JOURNAL of FOOD DISTRIBUTION RESEARCH

Volume 55 / Issue 3 / November 2024



Published by

FDRS
FOOD DISTRIBUTION RESEARCH SOCIETY

http://www.fdrsinc.org

Food Distribution Research Society

2024 Officers and Directors

Journal of Food Distribution Research
Volume 55, Issue 3
November 2024

ISSN 0047-245X

The *Journal of Food Distribution Research* has an applied, problem-oriented focus on the flow of food products and services through wholesale and retail distribution systems. Related areas of interest include patterns of consumption, impacts of technology on processing and manufacturing, packaging and transport, data and information systems in the food and agricultural industry, market development, and international trade in food products and agricultural commodities. Business, agricultural, and applied economic applications are encouraged. Acceptable methodologies include survey, review, and critique; analysis and synthesis of previous research; econometric or other statistical analysis; and case studies. Teaching cases will be considered. Issues on special topics may be published based on requests or on the editors' initiative. Potential usefulness to a broad range of agricultural and business economists is an important criterion for publication.

The *Journal of Food Distribution Research* (JFDR) is a publication of the Food Distribution Research Society, Inc. (FDRS). The journal is published three times a year (March, July, and November). JFDR is refereed in its July and November issues. A third, non-refereed issue contains Research Reports and Research Updates presented at FDRS's annual conference. Members and subscribers also receive the Food Distribution Research Society Newsletter, normally published twice a year.

JFDR is refereed by a review board of qualified professionals (see Editorial Review Board, at left). Manuscripts should be submitted to the FDRS editors (see back cover for Manuscript Submission Guidelines).

The FDRS accepts advertising of materials considered pertinent to the purposes of the Society for both the journal and the newsletter. Contact the V.P. for Membership for more information.

Lifetime membership is $400; one-year professional membership is $45; three-year professional membership is $120; student membership is $15 a year; junior membership (graduated in last five years) is $15 and company/business membership is $140.

Food Distribution Research Society
http://www.fdrsinc.org/

Indexing and Abstracting
Articles in *Journal of Food Distribution Research* are indexed or cataloged in Ag Econ Search, Google Scholar, RePEc, Econ Papers, SCOPUS, IDEAS, and CAB International

Send change of address notifications to
Samuel Zapata
Texas AgriLife Extension Service
2401 E. Business 83
Weslaco, TX 78596
Phone: (956) 5581;
e-mail: samuel.zapata@ag.tamu.edu

Food Distribution
Research Society

Journal of Food Distribution Research
Volume 55, Issue 3

Table of Contents

Journal of Food Distribution Research
Volume 55, Issue 3, pp. 1–20

The Effect of Amazon Prime on Sales Ranks

Chinonso E. Etumnu[a Φ] and Ashraf Noumir[b]

[a] *Assistant Professor, School of Agriculture, Environment, and the Sciences,*
400 E. Main St., Kentucky State University,
Frankford, KY 40601, USA

[b] *Lecturer, Faculty of Commerce,*
Damietta University,
Damietta, Egypt

Abstract

This study estimates the effect of the Amazon Prime status of beverage products on their sales ranks. Specifically, using fixed effects regressions and data from Amazon's platforms in the United States and Canada, we first show that Amazon Prime improves grocery sales ranks of ground coffee and black tea by 20% and 17% in the United States and by 8% and 17% in Canada, respectively. Then, we confirmed the validity of these results using product-level sales ranks. These findings suggest that Amazon Prime's economic success is observed in Amazon's marketplaces, as reported in the mass media.

Keywords: black tea; e-commerce; fixed effects; ground coffee

[Φ] Corresponding author: Tel: (415) 205-7395
Email: chinonso.etumnu@kysu.edu

Introduction

In his 2020 letter to shareholders, Jeff Bezos noted that Amazon Prime allows 200 million subscribers worldwide to receive fast and free delivery of millions of eligible online products within two, one, or the same business days (Smith, 2013; Weissmann, 2014; Del Rey, 2019; Hatch, 2020; Bezos, 2021; Spangler, 2021). Due to this logistical prowess, Amazon Prime has become very popular and extremely successful. The success of the subscription program has been widely examined by several media reports, including articles in the New York Times (Ovide, 2022) and the Wall Street Journal (Brown, 2022). Yet, few economic and marketing studies have carefully considered the managerial and regulatory implications of the subscription program (Ramadan, Farah, and Bou Saada, 2021; Snyder, Canaday, and Hughes, 2022), likely because the causal impact of Amazon Prime requires experimental or observational big data, which are not readily available. Given that Amazon is a nongovernmental enterprise, gathering experimental data on Amazon Prime violates the company's terms and conditions and could cause legal problems for researchers. In addition, collecting observational big data from Amazon's websites requires substantial investments in web scraping, and such data may suffer from nonrandom biases that could lead to incorrect research conclusions. In this study, we circumvent these challenges through an online Amazon database and fixed effects regression approaches, thereby addressing these limitations and contributing to the literature on subscription programs in economics and marketing.

There are at least three reasons why it is essential to consider the effect of Amazon Prime on sales. First, Amazon is a global economic powerhouse with billions of monthly visits to its websites across the globe. Therefore, every third-party seller wants to sell on Amazon and thus considers the Amazon Prime program as one of the ways to improve sales on the platform. However, justifying any investments in Amazon Prime-eligible products depends on understanding the subscription program's heterogeneous effects on underperforming, median-performing, and top-performing products. However, to our knowledge, limited economic and marketing studies have quantified the effect of Amazon Prime and ascertained its managerial importance to third-party sellers. Second, anti-trust investigations of Amazon in the United States have occurred (California Department of Justice-Office of the Attorney General, 2022). For example, the Attorney General of California, Rob Bonta, recently filed a lawsuit against Amazon, alleging that the company is stifling competition using its dominance in e-commerce through several mechanisms, including the Amazon Prime program. Thus, because of its regulatory importance, this study seeks to contribute to the understanding of anti-competition implications of the Amazon Prime program. Third, from Amazon's perspective, providing evidence of the effect of Amazon Prime on sales using data collected from Amazon's websites could confirm or dispute the success of the subscription program. Since Amazon is the predominant e-commerce firm, sustaining Amazon Prime's short- and long-term success is necessary for its brand to thrive in the highly competitive retail sector.

Most products sold in Amazon's global marketplaces have a sales rank. This Amazon-assigned number indicates a snapshot of the sales level of a product relative to other products at a particular time (Hanks and Spils, 2006). Thus, the sales rank is a good indicator of a product's sales performance. Amazon assigns a sales rank at the broad category level and the more specific product level. For example, Amazon can assign a sales rank of 2,000 to a ground coffee product at the grocery category level and a sales rank of 1 to the same ground coffee at the ground coffee product level. This example indicates that although the product is the highest-selling ground coffee, it is the 2,000th highest-selling grocery item. This information allows us to estimate the effect of Amazon Prime on the grocery and product sales ranks of ground coffee and black tea beverages, two of the most popular products in the grocery category. We chose both products for this study because they are nonperishable food items with enormous economic contributions, which can serve as gateway products to the emergence of online grocery shopping (Heng et al., 2018; Etumnu et al., 2020).

Using sales ranks as an indicator of sales performance is appropriate because Schnapp and Allwine (2001) have shown that the relationship between log sales and log sales ranks is close to linear. Thus, if we used log sales ranks as our dependent variable, we would adjust our estimation coefficients and standard errors by a constant value (Chevalier and Mayzlin, 2006). In addition, some studies (Chevalier and Mayzlin, 2006; Etumnu et al., 2020; Reimers and Waldfogel, 2021) have successfully used log sales ranks in place of log sales because, as a nongovernmental enterprise, Amazon is not required to share its product sales data with the public. Given the inverse relationship between the sales rank and product sales performance, our main objective is to address whether Amazon Prime products are associated with lower sales ranks (i.e., an increase in sales). This result is expected because e-commerce thrives on cheap, fast, and free delivery of products, which Amazon Prime promises its loyal subscribers (Del Rey, 2019; Brown, 2023). To test our hypotheses, we collected ground coffee and black tea data on sales ranks, Amazon Prime status, customer ratings, and prices from Amazon's platforms in Canada and the United States through an online Amazon database (Keepa.com). With this data, we estimated the effect of Amazon Prime on the sales ranks of beverage products using a fixed effects approach.

Specifically, we first show that Amazon Prime improves the grocery sales ranks of ground coffee and black tea by 20% and 17% in the United States and 8% and 17% in Canada, respectively. Next, we confirmed the validity of the results using product-level sales ranks. These findings suggest that the economic success of Amazon Prime reported in the mass media is observed at the product level on Amazon's websites (Brown, 2022; Ovide, 2022). The results contribute to a growing body of literature that uses internet data and theoretical modeling to investigate several aspects of e-commerce (Edelman, 2012; Einav et al., 2015; Richards, Hamilton, and Empen, 2017; Lu and Reardon, 2018; Harris-Lagoudakis, 2022 ; Reimers and Waldfogel, 2021; Teh, 2022). However, the value that e-commerce firms create in society has led to the growth of these studies. For example, according to the U.S. Department of Commerce-Bureau of the Census (2011, 2021), the e-commerce subsector led by Amazon contributed \$871 billion to the economy in 2021 compared to \$194 billion in 2011, making e-commerce one of the fastest-growing subsectors in the economy of the United States. Thus, this study expands upon a growing body of research by estimating the effect of Amazon Prime, one of the most iconic e-commerce brands in the 21st century.

This study also shows that Amazon Prime increases sales, and thus, it can provide insights that might guide marketing strategies and decisions (Ailawadi, Lehmann, and Neslin, 2003; Gilbert, 2021; Teh, 2022; Teh and Wright, 2022; Etumnu, 2022a). Measuring the success of Amazon Prime in terms of sales ranks only satisfies some of the attributes of an ideal brand equity measure (Ailawadi, Lehmann, and Neslin, 2003). Yet, sales ranks have several advantages that could be appealing to e-commerce managers. For example, the sales rank is an objective measure derived from sales readily available on Amazon's platforms across the globe. Thus, it reflects the overall health of Amazon Prime over time, and e-commerce managers can use it to assess the impact of their marketing strategies and decisions. Furthermore, the availability of databases like Keepa provides a resource for third-party e-commerce managers who may lack the expertise or tools required to monitor their products and those of their competitors continuously. Thus, the challenge of identifying and monitoring one's and competitors' Amazon Prime products can be quickly and adequately addressed through Keepa and similar databases and trackers, such as JungleScout and camelcamelcamel.

An Overview of Amazon Prime

Amazon Prime is a paid subscription program that, as of April 2021, gives 200 million subscribers exclusive access to additional services offered by Amazon (Howley, 2021; Spangler, 2021). These services include same-, one-, or two-day delivery of purchased products and access to video, music, reading, gaming, photos, grocery shopping, and exclusive deals such as Prime Day (Amazon 2022a). Amazon Prime is believed to be one of the most iconic retail inventions in the world, which has led to unprecedented success for Amazon (Brown, 2022; Ovide, 2022). Although there are several media reports on the economic success of Amazon Prime, only some empirical studies focus on these success stories (Ramadan, Farah, and Bou Saada, 2021; Snyder, Canaday, and Hughes, 2022). Given the rich history of Amazon Prime and the need to provide some context for this study, an overview of the subscription program follows.

Amazon Prime was launched in 2005 as a membership service in the United States, offering two-day free shipping on eligible products within the country (Del Rey, 2019). This service was initially provided to Amazon customers for an annual fee of $79, which increased to $139 as of December 31, 2022 (Weissmann, 2014; Del Rey, 2019; Amazon, 2022a). After launching in the United States, Amazon subsequently implemented the subscription service in other countries, starting with Germany and the United Kingdom in 2007 and Canada in 2013 (Smith, 2013; Weissmann, 2014; Amazon, 2022b). As of October 2021, Amazon Prime memberships are available to Amazon customers in 23 countries across the globe (see Table 1).

Table 1. What Is Amazon Prime?

Type of program	Subscription service
Founded	February 2, 2005
Revenue	$31.8 billion (2021)
Number of subscribers	200 million
Services offered	Fast, free delivery
	Prime video
	Prime music
	Prime day

Table 1. (cont)

Services offered	Prime gaming Exclusive deals Rx savings Prime reading Amazon photos Prime try before you buy Free Grubhub+ for a year
Countries available	Austria, Australia, Belgium, Brazil, Canada, China, France, Germany, India, Ireland, Italy, Japan, Luxembourg, Mexico, the Netherlands, Poland, Portugal, Saudi Arabia, Singapore, Spain, Sweden, Turkey, the UK, and the US
Subscription plans (US)	Prime monthly ($14.99) Prime annual ($139) Prime student monthly ($7.49) Qualified government assistance ($6.99)
Subscription plans (CA)	Prime monthly (CDN$9.99) Prime annual (CDN$99.00) Prime student monthly (CDN4.99)

Since its launch, Amazon Prime has contributed to Amazon's net sales—total revenue minus sales returns, allowances, and discounts (Kenton, 2024). The other sources of consolidated net sales include online stores, physical stores, third-party seller services, advertising services, Amazon Web Services (AWS), and others. However, records of Amazon Prime's contribution to net sales began to appear in Amazon's annual report in 2014. Amazon Prime (included under subscription services in the annual reports) contributed $2.8 billion in fees, which represents about 3.1% of Amazon's consolidated net sales in 2014 (Amazon, 2022c). As of 2021, this contribution level has increased almost 12 times to $31.8 billion, representing 6.8% of Amazon's total net sales (Amazon, 2022c). Despite this tremendous increase, Amazon Prime subscription fees still represent a small portion of Amazon's total revenue. In 2021, the major revenue contributors were online stores ($222 billion, 47%) and AWS ($62 billion, 13%) (Amazon, 2022c).

However, in recent years, Amazon Prime has also contributed to sales through Prime Day—an annual two-day sales event exclusive to Amazon Prime members (Amazon, 2022d). The event started in 2015, and its eighth edition took place on July 12–13, 2022. The 2022 event sold more than 300 million items, resulting in $1.7 billion in savings for Amazon Prime members (Amazon, 2022d). Although Amazon did not disclose the total sales from the 2022 Prime Day event, it was estimated to have generated approximately $12 billion (Morris, 2022; Reuters, 2022; Walk-Morris, 2022). Additionally, a survey of Amazon shoppers in the United States by Bank of America found that the average Amazon Prime member spent $1,968 per year on Amazon (Bain, 2021). This estimate is about four times the amount spent on Amazon by non-Prime members (Bain, 2021).

Therefore, Amazon Prime's contribution to Amazon's total revenue now includes three channels: subscription fees, Prime Day sales, and regular sales from Prime members, making it one of the main foundations of Amazon's business and of interest not only to academics but also to third-party sellers and regulators (Bain, 2021).

Two different fulfillment methods operate on Amazon for third-party sellers: (i) Fulfilled by Amazon (FBA)—independent third-party sellers who sell their products to the final consumer but pay Amazon a fulfillment fee to handle its sale, from storing and delivering and customer service to a possible return of the products, and (ii) fulfilled by Merchant (FBM)—independent third-party sellers who sell on Amazon without using the company's logistics. Amazon designates its products and those of FBA sellers as Amazon Prime, whereas the products of FBM sellers are undesignated as Amazon Prime.

Data

Our data consist of ground coffee and black tea attributes collected from the public websites of Amazon.com and Amazon.co.ca through the Amazon database Keepa. Our goal is to estimate the effect of a product's Amazon Prime status on its sales. However, we could not access their ground coffee and black tea sales data because Amazon is a non-governmental enterprise. Thus, we relied on the sales ranks reported on Amazon's websites in the United States and Canada. We used the sales ranks as they are in our analysis without attempting to approximate or derive ground coffee and black tea sales from them. We collected data on all ground coffee and black tea products on Keepa because the entire population of ground coffee and black tea products sold on Amazon's websites in the United States and Canada was less than 10,000, the maximum allowable number of products on Keepa's product finder.

We collected the ground coffee and black tea products on October 5, 2022, and October 5, 2023. We also obtained each product's current price, grocery and product level sales ranks, average star rating, the number of ratings, stockout rate, and seller type. We created an Amazon Prime variable using the seller type (Amazon, FBA, and FBM). In addition, we merged the datasets that appeared in the two periods for each country and product, creating panel datasets. In the United States, our panels comprise 5,931 products (1,1862 observations) for ground coffee and 5,355 products (10,710 observations) for black tea. In Canada, our panel samples comprise 1,014 products (2,028 observations) for ground coffee and 1,550 products (3,100 observations) for black tea. We attribute the significant differences in sample size across the two countries to at least three reasons. First, Amazon was founded in the United States in 1994 and in Canada in 2002. The lag between starting Amazon and establishing the Canadian branch correlates with how the company expends its resources, such as labor, marketing, and research and development. These factors bolster the success of Amazon United States compared to Amazon Canada. Second, the population of the United States (333 million) is almost 9 times more than the population of Canada (39 million), which implies that the market size of Amazon products in the United States, holding other factors constant, can be assumed to be 9 times the market size in Canada. With a higher beverage demand due to the U.S. population size, Amazon and its third-party sellers meet this demand by providing more products. Finally, the level of competition among sellers in the United States appears fiercer than in Canada. With more than 1 million third-party sellers competing for market share in the United States compared to about 58,000 in Canada, the number of sellers in each country reflects the number of products available (Chevalier, 2022; Keepa, 2024).

Tables 2 and 3 present the summary statistics of the main variables. The tables also compare the means of the variables and indicate whether or not they are Amazon Prime eligible. In Table 2 (United States), 67% of the ground coffee products and 46% of the black tea products are Amazon Prime eligible. However, in Table 3, 45% of the ground coffee products and 36% of the black tea products are Amazon Prime eligible. Aside from these percentages, comparing eligible and ineligible products reveals at least four relevant differences to our study.

First, the sales ranks of eligible products are significantly less than those of ineligible products. These differences suggest that the Amazon Prime products have better sales performance than the ineligible products. However, aside from the influence of the free and fast shipping enjoyed by Amazon Prime products, other factors, such as price, consumer ratings, and stockout rates, could also play a role in determining the differences in sales ranks. Second, the prices of Amazon Prime-eligible products are lower than those of Amazon Prime-ineligible products. Although these prices do not control for package size, they do reflect the law of demand, which suggests that the lower prices of Amazon Prime-eligible products command higher demand (sales) than those of the ineligible products. Third, the average rating and the number of ratings indicate that Amazon Prime-eligible products have better consumer-perceived quality and popularity in Amazon's marketplaces.[1] Finally, the stockout rates of Amazon Prime-eligible products are much lower than those of ineligible products, which suggests that eligible products may have higher sales because stockouts on Amazon's marketplaces are correlated with sales (Etumnu, Jaenicke, and Cheranades, 2024).

[1]This observation seems true for all products except ground coffee in Canada.

Table 2. Summary Statistics (United States)

	Ground Coffee				Black Tea			
	All	**Eligible**	**Ineligible**	***p*-value**	**All**	**Eligible**	**Ineligible**	***p*-value**
Amazon Prime	0.67	1	0	0.0000	0.46	1	0	0.0000
Grocery sales ranks	90,286	73,930	123,176	0.0000	249,713	144,195	339,816	0.0000
Coffee sales ranks	1,462	1,181	2,027	0.0000	2,808	1,604	3,835	0.0000
BuyBox price	$27.26	$23.96	$33.89	0.0000	$25.77	$22.02	$28.96	0.0000
Average rating	4.14	4.19	4.03	0.0000	3.17	3.86	2.58	0.0000
Number of ratings	1,993	2,187	1,603	0.0000	1,501	2,549	606	0.0000
Stockout rate	0.14	0.05	0.32	0.0000	0.12	0.04	0.19	0.0000
Number of products	5,931			5,355				
Observations	11,862			10,710				

Table 3. Summary Statistics (Canada)

	Ground Coffee				Black Tea			
	All	**Eligible**	**Ineligible**	***p*-value**	**All**	**Eligible**	**Ineligible**	***p*-value**
Amazon Prime	0.45	1	0	0.0000	0.36	1	0	0.0000
Grocery sales ranks	30,238	14,334	43,183	0.0000	30,317	16,505	38,216	0.0000
Tea sales ranks	1,109	446	1,649	0.0000	1,932	635	2,674	0.0779
BuyBox price	CA$50.56	CA$33.00	CA$64.85	0.0000	CA$40.02	CA$25.39	CA$48.40	0.0000
Average rating	3.98	4.12	3.86	0.0000	3.02	3.49	2.75	0.0000
Number of ratings	2,484	2,300	2,634	0.3293	1,457	2,353	944	0.0000
Stockout rate	0.19	0.82	0.28	0.0000	0.17	0.06	0.23	0.0000
Number of products		1,014				1,550		
Observations		2,028				3,100		

Although the comparison of means suggests that being Amazon Prime eligible has a causal impact, regression analysis is necessary to disentangle the effects of Amazon Prime status from the effects of any confounding factors. Thus, in the next section, we present the fixed-effects regressions used in the causal analysis.

