milk used for fluid products, as recently recommended by the USDA and (2) the removal of the longstanding above-market farm milk pricing regulation. Results show that increasing the price differentials would raise the price of conventional cow's milk by about 4% and reduce the retail quantity by about 3.6%, placing further pressure on the already declining consumption of fluid cow's milk. However, changes in the relative prices between cow's milk and plant-based milk, induced by the USDA price regulations, are not the primary factor for the declining consumption nor the key to revitalizing the demand. entries

Acknowledgments

I am deeply grateful to my dissertation committee members for their encouragement and guidance throughout the dissertation process. I would like to express my sincere appreciation to my advisor, Dan Sumner, for his unwavering support and inspiring guidance. His approach to analyzing economic problems has profoundly shaped the way I tackle challenges. I am also thankful to Rich Sexton and Julian Alston for thoughtful and constructive feedback on my dissertation research. I would also like to thank Kristin Kiesel for her valuable insights into my research. I am fortunate to have learned from and worked alongside each of them.

My friends and colleagues provided valuable moments in Davis. I am especially grateful to Julian Artega and Seunghyun Lee for their friendship and consistent encouragement. Their support helped me navigate difficult moments and brought joy to both my personal and academic journey. I am also grateful to Hanbin Lee and Ian Xu for their collaboration in research. Additionally, I would like to thank my other Korean friends, Hyunjung Lee, Jisang Yu, and Woongchan Jeon, for their encouragement. I would also like to express my gratitude to Hyunok Lee for her genuine support.

I am deeply thankful to my family in Korea for their love and support. Most importantly, I want to thank my wife, Kyoungeun Lee, for her steady support throughout the challenges and for taking care of my two lovely daughters, Serene and Teryn. Thank you, Serene and Teryn, for filling my life with joy and love.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Plant-based products have become prominent alternatives to traditional fluid dairy products, with sales in U.S. retail stores growing remarkably over the last two decades. The share of plant-based products, relative to the combined sales of fluid cow's milk and plant-based milk at retail stores, increased from 3.1% in 2006 to 11.2% in 2020 (NielsenIQ 2020). The introduction of almond milk to refrigerated shelves in retail stores contributed to the substantial growth of plant-based milk sales, with almond milk reaching about 70% of the quantity share of plant-based milk by 2020 (NielsenIQ 2020).

Fluid cow's milk consumption has steadily declined since at least the 1950s, with a remarkable acceleration of the percentage rate of decline in the 2010s (USDA ERS 2023). The U.S. dairy industry has responded to the expansion of plant-based milk with a variety of public policy initiatives. For example, the National Milk Producers Federation (NMPF) has urged the Food and Drug Administration to restrict plant-based milk manufacturers from using the term "milk" (NMPF 2023). Recent advertising efforts have highlighted the "realness" of cow's milk compared to plant-based milk (for example, see Wood Milk 2023). However, the effectiveness of these efforts to build consumption of cow's milk depends on the extent to which the recent decline in consumption has been caused by the rise of plant-based milk.

This dissertation explores the economic impacts of the introduction and the growing availability of plant-based milk on prices and quantities of cow's milk. The analysis first uses detailed econometric estimation to consider the implications of the geographic spread of almond milk into retail refrigerated shelves. The result was differential shifts in consumption of categories of cow's milk products and soymilk products. The discussion then turns to econometric estimation of how the availability of plant-based milk has driven the decline of fluid cow's milk consumption, using recent household purchase data. Plant-based milk products substitute differentially for particular categories of cow's milk products. I then use my econometrically estimated demand parameters and model of wholesale marginal costs to assess the implications of recent proposals for U.S. farm milk price policy adjustments.

Chapter 2 considers the impacts of the introduction of refrigerated almond milk on the retail price and quantity of fluid cow's milk and soymilk. Chapter 2 identifies the effects using temporal variations in the availability of refrigerated almond milk across geographically dispersed retail establishments. The month of the first launch of refrigerated almond milk differs across stores, with most stores experiencing the introduction between 2008 and 2010. The adoption timing of almond milk was not random but determined by interaction between almond milk manufacturers and retailers. With conditional parallel trend assumptions that account for the selection mechanism behind introduction timing, the empirical analysis uses recently developed econometric approaches to estimate the price and quantity effects of refrigerated almond milk.

Empirical results indicate that soymilk experienced a short-run 6% decline in quantity, while organic and lactose-free cow's milk each saw a 3% quantity decrease. In contrast, demand for conventional cow's milk remained mostly unaffected. These quantity effects align with expectations that products that are expected to be more substitutable for almond milk experienced greater reductions in quantity. Overall, on an annual per capita basis, the estimated effects imply that the quantity of cow's milk and soymilk decreased by a range of 0.055 to 0.086 gallons during the 2008–2010 period. During the same sample period, the annual per capita quantity of refrigerated almond milk increased by 0.152 gallons, indicating that refrigerated almond milk did not merely cannibalize soymilk and cow's milk but served to expand the overall milk market. In contrast to the economically significant quantity effects, the price effects were modest, remaining at less than 1% for all product categories.

Chapter 3 develops econometric estimates of the discrete choice demand model for many cow's milk and plant-based milk products. Household purchase data, providing detailed information on purchases of dairy and non-dairy products, including prices and quantities, is matched with store-level data to represent the households' choice sets. The estimated demand parameters for individual products generate own and cross elasticities of demand for plant-based milk and conventional, lactose-free, and organic cow's milk. The supply side is modeled as an oligopolistic market where processors follow Bertrand-Nash price competition. With all these models and estimates, the chapter conducts counterfactual simulations of the removal of all plant-based milk products from the choice sets.

Results from the demand estimates show that plant-based milk is more substitutable for organic and lactose-free cow's milk products than for conventional cow milk. Correspondingly, the counterfactual simulation reports that the removal of all plant-based milk products from the choice set would lead to a 23% increase in the equilibrium retail quantity for organic cow's milk, a 16% increase for lactose-free cow's milk, and an 11% increase for conventional cow's milk. On an annual per capita basis, the percentage changes correspond to an increase of 1.46 gallons for conventional cow's milk and 0.26 gallons each for organic and lactose-free cow's milk. The counterfactual results indicated that annual U.S. fluid cow's milk consumption would increase by 2 gallons per capita, from 16.3 gallons to 18.3 gallons in 2020, in the absence of availability of plant-based milk in the consumer's choice sets. Out of the 5.1-gallon decrease in annual per capita cow's milk consumption from 2006 to 2020, the counterfactual experiment reveals that the availability of plant-based milk is responsible for about 38% of this drop.

Chapter 4 explores the implications of modification of U.S. farm milk price policy. By using the demand parameters and the marginal costs of processing cow's milk and plant-based milk estimated in Chapter 3, this chapter conducts two milk price policy simulations: (1) an increase in the price of farm milk used for fluid products, as recently included in USDA Federal Milk Marketing Order (FMMO) reform proposals and (2) the removal of the FMMO price regulation that set above-market minimum prices for farm milk used for fluid products.

Results of these simulations show that increasing the farm price of milk for fluid consumer products would raise the retail price of conventional cow's milk products by about 4% and reduce the equilibrium retail quantity by about 3.6%, exacerbating the ongoing decline in the consumption of fluid cow's milk. In contrast, removing FMMO-regulated higher farm prices for fluid milk products would lower the retail price of conventional cow's milk by about 5.6% and increase the quantity used by 5.2%. These simulations indicate that by increasing the prices of cow's milk products relative to plant-based milk, farm milk price regulations have reduced the consumption of cow's milk and increased the consumption of plant-based milk, but the impacts have been small.

In conclusion, this dissertation research found that the introduction and growing availability of plant-based milk have had significant negative impacts on demand for cow's milk, especially organic and lactose-free cow's milk. Nonetheless, plant-based milk has not been the primary driver in the recent decline in consumption of fluid cow's milk. Additionally, even with plant-based milk well established in retail markets, the policy-induced increases in farm milk prices have limited impacts on the consumption of retail fluid cow's milk products.

References

National Milk Producers Federation. 2023. "RE: Docket No. FDA-2023-D-0451 Labeling of Plant-based Milk Alternatives and Voluntary Nutrient Statements." Available online at: https://www.regulations.gov/comment/FDA-2023-D-0451-0870.

U.S. Department of Agriculture, Economics Research Service. 2023. *Food Availability (per capita) Data System*. Washington DC. Available online at: https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/.

Wood Milk. 2023. "*Introducing Wood Milk: Aubrey Plaza's Newest Product* [Video]." YouTube. Available online at: https://www.youtube.com/watch?v=C4klWmwC2ds.

Chapter 2

Impacts of the Introduction of Refrigerated Almond Milk on the Retail Prices and Quantities of Cow's Milk and Soymilk

2.1 Introduction

Sales of plant-based milk products at retail stores in the United States grew remarkably over the past two decades. The share of plant-based milk, relative to the combined sales of liquid cow's milk and plant-based milk at retail stores, increased from 3.1% in 2006 to 11.2% in 2020 (NielsenIQ RMS). This growth was primarily accounted for by the introduction and expansion of refrigerated almond milk, which was first launched in late 2008.

The U.S. dairy industry has been attentive to the growth of the plant-based milk segment. Farm production of cow's milk remains a crucial part of U.S. agriculture. It is the fourth largest commodity in terms of annual farm revenue. Consumption of fluid cow's milk products has steadily decreased since at least the 1950s, with a marked acceleration in the downward trend during the 2010s. The recent steep decline in consumption of cow's milk coincided with the introduction of almond milk to the refrigerated dairy section of retail stores. Stakeholders in the U.S. dairy industry raised concerns about plant-based milk products, which compete directly next to fluid cow's milk in the refrigerated dairy section.

This chapter examines the implications of the introduction of refrigerated almond milk for the prices and quantities of the pre-existing fluid cow's milk and soymilk products sold in retail stores. The extent to which the prices and quantities of existing products are affected by the new product depends on demand and supply conditions, including the degree of substitutability between the existing and new products (Hausman and Leonard 2002, Gentzkow 2007, Vives 2008). Price and quantity responses are likely to vary among specific milk categories, namely, conventional cow's milk, organic cow's milk, lactose-free cow's milk, and soymilk. This study investigates empirically how the price and quantity effects differ across the four categories.

This research proposes a framework, building on Vives (2008), to illustrate how the price and quantity of traditional cow's milk and soymilk are influenced by the degree of substitutability between the new product and the existing products under two causes of the introduction of new product. Specifically, in the context of the introduction of refrigerated almond milk, two relevant causes of introduction are: (1) the reduction of an entry cost to the refrigerated dairy case in retail stores and (2) the growing demand for plant-based milk, broadly.

Extending Vives (2008)'s quantity competition model in a symmetric equilibrium, this research incorporates two asymmetric parameters to reflect the market dynamics of cow's milk and almond milk. First, the growing market demand unilaterally applies to plant-based milk. Second, the substitutability between cow's milk products and almond milk may differ among the products. The quantity competition model with the above asymmetric parameters indicates that, for both causes of the new product introduction, greater reductions in price and quantity are experienced for the existing products that are more substitutable for almond milk. Furthermore, when the growth in the demand for the new product is the primary driver of the introduction, the rate of the decline in the prices and quantities of existing products is steeper for those products that are more substitutable to the new product.

This chapter empirically estimates the price and quantity effects of the introduction of refrigerated almond milk. Using temporal variations in the availability of refrigerated almond milk across retail establishments, this study identifies the effects of the introduction of almond milk. The treatment is defined as the initial month of store-specific launch for the first refrigerated almond milk. The month of the first launch of refrigerated almond milk differs across stores, with most stores adopting the introduction between 2008 and 2010. The preferred specification employs the estimator proposed by Callaway and Sant'Anna (2021), which is robust to the heterogeneous treatment effects across different timing groups.

The key identifying assumption is the parallel trend assumption, conditional on store-specific demographics and retailer-time-fixed effects. The adoption of refrigerated almond milk at the store level was not random; rather, it was determined by the interaction between almond milk manufacturers and retailers. This selection into treatment requires rationales for the selection mechanism compatible with the conditional parallel trend assumption (Ghanem, Sant'Anna, and Wüthrich 2024; Marx, Tamer, and Tang 2024). This research argues that the timing of the adoption of refrigerated almond milk depends predominantly on store-specific characteristics, including demographic factors and the parent company that these stores belong to. For example, Blue Diamond, the first refrigerated almond milk manufacturer, chose Florida due to its significant Hispanic population with higher lactose intolerance rates (Chaker 2011; Franklin-Wallis 2019). Retailers' adoption decisions are less influenced by time-varying unobserved factors in part due to uncertainty faced by retailers about the demand for the new products, as reflected in higher slotting fees paid by companies that want to introduce new products in the U.S. grocery industry (Chu 1992; Lariviere and Padmanabhan 1997; Desai 2000; Sudhir and Rao 2006).

Given the selection mechanism, the conditional parallel trend assumption is plausible if time-varying unobservable factors affecting untreated outcomes have constant means over time (Ghanem, Sant'Anna, and Wüthrich 2024). During the 2008–2010 period, no manufacturers produced both cow's milk and almond milk, which reduces concerns about biased parameter estimates from the endogenous introduction of almond milk in the wholesale market. These data do not have the concern of simultaneous pricing and product launch by multiproduct manufacturers. Potential unobserved time-varying factors like retailer-specific promotions may influence cow's milk prices, but these are controlled through retailer-time-fixed effects, based on the observations of large retail chains' uniform pricing strategies across regions (DellaVigna and Gentzkow 2019).

The empirical analysis uses store-level retail scanner data from NielsenIQ, aggregated to monthly quantity sold and monthly average unit value for each store for each of the four product groups. The availability of refrigerated almond milk is readily available from the same NielsenIQ store data. Store-specific demographic variables are constructed using household demographics from NielsenIQ's Homescan Panel (HMS), combined with the 2007–2011 American Community Survey, following DellaVigna and Gentzkow (2019).

2.2 Overview of Results and Connection to the Empirical Literature

My empirical findings suggest that the introduction of refrigerated almond milk has a more substantial impact on the quantity of cow's milk and soymilk than the associated price effects. Estimated price effects are less than 1% for all product categories, with statistically significant price reductions estimated only in organic cow's milk and soymilk. Most quantity effects are substantial and statistically significant. Soymilk experienced a 6% quantity fall, while organic cow's milk and lactose-free cow's milk saw 3% reductions. In contrast, conventional cow's milk showed almost no quantity effect (in percentage terms) from the introduction of almond milk. These results on percentage quantity effects align with the theoretical model, which indicates that the quantities of existing products that are more substitutable for the new product will experience a more significant percentage reduction.

On an annual per capita quantity basis, the estimated percentage quantity effects imply a reduction in the quantity of cow's milk and soymilk, ranging from 0.055 gallons to 0.086 gallons. Soymilk quantity falls by 0.026 to 0.037 gallons while organic cow's milk falls by 0.021 to 0.036 gallons. Although conventional cow's milk shows almost no percentage change, its quantity accounts for 91% of the combined market. Consequently, the magnitude of changes in its quantity is comparable to other categories, ranging from a decrease of 0.025 gallons to an increase of 0.031 gallons.

During the same period, the annual per capita quantity of refrigerated almond milk increased by 0.152 gallons. This increase exceeds the total decline in cow's milk and soymilk quantities. Therefore, refrigerated almond milk did not merely cannibalize soymilk and cow's milk but served to expand the overall milk market, at least in the initial phase of its introduction from 2008 to 2010.

This research builds on a large body of literature on the effects of new-product introduction. Previous studies on the effects of new products estimate the price and quantity effects using structural demand estimation (Petrin 2002; Gentzkow 2007) or two-way fixed-effect estimation when pre- and post-introduction data are available (Choi, Wohlgenant, and Zheng 2013), or both approaches (Hausman and Leonard 2002). Recent econometric studies indicate that two-way fixed-effect estimation can yield biased estimates for average treatment effects when treatment effects are heterogeneous among timing groups and evolve over time (Callaway and Sant'Anna 2021; Goodman-Bacon 2021; Sun and Abraham 2021). This research contributes to the literature by providing econometric estimates of the price and quantity effects that are robust to such biases from two-way fixed-effect estimates, specifically in the context of cow's milk and plant-based milk.

This research adds to the expanding literature on the demand for plant-based foods and their competition with animal-based foods. Previous studies, including Tonsor, Lusk, and Schroeder (2023) and Zhao et al. (2023) on plant-based meats, as well as Alviola and Capps (2010) and Khanal and Lopez (2021) on plant-based milk, have demonstrated that plant-based products tend to substitute more for certain animal-based protein products than others. Specifically, plant-based meats are more likely to be substituted for chicken than ground beef, while plant-based milk products exhibit greater substitutability with lactose-free milk. Recent studies report a negative correlation in consumption between cow's milk and plant-based milk (Stewart et al. 2020; Slade 2023). By focusing on substitutability with the milk category, this research provides causal evidence on how the introduction of refrigerated almond milk differentially affected the prices and quantities of cow's milk and soymilk, depending on the substitutability.

The remainder of this paper proceeds as follows. Section 2.3 provides some background on the introduction of refrigerated almond milk and patterns of diffusion across different regions in the U.S. Section 2.4 describes the mechanisms of the introduction of a new product and illustrates the economic model of pricing and quantity competition among cow's milk manufacturers. Section 2.5 presents the econometric strategies. Section 2.6 describes the dataset used to estimate the treatment effects. Section 2.7 presents the empirical findings regarding the impact of the introduction of refrigerated almond milk.

2.3 Background to Introduction of Refrigerated Almond Milk

Plant-based milk has a long history across different cultures globally. For example, almond milk was documented as early as the 13th century, and soymilk in the 14th century. Despite the presence of shelf-stable plant-based milk brands, at least from the early 20th century, it was not until 1996 that White Wave's soymilk became the first to be sold in the refrigerated dairy section in retail grocery stores (Shurtleff and Aoyagi 2013).

The introduction of soymilk into the refrigerated section marked a significant milestone, facilitating the entry of other plant-based milk alternatives into the refrigerated dairy section of retail stores. A decade later, in late 2008, Blue Diamond introduced the first refrigerated almond milk, Almond Breeze. Subsequently, in early 2010, Silk, already known for its soymilk, launched its own variant of refrigerated almond milk. Around the same time, Whole Foods Market introduced its private-label organic almond milk into the refrigerated section, recognizing the substantial growth of refrigerated almond milk.

The U.S. plant-based milk industry has witnessed significant expansion since the introduction of refrigerated almond milk in 2008, with distinct growth trajectories across various plant sources. Figure 2.1 illustrates these divergent growth patterns within each plant-based milk category from 2006 to 2020. Initially, soymilk held the largest share of the market segment in the plant-based milk industry until 2008, when the emergence of almond milk products in retail stores began to reshape the industry landscape. Notably, Almond Breeze's debut in the refrigerated dairy section in 2008 catalyzed a swift proliferation in both the di-

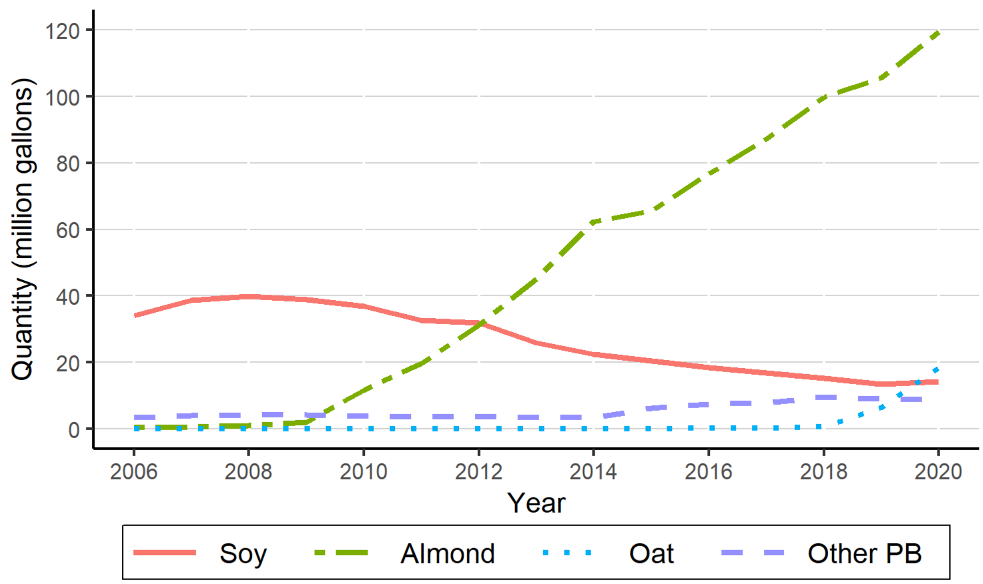


Figure 2.1. Quantity of plant-based milk sold in U.S. retail stores by plant source

Source: NielsenIQ RMS (2006–2020)

Note: The number of stores affiliated with NielsenIQ varies each year. To account for the varying number of stores each year, the annual quantity sold is adjusted by multiplying the average number of retail stores across all years divided by the number of stores each year. Two main retail chain types–grocery stores and mass merchandisers–are used for this calculation over the sample period from 2006 to 2020.

versity of almond milk offerings and their sales volumes. By 2013, almond milk had eclipsed soymilk as the preeminent plant-based milk option, accounting for 60% of the market in terms of sales volume. This trend has persisted, with almond milk consolidating its dominance over the ensuing years, reaching 70% of sales volume by 2020. More recently, the Oat milk category has been on the rise, becoming the second-largest category in the plant-based milk industry.

The expansion of plant-based milk products has been particularly notable within the refrigerated section of retail stores. Anecdotal evidence and discussions with industry experts suggest that the expansion of plant-based products coincided with their introduction to the refrigerated aisle alongside dairy milk. For example, WhiteWave's Silk refrigerated soymilk experienced 600 percent sales growth in 1999 (Shurtleff and Aoyagi 2013). A decade later, Blue Diamond's Almond Breeze achieved similar success following its debut in the

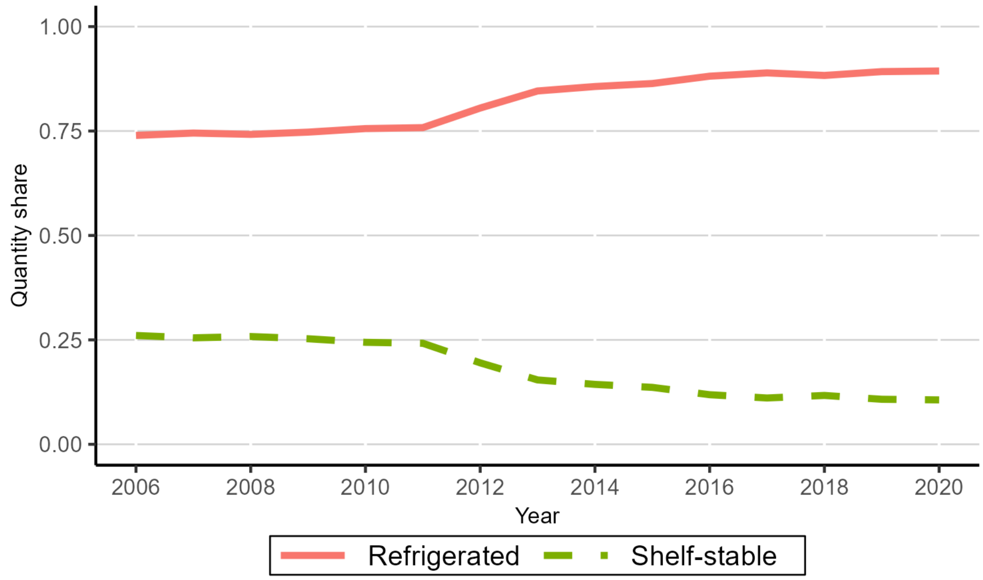


Figure 2.2. Quantity shares of plant-based beverages by storage type
Source: NielsenIQ RMS (2006–2020)

refrigerated case. Figure 2.2 illustrates the dominant share of refrigerated products in the plant-based milk industry. Refrigerated products account for about 75% of the market volume, with shelf-stable products making up the remaining 25% from 2006 to 2011. As refrigerated almond milk emerged as a leading category in the plant-based milk industry, the share of refrigerated products surged to 88% by 2020.

The timing of the introduction of refrigerated almond milk products varied across regions. Figure 2.3 illustrates the spread of refrigerated almond milk products by county from the third quarter of 2008 onward, using NielsenIQ's store scanner data. Refrigerated almond milk was initially introduced in late 2008 in Florida markets. Blue Diamond's decision to start in Florida was aimed at testing its almond milk products, targeting large Hispanic populations with a higher prevalence of lactose intolerance (Chaker 2011; Franklin-Wallis 2019). It then expanded to neighboring regions, such as Georgia and Alabama. Subsequently, in the second quarter of 2009, it was introduced to many counties in the western states and then to the eastern states in the third and fourth quarters of 2009.

Interestingly, counties where refrigerated almond milk was introduced earlier tend to

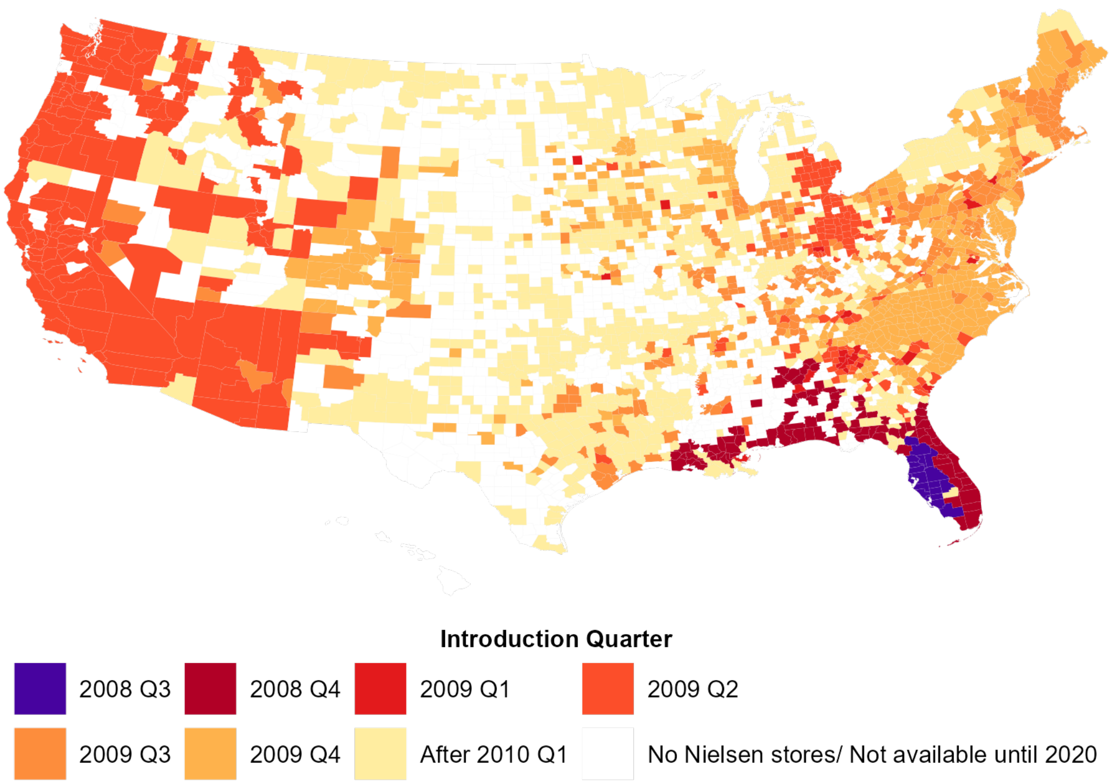


Figure 2.3. Timing of the introduction of refrigerated almond milk by county
Source: NielsenIQ RMS (2006–2020)

have a larger share of the overall plant-based milk in 2020, about a decade later. Figure 2.4 illustrates the quantity share of plant-based milk in the combined cow's milk and plant-based milk in 2020 using NielsenIQ's store scanner data. Many counties in Florida, such as Miami-Dade and Palm Beach, where refrigerated almond milk was introduced in the fourth quarter of 2008, have more than 20% of the quantity share of plant-based milk. Similar patterns are also found in San Francisco County (23%) and Marin County (21%) in California, where refrigerated almond milk was introduced in the second quarter of 2009. Although refrigerated almond milk was introduced relatively later in Boulder County in Colorado and New York County and Kings County in New York, plant-based milk made up more than 20% of the total milk market within each of these counties by 2020.[1]

[1]A simple linear regression of the quantity share of plant-based milk on the month of refrigerated almond milk's introduction indicates that introducing it one month earlier is associated with a statistically significant

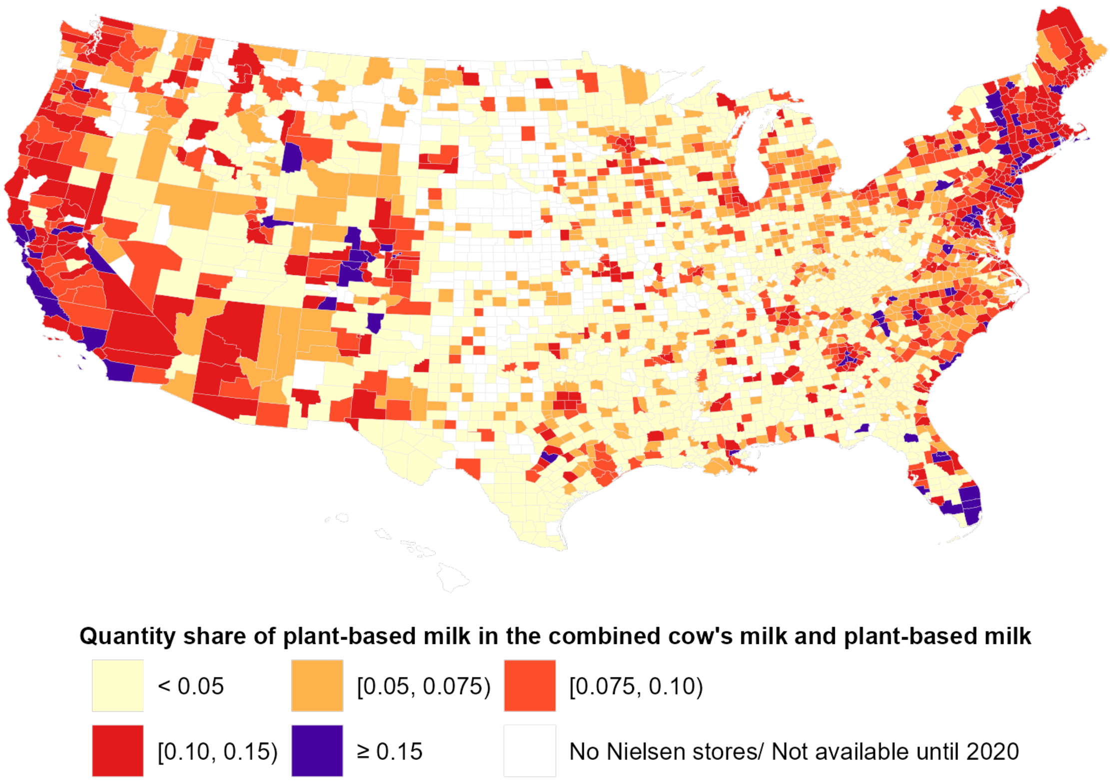


Figure 2.4. Quantity shares of plant-based milk by county in 2020
Source: NielsenIQ RMS (2006–2020)

2.4 Economic Model of Quantity Competition among Incumbent Manufacturers Following the Introduction of a New Product

This section explains the causes of the introduction of refrigerated almond milk and develops the economic model of pricing behavior of cow's milk and soymilk manufacturers in response to the introduction. I first discuss the two main causes of the introduction of refrigerated almond milk in 2008 and 2009: (1) new market and supply conditions that eased entry to the refrigerated dairy section in retail stores and (2) the growing demand for plant-based milk products. Next, I develop my quantity competition model to show how the prices and quantities of traditional products change and how the degree of substitutability affects the

0.0006 percentage point increase in the quantity share. The regression uses the annual total quantity as weights, with standard errors clustered at the state level.

changes in the equilibrium outcomes when the new product is introduced.[2]

2.4.1 Causes of the Introduction of Almond Milk

Previous studies that developed theoretical frameworks identified low entry barriers, increasing market demand, technological advancements, and changes in regulation as key drivers of product innovation and product introduction (Sutton 1991; Vives 2008). They provide theoretical insights into the effects of new product introduction on the prices and quantities of existing products. Broadly, the effects depend on demand and supply conditions, including these underlying drivers of the product introduction (Vives 2008).

