Browsing the Aisles or Browsing the App? How Online Grocery Shopping is Changing What We Buy

Sai Chand Chintala* Jūra Liaukonytė† Nathan Yang‡

August 3, 2023

Abstract

This paper investigates the systematic differences between online and offline grocery shopping baskets using data from approximately two million brick-and-mortar and Instacart trips. We apply unsupervised machine learning algorithms agnostic to the shopping channel to identify what constitutes a typical food shopping trip for each household. We find that food shopping basket variety is significantly lower for online shopping trips, as measured by the number of unique food categories and items purchased. Within a given household, the Instacart baskets are more similar to each other as compared to offline baskets, with twice as many overlapping items between successive trips to the same retailer. These results suggest a potential link between online grocery shopping environments and heightened consumer inertia, which may lead to stronger brand loyalty and pose challenges for new entrants in establishing a customer base. Furthermore, Instacart baskets have 13% fewer fresh vegetables and 5-7% fewer impulse purchases, such as candy, bakery desserts, and savory snacks, which are not compensated for by alternative or additional shopping trips. We discuss the implications of these systematic shopping basket differences for competition, product management, retailers, consumers, and online platforms.

Keywords: Digitization; Food Marketing; Omnichannel Retail; Grocery Industry; Variety

---

*Cornell University, SC Johnson College of Business. Email: sc2947@cornell.edu.

†Cornell University, SC Johnson College of Business. Email: jurate@cornell.edu.

‡Cornell University, SC Johnson College of Business. Email: ncy6@cornell.edu.

The author names are listed in alphabetical order. The authors thank conference and seminar participants from the Canadian Centre for Health Economics, AMA Winter Conference, Vienna University of Economics and Business, Marketing Science Conference, NBER Digitization Tutorial, Temple University, Instacart Data Science Seminar, Erasmus University, McGill University, Virtual Digital Economy Seminar, Marketing Modelers Group in New York, Arizona State University, eNumerate 2021, Baltic Economic Association, and Cornell University, as well as Anne Byrne, Daniel Hooker, Ben Leyden, Peggy Liu, Daniel McCarthy, Mike Palazzolo, Adithya Pattabhiramaiah, Davide Proserpio, Andrey Simonov, Anna Tuchman, Kosuke Uetake, Kenneth Wilbur, Alminas Zaldokas, and Xinrong Zhu for helpful comments and discussions. Instacart was not involved in this research.

1 Introduction

Online shopping has significantly transformed the retail landscape, with the grocery sector being among the most impacted in recent years. The Covid-19 pandemic has further accelerated this shift, leading to a sharp rise in demand for online grocery services. In the United States, around 54% of households placed an online grocery order in March 2021, which was a 328% increase from August 2019 (Mercatus 2020, 2021). Instacart, the leading online grocery delivery service and the focus of this paper, also saw a multifold year-on-year growth after the onset of the pandemic (Sorvino 2021). Experts predict that this trend is likely to continue in the post-pandemic era, as consumers become more accustomed to the convenience of online grocery shopping (Hussey 2021). Indeed, according to a recent post-Covid industry survey, consumers have expressed sustained interest in purchasing groceries online (McKinsey 2022), and Instacart sales further increased by 39% in 2022 (WSJ 2023).

Over the past few decades, conventional brick-and-mortar grocery stores have optimized their physical space including store layouts, shelf displays, and end-caps, all with the primary objective of maximizing revenue. In the digital sphere, the grocery shopping experience predominantly revolves around a mobile app-based interface, resulting in reduced reliance on physical store arrangements and increasing the importance of "digital shelves". As the shift towards online grocery shopping continues, it is important to understand how this channel shapes consumer behavior. Therefore, our paper aims to investigate the systematic differences between online and brick-and-mortar shopping baskets.

To address this question, we leverage a large-scale panel dataset that captures a household's entire online and offline food shopping activity across multiple retailers. This comprehensive "360 degree view" addresses a formidable data challenge in omnichannel research (Cui et al. 2021). We begin by examining the overall grocery shopping patterns across both online and offline channels, for those households that use both. We find that the households that adopted online grocery shopping make grocery purchases from three distinct online retailers, on average. Furthermore, households use Instacart services from one to two preferred retailers, typically coinciding with the retailers they frequent the most for their regular brick-and-mortar grocery shopping. Online shopping baskets generally consist of a higher proportion of bulk purchases and exhibit less diversity in their contents, suggesting that the choice of online channels may be primarily driven by convenience factors. Additionally, both Instacart and other online baskets show higher spending and a larger quantity of items bought in comparison to the average brick-and-mortar grocery baskets. While these overall shopping patterns provide valuable insights, they mask considerable heterogeneity in trip characteristics within individual households. Our data consist of various types of shopping trips, including routine restocking of commonly used grocery items and intermittent visits to convenience stores or gas stations. To address the trip heterogeneity, we utilize machine learning techniques to analyze the shopping patterns of each household with the goal of identifying their regular restocking grocery trips, which we refer to as *characteristic* trips. Our analysis demonstrates that our algorithm effectively distinguishes a plausible substitution of a brick-and-mortar characteristic trip with an Instacart trip.