Empirical Strategy

This section presents the fixed effects model that we employ to estimate the effects of Amazon Prime on grocery and product-level sales ranks. Consider the following fixed effects model:

$$y_{it} = \alpha + \beta Prime_{it} + \gamma \boldsymbol{X_{it}} + \rho_t + \sigma_i + \varepsilon_{it} \tag{1}$$

where i could be a ground coffee or black tea product in the United States or Canada. t is the period for data collection, with $t = 1$ representing October 2022, and $t = 2$ representing October 2023. y_{it} is the natural logarithm of the grocery sales ranks or the product-level (ground coffee or black tea) sales ranks of product i in period t. We prefer a log-transformed dependent variable for two reasons. First, the sales rank is an ordinal variable with an extensive range. For example, the grocery sales ranks range from 1 to hundreds of thousands. Thus, a log transformation of the sales ranks increases our chances of normalizing the distribution of the variable. Second, we enhance the interpretability of the effect of Amazon Prime as percentage change through the log transformation. Given the range and ordinal nature of sales ranks, this method of interpreting the effects seems much more intuitive. $Prime_{it}$ is a dummy variable indicating whether product i is eligible for Amazon Prime status in period t. We hypothesize that β is negative, suggesting that being eligible for Amazon Prime lowers (improves) the sales ranks. To ensure that the effects of Amazon Prime are disentangled from cofounding factors, we also control for other factors in the vector, $\boldsymbol{X_{it}}$. The variables in the vector include price, consumer ratings, and stockout rate. We hypothesize that increasing prices and stockout rates hurt sales ranks, whereas increasing consumer ratings improves sales ranks. ρ_t, σ_i, and ε_{it} are the time-fixed effects, product-fixed effects, and error terms, respectively.

Our specification in equation (1) has several advantages that enable us to identify the causal effects of Amazon Prime. However, it also has limitations, which we cannot address in this study. We first present the advantages and then the limitations. The most significant advantage of our model is that it includes time and product-fixed effects. Thus, we assume that our estimated Amazon Prime effects are causal conditional on the fixed effects. The time-fixed effects control for macroeconomic and weather-related issues that could influence the sales ranks (Ebbes, Papies, and van Heerde, 2021). Additionally, product-fixed effects control for omitted time-invariant attributes that could affect sales ranks (Ebbes, Papies, and van Heerde, 2021). Another advantage of our model is that it includes relevant control variables in Amazon's marketplaces. These variables include price, average rating, number of ratings, and stockout rate. Previous studies on Amazon have shown these variables to be relevant (Chevalier and Mayzlin, 2006; Sun, 2012; Etumnu, Jaenicke, and Cheranades, 2024). Hence, not controlling them in our model will lead to omitted variable bias.

Despite these advantages, we made one assumption in our fixed effects model that may be too optimistic. That is, we assumed that Amazon Prime is an exogenous variable. However, in a few situations Amazon Prime could be endogenous. For example, the probability of becoming Amazon Prime eligible may be due to observed product sales, while Amazon Prime impacts sales, leading to a reverse causality. Another example could be due to omitted variable bias, whereby unobserved attributes determine third-party seller enrollment in the FBA program. To correct this potential endogeneity challenge, the literature suggests using instrumental variables regressions (Angrist and Pischke, 2009; Cunningham, 2021; Ebbes, Papies, and van Heerde, 2021). However, we are cautious of using instrumental variables in this study as they may produce worse outcomes than our proposed fixed effects model for two reasons. The first is meeting the relevance criterion of an instrumental variable, and the second is meeting the validity criterion of an instrumental variable. Because of these challenges, we focused on our panel fixed effects regressions despite acknowledging the potential caveats of the model.

Results and Discussions

This section presents the study's results. Table 4 presents the effects of Amazon Prime in the United States, whereas Table 5 presents the effects of Amazon Prime in Canada. We also made two decisions that increased the generalizability of our findings. First, we estimated the effects of Amazon Prime in both the United States and Canada. Finding similar results in both countries suggests that Amazon Prime's effects may be universal. Second, we estimate the effect of Amazon Prime not only on the grocery sales ranks, but also on the product-level (ground coffee and black tea) sales ranks in the United States and Canada. If the results of the effects of Amazon Prime on the grocery sales ranks are like the effects of Amazon Prime on the product-level sales ranks, the validity of our findings will be enhanced. We present the specific results next, starting with the United States and then Canada.

Table 4 shows results in the United States. The first column presents the effect of Amazon Prime and the control variables on grocery sales ranks of ground coffee. We find that being Amazon Prime eligible reduces (improves) grocery sales ranks by about 20%. The second column presents the effect of Amazon Prime and control variables on ground coffee sales ranks. We find that Amazon Prime improves ground coffee sales ranks by about 15%. The third and fourth columns focus on black tea, presenting the effects of Amazon Prime on grocery sales ranks and black tea sales ranks, respectively. The results show that grocery sales ranks improved by about 17%, whereas black tea sales ranks improved by about 14%. The results found in each column confirm the positive effects of Amazon Prime on sales ranks and, thus, sales.

Aside from Amazon Prime, the effects of the control variables also have the expected signs. For example, the impact of price on sales ranks is positive, although with different effect sizes for each column. These results indicate (i) the positive relationship between price and sales ranks suggests an increase in price hurt sales ranks, leading to a downward-sloping demand curve. This also confirms the law of demand; and (ii) the price-sales ranks relationship suggests that the demand for ground coffee and black tea is inelastic because the elasticities range from 0.37 to 0.46. The second set of control variables are consumer ratings—average rating and number of ratings. As hypothesized, an increase in each consumer rating variable improves sales ranks.[2] The positive impacts occur because consumer ratings address information asymmetry. The average rating informs consumers of the perceived quality of the products as well as their associated services, such as delivery time, returns, and customer service. The number of ratings addresses information asymmetry by indicating which products are popular, thus increasing their visibility to other consumers. Both variables also reduce the possibility of omitted variable bias given how important consumer ratings are in online marketplaces, especially Amazon. The last control variable is the stockout rate, which also has the expected sign. An increase in the stockout rate hurts sales ranks of ground coffee and black tea products with a range of 24%–49%. This finding is also consistent with previous studies that show that stockout rates are crucial for the success of retailers, including Amazon and its third-party sellers.

Table 5 presents four columns that show the effects of Amazon Prime and control variables on ground coffee and black tea products in Canada. The effect of Amazon Prime on the columns ranges from 8%–21%. Specifically, in column 1, Amazon Prime improves grocery sales ranks for ground coffee by 8%, although insignificantly. In column 2, which also focuses on ground coffee, the effect is 8%, again insignificant. However, for black tea products, the effect of Amazon Prime on grocery sales ranks and black tea sales ranks are 17% and 21%, respectively. Both are statistically significant at $p < 0.01$. Each of these results confirms that Amazon Prime improves sales ranks, and thus, sales.

Furthermore, like the results found in the United States, the effect of each of the control variables has the expected signs. First, the price elasticity ranges from 0.37 to 0.46. Second, the effect of average rating on ground coffee ranges from 5%–7%, but is insignificant for black tea products. Third, the number of ratings improves sales ranks, with a range of 32%–40%. Finally, an increase in the stockout rate hurts sales ranks by 24%–49%. These control variables are crucial to addressing omitted variable biases.

[2]The effects of average rating on black tea are insignificant.

Table 4. Effect of Amazon Prime on Sales Ranks in the United States

	Ground Coffee		Black Tea	
	Grocery Sales Ranks (1)	**Coffee Sales Ranks (2)**	**Grocery Sales Ranks (3)**	**Tea Sales Ranks (4)**
Amazon Prime	-0.1998***	-0.1543***	-0.1652***	-0.1390***
	(0.0350)	(0.0311)	(0.0257)	(0.0282)
Log price	0.6799***	0.6233***	0.0949**	0.0821
	(0.0654)	(0.0581)	(0.0480)	(0.0525)
Average rating	-0.0457***	-0.0454***	-0.0096*	-0.0146***
	(0.0106)	(0.0094)	(0.0050)	(0.0055)
Number of ratings	-2.0314***	-1.6667***	-2.1183***	-1.7189***
	(0.0359)	(0.0319)	(0.0477)	(0.0522)
Stockout rate	0.5774***	0.5360***	0.3106***	0.2957***
	(0.0532)	(0.0473)	(0.0344)	(0.0377)
Constant	8.9033***	5.0398***	11.6624***	7.1556***
	(0.2112)	(0.1876)	(0.1486)	(0.1626)
N	1,1862	1,1862	10,710	1,0710
R2	0.9043	0.8922	0.9491	0.9316
Fixed effects	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Single, double, and triple asterisks (*, **, ***) indicate statistical significance at the 10%, 5%, and 1% levels.

Table 5. Effect of Amazon Prime on Sales Ranks in Canada

	Ground Coffee		Black Tea	
	Grocery Sales Ranks (1)	**Coffee Sales Ranks (2)**	**Grocery Sales Ranks (3)**	**Tea Sales Ranks (4)**
Amazon Prime	-0.0808	-0.0800	-0.1680***	-0.2062***
	(0.0678)	(0.0679)	(0.0413)	(0.0553)
Log price	0.4561***	0.3749***	0.3808***	0.4046***
	(0.1032)	(0.1034)	(0.0480)	(0.0642)
Average rating	-0.0711***	-0.0470**	0.0001	-0.0167
	(0.0219)	(0.0220)	(0.0079)	(0.0106)
Number of ratings	-0.3971***	-0.3228***	-0.3591***	-0.3322***
	(0.0667)	(0.0668)	(0.0749)	(0.1002)
Stockout rate	0.4898***	0.3577***	0.2373***	0.2903***
	(0.0832)	(0.0834)	(0.0489)	(0.0654)
Constant	8.1116***	4.7642***	8.4974***	5.1188***
	(0.3946)	(0.3954)	(0.1712)	(0.2288)
N	2,028	2,028	3,100	3,100
R_2	0.9163	0.8909	0.9303	0.9004
Fixed effects	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Single, double, and triple asterisks (*, **, ***) indicate statistical significance at the 10%, 5%, and 1% levels.

Based on the existing economic and marketing literature, there are at least three reasons why Amazon Prime improves product sales performance. First, online shoppers long for convenience to save time and to reduce their search costs, and Amazon Prime provides them with these benefits through its free shipping and free returns programs (Shehu, Papies, and Neslin, 2020; Patel et al., 2021; Etumnu, 2022a). Although the free shipping and free returns offered through Amazon Prime are not actually free—Amazon Prime members pay subscription fees—the benefits of the programs seem to outweigh the costs of the subscription fees. These benefits thus translate to more sales for products that have the Amazon Prime label. Second, Amazon has succeeded in building a solid reputation in retail (Marcotte, 2022). For example, in 2023, Amazon was ranked as the best place to work by LinkedIn and the second most admired company in the world by Fortune (Fortune, 2023; LinkedIn, 2023). Amazon's solid reputation, particularly its brand equity, brand visibility, and brand trust, significantly boosts sales. Third, Amazon Prime might be boosting sales because Amazon practices self-preference for its brand (European Commission, 2022; Farronato, Fradkin, and MacKay, 2023). Self-preferencing has been highlighted as one way through which digital platforms could stifle competition in the marketplace. Whether this is the case for Amazon Prime was raised by the European Commission (2022), a case which was eventually settled when Amazon made some commitments, including how to utilize seller data, "featured offer," and Amazon Prime. Self-preferencing has also been highlighted in terms of rankings of search results (Farronato, Fradkin, and MacKay, 2023). However, it is beyond the scope of this work to investigate whether Amazon self-prefers the Amazon Prime program in its platform and the reasons why the company might be involved in such practice.

Overall, our results suggest that every serious third-party seller on Amazon should strongly consider the Amazon Prime label. As we show, Amazon Prime improves sales across countries and products and appears to outweigh the program's costs.

Conclusions and Implications

Although several media reports have ascertained that Amazon Prime is economically successful, there has been a limited empirical investigation of the subscription program. Using a fixed-effects approach, this study quantifies the effect of Amazon Prime on the sales ranks of ground coffee and black tea products in the United States and Canada. We first show that Amazon Prime improves grocery sales ranks of ground coffee and black tea by 20% and 17% in the United States and by 8% and 17% in Canada, respectively. We also confirmed the validity of these results using product-level sales ranks. These findings suggest that the economic success of Amazon Prime reported in the mass media is observed at the product level on Amazon's marketplaces.

This study is potentially useful to Amazon, third-party sellers that operate on its online marketplaces, and policy makers and regulators who promulgate and execute innovation and competition laws. For Amazon, the results indicate that their iconic subscription program creates value for them as well as third-party sellers that adopt the Amazon Prime program. The results indicate that these third-party sellers are obtaining the value they pay for, which gives their products the privilege to become Amazon Prime-eligible and an advantage over Prime-ineligible products. This study could spur other sellers operating on Amazon's marketplace (Fulfilled by

Merchant—FBM) to contemplate adopting the FBA sales strategy (Lai, Liu, and Xiao, 2018; Etumnu, 2022b). But deciding whether to adopt the FBA sales strategy should entail a case-by-case analysis to weigh its potential costs and benefits.

For policy makers and regulators, there remains a question of whether the Amazon Prime business strategy is anticompetitive (Zhu and Liu, 2018; Hagiu, Teh, and Wright, 2020; Competitions and Market Authority [CMA], 2022), because products that are Amazon Prime eligible have an advantage over other products on Amazon's marketplaces. Whether this advantage is unfair is debatable. In addition, questions have been raised about how Amazon sets the eligibility criteria for third-party sellers to use the Amazon Prime label (CMA, 2022). But it is worth noting that Amazon Inc. defeated the Attorney General (AG) of Washington DC in court regarding a complaint that accused Amazon of stifling competition through its business strategies including Amazon Prime (Lamm, 2022). However, Amazon's victory over the AG has been short-lived as a U.S. anti-competition bill (the American Innovation and Online Choice Act) and the United Kingdom's Competition and Market Authority have recently targeted Amazon again (Baer, 2022; CMA, 2022; Huseman, 2022). Thus, this study contributes to an important and active debate on Amazon's dominance as a dual player (retailer and marketplace) in the growing e-commerce market.

Despite the usefulness of estimating the effect of Amazon Prime eligibility on sales ranks, e-commerce managers should be aware of the shortcomings of our approach. First, the valuation of Amazon Prime in terms of sales ranks is only one aspect of the entire valuation of Amazon Prime. Amazon Prime allows for subscriptions to e-books and videos, and the value Prime provides in those avenues was not included in this analysis (see Table 1). Thus, the estimate of the value of Prime, which is enormous but still incomplete, should only be interpreted for the products markets. Another shortcoming of this analysis for e-commerce managers is that valuing Amazon Prime in terms of sales ranks does not provide subjective reasons for diagnosing a brand. These reasons might include customer loyalty, pre-commitment, awareness, and attitudes toward Amazon Prime, which might be better captured through surveys, focus group discussions, in-depth interviews, and experiments (Ailawadi, Lehmann, and Neslin, 2003; Aaker, 2009; Bronnenberg, Dube, and Moorthy, 2019). Therefore, we recommend that e-commerce managers value Amazon Prime in terms of sales ranks and other metrics to create a more comprehensive picture of consumer behavior.

Furthermore, no financial value was assigned to capture the value of Amazon Prime. Corporations like financial values, which they can include in their financial statements, balance sheets, and reports to reflect the contribution of Amazon Prime and justify its use at the product level. Therefore, future research can focus on assigning a financial value (like changes in revenues or profits) to the Amazon Prime brand. Such value could be used to assess the current and future health of Amazon Prime. Future research can also extend this study to other e-commerce brands and subscription programs, such as Walmart+. As major e-commerce giants continue to compete and be dual players as a retailer and marketplace, assessing the viability and longevity of their most successful brands becomes even more crucial.

References

Aaker, D.A. 2009. *Managing Brand Equity*. New York, NY: The Free Press.

Ailawadi, K.L., D.R. Lehmann, and S.A. Neslin. 2003. "Revenue Premium As an Outcome Measure of Brand Equity." *Journal of Marketing* 67(4):1–17.

Amazon. 2022a. *Amazon.com: Amazon Prime*. Available online: https://www.amazon.com/amazonprime.

Amazon. 2022b. *History and Timeline: Amazon.Co.Uk Milestones*. Available online: https://amazonuk.gcs-web.com/history-timeline.

Amazon. 2022c. *2021 Annual Report*. Available online: https://ir.aboutamazon.com/annual-reports-proxies-and-shareholder-letters/default.aspx.

Amazon. 2022d. *Prime Day 2022 Was the Biggest Prime Day Event Ever. Here Are Some of the Best-Selling Items*. Available online: https://advertising.amazon.com/blog/prime-day-2022-best-selling-products#:~:text=Prime%20members%20purchased%20more%20than,Day%20event%20in%20Amazon's%20history.

Amazon. 2023. *Seller Fulfilled Prime*. Available online: https://sell.amazon.com/programs/seller-fulfilled-prime.

Angrist, J.D., and J.S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Baer, B. 2022. "Why Amazon Is Wrong about the American Innovation and Online Choice Act." *Brookings*. Available online: https://www.brookings.edu/blog/techtank/2022/06/14/why-amazon-is-wrong-about-the-american-innovation-and-online-choice-act/.

Bain, M. 2021. "Prime Has Never Been More Important to Amazon." *Quartz*. Available online: https://www.yahoo.com/tech/prime-never-more-important-amazon-170652103.html.

Bezos, J. 2021. *2020 Letter to Shareholders*. Available online: https://www.aboutamazon.com/news/company-news/2020-letter-to-shareholders.

Bronnenberg, B.J., J.-P. Dube, and S. Moorthy. 2019. "The Economics of Brand and Branding." In J.-P. Dube and P.E. Rossi, eds. *Handbook of the Economics of Marketing*. Amsterdam, Netherlands: North Holland, pp. 291-358.

Brown, D. 2022. "Amazon Prime Price Increase: Here's Why It's Still Worth It." *Wall Street Journal*. Available online: https://www.wsj.com/articles/amazon-prime-price-increase-value-11644011470?mod=Searchresults_pos1&page=1 [Accessed August 10, 2022].

Brown, R. 2023. *Amazon Prime Timeline: 2005–Present*. Available online: https://uk.pattern.com/blog/amazon-prime-timeline-to-present/.

California Department of Justice, Office of the Attorney General. 2022. *Attorney General Bonta Announces Lawsuit against Amazon for Blocking Price Competition*. Available online: https://oag.ca.gov/news/press-releases/attorney-general-bonta-announces-lawsuit-against-amazon-blocking-price.

Chevalier, S. 2022. "U.S. Amazon.com Third-party Seller Metrics 2019." *Statista.* Available online: https://www.statista.com/statistics/1086637/amazoncom-3p-seller-metrics-usa/.

Chevalier, J.A., and D. Mayzlin. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews." *Journal of Marketing Research* 43:245–354.

Competitions and Market Authority. 2022. *CMA Investigates Amazon Over Suspected Anti-competitive Practices*. Available online: https://www.gov.uk/government/news/cma-investigates-amazon-over-suspected-anti-competitive-practices.

Cunningham, S. 2021. *Causal Inference: The Mixtape*. New Haven, CT: Yale University Press.

Del Rey, J. 2019. "The Making of Amazon Prime, the Internet's Most Successful and Devastating Membership Program." *Vox.* https://www.vox.com/recode/2019/5/3/18511544/amazon-prime-oral-history-jeff-bezos-one-day-shipping?curator=TechREDEF.

Ebbes, P., D. Papies, and H.J. van Heerde. 2021. "Dealing with Endogeneity: A Nontechnical Guide for Marketing Researchers." In A. Vomberg, C. Homburg, and M. Klarmann, eds. *Handbook of Market Research.* New York, NY: Springer.

Edelman, B. 2012. "Using Internet Data for Economic Research." *Journal of Economic Perspectives* 26(2):189–206.

Einav L., T. Kuchler, J. Levin, and N. Sundaresan. 2015. "Assessing Sale Strategies in Online Markets Using Matched Listings." *American Economic Journal: Microeconomics* 7(2):215–247.

Etumnu, C.E. 2022a. "Free Shipping." *Applied Economics Letters* 30(16):1–4.