I argue that two sources of the introduction of refrigerated almond milk are especially relevant for understanding the economic effects of the new product. First, the reduction of entry costs in the refrigerated dairy sections facilitates the placement of almond milk in this competitive space. The space in the refrigerated section of retail stores is scarce and expensive. It typically requires higher slotting allowances – a per-unit-time charge made by manufacturers to retailers – compared to other food categories (Federal Trade Commission 2003). Retailers tend to charge higher slotting fees to stock a new product to compensate for the opportunity costs of stocking more established products (Sullivan 1997). Furthermore, leading suppliers within specific product categories, known in the retail business as "category captains," sometimes have a role in advising retailers on the selection and placement of new products (Federal Trade Commission 2003). Blue Diamond's director of marketing during the relevant time period, Al Greenlee, noted that Blue Diamond overcame these challenges and successfully placed its products in the refrigerated dairy section by establishing a partnership with the second-largest dairy company in the U.S. (Franklin-Wallis 2019).

The second source, the growing demand for plant-based milk products, also supported the introduction of refrigerated almond milk. Even before its widespread availability in 2009, the plant-based milk industry had been experiencing gradual growth, in contrast to

[2]In future work, I plan to examine the competitive market without assuming market power by milk processors. I anticipate a similar pattern, where existing products that are closer substitutes for the new product will experience greater quantity reductions.

the decline in per capita consumption of fluid cow's milk. Previous studies broadly highlight health concerns (such as, fat contents and calorie intakes), environmental awareness, and animal welfare as main motivations for purchasing plant-based milk (Ruby 2012; McCarthy et al. 2017; Schiano et al. 2020; Wolf, Malone, and McFadden 2020). Data from NielsenIQ's Homescan on annual consumption of cow's milk and plant-based milk indicates that, on average, household consumption of plant-based milk rose from 0.80 gallons in 2004 to 0.96 gallons in 2008, representing an annual growth rate of 4.6%. Meanwhile, the consumption of cow's milk per household decreased by 2.1% annually from 28.6 gallons to 26.2 gallons.

2.4.2 Model of Quantity Competition in Response to a New Product Introduction

This subsection describes the quantity and price decisions of the incumbents (cow's milk processors) after the introduction of an entrant (almond milk manufacturer) induced by two exogenous changes: (1) a decrease in a fixed entry cost to get into refrigerated shelf space in retail stores and (2) an increase in the demand for plant-based milk products. Then, it describes how the degree of substitutability between the existing product and the new product plays a role in the changes in the equilibrium price and quantity.

This subsection develops an illustrative model of Cournot competition based on the framework established by Vives (2008), incorporating two asymmetric model parameters to reflect the market dynamics of cow's milk and almond milk. First, the increase in demand unilaterally applies to almond milk. Recall that fluid milk consumption is decreasing while the plant-based milk market is growing. Second, the degree of substitutability between each of the three cow's milk products and the almond milk product may differ. For example, organic and lactose-free cow's milk products are closer substitutes for almond milk than conventional cow's is for almond milk.

In his model of product innovation and product introduction, Vives (2008) shows that, in a symmetric equilibrium with free entry, the direction of changes in prices and quantities of related products depends on which underlying factors lead to the product introduction. If the decrease in the entry barrier is the source of new products, both the price and quantity of existing products will fall. Conversely, if the growing demand leads to more product introduction, prices would fall while quantity per firm will rise.

Consider a Cournot competition model with three manufacturers, each producing a single differentiated product: one producing conventional cow's milk, another producing organic cow's milk, and a potential entrant producing almond milk.[3] For simplicity, I follow the literature and use a simple linear demand model derived from a quadratic utility function of a representative consumer, which is separable and quasi-linear in a numeraire good (Singh and Vives 1984). Therefore, this simple utility function assumes no income effects. Inverse demand functions faced by manufacturers are linear in quantities: $p_i = a_i - q_i - \sum_{j \neq i} \gamma_{ij} q_j$ for $i = 1, 2, 3$; $\gamma_{ij} = \gamma_{ji}$, where p_i denotes the price for firm i's product. The parameter γ_{ij} represents the degree of substitutability between products i and j. The model is confined to cases where products are gross substitutes ($\partial q_i / \partial p_j > 0$).[4] In the case of three products, this condition can be expressed as: $\gamma_{ij} - \gamma_{ik}\gamma_{jk} > 0$ for all i, j, and k (Amir, Erickson, and Jin 2017). The marginal costs of production are assumed to be constant and denoted by c_i. Profit functions are $\pi_i = (p_i - c_i)q_i - F_i$ where F_i represents the fixed entry cost.

When the almond milk manufacturer does not enter the market due to a high entry cost or a low demand for its product, the conventional cow's milk manufacturer and organic cow's milk manufacturer compete.[5] By setting $q_3 = 0$, the inverse demand functions for conventional cow's milk and organic cow's milk are expressed as: $p_i = a_i - q_i - \gamma_{ij}\ q_j$ for i=1,2 and $i \neq j$. The profit-maximizing quantities for conventional cow's milk and organic

[3]The model incorporating three products is used to illustrate how the degree of substitutability affects the changes in equilibrium outcomes from exogenous shocks, such as a reduction in entry costs or an increase in demand. Conversely, the model with only two products entails simultaneous changes in exogenous shocks and the degree of substitutability, making it challenging to isolate the effect of each factor.

[4]The model excludes two cases: 1) products i and j are perfect substitutes ($a_i = a_j$ and $\gamma_{ij} = 1$) and 2) they are complements ($\gamma_{ij} < 0$).

[5]Threshold levels of entry cost and intercept of demand for almond milk are derived in Appendix 2.B

cow's milk are given by:

$$q_1^D = \frac{2(a_1 - c_1) - \gamma_{12}(a_2 - c_2)}{4 - \gamma_{12}^2} \tag{2.1}$$

$$q_2^D = \frac{2(a_2 - c_2) - \gamma_{12}(a_1 - c_1)}{4 - \gamma_{12}^2} \tag{2.2}$$

When entry occurs because of a decreasing entry cost or an increase in the demand for plant-based milk, in the Cournot model, each firm resets the quantities of its products to maximize its profit, considering other firms' quantities to be constant. Solving the first-order conditions for profit maximization in the three-good case results in the following equilibrium quantities:

$$q_1^T = \frac{1}{A}\Big\{(4 - \gamma_{23}^2)(a_1 - c_1) - (2\gamma_{12} - \gamma_{13}\gamma_{23})(a_2 - c_2) - (2\gamma_{13} - \gamma_{12}\gamma_{23})(a_3 - c_3)\Big\} \tag{2.3}$$

$$q_2^T = \frac{1}{A}\Big\{(4 - \gamma_{13}^2)(a_2 - c_2) - (2\gamma_{12} - \gamma_{13}\gamma_{23})(a_1 - c_1) - (2\gamma_{23} - \gamma_{12}\gamma_{13})(a_3 - c_3)\Big\} \tag{2.4}$$

$$q_3^T = \frac{1}{A}\Big\{(4 - \gamma_{12}^2)(a_3 - c_3) - (2\gamma_{13} - \gamma_{12}\gamma_{23})(a_1 - c_1) - (2\gamma_{23} - \gamma_{12}\gamma_{13})(a_2 - c_2)\Big\} \tag{2.5}$$

where $A = 8 - 2\gamma_{12}^2 - 2\gamma_{13}^2 - 2\gamma_{23}^2 + 2\gamma_{12}\gamma_{13}\gamma_{23}$.

2.4.3 Role of the Degree of Substitutability in Changes in Prices and Quantities

We now illustrate the influence of the two exogenous shocks that induce the entry of almond milk manufacturers: a decrease in a fixed entry cost and an increase in the demand for almond milk. The purpose of the illustration is to show an example of the differential price and quantity effects depending on the substitutability. To fix the idea of the effect of the degree of substitutability, assume that $a_1 = a_2$ and $c_1 = c_2$ (meaning the inverse demand intercept and marginal cost for conventional and organic milk are the same for this exposition).[6] When almond milk is introduced in the market, the equilibrium prices and

[6]These examples do not reflect the actual market conditions for the cow's milk market, as the demand intercepts are set identically.

quantities of existing products would fall in both cases of the source of introduction, as shown in inequality (2.6).[7]

$$p_i^T < p_i^D \quad \text{and} \quad q_i^T < q_i^D \quad \text{for} \quad i = 1, 2. \tag{2.6}$$

Importantly, the decreases in prices and quantities of incumbents' products are larger for existing products that are more substitutable to the new product. Equations (2.7) and (2.8) demonstrate the role of substitutability in the price and quantity responses under the two exogenous changes that induce the entry of the new product.[8]

1) Decreasing entry cost: the almond milk manufacturer enters the market if the fixed entry cost is lower than a threshold level, $\hat{F}_3$. Then, the following inequalities hold:[9]

$$\begin{aligned} q_2^T - q_2^D \leq q_1^T - q_1^D \leq 0 \quad &\text{if and only if} \quad \gamma_{23} \geq \gamma_{13} \\ p_2^T - p_2^D \leq p_1^T - p_1^D \leq 0 \quad &\text{if and only if} \quad \gamma_{23} \geq \gamma_{13} \end{aligned} \tag{2.7}$$

Inequalities (2.7) show that the introduction of almond milk (product 3) causes the quantity of the incumbents' products to fall from q_i^D to q_i^T and the prices to fall from p_i^D to p_i^T. Moreover, these prices and quantities decrease more for cow's milk products that are more substitutable for almond milk. A numerical simulation shown in Panel A of Figure 2.5 illustrates this effect of substitutability on prices and quantities. The almond milk manufacturer enters the market when the entry cost is less than $\hat{F}_3$. Following the entry of almond milk, the quantity of both incumbent products decreases. However, the quantity of organic cow's milk (q_2), represented by the blue dotted line, decreases more than the quantity

[7]Proofs can be found in Appendix 2.B.

[8]Results from the Bertrand competition indicate prices would fall upon entry, but quantities do not necessarily decrease. When the new product is highly similar to incumbent products (indicated by large values of γ_{13} or γ_{23}), the price of incumbent products would fall close to marginal cost (Bertrand Paradox), while quantities increase, approaching the outcomes under perfect competition. For this reason, in the Bertrand competition model, $p_i^T < p_i^D$ for $i = 1, 2$. However, the sign of $q_i^T - q_i^D$ is indeterminate.

[9]Proofs can be found in Appendix 2.B.

Panel A. Decreasing entry cost

Panel B. Expanding demand for almond milk

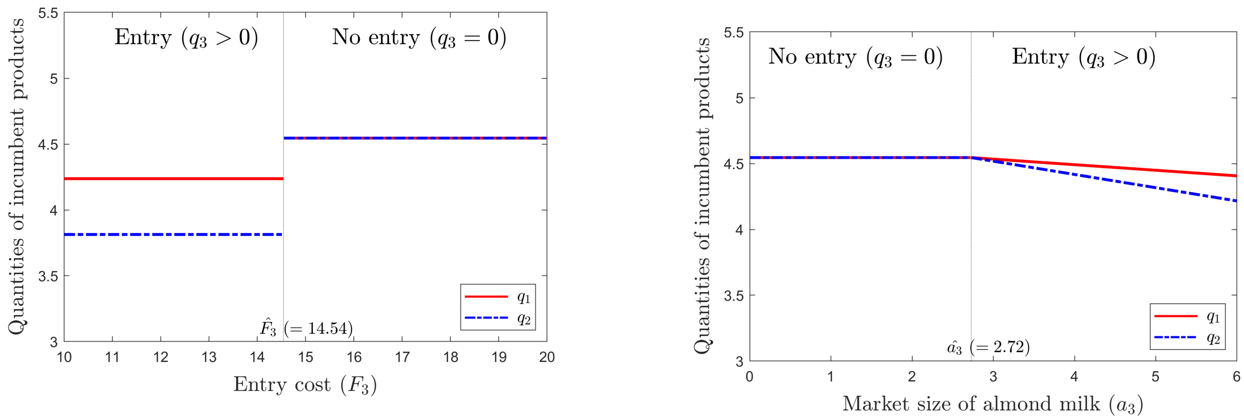


Figure 2.5. Change in quantity of incumbent products by entry cost and demand for almond milk
Note: The figures show the numerical simulation of a decreasing entry cost (Panel A) and an expanding demand for almond milk (Panel B) on the quantities of incumbent products. The entry cost simulation (Panel A) sets the parameter values as follows: a_1=a_2=$a_3 = 10$, $c_1 = c_2 = c_3 = 0$, $F_1 = F_2 = 0$, $\gamma_{12} = 0.2$, $\gamma_{13} = 0.2$, $\gamma_{23} = 0.4$. Under these parameter values, the threshold level of entry cost for almond milk manufacturer is $\hat{F}_3 = 14.54$. The expanding demand simulation (Panel B) sets the parameter values as follows: a_1=a_2=10, $c_1 = c_2 = c_3 = 0$, $F_1 = F_2 = F_3 = 0$, $\gamma_{12} = 0.2$, $\gamma_{13} = 0.2$, $\gamma_{23} = 0.4$. The threshold demand for almond milk manufacturer to enter is $\hat{a}_3 = 2.72$.

of conventional cow's milk (q_1), represented by the red line, when the substitutability of almond milk is greater for organic cow's milk ($\gamma_{23} = 0.4$) than for conventional cow's milk ($\gamma_{13} = 0.2$).[10]

2) Expanding demand for almond milk: the almond milk manufacturer enters the market as the demand for almond milk (represented by parameter a_3) increases above the threshold value. In this case, the following inequalities hold:[11]

[10] In this model setup where $c_i = 0$, the equilibrium prices are equivalent to the equilibrium quantities ($p_i^D = q_i^D$ or $p_i^T = p_i^T$). Since $c_i = 0$ for all i, the y-axis in Figure 2.4 can also be interpreted as the absolute markup ($p_i - c_i$). Upon entry, both Panel A and B of Figure 2.4 show a decrease in the markups for existing products.

[11] Proofs can be found in Appendix 2.B.

$$\begin{aligned}
\frac{\partial q_2^T}{\partial a_3} \leq \frac{\partial q_1^T}{\partial a_3} \leq 0 \quad \text{if and only if} \quad \gamma_{23} \geq \gamma_{13} \\
\frac{\partial p_2^T}{\partial a_3} \leq \frac{\partial p_1^T}{\partial a_3} \leq 0 \quad \text{if and only if} \quad \gamma_{23} \geq \gamma_{13}
\end{aligned} \tag{2.8}$$

The inequalities in (2.8) demonstrate that while both prices and quantities fall after the entry induced by the increasing market size for almond milk, the price and quantity of incumbent products that are more substitutable for almond milk experience greater reductions. Panel B of Figure 2.5 presents a numerical example illustrating the effects of an expanding market size for almond milk. The almond milk manufacturer enters the market when the potential market size exceeds $\hat{a}_3$. After the entry of almond milk, the quantity of both incumbent products declines linearly with the market size of almond milk. However, the quantity of the incumbent product that is more substitutable for almond milk (q_2) declines more steeply than that of the less substitutable (q_1), given the substitutability parameters $\gamma_{13} = 0.2$ and $\gamma_{23} = 0.4$.

2.5 Econometric Strategy

The empirical setting of this research involves many stores around the United States where refrigerated almond milk was introduced in different months over a roughly two-year period. The staggered nature of the introduction presents an empirical opportunity to estimate the effects of introduction on cow's milk and soymilk, but also a statistical challenge. Recent econometric advances suggest that the commonly used two-way fixed-effect (TWFE) estimators are likely to be biased in such staggered treatment settings (De Chaisemartin and d'Haultfoeuille 2020; Callaway and Sant'Anna 2021; Sun and Abraham 2021). To address this potential bias, this study adopts a recently developed econometric approach suitable for settings characterized by staggered treatment and heterogenous treatment effects, following Callaway and Sant'Anna (2021).

2.5.1 Two-Way Fixed-Effect Estimators

The TWFE regression, which controls for individual and time-fixed effects, became a standard method to evaluate a causal treatment effect. In the context of the introduction of refrigerated almond milk as a treatment, a static TWFE regression can be described by Equation (2.9). The outcome variable (Y_{it}) is either the log of the monthly average price or the log of the monthly quantity sold of categories of cow's milk and soymilk in store i and month t. The term, D_{it} is a binary variable equal to 1 if the refrigerated almond milk is available in store i and month t and equal to 0 otherwise.

$$Y_{it} = \alpha_i + \alpha_{\gamma t} + \beta D_{it} + \epsilon_{it} \tag{2.9}$$

where α_i is a specific store fixed effect, $\alpha_{\gamma t}$ is a retailer-specific time-fixed effect (recall, the term "retailer" refers to a group of stores operated together as a "chain"), and ϵ_{it} is an error term. The parameter of interest in (2.9) is β, which can be interpreted as the overall effect of the introduction of refrigerated almond milk on prices or quantities of cow's milk and soymilk.

I estimate both static and dynamic TWFE specifications separately for two outcome variables (log of price and log of quantity sold) for each of the four product categories (conventional cow's milk, organic cow's milk, lactose-free cow's milk, and soymilk).[12]

The potential bias of TWFE estimates arises from using previously treated units as comparison groups for later-treated units. In particular, the bias in the static TWFE estimates becomes severe when treatment effects change over time because the change in outcome variable of the comparison group is contaminated by the change in treatment effects over time (Goodman-Bacon 2021). Furthermore, Sun and Sun and Abraham (2021) show that the bias still remains in more general, dynamic settings.

[12]A dynamic version of the TWFE regression can be specified by including dummy variables for months relative to the initial treatment month. The coefficients on these dummy variables are interpreted as the treatment effect at different length of exposure to the treatment, capturing treatment heterogeneity over time.

In the context of almond milk introduction, treatment effects on the prices and quantities of cow's milk and soymilk are likely to evolve over time. A new product like almond milk typically entails a gradual process of consumer awareness and adoption. While almond milk witnessed a rapid surge in growth following its introduction in 2008, consumer demand for almond milk continued to grow at least until 2015 (Figure 2.1). Moreover, it is likely that the effects of the introduction grew over time as more almond milk brands became available.[13]

2.5.2 A Heterogeneity-Robust Difference-in-Differences Estimator

To address the challenges associated with staggered introduction and heterogenous treatment effects over time, this study adopts the approach proposed by Callaway and Sant'Anna (2021). This approach involves estimating group-time average treatment effects, allowing treatment heterogeneity across different timing groups and over time. Following Callaway and Sant'Anna (2021), the average treatment effect for timing group g at month t can be nonparametrically identified as Equation (2.10). The timing group g is defined as a group of stores into which almond milk was first introduced in month g.

$$ATT(g,t) = E[Y_{it}(g) - Y_{it}(0)|G_{ig} = 1] \tag{2.10}$$

where $Y_{it}(g)$ is the outcome of store i belonging to group g at month t, and $Y_{it}(0)$ denotes the untreated potential outcome of store i at month t. The term, G_{ig} is a binary indicator that equals one if store i first adopted refrigerated almond milk in month g.

The estimation of $ATT(g,t)$ in Equation (2.10) involves a two-step procedure. The first step parametrically estimates the conditional expectation of the evolution of the outcome among stores that are not yet treated until month t, written as $E[Y_{i,t} - Y_{i,g-1} \mid X_i, D_{it} = 0, G_{ig} = 0]$, where D_{it} equals one if store i had already introduced almond milk by month t or

[13] Future work plan to deal with the increased variety (intensity) of the introduction of refrigerated almond milk.

zero otherwise. This conditional expectation function can be estimated by linear regression of the evolution of the outcome $(Y_{i,t} - Y_{i,g-1})$ on store-specific demographics, including retailer-fixed effects (X_i), among the not-yet-treated stores. The fitted conditional expectation function can be written as $E[Y_{i,t} - Y_{i,g-1} \mid X_i, D_{it} = 0, G_{ig} = 0] = X_i' \hat{\gamma}$, where γ is a vector of coefficients on X_i.

In the second step, the fitted values of the conditional expectation function are plugged into the sample analog of the group-time ATTs using the empirical distribution of X_i among timing group g, as shown in Equation (2.11).

$$A\hat{T}T(g,t) = \frac{1}{N_g} \sum_{i:G_{i,g}=1} \Big\{ (Y_{i,t} - Y_{i,g-1}) - \hat{E}[Y_{i,t} - Y_{i,g-1} | X_i, D_{i,t} = 0, G_{ig} = 0] \Big\} \tag{2.11}$$

where N_g is the number of stores in timing group g.

The key assumption for identifying ATTs in Equation (2.10) is the conditional parallel trend assumption. This assumption, as applied to cow's milk prices, for example, states that the retail cow's milk prices in treatment and control stores would have followed the same paths in the absence of the availability of refrigerated almond milk, conditional on store-specific demographics and retailer-time-fixed effects. Store-specific demographic variables include median household income, percent of the White population, and percent of the Hispanic population. These characteristics are more thoroughly discussed in Section 2.6.

The rationale for the selection mechanism (whether and when stores adopt almond milk) assumes that the treatment timing of refrigerated almond milk over the 29-month period depends predominantly on time-stable store-specific characteristics, such as demographic factors and the parent retailer. As noted, early stores are a result of Blue Diamond's strategic decision. On the other hand, retailers often face uncertainty regarding the demand for and potential profits from new products. The prevalence of slotting fees in the U.S. retail grocery industry reflects this uncertainty faced by retailers,[14] suggesting that decisions to adopt new

[14]Many papers on theories on slotting fees point out that retailers charge slotting fees, in part, to balance the risk of the failure of new products between retailers and manufacturers (Chu 1992; Desai 2000; Lariviere and Padmanabhan 1997; Sudhir and Rao 2006). The risk of failure may arise from retailers' imperfect

products are driven by time-stable factors and less influenced by time-varying unobserved factors.[15]

Based on the selection mechanism just described, the conditional parallel trend assumption is plausible if time-varying unobservable factors affecting the untreated outcome have constant means over time (Ghanem, Sant'Anna, and Wüthrich 2024). For the case of almond milk introduction, retail prices of cow's milk can be decomposed into the wholesale price of cow's milk and retail margin. Note that no manufacturers produced both cow's milk and almond milk during the 2008–2010 introduction period. This absence of overlap in product ownership implies that, at least for the wholesale market, endogenous introduction, such as multiproduct manufacturers' simultaneous decisions regarding the pricing of their existing products and introducing new products, is less of a concern (Hausman 1996; Hausman and Leonard 2002). However, for soymilk, the exogenous claim only holds before the first quarter of 2010 when Silk, which produced soymilk products, launched its almond milk products.

Potential unobserved time-varying factors, such as retailer-specific promotional activities, may induce different paths of prices of cow's milk in the absence of almond milk. However, this time-varying factor is controlled by including retailer-time-fixed effects.

The conditional parallel trend assumption is more plausible than the unconditional counterpart when covariate-specific trends in outcome variables are expected (Callaway and Sant'Anna 2021). For example, stores with a higher share of Hispanic residents typically introduced almond milk earlier, and their market share tended to grow faster compared to stores with a lower share of Hispanic residents. Furthermore, large retail chains often implement uniform pricing strategies across their stores in different regions (DellaVigna and Gentzkow 2019). Incorporating the retailer-fixed effects controls for time-varying pricing dynamics and variation in product composition sold across retail chains.

information on new products (Sullivan 1997).

[15] Manufacturers' decisions about where to launch first and retailers' risk-sharing strategy in the new product acceptance are more likely short-term considerations, as the introduction of almond milk mostly occurred within two years. In the long run, product innovations tend to reflect gradual changes in lifestyles, such as increased labor force participation by women and an increase in eating away from home.

Alternative estimators have been proposed to address the potential bias of TWFE. (for review papers, see De Chaisemartin and d'Haultfoeuille (2023), and Baker, Larcker, and Wang (2022).) Sun and Abraham (2021) also propose estimating the group-time treatment effect. Another strand of studies, such as Borusyak, Jaravel, and Spiess (2024) and Gardner (2022), proposes an imputation estimation strategy where the potential untreated outcome for the treated group is imputed by predicted values from a TWFE regression of outcome on group- and time-fixed effects in the sample of never-treated observations. However, these estimators allow either for the never-treated or last-treated units as for the control group, whereas Callaway and Sant'Anna (2021) estimator allows for the not-yet-treated units as the control group. In the context of almond milk introduction, allowing the not-yet-treated stores is more appropriate because almond milk is eventually introduced to more than 99% of stores by the end of the sample year, 2010.

2.5.3 Summary Measures of Group-Time Treatment Effects

The group-time ATTs in (2.11) are estimated for each timing group g and month t. These group-time ATTs can be aggregated to form different summary measures of causal parameters. One straightforward aggregation to an overall treatment effect is to compute a weighted average of all group-time ATTs, weighted by the size of the timing group. The estimator for the simple overall treatment effect parameter can be written as shown in Equation (2.12).

$$\hat{\theta}_W = \frac{1}{\kappa} \sum_g \sum_t 1(t \geq g) A\hat{T}T(g,t) \hat{P}(G = g | G \leq T) \tag{2.12}$$

where $\kappa = \sum_g \sum_t 1(t \geq g) \hat{P}(G = g | G \leq T)$. The term, $\hat{P}(G = g | G \leq T)$, is the share of timing group g.

Alternatively, two partial aggregations can be used to highlight the treatment heterogeneity over time and across timing groups. The first partial aggregation parameter is a dynamic treatment effect for a specific number of months relative to the introduction month. Let e denote event-time, which counts the number of months since the introduction. The dynamic treatment effect is defined as the average treatment effect e months after the introduction, averaged across all timing groups using group size as weights, as shown in Equation (2.13). The event-time treatment effects, $\theta_{es}(e)$, can further be aggregated into the overall treatment effect, θ_{es}, by taking an average of them, as shown in Equation (2.14).

$$\hat{\theta}_{es}(e) = \sum_{g} 1g + e \leq T\hat{P}(G = g|G + e \leq T)A\hat{T}T(g, g + e) \tag{2.13}$$

$$\hat{\theta}_{es} = \frac{1}{T-1}\sum_{e=0}^{T-2}\hat{\theta}_{es}(e) \tag{2.14}$$

The second partial aggregation parameter is a timing-group-specific treatment effect, defined as the average treatment effect for each timing group across all their post-introduction periods, as shown in Equation (2.15). This timing-group-specific parameter helps to explain the heterogeneous treatment effect across stores with different timing of the introduction of refrigerated almond milk. The timing-group-specific treatment effect, $\theta_{gr}(\tilde{g})$, can further be aggregated into the overall treatment effect, θ_{gr}, by taking the weighted average of group-specific treatment effects, using group size as weights, as shown in Equation (2.16).[16]

$$\hat{\theta}_{gr}(\tilde{g})\frac{1}{T-\tilde{g}-1}\sum_{t\geq\tilde{g}} A\hat{T}T(\tilde{g}, t) \tag{2.15}$$

$$\hat{\theta}_{gr} = \sum_{g}\hat{\theta}_{gr}(g)\hat{P}(G = g|G \leq T) \tag{2.16}$$

[16]Three summary measures, θ_W, θ_{es}, and θ_{gr}, target the same overall treatment effect parameter, while they use different weights. Callaway and Sant'Anna (2021) recommend to use θ_{gr} because its interpretation is consistent with the canonical Difference-in-Differences setup. The simple aggregate measure, θ_W, puts more weight on the earlier-treated group. The aggregate measure based on dynamic treatment effects, θ_{es}, has limited appeal because its interpretation may be complicated by compositional changes of comparison groups at different e.

2.6 Data for Estimating the Effect of Refrigerated Almond Milk Introduction

The effects of the introduction of almond milk on cow's milk prices and quantities are estimated using data on retail stores from NielsenIQ's Retail Scanner data (RMS) for the period 2007–2010. As noted above, these dates align with the widespread introduction of almond milk across the United States. NielsenIQ's Homescan Panel (HMS), for the same timeframe, is used to help construct store-specific demographics by matching households who visited the affiliated stores.

2.6.1 Sample Stores

NielsenIQ's RMS provides weekly sales revenue and quantity sold for every product UPC sold across approximately 37,000 stores from 2007 to 2010. The dataset includes a range of store formats, including grocery stores, mass merchandisers, drug stores, and convenience stores.

Several data-cleaning processes, adjustments, and sample choices for RMS stores are applied to define the main analysis sample, as outlined in Table 2.1. First, attention is confined to grocery stores because about 87% of milk sales are concentrated in grocery stores, which comprise about 30% of all types of stores. Also, grocery stores typically have a wider selection of alternative cow's milk and plant-based milk products than the other store formats. Second, stores with no HMS consumer visits during the sample period are excluded. HMS consumer demographic variables are needed to construct the store-level demographic variables. Focusing the dataset on stores visited by HMS households excludes an additional 700 stores, accounting for only 5% of the total sales quantity of the products of interest (third row of Table 2.1). Third, stores that consistently recorded positive monthly sales every month from January 2008 to December 2010, the period when almond milk was introduced in most locations, were retained. Fourth, stores with unreliable sales records or prices were excluded. The main analysis sample excludes any store where the maximum

Table 2.1. Sample store formation

Store sample formation	Number of stores	Number of retailers	Total annual quantity (million gallons)
Initial sample	36,275	86	1,260.70
Grocery stores	10,694	66	1,096.14
Stores with HMS household visits	9,962	64	1,045.98
Consistently observed stores	9,363	63	954.96

Note: Numbers presented are based on NielsenIQ's retail scanner (RMS) dataset. "Grocery stores" are defined as stores with channel code of F according to NielsenIQ's store classification. "HMS" in the third row stands for NielsenIQ Homescan Panel. "Consistently observed stores" are identified based on three data-cleaning criteria: 1) stores that consistently recorded positive total monthly quantities of cow's milk and soymilk throughout January 2008 to December 2010, 2) stores with maximum monthly milk sales not exceeding ten times the minimum monthly milk sales, and 3) average monthly quantity of cow's milk and soymilk greater than or equal to 10 gallons. Total annual quantity is calculated as the sum of the total quantity sold in stores in each criterion divided by the number of years in the sample period.

monthly milk sales are more than ten times the minimum monthly milk sales or where the average milk sales are less than 10 gallons per month. The final sample consists of 9,363 stores with 954 million gallons sold of cow's milk and soymilk, accounting for 76% of the quantity sold in the initial sample.