When we restrict the comparison of Instacart and offline baskets focusing solely on the characteristic trips, we document four important results. First, we find that, on average, online baskets exhibit significantly lower variety than brick-and-mortar baskets, as measured by the number of unique food categories and items purchased. Specifically, we find that online basket variety is 9.6% lower at the category level and 14.1% lower at the item level. Second, we find that these Instacart trips are 27% more similar to each other than offline trips within the same household when comparing categories. A more granular item-level similarity analysis indicates that Instacart shopping trips have twice as much item overlap between any two successive trips within the same retailer. Additionally, we find that although households with more experience using the online channel tend to show less pronounced differences in variety and similarity, the gap in these measures persists even among the most experienced users. One potential explanation for these differences in similarity may be online shopping platforms' adoption of user-friendly interfaces that simplify the shopping process for consumers by recommending their previous purchases to them as they assemble their online baskets (IRI Worldwide 2020). Third, we document the systematic basket composition differences between characteristic online and brick-and-mortar baskets. Specifically, we find that Instacart shopping baskets have 13.6% fewer fresh vegetables compared to the brick-and-mortar shopping baskets. At the same time, we see fewer impulse purchase categories such as candy (7.1% fewer), bakery desserts (5.9%), and savory snacks such as chips (4.7%). Fourth, we investigate whether households compensate for the fewer fresh vegetable and impulse purchases by modifying their shopping behavior via alternative or additional trips. We find no evidence of compensatory shopping behavior or any adjustments in eating out or using food delivery services such as GrubHub or UberEats.

Our research contributions overlap with the proposed future research directions for offline-online retail research (Ratchford et al. 2022). Specifically, our work contributes to the stream of literature about omnichannel retail, the impact of digitization on shopping behaviors, as well as the literature on inertial and habitual behavior.

Omnichannel retail. Our research offers new insights about omnichannel strategies (Neslin 2022). Much of the past research in this area is focused on determining whether offline and online shopping help or hurt each other (Bell et al. 2018, Li 2020, Narang and Shankar 2019, Wang and Goldfarb 2017), differences in price sensitivity, and customer loyalty or search (Chu et al. 2008, Danaher et al. 2003, Degeratu et al. 2000), as well as strategies for omnichannel retail (Ertekin et al. 2021, Luo and Sun 2016). Furthermore, other research has shown a positive effect of offline shopping on online channels (Bell et al. 2018, Li 2020, Wang and Goldfarb 2017). Our research, on the other hand, highlights the overall differences between offline and online shopping behaviors, as well as differences in the shopping behavior when the primary goal for the shopper remains the same—to complete a routine grocery shopping trip.

Digitization and shopping behavior. The impact of digitization on shopping behavior has yielded mixed findings about its eventual effect on variety, as some research has found that online channels (Brynjolfsson et al. 2011, Choi and Bell 2011, Datta et al. 2018, Donnelly et al. 2023, Nagaraj and Reimers 2021, Zentner et al. 2013), popularity information (Tucker and Zhang 2011) and recommendation systems (Fleder and Hosanagar 2009, Li et al. 2022, Oestreicher-Singer and Sundararajan 2012) increase consumption variety or consumption of niche products (and consequently, welfare), while others have found the opposite effect (Holtz et al. 2020, Pozzi 2012). Our findings are more consistent with the latter, as we find that Instacart grocery purchases exhibit lower variety, and the past purchase shortcuts might be contributing towards creating filter bubbles and echo chambers in consumption patterns (Ge et al. 2020). Most related to us are studies that investigate the impact of online channels on grocery shopping behavior (Harris-Lagoudakis 2021, Huyghe et al. 2017, Milkman et al. 2010) in single retailer settings. We significantly depart from this past literature both in terms of our research design and key findings. Unlike the above-mentioned studies that focus on *one* retailer in a pre-pandemic setting, our sample covers a time period with significantly higher online grocery delivery services penetration as well as households across all 50 states and 98,851 unique store locations (6,760 unique stores with Instacart delivery). Importantly, our empirical setting provides us a comprehensive picture of household grocery shopping behavior across multiple retailers, allowing us to evaluate potential compensatory behavior across other trips to fully understand a wider timeline of consumption behavior.

Inertial purchasing behavior. Past research on state dependence in consumption behavior is mature and has provided many important insights about inertial behavior—stickiness in customer purchase or usage choices—stemming from customer loyalty or habit formation (Bronnenberg et al. 2012, Chintagunta 1998, Dubé et al. 2010, Erdem and Sun 2001). Understanding whether online channels might accelerate inertial behavior is particularly important as inertial shopping behavior is often used to justify various marketing interventions, including advertising (Freimer and Horsky 2012), free sample and freemium design (Bawa and Shoemaker 2004), product line decisions (Chintagunta 1998), and pricing promotions (Gupta et al. 1997). Indeed, Danaher et al. (2003) and Guo and Wang (2023) provide evidence that customers demonstrate higher brand loyalty in the online channel compared to offline. While establishing the causal link between online grocery shopping and accelerated inertial behavior remains a topic for future research, our results indicate a noteworthy pattern: online baskets tend to exhibit greater similarity than offline ones within the same household. This suggests the possibility of distinct inertial patterns across trips across different channels and may have significant implications for competition, as inertia can impede new entrants from establishing a customer base (Bornstein 2020, Pozzi 2012).

2 Empirical Setting

2.1 Data and Sample Construction

Our data come from Numerator, a market research company with a representative US consumer panel. The company uses a popular mobile app to capture the images of receipts from brick-and-mortar (hereafter referred to as BM) stores uploaded by the panelists and proprietary methods to access online grocery purchases of the same panelists. Our dataset includes information on each panelist and their shopping trips, including the date of the trip, the name of the retailer, the channel type, the total amount spent on an item, the total quantity purchased, and the category and department to which the item belongs. Numerator extracts this information via algorithms that classify items from receipts into standardized categories across all retailers. These categories are highly granular product groups such as "packaged cookies," "yogurt & yogurt drinks," and "stocks & broths." One limitation of our data is that Numerator provided us with a subset of data limited to the grocery sector (i.e., food) items. As a result, while we can observe all the food items and the overall basket totals, we do not observe non-food items purchased during the same shopping trips.[1]

As Numerator relies on an algorithm to encode individual items in an uploaded receipt, there are additional limitations to the information that we can obtain at an item level. While we can identify certain details such as item ID, brand name, parent brand name, and category name, we are unable to observe other specifics like package size or flavor, which traditional UPC codes in scanner data would typically provide. Additionally, an item ID assigned to BM receipts can only be linked to an item ID in Instacart receipts in 2.5% of cases. Because of these limitations, we are unable to calculate item-level price differences between Instacart and BM channels directly. Therefore, we conduct most of our analysis at the category level. Nevertheless, we have carried out robustness tests which indicate that pricing differences between the channels are unlikely to be the main explanation for our findings.