Etumnu, C.E. 2022b. "A Competitive Marketplace or an Unfair Competitor? An Analysis of Amazon and Its Best Sellers Ranks." *Journal of Agricultural Economics* 73(3):924–937.

Etumnu, C.E., K. Foster, N.O. Widmar, J.L. Lusk, and D.L. Ortega. 2020. "Does the Distribution of Ratings Affect Online Grocery Sales? Evidence from Amazon." *Agribusiness* 4:501–521.

Etumnu, C.E., E. Jaenicke, and L. Cheranades. 2024. "The Impact of Stockouts on Sales Performance across Seller Types in Amazon's Marketplaces." Manuscript submitted for publication.

European Commission. 2022. *Antitrust: Commission Accepts Commitments by Amazon Barring It from Using Marketplace Seller Data, and Ensuring Equal Access to Buy Box and Prime*. Available online: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7777.

Farronato, C., A. Fradkin, and A. MacKay. 2023. "Self-Preferencing at Amazon: Evidence from Search Rankings." In W.R. Johnson and T. Lippert, eds. *AEA Papers and Proceedings*. Nashville, TN: American Economic Association, vol. 113, pp. 239–243.

Fortune. 2023. *World's Most Admired Companies*. Available online: https://fortune.com/ranking/worlds-most-admired-companies/.

Gilbert, R. 2021. "Separation: A Cure for Abuse of Platform Dominance?" *Information Economics and Policy* 54:100876.

Hagiu, A., T. Teh, and J. Wright. 2020. "Should Platforms Be Allowed to Sell on Their Own Marketplaces?" *The RAND Journal of Economics* 53(2):297–327.

Hanks, S., and D. Spils. 2006. *Increases in Sales Rank As a Measure of Interest*. US Patent 7,058,599. Available online: https://patents.google.com/patent/US7058599B2/en.

Harris-Lagoudakis, K. 2022. "Online Shopping and the Healthfulness of Grocery Purchases." *American Journal of Agricultural Economics* 104(3):1050–1076.

Hatch, G. 2020. "What You Need to Know about Amazon Prime: 2005–Today." *Pattern*. Available online: https://pattern.com/blog/amazon-prime-a-timeline-from-2005-to-2020/.

Heng Y., Z. Gao, Y. Jiang, and X. Chen. 2018. "Exploring Hidden Factors behind Online Food Shopping from Amazon Reviews: A Topic Mining Approach." *Journal of Retailing and Consumer Services* 42:161–168.

Howley, D. 2021. "Amazon Prime Now Has 200 Million Members, Jumping 50 Million in One Year." *Yahoo!Finace*. Available online: https://news.yahoo.com/amazon-prime-has-200-million-members-142910961.html.

Huseman, B. 2022. "Antitrust Legislation and the Unintended Negative Consequences for American Consumers and Small Businesses." *Amazon Policy News and Views*. Available online: https://www.aboutamazon.com/news/policy-news-views/antitrust-legislation-and-the-unintended-negative-consequences-for-american-consumers-and-small-businesses.

Keepa. 2024. *Top Seller List: Canada*. Available online: https://keepa.com/#!topseller.

Kenton, W. 2024. "Net Sales: What They Are and How to Calculate Them." *Investopedia*. Available online: https://www.investopedia.com/terms/n/netsales.asp.

Lai, G., H. Liu, and W. Xiao. 2018. *Fulfilled by Amazon: A Strategic Perspective of Competition at the E-commerce Platform*. Available online: https://ssrn.com/abstract=3270958.

Lamm, G. 2022. "DC Court Won't Revive AG's Antitrust Suit Against Amazon." *Law360*. Available online: https://www.law360.com/articles/1517313.

LinkedIn. 2023. *Top Companies 2023: The 50 Best Workplaces to Grow Your Career in the U.S.* Available online: https://www.linkedin.com/pulse/top-companies-2023-50-best-workplaces-grow-your-career-us/.

Lu, L., and T. Reardon. 2018. "An Economic Model of the Evolution of Food Retail and Supply Chains from Traditional Shops to Supermarkets to E-commerce." *American Journal of Agricultural Economics* 100(5):1320–1335.

Marcotte, D. 2022. *A Look at the 2022 Top 50 Global Retailers*. National Retail Federation. Available online: https://nrf.com/blog/look-2022-top-50-global-retailers.

Morris, C. 2022. "Amazon Touts 2022 Prime Day As 'Biggest Ever.' Here Are the Items That Sold the Most." *Fortune.* Available online: https://fortune.com/2022/07/14/amazon-2022-prime-day-biggest-ever-12-billion/.

Ovide, S. 2022. "Amazon Prime: Loved at Almost Any Price." *New York Times*. Available online: https://www.nytimes.com/2022/01/12/technology/amazon-prime-price.html.

Patel, P.C., C. Baldauf, S. Karlsson, and P. Oghazi. 2021. "The Impact of Free Returns on Online Purchase Behavior: Evidence from an Intervention At an Online Retailer." *Journal of Operations Management* 67(4):511–555.

Ramadan, Z., M.F. Farah, and R. Bou Saada. 2021. "Fooled in the Relationship: How Amazon Prime Members' Sense of Self-Control Counter Intuitively Reinforces Impulsive Buying Behavior." *Journal of Consumer Behaviour* 20(6):1497–1507.

Reimers, I., and J. Waldfogel. 2021. "Digitization and Pre-Purchase Information: The Causal and Welfare Impacts of Reviews and Crowd Ratings." *American Economic Review* 111(6):1944–1971.

Reuters. 2022. *U.S. Online Sales during Amazon's Prime Day Jump to $12 bln- Report.* Available online: https://www.reuters.com/technology/us-online-sales-during-amazons-prime-day-jump-12-bln-report-2022-07-14/.

Richards, T.J., S.F. Hamilton, and J. Empen. 2017. "Attribute Search in Online Retailing." *American Journal of Agricultural Economics* 99(1):1–18.

Schnapp, M., and T. Allwine. 2001. *Mining of Book Data from Amazon.com*. Presented at the UCB/SIMS Web Mining Conference.

Shehu, E., D. Papies, and S.A. Neslin. 2020. "Free Shipping Promotions and Product Returns." *Journal of Marketing Research* 57(4):640–658.

Smith, M. 2013. *Amazon Prime Arrives in Canada: Free Two-Day Shipping, No Instant Video*. Available online: https://www.engadget.com/2013-01-08-amazon-prime-canada-no-instant-video.html.

Snyder, E.A., J. Canaday, and M. Hughes. 2022. "Amazon's Three Major Lines of Business." Chicago, IL: George J. Stigler Center for the Study of the Economy and the State Working Paper 319.

Spangler, T. 2021. "Amazon Prime Tops 200 Million Members, Jeff Bezos Says." *Variety*. Available online: https://variety.com/2021/digital/news/amazon-prime-200-million- jeff-bezos-1234952188/.

Sun, M. 2012. "How Does the Variance of Product Ratings Matter?" *Management Science* 58(4):696–707.

Teh, T. 2022. "Platform Governance." *American Economic Journal: Microeconomics* 14(3):213–254.

Teh, T., and J. Wright. 2022. "Intermediation and Steering: Competition in Prices and Commissions." *American Economic Journal: Microeconomics* 14(2):281–321.

U.S. Department of Commerce. 2011. *Quarterly Retail E-Commerce Sales: 4th Quarter 2011*. Washington, DC: U.S. Department of Commerce, Bureau of the Census.

U.S. Department of Commerce. 2021. *Quarterly Retail E-Commerce Sales: 4th Quarter 2021*. Washington, DC: U.S. Department of Commerce, Bureau of the Census.

Walk-Morris, T. 2022. "Adobe: Total Online Sales during Prime Day Near $12B." *Retail Dive.* Available online: https://www.retaildive.com/news/amazon-prime-day-online-sales-12-billion/628220/.

Weissman, J. 2014. "Amazon Is Jacking Up the Cost of Prime, and It's Still Cheap." *Slate*. Available online: https://slate.com/business/2014/03/amazon-prime-price-increase-it- s-still-inexpensive.html.

Zhu, F., and Q. Liu. 2018. "Competing with Complementors: An Empirical Look at Amzon.com." *Strategic Management Journal* 39(10):2618–2642.

Journal of Food Distribution Research
Volume 55, Issue 3, pp. 21–42

Reactions to Food Safety Recalls among Food Insecure and Food Secure Households

Mariah Beverly[a], Clinton L. Neill[b]ⓘ, and Kimberly L. Morgan[c]

[a]*Extension Economist, Center for Farm Financial Management, 130 Ruttan Hall, 1994 Buford Avenue, University of Minnesota, St. Paul, MN 55108, USA*

[b]*Assistant Professor, Dyson School of Applied Economics and Management, Warren Hall, Cornell University, Ithaca, NY 14853, USA*

[c]*Associate Professor and Extension Economist, Food and Resource Economics Department, 2685 SR 29 North, University of Florida, N. Immokalee, FL 34142, USA*

Abstract

Behavioral reactions to food safety concerns among food-insecure persons are understudied. The study of the intersection of food insecurity and food safety challenges is vital to provide more nuanced guidance on policy measures related to food safety. We use a vignette approach to examine the reactions of food-insecure individuals to a hypothetical food safety recall. Food-insecure persons are likelier to seek refunds for eggs, while Supplemental Nutrition Assistance Program (SNAP) recipients are more likely to consume romaine lettuce. We recommend policy makers use multiple channels to target food-insecure groups and to better reach consumers with information aimed at reducing the risk of illnesses in the event of a food safety recall.

Keywords: food safety, recalls, romaine lettuce, eggs, food insecurity

ⓘCorresponding author: Email: cln64@cornell.edu

Introduction

Food safety remains a paramount concern due to its significant economic impacts across entire food supply chains and on consumers globally. In 2021, the Food and Drug Administration (FDA) in the United States issued more than 500 recalls for various food and beverage products, including critical items like powdered infant formula and peanut butter (U.S. Food and Drug Administration, 2024). While food safety recalls are not uncommon, their effects can vary significantly across different populations. For instance, households experiencing food insecurity may respond differently to recalls compared to food-secure households, as they often prioritize food purchases over other health-related activities (Berkowitz, Seligman, and Choudhry, 2014). However, such nuances in behavioral responses along the spectrum of food security are largely understudied.

Food recalls are crucial to mitigate risks associated with contaminated or unsafe food products. Recalls typically initiated by regulatory agencies like the FDA or food manufacturers involve removing potentially harmful products from distribution and consumption channels. The process often begins with identifying a safety issue through surveillance systems, consumer complaints, or routine testing. Upon confirming the presence of a hazard, authorities or companies issue public notifications detailing the affected products, reasons for the recall, and recommended actions for consumers, which may include disposal, return, or refund. Subsequently, investigators trace the distribution and sale of the recalled items to minimize consumer exposure and prevent further harm. Effective communication and cooperation among stakeholders, including producers, retailers, and consumers, are essential for successfully executing recall protocols (U.S. Food and Drug Administration, 2022). Despite these measures, challenges such as incomplete product traceability and delayed responses can hinder the effectiveness of recalls, underscoring the importance of continuous improvement and vigilance in food safety management.

Previous research at the intersection of food safety and food security has primarily focused on Supplemental Nutrition Assistance Program (SNAP) participants' perceptions of risks and the role of food safety in enhancing the welfare of food-insecure populations (Neill and Holcomb, 2019; Kinsey, 2005). While SNAP participation, food insecurity, and low income are interrelated, we examine these factors individually. There are high proportions of households that are poor but food secure, and also households that are food insecure with incomes above the poverty line (Gundersen, Kreider, and Pepper, 2011). We find similar results in our sample, demonstrating with a Venn diagram the overlap among SNAP benefit recipients, food-insecure, and/or low-income respondents in Figure 1. While studies have examined the impacts of food safety recalls on various food products, particularly meat, poultry, and eggs, due to data availability the emphasis has been mainly on consumer demand and price reactions (McKenzie and Thomsen, 2001; Lusk and Schroeder, 2002; Neill and Chen, 2022; Marsh, Schroeder, and Mintert, 2004; Thomsen, Shiptsova, and Hamm, 2006). Despite the apparent impacts of food safety recalls on prices and demand, the welfare effects are not always straightforward. Factors, such as willingness-to-pay (WTP) changes for affected food items, can vary depending on consumer knowledge and preferences (Richards and Nganje, 2014).

Moreover, the consequences of a food safety recall extend beyond economic considerations. Socioeconomic and demographic factors play a crucial role in determining the magnitude and distribution of these impacts. We believe that a better understanding of how consumers may react to food safety recalls has two main pathways: limited income/price sensitivity and risk preferences. Households near the poverty line, who are more likely to experience food insecurity, spend a significant proportion of their income on food compared to more affluent households (Coleman-Jensen et al., 2019). Thus, they may value the consumption of food products differently and react differently to food safety recalls, possibly prioritizing immediate food needs over health considerations, often referred to as patience. An individual's willingness to accept the risk of becoming ill may influence their response to food safety recalls. We expect some may be more risk averse and choose to avoid potentially contaminated products, instead electing to either dispose of them or seek a refund from the place of purchase.

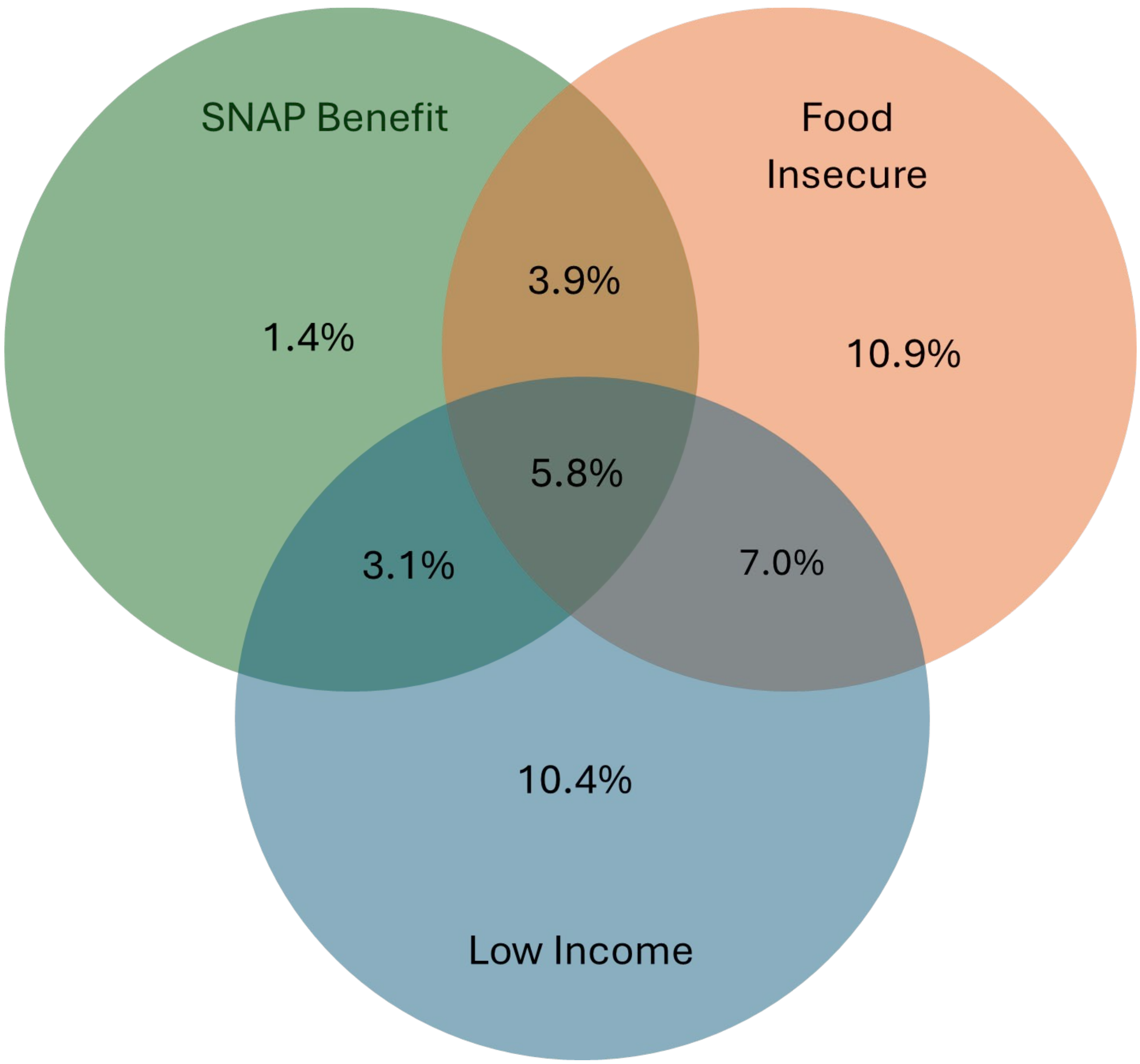

Figure 1. Percentage of SNAP Benefit Recipient, Food Insecure, and Low-Income Respondents

The specific effects of such pathways are likely to manifest most in food-insecure individuals as they are more likely to consume or return food items identified by a recall. Our work aims to directly observe the behavioral responses of food-insecure individuals to food safety recalls, shedding light on the factors influencing their decisions regarding recalled products. We do this by examining how time/patience and risk preferences interact with the choices of food-insecure individuals in a hypothetical experiment.

Our study contributes to the existing literature by employing a vignette approach to examine the reactions of food-insecure and food-secure individuals to hypothetical food safety recalls. Application of our results by companies and government agencies during a food recall event using targeted interventions that consider the food security status of households may reduce instances and severity of foodborne illnesses. Our findings suggest that attributes of specific food items, return policies, and demographic factors significantly influence consumers' responses to food safety recalls. We find that consumer reactions to recalls vary across demographics, food security status, and access to SNAP benefits. Food-insecure individuals appear more willing to accept the risk of becoming ill and choose to consume a recalled food. Nonwhite and relatively older respondents appear more likely to seek a refund for food purchased when a recall is announced. We recommend government and private sector entities use social media to distribute science-based information and risk-reducing actions available to vulnerable groups in the event of a food safety recall. Another option for SNAP recipients is to credit refunds to SNAP accounts, whether or not they used the funds to purchase the recalled food items.

In the remainder of this article, we discuss the survey data collected, the empirical model, study results, policy implications, and concluding remarks, and highlight avenues for future research.

Survey Design and Data Collection

We utilize stated preference methods to understand consumers' decisions to seek a refund, throw away, or consume eggs and romaine lettuce subject to a hypothetical food safety recall. The choice of food items for the study is not arbitrary. We identify eggs and romaine lettuce purchases based on participant consumption of each food and the representativeness of these two distinct food categories. Eggs must be cooked before consumption and are a relatively cheaper source of protein compared to meat and meat alternatives. Romaine lettuce is eaten in fresh form and is a vegetable households consume—albeit not the cheapest nor most expensive one. Both food items have been subject to several recalls over the past decade. By examining products with different dietary functions, we can determine how food attributes and demographic factors influence decision making in the case of a food safety recall. The primary contribution of our study to the existing literature is to improve understanding of behavioral reactions to food safety recalls between food-secure and food-insecure households.

The experimental design used for this study is the vignette method, which is a type of stated preference experiment where respondents make hypothetical decisions (regarding products, situations, etc.) with differing levels of attributes. Social psychology was the first field to use this methodology (Alexander and Becker, 1978) and has expanded to several fields, including marketing and management (Aguinis and Bradley, 2014), as well as economics (Kapteyn, Smith, and van Soest, 2007; Epstein, Mason, and Manca, 2008; Ellison and Lusk, 2018). The vignette method has proven to recover the actual effects of attributes of interest in real-world scenarios (Hainmueller, Hangartner, and Yamamoto, 2015).

The vignette in our analysis has three attributes—price, risk of sickness, and travel time to store, each varied at three levels. From the 27 possible vignettes ($3^3 = 27$), we selected a subset of nine vignettes such that each variable was uncorrelated with the others (an orthogonal, fractional factorial design).

We recognize that one potential limitation to using data from an online survey analysis is stated preferences rather than revealed preference. However, the authors are not aware of data showing revealed preferences that also include observations on whether consumers returned a recalled food item for a refund or threw away the affected product. Therefore, we feel the online survey data used in this analysis that uses stated preferences is justified, given that we also collect observations about a consumer's response to a hypothetical food safety recall.

We elicit each participant's consumption pattern for each food item and then randomly assign them to evaluate one of the nine vignettes. Each respondent answered a vignette for romaine and egg food safety recalls if they consumed each food item at least once a month. If they responded by indicating that they never consumed one of the food items, they were not presented with the vignette for that food item. They were not included in the experiment if they never consumed either food item.