Table 2.2 presents the number of stores that adopted refrigerated almond milk each month from August 2008 to December 2010. By late 2008, 6% of the final sample stores had adopted refrigerated almond milk. These stores were mostly located in Florida, Georgia, and Alabama. Three-quarters of stores had adopted refrigerated almond milk by 2009, and almost all stores had adopted refrigerated almond milk by 2010. Only 1 percent of stores had not yet introduced refrigerated almond milk by the end of 2010.

Store-specific demographic variables are constructed using HMS household demographics, following an approach similar to DellaVigna and Gentzkow (2019). HMS data provide 5-digit zip codes for the residential location of HMS panelists. Demographic variables asso-

Table 2.2. Store distribution for the month of the introduction of almond milk

Year	Month	Number of stores	Percentage share of stores	Cumulative share of stores
2008	8	80	0.85%	0.85%
2008	9	11	0.12%	0.97%
2008	11	82	0.88%	1.85%
2008	12	369	3.94%	5.79%
2009	1	23	0.25%	6.03%
2009	3	29	0.31%	6.34%
2009	4	376	4.02%	10.36%
2009	5	530	5.66%	16.02%
2009	6	1,098	11.73%	27.75%
2009	7	398	4.25%	32.00%
2009	8	261	2.79%	34.79%
2009	9	750	8.01%	42.80%
2009	10	2,117	22.61%	65.41%
2009	11	721	7.70%	73.11%
2009	12	126	1.35%	74.45%
2010	1	714	7.63%	82.08%
2010	2	565	6.03%	88.11%
2010	3	469	5.01%	93.12%
2010	4	194	2.07%	95.19%
2010	5	93	0.99%	96.19%
2010	6	51	0.54%	96.73%
2010	7	59	0.63%	97.36%
2010	8	39	0.42%	97.78%
2010	9	25	0.27%	98.05%
2010	10	21	0.22%	98.27%
2010	11	47	0.50%	98.77%
2010	12	19	0.20%	98.97%
After 2011		96	1.03%	100.00%

Note: Numbers presented are based on the final dataset. The total number of stores is 9,363.

Table 2.3. Store-specific demographic variables

Demographic variables	Mean	25th	Median	75th
Median household income	\$59,774	\$44,434	\$55,124	\$71,028
Percent of White population	75.6%	66.3%	79.8%	89.4%
Percent of Hispanic population	13.3%	3.5%	7.7%	17.2%

Note: Mean and percentile statistics are calculated based on 9,363 stores in the final sample over the sample period from 2007 to 2010. On average, each store is visited by 21 HMS households, with each household making 33 visits, resulting in a total of 668 HMS household visits over the sample period. Note that the demographics of HMS are not directly used; Instead, demographics of the 5-digit zip code areas where the HMS households reside are used (see the explanations in the main text).

ciated with each HMS household are based on the zip code of their residence, as measured in the 2007–2011 American Community Survey. The store-specific demographic variables for a given store are defined as the average demographic variables of HMS households who visit that store, weighted by the number of visits to that store and sampling weights in HMS data. Table 2.3 summarizes three store-specific demographic variables. For the median store, the annual median household income is \$55,081, and the percentage of the White population is 79.8%. The percentage of the Hispanic population varies substantially across stores, ranging from 3.5% at the 25^{th} percentile to 17.2% at the 75^{th} percentile.

2.6.2 Definition of Product Categories and Prices

In the final dataset, there are 1,669 UPCs for cow's milk and 82 UPCs for soymilk in the stores' refrigerated sections. These UPCs are classified into four product categories: conventional cow's milk, organic cow's milk, lactose-free cow's milk, and soymilk. The econometric estimation is confined to package sizes larger than 28 oz. Those products account for more than 99 percent of the quantity sold in retail stores. Only UPCs with positive monthly sales at a store in every month throughout the sample periods are retained in order to prevent compositional differences of products due to the entry or exit of cow's milk or soymilk products from a particular store. Finally, UPCs are eliminated as outliers if their minimum price over the sample period is below 1$ per gallon or if their maximum price exceeds $25 per gallon.

The monthly quantity for each of the four categories at a store is defined as the sum of the weekly quantities sold across all UPCs belonging to each category within that store for the given month.[17] [18] Sales data for the first week are prorated based on the number of days in that week. To account for variations in the number of days across months, the monthly quantity is divided by the number of days in that month and multiplied by 30. Monthly revenues are aggregated in the same manner as the monthly quantities. The monthly average unit value for each category, which is the ratio of the monthly revenue to the monthly quantity sold, is used for the monthly average prices for each category. All prices are deflated to 2020 dollars using the Consumer Price Index (CPI) for all items.

Table 2.4 describes the mean and standard deviation of the monthly average price and total quantity for each of the four categories. These prices are reported in dollars per gallon, even though some of the packages are in quantities less than one gallon. The share of quantity sold in packages of less than one gallon ranges from 22% for conventional cow's milk to 94% for soymilk. Conventional cow's milk is the least expensive category and accounts for 91% of the monthly quantity sold in an average store. Organic cow's milk and lactose-free cow's milk are more than twice as expensive as conventional cow's milk, with considerably less quantity sold in the average store. Soymilk accounts for 2.6% of the quantity share of these four categories, and the price of soymilk is significantly higher than conventional cow's milk but cheaper than organic or lactose-free cow's milk.

[17] NielsenIQ's RMS weekly quantities and revenues are reported for the period from Sunday to Saturday, with the week-ending code assigned to Saturday. The first week of each month includes weekly sales from both the previous month and the current month up to the week-ending code.

[18] Monthly average unit value and quantity for each category reflect all package sizes. Price and quantity effects estimated in this chapter represent the overall effects across different package sizes. Future work can explore the heterogeneous effects of package size.

Table 2.4. Monthly sample store quantity and average price (2008–2010)

	Monthly average quantity (gallons)	**Monthly average price ($/gallon)**
Conventional cow's milk	7,610 (91.1%)	4.04
Organic cow's milk	340 (4.1%)	8.42
Lactose-free cow's milk	184 (2.2%)	9.22
Soymilk	218 (2.6%)	7.67

Note: Numbers are based on the final sample of 9,363 stores. The monthly average quantity is calculated as the simple average of the monthly quantities sold across all stores. Numbers in parentheses are the quantity share of each product category among the four categories. The monthly average price is calculated as the weighted average of the monthly prices across all stores, weighted by the monthly quantity sold. Prices are deflated to 2020 dollars using CPI for all items.

2.7 Price and Quantity Effects of the Introduction of Refrigerated Almond Milk from 2008 through 2010

This section presents econometric estimates of the effects of the introduction of almond milk on the price and quantity of cow's milk and soymilk. Section 2.7.1 reports the average treatment effects that are separately estimated for four product categories: conventional cow's milk, organic cow's milk, lactose-free cow's milk, and soymilk. For each product category, the price and quantity effects are estimated using Callaway and Sant'Anna (2021)'s heterogeneity-robust Difference-in-Differences estimator (hereafter, CS estimator), along with the potentially biased TWFE estimator. Section 2.7.2 further explores the implications of different sample periods and different parallel trend assumptions on CS and TWFE estimates.

2.7.1 Results from the Main Specifications

This subsection reports and discusses the main results obtained from the econometric estimation of my baseline model specification. The CS estimates are obtained from a conditional parallel trend assumption that allows retailer-specific time trends. The TWFE estimates are obtained from the specification with store-fixed effects and retailer-by-month fixed effects.

Table 2.5 presents the estimates of the average treatment effects on the prices and quantities derived from the CS and TWFE estimators. The first three rows show the average treatment effects from CS estimators, using different summary measures of the group-time treatment effects corresponding to $\hat{\theta}_W$, $\hat{\theta}_{es}$, and $\hat{\theta}_{gr}$ in Equations (2.12), (2.14), and (2.16) in Section 2.5, Econometric Strategy. The final row presents the TWFE estimates corresponding to Equation (2.9). In the discussion of the effects estimated for each potential substitute product (the columns in Table 2.5), I will focus on the results from the CS estimators provided in the first three rows of Table 2.5 rather than the TWFE estimates that are likely to be biased.

Figures 2.6 and 2.7 present the event-study plots obtained from the CS partial summary estimates for dynamic treatment effects. These figures show the average treatment effect estimates by the number of months before and after the introduction of refrigerated almond milk, aggregated over all timing groups using the calculation procedure of Equation (2.13) in Section 2.5. Price effects are presented in Figure 2.6, and quantity effects are presented in Figure 2.7. The event-study plots obtained from the TWFE estimator can be found in Appendix Figures 2.A.1 and 2.A.2.

2.7.1.1 Price and Quantity Effects for Each Product Group

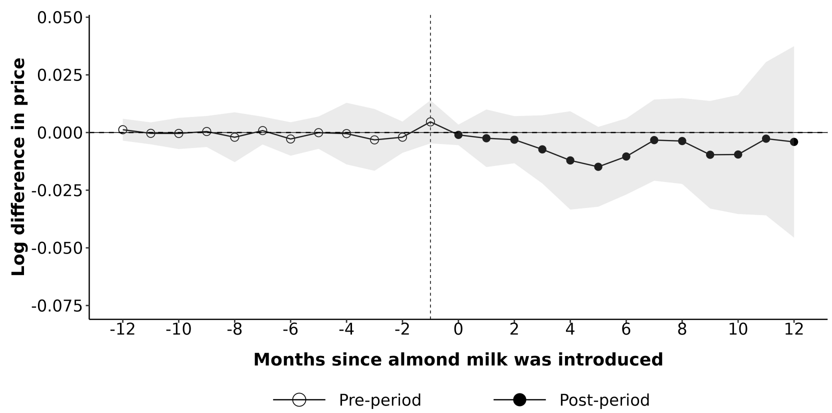
Conventional cow's milk: The first two columns of Table 2.5 present the average price and quantity impacts for conventional cow's milk. Overall, for all the alternative measures, the introduction of refrigerated almond milk has minimal impact on the price and quantity of conventional cow's milk. The average treatment effects obtained from the CS estimators indicate a slight decrease in the price of conventional cow's milk, ranging from 0.65% to 0.86%. However, these estimates are much smaller than their standard errors and not statistically significant from zero by any conventional criteria. The quantity effects estimated from the CS approach are almost negligible in percentage terms, ranging from -0.13% to 0.16%.

Table 2.5. Effect of almond milk introduction on cow's milk and soymilk

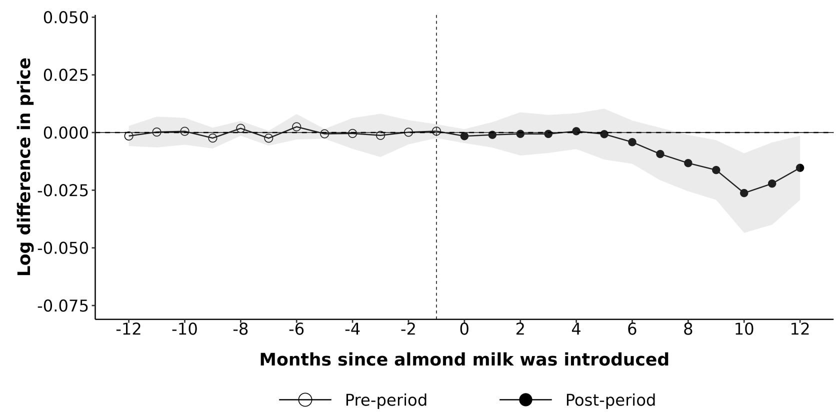
	Conventional cow's milk		Organic cow's milk		Lactose-free cow's milk		Soymilk	
Dependent variables	**Log price**	**Log quantity**	**Log price**	**Log quantity**	**Log price**	**Log quantity**	**Log price**	**Log quantity**
CS – Simple weighted avg ($\hat{\theta}_W$)	-0.0086 (0.0069)	0.0016 (0.0106)	-0.0075** (0.0031)	-0.0413** (0.0205)	0.0037 (0.0033)	-0.0331** (0.0131)	-0.0098* (0.0054)	-0.0659** (0.0139)
CS – Avg event-time TEs ($\hat{\theta}_{es}$)	-0.0065 (0.0053)	-0.0013 (0.0100)	-0.0085** (0.0025)	-0.0246** (0.0117)	0.0031 (0.0029)	-0.0281** (0.0115)	-0.0060 (0.0052)	-0.0472** (0.0125)
CS – Avg group-specific TEs ($\hat{\theta}_{gr}$)	-0.0077 (0.0049)	0.0012 (0.0102)	-0.0070** (0.0030)	-0.0345** (0.0193)	0.0019 (0.0034)	-0.0296** (0.0135)	-0.0080* (0.0041)	-0.0600** (0.0162)
TWFE ($\hat{\beta}$)	-0.0147* (0.0087)	0.0225** (0.0105)	0.0006 (0.0037)	0.0028 (0.0067)	0.0024 (0.0031)	-0.0083 (0.0073)	0.0027 (0.0054)	-0.0030 (0.0062)

Note: This table reports the average treatment effects of almond milk introduction on the price and quantity of four product categories, using Callaway and Sant'Anna (2021) and two-way fixed-effect estimators. Estimates based on Callaway and Sant'Anna (2021)'s estimator account for (store-level) covariate-specific trends and retailer-fixed effects. two-way fixed-effects estimates are calculated after controlling for store-fixed effects and retailer-year-month-fixed effects. All estimates are weighted by the monthly store quantity. As noted in the Data section, the monthly store quantities are adjusted to represent 30 days of sales. Standard errors are clustered at the state level and are shown in parentheses.

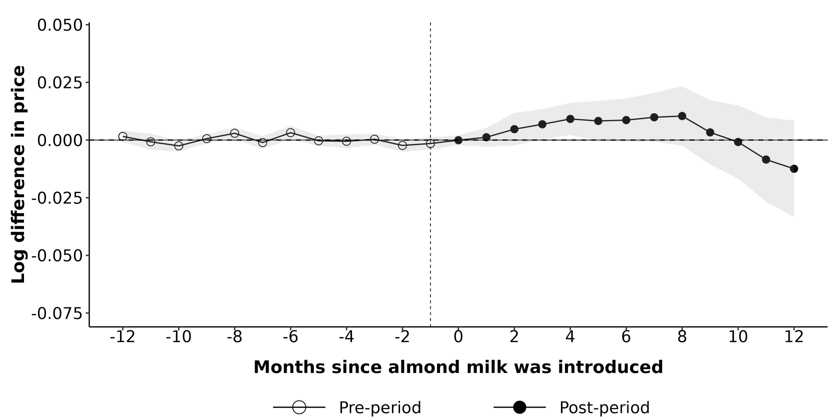
Panel A. Price of conventional cow's milk



Panel B. Price of organic cow's milk



Panel C. Price of lactose-free cow's milk



Panel D. Price of soymilk

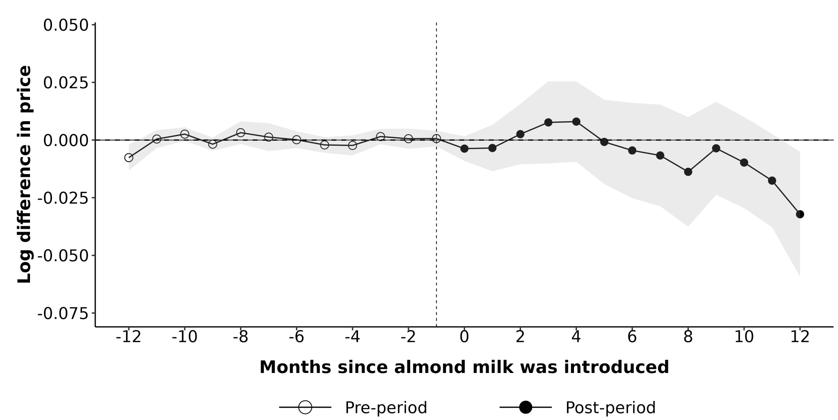
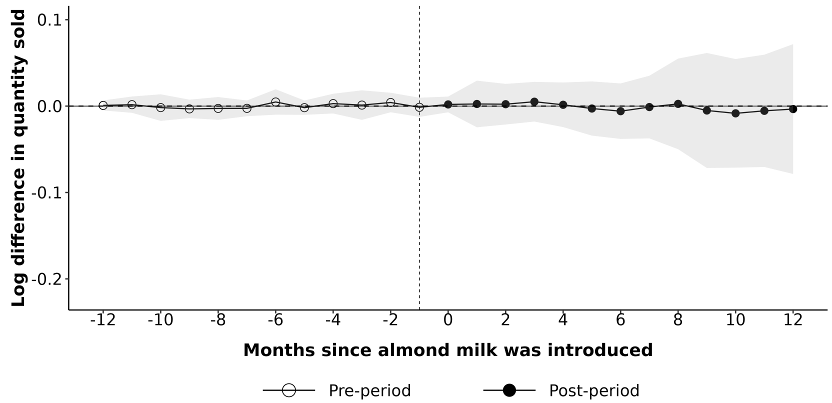


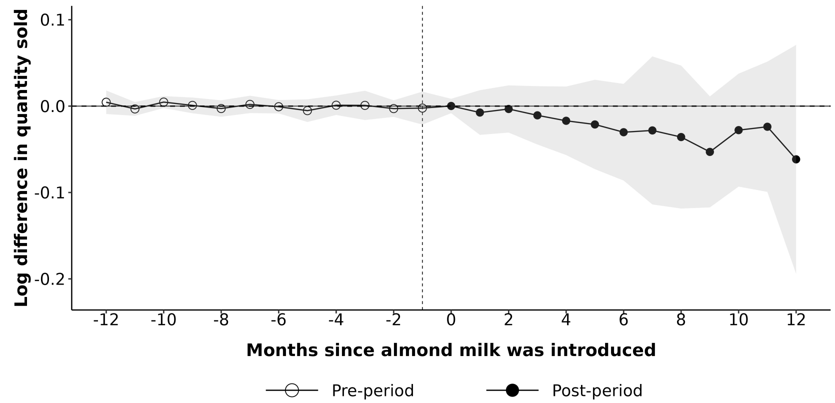
Figure 2.6. Event-study graphs for the price of categories of cow's milk and soymilk

Note: The figures show average treatment effects on the prices of categories of cow's milk and soymilk relative to one month before the initial introduction month, controlling for demographic-specific trends and retailer-specific fixed effects. The pre-trend coefficients are estimated relative to one month before the month under consideration. The vertical axis represents the log difference in the price of categories of cow's milk and soymilk. Circled dots represent pre-trend coefficients, while filled dots represent post-period treatment effects. Estimated treatment effects are based on the procedures developed by Callaway and Sant'Anna (2021)). Shaded areas indicate the 95% confidence bands based on standard errors clustered at the state level.

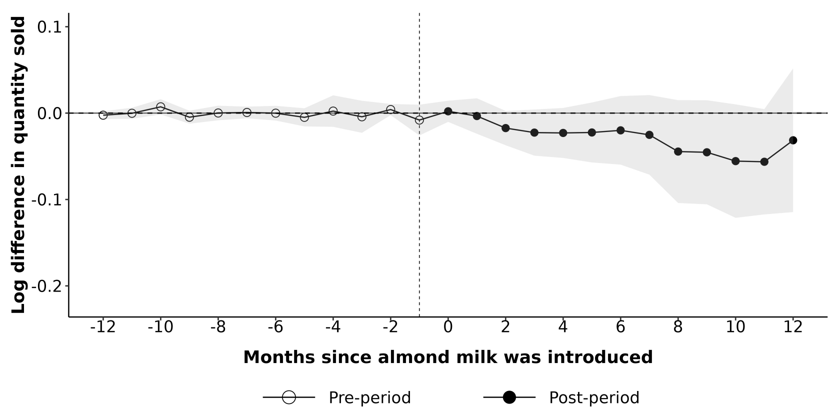
Panel A. Quantity of conventional cow's milk



Panel B. Quantity of organic cow's milk



Panel C. Quantity of lactose-free cow's milk



Panel D. Quantity of soymilk

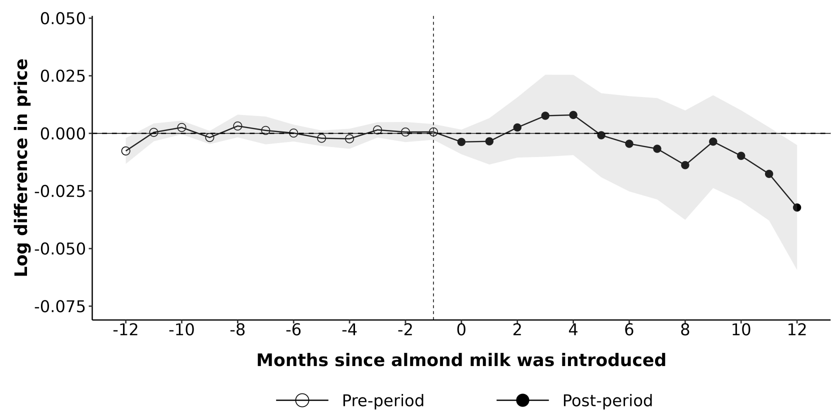


Figure 2.7. Event-study graphs for the quantity of categories of cow's milk and soymilk

Note: The figures show average treatment effects on the quantities of categories of cow's milk and soymilk relative to one month before the initial introduction month, controlling for demographic-specific trends and retailer-specific trends. The pre-trend coefficients are estimated relative to one month before the month under consideration. The vertical axis represents the log difference in the quantity of cow's milk and soymilk. Circled dots represent pre-trend coefficients, while filled dots represent post-period treatment effects. Estimated treatment effects are based on the procedures developed by Callaway and Sant'Anna (2021). Shaded areas indicate the 95% confidence bands based on standard errors clustered at the state level.

The event-study plots are fully consistent with the findings for the estimates of the averages that the introduction of almond milk had a negligible percentage impact on the price and quantity of conventional cow's milk sold in retail stores. Panel A of Figure 2.6 for prices shows a tiny and insignificant effect. Figure 2.7 shows that the treatment effects on the quantity of conventional cow's milk remained stable and close to zero over the 12 months following the introduction. We will see that the minimal impact on conventional cow's milk contrasts with the treatment effects on quantities of organic cow's milk, lactose-free cow's milk, and, especially, soy milk, which shows a decreasing quantity trend over time.

Organic cow's milk: The third and fourth columns of Table 2.5 present the average treatment effects on the price and quantity of organic cow's milk. Across all three weighting schemes of aggregate measures, the CS estimates indicate that the introduction of refrigerated almond milk caused an economically and statistically significant reduction in quantity by -2.46% to -4.13%. While the price effect for organic cow's milk is statistically significant, the percentage changes are small (between -0.70% and -0.85%).

As described in the economic model of price and quantity responses for incumbent firms, decreases in the price and quantity of incumbent firms' products are expected to be larger if the existing products are more substitutable with the new product. Table 2.5 indicates that the reduction in the quantity of organic cow's milk is considerably more negative compared to conventional cow's milk, suggesting that almond milk is more substitutable for organic cow's milk than for conventional cow's milk.

Event-study plots in Panel B of Figures 2.6 and 2.7 show that the treatment effects on the price and quantity of organic cow's milk become gradually more negative over time after the introduction of almond milk. The price of organic milk remained stable for five months after the introduction of almond milk but declined gradually starting six months later. Similarly, the quantity sold of organic cow's milk began to decline gradually two months after the introduction of almond milk.

Lactose-free cow's milk: The fifth and sixth columns of Table 2.5 present estimates of the average treatment effects for lactose-free cow's milk. The CS estimates indicate no economically or statistically significant changes in the prices of lactose-free milk, with small estimated increases ranging from 0.19% to 0.37%. The event-study plot in Panel C of Figure 2.6 illustrates that the price of lactose-free cow's milk gradually increased for up to eight months following the initial introduction of almond milk. The price of lactose-free cow's milk began to decline in the final few months of the year.

In contrast, the CS estimates indicate a statistically significant decline in the quantity of lactose-free cow's milk, ranging from -2.8% to -3.3%. The treatment effect on the quantity is gradually decreasing but stabilizes eight months post-introduction (Panel B of Figure 2.7). The significant reduction in the quantity of lactose-free cow's milk is comparable to that of organic cow's milk. Both lactose-free and organic cow's milk show approximately a 3 percent decline in quantity, indicating that almond milk is a closer substitute for organic cow's milk and lactose-free cow's milk than for conventional cow's milk. This is consistent with substitutability between these cow's milk products and almond milk.

Soymilk: The final two columns of Table 2.5 present the average treatment effects for soymilk. Since both soymilk and almond milk are categorized as plant-based, we might expect the negative price and quantity impacts on soymilk to be more economically and statistically significant than those for cow's milk categories. This is indeed what the results show. Notably, the reductions in both the price and quantity of soymilk are the most substantial among the four categories. Specifically, the price of soymilk is estimated to fall by between 0.6% and 1.0%, while its quantity shows significant average declines, ranging from 4.7% to 6.6%.

Event-study plots in Panel D of Figures 2.6 and 2.7 also indicate that the introduction of almond milk significantly reduces both the price and quantity of soymilk. Although there is a slight increase in the price of soymilk during the first four months following the introduction of almond milk, it decreases notably after four months. Conversely, the quantity of soymilk gradually decreases from the initial month, indicating that the quantity effect on soymilk is more pronounced compared to cow's milk categories.

2.7.1.2 Overall Quantity Effects

This subsection further explores the overall quantity effects of the introduction of refrigerated almond milk. Table 2.6 presents the changes in the annual per capita quantity for each product category driven by the introduction. During the 2008–2010 period, the annual per capita consumption of fluid cow's milk was, on average, 20.88 gallons, with that total quantity distribution as shown in Panel A.

In Panel B of Table 2.6, changes in annual per capita quantities are calculated by multiplying the annual per capita quantities (the second row of Panel A in Table 2.6) by the percentage quantity effects obtained from the CS estimates (Table 2.5). The annual per capita quantity for soymilk falls by between 0.026 and 0.037 gallons, depending on the summary measures of the CS estimates. The quantity of organic cow's milk falls by around 0.021 to 0.036 gallons, while lactose-free cow's milk falls by about 0.013 to 0.016 gallons. Notice the magnitude of changes in the annual per capita quantity of conventional cow's milk is comparable to those for other products due to its significant share (91%), despite its tiny percentage changes in quantity.

Across all product categories, the decrease in the annual per capita quantity driven by the introduction of refrigerated almond milk ranges from -0.055 to -0.086 gallons between 2008 and 2010. During the same period, the annual per capita quantity of refrigerated almond milk increased by 0.152 gallons. Therefore, refrigerated almond milk did not solely cannibalize soymilk and cow's milk but expanded the overall milk market, at least in the initial phase of its introduction from 2008 to 2010.

2.7.2 Implications of Different Sample Periods and Alternative Parallel Trends Assumptions

This subsection addresses potential concerns regarding the size of control groups and explores alternative parallel trend assumptions. Tables 2.A.1 through 2.A.4 for each product category

Table 2.6. Effect of almond milk introduction on the per capita quantities of cow's milk and soymilk

	Conventional cow's milk	Organic cow's milk	Lactose-free cow's milk	Soymilk	Total
Panel A: Average quantity shares and per capita quantities					
Quantity shares (2008–2010 average)	91.1%	4.1%	2.2%	2.6%	100%
Annual per capita quantity (2008-2010 average, gallons)	19.533	0.872	0.472	0.560	21.436
Panel B: Projected changes in per capita quantities, based on econometric estimates[a]					
Simple weighted avg	0.031	-0.036	-0.016	-0.037	-0.058
Avg event-time TEs	0.023	-0.030	-0.014	-0.034	-0.055
Avg group-specific TEs	-0.025	-0.021	-0.013	-0.026	-0.086

Source: Quantity shares are taken from Table 2.4. Per capita consumption of cow's milk from the USDA ERS (USDA ERS (2023)).

Note: The annual per capita consumption of cow's milk was, on average, 20.88 gallons from 2008 to 2010. The per capita quantity for each cow's milk category is prorated according to their quantity shares. The sum of per capita quantities across conventional, organic, and lactose-free cow's milk equals to 20.88 gallons. The per capita quantity for soymilk is prorated based on its quantity share in the combined market.

[a] In Panel B, changes in per capita quantities are calculated by multiplying the annual per capita quantity in Panel A by Callaway and Sant'Anna (CS) estimates of percentage changes from Table 2.5.

in Appendix 2.A provides CS estimates using different sample periods and alternative parallel trend assumptions.

2.7.2.1 Robustness of the CS Estimates to Different Sample Periods

One concern regarding the CS estimates relying on stores that had not yet adopted refrigerated almond milk as control groups is that the number of stores serving as the control group gets substantially smaller toward the end of the sample period. Smaller control groups may lead to less efficient inference procedures (Callaway and Sant'Anna 2021; Marcus and Sant'Anna 2021).

By the end of 2008, only 6% of sample stores had adopted refrigerated almond milk. The remaining 94% of stores can serve as a control group for stores adopting almond milk in 2008, at least for one month. However, by the end of 2009, 74% of stores had adopted refrigerated almond milk, significantly reducing the number of potential control stores in 2010. For example, for stores that adopted almond milk in October 2010, there are 3,239 control stores available to measure the immediate treatment effect, 450 stores for the effect after six months, and 162 stores for the treatment effect after twelve months. For details on the number of stores adopting almond milk each month, refer to Table 2.2.

One way to check the robustness of the CS estimates to the number of control stores toward the end of 2010 is to stop the estimation in June 2010 in the sample period so that the number of control stores exceeds at least 3% of the total stores (around 300 stores). Across all four product categories, the price effects estimated using the shortened sample period show no significant differences compared to those derived from the original sample period. For quantity effects, the CS estimates from the shortened sample are largely consistent with those from the original sample, except for a slightly greater reduction in the quantity of conventional cow's milk when the sample period ends in June 2010. However, the magnitude of this effect is smaller than that observed for organic cow's milk, lactose-free cow's milk, and soymilk. Importantly, the order of magnitude for the quantity effects remains consistent with the original sample. Column (3) of Appendix Tables 2.A.1 through 2.A.4 reports the detailed CS estimates using a shortened sample period ending in June 2010 while maintaining the parallel trend assumption used in the main specification.

2.7.2.2 Alternative Parallel Trend Assumptions

A potential threat to the causal effects of the introduction of refrigerated almond milk would be time-varying confounders that affect the paths of the prices and quantities of cow's milk and soymilk. As noted, the main CS specification relies on the parallel trend assumption, conditional on store-specific time-invariant demographics and retailer-fixed effects. Although this conditional parallel trend assumption cannot be directly tested, it is more credible if the prices and quantities of cow's milk and soy milk generally move together prior to the introduction of refrigerated almond milk. Some support for this movement is that the pre-treatment event-study coefficients, depicted in Figures 2.6 and 2.7, are predominantly insignificant and close to zero, with only minor deviations for the price and quantity of soymilk 12 months before the introduction.

The plausibility of the conditional parallel trend assumption in the main specification can be further evaluated by comparing the CS estimates based on alternative parallel trend assumptions. Two alternative parallel trend assumptions are considered: 1) the parallel trend assumption conditional on store-specific demographics and state-fixed effects, and 2) the parallel trend assumption conditional only on store-specific demographics. The first alternative assumes that prices and quantities of cow's milk and soymilk would follow the same paths in treated and control stores located in the same states. The second alternative imposes a stronger assumption that no time-varying confounders exist after conditioning for store-specific demographics.