In addition to the shopping data, we also have some demographic information about the

[1] In this paper, we use the terms "grocery" and "food" interchangeably.

panelists. This information includes the income bracket of the household, age of the panelist, ethnic/racial background, education level, and household zip code. The descriptive statistics for these demographics are presented in Appendix Table A1.

Instacart is the dominant grocery delivery service in the online space, with a market share of approximately 50% (Damiani 2020). Unlike retailer-specific online delivery services, such as Walmart.com and Kroger.com, Instacart is a third-party service that partners with over 700 different retail chains. The grocery retail market is highly localized, with prominent chains like Stop & Shop in the Northeast, Publix in the South, and Safeway in the West. Focusing on Instacart purchases enables us to capture households' typical shopping behavior for their regular grocery needs because most popular local retail chains do not typically offer regional-level alternative delivery services. Our data sample covers a comprehensive national selection of significant regional and local grocery retailers, allowing for greater generalizability of our findings regarding omnichannel basket composition and variety.

Figure 1: Instacart Sales Trend

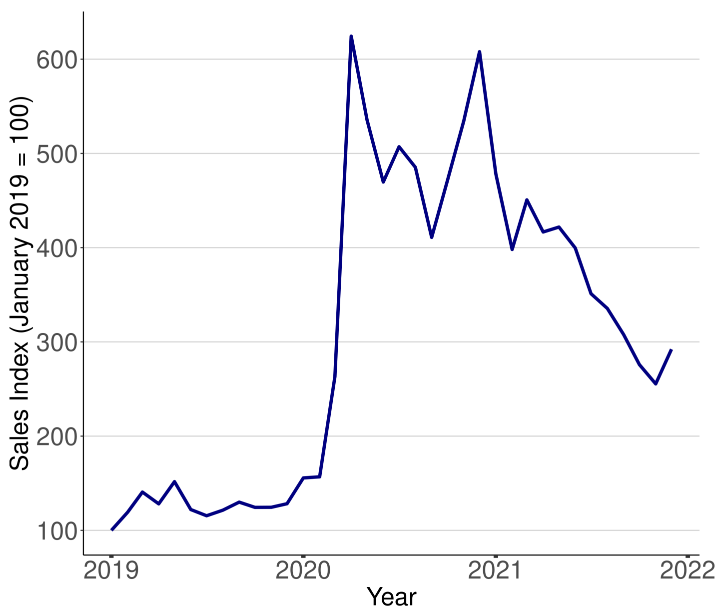


Figure 1 presents the Instacart sales index data from Numerator, where January 2019 sales are used as the normalization benchmark at 100. This sales trend displays the distinctive e-commerce "Covid bump" pattern.[2] Post-bump sales are notably higher than pre-

[2]We refer readers to Oblander and McCarthy (2023) for a more comprehensive investigation of the impact of Covid on consumer behavior across various digital/mobile platforms, including Instacart.

pandemic levels, indicating the continued increased utilization of grocery delivery services among US households (WSJ 2023).

Our dataset includes shopping trip information from 86,684 households between 2019 and 2021. However, not all panelists are active in all three years (see Appendix A for details). We are only able to observe Instacart sales for panelists who gave permission for Numerator to track their Instacart purchases, resulting in a sample of $N = 4,388$ households with at least one Instacart purchase. These households had a total of $N = 1,968,392$ distinct online and offline shopping trips during 2019-2021 and form the primary sample for our study. While the permission sample limitation is a drawback of our data, we find that the sales patterns of this sample closely mirror overall Instacart sales patterns from other proprietary data sources, such as Earnest Research (see Appendix Figure A2). Additionally, our primary sample's demographic composition is similar to that of the overall managed panel sample, which is designed to be representative of US consumers (see Appendix A.3).

Table 1: Roadmap of Empirical Analysis

Section	**Scope**
3 All online vs. all offline trips:	
	Describe the general patterns of shopping behavior for online vs. offline groceries
4 Characteristic Instacart vs. Characteristic Offline trips:	
4.1 Identifying Characteristic Trips	Introduce an algorithm that detects a likely substitution of an offline trip with an Instacart trip at a household level
4.2 Basket Variety	Show that the variety of Instacart baskets is systematically lower than that of offline baskets
4.3 Basket Similarities	Show that Instacart baskets within the same household exhibit higher similarity than offline baskets
4.4 Basket Composition	Show systematic differences in basket composition, such as lower produce and impulse purchases on Instacart compared to offline trips
4.5 Compensation and Adjustment	Demonstrate that households do not compensate for the differences in their online characteristic trips by altering their behavior in other shopping trips

2.2 Roadmap of Empirical Analysis

Broadly, our empirical analysis proceeds in two main steps, as summarized in Table 1. In section 3, we start by examining the overall grocery shopping patterns across both online and offline channels. The majority of the empirical investigation in this paper, covered in section 4, is dedicated to understanding shopping patterns when households are likely substituting their regular BM trip with an online Instacart trip (as opposed to using online channels as supplementary shopping). We first introduce the method to identify such trips, and then proceed to compare the variety, similarity, and composition of online versus offline characteristic baskets.