Below are examples of the basic vignette for romaine lettuce and eggs:

Romaine Vignette

Imagine you just found out about a food safety recall for romaine lettuce you recently purchased due to the risk of *E. coli*. The estimated risk of *E. coli* from the consumption of the lettuce is about **[1 in 3 (33%); 1 in 6 (17%); 1 in 9 (11%)]**. The lettuce cost you **[$2.00; $2.70; $3.30]** per pound. Assuming the grocery store where you can return the lettuce for a refund is a **[20; 30; 40]** minute round trip, what would you do?

Eggs Vignette

Imagine you just found out about a food safety recall for large, Grade A eggs you recently purchased due to the risk of salmonella. The estimated risk of salmonella from the consumption of eggs is about **[1 in 100 (1%); 1 in 200 (0.5%); 1 in 300 (0.33%)]**. The eggs cost you **[$1.60; $1.80; $2.00]** per dozen. Assuming the grocery store where you can return the eggs for a refund is a **[20; 30; 40]** minute round trip, what would you do? The respondent had three options: throw away the food item, return it to the store for a refund, or consume it.

We utilize data from Centers for Disease Control and Prevention (2022a,b) to create realistic probabilities of sickness from *E. coli* and salmonella. We first gather the total number of cases of foodborne illness of interest over 2017–2020, regardless of the source of contamination. Then, we gather the total number of illnesses for the foodborne illness of interest and the specific food that caused the illness. For example, the risk of illness from romaine facing a food safety recall for *E. coli* is calculated as follows:

$$\text{Risk of illness from romaine} = \frac{\text{\# of } E.\ coli \text{ cases caused by consuming romaine lettuce contaminated by } E.\ coli}{\text{\# of illnesses caused by } E.\ coli}$$

To examine risk preference specifically related to foodborne illnesses, one of the choice attributes within the experiment is the risk of illness caused by consuming romaine and eggs. The specific calculation for the risk of illness caused by consuming romaine contaminated by *E. coli* to be $\frac{617}{3,588} \approx 17\%$ and the risk of illness caused by consuming eggs contaminated by salmonella to be $\frac{103}{13,472} \approx 0.7\%$. From these base calculations, we choose the three levels of risk to be 1 in 3, 1 in 6, and 1 in 9 for romaine lettuce and 1 in 100, 1 in 200, and 1 in 300 for shell eggs. We acknowledge that this risk estimation is the probability of eggs/romaine being the source of contamination given an illness due to the specified bacteria and not the risk of becoming ill. The likelihood of becoming sick is much smaller (i.e., $\leq 0.01\%$ for eggs). While our estimates exceed the documented infection rates, we argue that this approach is still worthwhile in understanding how different consumers respond to the recall. We suggest that future research focus on testing the hypothesis with smaller probabilities. Moreover, the average consumer likely has little or no frame of reference for the risk of illness from recalled food products, given the small probabilities associated with such rare events (Burns, Chiu, and Wu, 2010).

Before completing the choice experiment, we informed participants that food is occasionally recalled due to the risk of contamination of a foodborne illness, and that consumers who have purchased a recalled food item have three options: to return the product to the store from where it was purchased for a refund, dispose of the product properly so that people cannot eat it, or ignore the recall and consume the product. We did not inform them of the potential consequences of consuming contaminated foods as we were interested in extracting inherent risk perceptions. For SNAP benefit recipients, all refunds go directly to the recipients' Electronic Benefit Transfer (EBT) card. Cash refunds are not allowed for food items purchased with SNAP benefits.

To determine the price-level attributes, we gathered price data from the Federal Reserve Bank of Saint Louis and rounded to the nearest 10 cents. For romaine, we used the average price of romaine from February 2020 and December 2021 (this data range contained the lowest and highest price for romaine lettuce over the past five years) to determine a midpoint price point of $2.70 (U.S. Bureau of Labor Statistics, 2022b). We used the average price for shell eggs between February 2021 and February 2022 to determine a midpoint price point of $1.80 (U.S. Bureau of Labor Statistics, 2022a). Travel time to a store to obtain a refund for a food item facing a food safety recall is based on research showing the average time individuals in low-income areas spend traveling to a grocery store, which is 19.5 minutes (Ver Ploeg et al., 2009; Hamrick and Hopkins, 2012).

Utilizing the 18 survey items from the U.S. Household Food Security Survey Module (USDA-ERS, 2012), we calculated the food insecurity status of respondents. Respondents who answer affirmatively to three or more questions in the survey meet the definition of food-insecure households.[1]

Given our interest in time preferences, we utilize survey questions and methods outlined in previous studies by Falk et al. (2023) and Falk et al. (2018). For the patience measure, each survey respondent was asked about their willingness to give up something beneficial today to gain something more valuable in the future and answered five questions about their choice between differing amounts of money today versus in the future.[2] We elicit a measure of personal risk preference using the methods above. Each survey respondent was asked about their willingness to take risks and answered five questions about their preference for a 50/50 chance of receiving different amounts of money as a sure payment. We normalized both scores to the average, where a negative value would indicate a lower-than-average risk/patience measure and vice versa for a positive measure.

We ask respondents questions about age, gender, race, education, political affiliation, income, whether children are present in the household, and whether they receive SNAP benefits. We collected a national sample of consumers in the United States via an online panel.[3] We incentivized respondents via payment to complete the survey and provided accurate responses through an online panel maintained by a third party (Qualtrics), resulting in 1,050 completed responses after removing inconsistent responses based on an inattention question.[4] The food insecurity rate in our sample is approximately 28% higher than the national average. Compared to the latest U.S. Census data, 24.7% of people in the United States identify as nonwhite, 50.5% are female, and the average age is 38.9. Fewer of our respondents identify as nonwhite (9.3%), 51.3% are female, and the average age is 41.9. We were left with 860 responses for analysis. Comparing food insecure (N = 238) and food secure (N = 622) individuals, we find several differences in characteristics. Notably, food-secure individuals in our study have lower measures of personal risk preference (indicating a person is more risk averse). Sociodemographic information about the sample can be viewed in Table 1.

[1]Note: While the food insecurity questions do produce a categorical measure of food insecurity, we follow how the USDA Economic Research Service (2012) typically report food insecurity in a binary measure for ease of interpretation.
[2]Others may download these survey questions for U.S. residents from https://www.briq-institute.org/global-preferences/downloads.
[3]We received appropriate university IRB approval before data collection.
[4]The inattention question was “I/my household ate at least once in the last 12 months. For this question, please select ‘often true.’”

Table 1: Socio-Demographic Variables and Definitions

Variable	Definition	Food Insecure Mean	Food Secure Mean	Full Sample Mean
Personal risk preference	Measure of personal risk preference	0.145	-0.151	-0.068
Patience	Measure of patience	-0.246	0.103	0.007
Female	1 if female; 0 otherwise	0.643	0.463	0.513
SNAP	1 if current SNAP recipient; 0 otherwise	0.353	0.068	0.147
Age	Current age	35.042	44.506	41.887
Food budget > $100	1 if weekly food budget > $100; 0 otherwise	0.592	0.672	0.650
Nonwhite	1 if respondent identified as nonwhite; 0 otherwise	0.118	0.084	0.093
College	1 if obtained a college degree; 0 otherwise	0.462	0.641	0.592
Democrat	1 if identifies as a Democrat; 0 otherwise	0.378	0.342	0.352
Children in HH	1 if children under 18 are in household; 0 otherwise	0.424	0.172	0.242
Low income	1 if income is less than $40,000; 0 otherwise	0.471	0.190	0.267
Medium income	1 if income is between 40, 000−99,999; 0 otherwise	0.332	0.476	0.436
High income	1 if income is $100,000 or more; 0 otherwise	0.197	0.334	0.297
Number of observations		238	622	860

Note: This table presents means for the combined sample of respondents for both food safety recall food types.

Econometric Methods

Our analysis makes use of a multinomial logistic regression to determine how attributes, such as price, risk of illness, travel time to a store, and demographic variables, affect a person's decision to obtain a refund, dispose of, or consume a food facing a food safety recall. Given the three possible outcomes, the corresponding probability P that a person i chooses a specific outcome j (to obtain a refund, throw away, or consume a food item facing a food safety recall) are as follows (Greene, 2012):

$$P(Y_i = j) = \frac{\exp(X\beta^j)}{\sum_{j=1}^{3} \exp(X\beta^j)} \tag{1}$$

where X are explanatory variables and β^j is a set of estimated coefficients corresponding to each outcome j. To identify our model, we set the base outcome as the decision to throw away the food item. Therefore, all of the coefficient estimates are relative to the decision to discard the recalled food.

Specifically, we model each person's decision to obtain a refund, throw away, or consume a food item that has a food safety recall through a multinomial logistic regression with the following covariates:

$$X\beta^j = \beta_{0i} + \beta_1 Price_i + \beta_2 Store_i + \beta_3 Sick_i + \beta_4 FI_i \tag{2}$$

$$+ \beta_5 Patience_i + \beta_6 FIPatience_i + \beta_7 FISick_i + \alpha \mathbf{Z_i}$$

where we note the vignette variables by *Price, Store,* and *Sick*. *FI* is the food insecure dummy variable, and *Patience* measures the respondent's patience.

Given our hypotheses that food-insecure persons are likely more sensitive to their own time and risk preferences regarding foodborne illnesses, we utilize interaction terms. *FIPatience* and *FISick*, are the food insecurity binary variable and the respondent's patience measure and the risk of sickness attribute from the vignette, respectively. These terms directly test whether individuals are more concerned about how they spend their time dealing with food recalls, as they may be less willing to return the items for a refund. If food-insecure individuals are more concerned about the inherent risk of illness, they should be less likely to consume. Given that we consider a multinomial option in response, we expect the *FIPatience* variable to be positive in the Refund option and the *FISick* variable to be negative in the Consume option. Prior literature suggests that because food insecurity is so stressful (Laraia et al., 2017), individuals suffering from it may be unwilling to risk sickness that will possibly cause increased stress or expenditures from medical treatment. But, as previously mentioned, whether this manifests in terms of time or risk preferences (or both) has yet to be determined.

We denote the matrix of demographic variables as **Z**, which includes the following: *Child* is a dummy variable, indicating the presence of children in the household; *Female* is an indicator for whether the respondent identifies as female; *SNAP* is an indicator for whether the respondent is a SNAP benefit recipient; *Age* is the age of the respondent; *Nonwhite, College, Dem, Medium Income, High Income* are the indicator variables for whether the respondent identifies as nonwhite, has a college degree, identifies with the Democratic party, has a medium level of income, or has a high level of income, respectively. We model the choice to obtain a refund, throw away, or consume each food item separately (i.e., eggs (N = 838) impacted by a food safety recall due to the risk of salmonella and a second model for romaine lettuce (N = 742) impacted by a food safety recall due to the risk of *E. coli.*) We tested the model for multicollinearity given the various types of risk controls and found these variables were uncorrelated.

Our study design has several assumptions. First, we assume consumers' reactions to food safety recalls are unaffected by attributes outside our experimental design, such as the recall timing relative to the purchase. For example, enough time may have lapsed between purchasing a food item and a food recall event that the consumer may have already consumed or disposed of the item due to spoilage. Our experiment also assumes the respondent is aware of the recall because we explicitly informed them. Consumers who have purchased a food item and are facing a food safety recall may have varying amounts of information regarding the recall.

Media coverage around the time of the recall event has impacted consumers' decisions (Neill and Chen, 2022). Also, we do not have a proper measure of respondents' time use, but rather a measure for patience, which, while not equivalent, is more straightforward to extract in a survey. Respondents with less leisure time may react differently to a food safety recall than those with ample leisure time. Our results may depend on the choice of food items in our analysis and may not be comparable outside of recalls for romaine lettuce or eggs. However, our analysis offers new insight to policy makers and researchers on the reactions to food safety recalls across groups of individuals. Finally, there are only 16 households that consume romaine lettuce after a recall in our sample. Thus, the results observed are driven by a small number of observations, which is a potential threat to proper identification.

Results

Of initial interest are the respondents' preferences of risk and time preferences. We find that food-insecure individuals are less patient and more willing to take risks than food-secure individuals, as shown in Figure 2. Our findings are similar to Neill and Holcomb (2019), where SNAP recipients had a lower perceived risk of the presence of *E. coli* in fresh produce from smaller farms. Given the challenges food-insecure people face and differences in risk preference and patience measures, we hypothesize that food-insecure individuals will react differently to food safety recalls than food-secure households.

We summarize the survey respondents' decisions by food recall type and food security status in Figure 3. Food-insecure respondents had a higher percentage of seeking a refund for both recalled eggs and romaine compared to food-secure respondents.

We present the multinomial logistic (MNL) regression results for romaine lettuce and shell eggs in Table 2. All regressions use the decision to throw away affected food items as the base outcome. We discuss results predominately via marginal effects for the MNL regressions in Table 3 for eggs and in Table 4 for romaine. In our MNL regressions, all three choice attributes (price, travel time to store, and risk of sickness) are significant factors in the decision to refund or consume food items facing a food safety recall. The price variable is significant for the decision to consume purchased eggs despite a recall event compared to disposal of the eggs. The price variable is significant for obtaining a refund and consuming recalled romaine relative to the disposal of the romaine. A one-dollar increase in the price of romaine results in a 29% increase in the log-odds of an individual choosing to seek a refund for the purchased romaine. The marginal effect of price presents similar findings. As price increases, a consumer's probability of throwing away eggs or romaine decreases by 19 percentage points.

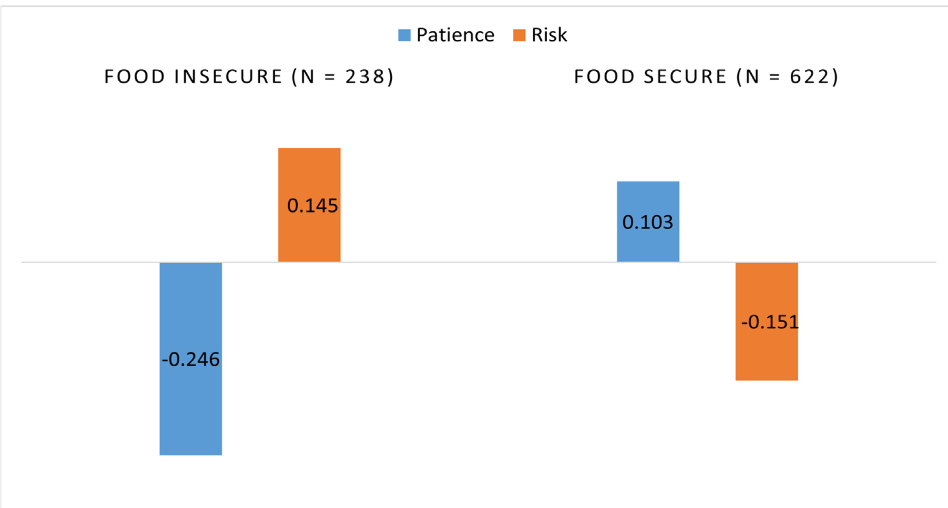

Figure 2. Risk and Patience among Food-Secure vs. Food-Insecure Households (Normalized Z-score Values)

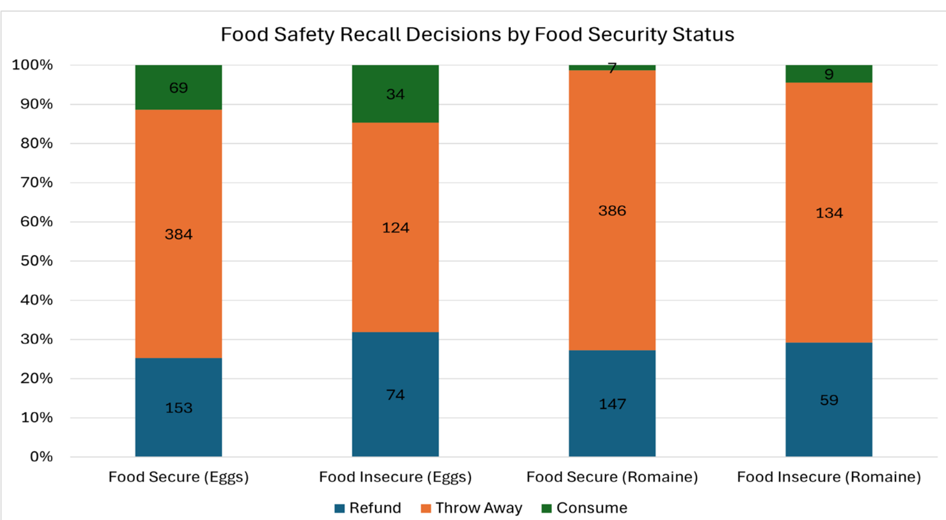

Figure 3. Survey Responses to Seek Refunds, Throw Away, or Consume Food Under a Food Safety Recall

Table 2. Multinomial Logistic Regression Results from Egg and Romaine Lettuce Vignette (Base = Throw Away Food Item)

	Eggs		**Romaine**	
Variables	**Refund**	**Consume**	**Refund**	**Consume**
Price	0.590	1.339*	0.288*	0.772*
	(0.514)	(0.743)	(0.160)	(0.436)
Travel time to store	-0.047***	0.015	-0.035***	-0.003
	(0.010)	(0.015)	(0.011)	(0.039)
Risk of sickness	0.338	1.077**	-0.001	-0.017
	(0.344)	(0.452)	(0.011)	(0.039)
Children in HH	0.155	0.588*	0.396	0.780
	(0.241)	(0.305)	(0.246)	(0.570)
Female	-0.167	-0.225	-0.407**	-0.807
	(0.184)	(0.249)	(0.193)	(0.581)
SNAP benefit recipient	0.228	0.503	0.204	1.847**
	(0.272)	(0.340)	(0.285)	(0.772)
Age	0.023***	0.006	0.028***	-0.015
	(0.008)	(0.011)	(0.008)	(0.032)
Nonwhite	0.553**	-1.587**	0.017	-13.747***
	(0.270)	(0.737)	(0.303)	(0.415)
College	-0.132	-0.134	0.107	0.118
	(0.186)	(0.249)	(0.190)	(0.691)
Democrat	0.280	0.207	0.398**	-0.157
	(0.176)	(0.235)	(0.178)	(0.649)
Medium income	-0.084	0.246	-0.324	0.728
	(0.231)	(0.349)	(0.245)	(0.863)
High income	-0.224	0.112	-0.334	0.880
	(0.277)	(0.382)	(0.283)	(1.025)
Food insecure	1.048**	2.079***	0.769	1.292
	(0.463)	(0.615)	(0.469)	(1.264)
Patience	0.012	0.281	-0.012	0.447
	(0.123)	(0.176)	(0.125)	(0.500)
Food insecure ×	0.579**	0.084	0.336	-0.060
patience	(0.251)	(0.312)	(0.258)	(0.653)
Food insecure ×	-0.701	-2.790***	-0.020	-0.035
risk of sickness	(0.642)	(0.903)	(0.022)	(0.061)
Constant	-1.747	-5.721***	-1.827**	-6.029**
	(1.150)	(1.527)	(0.781)	(2.634)
Observations	838		742	

Note: Robust standard errors presented in parentheses. Single, double, and triple asterisks (*, **, ***) indicate statistical significance at the 10%, 5%, and 1% level.

Table 3. Average Marginal Effects for Eggs

	Refund		Throw Away		Consume	
	Avg. ME	Std. Err.	Avg. M.E.	Std. Err.	Avg. M.E.	Std. Err.
Price	0.068	0.092	-0.187*	0.104	0.119	0.074
Travel time to store	-0.009***	0.002	0.006***	0.002	0.003**	0.001
Risk of sickness	0.029	0.061	-0.129*	0.067	0.100**	0.045
Children in HH	0.011	0.043	-0.066	0.047	0.055*	0.030
Female	-0.024	0.033	0.042	0.036	-0.018	0.025
SNAP benefit recipient	0.027	0.048	-0.071	0.054	0.045	0.033
Age	0.004***	0.001	-0.004**	0.002	0.000	0.001
Nonwhite	0.150***	0.049	0.029	0.073	-0.179**	0.074
College	-0.020	0.033	0.030	0.037	-0.010	0.025
Democrat	0.045	0.031	-0.058*	0.035	0.013	0.023
Medium income	-0.023	0.042	-0.005	0.048	0.028	0.035
High income	-0.045	0.049	0.026	0.056	0.018	0.038
Food insecure	0.129	0.081	-0.310***	0.090	0.181***	0.060
Patience	-0.006	0.022	-0.022	0.025	0.028	0.018
Food insecure × patience	0.104**	0.044	-0.095*	0.050	-0.009	0.030
Food insecure × risk of sickness	-0.044	0.114	0.308**	0.128	-0.264***	0.090

Note: Single, double, and triple asterisks (*, **, ***) indicate statistical significance at the 10%, 5%, and 1% level.