The quantity effects under the alternative assumption are similar to those with the main specification, maintaining the same order of magnitude across the four categories. In contrast, the price effect estimates for lactose-free cow's milk and soymilk differ significantly from those in the main specification. However, the pre-treatment event-study plots on the price effects for lactose-free cow's milk and soymilk using the alternative parallel trend assumption indicate idiosyncratic pre-trends, which makes the baseline CS estimates more preferable. Column (5) of Appendix Tables 2.A.1 to 2.A.4 presents the CS estimates based on the first alternative parallel trend assumption, using the same sample period as the main specification. For comparison, Column (6) in these tables repeats the CS estimates from the main specification.

The estimates for the price and quantity effect under the stronger parallel trend assumption, which relies only on the store-specific demographics, differ significantly from those in the main specification and from those under the first alternative parallel trend assumption.

Notably, more pre-trend coefficients are estimated to be significantly different from zero than those under the first alternative. Additionally, the standard errors of the aggregate CS measure are larger compared to those in the main specification. The CS estimates under the second parallel trends assumption are reported in Column (4) of Appendix Tables 2.A.1 to 2.A.4.

2.8 Summary and Concluding Remarks

This chapter studies the effects of the introduction of refrigerated almond milk on the prices and quantities of traditional cow's milk and soymilk. The rapid growth of the plant-based milk market in the 2010s has been primarily driven by almond milk following its introduction into the refrigerated dairy section in 2008.

This chapter describes how almond milk entered into the refrigerated section of retail stores gradually around the United States and presents a corresponding economic model of the price and quantity responses of existing products. The theoretical framework, a Cournot product entry model, demonstrates that the degree of substitutability between existing and new products is an important factor in determining the changes in the price and quantity of existing products.

This chapter evaluates econometrically the price and quantity effects of the introduction of refrigerated almond milk by using the recently developed innovations to the now traditional difference-in-differences framework. Empirical findings also point to the importance of substitutability between existing and new products. During the initial introductory phase, the spread of refrigerated almond milk led to a noticeable short-run decline in the quantity sold in retail stores, with soymilk experiencing the sharpest drop (6%), followed by organic cow's milk (3%) and lactose-free cow's milk (3%). In contrast, conventional cow's milk showed minimal percentage change in both price and quantity.

Overall, the introduction of almond milk caused the annual per capita quantity of cow's milk and soymilk to fall by 0.055 to 0.086 gallons between 2008 and 2010. During the same period, the annual per capita quantity of refrigerated almond milk increased by 0.152 gallons. Therefore, the spread of almond milk to the refrigerated dairy section expanded the overall milk market.

The limited impacts on the quantity of conventional cow's milk—representing about 90% of the combined market—raise further questions about why the U.S. dairy industry has expressed concerned over the rise of plant-based milk. Note that the results in this chapter reflect short-run effects. During the introductory phase, consumers who view the new refrigerated almond milk as a closer substitute for the existing cow's milk are more likely to switch to the new product, including those who previously did not buy any cow's milk or soymilk products, such as vegans and certain consumer groups. Beyond the introductory phase, as the market for plant-based milk expands and consumer awareness of the new product increases, the quantity effects are expected to become more significant, even affecting conventional cow's milk.

The econometric results on the short-run effects during the introductory phase presented in this chapter naturally lead to further econometric study of the demand relationships between almond milk and other milk products once almond milk and other plant-based milk alternatives are more established. The next two chapters pursue such econometric demand analysis, the implications for market evolution and the U.S. cow's milk policy.

References

Alviola, P.A., IV, and O. Capps, Jr. 2010. "Household Demand Analysis of Organic and Conventional Fluid Milk in the United States Based on the 2004 Nielsen Homescan Panel." *Agribusiness* 26(3):369–388.

Amir, R., P. Erickson, and J. Jin. 2017. "On the Microeconomic Foundations of Linear Demand for Differentiated Products." *Journal of Economic Theory* 169:641–665.

Baker, A.C., D.F. Larcker, and C.C. Wang. 2022. "How Much Should We Trust Staggered Difference-in-Differences Estimates?" *Journal of Financial Economics* 144(2):370–395.

Borusyak, K., X. Jaravel, and J. Spiess. 2024. "Revisiting Event-Study Designs: Robust and Efficient Estimation." *Review of Economic Studies* Forthcoming.

Callaway, B., and P.H. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225(2):200–230.

Chaker, A.M. 2011. "Almond Breeze, Silk Pure Almond Spark Fight in Dairy Case." *Wall Street Journal*. Available online at: https://www.wsj.com/articles/SB10001424052748703791904576075790255426176.

Choi, H.J., M.K. Wohlgenant, and X. Zheng. 2013. "Household-Level Welfare Effects of Organic Milk Introduction." *American Journal of Agricultural Economics* 95(4):1009–1028.

Chu, W. 1992. "Demand Signalling and Screening in Channels of Distribution." *Marketing Science* 11(4):327–347.

De Chaisemartin, C., and X. d'Haultfoeuille. 2023. "Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey." *The Econometrics Journal* 26(3):C1–C30.

—. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *Amer-*

ican Economic Review 110(9):2964–2996.

DellaVigna, S., and M. Gentzkow. 2019. "Uniform Pricing in US Retail Chains." *The Quarterly Journal of Economics* 134(4):2011–2084.

Desai, P.S. 2000. "Multiple Messages to Retain Retailers: Signaling New Product Demand." *Marketing Science* 19(4):381–389.

Federal Trade Commission. 2003. *Slotting Allowances in the Retail Grocery Industry: Selected Case Studies in Five Product Categories*. Washington DC. Available online at: `https://www.ftc.gov/sites/default/files/documents/reports/use-slotting-allowances-retail-grocery-industry/slottingallowancerpt031114.pdf`.

Franklin-Wallis, O. 2019. "White Gold; the Unstoppable Rise of Alternative Milks." *The Guardian*. Available online at: https://www.theguardian.com/news/2019/jan/29/white-gold-the-unstoppable-rise-of-alternative-milks-oat-soy-rice-coconut-plant.

Gardner, J. 2022. "Two-Stage Differences in Differences." arXiv preprint arXiv:2207.05943.

Gentzkow, M. 2007. "Valuing New Goods in a Model with Complementarity: Online newspapers." *American Economic Review* 97(3):713–744.

Ghanem, D., P.H. Sant'Anna, and K. Wüthrich. 2024. "Selection and Parallel Trends." arXiv preprint arXiv:2203.09001v11.

Goodman-Bacon, A. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225(2):254–277.

Hausman, J.A. 1996. "Valuation of New Goods under Perfect and Imperfect Competition." In *The economics of new goods*. University of Chicago Press, pp. 207–248.

Hausman, J.A., and G.K. Leonard. 2002. "The Competitive Effects of a New Product Introduction: A Case Study." *The Journal of Industrial Economics* 50(3):237–263.

Khanal, B., and R.A. Lopez. 2021. "Demand for Plant-Based Beverages and Competition in Fluid Milk Markets." Available online at SSRN https://ssrn.com/abstract=3888564.

Lariviere, M.A., and V. Padmanabhan. 1997. "Slotting Allowances and New Product Introductions." *Marketing Science* 16(2):112–128.

Marcus, M., and P.H. Sant'Anna. 2021. "The Role of Parallel Trends in Event Study Settings: An Application to Environmental Economics." *Journal of the Association of Environmental and Resource Economists* 8(2):235–275.

Marx, P., E. Tamer, and X. Tang. 2024. "Parallel Trends and Dynamic Choices." *Journal of Political Economy Microeconomics* 2(1):129–171.

McCarthy, K., M. Parker, A. Ameerally, S. Drake, and M. Drake. 2017. "Drivers of Choice for Fluid Milk versus Plant-Based Alternatives: What Are Consumer Perceptions of Fluid Milk?" *Journal of Dairy Science* 100(8):6125–6138.

Petrin, A. 2002. "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy* 110(4):705–729.

Ruby, M.B. 2012. "Vegetarianism. A blossoming Field of Study." *Appetite* 58(1):141–150.

Schiano, A., W. Harwood, P. Gerard, and M. Drake. 2020. "Consumer Perception of the Sustainability of Dairy Products and Plant-Based Dairy Alternatives." *Journal of Dairy Science* 103(12):11228–11243.

Shurtleff, W., and A. Aoyagi. 2013. "History of Soymilk and Other Non-Dairy Milks (1226-2013): Including Infant Formulas, Calf Milk Replacers, Soy Creamers, Soy Shakes, Soy Smoothies, Almond Milk, Coconut Milk, Peanut Milk, Rice Milk, Sesame Milk, etc." *Soyinfo Center.*

Singh, N., and X. Vives. 1984. "Price and Quantity Competition in a Differentiated Duopoly."

The Rand Journal of Economics, pp. 546–554.

Slade, P. 2023. "Does Plant-Based Milk Reduce Sales of Dairy Milk? Evidence from the Almond Milk Craze." *Agricultural and Resource Economics Review* 52(1):112–131.

Stewart, H., F. Kuchler, J. Cessna, and W. Hahn. 2020. "Are Plant-Based Analogues Replacing Cow's Milk in the American Diet?" *Journal of Agricultural and Applied Economics* 52(4):562–579.

Sudhir, K., and V.R. Rao. 2006. "Do Slotting Allowances Enhance Efficiency or Hinder Competition?" *Journal of Marketing Research* 43(2):137–155.

Sullivan, M.W. 1997. "Slotting Allowances and the Market for New Products." *The Journal of Law and Economics* 40(2):461–494.

Sun, L., and S. Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225(2):175–199.

Sutton, J. 1991. *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. MIT press.

Tonsor, G.T., J.L. Lusk, and T.C. Schroeder. 2023. "Market Potential of New Plant-Based Protein Alternatives: Insights from Four US Consumer Experiments." *Applied Economic Perspectives and Policy* 45(1):164–181.

U.S. Department of Agriculture, Economics Research Service. 2023. *Food Availability (per capita) Data System*. Washington DC. Available online at: `https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/`.

Vives, X. 2008. "Innovation and Competitive Pressure." *The Journal of Industrial Economics* 56(3):419–469.

Wolf, C.A., T. Malone, and B.R. McFadden. 2020. "Beverage Milk Consumption Patterns in the United States: Who Is Substituting from Dairy to Plant-Based Beverages?" *Journal* *of Dairy Science* 103(12):11209–11217.

Zhao, S., L. Wang, W. Hu, and Y. Zheng. 2023. "Meet the Meatless: Demand for New Generation Plant-Based Meat Alternatives." *Applied Economic Perspectives and Policy* 45(1):4–21.

Appendices

2.A Additional Tables and Figures

Table 2.A.1. CS estimates from alternative specifications for conventional cow's milk

	Dependent variable: log of price					
	(1)	(2)	(3)	(4)	(5)	(6)
Simple weighted avg ($\hat{\theta}_w$)	0.0060 (0.0082)	-0.0082 (0.0054)	-0.0056 (0.0073)	0.0139 (0.0109)	-0.0058 (0.0075)	-0.0086 (0.0069)
Avg event-time TEs ($\hat{\theta}_{es}$)	0.0102 (0.0091)	-0.0118 (0.0057)	-0.0006 (0.0103)	0.0110 (0.0100)	-0.0023 (0.0078)	-0.0065 (0.0053)
Avg group-specific TEs ($\hat{\theta}_{gr}$)	0.0064 (0.0076)	-0.0040 (0.0072)	-0.0050 (0.0055)	0.0138 (0.0110)	-0.0020 (0.0085)	-0.0077 (0.0049)
	Dependent variable: log of quantity					
Simple weighted avg ($\hat{\theta}_w$)	0.0047 (0.0125)	0.0015 (0.0109)	-0.0084 (0.0088)	-0.0103 (0.0225)	-0.0108 (0.0156)	0.0016 (0.0106)
Avg event-time TEs ($\hat{\theta}_{es}$)	0.0006 (0.0142)	0.0009 (0.0099)	-0.0246 (0.0108)	-0.0053 (0.0167)	-0.0061 (0.0158)	-0.0013 (0.0100)
Avg group-specific TEs ($\hat{\theta}_{gr}$)	-0.0002 (0.0160)	-0.0025 (0.0164)	-0.0073 (0.0082)	-0.0133 (0.0224)	-0.0135 (0.0176)	0.0012 (0.0102)
Observations	280,500	280,500	280,500	336,600	336,600	336,600
Sample: 2008m1–2010m6	X	X	X			
Sample: 2008m1–2010m12				X	X	X
PT 1: store demographics	X			X		
PT 2: store demographics + state		X			X	
PT 3: store demographics + retailer			X			X

Note: The 95% confidence intervals are shown in parentheses.

Table 2.A.2. CS estimates from alternative specifications for organic cow's milk

	Dependent variable: log of price					
	(1)	(2)	(3)	(4)	(5)	(6)
Simple weighted avg ($\hat{\theta}_w$)	-0.0182 (0.0150)	-0.0042 (0.0039)	-0.0070 (0.0024)	-0.0301 (0.0246)	-0.0067 (0.0035)	-0.0075 (0.0031)
Avg event-time TEs ($\hat{\theta}_{es}$)	-0.0237 (0.0164)	-0.0045 (0.0037)	-0.0104 (0.0025)	-0.0210 (0.0203)	-0.0037 (0.0039)	-0.0085 (0.0025)
Avg group-specific TEs ($\hat{\theta}_{gr}$)	-0.0158 (0.0146)	-0.0025 (0.0031)	-0.0053 (0.0022)	-0.0277 (0.0253)	-0.0054 (0.0035)	-0.0070 (0.0030)
	Dependent variable: log of quantity					
Simple weighted avg ($\hat{\theta}_w$)	0.0232 (0.0136)	-0.0339 (0.0152)	-0.0307 (0.0142)	0.0564 (0.0183)	-0.0249 (0.0199)	-0.0413 (0.0205)
Avg event-time TEs ($\hat{\theta}_{es}$)	0.0332 (0.0192)	-0.0529 (0.0164)	-0.0347 (0.0154)	0.0337 (0.0152)	-0.0259 (0.0177)	-0.0246 (0.0117)
Avg group-specific TEs ($\hat{\theta}_{gr}$)	0.0238 (0.0127)	-0.0283 (0.0128)	-0.0263 (0.0125)	0.0537 (0.0189)	-0.0196 (0.0186)	-0.0345 (0.0193)
Observations	270,210	270,210	270,210	324,252	324,252	324,252
Sample: 2008m1–2010m6	X	X	X			
Sample: 2008m1–2010m12				X	X	X
PT 1: store demographics	X			X		
PT 2: store demographics + state		X			X	
PT 3: store demographics + retailer			X			X

Note: The 95% confidence intervals are shown in parentheses.

Table 2.A.3. CS estimates from alternative specifications for lactose-free cow's milk

	Dependent variable: log of price					
	(1)	(2)	(3)	(4)	(5)	(6)
Simple weighted avg ($\hat{\theta}_w$)	-0.0096 (0.0053)	-0.0146 (0.0041)	0.0074 (0.0028)	-0.0195 (0.0093)	-0.0157 (0.0049)	0.0037 (0.0033)
Avg event-time TEs ($\hat{\theta}_{es}$)	-0.0135 (0.0067)	-0.0183 (0.0044)	0.0077 (0.0029)	-0.0154 (0.0077)	-0.0162 (0.0051)	0.0031 (0.0029)
Avg group-specific TEs ($\hat{\theta}_{gr}$)	-0.0109 (0.0046)	-0.0152 (0.0034)	0.0047 (0.0022)	-0.0200 (0.0086)	-0.0157 (0.0046)	0.0019 (0.0034)
	Dependent variable: log of quantity					
Simple weighted avg ($\hat{\theta}_w$)	-0.0044 (0.0082)	-0.0072 (0.0080)	-0.0298 (0.0099)	0.0059 (0.0148)	-0.0316 (0.0158)	-0.0331 (0.0131)
Avg event-time TEs ($\hat{\theta}_{es}$)	-0.0013 (0.0098)	-0.0045 (0.0091)	-0.0365 (0.0127)	0.0010 (0.0108)	-0.0186 (0.0138)	-0.0281 (0.0115)
Avg group-specific TEs ($\hat{\theta}_{gr}$)	-0.0089 (0.0079)	-0.0034 (0.0076)	-0.0250 (0.0081)	0.0031 (0.0140)	-0.0326 (0.0141)	-0.0296 (0.0135)
Observations	277,440	277,440	277,440	332,928	332,928	332,928
Sample: 2008m1–2010m6	X	X	X			
Sample: 2008m1–2010m12				X	X	X
PT 1: store demographics	X			X		
PT 2: store demographics + state		X			X	
PT 3: store demographics + retailer			X			X

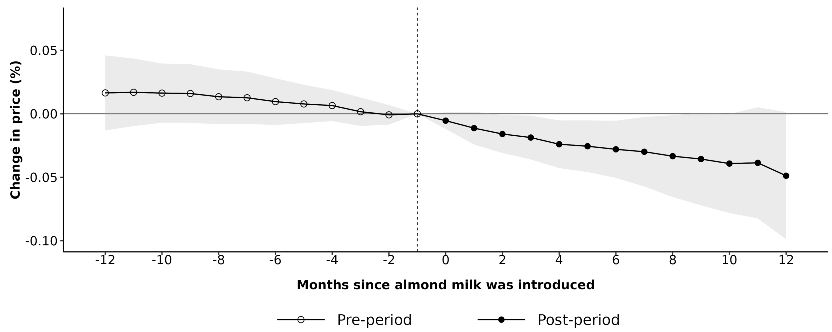
Note: The 95% confidence intervals are shown in parentheses.

Table 2.A.4. CS estimates from alternative specifications for soymilk

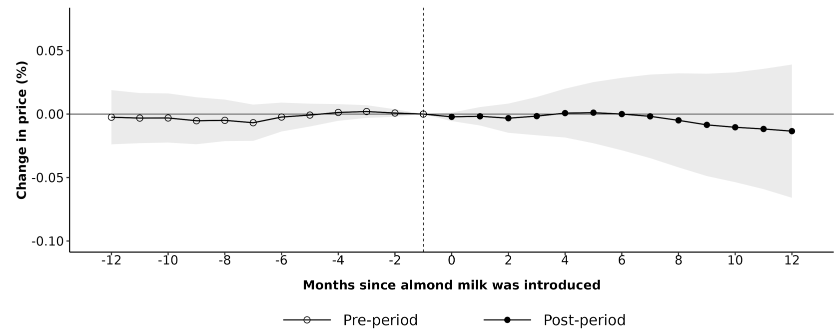
	Dependent variable: log of price					
	(1)	(2)	(3)	(4)	(5)	(6)
Simple weighted avg ($\hat{\theta}_w$)	-0.0228 (0.0085)	-0.0267 (0.0077)	-0.0074 (0.0057)	-0.0290 (0.0133)	-0.0295 (0.0093)	-0.0098 (0.0054)
Avg event-time TEs ($\hat{\theta}_{es}$)	-0.0309 (0.0083)	-0.0329 (0.0067)	-0.0115 (0.0062)	-0.0204 (0.0114)	-0.0297 (0.0080)	-0.0060 (0.0052)
Avg group-specific TEs ($\hat{\theta}_{gr}$)	-0.0177 (0.0074)	-0.0207 (0.0065)	-0.0045 (0.0052)	-0.0248 (0.0141)	-0.0275 (0.0087)	-0.0080 (0.0041)
	Dependent variable: log of quantity					
Simple weighted avg ($\hat{\theta}_w$)	0.0189 (0.0150)	-0.0191 (0.0109)	-0.0428 (0.0101)	0.0236 (0.0365)	-0.0496 (0.0199)	-0.0659 (0.0139)
Avg event-time TEs ($\hat{\theta}_{es}$)	0.0269 (0.0155)	-0.0257 (0.0093)	-0.0597 (0.0118)	0.0176 (0.0240)	-0.0149 (0.0168)	-0.0472 (0.0125)
Avg group-specific TEs ($\hat{\theta}_{gr}$)	0.0129 (0.0161)	-0.0170 (0.0105)	-0.0431 (0.0096)	0.0197 (0.0400)	-0.0417 (0.0188)	-0.0600 (0.0162)
Observations	279,600	279,600	279,600	335,520	335,520	335,520
Sample: 2008m1–2010m6	X	X	X			
Sample: 2008m1–2010m12				X	X	X
PT 1: store demographics	X			X		
PT 2: store demographics + state		X			X	
PT 3: store demographics + retailer			X			X

Note: The 95% confidence intervals are shown in parentheses.

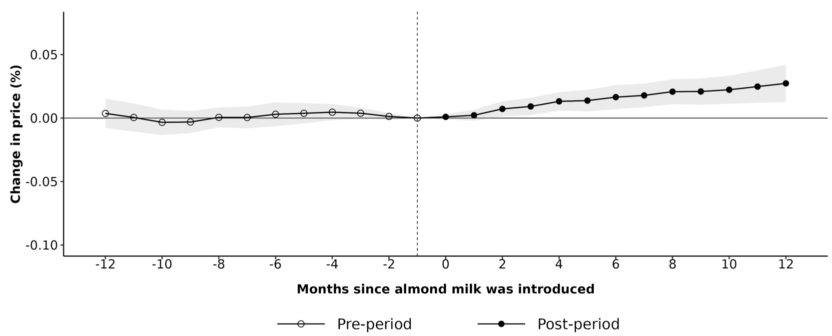
Panel A. Price of conventional cow's milk



Panel B. Price of organic cow's milk



Panel C. Price of lactose-free cow's milk



Panel D. Price of soymilk

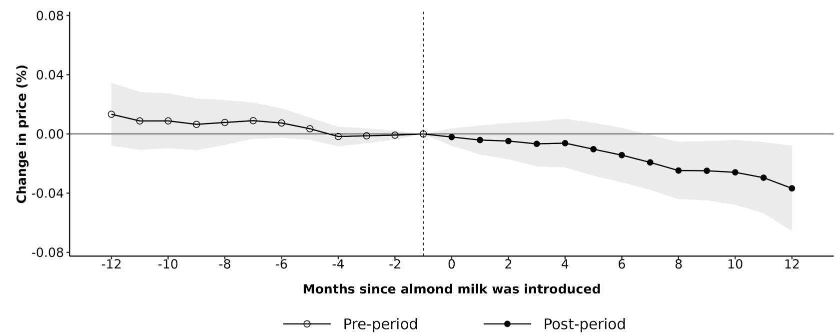
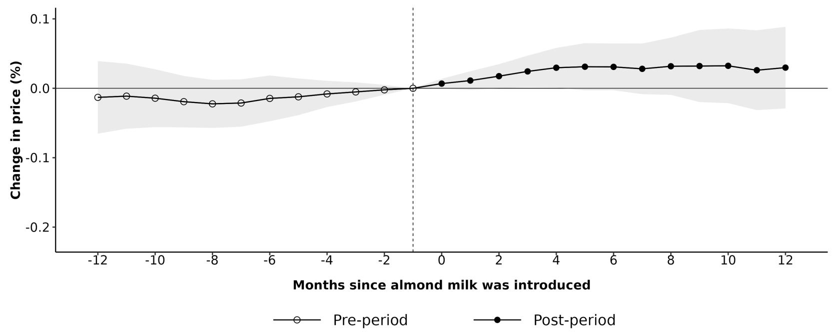
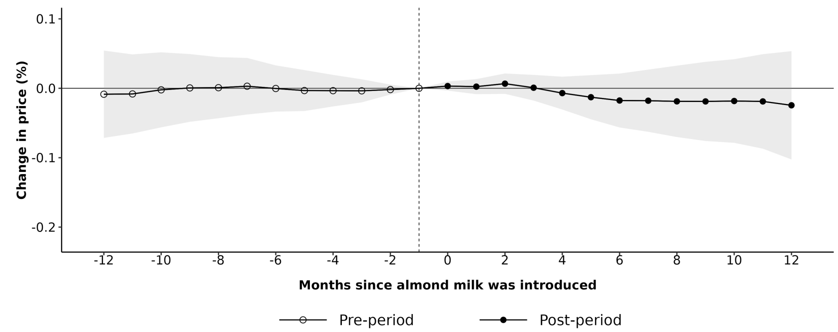


Figure 2.A.1. Dynamic TWFE estimates for the price of cow's milk and soymilk
Note: The figures show average treatment effects relative to one month before the initial introduction month obtained from the two-way fixed-effect estimator, controlling for demographic-specific trends and retailer-specific trends. The vertical axis represents the percentage change in the price. Circled dots represent pre-period treatment effect estimates, while filled dots represent post-period treatment effects.

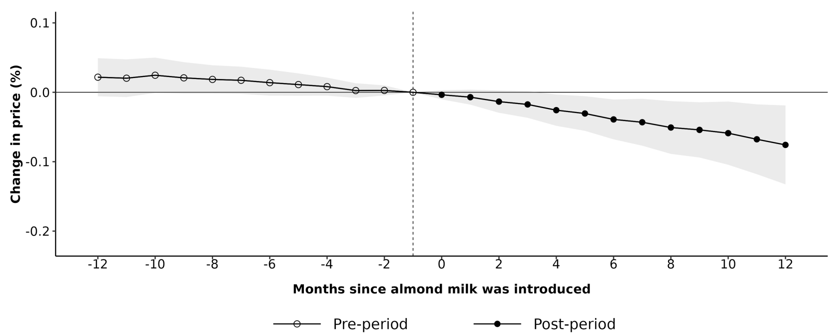
Panel A. Quantity of conventional cow's milk



Panel B. Quantity of organic cow's milk



Panel C. Quantity of lactose-free cow's milk



Panel D. Quantity of soymilk

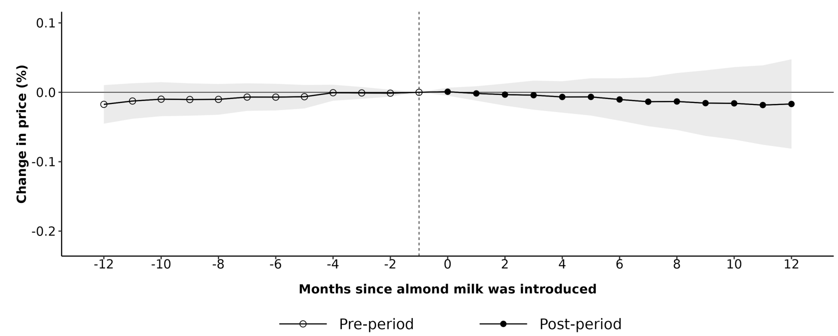


Figure 2.A.2. Dynamic TWFE estimates for the quantity of cow's milk and soymilk
Note: The figures show average treatment effects relative to one month before the initial introduction month obtained from the two-way fixed-effect estimator, controlling for demographic-specific trends and retailer-specific trends. The vertical axis represents the percentage change in the price. Circled dots represent pre-period treatment effect estimates, while filled dots represent post-period treatment effects.

2.B Mathematical Appendix

Define the inverse demand model as: $p_i = a_i - q_i - \sum_{i \neq j} \gamma_{ij} q_j$ for $i = 1, 2, 3$; $\gamma_{ij} = \gamma_{ji}$, where $\gamma_{ij} \in (0, 1)$ and $\gamma_{ij} - \gamma_{ik}\gamma_{jk} > 0$.

Equantion (2.3) through (2.5): Solving for profit maximization for each firm under Cournot competition gives the following first-order conditions: $2q_i + \gamma_{ij} q_j + \gamma_{ik} q_k = a_i - c_i$, $i \neq j, k$. Solving for q_i gives the following equilibrium quantities under the Cournot triopoly.

$$\begin{pmatrix} q_1^T \\ q_2^T \\ q_3^T \end{pmatrix} = \frac{1}{A} \begin{pmatrix} 4 - \gamma_{23}^2 & -2\gamma_{12} + \gamma_{13}\gamma_{23} & -2\gamma_{13} + \gamma_{12}\gamma_{23} \\ -2\gamma_{12} + \gamma_{13}\gamma_{23} & 4 - \gamma_{13}^2 & -2\gamma_{23} + \gamma_{12}\gamma_{13} \\ -2\gamma_{13} + \gamma_{12}\gamma_{23} & -2\gamma_{23} + \gamma_{12}\gamma_{13} & 4 - \gamma_{12}^2 \end{pmatrix} \begin{pmatrix} a_1 - c_1 \\ a_2 - c_2 \\ a_3 - c_3 \end{pmatrix} \tag{2.B.1}$$

where $A = 8 - 2\gamma_{12}^2 - 2\gamma_{13}^2 - 2\gamma_{23}^2 + 2\gamma_{12}\gamma_{13}\gamma_{23}$. Note that using FOCs, the equilibrium prices can be expressed as: $p_i^T = q_i^T + c_i$.

Equation (2.6): assume that $a \equiv a_1 = a_2$ and $c \equiv c_1 = c_2$. By comparing the quantities of incumbent products before and after the introduction of the third product, one can find the following inequality.

$$q_1^T - q_1^D = \underbrace{\frac{1}{A}}_{>0} \underbrace{\frac{2\gamma_{13} - \gamma_{12}\gamma_{23}}{2 + \gamma_{12}}}_{>0} \underbrace{\{(\gamma_{13} + \gamma_{23})(a - c) - (2 + \gamma_{12})(a_3 - c_3)\}}_{<0} < 0 \tag{2.B.2}$$

where the first term on the right-hand side of Equation (2.B.2) is positive when the second-order conditions for the profit maximization hold, and the second term is positive because products are gross substitutes. The negative sign of the third term in Equation (2.B.2) follows from $q_3^T = (2 - \gamma_{12})\{-(\gamma_{13} + \gamma_{23})(a - c) + (2 + \gamma_{12})(a_3 - c_3)\}/A > 0$. The derivation for the inequality, $q_2^T - q_2^D < 0$, is analogous to Equation (2.B.2) and therefore omitted. Finally, since $p_i^T = q_i^T + c_i$, the prices of existing products after the introduction are lower than those before the introduction, $p_i^T - p_i^D < 0$ for $i = 1, 2$.

Equation (2.7): assume that $a \equiv a_1 = a_2$ and $c \equiv c_1 = c_2$. When the third firm does not enter due to a high entry cost or a low market size for the third product, the duopoly equilibrium quantities of incumbent products are the same. $q_1^D = q_2^D = (a - c)/(2 + \gamma_{12})$. When the third firm enters, equilibrium quantities are as in Equation (2.B.2). Then the following relationship holds:

$$\begin{aligned} q_2^T - q_1^T = & \underbrace{\frac{1}{A}}_{>0} (\gamma_{23} - \gamma_{13}) \underbrace{\{(\gamma_{13} + \gamma_{23})(a - c) - (2 + \gamma_{12})(a_3 - c_3)\}}_{<0} \leq 0 \\ \iff & \quad \gamma_{23} \geq \gamma_{13} \end{aligned} \tag{2.B.3}$$

Equation (2.7) holds since Equation (2.B.3) holds and $q_1^D = q_2^D$ under the maintained assumption that $a \equiv a_1 = a_2$ and $c \equiv c_1 = c_2$. Note that using the first-order conditions, the equilibrium prices can be expressed as: $p_i^T = q_i^T + c_i$, and $p_2^T - p_1^T \leq 0 \;\; \gamma_{23} \geq \gamma_{13}$.

Equation (2.8): taking the partial derivative of Equation (2.B.3) with respect to a_3 gives the following inequality.

$$\begin{aligned} \frac{\partial (q_2^T - q_1^T)}{\partial a_3} = & \underbrace{-\frac{1}{A}}_{<0} (\gamma_{23} - \gamma_{13}) \underbrace{\{2 + \gamma_{12}\}}_{>0} \leq 0 \\ \iff & \quad \gamma_{23} \geq \gamma_{13} \end{aligned} \tag{2.B.4}$$

Chapter 3

Estimating the Effects of Plant-Based Milk on Retail Cow's Milk Prices and Quantities

3.1 Introduction

Cow's milk was the fourth largest U.S. farm commodity industry by revenue in 2022, with about $57 billion in annual farm sales. However, per capita consumption of fluid milk products—about 30% of all use of farm milk—has steadily declined since at least the 1950s, with a notable acceleration of the percentage rate of decline in the 2010s. As fluid cow's milk consumption has declined, plant-based milk has experienced substantial growth over the past two decades.