3 Comparison of All Online and Offline Purchases

Table 2 presents a summary of the shopping trip data for 4,388 households across BM, Instacart, and other online channels. The BM channel accounts for 87.2% of all expenditure-weighted trips, while Instacart and other online channels constitute 3.9% and 8.9%, respectively. The other online channels include Amazon.com (47.7% of all other online trips), Walmart.com (22.4%), Target.com (12.7%), Kroger.com (5.2%), and a long tail of other retailers such as Sheetz.com (0.1%) and Wine.com (0.1%). On average, households spend less on food in the BM channel ($29.89) compared to Instacart ($50.46) and other online channels ($34.5), and make significantly more trips using the BM channel than the online channels. The BM channel trips include visits to various types and sizes of retailers, ranging from mass merchandisers such as Sam's Club to traditional grocery stores such as Kroger, as well as small convenience stores and gas stations. Overall, our sample covers 8,117 BM retailers with 98,209 unique store locations across 3,350 zip codes. 134 retailers with 6,760 unique store locations have serviced Instacart purchases.

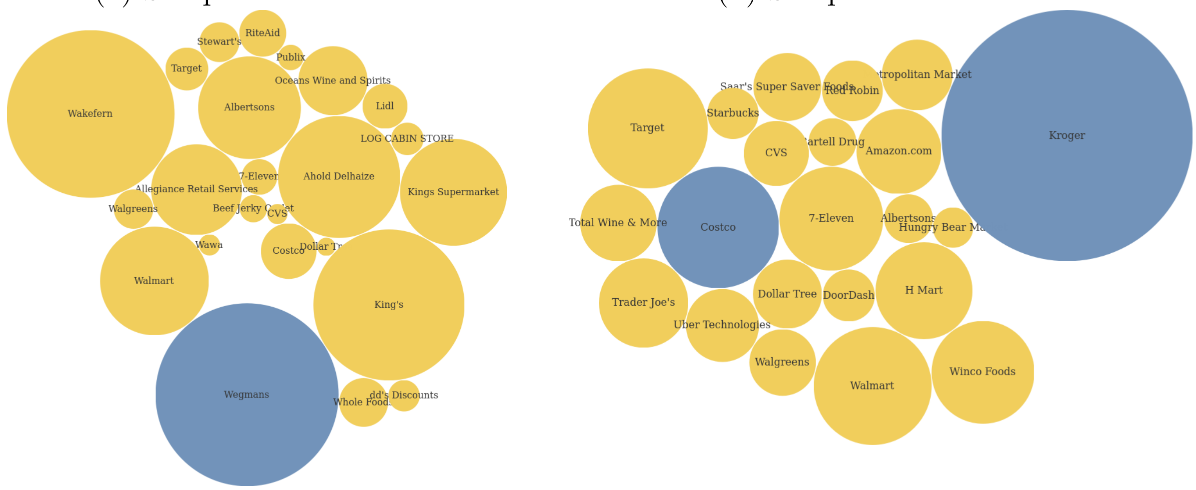
Table 2: Summary Statistics of All Shopping Trips Across both Channels

	Brick & Mortar		Instacart		Other Online	
	Mean	SD	Mean	SD	Mean	SD
Grocery amount per trip (in $)	29.89	39.73	50.46	40.99	34.53	40.24
Number of unique grocery items bought per trip	6.99	8.57	11.23	8.70	6.75	10.16
Number of grocery categories bought per trip	5.06	5.41	8.17	5.52	4.80	6.36
Number of trips made by a household	404.89	226.07	10.16	20.55	35.94	51.04
Average Expenditure Share by Channel	87.23%		3.87%		8.90%	
% of trips to Food, Mass, Club retailers with BM presence	76.07%		97.25%		44.19%	
N Retailer Chains	8,117		134		118	
N Unique Store Locations	98,209		6,760		11,349	

Notes: Food and Mass retailers are typical retailers such as Wegmans, Walmart, and Kroger where consumers usually shop for groceries. Club retailers include retailers such as Costco, Sam's club. The % trips to Food, Mass, Club retailers does not include trips to Dollar stores, Drug stores, Gas stations.

Figure 2: Grocery Spend by Retailer for Two Sample Households

(a) Sample Household 1 (b) Sample Household 2



Notes: This figure shows the grocery spending at different retailers for two sample households. The size of the bubble is proportional to the total amount spent at a retailer. The blue (darker shaded) bubble indicates that the household used both the BM and Instacart channels at that particular retailer, whereas the yellow (lighter shaded) bubble indicates that the household made only BM trips.

Next, we examine the descriptive statistics associated with household shopping patterns across different retailers. Figure 2 illustrates the grocery spending patterns of two sample households in our sample, where the size of each bubble represents the amount spent on food at a specific retailer. The blue bubbles indicate retailers where the household made purchases using both BM and Instacart channels. As we note above, we observe that in many cases, the household's top primary retailer is utilized via *both* BM and Instacart channels. Table 3 indicates that during our sample period, an average household in our sample frequented approximately 25 distinct BM retailers, utilized Instacart with 1.7 retailers, and shopped online with three other online retailers. We designate as primary retailers the top two retailers for each household that command the highest grocery expenditure shares in each channel. We find that households, on average, spend 45.38% of their grocery expenditures at their primary BM retailer and 65.36% of their expenditures at their top two primary BM retailers. When it comes to Instacart purchases, household spending is even more concentrated among the primary retailers, with an average of about 92% of Instacart purchases coming from the top two primary retailers.