Table 4. Average Marginal Effects for Romaine

	Refund		Throw Away		Consume	
	Avg. ME	Std. Err.	Avg. M.E.	Std. Err.	Avg. M.E.	Std. Err.
Price	0.050	0.030	-0.062**	0.030	0.013	0.009
Travel time to store	-0.007***	0.002	0.006***	0.002	0.000	0.001
Risk of sickness	0.000	0.002	0.000	0.002	0.000	0.001
Children in HH	0.070	0.046	-0.082*	0.046	0.012	0.011
Female	-0.072**	0.036	0.085**	0.036	-0.013	0.011
SNAP Benefit Recipient	0.027	0.053	-0.061	0.053	0.034**	0.016
Age	0.005***	0.001	-0.005***	0.002	0.000	0.001
Nonwhite	0.089	0.060	0.173**	0.070	-0.262***	0.062
College	0.019	0.036	-0.021	0.036	0.002	0.013
Democrat	0.076**	0.033	-0.071**	0.034	-0.005	0.012
Medium income	-0.066	0.046	0.050	0.046	0.016	0.017
High income	-0.069	0.053	0.050	0.054	0.019	0.020
Food insecure	0.137	0.088	-0.157*	0.088	0.020	0.024
Patience	-0.005	0.024	-0.004	0.024	0.009	0.010
Food insecure × patience	0.064	0.048	-0.061	0.049	-0.003	0.012
Food insecure × risk of sickness	-0.003	0.004	0.004	0.004	-0.001	0.001

Note: Single, double, and triple asterisks (*, **, ***) indicate statistical significance at the 10%, 5%, and 1% level.

Obtaining a refund for eggs and romaine lettuce requires significant travel time to the store. For eggs, we find that for each minute the travel time to the store increases, the probability a consumer seeks a refund decreases by approximately 1 percentage point. We also find that the probability the respondent disposes of or consumes the eggs increases by 0.6 percentage points and 0.3 percentage points, respectively. The marginal effect of travel time to the store for the food safety recall for romaine is similar. As travel time to the store increases, the consumer's probability of returning the affected romaine decreases by 0.7 percentage points. Respondents with children present in the household are more likely to consume eggs recalled due to a food safety issue than dispose of them. If children are present in the household, the probability of consuming the eggs despite the recall increases by 6%. The marginal effect of having children present in the household decreases the probability of throwing away the affected romaine by 8 percentage points. Eggs are relatively low cost compared to other protein sources (Farrell, 2013; Conrad et al., 2017) and are rich in nutrients, such as amino acids, choline, vitamins A, B, and D, and iron (FAO, 1985; Griffin, 2016; USDA-ARS, 2019). Given the reduced chance of getting sick from eggs cooked until the white and yolk are firm (CDC, 2022), parents may assume that cooking the eggs results in an acceptable reduction of the risk of using eggs under an active recall. Additionally, children do not often prefer vegetables (Skinner et al., 2002). It is possible that households with children may not be as concerned with a food safety recall for romaine lettuce because their children prefer not to eat vegetables, thus leading to a lower probability of throwing the romaine away due to the recall. Lastly, it is also possible that because romaine recalls are more prevalent than egg recalls, consumers may be more aware of romaine recalls in recent years.

For romaine lettuce facing a food safety recall, the MNL regression shows female consumers are less likely to seek a refund than they are to elect to dispose of the recalled romaine. For female romaine consumers, the probability of seeking a refund decreases by 7 percentage points, and the probability of disposal decreases by 9 percentage points compared to male consumers. This result is likely attributable to the fact that women have documented less leisure time than men given traditional gender roles and the overall differences in time use between men and women (Thrane, 2000; Sayer, 2005; Van der Lippe et al., 2011).

Being a SNAP benefit recipient is significant and increases the probability a consumer will choose to consume the romaine despite the food safety recall by 3 percentage points. SNAP recipients are the most price conscious and employ price-saving efforts soon after receiving their benefits (Zaki and Todd, 2021). This fact, coupled with the relatively short window of consumption before romaine lettuce spoils, likely drives SNAP recipients to consume rather than throw away recalled romaine lettuce.

As age increases, the probability of seeking a refund increases by 0.4 percentage points, whereas the probability of throwing away recalled eggs decreases by 0.4 percentage points. For romaine under a food safety recall, the probability of seeking a refund increases by 0.5 percentage points, and the probability of throwing away the romaine decreases by 0.5 percentage points. Our results are similar to Schafer et al. (1993), who found that age is related to food safety behavior. In addition, consumer expenditures vary by age (Foster, 2015). For example, the share of the food budget spent on food at home increases with age (Foster, 2015). It is likely that as age increases, respondents are more likely to seek a refund relative to throwing away a food item or when facing a food safety recall due to spending habit differences among different age groups.

For nonwhite consumers, the probability of seeking a refund increases by 15 percentage points, and the probability of consuming the recalled eggs decreases by 18 percentage points. The likelihood of nonwhite respondents who choose to consume romaine under a food safety recall is 26 percentage points lower than the choice to dispose of the lettuce.

Democrats are more likely to seek a refund than to dispose of romaine lettuce. The probability of seeking a refund increases by 7.6 percentage points relative to disposing of romaine lettuce impacted by a food safety recall when the consumer identifies with the Democratic party. Identifying as a Democrat decreases the probability that the consumer throws away recalled eggs by 6% and recalled romaine by about 7 percentage points. This finding may be due to the link between personality and political choice (Capara, Barbaranelli, and Zimbardo, 1999; Caprara et al., 2006).

Our results indicate that food-insecure persons are more likely than food-secure persons to seek a refund of eggs or consume the eggs under a food safety recall. Being food insecure decreases the probability a consumer will throw away recalled eggs by 31 percentage points and increases the probability of choosing to consume the eggs despite the recall by 18 percentage points. Additionally, food insecurity decreases the probability of throwing away recalled romaine by 16 percentage points. Since food insecurity is stressful to individuals (Laraia et al., 2017), and food-insecure persons may focus all their efforts on finding food (Hadley and Crooks, 2012), it is plausible that food-insecure individuals are less likely to throw away food items or consume them despite the recall. When interpreting the effect of being food insecure on reactions to food safety recalls, patience and risk of sickness must also be considered as we included the interaction of these variables and food insecurity status. As a food-insecure person's patience measure increases, a food-insecure person is more likely to seek a refund and less likely to throw away recalled eggs. Additionally, as the risk of sickness increases for a food insecure person, the more likely they are to throw away recalled eggs and the less likely they are to consume recalled eggs.

Testing our empirical hypothesis about food-insecure persons, we find the interaction terms between food insecurity and patience significant in the egg model. Food-insecure individuals with higher patience measures have an increased probability of seeking a refund increase by 10 percentage points, and the probability of throwing away recalled eggs decreases by 9.5 percentage points. More patient, food-insecure individuals may have an inherent ability to devote time to seeking safe food options. As such, they are more likely to pursue a refund and are less likely to throw away contaminated eggs due to the opportunity to buy uncontaminated eggs or another cheap protein source with the refund given. Our second interaction term between food-insecure persons and the risk of sickness from the experiment is also statistically significant in the egg model. We find that food-insecure individuals have a lower probability (26.4 percentage points) of consuming recalled eggs than to dispose of them as the risk of sickness associated with the recall increases. Similarly, as the risk of illness increases during a food safety recall, the probability of a food-insecure person disposing of recalled eggs increases by 30.8 percentage points. However, neither interaction term was significant in the romaine models, which supports our initial hypothesis that time and risk preferences are not universally important across all categories of products identified in food recalls.

A summary of our general findings for the vignette attributes, SNAP benefit recipients, and food insecure individuals are given in Table 5.

Table 5. Summary of Findings for Vignette Attributes, SNAP Recipients, and Food-Insecure Individuals

General Findings	Possible Explanation
As price increases, consumers are less likely to choose to throw away both eggs and romaine.	Consumers save money by reducing food waste, so they are less likely to choose to throw away food affected by a recall.
As travel time to the store increases, consumers are less likely to choose seeking a refund for recalled eggs and romaine and more likely to throw away recalled eggs and romaine.	Travel costs increase as travel time, discouraging consumers to seek a refund for recalled items and encouraging them to throw away recalled items.
SNAP benefit recipients are more likely to choose to consume recalled romaine lettuce.	SNAP benefit recipients are most price conscious after receiving their benefits and may be taking advantage of the short consumption window for romaine.
Food insecure individuals are less likely to choose to throw away and more likely to consume recalled eggs.	Food insecurity is stressful and those individuals may focus much of their effort into finding and keeping food.
As a patience score for a food insecure individual increases, the individual is more likely to choose seeking a refund and less likely to throw away recalled eggs.	Food insecure individuals with a higher patience score may be taking advantage of the opportunity to receive uncontaminated eggs or another cheap protein source with the refund.
As risk of sickness from a recall increases for a food insecure individual, they are more likely to choose to throw away recalled eggs and less likely to choose to consume recalled eggs.	Food insecure individuals may be unwilling to risk a sickness that could cause more stress or more health-related expenditures, given the increasing risk of sickness from a food safety recall.

Discussion and Implications

Our results offer several insights to researchers studying food safety recalls and policy makers seeking to implement effective strategies surrounding the consumer decision to heed food safety recalls. For researchers, we find several factors that should be considered when studying the reactions to food safety recalls in the future. Because travel time to stores is a significant factor in deciding whether to obtain a refund, further studies should include this attribute to accurately model the decision-making process of consumers facing a food safety recall. Additionally, public messaging from retailers indicating that the recalled product itself need not be returned to the store; rather, they honor the refund based on a receipt showing purchases made within the time frame of the recall may improve accessibility to the food without costing more time and resources.

For policy makers, the goal of a food recall is ”to protect the public from products that may cause health problems or possible death” by removing ” food products from commerce when there is reason to believe the products may be adulterated or misbranded” (USDA Food Safety Inspection Service, 2015). Thus, a “successful” consumer reaction to a food safety recall would result in consumers who have purchased a potentially harmful food item throwing the item away or returning the item to the store for a refund. Our results indicate that not all consumers would be willing to throw away or obtain a refund for eggs or romaine if they face a food safety recall. Specifically, being a SNAP benefit recipient increases the probability of choosing to consume romaine under a food safety recall. As indicated previously, several explanations exist for this phenomenon, including that SNAP benefit recipients view risk differently than consumers who do not receive SNAP benefits.

Furthermore, we do not find that our hypotheses about SNAP recipients are universally true, given the non-statistically significant effects of a romaine lettuce recall. Policy makers can focus on targeting SNAP recipients during a food safety recall to discourage consumption instead of focusing on their time and risk preferences, which is unlikely to be ineffective.

Other notable demographic groups in our analysis include nonwhite respondents. Survey respondents who identified as nonwhite were more likely to react successfully to a food safety recall for eggs and romaine lettuce (i.e., they were more likely to seek a refund or throw away an item under a food safety recall and less likely to consume an impacted product). Additionally, food-insecure persons have a higher probability of consuming recalled eggs and a lower probability of choosing to throw away recalled eggs or romaine lettuce. Given the success of social media tools in disseminating public health messages (Mayer and Harrison, 2012), we recommend public and private sectors cooperate to circulate relevant information regarding food safety recalls to consumers using these channels. Communicating the importance of reporting and removing unsafe food items is critical to decreasing foodborne illnesses and costs. In addition, policy support for increasing traceability from food production to households that purchased unsafe foods will assist in tracking food safety recalls more accurately.

Conclusions

There are several nuances in the decision-making process when consumers face a hypothetical food safety recall of romaine lettuce and eggs. In our study, we attempted to determine how decision-making is similar or different across food-insecure and food-secure persons. Using the vignette method and multinomial logistic regression, we find the outcome depends on contextual factors, such as price, travel time to a store, and socioeconomic and demographic factors. We also show that food-insecure individuals react to food safety recalls differently than food-secure individuals as the risk of sickness from consuming recalled eggs or romaine lettuce increases. Again, we acknowledge our limitation in using probabilities larger than the actual risk of illness. However, this approach can provide opportunities for further understanding of food-insecure consumers' reactions to food safety concerns. Also, many consumers likely have no frame of reference for the probability of illness from recalled products, so extracting such a measure may also prove worthwhile.

We add to the food insecurity and the food safety literature by determining attributes affecting a decision across individuals who have purchased items subject to a food safety recall. We show differences in how food-insecure persons react to recalls of shell eggs, and we find that other demographic groups respond differently to food safety recalls. Specifically, we find that being a SNAP benefit recipient increases the likelihood of choosing to consume romaine lettuce when facing a food safety recall. We also find that compared to white consumers, nonwhite consumers have a higher probability of seeking a refund for eggs affected by a food safety recall and a lower probability of consuming recalled eggs. Additionally, nonwhite consumers are more likely to decide to throw away and less likely to choose to consume romaine lettuce under a food safety recall compared to white consumers. Our findings are relevant to researchers and policy makers, as decisions on how best to react to a food safety recall differ based on demographics and product-specific factors.

Our findings set the stage for further research surrounding the factors that influence decision-making under a food safety recall. We demonstrate that attributes regarding a food safety recall are essential to how consumers react to food safety recalls. We determine that these decisions may differ based on demographic factors. Future work should focus on other variables not utilized in this analysis, such as the timing of the recall event relative to the purchase date or the amount of leisure time available to consumers. For example, policy analysis often fails to consider how SNAP benefit recipients use available time (Davis and You, 2011; You and Davis, 2019). Another option could elicit participants' actual travel times and frequency of visits to their preferred food stores and utilize this information within the experiment. Capturing these metrics in future work may provide a better understanding of the decision-making process consumers undergo when faced with a food safety recall and better inform policy makers on the best practices to reduce the risks of foodborne illness among consumers. Our research motivates the importance of incorporating the link between food safety and food waste in future research. Food waste is a natural part of the food system, predominately due to supply chain concerns, such as spoilage, that render it unfit for safe human consumption.

References

Aguinis H., and K.J. 2014. "Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies." *Organizational Research Methods* 17:351–371.

Alexander, C.S., and H.J. Becker. 1978. "The Use of Vignettes in Survey Research." *The Public Opinion Quarterly* 42:93–104,.

Berkowitz, S.A., H.K. Seligman, and N.K. Choudhry. 2014. "Treat or Eat: Food Insecurity, Cost-related Medication Underuse, and Unmet Needs." *The American Journal of Medicine* 127:303–310.e3.

Burns, Z., A. Chiu, and G. Wu. 2010. "Overweighting of Small Probabilities." In J.J. Cochran, L.A. Cox, P. Keskinocak, J.P. Kharoufeh, and J.C. Smith, eds. *Encyclopedia of Operations Research and Management Science*. Hoboken, NJ: Wiley.

Capara, G.V., C. Barbaranelli, and P.G. Zimbardo. 1999. "Personality Profiles and Political Parties." *Political Psychology* 20:175–197.

Caprara, G.V., S. Schwartz, C. Capanna, M. Vecchione, and C. Barbaranelli. 2006. "Personality and Politics: Values, Traits, and Political Choice." *Political Psychology* 27:1–28.

Centers for Disease Control and Prevention. 2022a. *Salmonella and Eggs*. Atlanta, GA: CDC. Available online: https://www.cdc.gov/foodsafety/communication/salmonella- and-eggs.html.

Centers for Disease Control and Prevention. 2022b. *National Outbreak Reporting System (NORS) Dashboard*. Atlanta, GA: CDC. Available online: https://wwwn.cdc.gov/norsdashboard/.

Coleman-Jensen, A., M.P. Rabbitt, C.A. Gregory, and A. Singh. 2019. *Household Food Security in the United States in 2018*. Washington, DC: U.S. Department of Agriculture, Economic Research Service Report No. 270.

Conrad, Z., L.K. Johnson, J.N. Roemmich, W. Juan, and L. Jahns. 2017. "Time Trends and Patterns of Reported Egg Consumption in the U.S. by Sociodemographic Characteristics." *Nutrients* 9:333.

Davis, G.C., and W. You. 2011. "Not Enough Money or Not Enough Time to Satisfy the Thrifty Food Plan? A Cost Difference Approach for Estimating a Money–Time Threshold." *Food Policy* 36:101–107.

Ellison, B., and J.L. Lusk. 2018. "Examining Household Food Waste Decisions: A Vignette Approach." *Applied Economic Perspectives and Policy* 40:613–631.

Epstein, D., A. Mason, and A. Manca. 2008. "The Hospital Costs of Care for Stroke in Nine European Countries." *Health Economics* 17:S21–S31.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde. 2018. "Global Evidence on Economic Preferences." *Quarterly Journal of Economics* 133:1645–1692.

Falk, A., A. Becker, T. Dohmen, D. Huffman, and U. Sunde. 2023. "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences." *Management Science* 69:1935–1950.

Farrell, D. 2013. "The Role of Poultry in Human Nutrition." *Poultry Development Review.*

Food and Agriculture Organization of the United Nations. 1985. *Energy and Protein Requirements: Report of a Joint FAO/WHO/UNU Expert Consultation*. World Health Organization Technical Report Series 724. Rome, Italy: FAO.

Foster, A.C. 2015. "Consumer Expenditures Vary by Age." *Beyond the Numbers: Pricing & Spending* 4. Washington, DC: U.S. Bureau of Labor Statistics.

Greene, W.H. 2012. *Econometric Analysis*, 7th ed. Hoboken, NJ: Prentice Hall.

Griffin, B.A. 2016. "Eggs: Good or Bad?" *Proceedings of the Nutrition Society* 75:259–264.

Gundersen, C., B. Kreider, and J. Pepper. 2011. "The Economics of Food Insecurity in the United States." *Applied Economic Perspectives and Policy* 33:281–303.

Hadley, C., and D.L. Crooks. 2012. "Coping and the Biosocial Consequences of Food Insecurity in the 21st Century." *American Journal of Physical Anthropology* 149:72–94.

Hainmueller, J., D. Hangartner, and T. Yamamoto. 2015. "Validating Vignette and Conjoint Survey Experiments Against Real-World Behavior." *Proceedings of the National Academy of Sciences* 112:2395–2400.

Hamrick, K.S., and D. Hopkins. 2012. "The Time Cost of Access to Food—Distance to the Grocery Store as Measured in Minutes." *Electronic International Journal of Time Use Research* 9:28–58.

Kapteyn, A., J.P. Smith, and A. van Soest. 2007. "Vignettes and Self-Reports of Work Disability in the United States and the Netherlands." *The American Economic Review* 97:461–473.

Kinsey, J. 2005. "Will Food Safety Jeopardize Food Security?" *Agricultural Economics* 32:149–158.

Laraia, B.A., T.M. Leak, J.M. Tester, and C.W. Leung. 2017. "Biobehavioral Factors That Shape Nutrition in Low-Income Populations: A Narrative Review." *American Journal of Preventive Medicine* 52:S118–S126.

Lusk, J.L., and T.C. Schroeder. 2002. "Effects of Meat Recalls on Futures Market Prices." *Agricultural and Resource Economics Review* 31:47–58.

Marsh, T.L., T.C. Schroeder, and J. Mintert. 2004. "Impacts of Meat Product Recalls on Consumer Demand in the USA." *Applied Economics* 36:897–909.

Mayer, A.B., and J.A. Harrison. 2012. "Safe Eats: An Evaluation of the Use of Social Media for Food Safety Education." *Journal of Food Protection* 75:1453–1463.

McKenzie, A.M., and M.R. Thomsen. 2001. "The Effect of 'E. Coli' O157:H7 on Beef Prices." *Journal of Agricultural and Resource Economics* 26:431–444.

Neill, C.L., and S.E. Chen. 2022. "Food Safety Events versus Media: Nonlinear Effects of Eggs Recalls on U.S. Egg Prices." *Journal of Agricultural and Resource Economics* 47:23–37.

Neill, C.L., and R.B. Holcomb. 2019. "Does a Food Safety Label Matter? Consumer Hetero-Geneity and Fresh Produce Risk Perceptions under the Food Safety Modernization Act." *Food Policy* 85:7–14.

Richards, T.J., and W. Nganje. 2014. "Welfare Effects of Food Safety Recalls." *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 62:107–124.

Sayer, L.C. 2005. "Gender, Time and Inequality: Trends in Women's and Men's Paid Work, Unpaid Work and Free Time." *Social Forces* 84:285–303.

Schafer, R.B., E. Schafer, G.L. Bultena, and E.O. Hoiberg. 1993. "Food Safety: An Application of the Health Belief Model." *Journal of Nutrition Education* 25:17–24.