This chapter seeks to evaluate to what extent the decline in consumption of fluid cow's milk is attributable to the availability of plant-based milk. The findings from this evaluation have significant implications for the U.S. dairy industry's strategy to reverse the declining trend. The U.S. dairy industry has made efforts to position cow's milk relative to plant-based milk. For example, the National Milk Producers Federation (NMPF) has urged the Food and Drug Administration to prevent plant-based milk manufacturers from using the term "milk" (NMPF 2023). Recent promotional campaigns have emphasized the "realness" of cow's milk compared to plant-based milk (for example, Wood Milk 2023). These strategies could be effective if the decline in consumption is primarily due to the rise of plant-based milk. However, if plant-based milk has contributed little to the decline, such efforts would be ineffective in expanding cow's milk sales.

This chapter estimates the impact of the availability of plant-based products on the prices and quantities of cow milk products in the United States, using a discrete-choice demand model with random coefficients. The demand model is flexible in that it allows taste for product characteristics (or other unobserved characteristics) to vary among consumers and permits these characteristics to be correlated with each other, following Train (1998) and Revelt and Train (1998).

The crucial ingredient for evaluating the impact on the market shares of cow's milk is to estimate the substitution pattern among the several products. Price and quantity responses to plant-based milk expansion are likely to be heterogeneous among product groups within the category of cow's milk, namely, conventional milk, organic milk, and lactose-free milk. Purchases observed in household scanner data indicate that households that purchased organic milk or lactose-free milk are more likely also to buy plant-based milk than are conventional milk buyers.

Household scanner data, providing detailed information on purchases of dairy and non-dairy products, including prices and quantities, is matched with store-level data to represent the choice sets faced by each household. The estimation process retrieves the demand parameters that govern consumer behavior and the valuation of different product characteristics. The supply side of the retail market is modeled as an oligopolistic market structure where cow's milk processors and plant-based milk manufacturers reach price decisions following a Bertrand-Nash price competition. Product-level marginal costs per unit are derived in each geographical market from the equilibrium conditions of the demand and supply models.

With all of these elements at hand, to illustrate implications, the chapter ends by evaluating the impact of the spread of plant-based milk with an extreme counterfactual simulation: the removal of all plant-based milk products from the choice set. Exploring this counterfactual scenario shows the overall effects of plant-based milk on the prices and quantities demanded of incumbent products.

3.2 Overview of Results and Connections to the Empirical Literature

Results of the demand estimation show important heterogeneities across different types of cow's milk products, with organic and lactose-free products being closer substitutes for plant-based products than conventional cow milk is. Correspondingly, the counterfactual analysis shows that the removal of all plant-based milk products from the choice set would lead to a 23% increase in the equilibrium retail quantity for organic cow's milk, a 16% increase for lactose-free cow's milk, and an 11% increase for conventional cow's milk. The counterfactual results imply that the removal of plant-based milk would raise the annual per capita consumption of conventional cow's milk by 1.46 gallons and by 0.26 gallons each for organic cow's milk and lactose-free cow's milk. These results, plus the fact that a large share of organic farm milk is devoted to fluid products, imply that the increase in retail competition faced by cow's milk due to the spread of plant-based milk was particularly negative for organic dairy farms.

The counterfactual results are useful for evaluating the extent to which the decline in the consumption of cow's milk is attributable to the rise in plant-based milk. Out of the 5.1-gallon decrease in annual per capita cow's milk consumption from 2006 to 2020, the counterfactual experiment reveals that the availability of plant-based milk is responsible for 38% of this drop. This result is broadly in line with previous studies indicating that plant-based milk did not lead to a one-for-one displacement of cow's milk. The magnitude of the impact, however, is larger than results based on correlations or summary data (Stewart et al. 2020; Slade 2023).

This research builds on a large body of literature on the effects of new product introduction. Economists have studied the effect of new product introduction on competition (Petrin 2002; Goolsbee and Petrin 2004) and consumer welfare (Hausman and Leonard 2002; Gentzkow 2007; Choi, Wohlgenant, and Zheng 2013) among related topics. This research contributes to this literature by estimating the impact of the expansion of plant-based milk products on the market for cow's milk products. The detailed procedures outlined in the data section, involving the matching of household purchase data with store-level data, contribute to the literature by showing how to better recover consumer choice sets in demand estimation. Notably, this approach avoids susceptibility to measurement errors associated with the imputation of prices of alternatives.

The findings in this research also contribute to the burgeoning literature on demand for plant-based products. Several studies have attempted to estimate substitution relationships between animal-based protein and plant-based protein (Alviola and Capps 2010; Dharmasena and Capps 2014; Khanal and Lopez 2021; Tonsor, Lusk, and Schroeder 2023; Zhao et al. 2023). A few recent studies focus on the impacts of plant-based milk on the dairy industry (Stewart et al. 2020; Slade 2023). This research enhances the literature by offering a more rigorous estimation of detailed substitution effects and by providing the first structural estimates of the impact of plant-based milk.

This chapter proceeds as follows. Section 3.3 provides some background on cow's milk and plant-based milk markets and describes purchase patterns observed in household data. Section 3.4 describes the demand and supply model for cow's milk and plant-based milk. Section 3.5 describes the construction of the dataset for estimating the demand model. Section 3.6 presents the results of the demand estimation and substitution patterns. Section 3.7 presents the results of the counterfactual simulation.

3.3 Backgrounds on Cow's Milk and Plant-Based Milk Market

This section employs household scanner data to describe consumption patterns for cow's milk and plant-based milk products.

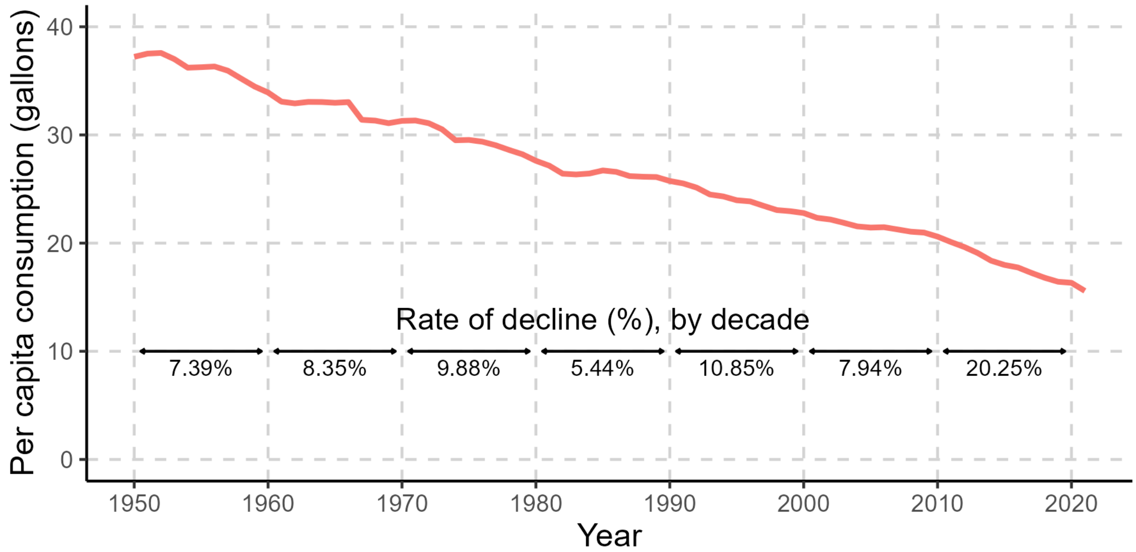


Figure 3.1. U.S. per capita consumption of fluid milk (1950–2021)

Source: USDA Food Availability System, Fluid milk, 1950–2021.

Note: The rates of decline are calculated using a 5-year moving average to smooth out the fluctuation around the beginning or the end of the time frame.

3.3.1 Long-Term Decline in Cow's Milk Fluid Use Compared to the Past Decade

The long-term decline in per capita consumption of fluid cow's milk in the United States is not new and, indeed, has been underway for seven decades (Figure 3.1). However, the rate of decline has accelerated, with a 20% decrease from 2010 to 2019.

The rise of plant-based milk products has garnered attention as a potential contributor to the accelerated decline in fluid cow's milk consumption (Stewart et al. 2020; Wolf, Malone, and McFadden 2020; Slade 2023). These studies argue that plant-based milk accounts for part of the recent decline in cow's milk consumption as the quantity share of plant-based milk in the combined market has risen to 11% by 2020. The plant-based milk segment has attracted significant product development and innovation, as investors saw the potential for further expansion of the market for plant-based substitutes for cow's milk.

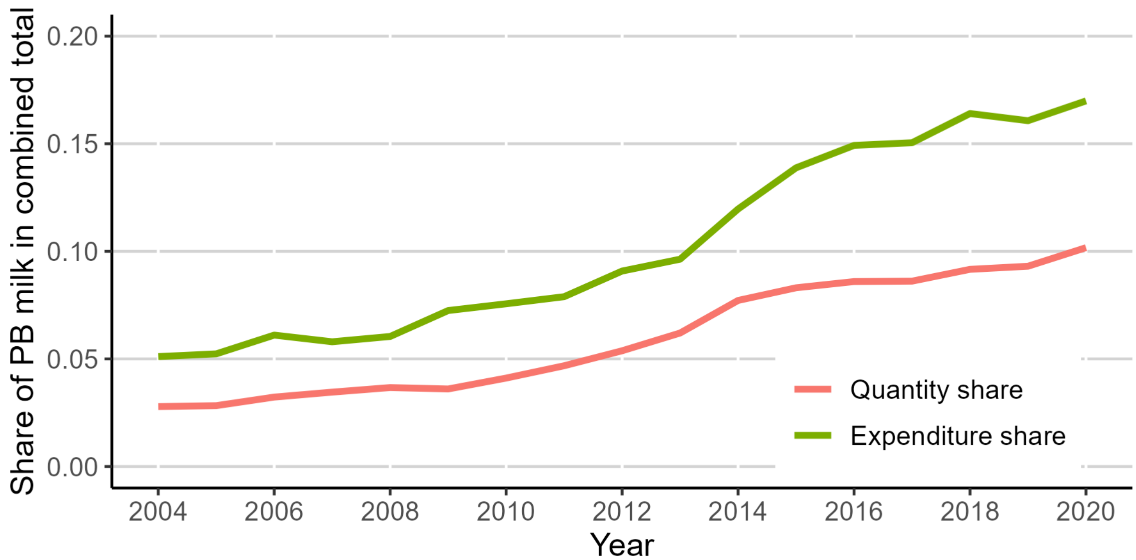


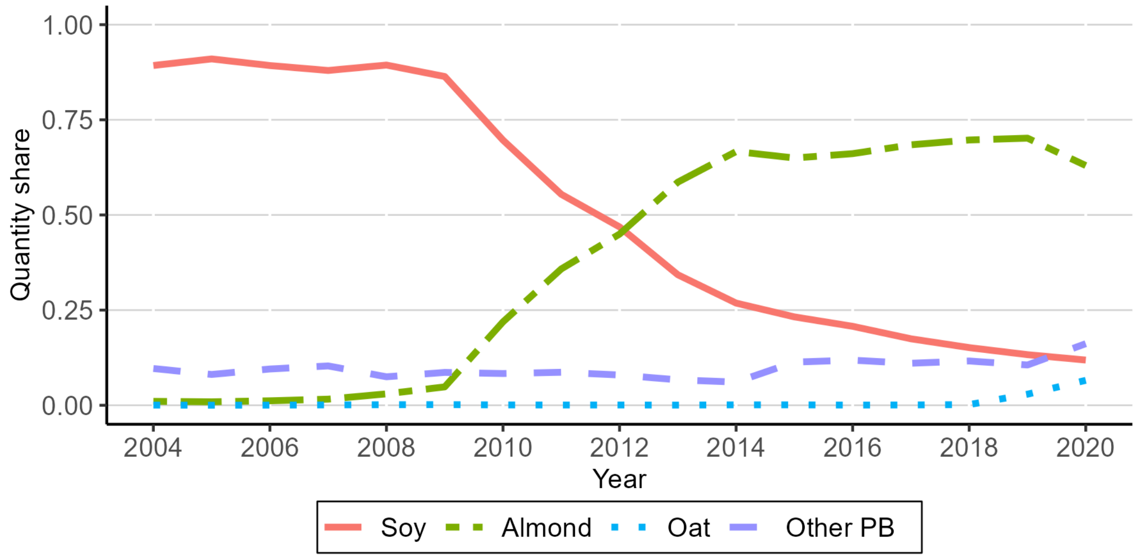
Figure 3.2. Quantity and revenue shares of plant-based milk in the U.S. (2004–2020)
Source: NielsenIQ Homescan Panel (2004–2020)

3.3.2 Growth of Plant-Based Milk

Plant-based milk or similar beverages have a rich historical tradition in various cultures around the world. Nevertheless, it was not until 1996 that the first brand of plant-based milk alternatives, White Wave's Silk soy milk, made its debut in milk cartons within the refrigerated dairy section of U.S. grocery stores. Although plant-based milk alternatives were available at some stores, the market for these alternatives remained niche, primarily catering to vegan consumers and often placed in the health food section, typically distant from the dairy aisle.

Plant-based milk quantity and revenue experienced substantial growth in the 2010s (Figure 3.2). The rapid expansion of sales of plant-based milk products coincided with the introduction of refrigerated almond milk in late 2008. Figure 3.3, Panel A, illustrates the rise of almond milk. Almond milk surpassed soy milk in 2013, capturing 59% of the sales volume and growing to 70% by 2019. The growth of plant-based milk products has been prominent in the refrigerated shelves, where they compete directly next to cow's milk (Panel

Panel A: Quantity share by plant source



Panel B: Quantity share by storage type

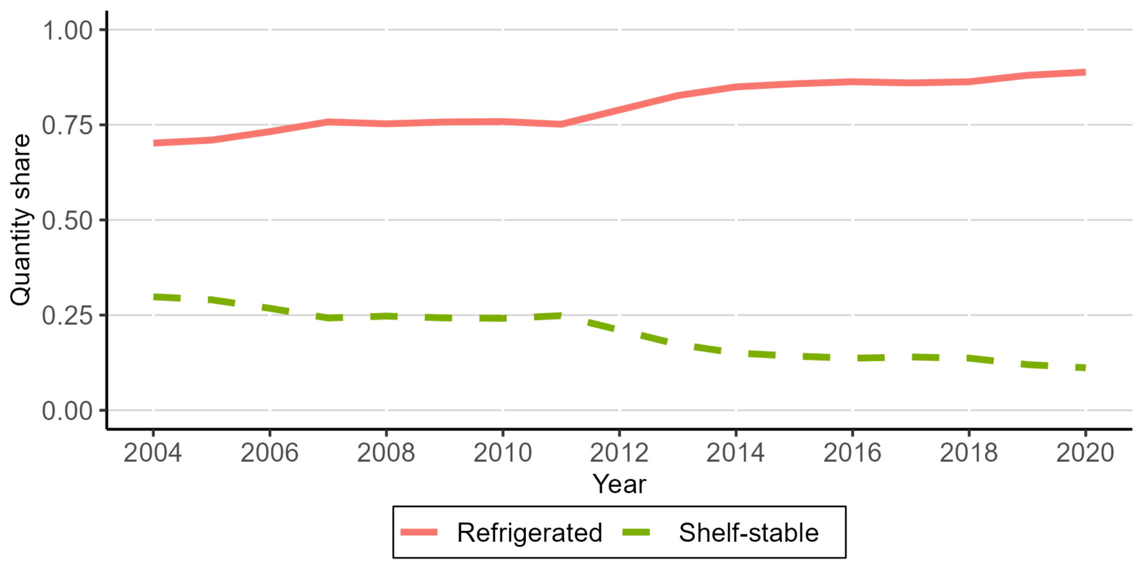


Figure 3.3. Quantity shares of plant-based milk by plant source and storage type in the U.S. (2004–2020)

Source: NielsenIQ Homescan Panel (2004–2020)

B of Figure 3.3).

3.3.3 Consumption of Cow's Milk and Plant-Based Milk

Examining household purchase data is useful for understanding how cow's milk and plant-based milk consumption vary across different demographic profiles of households. Household data also helps identify patterns of households switching between these products. NielsenIQ Homescan (HMS) provides basic facts on these and other consumption patterns.[1]

3.3.3.1 Consumption of Cow's Milk and Plant-Based Milk across Demographic Groups

Table 3.1 displays the consumption (purchase) patterns of cow's milk and plant-based milk among various demographic groups during the most recent sample period from 2018 to 2020. Most of the subsample means in Table 3.1 are statistically different at 5% from the other subsample means in the relevant. I also note some partial correlations from a series of linear regressions of quantities of products on household income and demographic variables.

Several differences in consumption patterns across groups are noteworthy. First, households tend to buy more organic cow's milk and plant-based milk as income rises, while middle-income households buy more conventional milk. Second, households with younger heads tend to purchase more of all cow's milk and plant-based milk product categories except lactose-free cow's milk. Households with older heads buy more lactose-free cow's milk compared to those with middle-aged heads. Third, households with more educated heads tend to purchase more organic cow's milk and plant-based milk than households without college degrees. Positive association of higher income, younger head age, and higher level of education attainment with organic cow's milk and plant-based milk are generally consistent with previous demand studies on organic milk (Dhar and Foltz 2005; Alviola and Capps 2010; Choi, Wohlgenant, and Zheng 2013) and on plant-based milk (Dharmasena and Capps 2014; Wolf, Malone, and McFadden 2020).

[1] Section 3.5 provides details on NielsenIQ HMS data.

Table 3.1. Consumption of cow's and plant-based milk by demographics (2018–2020)

Demographics		Total	By product group: Conventional cow	Organic cow	Lactose-free cow	Plant-based
			(unit: gallons per year)			
All sample households		18.14	14.91	0.72	0.78	1.73
By annual household income						
	$< \$50,000$	16.80	14.47	0.32	0.66	1.35
	$\$50,000$-$100,000$	19.14	15.86	0.66	0.80	1.83
	$> \$100,000$	19.00	14.59[a]	1.32	0.92	2.17
By demographic variable						
Age of head	< 45	19.81	15.82	1.06	0.84	2.08
	45-65	17.90	14.83	0.65	0.71	1.72
	> 65	16.53	13.94	0.43	0.82[a]	1.34
Education of head	No college	18.43	16.28	0.31	0.67	1.18
	College	18.03	14.41	0.87	0.82[a]	1.94
Race of head	White	19.43	16.46	0.67	0.68	1.62
	Black	10.73	7.66	0.39	0.97	1.72[a]
	Other race	17.89	13.04	1.30	1.14	2.42
Ethinicity	Non-Hispanic	18.01	14.99	0.69	0.69	1.65
	Hispanic	18.95[a]	14.37[a]	0.91[a]	1.36	2.31

Source: NielsenIQ Homescan Panel (2018–2020).

Note: This table shows the annual quantity purchased of cow's milk and plant-based milk by demographic group based on 80,218 households. The household demographic information is self-reported as applying to the respondent for the household.

[a] All coefficients, except those with a superscript "a," are statistically significant at a 5% significance level from a series of regression of the quantity purchased of each product category on demographic variables.

White households purchased more cow's milk than Black and other race households, especially for conventional cow's milk. Black households consumed less than half of the quantity of conventional cow's milk compared to White households, while their consumption of lactose-free milk was higher. Finally, Hispanic households consumed more lactose-free cow's milk and plant-based milk than non-Hispanic households.

In additional descriptive analyses of the relationships between product category purchases and demographic variables, the annual quantity purchased of each category was regressed on annual household income and demographic variables (age, education, race, and ethnicity of household heads) to assess whether differences in the annual quantity purchased among demographic groups, holding constant the other regressors, are statistically significant. For each demographic variable, the first category is used as a reference (for example, "Less than $50,000" for the annual household income). The regression results, with standard errors clustered on households, show that most coefficients are statistically significant at a 5% significance level, except those marked with superscripts "a" in Table 3.1.

3.3.3.2 Substitution Pattern between Cow's Milk and Plant-Based Milk

Authors outside of economics describe four main consumer motivations for buying organic products: healthiness, environmental concern, food safety, and animal welfare (for example, Hughner et al. 2007). Similarly, consumers' stated motivations for a plant-based diet often include animal welfare, healthiness, and environmental concerns (for example, Ruby 2012). Plant-based milk products are also lactose-free and thus provide an option for lactose-intolerant consumers.

Table 3.2 presents how the purchase experience and quantity purchased of plant-based milk vary with the purchase frequency of four consumer groups. Consumers are categorized into four groups depending on how often they purchase each of the four product categories: three types of cow's milk (conventional, organic, and lactose-free) and plant-based milk. For example, the "Organic cow" group comprises households that purchased organic cow's milk more frequently than the average number of purchases across all consumers. The remaining three groups—"Conventional cow," "Lactose-free cow," and "Plant-based milk"—are defined similarly.

Households that frequently buy organic and lactose-free cow's milk are more likely to buy plant-based milk compared to those who frequently buy conventional cow's milk. Specifically,

Table 3.2. Purchases of plant-based milk by the frequency of purchases of categories of cow's milk products

		Consumer groups by purchase frequency			
	All consumers	Conventional cow's milk	Organic cow's milk	Lactose-free cow's milk	Plant-based milk
Ever bought PB milk (%)	34.5%	19.2%	47.5%	41.0%	100%
Annual quantity of PB milk (gallons)	1.73	0.29	2.08	1.61	6.99
Number of households	80,218	52,742	6,607	9,116	18,384

Source: NielsenIQ Homescan Panel (2018–2020).

Note: This table shows the purchase experience and annual average quantity purchased of plant-based milk by category of households based on purchase frequency, exceeding the average frequency, among the three cow's milk product categories (conventional, organic, and lactose-free cow's milk) and plant-based milk. The sum of the number of households across the four groups does not equal to one because some households purchase products at above average frequency for two or more categories.

households with a high frequency of organic cow's milk have a 48% probability of purchasing plant-based milk at least once a year compared to only 19% of those frequently buying conventional cow's milk. This substitution pattern is also evident in the annual average quantity of plant-based milk purchased.

3.4 Model of Demand and Supply and the Estimation Approach

This section describes a discrete choice model of cow's milk and plant-based milk demand and lays out the procedures for estimating the demand model. The supply side is modeled as Bertrand-Nash competition among manufacturers following Bresnahan (1987) and Berry, Levinsohn, and Pakes (1995). The equilibrium conditions are derived from the demand and supply models, and the marginal costs are recovered from the equilibrium conditions and the observed prices and quantities.

3.4.1 Discrete Choice Model of Household Demand

The demand for cow's milk and plant-based milk is derived from a standard discrete choice demand model following McFadden and Train (2000), Train (1998), and Revelt and Train (1998). Each consumer faces a set of alternative products and chooses the single alternative that generates the highest utility. Specific products are defined as unique bundles of observed product characteristics. For the milk choice, these characteristics are fat content, lactose-free status, organic status, package size, plant-based status, and brand category.

3.4.1.1 Formal Specification

Consumers are indexed by $i \in \{1, \ldots I\}$, and each faces many choice occasions, indexed by $\tau \in \{1, \ldots T\}$. Each choice occasion refers to a shopping occasion by a consumer to a specific retail store at a location and time, denoted as market $t(\tau)$. Each market is defined as a combination of the geographic Designated Market Area (DMA) and the years from 2018 through 2020. Products are indexed by $j \in \{1, \ldots J_\tau\}$, where J_τ denotes the set of cow's and plant-based milk products available to consumers at choice occasion τ. The indirect utility that consumer i obtains from purchasing product j in the specific retail store at choice occasion τ can be represented by Equation (3.1). Indirect utility is a function of the price of product j at choice occasion τ ($p_{j\tau}$), observed product characteristics for product j at occasion τ ($X^k_{j\tau}$) with k^{th} characteristic denoted by the superscript k, brand dummies indicating product j belongs to brand $b(j)$ among a total of B brands ($\xi_{b(j)}$), and a vector of consumer characteristics ($Z_{i\tau}$). The term, $\lambda\hat{v}_{j\tau}$, is called the control function, which will be discussed in the sub-section on price endogeneity.

$$U_{ij\tau} = -\beta^p_i p_{j\tau} + \sum_{k=1}^{K}\{(\beta^k_i + Z'_{i\tau}\gamma_k)X^k_{j\tau}\} + \lambda\hat{v}_{j\tau} + \xi_{b(j)} + \epsilon_{ij\tau} \tag{3.1}$$

The vector of household characteristics, $Z'_{i\tau}$, include income, age, education, race, Hispanic origin, and household size. Especially, these are defined as follows: household income groups (Low-income: less than \$50,000, Middle-income: between \$50,000 and \$100,000, and High-income: more than \$100,000); the age of household head (Younger-age: under 40 years old, Middle-age: 40-64 years old, and Older-age: 65 years and older); education: college degree dummy; race groups (White, Black, and other race); Hispanic dummy; and household size.

In Equation (3.1), β_i^p denotes household i's marginal utility of income. The term, $(\beta_i^k + Z'_{i\tau}\gamma^k)$ is the taste parameter associated with the k^{th} observed product characteristics for household i, which varies across the demographic group to which household i belongs. These taste parameters vary across households but are the same over choice situations for each household. Early research showed that incorporating consumer-specific heterogeneity through random coefficients was useful for estimating realistic substitution patterns in the framework of logit-type models (Berry, Levinsohn, and Pakes 1995; McFadden and Train 2000; Nevo 2001).

The random coefficients on product characteristics are assumed to follow a joint normal distribution with correlation across coefficients. This assumption reflects the expectation that consumers' valuations of product characteristics, such as organic status and plant-based status, are correlated (Revelt and Train 1998). In the context of cow's milk and plant-based milk demand, this specification helps researchers to learn whether consumers' valuations of product attributes, such as the organic attribute, the lactose-free attribute, and fat content, are positively or negatively correlated with consumers' valuation of the plant-based attribute.

The empirical application excludes shopping occasions during which consumers do not purchase any fluid cow's milk or plant-based milk products unless they buy a designated "outside option" product category that is considered to be related to the milk products. For the econometric estimation, the outside option is the purchase of other dairy products, including yogurt, cottage cheese, or hard cheese, which may be regarded as potential substitutes for fluid milk products. This outside option is denoted by $j = 0$. The utility from selecting the outside option is given by:

$$U_{i0\tau} = \epsilon_{i0\tau} \tag{3.2}$$

where, along with the error terms for inside goods, the vector of error terms $(\epsilon_{i0\tau}, \epsilon_{i1\tau}, ..., \epsilon_{iJ\tau})$ is independently and identically distributed with the Type I extreme value distribution.

3.4.1.2 Estimation of Demand Model

The indirect utility in Equations (3.1) and (3.2) can be decomposed into an observed part to econometrician and an unobserved component as in Equation (3.3).

$$U_{ij\tau} = V_{ij\tau} + \epsilon_{ij\tau} \tag{3.3}$$

where the observed component, $V_{ij\tau}$, is a function of observed characteristics and linear in parameters in Equation (3.4).

$$V_{ijt} = -\beta_i^p p_{j\tau} + \sum_{k=1}^{K} \{(\beta_i^k + Z'_{i\tau}\gamma_k)X_{j\tau}^k\} + \lambda\hat{v}_{j\tau} + \xi_{b(j)} \tag{3.4}$$

Under the assumption that the vector of $(\epsilon_{i0\tau}, \epsilon_{i1\tau}, ..., \epsilon_{iJ\tau})$ is independently and identically distributed Type I extreme value, it is well known that conditional on $\boldsymbol{\beta}_i$=$(\beta_i^p, \beta_i^1, ..., \beta_i^K)$, the choice probability that household i chooses product j at choice occasion τ is the standard logit choice probability shown in Equation (3.5).

$$L_{ij\tau}(\boldsymbol{\beta_i}, \boldsymbol{\gamma}, \lambda, \boldsymbol{\xi}) = \frac{exp(V_{ij\tau})}{1 + \sum_{l=1}^{J_\tau} exp(V_{ilt})} \tag{3.5}$$

where the vector of random coefficients is denoted as $\boldsymbol{\beta}_i$=$(\beta_i^p, \beta_i^1, ..., \beta_i^K)$, demographic interaction as $\boldsymbol{\gamma}$=($\gamma_1, ...\gamma_K$), and the vector of brand dummies as $\boldsymbol{\xi}$=($\xi_1, ..., \xi_B$).

Let $j(i,\tau)$ denote the product that household i chooses at τ. Then, conditional on the household i's taste parameters, $\boldsymbol{\beta_i}$, the probability that household i makes the sequence of choices, $(j(i,1),\ldots,j(i,T))$ across all choice occasions can be represented as in Equation (3.6).

$$L_i(\boldsymbol{\beta_i},\boldsymbol{\gamma},\lambda,\boldsymbol{\xi}) = \prod_{\tau=1}^{T} L_{ij(i,\tau)\tau}(\boldsymbol{\beta_i},\boldsymbol{\gamma},\lambda,\boldsymbol{\xi}) \tag{3.6}$$

The unconditional probability is the integral of the conditional probability over all possible values of $\boldsymbol{\beta_i}$ as in Equation (3.7).

$$P_i(\boldsymbol{\theta},\boldsymbol{\gamma},\lambda,\boldsymbol{\xi}) = \int L_i(\boldsymbol{\beta_i},\boldsymbol{\gamma},\lambda,\boldsymbol{\xi}) f(\boldsymbol{\beta_i}|\boldsymbol{\theta}) d\boldsymbol{\beta_i} \tag{3.7}$$

Here, the function f is the joint distribution of random coefficients, $\boldsymbol{\beta_i}$ where $\boldsymbol{\theta}$ refers to the parameters of joint normal distribution of f, including mean and covariance of $\boldsymbol{\beta_i}$.

Since the unconditional choice probability in Equation (3.7) cannot be solved analytically, it is approximated for a given value of $\boldsymbol{\theta}$ through the simulation method of Train (2009) and Revelt and Train (1998). Specifically, for a given value of $\boldsymbol{\theta}$, a total of R values of $\boldsymbol{\beta_i}$ are randomly drawn. This study sets the number of draws to R=50. For r^{th} draw of $\boldsymbol{\beta_i}$, the probability of a sequence of observed choices in Equation (3.7) is calculated and denoted by $L_i^r(\boldsymbol{\beta_i},\boldsymbol{\gamma},\lambda,\boldsymbol{\xi})$. The process is repeated over a total of R draws. The average choice probability over all draws is the simulated probability that household i makes the sequence of choices.

$$SP_i(\boldsymbol{\theta}, \boldsymbol{\gamma}, \lambda, \boldsymbol{\xi}) = \frac{1}{R} \sum_{r=1}^{R} L_i^r(\boldsymbol{\beta_i}, \boldsymbol{\gamma}, \lambda, \boldsymbol{\xi}) \tag{3.8}$$

The simulated likelihood function is the product of individual simulated choice probability in Equation (3.8). The associated log-likelihood function is then represented in Equation (3.9):

$$SLL(\boldsymbol{\theta}, \boldsymbol{\gamma}, \lambda, \boldsymbol{\xi}) = \sum_{i=1}^{N} lnSP_i(\boldsymbol{\theta}, \boldsymbol{\gamma}, \lambda, \boldsymbol{\xi}) \tag{3.9}$$

3.4.1.3 Price Endogeneity and Identification

The discrete choice model discussed in section 3.4.1.3 (equations (3.1) through (3.9)) can be consistently estimated under the standard independence assumption that any product characteristics not included in the model are not correlated with product characteristics that are included, including price. Of course, researchers always fail to observe some product characteristics, food these may include, for example, the level of fortified micronutrients, brand reputation, and stylishness of packages. The presence of unobserved product characteristics that tend to be correlated with price likely introduces bias into the price coefficient (Petrin and Train 2010).