Next, we conduct a comprehensive comparison of the variety of grocery items that are purchased online versus offline within a specific time frame. Here, we estimate the following

Table 3: Summary Statistics of Shopping Trips to the Primary Retailers by Channel Type

	Brick & Mortar		Instacart		Other Online	
	Mean	SD	Mean	SD	Mean	SD
Number of retailers per household	25.26	11.02	1.70	1.32	3.08	2.03
Basket size (in $) at:						
#1 retailer	41.77	47.54	52.42	42.20	39.66	43.62
#2 retailer	34.27	41.17	46.80	37.85	26.45	31.23
Number of trips to:						
#1 retailer	131.48	102.27	7.55	14.69	24.23	37.09
#2 retailer	70.57	61.09	4.64	7.25	10.47	22.43
Basket Expenditure Share of:						
#1 retailer	45.38%		77.14%		77.40%	
#2 retailer	19.98%		14.66%		22.31%	

specification:

$$y_{it} = \beta \times \text{Online}_{it} + \gamma_{iq(t)} + \varepsilon_{it} \tag{1}$$

The variable of interest, y_{it}, represents the cumulative variety outcome for household i in time period t. To measure variety, we use two primary variables: (i) log(Unique Number of Categories), which is a metric also used in previous research (Haws et al. 2017) and (ii) log(Unique Number of Items). We calculate the number of unique categories and items purchased by a household within a given month, quarter, or year. On the right-hand side, the primary variable of interest is a dummy indicating whether the cumulative variety measure is for an online channel (or Instacart channel), represented by the indicator variable Online_{it} or (Instacart_{it}). We also include $Household \times TimePeriod$ or $Household \times TimePeriod \times Retailer$ fixed effects. Therefore, the coefficient β measures the average cumulative variety differences purchased online vs. offline across all households and all time periods.

Panel (1) of Table 4 presents the variety differences across the BM and online baskets using $Household \times TimePeriod$ fixed effects. In Panel (2) of the same table, we also report a specification that includes a triple interaction $HH \times Retailer \times Time\ period$ fixed effects to compare the broader variety measures of online vs. BM shopping patterns *within* any given retailer, conditional on a retailer having presence in both online and BM channels. Note

Table 4: Variety Effects Including all Online, all BM, and all Instacart Trips

	(A) Number of Categories			(B) Number of Items		
	Monthly	Quarterly	Yearly	Monthly	Quarterly	Yearly
			(1) *All Online and all BM trips*			
Online	-1.580***	-1.918***	-1.814***	-1.960***	-2.635***	-3.050***
	(0.018)	(0.016)	(0.015)	(0.023)	(0.021)	(0.022)
Num.Obs.	189,796	72,081	20,716	189,796	72,081	20,716
R^2	0.506	0.632	0.656	0.518	0.666	0.739
			(2) *All Online and all BM trips with Retailer fixed effects*			
Online	-0.203***	-0.561***	-1.046***	-0.207***	-0.664***	-1.392***
	(0.015)	(0.013)	0.012	(0.017)	(0.016)	(0.016)
Num.Obs.	791,010	449,912	208,724	791,010	449,912	208,724
R^2	0.022	0.145	0.389	0.017	0.136	0.391
			(3) *All Instacart and all BM trips*			
Instacart	-0.914***	-1.516***	-2.027***	-1.195***	-2.178***	-3.383***
	(0.020)	(0.017)	(0.016)	(0.025)	(0.024)	(0.022)
Num.Obs.	150,041	54,712	16,917	150,041	54,712	16,917
R^2	0.380	0.660	0.777	0.415	0.696	0.837
			(4) *All Instacart and all BM trips with Retailer fixed effects*			
Instacart	-0.239***	-0.562***	-1.070***	-0.299***	-0.728***	-1.505***
	(0.016)	(0.015)	(0.014)	(0.020)	(0.020)	(0.020)
Num.Obs.	730,389	414,890	192,417	730,389	414,890	192,417
R^2	0.044	0.184	0.440	0.049	0.195	0.459

Notes: This table reports the cumulative basket variety results for all BM and all Online trips (panels (1) and (2)) as well as all BM and all Instacart trips (panels (3) and (4)) —not just the characteristic ones in a given time period. The outcome variable is *log(Number of Categories)* for column A and *log(Number of Items)* for column B. For panels (1) and (3), outcome variables are computed at the time period (Month/Quarter/Year) level, and for panels (2) and (4), outcome variables are computed at the Time period (Month/Quarter/Year)×Retailer level. Models in panels (1) and (3) use *HH×Month/Quarter/Year* fixed effects. Models in panels (2) and (4) use *HH×Month/Quarter/Year×Retailer* fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.

that this specification excludes purchases from Amazon.com (since there is no Amazon BM counterpart in our data), which account for the majority of online purchases. In Panels (3) and (4), we focus specifically on Instacart purchases and analyze the variety differences between all BM trips and all Instacart trips.

In summary, our results show that online purchases exhibit a significantly lower variety than those made in BM stores. For instance, when we analyze all shopping trips, the cumulative variety of items purchased within a month is 158% (91%) lower online (on Instacart) than offline. Focusing on retailers present in both online and offline domains, the variety decreases by 20% (24%) online (on Instacart).[3]

Finally, we also investigate whether online trips are more likely to involve bulk purchases or stockpiling trips. Our analysis reveals that all online (Instacart) trips are 4.6% (5%) more likely to include bulk purchases than BM trips to the same retailer. The detailed results of this analysis are provided in Appendix Table D6.