Skinner, J.D., B.R. Carruth, W. Bounds, and P.J. Ziegler. 2002. "Children's Food Preferences: A Longitudinal Analysis." *Journal of the American Dietetic Association* 102:1638–1647.

Thomsen, M.R., R. Shiptsova, and S.J. Hamm. 2006. "Sales Responses to Recalls for Listeria monocytogenes: Evidence from Branded Ready-to-Eat Meats." *Review of Agricultural Economics* 28:482–493.

Thrane, C. 2000. "Men, Women, and Leisure Time: Scandinavian Evidence of Gender Inequality." *Leisure Sciences* 22:109–122.

U.S. Bureau of Labor Statistics. 2022a. *Average Price: Eggs, Grade A, Large (Cost per Dozen) in U.S. City Average* [APU0000708111]. Available online: https://fred.stlouisfed.org/series/APU0000708111.

—. 2022b. *Average Price: Lettuce Romaine (Cost per Pound/453.6 Grams) in U.S. City Average* [APU0000FL2101]. Available online: https://fred.stlouisfed.org/series/APU0000FL2101.

U.S. Department of Agriculture. 2012. *U.S. Household Food Security Module: Three-Stage Design, with Screeners*. Washington, DC: U.S. Department of Agriculture, Economic Research Service. Available online: https://www.ers.usda.gov/media/8271/hh2012.pdf.

—.2015. *Understanding FSIS Food Recalls*. Washington, DC: Department of Agriculture, Food Safety Inspection Service. Available online: https://www.fsis.usda.gov/food-safety/safe-food-handling-and-preparation/food- safety-basics/understanding-fsis-food-recalls.

—.2019. *FoodData Central*. Washington, DC: U.S. Department of Agriculture, Agricultural Research Service. Available online: fdc.nal.usda.gov.

U.S. Food and Drug Administration. 2022. *Food Recalls: What You Need to Know*. Silver Springs, MD: U.S. Food and Drug Administration. Available online: https://www.fda.gov/food/buy-store-serve-safe-food/food-recalls-what-you-need-know.

—.2024. *Recalls, Market Withdrawals, Safety Alerts*. Silver Springs, MD: U.S. Food and Drug Administration. Available online: https://www.fda.gov/safety/recalls-market-withdrawals-safety-alerts.

Van der Lippe, T., J. de Ruijter, E. de Ruijter, and W. Raub. 2011. "Persistent Inequalities in Time Use between Men and Women: A Detailed Look at the Influence of Economic Circumstances, Policies, and Culture." *European Sociological Review* 27:164–179.

Ver Ploeg, M., V. Breneman, T. Farrigan, K. Hamrick, D. Hopkins, P. Kaufman, B.H. Lin, M. Nord, T. Smith, R. Williams, K. Kinnison, C. Olander, A. Singh, E. Tuckermanty, R. Krantz-Kent, C. Polen, H. McGowan, and S. Kim. 2009. *Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences*. Washington, DC: U.S. Department of Agriculture, Economic Research Service.

You, W., and G.C. Davis. 2019. "Estimating Dual Headed Time in Food Production with Implications for SNAP Benefit Adequacy." *Review of Economics of the Household* 17:249–266.

Zaki, M., and J.E. Todd. 2023. "Price Consciousness at the Peak of 'Impatience.'" *Journal of Human Resources* 58(3):1003–1027.

Journal of Food Distribution Research
Volume 55, Issue 3, pp. 43–64

Predicting Firm Diversification in Agri-Food Value Chains

Andrew W. Stevens[a①], Jim Teal[b], Christa D. Court[c], Gigi DiGiacomo[d], Michelle Miller[e], and Hikaru Hanawa Peterson[f]

[a]*Assistant Professor, Department of Agricultural and Applied Economics, 427 Lorch Street, University of Wisconsin–Madison, Madison, WI 53706, USA*

[b]*Research Assistant, Department of Agricultural and Applied Economics, 427 Lorch Street, University of Wisconsin–Madison, Madison, WI 53706, USA*

[c]*Assistant Professor, Food and Resource Economics Department, 1113 McCarty Hall B, University of Florida, Gainesville, FL 32611, USA*

[d]*Research Fellow, Department of Applied Economics, 146 Ruttan Hall, University of Minnesota, St. Paul, MN 55108, USA*

[e]*Senior Researcher III, Center for Integrated Agricultural Systems, 1535 Observatory Drive, University of Wisconsin–Madison, Madison, WI 53706, USA*

[f]*Professor, Department of Applied Economics, 231 Ruttan Hall, University of Minnesota, St. Paul, MN 55108, USA*

Abstract

Diversifying agri-food value chains and the firms within them is one proposed strategy for increasing resilience within the global agri-food sector; however, adapting policy to specific firms based on their level of diversification is challenging in practice due to frequent data limitations. We investigate whether more easily observable firm characteristics can predict diversification for firms in the agri-food value chain, thereby facilitating policy targeting. Using regression analysis of survey-based data from roughly 200 agri-food firms in the United States, we find that few firm characteristics reliably predict diversification, but engagement in direct-to-consumer sales is positively correlated with firm diversification.

Keywords: diversification, agri-food value chains, policy targeting, direct-to-consumer sales

[①]Corresponding author: Tel: (608) 263–3212
Email: awstevens@wisc.edu

Introduction

In recent years, the global agri-food sector has become increasingly organized around complex and interconnected global agri-food value chains (Barrett et al., 2022; Bellemare, Bloem, and Lim, 2022; Lim and Kim, 2022; Montalbano and Nenci, 2022; Lim, 2023). The COVID-19 pandemic has drawn attention to these value chains and raised questions about their efficiency, sustainability, resilience, and ability to innovate in the face of large market shocks (Coopmans et al., 2021; Hobbs, 2021; Mishra, Singh, and Subramanian, 2021; Nordhagen et al., 2021; Weersink et al., 2021; Arita et al, 2022; Ahn and Steinbach, 2023; Azzam, Gren, and Andersson, 2023; DiGiacomo et al., 2023; Ramsey, Goodwin, and Haley, 2023; Hadachek, Ma, and Sexton, 2024). Policy makers are now particularly interested in how to increase the resilience of food supply networks and the agri-food value chains that comprise them: Should agri-food firms be encouraged to specialize and consolidate? Should they be encouraged to diversify? Should they be encouraged to participate in international trade?

Of the many possible approaches to strengthen agri-food value chain resilience, increasing diversification is a frequently proposed strategy, especially in food systems that are highly specialized and efficient. Previous research has explored diversification at the levels of entire supply networks (Choi, 2023; Karakoc et al., 2023), separate supply chains (Stone and Rahimifard, 2018; Hertel et al., 2021), and individual firms (Dorsey and Boland, 2009; Rawley and Simcoe, 2010). At the firm level, Stevens and Teal (2024) document an important distinction between what they call “vertical diversification” (a firm participating in multiple different segments of the agri-food supply chain) and “horizontal diversification” (a firm participating in multiple different activities within individual segments of the supply chain) (see Figure 1).[1] Specifically, they find that vertical diversification reduces firm resilience among small- and medium-sized agri-food firms whereas horizontal diversification increases resilience.

[1]Vertical diversification and horizontal diversification are related to vertical integration and horizontal integration, respectively, but differ in that diversification does not necessarily imply product-level linkages across different segments of the supply chain. For example, a farm that is both vertically diversified and vertically integrated might grow its own corn that it feeds to its own cattle that it raises for beef. However, a different farm could grow corn that it sells as grain and separately raise cattle using feed it buys from other suppliers. This second farm would be vertically diversified, but not vertically integrated.

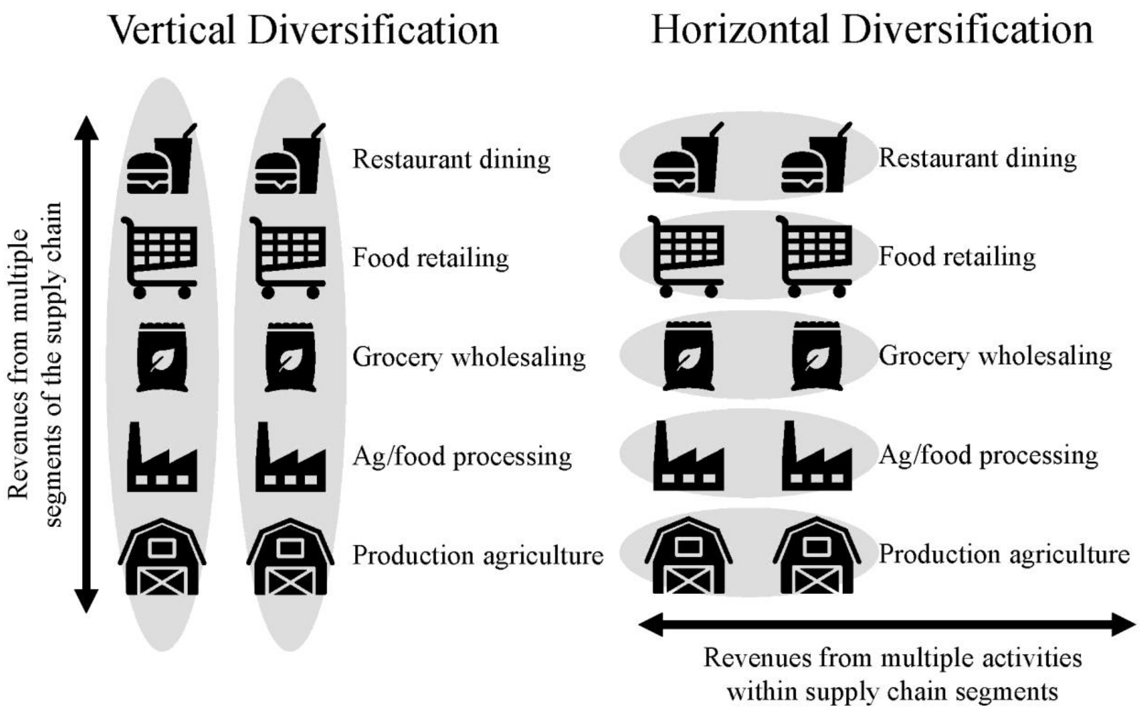

Figure 1. Vertical and Horizontal Diversification in the Agri-Food Value Chain

A shortcoming of much of the research on diversification in the agri-food value chain is that diversification is frequently difficult to observe. At the firm level, for instance, measuring diversification frequently requires detailed and proprietary information about firm expenditures, revenues, or activities. Such information is not generally available, especially to policy makers who might want to target their policies to firms based on levels of diversification. To overcome this challenge, we investigate whether other, more easily observable firm characteristics can consistently predict the degree of firm-level diversification in the agri-food sector. If suitable proxy variables exist for firm diversification, they can be leveraged by policy makers to implement targeted policies based on feasibly observable data.

Existing research on the predictors of diversification for firms operating throughout the agri-food sector is scarce. For large corporate firms, the existing literature in the fields of finance and management on diversification focuses on things like market power, principal-agent problems, or financing constraints (Montgomery, 1994). However, many firms in the agri-food value chain are considerably smaller than the corporations studied in this literature. Within the agri-food sector, research has focused largely on farms rather than on processors or manufacturers. In this literature, farm size, ownership structure, and owner characteristics are frequently identified as factors influencing diversification (Mishra, El-Osta, and Sandretto, 2004; Khanal, 2020; Khanal and Ojha, 2023).

We analyze the data from Stevens and Teal (2024) to determine which firm characteristics—if any—can effectively and consistently predict firm-level vertical diversification and horizontal diversification. Among a sample of U.S. agri-food firms from California, Florida, Minnesota, and Wisconsin, we examine many possible predictors of diversification including firm size (sales revenues, number of employees, etc.), firm ownership (women-owned, veteran-owned, cooperative-owned, etc.), and firm owner characteristics (education level, years of experience, etc.). We analyze potential predictors both individually through unidimensional difference-in-means *t*-tests and collectively through ordinary least squares (OLS) regressions.

Our analyses yield four main findings. First, surprisingly few firm characteristics can consistently predict either vertical or horizontal diversification with any statistical significance. Second, the most consistent predictor of being vertically or horizontally diversified is whether a firm is engaged in direct-to-consumer sales. Third, being located in Florida and being an organic certified firm, respectively, are consistently negatively correlated with firms' levels of vertical diversification. And fourth, engaging in food and beverage retailing is consistently positively correlated with firms' levels of horizontal diversification.

The remainder of the article is organized as follows. First, we briefly describe how we measure firm diversification. Then, we summarize our data. Next, we describe our empirical framework. After that, we present our results and discuss the lessons we can draw from them. We then discuss the policy implications of our findings. Finally, we conclude.

Measuring Firm Diversification

To measure the extent to which firms are diversified across and within segments of the agri-food supply chain, we adapt Stevens and Teal's (2024) normalized measures of vertical and horizontal diversification. These measures are a generalization of the Herfindahl-Hirschman Index but differ in that small values indicate concentration while large values indicate diversification. In our empirical context, which we share with Stevens and Teal (2024), firms were asked in a survey about how their revenues in a typical year were split between six different segments of the agri-food supply chain: production agriculture, processing/manufacturing, grocery wholesaling, food and beverage retailing, restaurant dining, and other. Then, firms were asked about how their revenues within each segment were split across different activities.

Specifically, we define our vertical diversification index VD as

$$VD = \frac{1 - \sum_{i=1}^{n} \left(\frac{r_i}{R}\right)^2}{1 - \frac{1}{n}} \quad (1)$$

where r_i is a firm's revenue from each i of n supply chain segments, n is taken as given and is strictly greater than 1 (in our application, $n = 6$), and R is the firm's total revenue. A VD value of zero signifies a vertically specialized firm whose revenue all comes from a single supply chain segment. A VD value of 1 signifies a "maximally diversified" firm whose revenue is equally split across all possible supply chain segments.

We define our horizontal diversification index *HD* as:

$$HD = \sum_{j=1}^{5}\left(\frac{r_j}{R} \times \left(\frac{1 - \sum_{i=1}^{n_j}\left(\frac{r_i}{r_j}\right)^2}{1 - \frac{1}{n_j}}\right)\right) \quad (2)$$

where j indexes the five named supply chain segments in our data, r_j is the revenue generated from segment j, n_j is the number of activities in segment j, and r_i is the revenue generated from activity i (in segment j). We omit the "other" category from our calculation of *HD* because firms were not asked about how their revenues from their "other" supply chain segment(s) were split across different activities. For additional details about the activities i in each segment j in our empirical application, see Stevens and Teal (2024).

Both *VD* and *HD* are ordinal measures of firm diversification, allowing for firm-to-firm comparisons even if firms are active in different markets or supply chain segments. However, these indices are not cardinal; that is, a firm with a *VD* value of 0.5 is not necessarily "twice as diversified" as a firm with a *VD* value of 0.25. The real strength of these measures for our purposes is that they capture the *extent* to which a firm is diversified. For instance, consider two firms (A and B) that are both active in four supply chain segments. Suppose 85% of firm A's revenue comes from one segment, with the remaining 15% of its revenue split among the remaining three. Then suppose firm B's revenues are split evenly among the four segments: 25% each. In this case, firm B would have a higher value of *VD* than firm A, reflecting the more even distribution of its revenues across different segments.

Data

We analyze firm-level data collected by an online survey conducted in the spring of 2021. The survey targeted firms in four states within the United States—California, Florida, Minnesota, and Wisconsin—and was designed to assess how firms in the agri-food supply chain were impacted by the COVID-19 pandemic. Different firms answered different subsets of questions depending on their business status (closed, temporarily closed, or open) and the supply chain segments in which they operated (production agriculture, processing/manufacturing, grocery wholesaling, food and beverage retailing, restaurant dining, and other). We use information about firms' self-reported pre-pandemic revenue to calculate our measures of *VD* and *HD* as described in equations (1) and (2), respectively. We also observe a variety of other pre-pandemic firm characteristics. A complete list of variables and their definitions can be found in Table 1. For additional information about our data source, see Peterson et al. (2023).

Table 1. Variable Descriptions

Variable	Description
Diversification variables	
VD	Vertical diversification index, see equation (1)
HD	Horizontal diversification index, see equation (2)
Supply chain segments	
productionAg	Dummy for if the firm was active in production agriculture
processing	Dummy for if the firm was active in processing and manufacturing
groceryWholesaling	Dummy for if the firm was active in grocery wholesaling
foodBeverageRetail	Dummy for if the firm was active in food and beverage retailing
restaurant	Dummy for if the firm was active in restaurant dining
other	Dummy for if the firm was active in an unlisted agri-food supply chain segment
Other binary firm characteristics	
WI	Dummy for if the firm is located in Wisconsin
MN	Dummy for if the firm is located in Minnesota
FL	Dummy for if the firm is located in Florida
CA	Dummy for if the firm is located in California
womenOwned	Dummy for if the firm is majority-owned by women
minorityOwned	Dummy for if the firm is majority-owned by ethnic minorities
veteranOwned	Dummy for if the firm is veteran-owned
LGBTOwned	Dummy for if the firm is LGBT-owned
firstGenOwned	Dummy for if the firm owner is first-generation
multiGenOwned	Dummy for if the firm is a multi-generation business
familyOwned	Dummy for if the firm is majority-owned by a single family
franchised	Dummy for if the firm is franchised
cooperative	Dummy for if the firm is a cooperative
organic	Dummy for if the firm is certified organic
LEED	Dummy for if the firm is LEED-certified
BCorp	Dummy for if the firm is a B Corporation
hiringVisa	Dummy for if the firm is authorized to hire H-2A visa workers
ebtPurchases	Dummy for if the firm allows SNAP, WIC, or EBT purchases[1]
onSiteSales	Dummy for if the firm makes on-site sales
directSales	Dummy for if the firm makes retail or direct-to-consumer sales
exportSales	Dummy for if the firm exports any of its products
someCollege	Dummy for if the firm owner (survey respondent) has completed at least some college education
associates	Dummy for if the firm owner (survey respondent) has completed at least an associate's degree
bachelor	Dummy for if the firm owner (survey respondent) has completed at least a bachelor's degree

Table 1. (cont.)

Variable	Description
Other continuous firm characteristics	
salesRevenue	Firm's self-reported sales revenue in 2019, measured in USD
lnSalesRevenue	Natural logarithm of *salesRevenue*
fullTime	Number of full-time employees employed in 2019
partTime	Number of part-time employees employed in 2019
contractLabor	Number of contract labor employees employed in 2019
ownerAge	Age of firm owner (survey respondent)
yearsInOperation	Number of years the firm has been in business
yearsInIndustry	Number of years the firm owner (survey respondent) has worked in their industry

Note: [1]SNAP = supplemental nutrition assistance program, WIC = women, infants, and children program, EBT = electronic benefits transfer.

Although more than 800 firms provided responses to the survey described above, not all responses are usable for our analysis. Specifically, to construct our measures of *VD* and *HD*, a firm must have provided sufficient information about the distribution of its pre-pandemic revenues across different supply chain segments and economic activities. We therefore focus on two samples of our data: our "vertical diversification" sample includes firms for which we can calculate a value for *VD*, and our "horizontal diversification" sample includes firms for which we can calculate a value for *HD*. Although there is considerable overlap between these two samples, they are not identical. We further restrict our sample by omitting firms that: (i) did not report the current status of their business at the time of the survey, (ii) reported a pre-pandemic annual sales revenue of zero dollars or over 98 million U.S. dollars (USD), (iii) were not located in one of the four targeted states (California, Florida, Minnesota, and Wisconsin), (iv) reported having more than 300 full-time employees, (v) reported having more than 200 part-time employees, (vi) reported having more than 40 contract employees, or (vii) reported having zero full-time, part-time, and contract employees. Enforcing these criteria leads us to drop a handful of outlier firms that are not readily comparable to the rest of our sample, which is comprised largely of small- and medium-sized agri-food firms.

After restricting our sample as described above, we are left with 349 firms in our vertical diversification sample and 248 firms in our horizontal diversification sample. However, within each of these samples, not all firms have valid values for all observable characteristics. If we restrict our samples further to only those firms with complete information (as we do in our regression analysis described below), there are 211 firms in the vertical diversification "complete case" sample and 196 firms in the horizontal diversification "complete case" sample representing all six supply chain segments. All firms in the horizontal diversification complete case sample are included in the vertical diversification complete case sample. Summary statistics for both complete case samples are reported in Table 2.