Two strategies are employed to tackle this endogeneity of price. First, brand-specific fixed effects, $\xi_{b(j)}$, are included in the demand model. Each brand-specific dummy captures the brand-specific unobserved factors, such as package stylishness and reputation, that do not vary across households and accounts for correlations between prices and brand-specific means of unobserved characteristics (Nevo 2000). After controlling for the brand effects, the remaining variations in prices are 1) the variation in relative prices of different products across stores and 2) the price variation within the brand across different package sizes.

Second, many papers in the relevant consumer demand literature employ a control function approach for discrete choice models to further control for the demand shocks that may be correlated with price after conditioning out the brand-specific fixed effects. (For recent examples using this approach, see Dubois, Griffith, and O'Connell 2018; Oh and Vukina 2022.) The control function approach involves two steps following Petrin and Train (2010). In the first step, the endogenous variable (price) is regressed on a vector of observed product characteristics ($X_{j\tau}$), price instruments ($p_{j\tau}^{IV}$), and brand dummies ($\xi_{b(j)}$) that are assumed to be exogenous:

$$p_{j\tau} = X'_{j\tau}\delta + \alpha p_{j\tau}^{IV} + \xi_{b(j)} + \upsilon_{j\tau} \tag{3.10}$$

The residual term, $\upsilon_{j\tau}$, follows from the first-stage regression in Equation (3.10) and enters the demand model in Equation (3.1). The inclusion of this additional residual term serves to control for the source of dependence between prices and demand error terms. As with other applications, this chapter uses a Hausman-Nevo- price instrument, which is constructed as the average price of product j at nearby times and places (Hausman 1996; Nevo 2001). The application of the Hausman-Nevo instrument is especially pertinent in the context of the cow's milk market, as neighboring geographical markets delineated by Designated Market Area (DMA)[2] exhibit comparable fluid use farm milk prices that are regulated by the USDA.

3.4.2 Supply Model

The supply side of the market follows the standard approach of Bresnahan (1987) and Berry, Levinsohn, and Pakes (1995) that is used in much of the retail demand literature (for example, Dubois, Griffith, and O'Connell 2020). Here, a market is defined as the combination of

[2]DMA is commonly used as a geographic market designation as in, for example, Gillingham, Houde, and Van Benthem (2021), Hart and Alston (2020), Hitsch, Hortacsu, and Lin (2021)

a year and a DMA. To develop the intuition of the supply model, suppose that the collection of cow's milk processors and plant-based milk manufacturers compete in each market by simultaneously setting prices in a Bertrand-Nash game. Privately labeled (store-brand) products have a large (80%) quantity share in the combined fluid milk market. Retailers (chains) selling their privately labeled products are affiliated with processors, assuming they act as a vertically integrated firm competing with national and local brand milk processors and plant-based milk manufacturers.

Allowing market power among cow's milk processors and plant-based milk manufacturers in selling products to competitive retailers is the standard approach to wholesale market behaviors in this literature. This may be appropriate when privately labeled brands, which are owned by retailers, account for a dominant share of the milk market, even though store brands may be processed by the same company that supplies the national or local brands. For national brand suppliers, however, it is reasonable that there is a relationship between retail price and wholesale price. One common approach is to assume that retailers charge a constant percentage markup over the wholesale price set by manufacturers, as used by Hausman and Leonard (2002) for the bath tissue industry. That is the approach used in this research.[3]

Formally, markets are indexed by m, and a manufacturer by $f \in \{1, \ldots, F\}$. The set of products sold by firm f in market m is denoted as J_{fm}. According to the standard supply-side model of Bertrand-Nash competition in the retail demand literature, the firm f in market m sets the prices of products in J_{fm} and takes the prices of other products in the market m as given. The firm f's profit can be written as:

[3]The constant percentage markup is a strong and limiting assumption. In further work, I plan to also consider an alternative supply-side model where retailers may have market power while assuming wholesale prices of cow's milk are equal to wholesaler marginal costs, following, for example, scenario 3.1 in Villas-Boas (2007).

$$\Pi_{fm} = M_m \times \left\{ \sum_{j \in J_{fm}} (p_{jm} - mc_{jm}) \times s_{jm}(\boldsymbol{p_m}) - C_{fm} \right\} \tag{3.11}$$

where M_m is the size of market m, p_{jm} is the price of product j in market m, mc_{jm} is the marginal cost of product j in market m, and C_{fm} is a fixed cost of production by firm f in market m. The term, p_m denotes the vector of prices of all products in market m, and the term, s_{jm} is the quantity share of product j in market m, which is a function of the prices of all products in market m.

Note that the Nielsen data report retail quantities and prices, whereas, in the supply model, the relevant prices are in the wholesale market and wholesale price are not observed. The wholesaler marginal costs in the model are the retail price minus the implied wholesale to retail markup, which depends in part on market power of the processors and manufacturers.

3.4.3 Equilibrium Condition and Implied Marginal Costs

The supply-side modeling and demand specification enable us to describe the market equilibrium condition. The firm's first-order conditions are obtained by Equation (3.12).

$$s_{jm}(\boldsymbol{p_m}) + \sum_{k \in J_{fm}} (p_{km} - mc_{km}) \frac{\partial s_{km}(\boldsymbol{p_m})}{\partial p_{jm}} = 0 \qquad \forall j \in J_{fm} \tag{3.12}$$

Under the maintained assumption of Bertrand competition, the marginal cost can be recovered using the observed market prices and the estimated demand parameters. Specifically, the marginal cost of products in market m can be solved using Equation (3.13):

$$\boldsymbol{mc_m} = \boldsymbol{p_m} + \Omega^{-1} \boldsymbol{s_m}(\boldsymbol{p_m}) \tag{3.13}$$

where mc_m is a vector of marginal costs of all products in market m, s_m is a vector of quantity share functions of all prices in market m. The matrix, Ω_m contains own- and cross-price share derivatives that have elements $I_m\ (j,k)(\partial s_{km}/\partial p_{jm})$ where $I_m(j,k)$ equals one if product k and j are owned by the same firm and zero otherwise.[4]

3.5 Dataset for Estimation of the Demand Model

Retail stores from the NielsenIQ Retail scanner data (RMS) and households from the NielsenIQ Homescan Panel (HMS), for the years 2018–2020, are used to estimate the demand model.

3.5.1 Sample of Stores, Households, and Purchases

In total across all states, the NielsenIQ RMS contains weekly sales revenue and quantity sold for each UPC at about 40,000 affiliated stores. This dataset covers a range of store formats, including grocery stores [5] (e.g., Safeway), mass merchandisers (e.g., Walmart), drug stores (e.g., CVS), and convenience stores. Furthermore, NielsenIQ RMS contains detailed information on product characteristics of milk, such as fat content, the organic attribute, package size, and brand name.

In total across all states, NielsenIQ HMS captures data regarding the grocery shopping of about 60,000 nationally representative households each year. Panelists in HMS maintain records of their grocery purchases. These panelists are provided with an electronic scanner with which they report the specific grocery store they visited, the purchase date, the quantity of each UPC item, and the price paid for the UPC item. On an annual basis, HMS households report their demographic information, encompassing details such as geographic location (5-digit zip code), household income (in 16 bins), age, race/ethnicity, and educational attainment of household heads, and the number of household members.

For the estimation of the discrete choice demand model, the term "shopping occasion"

[4]Further work will calculate the distribution of the implied marginal costs of cow's milk processing (and markup) consistent with the estimated parameters.

[5]These retail chains are categorized as "Food" in NielsenIQ RMS. Details on purchases of cow's milk and plant-based milk are provided by retail chain type in Appendix 3.B.

is used in a distinct and specialized way. The shopping occasion is defined as the purchase of a single package of cow's milk, plant-based milk, or the outside option (here, other dairy products such as yogurt, cheese, and butter) from a specific grocery store. Some milk purchases involve either multiple units of a single product or multiple products. These purchases are split and treated as separate "shopping occasions." For example, if a household purchases one unit of cow's milk and one unit of plant-based milk, these purchases are considered two separate "shopping occasions." In the dataset for demand estimation below, the Nielsen HMS households choose one unit of a single product in 78% of the visits to retail stores. In 12% of visits, households purchase multiple units of the same product, while in 10% of the visits, they buy more than one individual distinct product. Appendix Table 3.B.4 further illustrates the multiple products purchased during store visits.

The sample of retail stores and households used for econometric estimation is confined to the eight states with the most HMS shopping occasions to reduce the computational burden in the estimation. These eight states – California, Texas, Ohio, Pennsylvania, New York, Michigan, Illinois, and North Carolina – represent several regions in the country. Only purchases made by HMS households in these selected states purchased from RMS retail stores in those states are included. As shown in Appendix 3.B.1, the eight-state sample is generally representative of all choice occasions across all states, with households in this sample purchasing slightly more organic cow's milk, lactose-free cow's milk, and plant-based milk. Further explanations on the rationale for selecting these eight states are provided in Appendix 3.B.1, with comparisons of price and product characteristics between the dataset for demand estimation and all choice occasions across all states.

3.5.2 Construction of the Choice Set

The NielsenIQ HMS does not include data regarding alternative products that panelists did not purchase. This limitation means that the HMS does not contain information, including price, about relevant products that are available in the markets frequented by the panelist but that are not chosen during the sample period.

Economists using household purchase data have commonly imputed product market prices of a range of alternatives by averaging the prices of each item for all transactions, across many households (for example, Choi, Wohlgenant, and Zheng (2013) and Dubois, Griffith, and O'Connell (2020) use this procedure). However, recent research by Blundell, Horowitz, and Parey (2022) has shown that substituting imputed prices for actual transaction prices can introduce bias in demand estimation. Some recent studies avoid this imputation concern by matching the household purchase data with retail store data based on information about when and in which store the transaction occurred (Oh and Vukina 2022; Joo 2023).

Following this matching approach, my data procedure recovers an appropriate household choice set for each household on each occasion by matching HMS shopping occasions that included any cow's milk and plant-based milk purchases with the RMS weekly store data. NielsenIQ RMS includes the item chosen by a panelist and alternatives that were available at the choice occasion but not chosen by that panelist. Because the HMS shopping occasions are dated specifically, while RMS contains weekly sales revenue and quantity sold of products, the recorded prices of alternative products available on a particular date are calculated as the weekly average prices for the week in which the shopping occasion took place.

The matching and cleaning of data proceeds as follows. First, panelists' shopping occasions in HMS are matched with stores in the RMS. NielsenIQ panelists visit grocery stores affiliated with NielsenIQ (RMS stores) as well as other stores not covered by RMS. Since the price and product information are available only for products sold in RMS stores, HMS shopping occasions that occur in RMS stores are preserved.[6] Second, attention is confined to grocery stores among various types of retailers. Grocery stores ("Food" chain in Nielsen categorization) account for more than 88% of the total milk quantity sold from 2018 to 2020. Another 7% is sold in "Mass-merchandisers," which tend to have offerings similar to the

[6]Unfortunately, excluding these occasions leads to the omission of a few regions, mainly Florida, where Nielsen households predominantly shopped at non-RMS stores, as detailed in Appendix 3.B.1. Nonetheless the benefit of having household purchases matched with RMS stores is worth the cost of losing data from certain regions.

grocery stores. Drug stores account for 3% and convenience stores for 2% (Table 3.B.1 in Appendix 3.B). Data are not used from mass-merchandisers, convenience stores, and drug stores, because these account for a small share of milk sales and often do not carry some product categories, such as organic cow's milk.

Third, in order to eliminate many products that may have specialized demand attributes, the sample includes only quart, half-gallon, and one-gallon package sizes and larger. These package sizes account for more than 99 percent of the quantity sold in retail stores. Fourth, in order to reduce computational burdens, ten shopping occasions per sample household during the 2018–2020 sample period were randomly chosen for inclusion in the dataset. Finally, observations of milk purchases are eliminated as outliers if the transacted price per gallon is lower than 50 cents or higher than $40. Further details on the data-matching strategy and sample store construction process are provided in Appendix 3.B.

Recall that the demand model, as described in Section 3.4, specified an outside option as shopping occasions in which households do not purchase any cow's milk or plant-based milk. Here, I define the outside option empirically as shopping occasions in which the household buys other dairy products but not one of the fluid milk products. Moreover, such shopping occasions are included only if they occur 14 days or more before a shopping occasion in which any cow's milk or plant-based milk products were purchased.[7]

3.5.3 Definition of Products for the Demand Specification

In the final dataset, there are 1,371 UPCs for cow's milk and 530 UPCs for plant-based milk. These are far too detailed and numerous for useful empirical demand analysis. To make the empirical analysis feasible, individual UPCs are aggregated into a single product if they share the same observed product characteristics and are labeled with the same brand name. For example, almond milk UPCs that are labeled as Almond Breeze with a half-gallon

[7]Households that have recently purchased cow's milk, plant-based milk, or other dairy products might not make additional purchases of these products in subsequent shopping trips within a week or two. Considering that the shelf life of fluid milk is typically around two weeks (Barbano, Ma, and Santos 2006), shopping occasions where no milk purchases are made within two weeks after the most recent purchase are not considered outside shopping occasions.

Table 3.3. Number of products, brands, and choices in HMS shopping occasions

	Conventional cow's milk	Organic cow's milk	Lactose-free cow's milk	Plant-based milk	Total
Number of products	128	72	56	102	358
Number of brands	15	8	10	12	31
Number of shopping occasions	95,349	6,715	8,753	19,420	200,935

Note: The numbers are based on the final data set for demand estimation. The column sum of brands may not necessarily equal the total number of brands in the last column because some brands offer multiple products across different categories. Among 200,935 shopping occasions, either cow's milk or plant-based milk was chosen on 130,237 occasions, while neither was chosen on 70,698 occasions.

size, non-organic, and original flavor are aggregated into one product. Furthermore, there are many niche brands that appear in a few markets but not others. These minor brands are aggregated into a single composite "local" brand. Products within this composite local brand are defined based on their observed characteristics. For example, any conventional cow's milk products in a half-gallon container from various niche brands are grouped together as an aggregated product. A similar aggregation applies to organic cow's milk, lactose-free cow's milk, and plant-based milk.

Table 3.1 describes the cow's milk and plant-based milk products in the cleaned dataset for NielsenIQ HMS data. There are a total of 358 products, including 128 conventional cow's milk products, 72 organic cow's milk products, 56 lactose-free cow's milk, and 102 plant-based milk products, all of which are lactose-free and some of which are organic. A total of 31 "brands" are observed. Any private-label store brand is considered a single brand, no matter which retailer-brand is on the product. The composite of local brands is also treated as a single brand.

3.5.4 Summary of Product Characteristics

The final dataset contains information about the behavior of 24,540 household panelists during the years 2018 to 2020. The cleaning and filtering process retains a total of 200,935 shopping occasions, including 130,237 shopping occasions in which panelists purchased cow's milk or plant-based milk and 70,698 shopping occasions in which they purchased neither. Panelists had an average of 69 milk products, as defined in Section 3.5.3, during each shopping occasion. With a total of 200,935 shopping occasions, this results in a total of 13,790,537 observations of milk products across all shopping occasions.[8]

Table 3.4 describes the product characteristics of cow's milk and plant-based milk chosen by NielsenIQ HMS households. Cow's milk was purchased on about 85% of shopping occasions, with conventional cow's milk purchased on 73% of occasions. Plant-based milk products were chosen for 15% of the shopping occasions[9], and the average price paid was \$6.51 per gallon. Plant-based milk products are about twice as expensive as cow's milk products. However, the average price paid for organic and lactose-free cow's milk was even higher than that for plant-based milk. Cow's milk and plant-based milk are predominantly purchased in half-gallon and one-gallon containers, with nearly equal distribution. Although the one-gallon container comprises 66% of conventional cow's milk purchases, the half-gallon container is more prevalent in the case of organic cow's milk (82%), lactose-free cow's milk (80%), and plant-based milk (83%). Recall that the demand estimation accounts for the price differences by the package size by including dummy variables for package sizes.

[8]These 13 million observations are used to calculate the choice probabilities in Equations (3.5) through (3.7) in demand estimation using the simulated maximum likelihood function. Despite major restrictions on the sample to reduce the computational burden—limiting it to eight states with the most shopping occasions and randomly selecting choice occasions up to ten per each panelist—the optimization process in MATLAB still required several tens of hours.

[9]Recall that the term, "shopping occasion" is used here as defined in Section 3.5.1 Purchases of multiple units or multiple products are separated and treated as separate "shopping occasions." Appendix 3.B presents further details on such shopping occasions.

Table 3.4. Descriptive statistics of product characteristics for chosen products

Price and product characteristics	Mean	SD	Min	Max
Price of conventional cow's milk ($/gallon)	3.06	1.48	0.56	17.96
Price of organic cow's milk ($/gallon)	7.38	1.98	2.14	19.92
Price of lactose-free cow's milk ($/gallon)	8.01	1.55	1.07	15.06
Price of plant-based ($/gallon)	6.51	2.32	1.40	31.95
Plant-based (0/1)	0.149	0.356	0	1
Organic (0/1)	0.070	0.255	0	1
Lactose-free-cow (0/1)	0.067	0.250	0	1
Fat contents	1.771	0.961	0	3
Package size (quart) (0/1)	0.070	0.255	0	1
Package size (half-gallon) (0/1)	0.453	0.498	0	1
Package size (one-gallon) (0/1)	0.477	0.499	0	1

Note: The numbers are based on the final data set for demand estimation. Organic attributes applied to both cow's milk and plant-based milk. Prices are calculated using a volume-weighted average across all package sizes. The maximum prices observed in four product categories correspond to products sold in a quart container. For example, a local brand's conventional cow's milk, packaged in a quart container, was available in Los Angeles in 2020 at $4.49 per product. This is equivalent to $17.96 per gallon.

3.6 Demand Estimation Results and Substitution Patterns

Table 3.5 summarizes the distribution of estimated random coefficients of price and product characteristics in the indirect utility function of Equation (3.1). The estimation result is obtained by the simulated maximum likelihood estimation based on 24,540 panelists and 200,935 choice occasions, controlling for 31 brand fixed effects.

Table 3.5. Demand estimation result

Variable	**Moment**	**Estimates**	**Standard Error**
Panel A: Moments of distribution of consumer-specific preference parameters			
Price	Mean	-0.745	0.003
	Standard Deviation	0.259	0.001
Plant-based	Mean	1.303	0.078
	Standard Deviation	3.402	0.425
Lactose-free	Mean	-2.232	0.053
	Standard Deviation	3.281	0.103
Fat content	Mean	0.238	0.013
	Standard Deviation	0.670	0.014
Organic	Mean	-2.500	0.086
	Standard Deviation	2.613	0.204
Package size (quart)	Mean	1.256	0.047
	Standard Deviation	1.999	0.124
Package size (half-gallon)	Mean	0.821	0.023
	Standard Deviation	1.571	0.041
Residual from the first-stage regression for price ($\hat{v}_{j\tau}$)		0.232	0.005
Panel B: Demographic interaction parameters			
Plant-based			
X Low-income		-0.017	0.036
X High-income		-0.016	0.036
X Younger-age		0.353	0.037
X Older-age		-0.455	0.036
X College		0.285	0.047
X Black		0.800	0.050
X Other race		0.283	0.045
X Hispanic		0.149	0.053

Table 3.5. Demand estimation result (Continued)

Variable	Moment	Estimates	Standard Error
Panel B: Demographic interaction parameter (Continued)			
Organic attribute			
X Low-income		-0.276	0.051
X High-income		0.410	0.048
X Younger-age		0.465	0.052
X Older-age		-0.288	0.052
X College		0.781	0.067
X Black		-0.227	0.074
X Other race		0.969	0.060
X Hispanic		-0.097	0.070
Fat content			
X Low-income		0.078	0.009
X High-income		-0.069	0.010
X Younger-age		0.102	0.010
X Older-age		-0.036	0.009
X College		-0.164	0.011
X Black		0.037	0.014
X Other race		0.124	0.013
X Hispanic		0.039	0.015
Package size (quart)			
X Household size		-0.682	0.017
Package size (half-gallon)			
X Household size		-0.280	0.008
Brand fixed effects		Yes	
Number of Households		24,540	
Number of choice occasions		200,935	

3.6.1 Product Characteristics Coefficient Estimates

Panel A of Table 3.5 reports the means and standard deviations of random coefficients. Given the large sample size, these estimates are highly precise, as shown by the very small size of the estimated standard error relative to the magnitude of the coefficients.

The means of random coefficients show the effect relative to the reference group for incomes and each demographic variable. The reference group is defined as households with "middle-income" (between \$50,000 and \$100,000 per year), "middle-age" (40–64 years old respondent), a non-college degree respondent, and a non-Hispanic White respondent.

The mean coefficient on the price is estimated to be negative as expected (-0.745), indicating that as the price of a product increases, the corresponding net utility derived from the purchase of that product decreases.

The mean coefficient of the plant-based dummy is estimated to be 1.303. Recall that all plant-based milk products are lactose-free, and some cow's milk is also lactose-free. The positive mean coefficient on the plant-based dummy implies that households, on average, prefer plant-based products to lactose-free cow's milk products. However, plant-based products are less preferred, on average, than conventional cow's milk products, as the sum of the mean coefficient on the plant-based attribute and lactose-free attribute is negative. Both the mean coefficients on lactose-free and organic attributes are estimated to be negative, reflecting a small market share relative to conventional cow's milk products. Panel A also displays parameter estimates for the other product attributes, and each shows the impact on average. The fat content has a positive impact, the organic attribute has a negative impact, and the smaller package sizes have positive attributes.

The estimated standard deviation in product characteristics reveals substantial heterogeneity among households. Notably, the standard deviation of price coefficients is relatively modest, whereas the standard deviations of product characteristics exceed the means of coefficients in absolute terms. This means, for example, that while fat content has a positive impact on average, many households value fat negatively. Although the organic attribute is negative on average, reflected in the small market share, some households gain from buying organic products. The coefficient on the residual from the first-stage regression for price is estimated to be positive (0.232) and statistically significant. The positive sign is expected because unobserved characteristics—that make the observed price of a product higher than the predicted price explained by observed characteristics—are desirable. Table **??** in Appendix 3.A reports the first-stage regression results of the observed prices on product characteristics, brand fixed effects, and the price instrument. The coefficient on the price instrument is statistically significant, affirming a strong correlation between the price instrument and the potentially endogenous price variable. All other coefficients on explanatory variables, except for plant-based attributes and fat content, are estimated to be significant.

3.6.2 Demographic Interactions

Panel B of Table 3.5 reports the estimated coefficients on demographic interactions with product characteristics. Most of the coefficients on demographic interaction are statistically significant, with the exception of income dummies interacting with the plant-based dummy. The plant-based attribute is preferred by younger and more educated (college degree) households, which is in line with the household consumption pattern presented earlier in Table 3.1. The plant-based attribute is preferred more by Black and Hispanic households compared to White households.

The demographic group that is younger and has a college degree exhibits a higher preference for the organic attribute, which is consistent with previous studies (Alviola and Capps 2010; Choi, Wohlgenant, and Zheng 2013). However, the organic attribute is less preferred by Black and Hispanic households than it is by White households. Finally, households with more family members prefer the one-gallon container over half-gallon or quart containers.

3.6.3 Correlation among Product Characteristics

As stated in the model description in Section 3.4, the random coefficients are assumed to follow a joint normal distribution with correlations across coefficients. The estimated correlations among these estimated random coefficients (with means and standard deviations in Table 3.5) are reported in Table 3.6. The estimated coefficient on the plant-based dummy is positively correlated with the coefficient on the lactose-free dummy for cow's milk.[10] (Recall that all plant-based products are lactose-free). This positive correlation is expected because plant-based milk and lactose-free cow's milk are considered substitutes for lactose-intolerant consumers. The plant-based coefficient is also positively correlated with the organic product characteristic. Households that put a higher value on the organic characteristic and the lactose-free attribute tend to have a stronger preference for plant-based products. This is consistent with the findings in Section 3.3 that organic cow's milk consumers are more likely to buy plant-based milk. Similarly, the correlation between the coefficient on the lactose-free attribute and the organic attribute is estimated to be positive.

Finally, the coefficients for plant-based, lactose-free, and organic attributes are all negatively correlated with the coefficient for fat content. This means that households with a low preference for fat content have a higher preference for plant-based, lactose-free, and organic attributes, which are positively correlated with each other. These negative correlations align with previous findings that consumers who drink both cow's milk and plant-based milk are more conscious about fat contents and calories (McCarthy et al. 2017).

3.6.4 Price Elasticities

The econometric procedures employed above do not generate price elasticities of demand directly for detailed individual products. The own- and cross-price elasticities are calculated from several parameters estimated econometrically. They measure the degree to which households change purchases of each aggregated cow's milk category of products and plant-based milk category of products in response to changes in relative prices of product categories. The calculated elasticities apply to implied responses of the purchases of the product groups

[10]The taste parameter associated with plant-based milk products is the sum of the coefficient on the plant-based dummy (β_{pb}) and the coefficient on the lactose-free dummy (β_{lcf}). The taste parameter associated with lactose-free cow's milk is β_{lcf}. Therefore, the correlation in taste parameters between the plant-based dummy and lactose-free dummy is calculated as $\text{corr}(\beta_{pb} + \beta_{lcf}, \beta_{lcf})$.

Table 3.6. Correlation matrix between product characteristics

	Plant-based	Lactose-free (cow)	Fat content	Organic
Plant-based	1	.	.	.
Lactose-free (cow)	0.281 (0.030)	1	.	.
Fat content	-0.586 (0.027)	-0.332 (0.008)	1	.
Organic	0.050 (0.027)	0.059 (0.009)	-0.135 (0.016)	1

Note: The correlations between product characteristics are calculated from the variance-covariance matrix of random coefficients. Standard errors for correlations are shown in the parenthesis and calculated using the delta method.

in response to changes in the average prices of the groups, holding constant relative prices within aggregated product groups.

Econometric studies of demand estimation with many differentiated products are used to generate product-level elasticities when those are useful for economic issues considered (For an early, well-known example of breakfast cereal, see Nevo (2000)). While individual specific product-level elasticities may be useful for some questions, they may not provide the measurements of the substitutability that are most useful in considering broader product groups. Therefore, this research calculates product group-level elasticities (and confidence intervals) using a procedure based on Dubois, Griffith, and O'Connell (2020). Specifically, for each shopping occasion, the change in quantities for four product groups is calculated in response to a simulated 1 percent change in the prices of all products within each product group. The predicted change in quantities for these four products is then averaged across all shopping occasions.

Table 3.7 reports a full set of product group-level elasticities along with confidence intervals.[11] Note that, given the large sample size, all the confidence intervals are very

[11]Confidence intervals on elasticities are constructed following Krinsky and Robb (1986). The covariance

narrow. In Table 3.7, the own-price elasticity for conventional cow's milk is the smallest among the four product groups. This is the expected result because conventional cow's milk has a large market share, and the prices of conventional products are much lower than the prices of the other milk products (as shown in Table 3.8). Since prices of other product groups are more than twice as expensive as conventional cow's milk, most consumers do not switch to or from conventional cow's milk to other product groups in response to observed variations in the price of conventional milk.[12] The demand functions for organic milk, lactose-free milk, and plant-based milk are all highly elastic (between -2.19 and -2.72). Evidently, consumers are willing to adjust buying behavior substantially in response to the prices of these products.

Table 3.7 reports the retail demand elasticities. The own-price elasticity for farm milk used for fluid products tends to be smaller in magnitude. For example, Zhang and Alston (2018) estimated the farm milk elasticity to be between -0.16 and -0.39. Assuming a fixed proportion production technology in processing fluid milk products using farm milk as an input, and under certain assumptions, the derived demand elasticity for farm milk used for fluid products can be approximated by multiplying the cost share of farm milk in fluid milk processing times the elasticity of retail demand for fluid products (Bronfenbrenner 1961; George and King 1971; Gardner 1975).

The share of the cost of farm milk in the retail price of conventional cow's milk in regions

matrix for demand parameters is obtained based on the asymptotic distribution. A total of 200 draws are taken from the jointly normal asymptotic distribution of parameters. The category-level elasticities are calculated for each draw. The lower and upper limits of a 95% confidence interval are given by the 6th and 195th sorted estimates of the elasticity.

[12]The retail own-price elasticities calculated in this research are comparable to previous demand studies using store- or household-scanner data (Gould 1996; Dhar and Foltz 2005; Alviola and Capps 2010; Chouinard et al. 2010; Choi, Wohlgenant, and Zheng 2013; Dharmasena and Capps 2014). The own price elasticity of conventional cow's milk (-0.84) falls within the range observed in previous studies, which varied from -0.51 to -1.03. For organic cow's milk, the own-price elasticities in previous studies ranged from -1.37 to -2.05. The own price elasticity for organic cow's milk presented in this research is slightly higher (-2.72). This can be attributed, in part, to inclusion of plant-based milk as a viable substitute for cow's milk products in the econometric specification. It is important to note that during the 2000s, when most earlier studies were conducted or relied on scanner data, plant-based milk had a smaller market share than in the 2018 to 2020 period.

Table 3.7. Product group elasticities

	Conventional cow	Organic cow	Lactose-free cow	Plant-based
Conventional cow	-0.841 [-0.851, -0.832]	0.126 [0.121, 0.130]	0.165 [0.160, 0.170]	0.258 [0.254, 0.264]
Organic cow	0.694 [0.677, 0.710]	-2.715 [-2.757, -2.675]	0.177 [0.162, 0.190]	0.424 [0.404, 0.439]
Lactose-free cow	0.510 [0.494, 0.527]	0.095 [0.089, 0.102]	-2.364 [-2.411, -2.311]	0.313 [0.293, 0.331]
Plant-based	0.563 [0.551, 0.573]	0.157 [0.151, 0.164]	0.227 [0.215, 0.242]	-2.188 [-2.216, -2.155]

Source: Author calculations based on estimated demand parameters.
Note: The table reports the percentage change in the quantity of each product group (row) with respect to a 1 percent increase in price for the alternative product group (column). The 95% confidence intervals based on Monte Carlo simulation are shown in brackets.

covered by the data for demand estimation is, on average, 48.1% from 2018 to 2020.[13] Based on this farm-to-retail price spread, the retail own-price elasticity of conventional cow's milk, estimated at -0.84, translates to a farm milk elasticity of -0.40 for fluid milk products. The implied farm milk elasticity from the retail estimate is very close to the upper range of estimates (-0.39) suggested by Zhang and Alston (2018).

Cross-price elasticities allow us to analyze the specific substitution patterns between the categories of cow's milk and plant-based milk products. Several comments are useful. First, households switch more readily from organic cow's milk to plant-based milk and from lactose-free cow's milk to plant-based milk than to conventional cow's milk. In the second and third columns, the cross-price elasticities are larger for plant-based milk. Second, plant-

[13]Farm milk prices are calculated using the "Class I minimum price" published each month by USDA AMS (2024) for each regional milk marketing order. Retail and farm prices are weighted by the quantity sold in RMS stores. Farm-to-retail price spread was 46.0% in 2018, 50.6% in 2019, and 47.8% in 2020. Across all years from 2018 to 2020, the spread based on weighted prices was 48.1%. Similar farm-to-retail spreads can be found in USDA ERS (2024) for "one-gallon" conventional milk, which are larger than those for all package sizes because one-gallon is the least expensive package size.