4 Comparison of Online and Offline Characteristic Trips

4.1 Identifying Characteristic Trips

As noted above, we observe close to two million shopping trips for our sample of 4,388 households. These shopping trips are comprised of grocery trips in which a given household makes purchases that reflect regular restocking of commonly used household grocery items; and non-regular (intermittent) shopping trips, such as fill-in trips or runs to convenience stores or gas stations. Our objective is to examine the shopping histories of each household to identify their routine restocking grocery trips, irrespective of the channel. In other words, we aim to identify those Instacart trips that serve as likely substitutes for a BM trip. We refer to such trips as *characteristic* trips.

There are various approaches to analyzing household-level shopping patterns, each with its advantages and disadvantages, depending on the empirical context. For example, Manchanda et al. (1999) propose a hierarchical Bayesian multi-category purchase incidence model, but such models become increasingly complex with numerous category combinations, as in our setting. Ruiz et al. (2020) adopt the sequential choice probabilistic SHOPPER model, while Platzer and Reutterer (2016) and Reutterer et al. (2021) investigate household-specific inter-purchase timing. Jindal et al. (2020) use clustering of shopping trips based on total dollars spent, as well as the number of items and categories purchased, to identify comparable trips. In contrast, our approach employs a combination of representation

[3] We note that, when limiting the analysis to multichannel retailers, the assortment differences between the two channels are usually minimal. This likely stems from these retailers maintaining similar product ranges across their platforms to ensure a uniform customer experience and explains the magnitude differences between the results reported panels (1) vs. (3) and panels (2) vs. (4) in Table 4.

learning using embeddings and clustering techniques from machine learning to analyze the shopping behavior of each household and determine what constitutes a regular restocking shopping trip for that household. Our embeddings-based approach is more flexible, easily scalable, and allows us not only to identify relevant similar online and offline grocery trips but also generates vector representations for each observed trip, enabling us to evaluate trip similarities.[4]

We refer the reader to Appendix B for a detailed explanation of this methodology. Specifically, Appendix B.1 summarizes the technical details for our two-step approach, B.2 provides examples to build intuition, B.3 describes the different classification approaches we consider to establish the robustness and generalizability, B.4 presents a validation exercise with human coders, and B.5 provides technical details about identifying the optimal number of clusters for classification.

Table 5: Summary Statistics by Trip Type

	Characteristic trip		Non-characteristic trip	
	Mean	SD	Mean	SD
Grocery amount per trip (in $)	72.64	52.69	17.64	22.54
Number of items bought per trip	23.93	17.05	4.92	6.05
Number of categories bought per trip	12.50	6.39	2.81	2.20
Average spacing between trips (in days)	10.56	5.39	3.45	1.91
Number of trips	467,543		1,500,849	
Percentage of trips to Food, Mass, Club retailers	90.15		65.12	

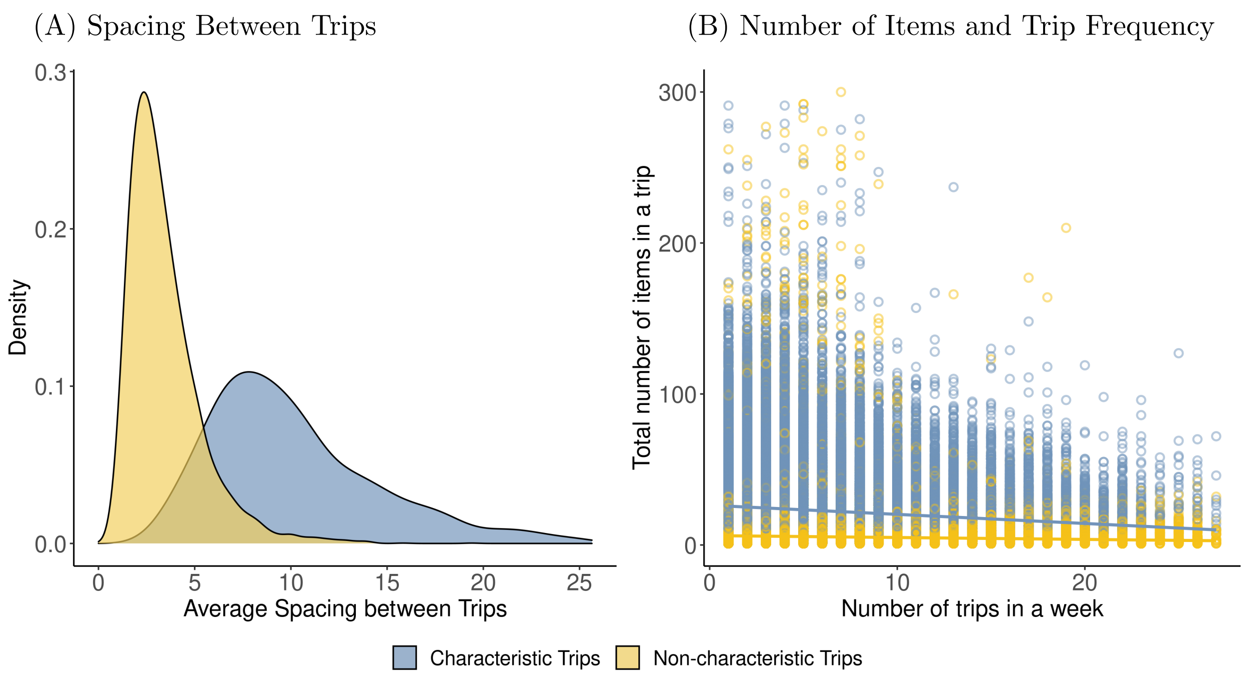
Notes: Total number of trips classified N=1,968,392

After classifying all trips into characteristic and non-characteristic trips, we present summary statistics for each trip type to further demonstrate the distinction between them in terms of shopping objectives and contexts. Table 5 displays key summary statistics for shopping basket outcomes of interest for both trip types, while Figure 3 visually highlights the