Table 2. Summary Statistics

	***VD* Complete Case Sample (*n* = 211)**				***HD* Complete Case Sample (*n* = 196)**			
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
VD	0.14	0.25	0.00	0.94				
HD					0.29	0.27	0.00	0.94
productionAg	0.24	0.43	0.00	1.00	0.26	0.44	0.00	1.00
processing	0.16	0.36	0.00	1.00	0.17	0.38	0.00	1.00
groceryWholesaling	0.12	0.33	0.00	1.00	0.13	0.34	0.00	1.00
foodBeverageRetail	0.34	0.48	0.00	1.00	0.37	0.48	0.00	1.00
restaurant	0.42	0.49	0.00	1.00	0.45	0.50	0.00	1.00
other	0.18	0.38	0.00	1.00	0.11	0.32	0.00	1.00
WI	0.08	0.27	0.00	1.00	0.08	0.27	0.00	1.00
MN	0.35	0.48	0.00	1.00	0.36	0.48	0.00	1.00
FL	0.04	0.19	0.00	1.00	0.03	0.17	0.00	1.00
CA	0.54	0.50	0.00	1.00	0.54	0.50	0.00	1.00
womenOwned	0.35	0.48	0.00	1.00	0.35	0.48	0.00	1.00
minorityOwned	0.20	0.40	0.00	1.00	0.21	0.41	0.00	1.00
veteranOwned	0.09	0.28	0.00	1.00	0.09	0.29	0.00	1.00
LGBTOwned	0.05	0.21	0.00	1.00	0.05	0.22	0.00	1.00
firstGenOwned	0.44	0.50	0.00	1.00	0.43	0.50	0.00	1.00
multiGenOwned	0.14	0.35	0.00	1.00	0.15	0.36	0.00	1.00
familyOwned	0.51	0.50	0.00	1.00	0.54	0.50	0.00	1.00
franchised	0.06	0.23	0.00	1.00	0.06	0.24	0.00	1.00
cooperative	0.01	0.12	0.00	1.00	0.02	0.12	0.00	1.00
organic	0.14	0.35	0.00	1.00	0.14	0.35	0.00	1.00
LEED	0.01	0.10	0.00	1.00	0.01	0.07	0.00	1.00
BCorp	0.02	0.14	0.00	1.00	0.02	0.14	0.00	1.00
hiringVisa	0.01	0.12	0.00	1.00	0.02	0.12	0.00	1.00
ebtPurchases	0.11	0.32	0.00	1.00	0.11	0.31	0.00	1.00
onSiteSales	0.55	0.50	0.00	1.00	0.55	0.50	0.00	1.00
directSales	0.72	0.45	0.00	1.00	0.73	0.45	0.00	1.00
exportSales	0.05	0.22	0.00	1.00	0.05	0.21	0.00	1.00
someCollege	0.87	0.34	0.00	1.00	0.86	0.35	0.00	1.00
associates	0.69	0.46	0.00	1.00	0.68	0.47	0.00	1.00
bachelor	0.63	0.49	0.00	1.00	0.62	0.49	0.00	1.00
graduate	0.24	0.43	0.00	1.00	0.24	0.43	0.00	1.00
salesRevenue	2,968,120	7,297,821	50.00	49,000,000	3,018,561	7,376,485	50.00	49,000,000
lnSalesRevenue	13.10	2.17	3.93	17.71	13.10	2.21	3.93	17.71
fullTime	10.34	20.93	0.00	200.00	10.82	21.58	0.00	200.00
partTime	16.82	28.51	0.00	175.00	17.63	29.26	0.00	175.00
contractLabor	0.73	2.61	0.00	20.00	0.69	2.66	0.00	20.00
ownerAge	52.64	11.80	23.00	81.00	52.73	12.00	23.00	81.00
yearsInOperation	22.24	19.33	2.00	106.00	22.15	19.51	2.00	106.00
yearsInIndustry	23.88	13.95	2.00	65.00	23.68	14.08	2.00	65.00

Notes: "Complete Case Sample" refers to observations in the *VD* and *HD* samples, respectively, for which we observe data for all listed variables. These "complete case" samples of 211 and 196 observations, respectively, are the same samples used in the regression analyses reported in Table 7.

Empirical Framework

Our objective is to determine which observable firm characteristics—if any—predict a firm's level of vertical or horizontal diversification. We take two different approaches: first, we analyze each firm characteristic in isolation to determine whether it has a statistically significant relationship with either *VD* or *HD* in the relevant sample. Second, we include all observable firm characteristics in a single OLS regression of each index.

In our first approach, we handle binary firm characteristics differently than continuous firm characteristics. For binary characteristics, we compare the mean value of *VD* (or *HD*) among firms that share a particular characteristic to the mean value of *VD* (or *HD*) among firms that do not share the characteristic. We then calculate a *t*-test on the difference in means to determine whether it is statistically significantly different from zero. For continuous characteristics, we compare the mean value of the characteristic among specialized firms ($VD = 0$ or $HD = 0$) to the mean value of the characteristic among diversified firms ($VD > 0$ or $HD > 0$, respectively). We then calculate a *t*-test on the difference in means to determine whether it is statistically significantly different from zero.

In our second approach, we include all binary and continuous firm characteristics in a single OLS regression where the dependent variable is either *VD* or *HD*. Specifically, we estimate equation (3):

$$D_i = \alpha + \boldsymbol{B_i\beta} + \boldsymbol{C_i\gamma} + \varepsilon_i \tag{3}$$

where D_i is either the value of either *VD* or *HD*, as appropriate, for firm i, $\boldsymbol{B_i}$ is a vector of binary characteristics for firm i, $\boldsymbol{C_i}$ is a vector of continuous characteristics for firm i, and ε_i is an error term. Within $\boldsymbol{C_i}$, we include the level and square of four different variables to capture potential non-linear effects: *lnSalesRevenue*, *fullTime*, *partTime*, and *contractLabor*.

When we estimate equation (3) for our horizontal diversification sample, we include binary variables for the six different supply chain segments: production agriculture, processing/manufacturing, grocery wholesaling, food and beverage retailing, restaurant dining, and other. However, we omit these variables in our analysis of the vertical diversification sample since they enter directly into the construction of *VD*. In a supplemental regression for the vertical diversification sample, we include the variable *numSegments*, which is an integer counting the number of supply chain segments in which a firm is active.

Our coefficients of interest from equation (3) are $\boldsymbol{\beta}$ and $\hat{\boldsymbol{\gamma}}$. If any of these coefficients are statistically significantly different from zero, we conclude that they are effective predictors of firm diversification. Importantly, we do not argue that any of these coefficients capture causal effects; we are merely interested in whether observable firm characteristics can reliably predict a firm's level of diversification—not whether these characteristics are the cause of any such diversification.

Results

We begin by presenting our unidimensional findings for firms in our vertical diversification sample. Table 3 presents differences-in-means for binary firm characteristics, and Table 4 presents differences-in-means for continuous firm characteristics. Figure 2 further summarizes the results from Table 3 and includes 95% confidence intervals.

Table 3. Difference in Means Of Vertical Diversification Index (*VD*) by Binary Firm Characteristics

Variable	Mean if False	Mean if True	Difference in Means	*p*-Value of Difference	Sample Size	Number True
WI	0.121	0.224	0.102	0.173	349	20
MN	0.127	0.128	0.000	0.988	349	129
FL	0.133	0.057	-0.075	0.016	349	24
CA	0.129	0.126	-0.003	0.897	349	176
womenOwned	0.105	0.164	0.059	0.079	252	89
minorityOwned	0.137	0.085	-0.052	0.152	252	51
veteranOwned	0.121	0.180	0.059	0.394	252	20
LGBTOwned	0.126	0.132	0.006	0.928	252	13
firstGenOwned	0.127	0.125	-0.002	0.952	252	111
multiGenOwned	0.122	0.153	0.031	0.526	252	35
familyOwned	0.129	0.123	-0.006	0.856	252	129
franchised	0.129	0.078	-0.051	0.537	252	12
cooperative	0.128	0.000	-0.128	0.000	252	4
organic	0.148	0.068	-0.080	0.032	217	31
hiringVisa	0.138	0.054	-0.084	0.219	217	4
ebtPurchases	0.133	0.161	0.027	0.667	217	25
onSiteSales	0.108	0.160	0.052	0.125	217	118
directSales	0.072	0.162	0.089	0.009	217	156
exportSales	0.135	0.158	0.023	0.774	217	11
someCollege	0.113	0.128	0.015	0.713	252	216
associates	0.155	0.113	-0.042	0.211	252	172
bachelor	0.167	0.101	-0.067	0.042	252	156
graduate	0.139	0.085	-0.054	0.094	252	59

Notes: "Number True" refers to the number of firms for which the relevant variable is equal to 1. Statistics in this table calculated using the vertical diversification sample.

Table 4. Difference in Means of Continuous Firm Characteristics by Vertical Diversification

Variable	Mean if *VD* = 0	Mean if *VD* > 0	Difference in Means	*p*-value of Differences	Sample Size	Number *VD* > 0
HD	0.299	0.287	-0.011	0.733	248	72
lnSalesRevenue	13.238	12.873	-0.365	0.143	349	99
salesRevenue	3,150,611	1,816,560	-1,334,051	0.029	349	99
fullTime	14.312	11.990	-2.322	0.475	349	99
partTime	15.660	15.465	-0.195	0.954	349	99
contractLabor	0.677	0.770	0.093	0.776	349	99
ownerAge	22.135	20.963	-1.172	0.688	252	67
yearsInOperation	24.651	21.500	-3.151	0.102	251	66
yearsInIndustry	52.613	51.569	-1.044	0.520	246	65

Notes: Statistics in this table calculated using the vertical diversification sample.

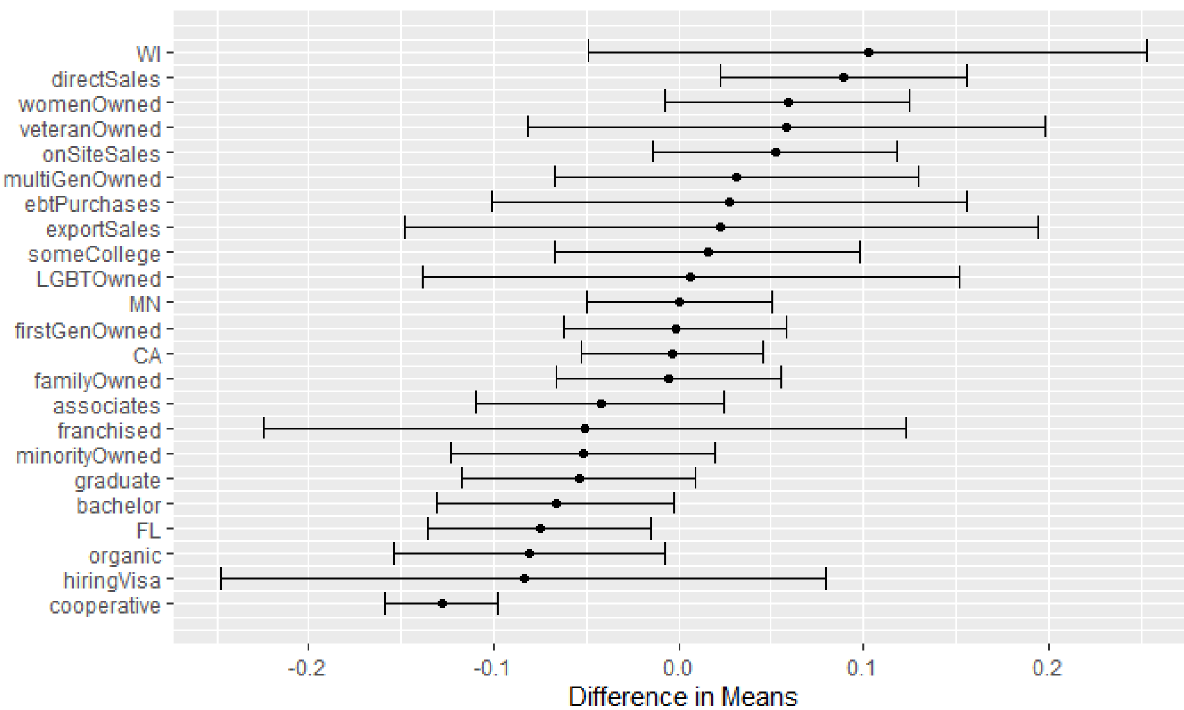

Figure 2. Difference in Means of Vertical Diversification Index (*VD*) by Firm Characteristics

Note: Error bars report 95% confidence intervals.

Overall, we find that few binary firm characteristics seem to be statistically significantly correlated with firms' vertical diversification. Only being engaged in direct sales and being woman-owned seem to be positively correlated with a firm's *VD*, while operating in Florida, being a cooperative, being certified organic, and having a firm owner with a bachelor's degree or graduate education seem to be negatively correlated with a firm's *VD.* Among continuous firm characteristics, only sales revenue (but not its natural logarithm) is statistically significantly correlated with *VD*: Vertically diversified firms have lower sales revenue than vertically specialized firms.

Next, we present our unidimensional findings for firms in our horizontal diversification sample. Table 5 presents differences-in-means for binary firm characteristics, and Table 6 presents differences-in-means for continuous firm characteristics. Figure 3 further summarizes the results from Table 5 and includes 95% confidence intervals.

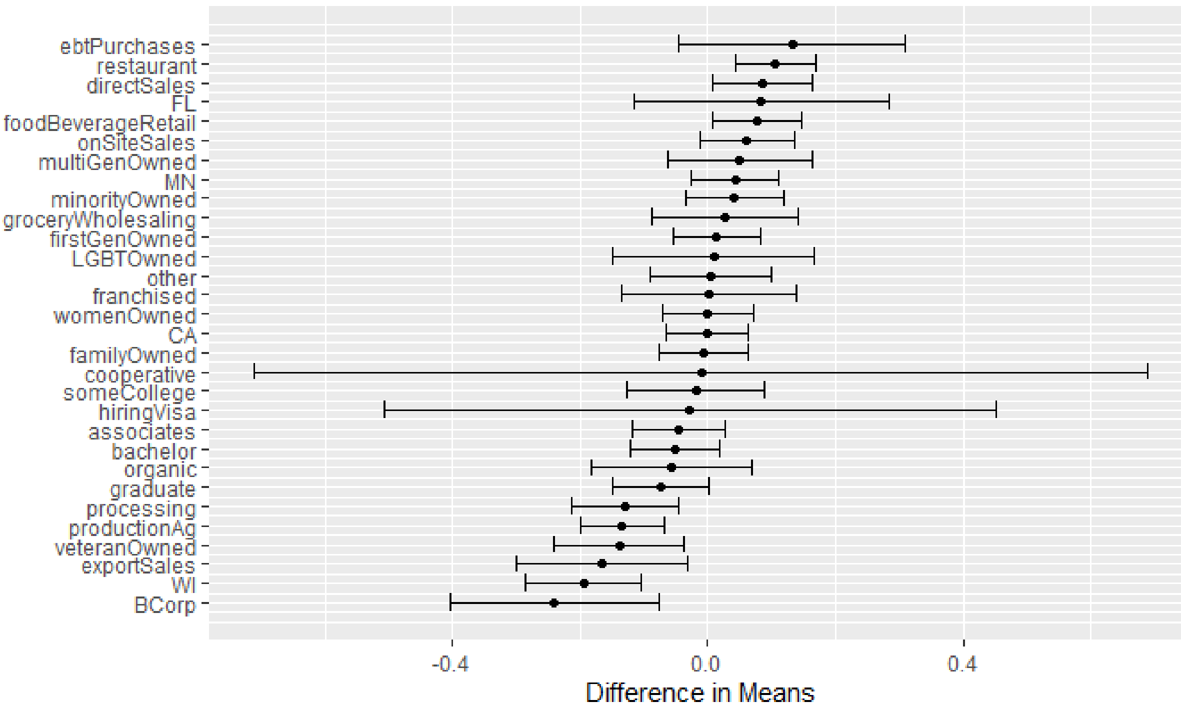

Figure 3: Difference in Means of Horizontal Diversification Index (HD) by Firm Characteristics

Note: Error bars report 95% confidence intervals.

Overall, we find that more firm characteristics seem to be statistically significantly correlated with firms' horizontal diversification. Characteristics that are positively correlated with *HD* include being engaged in direct sales, food and beverage retailing, and restaurant services. Characteristics that are negatively correlated with *HD* include operating in Wisconsin, being engaged in production agriculture, being engaged in agri-food processing or manufacturing, being veteran-owned, being a B Corporation,[2] engaging in export sales, and having a firm owner with a graduate-level education.

Among continuous firm characteristics, those that are positively correlated with *HD* include *VD*, the natural logarithm of sales revenue (but not its level), the number of full-time employees, and

[2]B Corporations are businesses that have received a certification for meeting "high standards of verified performance, accountability, and transparency on factors from employee benefits and charitable giving to supply chain practices and input materials." For more information, see https://bcorporation.net/en-us/certification.

the number of part-time employees. Those that are negatively correlated with *HD* include the number of years the firm owner has worked in their industry and the firm owner's age.

Table 5. Difference in Means of Horizontal Diversification Index (*HD*) by Binary Firm Characteristics

Variable	Mean if False	Mean if True	Difference in Means	*p*-value of Differences	Sample Size	Number True
WI	0.309	0.114	-0.195	0.000	248	17
MN	0.279	0.321	0.043	0.217	248	96
FL	0.293	0.376	0.083	0.363	248	8
CA	0.296	0.295	-0.001	0.973	248	127
productionAg	0.334	0.199	-0.135	0.000	248	71
processing	0.314	0.185	-0.129	0.003	248	36
groceryWholesaling	0.292	0.318	0.026	0.648	248	32
foodBeverageRetail	0.268	0.344	0.076	0.034	248	88
restaurant	0.250	0.355	0.105	0.001	248	106
other	0.295	0.299	0.004	0.936	248	24
womenOwned	0.294	0.294	0.000	0.997	233	81
minorityOwned	0.286	0.327	0.041	0.292	233	49
veteranOwned	0.306	0.167	-0.139	0.009	233	20
LGBTOwned	0.294	0.303	0.009	0.907	233	13
firstGenOwned	0.288	0.302	0.014	0.688	233	101
multiGenOwned	0.287	0.336	0.049	0.381	233	35
familyOwned	0.298	0.291	-0.007	0.848	233	125
franchised	0.294	0.295	0.000	0.994	233	12
cooperative	0.295	0.284	-0.011	0.964	233	4
organic	0.298	0.241	-0.057	0.363	201	29
BCorp	0.294	0.054	-0.240	0.015	201	4
hiringVisa	0.290	0.262	-0.028	0.866	201	4
ebtPurchases	0.275	0.407	0.132	0.139	201	22
onSiteSales	0.256	0.317	0.061	0.108	201	111
directSales	0.227	0.312	0.085	0.034	201	147
exportSales	0.297	0.131	-0.166	0.02	201	9
someCollege	0.311	0.292	-0.019	0.719	233	199
associates	0.325	0.279	-0.046	0.212	233	157
bachelor	0.326	0.274	-0.051	0.149	233	142
graduate	0.311	0.237	-0.074	0.052	233	53

Notes: "Number True" refers to the number of firms for which the relevant variable is equal to 1. Statistics in this table calculated using the horizontal diversification sample.

Table 6. Difference in Means of Continuous Firm Characteristics by Horizontal Diversification

Variable	**Mean if *HD* = 0**	**Mean if *HD* > 0**	**Difference in Means**	***p*-value of Difference**	**Sample Size**	**Number *HD* > 0**
VD	0.0880	0.153	0.065	0.042	248	182
lnSalesRevenue	12.332	13.322	0.990	0.005	248	182
salesRevenue	3,038,009	2,510,552	-527,456	0.667	248	182
fullTime	7.955	12.945	4.991	0.058	248	182
partTime	8.348	19.549	11.201	0.000	248	182
contractLabor	1.000	0.507	-0.493	0.243	248	182
ownerAge	25.841	20.338	-5.503	0.080	233	170
yearsInOperation	24.913	23.148	-1.765	0.467	232	169
yearsInIndustry	55.806	51.084	-4.722	0.014	228	166

Notes: Statistics in this table calculated using the horizontal diversification sample.

Figure 4 presents kernel density plots of the natural logarithm of sales revenue for firms in our *HD* sample (top panel) and *VD* sample (bottom panel). In each case, we report separate density plots for specialized firms (*HD* = 0 and *VD* = 0) and diversified firms (*HD* > 0 and *VD* > 0). This figure helps visualize the relationship between sales revenue and firm diversification and emphasizes that horizontally specialized firms tend to have lower revenue than horizontally diversified firms when measured in natural logarithms. It is notable, however, that we do not find the same relationship in levels, highlighting the statistical importance of firms with particularly large sales revenues when analyzing our data in levels.