Table 3.8. Prices and volume shares of product categories, averages across all households for all shopping occasions from 2018 through 2020

	Conventional cow's milk	Organic cow's milk	Lactose-free cow's milk	Plant-based milk
Price ($/gallon)	3.06	7.38	8.01	6.51
Volume share	79.7%	4.4%	5.2%	10.6%

Source: Author calculations based on the final dataset for demand estimation.
Note: Prices are calculated using a volume-weighted average of nominal prices.

based milk consumers more readily switch to organic cow's milk than they do to other product categories. In the last column, the cross-price elasticity of organic cow's milk with respect to plant-based milk is 0.42, which is greater than for lactose-free cow's milk (0.31) and conventional cow's milk (0.26). Table 3.A.2 in the Appendix shows similar substitution patterns when plant-based milk products are categorized into three sub-groups: soy milk, almond milk, and other plant-based milk.

The estimated elasticities are best interpreted as representing consumer behavior in the very short run, such as during a particular shopping trip. The data are generated by observing purchases on occasions with different prices. The demand model does not estimate gradual changes in consumption or purchases in response to higher average or sustained prices. The empirical demand model does not have data to link purchases over time to differentiate between intertemporal substitution, such as households postponing their milk purchase until their next shopping occasion due to high prices or purchasing more while observing lower prices only to reduce purchases subsequently. It is well documented in the literature that such dynamic demand behavior causes more elastic price behavior. It would be a mistake to interpret the relatively high elasticities as only reflecting substitution among the products without considering substitution across time. (Erdem, Imai, and Keane 2003; Hendel and Nevo 2006). Nonetheless, it is also worth noting that substitution over time may have smaller impacts on refrigerated products than other grocery products, such as laundry detergent and soft drinks, which were considered by Hendel and Nevo (2006).

3.7 Impacts of the Availability of Plant-Based Milk on Cow's Milk

This section assesses the impact of the availability of plant-based milk on the retail equilibrium price and quantity of cow's milk. This evaluation is conducted through a counterfactual simulation under which plant-based milk products are "removed" from the consumer's choice set.

The estimated demand parameters, along with the supply side of the Bertrand-Nash model, enable us to obtain the product price responses to changes of choice sets under the counterfactual scenarios. Equation (3.12) in Section 3.4 shows how the equilibrium prices are determined by the share function through demand parameters and the marginal costs of each competing product. Without data on firms' marginal costs or estimation of marginal cost functions, the estimated marginal cost dollars per unit of each product is recovered from the equilibrium condition in the Bertrand-Nash model using the observed prices and the relationship between market price and the amount of marginal cost.

Using this model, which includes a fixed percentage markup and no shift in farm marginal costs, the counterfactual scenario isolates the equilibrium price impacts caused by a hypothetical removal of plant-based milk options from the consumer choice set. Note that the same processors rarely produce cow's milk and plant-based milk products, and the two products are not produced with the same raw materials or processes, except packaging. Under these modeling assumptions and using the notation from Section 3.4, I solve for the equilibrium prices using Equation (3.14), which includes quantity shares as a function of marginal costs.

In Equation (3.14), p_m^{cf1} is a vector of counterfactual prices, and J_{fm}^{cf1} is a set of products produced by firm f in market m under the first counterfactual scenario. The counterfactual product share of each product in the counterfactual choice set is calculated using the share function in Equation (3.5) in Section 3.4.

Table 3.9 contains the results of the counterfactual scenario for the removal of plant-based milk products, which shifts the demand for cow's milk products. On average across markets, the removal of plant-based milk raises the price of conventional cow's milk by around 1.3% while raising the equilibrium quantity of conventional cow's milk by about 11% from the quantity to which it had declined. Of course, most of the increase in cow's milk overall is attributed to conventional cow's milk since it accounts for the bulk of cow's milk affected. Recall that per capita consumption of cow's milk was 16.3 gallons in 2020. Based on the quantity share of the four product categories, per capita consumption of conventional cow's milk is 13.6 gallons. The removal of plant-based milk leads to conventional cow's milk by 1.46 gallons per capita per year. However, the relative quantity effects are larger for organic cow's milk (23%) and lactose-free milk (16%) than for conventional cow's milk (11%). The result of the changes accords with the substitution patterns described in Section 3.3, where organic cow's milk and lactose-free milk consumers are more likely to buy plant-based milk.

The removal of plant-based milk reduced the combined total amount of all "milk" sold. If plant-based milk products were unavailable (counterfactual), 3.25% of the previously purchased combined products would not be sold.

The counterfactual results indicated that annual per capita fluid cow's milk consumption would increase by 2 gallons, from 16.3 gallons to 18.3 gallons in 2020, in the absence of availability of plant-based milk.[14] Annual per capita consumption of cow's milk fell by 5.14 gallons, from 21.5 gallons in 2006 to 16.3 gallons in 2020. Hence, the increase in cow's milk consumption from removing plant-based milk is equivalent to 38.4% of the decrease from 2006 to 2020. The estimated magnitude of the effect based on the structural estimates in this study is somewhat larger than correlational results from previous studies, such as 20% by Stewart et al. (2020) and 5% by Slade (2023).[15]

[14]From Table 3.9, the counterfactual quantity share of all cow's milk groups would be increased by 12.1% (=96.75/86.30-1). The counterfactual per capita annual consumption of fluid milk would then be 18.31 gallons (=16.33 $\times$ 112.1%) compared to the actual consumption of 16.33 gallons in 2020.

[15]Comparing the weekly household consumption of cow's milk and plant-based milk from 2013 to 2017, Stewart et al. (2020) illustrate that the increase in plant-based milk purchase accounts for 20% of the decrease in cow's milk purchase. Slade (2023) estimates the correlational relationship between the change

Table 3.9. Price and quantity effects of removal of plant-based milk availability

	Price effect		Quantity effect		
	Initial price ($/gal)	Percentage change in price	Initial Initial share (%)	Percentage change in quantity	Implied changes in annual per capita consumption (gallons)
Conventional cow's milk	3.20	1.32%	71.81	10.77%	1.46
Organic cow's milk	7.30	1.14%	5.99	22.60%	0.26
Lactose-free cow's milk	8.10	0.97%	8.50	16.01%	0.26
Plant-based milk	6.68	-	13.70	-100.00%	-2.59
Total	-	-	100.00	-3.25%	-0.61

Source: Author's calculation based on the counterfactual simulation of the removal of plant-based milk products.

Note: Prices are volume-average across 87 markets defined by the combination of quarter and DMA regions. Quantities are aggregated across all 87 markets. The last column calculates the implied changes in per capita annual consumption of each product category based on the 2020 annual per capita consumption of cow's milk (16.3 gallons).

3.8 Summary and Concluding Remarks

This chapter examined the implications of the rise of plant-based milk products which have been designed to substitute for cow's milk. Although conventional cow's milk continues to make up the lion's share of consumption, the rapid expansion of plant-based milk products places additional pressure on the consumption of cow's milk.

Demand estimation results indicate that plant-based alternatives are substitutes, especially for organic cow's milk and lactose-free cow's milk. Counterfactual scenarios using the econometrically estimated parameters show that removing plant-based milk availability

in cow's milk quantity bought and the change in plant-based milk quantity bought and shows that a 1-gallon decline in plant-based milk consumption is associated with a 0.4-0.6 gallon decrease in cow's milk consumption, implying only a small portion (5%) of the decline in cow's milk consumption is associated with the increased consumption of plant-based milk.

would result in a larger increase in the quantity purchased for organic cow's milk (23%) and lactose-free cow's milk (16%) compared to conventional cow's milk (11%). Overall, cow's milk consumption would increase by 2 gallons per capita in the absence of plant-based milk, 38% of the decline in consumption from 2006 to 2020.

In conclusion, the decline in per capita consumption of cow's milk is not primarily due to the rise of plant-based milk. Although organic cow's milk and lactose-free cow's milk have been more affected by the availability of plant-based milk, the impact on conventional cow's milk has been limited.

References

Alviola, P.A., IV, and O. Capps, Jr. 2010. "Household Demand Analysis of Organic and Conventional Fluid Milk in the United States Based on the 2004 Nielsen Homescan Panel." *Agribusiness* 26(3):369–388.

Barbano, D.M., Y. Ma, and M.V.d. Santos. 2006. "Influence of Raw Milk Quality on Fluid Milk Shelf Life." *Journal of Dairy Science* 89:E15–E19.

Berry, S., J. Levinsohn, and A. Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63(4):841–890.

Blundell, R., J. Horowitz, and M. Parey. 2022. "Estimation of a Heterogeneous Demand Function with Berkson Errors." *Review of Economics and Statistics* 104(5):877–889.

Bresnahan, T.F. 1987. "Competition and Collusion in the American Automobile Industry: The 1955 Price War." *The Journal of Industrial Economics*, pp. 457–482.

Bronfenbrenner, M. 1961. "Notes on the Elasticity of Derived Demand." *Oxford Economic Papers* 13(3):254–261.

Choi, H.J., M.K. Wohlgenant, and X. Zheng. 2013. "Household-Level Welfare Effects of Organic Milk Introduction." *American Journal of Agricultural Economics* 95(4):1009–1028.

Chouinard, H.H., D.E. Davis, J.T. LaFrance, and J.M. Perloff. 2010. "Milk Marketing Order Winners and Losers." *Applied Economic Perspectives and Policy* 32(1):59–76.

Dhar, T., and J.D. Foltz. 2005. "Milk by Any Other Name… Consumer Benefits from Labeled Milk." *American Journal of Agricultural Economics* 87(1):214–228.

Dharmasena, S., and O. Capps, Jr. 2014. "Unraveling Demand for Dairy-Alternative Beverages in the United States: The Case of Soymilk." *Agricultural and Resource Economics Review* 43(1):140–157.

Dubois, P., R. Griffith, and M. O'Connell. 2018. "The Effects of Banning Advertising in Junk Food Markets." *The Review of Economic Studies* 85(1):396–436.

—. 2020. "How Well Targeted Are Soda Taxes?" *American Economic Review* 110(11):3661–3704.

Erdem, T., S. Imai, and M.P. Keane. 2003. "Brand and Quantity Choice Dynamics under Price Uncertainty." *Quantitative Marketing and Economics* 1:5–64.

Gardner, B.L. 1975. "The Farm-Retail Price Spread in a Competitive Food Industry." *American Journal of Agricultural Economics* 57(3):399–409.

Gentzkow, M. 2007. "Valuing New Goods in a Model with Complementarity: Online newspapers." *American Economic Review* 97(3):713–744.

George, P.S., and G.A. King. 1971. *Consumer Demand for Food Commodities in the United States with Projections for 1980*. California Agricultural Experiment Station.

Gillingham, K.T., S. Houde, and A.A. Van Benthem. 2021. "Consumer Myopia in Vehicle Purchases: Evidence from a Natural Experiment." *American Economic Journal: Economic Policy* 13(3):207–238.

Goolsbee, A., and A. Petrin. 2004. "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV." *Econometrica* 72(2):351–381.

Gould, B.W. 1996. "Factors Affecting US Demand for Reduced-Fat Fluid Milk." *Journal of Agricultural and Resource Economics* 21(1):68–81.

Hart, J., and J.M. Alston. 2020. "Evolving Consumption Patterns in the US Alcohol Market: Disaggregated Spatial Analysis." *Journal of Wine Economics* 15(1):5–41.

Hausman, J.A. 1996. "Valuation of New Goods under Perfect and Imperfect Competition." In *The economics of new goods*. University of Chicago Press, pp. 207–248.

Hausman, J.A., and G.K. Leonard. 2002. "The Competitive Effects of a New Product Introduction: A Case Study." *The Journal of Industrial Economics* 50(3):237–263.

Hendel, I., and A. Nevo. 2006. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica* 74(6):1637–1673.

Hitsch, G.J., A. Hortacsu, and X. Lin. 2021. "Prices and Promotions in US Retail Markets." *Quantitative Marketing and Economics* 19:289–368.

Hughner, R.S., P. McDonagh, A. Prothero, C.J. Shultz, and J. Stanton. 2007. "Who Are Organic Food Consumers? A Compilation and Review of Why People Purchase Organic Food." *Journal of Consumer Behaviour: An International Research Review* 6:94–110.

Joo, J. 2023. "Rational Inattention as an Empirical Framework for Discrete Choice and Consumer-Welfare Evaluation." *Journal of Marketing Research* 60(2):278–298.

Khanal, B., and R.A. Lopez. 2021. "Demand for Plant-Based Beverages and Competition in Fluid Milk Markets." Available online at SSRN https://ssrn.com/abstract=3888564.

Krinsky, I., and A.L. Robb. 1986. "On Approximating the Statistical Properties of Elasticities." *The Review of Economics and Statistics* 68(4):715–719.

McCarthy, K., M. Parker, A. Ameerally, S. Drake, and M. Drake. 2017. "Drivers of Choice for Fluid Milk versus Plant-Based Alternatives: What Are Consumer Perceptions of Fluid Milk?" *Journal of Dairy Science* 100(8):6125–6138.

McFadden, D., and K.E. Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15(5):447–470.

National Milk Producers Federation. 2023. "RE: Docket No. FDA-2023-D-0451 Labeling of Plant-based Milk Alternatives and Voluntary Nutrient Statements." Available online at: `https://www.regulations.gov/comment/FDA-2023-D-0451-0870`.

Nevo, A. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Economet-*

rica 69(2):307–342.

—. 2000. “A Practitioner’s Guide to Estimation of Random-Coefficients Logit Models of Demand.” *Journal of Economics & Management Strategy* 9(4):513–548.

Oh, S.E., and T. Vukina. 2022. “The Price of Cage-Free Eggs: Social Cost of Proposition 12 in California.” *American Journal of Agricultural Economics* 104(4):1293–1326.

Petrin, A. 2002. “Quantifying the Benefits of New Products: The Case of the Minivan.” *Journal of Political Economy* 110(4):705–729.

Petrin, A., and K. Train. 2010. “A Control Function Approach to Endogeneity in Consumer Choice Models.” *Journal of Marketing Research* 47(1):3–13.

Revelt, D., and K.E. Train. 1998. “Mixed Logit with Repeated Choices: Households’ Choices of Appliance Efficiency Level.” *Review of Economics and Statistics* 80(4):647–657.

Ruby, M.B. 2012. “Vegetarianism. A blossoming Field of Study.” *Appetite* 58(1):141–150.

Slade, P. 2023. “Does Plant-Based Milk Reduce Sales of Dairy Milk? Evidence from the Almond Milk Craze.” *Agricultural and Resource Economics Review* 52(1):112–131.

Stewart, H., F. Kuchler, J. Cessna, and W. Hahn. 2020. “Are Plant-Based Analogues Replacing Cow’s Milk in the American Diet?” *Journal of Agricultural and Applied Economics* 52(4):562–579.

Tonsor, G.T., J.L. Lusk, and T.C. Schroeder. 2023. “Market Potential of New Plant-Based Protein Alternatives: Insights from Four US Consumer Experiments.” *Applied Economic Perspectives and Policy* 45(1):164–181.

Train, K.E. 2009. *Discrete Choice Methods with Simulation*, 2nd ed. Cambridge university press.

—. 1998. “Recreation Demand Models with Taste Differences over People.” *Land Economics*

74(2):230–239.

U.S. Department of Agriculture, Agricultural Marketing Service. 2024. *Federal Milk Marketing Order Statistics: Final Class and Component Prices by Order*. Washington DC. Available online at:
https://www.ams.usda.gov/resources/marketing-order-statistics/.

U.S. Department of Agriculture, Economics Research Service. 2023. *Food Availability (per capita) Data System*. Washington DC. Available online at: `https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/`.

—. 2024. *Price Spreads from Farm to Consumer, Whole milk, one gallon.* Washington DC. Available online at: `https://www.ers.usda.gov/data-products/price-spreads-from-farm-to-consumer/price-spreads-from-farm-to-consumer/`.

Villas-Boas, S.B. 2007. "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data." *The Review of Economic Studies* 74(2):625–652.

Wolf, C.A., T. Malone, and B.R. McFadden. 2020. "Beverage Milk Consumption Patterns in the United States: Who Is Substituting from Dairy to Plant-Based Beverages?" *Journal of Dairy Science* 103(12):11209–11217.

Wood Milk. 2023. "*Introducing Wood Milk: Aubrey Plaza's Newest Product* [Video]." YouTube. Available online at: https://www.youtube.com/watch?v=C4klWmwC2ds.

Zhang, W., and J.M. Alston. 2018. "The Demand for Inputs and Technical Change in the US Dairy Manufacturing Industry." *Agricultural and Resource Economics Review* 47(3):533–567.

Zhao, S., L. Wang, W. Hu, and Y. Zheng. 2023. "Meet the Meatless: Demand for New Generation Plant-Based Meat Alternatives." *Applied Economic Perspectives and Policy* 45(1):4–21.

Appendices

3.A Additional Tables

Table 3.A.1. First-stage regression in control function estimation

Dependent variable: Price	Estimate	Standard error
Plant-based	-0.005	0.102
Lactose-free	0.279	0.120
Fat content	0.012	0.008
Organic attribute	0.379	0.130
Package size (quart)	0.801	0.293
Package size (half-gallon)	0.246	0.081
Price IV	0.818	0.052
Brand fixed effects	Yes	
R^2	0.8467	
Number of observations	13,589,053	

Note: Price instruments are constructed as the average prices of the same product across nearby markets. Standard errors are clustered by retailer.

Table 3.A.2. Product group elasticity - separate plant-based milk

	CC	OC	LC	Soy	Almond	Other PB
CC	-0.841 [-0.851, -0.832]	0.126 [0.121, 0.130]	0.165 [0.160, 0.170]	0.054 [0.052, 0.055]	0.154 [0.151, 0.157]	0.049 [0.047, 0.050]
OC	0.694 [0.678, 0.711]	-2.715 [-2.762, -2.679]	0.177 [0.162, 0.188]	0.202 [0.191, 0.213]	0.165 [0.157, 0.171]	0.054 [0.051, 0.056]
LC	0.510 [0.492, 0.524]	0.095 [0.088, 0.102]	-2.364 [-2.406, -2.312]	0.058 [0.054, 0.062]	0.191 [0.179, 0.203]	0.062 [0.058, 0.066]
Soy	0.587 [0.576, 0.598]	0.397 [0.381, 0.412]	0.217 [0.200, 0.231]	-3.249 [-3.278, -3.223]	0.820 [0.803, 0.838]	0.267 [0.261, 0.276]
Almond	0.549 [0.539, 0.559]	0.102 [0.097, 0.104]	0.226 [0.210, 0.240]	0.263 [0.253, 0.273]	-2.710 [-2.745, -2.682]	0.290 [0.284, 0.300]
Other PB	0.589 [0.578, 0.601]	0.104 [0.101, 0.108]	0.240 [0.224, 0.252]	0.276 [0.266, 0.285]	0.942 [0.922, 0.964]	-3.492 [-3.538, -3.454]

Note: The table reports the percentage change in quantity of each product group (row) with respect to a 1 percent increase in price for alternative product group (column). The 95% confidence intervals based on Monte Carlo simulation are shown in brackets. CC represents conventional cow's milk, OC represents organic cow's milk, and LC represents lactose-free cow's milk.

Table 3.A.3. Price and quantity effects of removal of almond milk

	Price effect		Quantity effect	
	Initial Price ($/gal)	Percentage change in price	Initial quantity share (%)	Percentage change in quantity
Conventional cow's milk	3.20	0.51%	71.81%	5.03%
Organic cow's milk	7.30	0.19%	5.99%	6.43%
Lactose-free cow's milk	8.10	0.32%	8.50%	7.40%
Almond milk	6.51	.	9.02%	-100.00%
Other plant-based milk	7.00	1.44%	4.68%	37.97%
Total	.	.	100.00%	-2.62%

Source: Author's calculation based on the counterfactual simulation of the removal of almond milk products.

Note: Prices are volume-average across 87 markets defined by the combination of quarter and DMA regions. Quantities are aggregated across all 87 markets.

3.B Data Appendix

3.B.1 Store Selection

In the RMS data from NielsenIQ, affiliated stores are categorized into five retail channels using a "Channel Codes" variable: Convenience, Drug, Food, Mass Merchandiser, and Liquor. As noted in Section 3.5, only Food retail chains are considered, as this channel is where the majority of milk sales occur. During the main sample period, from 2018 to 2020, cow's milk sales within these "Food" retail chains account for more than 88% of the total milk quantity sold (Table 3.B.1).

Table 3.B.1. Quantity shares of cow's milk by channel type

Retail channels	**Food**	**Mass Merchandiser**	**Drug**	**Convenience**	**Total**
Quantity share	88.07%	7.13%	2.75%	2.05%	100%

NilsenIQ's HMS data includes approximately 1.8 million shopping occasions related to purchases of cow's milk, plant-based milk, and other dairy products, such as yogurt, cheese, and butter (outside options), all made in stores affiliated with NielsenIQ (RMS data). Recall that shopping occasions that occur in RMS stores are preserved because price and product information of unchosen products are available only for shopping occasions in RMS stores.

As stated in the main text, eight states were selected based on their high milk sales to reduce the computational burden, while ensuring that these states also represent different regions of the U.S. Table 3.B.2 presents the number of milk shopping occasions by state in descending order. The most shopping occasions occurred in California, accounting for 10%, followed by Ohio, Michigan, Texas, Pennsylvania, New York, Illinois, and North Carolina. Together, these eight states represent 45.5% of the total shopping occasions made in RMS stores. Including North Carolina, which ranks eighth in shopping occasions, ensures that the sample stores cover the southern U.S. region, where the production of farm milk is

Table 3.B.2. Number of shopping occasions made by HMS panelists in RMS stores by state

State	**Shopping occasions**	**Cumulative share**	**State**	**Shopping occasions**	**Cumulative share**
California	178,430	9.5%	Maine	19,118	88.8%
Ohio	127,552	16.2%	Kansas	18,999	89.8%
Michigan	106,313	21.9%	Missouri	17,485	90.7%
Texas	102,022	27.3%	Connecticut	17,246	91.6%
Pennsylvania	95,497	32.4%	Utah	17,010	92.5%
New York	90,969	37.2%	South Carolina	16,921	93.4%
Illinois	80,028	41.5%	Louisiana	14,458	94.2%
North Carolina	76,665	45.5%	Nebraska	12,904	94.9%
Massachusetts	71,890	49.4%	Arkansas	11,493	95.5%
Arizona	68,927	53.0%	West Virginia	10,729	96.1%
Indiana	68,798	56.7%	Idaho	9,575	96.6%
Virginia	67,119	60.2%	Mississippi	8,911	97.1%
Washington	58,644	63.3%	New Mexico	8,812	97.5%
Wisconsin	55,719	66.3%	Rhode Island	7,216	97.9%
Colorado	51,702	69.1%	Delaware	5,918	98.2%
Georgia	50,806	71.7%	Montana	5,379	98.5%
Tennessee	47,055	74.2%	Oklahoma	4,910	98.8%
Kentucky	43,214	76.5%	Vermont	4,254	99.0%
Iowa	40,447	78.7%	Wyoming	4,186	99.2%
Maryland	38,038	80.7%	South Dakota	4,157	99.4%
Oregon	37,721	82.7%	Alabama	3,838	99.6%
Minnesota	25,555	84.1%	North Dakota	2,935	99.8%
New Hampshire	24,845	85.4%	D.C.	2,283	99.9%
New Jersey	22,807	86.6%	Florida	1,507	100.0%
Nevada	22,042	87.8%			
Total	**1,883,049**	**100.0%**			

Source: NielsenIQ HMS and RMS, 2018-2020.

Note: This table presents the number of shopping occasions in which HMS panelists purchased cow's milk, plant-based milk, or other dairy products, such as yogurt, cheese, and butter, in stores that are affiliated with NielsenIQ RMS. States are listed in descending order based on the number of shopping occasions.

Table 3.B.3. Comparison of descriptive statistics of product characteristics between the final dataset and all choice occasions

	Final dataset from eight states		**All choice occasions across all states**	
Price and product characteristics	Mean	SD	Mean	SD
Price of conventional cow's milk ($/gallon)	3.06	1.48	3.21	1.43
Price of organic cow's milk ($/gallon)	7.38	1.98	7.50	1.87
Price of lactose-free cow's milk ($/gallon)	8.01	1.55	8.02	1.57
Price of plant-based ($/gallon)	6.51	2.32	6.60	2.20
Plant-based (0/1)	0.149	0.356	0.141	0.350
Organic (0/1)	0.070	0.255	0.066	0.250
Lactose-free-cow (0/1)	0.067	0.250	0.057	0.230
Fat contents	1.771	0.961	1.679	0.990
Package size (quart) (0/1)	0.070	0.255	0.056	0.230
Package size (half-gallon) (0/1)	0.453	0.498	0.443	0.500
Package size (one-gallon) (0/1)	0.477	0.499	0.501	0.500

Note: Organic attributes applied to both cow's milk and plant-based milk. Prices are calculated using a volume-weighted average across all package sizes.

more costly than in other regions. Although Florida is the third largest populous state and similarly ranks third in NielsenIQ's HMS panelist count, it had the fewest shopping occasions in Table 3.B.2 due to most of its shopping occasions occurring at stores not affiliated with NilelsenIQ's RMS.

Table 3.B.3 compares product characteristics between the final dataset and all choice occasions across all states. Overall, price and product characteristics in cow's milk and plant-based milk are similar between the two datasets. Prices are slightly lower in the final dataset across all product categories. Households in the final dataset chose less conventional cow's milk but selected more organic products (0.4 percentage point), lactose-free cow's milk (1 percentage point), and plant-based milk (0.8 percentage point).

3.B.2 Product Definition

Among these NielsenIQ RMS stores classified under the "Food" chains in the selected eight states, there are 1,520 unique UPCs representing 269 brands in cow's milk and 697 unique UPCs representing 99 brands in plant-based milk. I exclude small package sizes under 28 oz because those products account for less than 1% of the total quantity sold. After excluding these small packages, the dataset contains 1,371 UPCs consisting of 159 brands in cow's milk and 530 UPCs consisting of 69 brands in plant-based milk.

Since the number of UPCs is still too large for empirical analysis, I aggregate them into two "local" niche brands for cow's milk ("local cow") and plant-based milk (local plant-based"). For conventional cow's milk, I aggregate small brands whose brand-specific cumulative quantity share falls below 97%. The same criterion applies to organic cow's milk and lactose-free cow's milk together, and plant-based milk. This process results in 14 cow's conventional brands, 11 organic/lactose-free cow's milk brands, 11 plant-based milk brands, and two local composite brands (cow's milk and plant-based milk local composite brands).

I then further aggregate these UPCs into "products" based on their product characteristics. For example, there are twelve UPCs for Blue Diamonds' Almond Breeze, all of which are 32 oz package, non-organic, and refrigerated products. These twelve different versions of UPCs, which share the same observed product characteristics, are aggregated into a single "product." After this aggregation, the final dataset contains a total of 358 products: 128 cow's conventional milk products from 975 UPCs, 72 organic cow's milk products from 302 UPCs, 56 lactose-free milk products from 116 UPCs, and 102 plant-based milk products from 530 UPCs.

3.B.3 Treatment of Multiple Purchases of Products in Shopping Occasions

Nielsen HMS households choose one unit of a single product in 78% of the store visits. In some cases, households purchase multiple units of a single product in 12% of the visits,

Table 3.B.4. Number of products purchased in shopping occasions

Number of products	Frequency	Share (%)	Purchase types		
			Cow only	PB only	Both Cow and PB
1	113,678	89.74	99,703	13,975	0
2	11,989	9.46	5,535	1,328	5,126
3	940	0.74	244	122	574
4	69	0.05	7	24	38
5	4	0.003	0	0	4
Total	126,680	100	105,489	15,449	5,742

Source: NielsenIQ HMS, 2018–2020.

Note: This table shows the number of products purchased during shopping occasions before they are split. There are a total of 126,680 unique store visits in the final dataset. Note that Table 3.3 lists 130,237 milk shopping occasions, some of which are treated as separate occasions from 126,680 store visits. "Cow Only" includes all types of conventional, organic, and lactose-free cow's milk, while "PB Only" represents purchases of plant-based milk without buying any cow's milk.

while other households purchase multiple products in 10% of the visits. As noted in the main text, these shopping occasions are split and treated as separate shopping occasions. The random sampling of choice occasions up to 10 per household applies to these separate shopping occasions.

Table 3.B.4 presents the number of products that Nielsen households purchased on a shopping occasion. When households choose multiple products on a shopping occasion, they typically purchase two products (9.46% out of 10.26%). Among 120,938 cow's milk shopping occasions (the sum of "Cow only" and "Both Cow and PB"), plant-based milk was also purchased 5% of the time (=5,742/120,938). For 21,191 plant-based milk shopping (the sum of "PB only" and "Both Cow and PB"), cow's milk was also purchased 27% (=5,742/21,191) of the time.

Chapter 4

Effects of Potential Changes in U.S. Farm Milk Price Policies on Prices and Quantities of Cow's Milk and Plant-based Milk

4.1 Introduction

Federal and state regulations of farm milk prices have governed milk marketing in the United States since the 1930s. In particular, for about nine decades, a set of complicated Federal Milk Marketing Orders (FMMOs) has regulated minimum milk prices, where milk is marketed, and how milk revenue is distributed across farms. This system raises the minimum price paid for farm milk that is used for fluid products relative to prices paid to farms for identical milk used for other products, such as cheese, dry milk powders, and butter. The higher cost of farm milk caused by government regulation causes retail fluid milk products to become more expensive and thereby contributes to reduced consumption of fluid milk products.

The degree to which regulation-caused price increases suppress fluid milk consumption depends on the own-price elasticity of demand and hence on the prevalence of substitutes in the market. Therefore, the recent spread of plant-based milk substitutes and their importance to households that would otherwise buy cow's milk may increase the negative impact of the FMMO price discrimination rules on the quantity of cow's milk demanded. This chapter explores the implications of regulations in the FMMO system on prices and quantities of cow's milk in the presence of plant-based milk products.

Many economists have examined the economic implications of the FMMO system, but the existing literature has given relatively little attention to the roles played by potential substitutes for fluid milk (Ippolito and Masson 1978; Hanon 2023; Hanon, Mérel, and Sumner

2024). Results of Chapter 3 and research in the literature highlight that plant-based products have established themselves as viable alternatives for animal-based products (Stewart et al. 2020; Tonsor, Lusk, and Schroeder 2023; Slade 2023). In particular, consumption of plant-based fluid milk substitute products now accounts for about 10% of the quantity share in the combined cow's milk and plant-based milk sales. Considering the role of the fluid alternatives to cow's milk is important when evaluating FMMO price regulations because the impacts of these regulations depend on the substitutability between cow's milk and substitutes, such as plant-based milk.

This chapter reports on two milk price policy simulations: (1) An increase in the minimum price of farm milk used for fluid products, and (2) the removal of the longstanding above-market minimum prices for farm milk used for fluid products under the FMMO system. The first simulation reflects the recently proposed FMMO reform (USDA AMS 2024a). The second simulation reflects the price projections from the FMMO deregulation scenario from Kawaguchi, Suzuki, and Kaiser (2001).