[4]Based on our informal conversations with store managers, the application of embeddings to summarize high-dimensional information from shopping baskets is also starting to be used by in-house data analysts in some grocery retail chains (see, for example, Walmart Labs application in Mantha et al. (2020)). Indeed, recent computer and data science literature has shown several applications of embeddings-based approaches for studying consumer behavior (see, for example, Entezari et al. (2021) and Vijjali et al. (2022). In general, the use of embeddings remains somewhat nascent in marketing literature. The exceptions include using embeddings to characterize market structures (Gabel et al. 2019), predict consumer responses to marketing actions (Gabel and Timoshenko 2021), and study product-level competition (Chen et al. 2022, 2020).

differences between them. Specifically, Panel (A) shows density plots of the spacing (in days) between two successive trips made by a household. Characteristic trips have significantly longer spacing between successive trips compared to non-characteristic trips (206% higher; $t = 82.3$, $p < 0.000$). This suggests that non-characteristic trips may represent unplanned runs to the shops with a higher sense of urgency. This assertion is further supported by the data patterns depicted in Panel (B), where we see that the average number of items in a basket is significantly lower for non-characteristic trips and that the basket size of non-characteristic trips does not vary with the total frequency of household trips. In contrast, the size of the basket for characteristic trips varies significantly with the frequency of trips: households that shop less frequently have larger characteristic baskets, and the number of items in the characteristic baskets decreases as households increase the shopping frequency. This might suggest that non-characteristic trips satisfy some urgent needs of the household, while characteristic trips tend to satisfy the less urgent grocery needs that can be achieved with varying frequency.

Figure 3: Characteristic vs. Non-Characteristic Trips



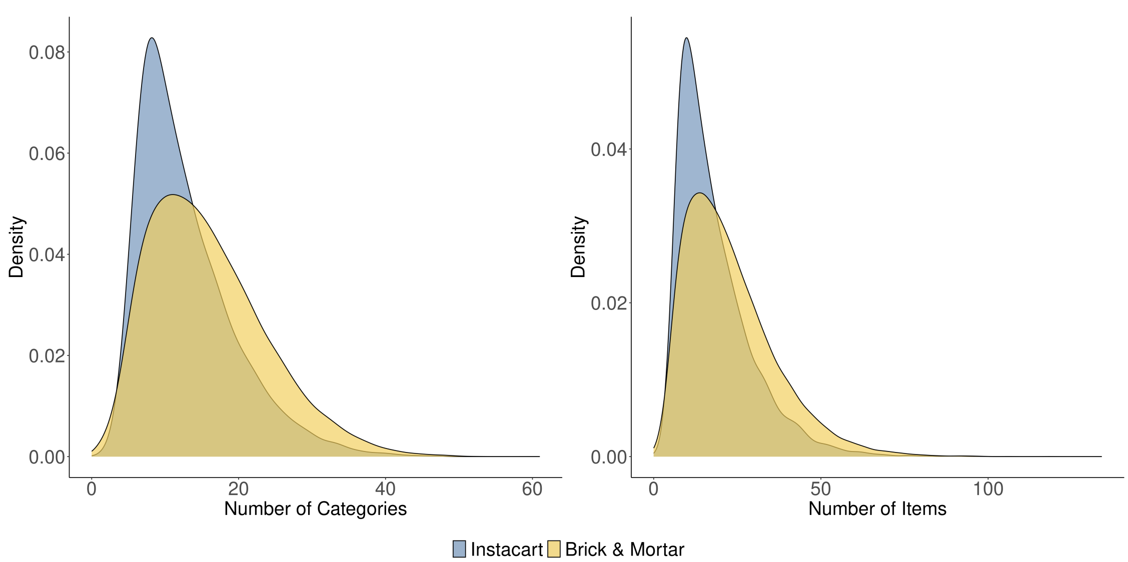
Notes: Panel A shows the density plot of the average spacing in days between two consecutive characteristic trips or two consecutive Non-characteristic trips. Panel B shows the scatter plot and regression lines describing the relationship between the total number of items purchased on a trip and the total number of trips in a week by trip type.

In Appendix B.4, we report the results of the validation exercise with human coders that confirm the effectiveness of our embeddings classification approach in identifying trips with different shopping objectives. Furthermore, in Appendix D.8, we conduct an analysis to evaluate the degree of substitutability between characteristic Instacart and BM trips and find that they are almost perfect substitutes. This analysis further validates our characteristic trip construct as our goal was to identify trips that could substitute for one another. Taken together, these findings provide compelling evidence that our classification methodology effectively differentiated between characteristic and non-characteristic trips, and successfully identified Instacart and BM characteristic trips that serve as substitutes for one another.

4.2 Characteristic Basket Variety

Our first outcome of interest is the variety of characteristic baskets. We use the same two measures to proxy for variety as in section 3: the number of distinct categories and the number of items in a given trip. Figure 4 shows the density plots of the number of unique categories and items for characteristic trips across the two channels. On average, Instacart characteristic baskets have fewer distinct categories (items), and the dispersion of the number of categories (items) is significantly lower than for BM trips.

Figure 4: Density plot of Unique Number of Categories and Items in Characteristic Trips



Notes: The figure shows the distribution of the number of unique categories and the number of unique items for BM and Instacart characteristic trips.