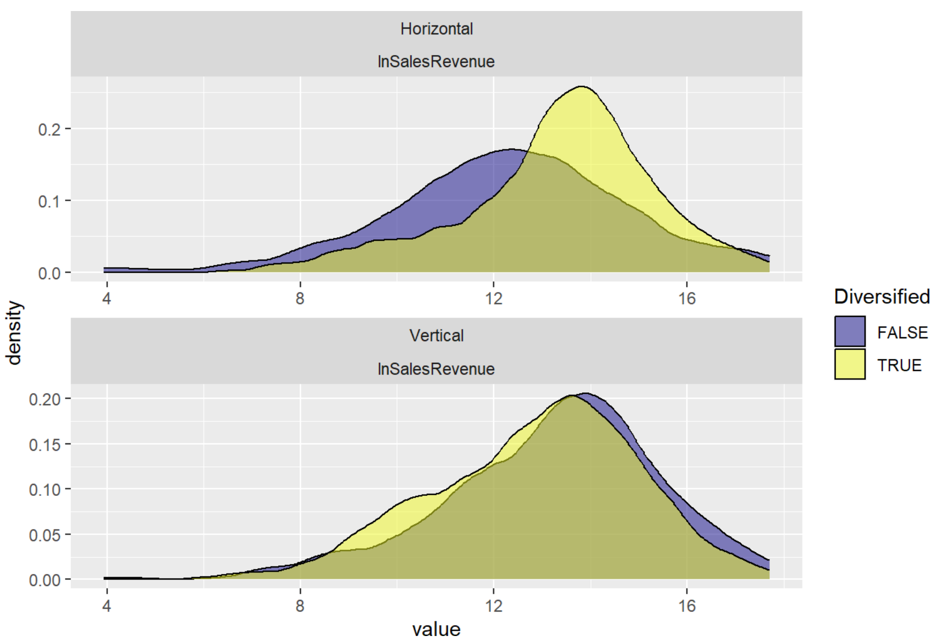

Figure 4. Kernel Density of Sales Revenue by Diversification Type

Note: Both panels report the density of firms in our sample measured by the natural logarithm of sales revenue (USD). The top panel includes firms in the horizontal diversification sample and separates firms by whether *HD* = 0 (Diversified = FALSE) or *HD* > 0 (Diversified = TRUE). The bottom panel includes firms in the vertical diversification sample and separates firms by whether *VD* = 0 (Diversified = FALSE) or *VD* > 0 (Diversified = TRUE).

Finally, we present our findings from OLS regressions of *VD* and *HD* on the full set of binary and continuous firm characteristics as described in equation (3). Table 7 contains our results with columns (1) and (2) analyzing *VD* and column (3) analyzing *HD*. Neither column (1) nor column (2) includes the supply chain segment dummy variables because they enter directly into the construction of *VD*: Being active in any particular supply chain segment increases a firm's level of *VD* mechanically. However, in column (2), we include the variable *numSegments* to try and explain more of the variation in *VD* without attributing importance to any segment over another. Unsurprisingly, including *numSegments* dramatically increases our model fit; however, because this variable also enters directly into the construction of *VD* and therefore alters the underlying assumptions of the functional form in our regression model, we are hesitant to over-rely on these results. We therefore present columns (1) and (2) as complementary analyses that should be interpreted together.

In columns (1) and (2) of Table 7, we find that relatively few firm characteristics are statistically significant predictors of firms' vertical diversification. Characteristics that are positively correlated with *VD* include the firm engaging in direct sales, engaging in export sales, and being veteran-owned. Characteristics negatively correlated with *VD* include operating in Florida, being certified organic, and the firm owner having relatively more years of experience in their industry. We also find that the number of contract labor employees is a statistically significant predictor of *VD*: depending on the specification, the relationship is positive when the number of contract labor employees is below seven (column [1]) or nine (column [2]) and negative when the number of contract labor employees is above these respective values.

In column (3) of Table 7, we find that even fewer firm characteristics are statistically significant predictors of firms' horizontal diversification. Characteristics that are positively correlated with *HD* include the firm engaging in direct sales, being active in the food and beverage retailing sector, and allowing for SNAP, WIC, and EBT purchases.[3] The only characteristic that is negatively correlated with *HD* is the number of years the firm has been in operation.

In general, we prefer the regression analyses reported in Table 7 over the unidimensional analyses reported in Tables 3–6 because the regression analyses account for correlations among different firm-level characteristics. Nonetheless, comparing the unidimensional results to the OLS results can help paint a more complete picture of how different firm characteristics relate to one another and firms' diversification levels.

[3]SNAP = supplemental nutrition assistance program, WIC = women, infants, and children program, EBT = electronic benefits transfer.

Table 7. Ordinary Least Squares Regression of Diversification Indices on Firm Characteristics

Variable	***VD*** **(1)**	***VD*** **(2)**	***HD*** **(3)**
(Intercept)	−0.142	−0.244	0.31
	(0.42)	(0.21)	(0.45)
WI	0.03	0.02	−0.059
	(0.07)	(0.04)	(0.08)
MN	−0.003	−0.006	−0.004
	(0.04)	(0.02)	(0.05)
FL	−0.228**	−0.054	0.18
	(0.10)	(0.05)	(0.12)
lnSalesRevenue	0.05	0.01	−0.038
	(0.07)	(0.03)	(0.07)
lnSalesRevenueSQ	−0.002	0.00	0.00
	(0.00)	(0.00)	(0.00)
fullTime	0.00	0.00	−0.002
	(0.00)	(0.00)	(0.00)
fullTimeSQ	0.00	0.00	0.00
	0.00	0.00	0.00
partTime	−0.002	0.00	0.00
	(0.00)	(0.00)	(0.00)
partTimeSQ	0.00	0.00	0.00
	0.00	0.00	0.00
contractLabor	0.040**	0.017*	0.01
	(0.02)	(0.01)	(0.02)
contractLaborSQ	−0.003**	−0.001*	−0.001
	(0.00)	(0.00)	(0.00)
yearsInOperation	0.00	0.00	−0.002*
	(0.00)	(0.00)	(0.00)
yearsInIndustry	−0.003*	−0.001	0.00
	(0.00)	(0.00)	(0.00)
womenOwned	0.06	0.01	0.00
	(0.04)	(0.02)	(0.04)
minorityOwned	−0.042	−0.028	−0.013
	(0.05)	(0.03)	(0.05)
veteranOwned	0.134*	0.01	−0.018
	(0.07)	(0.04)	(0.07)
LGBTOwned	0.05	0.04	−0.079
	(0.09)	(0.04)	(0.09)
firstGenOwned	0.02	−0.002	0.07
	(0.04)	(0.02)	(0.04)
multiGenOwned	0.07	0.01	0.09
	(0.06)	(0.03)	(0.06)

Table 7. (cont.)

Variable	***VD*** **(1)**	***VD*** **(2)**	***HD*** **(3)**
familyOwned	−0.015	0.01	−0.044
	(0.04)	(0.02)	(0.04)
franchised	−0.022	−0.001	−0.118
	(0.08)	(0.04)	(0.08)
cooperative	−0.100	−0.012	0.16
	(0.18)	(0.09)	(0.19)
ownerAge	0.00	0.00	−0.002
	(0.00)	(0.00)	(0.00)
organic	−0.130**	−0.028	−0.049
	(0.06)	(0.03)	(0.06)
LEED	−0.203	−0.079	0.11
	(0.22)	(0.11)	(0.42)
BCorp	−0.010	0.01	−0.059
	(0.15)	(0.07)	(0.16)
hiringVisa	−0.045	−0.001	−0.129
	(0.17)	(0.08)	(0.19)
ebtPurchases	0.05	0.00	0.190***
	(0.06)	(0.03)	(0.06)
onSiteSales	0.03	0.02	0.05
	(0.04)	(0.02)	(0.04)
directSales	0.084**	0.050**	0.087*
	(0.04)	(0.02)	(0.05)
exportSales	0.15	0.092**	−0.094
	(0.09)	(0.05)	(0.11)
someCollege	0.05	0.01	0.01
	(0.07)	(0.03)	(0.07)
associates	0.02	−0.013	−0.045
	(0.09)	(0.04)	(0.09)
bachelor	−0.112	−0.034	0.01
	(0.08)	(0.04)	(0.08)
graduate	−0.023	−0.002	−0.039
	(0.05)	(0.02)	(0.05)
numSegments		0.233***	
		(0.01)	
productionAg			−0.029
			(0.06)
processing			−0.084
			(0.06)
groceryWholesaling			0.00
			(0.06)
foodBeverageRetail			0.090*
			(0.05)
restaurant			0.08
			(0.05)
other			−0.086
			(0.06)

Table 7. (cont.)

Variable	*VD* (1)	*VD* (2)	*HD* (3)
Num.Obs.	211	211	196
R2	0.21	0.81	0.35
R2 Adj.	0.05	0.77	0.17
AIC	40.00	−257.9	39.00
BIC	164.10	−130.5	180.00
Log.Lik.	16.98	166.94	23.49
RMSE	0.22	0.11	0.21

Notes: For columns (1) and (2), we use the vertical diversification sample and *VD* is the dependent variable. For column (3), we use the horizontal diversification sample and *HD* is the dependent variable. In column (2), *numSegments* is an integer counting the number of supply chain segments (*productiongAg*, *processing*, etc.) in which a firm is active. OLS standard errors are reported in parentheses. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Discussion

Taken together, our results point to four conclusions about the potential for using firm characteristics to predict firm-level diversification in the agri-food supply chain:

Overall, very few firm characteristics predict firm-level diversification.

Perhaps surprisingly, we do not find consistent evidence that firms with greater sales revenue or larger workforces are any more or less likely to be diversified than smaller firms. Relatedly, we do not find evidence that being woman-owned, minority-owned, or cooperatively owned consistently predicts a firm's level of diversification. In this sense, many of the variables that would be natural candidates to be proxies for firm diversification in the agri-food supply chain fall short.

More broadly, the general lack of statistical significance among firm characteristics and low measures of model fit from our regression analyses suggest that firm diversification is difficult to predict even with a rich set of firm characteristics. Our results suggest there are few, if any, good ways to assess a firm's level of diversification without measuring it directly.

Engaging in direct sales is the most consistent predictor of increased firm diversification.

The only observable firm characteristic that is a statistically significant predictor of both vertical and horizontal diversification in both our unidimensional and regression analyses is whether a firm is engaged in retail or direct-to-consumer sales. This result is perhaps unsurprising, but it highlights how direct sales can be complementary to other activities throughout the agri-food supply chain, such as production or processing. Furthermore, firms engaged in direct sales likely have a sales infrastructure that can be readily adapted to various product categories to take advantage of different market opportunities or to engage consumers in using different sales strategies.

Geographic location (being located in Florida) and organic certification are consistently negatively correlated with firms' levels of vertical diversification.

Beyond being engaged in direct sales, the only two firm characteristics that are statistically significant in both the unidimensional and regression analyses of VD are whether a firm is located in Florida and whether a firm is certified organic. Although these results might not be very generalizable (our dataset only includes firms from four U.S. states and is certainly not nationally representative), they suggest that some specialization in the agri-food supply chain might be predictable. Florida agri-food firms and organic firms—both more likely focused on production agriculture of specialty crops—are less likely to expand to other segments of the supply chain, perhaps due to inefficiencies of scope (Rawley and Simcoe, 2010; Court et al., 2023) or restrictive growing contracts for specific crops.

Being engaged in food and beverage retailing is consistently positively correlated with firms' levels of horizontal diversification.

Beyond being engaged in direct sales, being engaged in food and beverage retailing is the only other firm characteristic that is a statistically significant predictor of horizontal diversification in both our unidimensional and regression analyses. Although this result is not terribly surprising—it is easy to imagine food retailers leveraging their experience and infrastructure to sell a variety of different goods—it is notable that food and beverage retailing is such a strong predictor of horizontal diversification apart from and in addition to the effects of being engaged in direct sales.

Conclusion

Understanding how firms' diversification decisions impact their resilience and the resilience of their supply chains is important for understanding how global agri-food value chains function. We extend the analysis in Stevens and Teal (2024) to investigate which observable firm characteristics—if any—can consistently predict firms' levels of vertical and horizontal diversification. In the U.S. context, we find that surprisingly few characteristics have strong predictive power. The most consistent predictor is whether a firm engages in direct-to-consumer sales. Firms that do tend to be more vertically diversified and more horizontally diversified.

Our findings suggest that conditions or policies that increase the number of firms engaged in direct-to-consumer sales might also increase firm diversification both across and within supply chain segments, thereby increasing value chain resilience. It is important to note that our empirical results are not causal, meaning we cannot conclude that increasing firms' adoption of direct-to-consumer sales will necessarily increase their diversity; however, direct-to-consumer sales are the single most consistent predictor of firm diversification across all our analyses. We also emphasize that our findings may be limited in their external validity given the limited geographic and temporal scope of our data.

Nonetheless, the over-arching implication of our analysis is that there are no good proxy variables for diversification among firms in the agri-food sector. Policies that intend to target diversified or specialized firms will need to consider strategies for observing and analyzing firms' levels of diversification. Given the necessity of detailed and proprietary information in such analyses, the feasibility of such policies is questionable.

Acknowledgements

We thank our collaborators Lauri Baker, Gustavo Oliveira, and Li Zhang. The survey questionnaire described in this article was reviewed and approved by the institutional review boards of the University of California, Irvine (protocol UCI IRBAPP #15202), the University of Florida (protocol IRB202002280), and the University of Minnesota (protocol STUDY00010619).

Funding Information

This work was supported by the Agriculture and Food Research Initiative, grant no. 2020-68006-33037, from the USDA National Institute of Food and Agriculture.

References

Ahn, S., and S. Steinbach. 2023. "Agri-food Trade Resilience among Food-Deficit Countries during the COVID-19 Pandemic." *International Food and Agribusiness Management Review* 26:397–408.

Arita, S., J. Grant, S. Sydow, and J. Beckman. 2022. "Has Global Agricultural Trade Been Resilient under Coronavirus (COVID-19)? Findings from an Econometric Assessment of 2020." *Food Policy* 107:102204.

Azzam, A., I-M. Gren, and H. Andersson. 2023. "Comparative Resilience of US and EU Meat Processing to the COVID-19 Pandemic." *Food Policy* 119:102517.

Barrett, C.B., T. Reardon, J. Swinnen, and D. Zilberman. 2022. "Agri- food Value Chain Revolutions in Low- and Middle-Income Countries." *Journal of Economic Literature* 60:1316–1377.

Bellemare, M.F., J.R. Bloem, and S. Lim. 2022. "Chapter 89 - Producers, Consumers, and Value Chains in Low- and Middle-income Countries." In C.B. Barrett and D.R. Just, eds. *Handbook of Agricultural Economics,* vol. 6. Amsterdam, Netherlands: Elsevier pp. 4933–4996.

Choi, T.Y. 2023. *The Nature of Supply Networks*. Oxford, England: Oxford University Press.

Coopmans, I., J. Bijttebier, F. Marchand, E. Mathijs, L. Messely, E. Rogge, A. Sanders, and E. Wauters. 2021. "COVID-19 Impacts on Flemish Food Supply Chains and Lessons for Agri-Food System Resilience." *Agricultural Systems* 190:103–136.

Court, C.D., D. Outerbridge, L. Baker, L. Birou, C. Campbell, G. DiGiacomo, S. Galindo, J. Lai, A. Magnier, M. Miller, G. de L. T. Oliveira, E. Orlando, H.H. Peterson, X. Qiao, F. Roka, A. Ropicki, B.B. Saha, A.W. Stevens, and L. Zhang. 2023. "Pandemic Produce: Impacts of COVID-19 on Florida's Fruit and Vegetable Industries." *Journal of Agricultural & Food Industrial Organization* 21:69–88.

DiGiacomo, G., L. Baker, C.-X. Yang, and H. Peterson. 2023. "Research Report: Innovation among Businesses across the Agri-Food Supply Chain during COVID-19." *Journal of Food Distribution Research* 54:8–16.

Dorsey, S., and M. Boland. 2009. "The Impact of Integration Strategies on Food Business Firm Value." *Journal of Agricultural and Applied Economics* 41:585–598.

Hadachek, J., M. Ma, and R.J. Sexton. 2024. "Market Structure and Resilience of Food Supply Chains under Extreme Events." *American Journal of Agricultural Economics* 106:21–44.

Hertel, T., I. Elouafi, M. Tanticharoen, and F. Ewert. 2021. "Diversification for Enhanced Food Systems Resilience." *Nature Food* 2:832–834.

Hobbs, J.E. 2021. "Food Supply Chain Resilience and the COVID-19 Pandemic: What Have We Learned?" *Canadian Journal of Agricultural Economics* 69:189–196.

Karakoc, D.B., M. Konar, M.J. Puma, and L.R. Varshney. 2023. "Structural Chokepoints Determine the Resilience of Agri-Food Supply Chains in the United States." *Nature Food* 4:607–615.

Khanal, A.R. 2020. "Interlinked Diversification Strategies: Evidence from the US Farm Business Households." *Journal of Agribusiness in Developing and Emerging Economies* 10:253–268.

Khanal, A.R., and R. Ojha. 2023. "The Financial Implications of Specialization, Diversification, or Alternative Enterprises on Small Farms: Evidence from Tennessee." *Agricultural and Resource Economics Review* 52:604–622.

Lim, S. "Global Agricultural Value Chains and Structural Transformation." 2023. In P. Antràs and D. Zilberman, eds. *Risks in Agricultural Supply Chains*. Chicago, IL: University of Chicago Press, pp. 29–57.

Camerer, C.F. 1992. "Recent Tests of Generalizations of Expected Utility Theory." In W. Edwards, ed. *Utility Theories: Measurements and Applications*. Dordrecht, Netherlands: Springer, pp. 207–251.

Lim, S., and S.W. Kim. 2022. "Global Agricultural Value Chains and Employment Growth." *Journal of the Agricultural and Applied Economics Association* 1:402–418.

Mishra, A.K., H.S. El-Osta, and C.L. Sandretto. 2004. "Factors Affecting Farm Enterprise Diversification." *Agricultural Finance Review* 64:151–166.

Mishra, R., R.K. Singh, and N. Subramanian. 2021. "Impact of Disruptions in Agri-Food Supply Chain Due to COVID-19 Pandemic: Contextualized Resilience Framework to Achieve Operational Excellence." *The International Journal of Logistics Management* 33:926–954.

Montalbano, P., and S. Nenci. 2022. "Does Global Value Chain Participation and Positioning in the Agriculture and Food Sectors Affect Economic Performance? A Global Assessment." *Food Policy* 108:102235.

Montgomery, C.A. 1994. "Corporate Diversification." *Journal of Economic Perspectives* 8:163–178.

Nordhagen, S., U. Igbeka, H. Rowlands, R.S. Shine, E. Heneghan, and J. Tench. 2021. "COVID-19 and Small Enterprises in the Food Supply Chain: Early Impacts and Implications for Longer-Term Food System Resilience in Low- and Middle- Income Countries." *World Development* 141:105405.

Peterson, H.H., G. DiGiacomo, C.D. Court, M. Miller, G. Oliveira, A.W. Stevens, L. Zhang, L.M. Baker, J. Nowak, E. Orlando, and B.B. Saha. 2023. "Impacts of COVID-19 on US Agri-Food Supply Chain Businesses: Regional Survey Results." *PLoS One* 18:e0281930.

Ramsey, A.F, B.K. Goodwin, and M.M. Haley. 2023. "Labor Dynamics and Supply Chain Disruption in Food Manufacturing." In P. Antràs and D. Zilberman, eds. *Risks in Agricultural Supply Chains*. Chicago, IL: University of Chicago Press, pp. 187–216.

Rawley, E., and T.S. Simcoe. 2010. "Diversification, Diseconomies of Scope, and Vertical Contracting: Evidence from the Taxicab Industry." *Management Science* 56:1534–1550.

Stevens, A.W., and J. Teal. 2024. "Diversification and Resilience of Firms in the Agrifood Supply Chain." *American Journal of Agricultural Economics* 106:739–778.

Stone, J., and S. Rahimifard. 2018. "Resilience in Agri-Food Supply Chains: A Critical Analysis of the Literature and Synthesis of a Novel Framework." *Supply Chain Management* 23:207–238.

Weersink, A., M. von Massow, N. Bannon, J. Ifft, J. Maples, K. McEwan, M. G.S. McKendree, C. Nicholson, A. Novakovic, A. Rangarajan, T. Richards, B. Rickard, J. Rude, M. Schipanski, G. Schnitkey, L. Schulz, D. Schuurman, K. Schwartzkopf-Genswein, M. Stephenson, J. Thompson, and K. Wood. 2021. "COVID-19 and the Agri-Food System in the United States and Canada." *Agricultural Systems* 188:103039.