The policy simulations leverage the demand parameters and marginal costs of processing cow's milk and plant-based milk, as estimated in Chapter 3. The two policy alternatives consider policy-driven exogenous changes in the price of farm milk used for fluid products, which lead to changes in the marginal costs of certain cow's milk products. The implied changes in equilibrium retail prices and quantities of cow's milk and plant-based milk are calculated in response to the exogenous marginal cost changes.

The following section includes a summary of the provisions of Federal Milk Marketing Orders and a discussion of alternative policies. Section 4.3 lays out the simulation methods and results.

4.2 Review of Economic Consequences of the Federal Milk Marketing Orders

This section provides the needed background information on the farm milk pricing regulations in the Federal Milk Marketing Orders (FMMOs) and their economic implications.

4.2.1 Summary of Federal Milk Marketing Order Rules

The Federal Milk Marketing Order (FMMO system), comprised of 11 regional sets of related regulations, has played a large role in the dairy product markets since its creation in the 1930s. Each of the 11 separate marketing orders regulates milk processing plants in its geographic territory, and each has different specific regulations. Some of these affect the movement of farm milk across order boundaries, and others apply different price provisions (Sumner 2018; Hanon 2023; USDA AMS 2024b). There are two main pricing features in the FMMO system.

First, FMMOs specify minimum prices that buyers must pay for farm milk by the end-use product category for which the milk will be used. The FMMO system recognizes four end-use product classes. Class I milk is used for fluid milk products. Class II milk is used for soft and frozen products, including ice cream, cottage cheese, and yogurt. Class III milk is used for hard cheese, cream cheese, and whey products. Class IV milk is used for butter, dry milk powder, and other dried dairy products. Although minimum prices differ, the milk sold in these classes may be identical. Farm milk comprises three separable solid components to which regulated prices apply: fat, protein, and other non-fat solids. Price rules apply to the components. Specifically, all orders set a positive price differential for milk used for fluid products, and the amount of that Class I price differential varies across the 11 regional orders (Hanon 2023; Hanon, Mérel, and Sumner 2024; USDA AMS 2024b, 2024c).[1]

[1]Class I price is based on the prices of two components – butterfat and skim milk. The Class I butterfat price is calculated using a forward projection of the national butter price. The Class I skim price is calculated as the average projected Class III and Class IV skim milk pricing factors plus $0.74 per hundredweight and regional order-specific Class 1 differentials. USDA AMS publishes monthly Class I prices for each order, based on the fixed proportions of butterfat (3.5%) and skim milk (96.5%). The regional order-specific Class

Second, each regional order requires a detailed accounting of revenue generated by the minimum price system and pools all milk revenue generated by the minimum prices of farm milk with all the end-use product categories. The FMMO system does not regulate farms but rather regulates participating dairy handlers or processors that procure farm milk. Processors may operate outside the system, but to do so, they must offer payments to farms that are expected to be at least as attractive as participation in the FMMO system; otherwise, farms would deliver milk to participating processing plants (USDA AMS 2024b).

The dairy handlers or processors are the entities that contribute revenue to the pool. The amount of money that they are required to pay into the pool depends on the amount of each component they buy to make each of the four classes of dairy products. The total amount of pooled milk receipts is the sum of the total quantity of each milk component purchased times the minimum price for each milk component for each class of milk. Farm milk producers are paid from the pool a weighted average price as a minimum price for their milk, no matter how their farm's milk is used. Usually, the minimum price is not equal to the actual price that farm milk producers are paid. Often, handlers or processors pay more than the minimum price (over-order premiums) based on competitive conditions in local milk procurement and milk quality characteristics (Sumner 2018; Hanon 2023).

4.2.2 Federal Milk Marketing Orders Direct Implications

While many specifics have changed, the key features of FMMOs, price discrimination by end-use and revenue pooling, have continued since the FMMO system began in the 1930s. Many economists have explored the economic consequences of the FMMO system (and similar state milk marketing orders) for farm milk price, quantity produced, social costs, and quality standards of fluid milk (Ippolito and Masson 1978; LaFrance and De Gorter 1985; Sumner and Wolf 1996; Cox and Chavas 2001; Kawaguchi, Suzuki, and Kaiser 2001; Balagtas, Smith, and Sumner 2007; Ahn and Sumner 2009; Sumner 2018; Hanon 2023; Hanon, Mérel, and Sumner 2024).

I prices and potential changes are used in the simulations of the effects of FMMO policy changes.

This literature has described and measured some economic effects of milk marketing orders. First, because price discrimination under the FMMO system raises the relative price of Class I milk to other classes, it leads to a decrease in the consumption of fluid products and a decline in the consumer surplus of fluid milk buyers. Second, price discrimination, together with revenue pooling, likely raises the average price that farm milk producers get paid, increases the quantity of milk supplied, and likely generates a larger farm milk revenue than without the FMMO system. Third, since the FMMO system reduces fluid milk consumption and increases farm milk production, the quantity and share of milk used for manufactured products (Class III and IV products) are larger under the FMMO system. The articles cited above each present similar overall results as summarized and reviewed in Sumner (2018) and Hanon (2023).

Previous studies in this literature have assessed to what extent the presence of the FMMOs raised the farm milk prices used for fluid products. Modeling an aggregate milk price policy with price discrimination and pooling, Ippolito and Masson (1978) calculate a 9.3% increase in Class I price due to FMMOs. LaFrance and De Gorter (1985) found that, on average, from 1965 to 1980, Class I prices would have decreased by 16.2% in competitive market conditions. Cox and Chavas (2001), in their second scenario, computed a 14% reduction in Class I price in the absence of FMMOs. Kawaguchi, Suzuki, and Kaiser (2001) projected the Class I prices that would have occurred in each marketing order in the absence of the FMMO system. They found declines ranging from 5.6% lower in the Southeast order to 15.6% lower in the Northeast order compared to the Class I prices in 1997.

4.2.3 Recent Recommended Increases in Class I Price Differentials

The price differentials under the FMMO system have remained largely unchanged since 2000. The single adjustment was a one-time increase of 10 to 20 percent in the differential in 2008 for three orders in the southeastern United States. In addition, California switched from a state-run regulation to the federal system in 2018.

After an elaborate and extended set of formal hearings, on July 15, 2024, the USDA Agricultural Marketing Service (AMS) released several recommended changes to the FMMO system for public comment, including adjustments to Class I price differentials and certain classified price formulas (USDA AMS 2024a). During the hearing process that began in 2023, the National Milk Producer Federation (NMPF) proposed county-specific increases in Class I price differentials, averaging approximately 12.5 cents per gallon (USDA AMS 2023).[2] The increases recommended by USDA AMS in July 2024 are slightly adjusted but closely aligned with those proposed by NMPF, averaging about 10.4 cents per gallon.[3] After the conclusion of the formal comment period, dairy farmers in each FMMO will vote on proposed amendments, and if approved, the amended rules will take effect. If the proposed amendments are rejected, the orders will be eliminated (USDA AMS 2024a, 2024b, and 2024c).

4.3 Implications of Changes in Federal Milk Marketing Orders for Farm Milk Prices

This sub-section presents the details of two policy-related counterfactuals concerning policy-induced changes in farm milk prices and explains how these changes cause expected changes in retail prices and quantities of cow's milk and plant-based milk products.

4.3.1 Alternative Farm Milk Pricing Policies

Class I price differentials differ across regions. The regional order-specific differential tends to be large in milk deficit regions, such as Florida, and tends to be small in major milk

[2]NMPF proposed the adjustment of Class I price differentials in all regional orders for 3,108 counties, parishes, and independent cities. The simple average of differences between the current and proposed Class I price differentials per hundredweight across these regions is calculated and converted to dollars per gallon (divided by 11.98 gallons per hundredweight).

[3]This represents the simple average of the difference between the current and the USDA AMS proposed Class I price differentials across 3,108 counties.

production regions, such as California and the Upper Midwest. Even within each regional order, the regional differential varies depending on the location of the processing plant. The second to fourth columns of Table 4.1 report the simple average of Class I milk prices and Class I price differentials across counties within each regional order. The average Class I price differential is the largest in the Florida order ($0.44 per gallon) and the smallest in the Upper Midwest order ($0.14 per gallon).

The counterfactual scenarios examine the effects of the changes in Class I price differentials. The first scenario considers retail responses based on the recent USDA marketing order reform draft proposal (USDA AMS 2024a), which specifies changes in Class I price differentials across 3,108 counties. The fifth column of Table 4.1 shows the simple average of the recommended differentials across counties within each of the 11 regional orders.[4] The sixth column presents the percentage change in the Class I price resulting from changes in price differential, relative to the 2018–2020 average Class I prices in the third column. The recommended increases in price differentials are larger in the Southern regional orders and the Northeast order and smaller in the Western regional orders.

The second scenario simulates the retail impacts of the removal of the Class I differentials based on projections of the Class I price changes derived from the FMMO deregulation scenario of Kawaguchi, Suzuki, and Kaiser (2001). Their price simulations are applied to each regional order using the Class I price differentials in the FMMO system that has been in place since 2000. The last column of Table 4.1 presents the projected percentage decreases in Class I prices for each regional order.[5] Since California was not part of the federal order

[4]Section 4.3 provides more details on how the price differentials across 3,108 counties are matched with geographic market definitions used in the counterfactual analysis. The absolute changes in price differential, not in percentage terms, are used in the simulation of the 2024 FMMO reform proposal.

[5]Section 4.3. provides more details on how the percentage price differentials from Kawaguchi, Suzuki, and Kaiser (2001) apply to the counterfactual analysis on the elimination of the FMMOs. The percentage changes in price differentials are used in this counterfactual analysis because the absolute price differentials used in Kawaguchi, Suzuki, and Kaiser (2001) are measured based on the 1997 prices.

Table 4.1. Class I prices (2018-2020), USDA AMS proposed Class I price differentials and FMMO deregulation scenarios

	Class I price		**Current Class I price**	**USDA AMS proposed increase in Class I price**		**Decrease in Class I price under FMMO deregulation from Kawaguchi et al. (2001)**	
	($/cwt)	($/gal)	($/gal)	($/gal)	(% increase from Class I price)	($/gal)	(% decrease from Class I price)
Northeast (1)	19.48	1.63	0.24	0.14	8.50%	-0.25	-15.58%
Appalachian (5)	19.64	1.64	0.26	0.16	9.53%	-0.25	-15.32%
Florida (6)	21.63	1.81	0.44	0.12	6.42%	-0.15	-8.26%
Southeast (7)	20.00	1.67	0.29	0.14	8.12%	-0.09	-5.59%
Upper Midwest (30)	17.99	1.50	0.14	0.09	5.98%	-0.12	-7.79%
Central (32)	18.24	1.52	0.17	0.07	4.88%	-0.14	-9.04%
Mideast (33)	18.29	1.53	0.16	0.14	9.30%	-0.14	-9.04%
California (51)	18.28	1.53	0.15	0.04	2.65%	-0.13	-8.42%
Pacific Northwest (124)	18.12	1.51	0.15	0.04	2.52%	-0.13	-8.42%
Southwest (126)	19.25	1.61	0.24	0.05	3.23%	-0.24	-14.78%
Arizona (131)	18.60	1.55	0.17	0.03	1.88%	-0.13	-8.42%

Source: FMMO statistics (2018-2020) from USDA AMS (2024c), Proposed Class I price differentials from USDA AMS (2024a), and Tables 1 and 4 in Kawaguchi, Suzuki, and Kaiser (2001).

Note: Class I prices in the second and third columns and Class I price differentials in the fourth column are averaged across all months and all counties within each marketing order between 2018 and 2020. The fifth column reports the simple average of USDA AMS proposed differentials across counties within each of the 11 regional orders. In the last two columns for the California order, the percentage change is imputed using the nearby orders, Arizona and the Pacific Northwest.

in 2000, its percentage decrease in Class I price was not included in their projections. I adopted the percentage change for California order from nearby Arizona and the Pacific Northwest orders (-8.42%). The decreases in the Class I price in the second to the last column are calculated by multiplying the 2018-2020 average Class I price in the third column by the percentage declines.

4.3.2 Implications for Retail Prices and Quantities under Counterfactual Policies

Recall from Chapter 3 that the econometric estimates of demand parameters were obtained from the estimation of the discrete choice demand model using data on households' shopping occasions in the eight states where most shopping trips occurred. These estimated demand parameters, together with the supply side of the Bertrand-Nash model, allow us to calculate the implied marginal costs facing retailers for each cow's milk and plant-based milk products in each market, defined as the combination of a year and Designated Marketing Area (DMA). The eight states included in the main sample of Chapter 3 contain 29 DMAs that represent seven of 11 regional FMMOs.[6]

The two policy simulations consider policy-driven exogenous changes in the Class I farm milk prices, which lead to changes in the marginal costs of processing certain cow's milk products. Recall that USDA AMS recommendations, which drive the first policy alternative, specify the price differentials for 3,108 counties. These counties are matched with 29 DMAs to calculate the DMA-specific increases in Class I farm milk prices for those locations. For the second scenario, which involves the elimination of Class I price differentials, the projected decreases in Class I farm milk prices, specific to each regional order, are applied to DMAs within those regional orders.

The changes in FMMO minimums apply to the costs of procuring farm milk for conventional and lactose-free milk products in all potential markets. Farm prices of organic milk

[6]These seven orders include the Northeast, Appalachian, Upper Midwest, Central, Mideast, California, and Southwest orders.

are so far above the FMMO minimums that market prices are unaffected by the FMMO minimums or by the two policy-driven price changes.[7] Of course, the FMMO minimums do not apply to plant-based milk products.

The marginal costs of processing wholesale milk under the proposed change in Class I farm price differential are specified in Equation (4.1). In Equation (4.1), mc_{jm} represents the marginal cost of product j in market m defined by DMA and year. Recall from Chapter 3 that these marginal costs apply to a total of 87 markets, comprising 29 DMAs over three years. An average of approximately 167 products are sold in each market, leading to a total of 14,563 implied marginal costs (mc_{jm}) estimated across all markets and products.

$$mc^*_{j,m} = \begin{cases} mc_{j,m} & \forall j \notin J_{CC,LF} \\ mc_{j,m} + \Delta PD_m & \forall j \in J_{CC,LF} \end{cases} \tag{4.1}$$

The alternative farm milk pricing policies would change the marginal cost of processing a set of conventional cow's milk and lactose-free cow's milk products, denoted by $J_{(CC,LF)}$. These marginal costs are assumed to change by the amount that the policy-induced shifts in Class I farm milk prices cause, while other marketing costs are assumed to be constant.[8] In the first counterfactual scenario, which considers the USDA's proposed changes to price differentials, the absolute changes in price differentials for each geographical market (ΔPD_m) are applied, after matching 3,108 county-specific differentials to 29 DMAs. For the second counterfactual scenario of the elimination of the FMMOs, the changes in marginal costs (ΔPD_m) are calculated by multiplying the 2018-2020 average Class I price in each market ($PC1_m$) by the market-specific percentage changes in Class I price (α_m). That is $\Delta PD_m = PC1_m \times \alpha_m$. The conversion from the percentage to the absolute changes in the second

[7]Note that organic fluid processors are required to pay into the marketing order pools.

[8]Farm milk constitutes a significant portion of the total input costs for fluid milk products (Zhang and Alston 2018). For conventional cow's milk, in particular, the Class I minimum price accounts for more than half of the retail price.

scenario is necessary because the absolute changes in Kawaguchi, Suzuki, and Kaiser (2001) are measured based on the 1997 prices.

Using the calculated counterfactual marginal costs (mc^*_{jm}) along with the estimated demand parameters, the vector of equilibrium prices ($\boldsymbol{p^*_m}$) under the counterfactual scenarios are found by solving Equation (4.2). The notation in Equation (4.2) follows from Chapter 3. The term, s_{jm} is the quantity share of product j in market m, which depends on the prices of all products in market m. The set of products sold by firm f in market m is denoted by J_{fm}.

$$s_{jm}(\boldsymbol{p^*_m}) + \sum_{k \in J_{fm}} (p^*_{km} - mc^*_{km}) \frac{\partial s_{km}(\boldsymbol{p^*_m})}{\partial p_{jm}} = 0 \qquad \forall j \in J_{fm} \tag{4.2}$$

The corresponding counterfactual quantity shares are obtained by substituting the counterfactual prices ($\boldsymbol{p^*_m}$) into share functions (s_{km}). The equilibrium solutions using the demand parameters and the marginal cost specifications allow us to assess the effects of changes in FMMO policy on retail market outcomes.

Table 4.2 reports the implied effects on prices and quantities of the increase in the Class I price differential in the 2024 FMMO reform proposal. Prices of organic cow's milk and plant-based milk products are almost unchanged. The implied retail price of conventional cow's milk increases by 4.0%, and lactose-free cow's milk increases by 1.6%. When that happens, the retail equilibrium quantity of conventional cow's milk would decrease by 3.6% and lactose-free cow's milk by 2.2% in a short-run horizon, using the interpretation of the demand parameter estimates. In contrast, the quantity of organic cow's milk would rise by 3.0%, and plant-based milk would rise by 2.6% because these products would become less expensive relative to conventional and lactose-free cow's milk.

The policy simulation results in Table 4.2 imply that the USDA-recommended increases in Class I price differentials are likely to further reduce the retail market share of cow's milk,

Table 4.2. Equilibrium simulated price and quantity effects of an increase in Class I price differential (FMMO reform proposal 2024)

	Price effect		**Quantity effect**	
	Initial price ($/gal)	% change in price	Initial quantity share (%)	% change in quantity demanded
Conventional cow's milk	3.20	4.00%	71.81	-3.56%
Organic cow's milk	7.30	-0.13%	5.99	3.04%
Lactose-free cow's milk	8.10	1.64%	8.50	-2.17%
Plant-based milk	6.68	-0.08%	13.70	2.57%
Total	.	.	100	-2.20%

Source: Author's calculation based on the counterfactual simulation of Class I price differential change in FMMOs.

Note: Percentage changes in prices are volume-weighted averages across 87 markets defined by the combination of three years and 29 DMA regions, covering seven regional orders. Percentage changes in quantities are aggregated across all 87 markets.

particularly for conventional cow's milk products, given that the farm milk cost accounts for a substantial portion of the retail prices of conventional cow's milk and a much smaller increase in the retail price of lactose-free cow's milk. However, the magnitude of the effects on retail prices and quantities of the relevant products overall are small, both because the initial retail price shocks are small, and the retail demand elasticities remain relatively low for conventional milk despite the prevalence and growing shares of plant-based milk in retail venues.

Table 4.3 reports the results of the counterfactual removal of FMMO price differentials. The retail price of conventional cow's milk would fall by 5.6% and lactose-free cow's milk by 2.4%, on average across markets. Prices of organic cow's milk and plant-based milk are almost unaffected by the deregulation scenario. The quantity share of conventional milk would rise the greatest in both absolute and relative terms. The quantity of conventional cow's milk would rise by 5.2% and lactose-free milk by 3.3%. Conversely, the quantity of organic cow's milk and plant-based milk would decrease by 4.3% and 3.8%, respectively.

Table 4.3. Equilibrium simulated price and quantity effects of deregulation of the FMMO system

	Price effect		Quantity effect	
	Initial price ($/gal)	% change in price	Initial quantity share (%)	% change in quantity demanded
Conventional cow's milk	3.20	-5.57%	71.81	5.16%
Organic cow's milk	7.30	0.13%	5.99	-4.31%
Lactose-free cow's milk	8.10	-2.42%	8.50	3.25%
Plant-based milk	6.68	0.12%	13.70	-3.75%
Total	.	.	100	3.21%

Source: Author's calculation based on the counterfactual simulation of Class I price differential change in FMMOs.

Note: Percentage changes in prices are volume-weighted averages across 87 markets defined by the combination of three years and 29 DMA regions, covering seven regional orders. Percentage changes in quantities are aggregated across all 87 markets.

The results in Table 4.3 imply that eliminating the fluid price differentials set by FMMOs would have limited impacts on boosting the retail quantity of conventional cow's milk. Recall from Chapter 3 that the per capita consumption of fluid cow's milk fell by 25% from 2006 to 2020. The projected 5% increase in the retail quantity of conventional cow's milk is insufficient to reverse this downward trend and is even smaller than the 11% decrease in the equilibrium quantity of conventional cow's milk attributed to the presence of plant-based milk substitutes.

4.4 Summary of Policy Implications and Concluding Remarks

This chapter presented the changes in the prices and quantities of cow's milk under alternative farm milk price regulations, considering the presence of plant-based milk as a viable substitute, recognizing the significant role plant-based milk now plays is important in evaluating these policy alternatives. By leveraging estimated demand parameters along with the supply-side model from Chapter 3, this chapter evaluated the impacts of two alternative farm milk price regulations.

The results suggest that relative price changes between cow's milk and plant-based milk, induced by the USDA price regulations, are not the primary factor for the declining consumption or the key to revitalizing the demand. Increasing the price differentials for farm milk used for fluid products, as proposed by USDA AMS, further pressures on the already declining consumption of fluid cow's milk, though the decrease is modest. While producers of organic cow's milk and plant-based milk would benefit from this change, the increases in quantity would be small.

The proposed increase in price differentials would have small impacts on farm milk production and revenues. The higher Class I price differentials would induce a slight increase in farm milk production. Given that the average share of farm milk used for fluid products was 28% of the total milk production from 2018 to 2020, farm milk revenues would see a modest increase from the proposed adjustment in price differentials.

References

Ahn, B.I., and D.A. Sumner. 2009. "Political Market Power Reflected in Milk Pricing Regulations." *American Journal of Agricultural Economics* 91(3):723–737.

Balagtas, J.V., A. Smith, and D.A. Sumner. 2007. "Effects of Milk Marketing Order Regulation on the Share of Fluid-Grade Milk in the United States." *American journal of agricultural economics* 89(4):839–851.

Cox, T.L., and J.P. Chavas. 2001. "An Interregional Analysis of Price Discrimination and Domestic Policy Reform in the US Dairy Sector." *American Journal of Agricultural Economics* 83(1):89–106.

Hanon, T.M. 2023. *Mapping Milk: Investigating the Effects of Federal Milk Marketing Orders on the Geography of Milk Production and Inter-Regional Trade in Milk and Dairy Products*. PhD dissertation, University of California, Davis.

Hanon, T.M., P. Mérel, and D.A. Sumner. 2024. "Effects of U.S. Federal Milk Marketing Orders on Inter-Regional Trade and the Dairy Supply Chain." Working paper, Department of Agricultural and Resource Economics, University of California, Davis.

Ippolito, R.A., and R.T. Masson. 1978. "The Social Cost of Government Regulation of Milk." *The Journal of Law and Economics* 21(1):33–65.

Kawaguchi, T., N. Suzuki, and H.M. Kaiser. 2001. "Evaluating Class I Differentials in the New Federal Milk Marketing Order System." *Agribusiness: An International Journal* 17(4):527–538.

LaFrance, J.T., and H. De Gorter. 1985. "Regulation in a Dynamic Market: The US Dairy Industry." *American Journal of Agricultural Economics* 67(4):821–832.

Slade, P. 2023. "Does Plant-Based Milk Reduce Sales of Dairy Milk? Evidence from the Almond Milk Craze." *Agricultural and Resource Economics Review* 52(1):112–131.

Stewart, H., F. Kuchler, J. Cessna, and W. Hahn. 2020. "Are Plant-Based Analogues Replacing Cow's Milk in the American Diet?" *Journal of Agricultural and Applied Economics* 52(4):562–579.

Sumner, D.A. 2018. "Dairy Policy Progress: Completing the Move to Markets." *American Enterprise Institute*, Washington, DC.

Sumner, D.A., and C.A. Wolf. 1996. "Quotas Without Supply Control: Effects of Dairy Quota Policy in California." *American Journal of Agricultural Economics* 78(2):354–366.

Tonsor, G.T., J.L. Lusk, and T.C. Schroeder. 2023. "Market Potential of New Plant-Based Protein Alternatives: Insights from Four US Consumer Experiments." *Applied Economic Perspectives and Policy* 45(1):164–181.

U.S. Department of Agriculture, Agricultural Marketing Service. 2023. *Appendix A: National Milk Producers Federation – Proposed Class I Differentials.*. Washington DC. Available online at: https://www.ams.usda.gov/sites/default/files/media/NationalHearing NMPFProposedDifferentials.pdf.

—. 2024a. *Milk in the Northeast and Other Marketing Areas; Proposed Amendments to Marketing Agreements and Orders*. Washington DC. Available online at: https://www.federalregister.gov/documents/2024/07/15/2024-14769/milk-in-the-northeast-and-other-marketing-areas-proposed-amendments-to-marketing-agreements-andopen-comment.

—. 2024b. *Federal Milk Marketing Orders*, various issues. Washington DC. Available online at: https://www.ams.usda.gov/rules-regulations/moa/dairy.

—. 2024c. *Federal Milk Marketing Order Statistics: Final Class and Component Prices by Order*. Washington DC. Available online at: https://www.ams.usda.gov/resources/marketing-order-statistics/.

Zhang, W., and J.M. Alston. 2018. "The Demand for Inputs and Technical Change in the US Dairy Manufacturing Industry." *Agricultural and Resource Economics Review* 47(3):533–567.

Chapter 5

Summary and Concluding Remarks

This dissertation investigates the economic impacts of the introduction and growing availability of plant-based milk on prices and quantities of cow's milk. Although conventional cow's milk continues to capture the largest share of consumption, the rapid expansion of plant-based milk products places additional pressure on the consumption of cow's milk, especially for organic and lactose-free cow's milk.

Chapter 2 empirically analyzes the impacts of the introduction of refrigerated almond milk from 2008 to 2010 on the retail prices and quantities of cow's milk and soymilk. Analyses of the impacts of market entry of new products are challenging both conceptually and empirically. This chapter uses models developed in the industrial organization literature, which assume that the wholesale suppliers of original plant-based and cow's milk had market power. Also, almond milk suppliers determined the gradual introduction of almond milk into geographically dispersed markets. Empirically, Chapter 2 leverages the temporal variations in the availability of refrigerated almond milk across U.S. retail stores to identify the price and quantity effects on competing products.

By employing the conditional parallel trend assumptions that carefully considered the selection in adoption timing, this chapter used the recently developed econometric strategies to estimate the price and quantity effects. Empirical results indicate that the introduction of refrigerated almond milk led to a short-run (within several months) 6% decline in the quantity of soymilk sold in retail stores and a 3% decline in the quantities of organic cow's milk and lactose-free cow's milk. In contrast, the short-run demand for conventional cow's milk remained largely unaffected. These estimated quantity effects are consistent with expectations that products that are expected to be more substitutable for almond milk experienced larger reductions in quantity. The estimated effects imply that, on an annual per capita basis, the quantity of cow's milk and soymilk decreased by between 0.055 and 0.086 gallons. During the 2008–2010 period, the annual per capita quantity of refrigerated almond milk increased by 0.152 gallons, indicating that refrigerated almond milk did not merely replace soymilk and cow's milk but also contributed to expanding the overall milk market. However, price effects were minimal for all the product categories.

Chapter 3 develops and estimates a discrete choice demand model for many individual retail cow's milk and plant-based milk products. I use a novel data-matching process to represent consumers' choice sets using household purchase data with store-level data. The estimation process retrieves demand parameters on valuations of individual product characteristics. To summarize important demand relationships, these parameters are aggregated into implied own- and cross-price elasticities of demand for product groups: (1) conventional, (2) organic, (3) lactose-free cow's milk, and (4) plant-based milk. The supply side is modeled as an oligopolistic market where milk processors follow Bertrand-Nash price competition in setting prices facing local groceries in geographically defined markets.

Demand estimation results provide the own- and cross-price elasticity for the four aggregate product groups. Own-price elasticities range from about -0.84 for conventional, -2.75 for organic, -2.36 for lactose-free, and -2.19 for plant-based milk. The econometric results also show, through the estimated positive cross-elasticities of demand, that plant-based milk is a closer substitute for organic and lactose-free than for conventional cow's milk.

Using the estimated retail consumer demand parameters and wholesale marginal costs, I conducted counterfactual simulations of removing all plant-based milk options from consumers' choice sets. In line with findings on the cross-price elasticities, the counterfactual simulation shows that removing all plant-based milk products from consumers' choice sets would lead to a 23% increase in the quantity of organic cow's milk, a 16% increase for lactose-free cow's milk, and an 11% increase for conventional cow's milk. The counterfactual results imply that the removal of plant-based milk would raise the annual per capita consumption of conventional cow's milk by 1.46 gallons and by 0.26 gallons each for organic cow's milk and lactose-free cow's milk. Out of the 5.1-gallon decrease in annual per capita cow's milk consumption observed between 2006 and 2020, the counterfactual experiment suggests that approximately 38% of this decrease can be attributed to the availability of plant-based milk.

Chapter 4 uses the demand estimates and supply model from Chapter 3 to explore some retail milk price implications of modifications in U.S. farm milk price regulations. This chapter considers two policy simulations: (1) an increase in farm milk used for fluid products and (2) the elimination of the longstanding USDA regulations that raise the price of farm milk used for fluid products.

The first simulation used the July 2024 "reform" proposal of USDA to raise the minimum prices required for farm milk processed in the sorts of fluid products studied in Chapter 3. This proposal was strongly supported by the main dairy farmer organization on the grounds that it would add to dairy farm revenue. The results of this simulation show that raising the farm price of milk for fluid consumer products would increase the retail price of conventional cow's milk by about 4% and decrease the retail quantity by around 3.6%, further reducing cow's milk consumption and use of farm milk for those uses that command higher prices.

Conversely, my simulations show that eliminating USDA-regulated minimum farm prices for milk used for fluid products would decrease retail prices of conventional cow's milk by about 5.6% and increase the retail quantity by 5.2%. Overall, these findings indicate that farm milk price regulations benefit the prices and quantity of plant-based milk products and cause lower retail consumption of fluid cow's milk and, consequently, less farm milk used for fluid products.

This dissertation provides useful insights into the retail demand for plant-based milk and cow's milk products. The empirical results strongly support the hypothesis that plant-based milk availability has affected and continues to affect the demand for retail cow's milk products, especially the higher-priced organic and lactose-free products. However, plant-based milk has not been the primary driver in the recent accelerated decline in consumption of fluid cow's milk. Finally, despite the widespread retail availability of plant-based milk, which substitutes for fluid cow's milk products, policy-driven increases in the farm prices of cow's milk raw material have only moderate negative impacts on reducing the quantity of farm milk used for fluid retail products.

Building on my dissertation research, several areas merit further investigation. First, Chapter 2 presents short-run estimates of the price and quantity effects of the introduction of refrigerated almond milk. Future work can use econometric estimates to project longer-term price and quantity effects, drawing on the evolution observed in the event-study estimates.

My estimates of counterfactual impacts of the removal of plant-based milk from consumers' choice sets and the comparison with the actual decline in per capita consumption of fluid cow's milk are based on data from 2018-2020. With more data and computational efficiencies, future research can use NielsenIQ household purchase data from as early as 2006 through 2022 to map the trajectory of the effects of plant-based milk availability over time.

Chapter 4 explores the retail price and quantity implications of U.S. farm milk price regulations, specifically the Federal Milk Marking Order above market minimum prices for farm milk used for fluid milk products. Future work can consider the farm revenue implications of the growing availability of milk substitutes. By extending the demand estimation data back to 2006, this research can investigate how changes in the price elasticities of retail cow's milk have influenced the effectiveness of raising the price of fluid-use farm milk in generating higher dairy farm revenue.