Motivated by these apparent descriptive differences in basket varieties, we investigate these patterns in a systematic manner using the following econometric framework:

$$y_{irt} = \beta \times \text{Instacart}_{irt} + \gamma_{irq(t)} + \theta_{im(t)} + \varepsilon_{irt} \tag{2}$$

Here, y_{irt}, represents a basket variety for household i at retailer r on shopping occasion t. We consider two variety measures: (i) the number of distinct categories purchased and (ii) the number of distinct items purchased. On the right-hand side, the main variable of interest is whether any given shopping trip was made using Instacart, which is represented by the indicator $Instacart_{irt}$. As such, β measures the systematic differences of Instacart characteristic baskets relative to BM characteristic baskets in the outcome variable of interest. For our baseline specification, we use two sets of highly granular fixed effects: (i) $\gamma_{irq(t)}$ is the triple interaction for *Household*×*Retailer*×*Quarter* fixed effects, which controls for any household-, retailer-, and quarter- level unobserved differences while (ii) $\theta_{im(t)}$ represents the *Household*×*Month* fixed effects, which account for time-varying household consumption changes across different channels. In other words, these fixed effects ultimately allow us to compare Instacart baskets to BM baskets *within* a given household, retailer, and quarter combination while simultaneously controlling for household-specific seasonality effects.

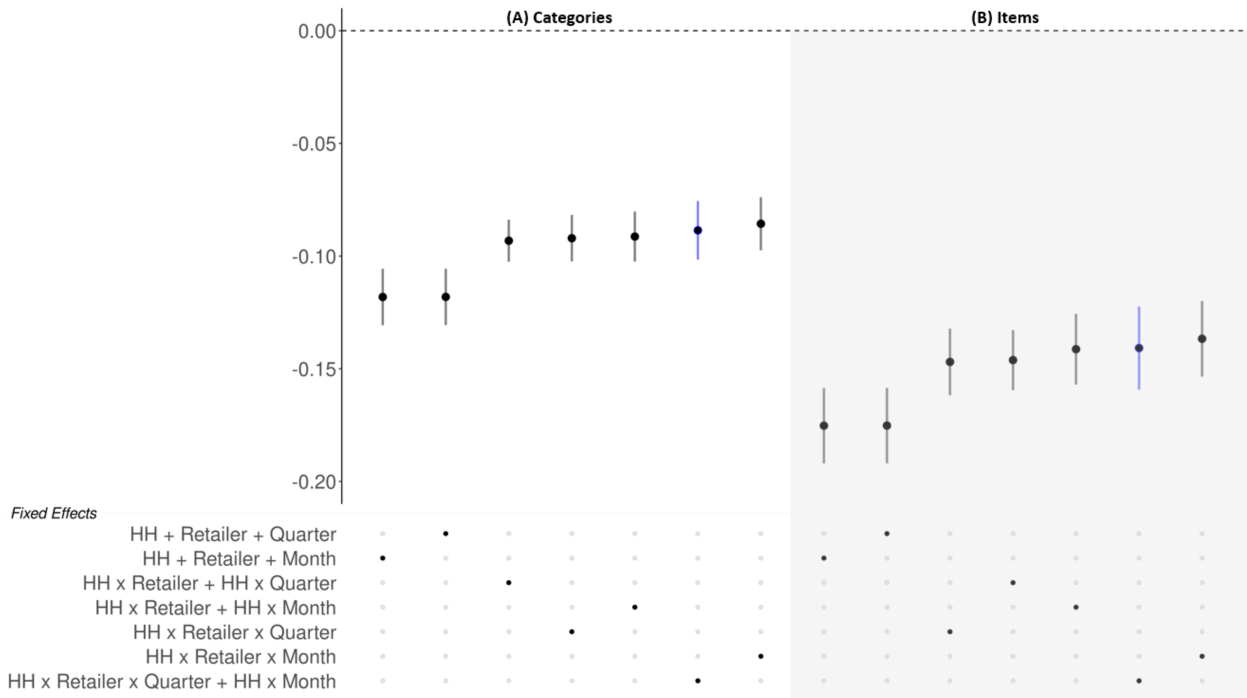
In Figure 5, we present the specification chart for the estimates of β using *log(Number of Categories)* and *log(Number of Items)* as outcome variables. We consider seven different specifications, each with varying sets of fixed effects, resulting in a total of 14 estimates. This approach allows us to examine the sensitivity of our results to different model specifications.[5] The baseline specifications (highlighted in blue and specified in Equation 2) include *Household*×*Retailer*×*Quarter*, and *Household*×*Month* fixed effects, with standard errors clustered at *Household*×*Retailer* level. The most restrictive specification includes the triple interaction fixed effects at a month level—*Household*×*Retailer*×*Month*.[6]

Overall, we find that Instacart characteristic baskets have a lower variety compared to

[5] For robustness, Appendix Table D3 also reports the estimates using count data (number of categories) as a dependent variable with Poisson regression. The point estimates and resulting interpretations are very similar.

[6] Appendix D.2 presents the results of our sensitivity analysis, where we estimate the same specifications using three different data samples from the years 2019, 2020, and 2021, respectively. This exercise helps demonstrate that our findings are robust across time. Appendix Figure D1 shows the specification plot of the estimates and Appendix Figure D2 reports estimates across all five classification approaches discussed in Appendix B.3.

Figure 5: Specification Chart for Variety Effects



Notes: This figure presents the specification chart for 7 different specifications each for categories (Panel A) and items (Panel B) with different sets of fixed effects. The outcome variables are *log(Number of Categories)* and *log(Number of Items)*. The baseline specifications as specified in Equation 2 are highlighted in blue. Error bars represent 95% CI using clustered standard errors. Please see Table D1 in the Appendix for regression results in the tabular format.

BM baskets, and this finding is robust across all 14 specifications. Specifically, using our preferred specification in Equation 2, we find that the variety of Instacart baskets is around 9.6% (14.1%) lower than the variety of BM baskets when considering distinct categories (items). Furthermore, we find that the lower variety estimates range between 5.8 and 12.8% (13.7 and 17.5%) for distinct categories (items) depending on the classification and specification used. Of note, the embeddings-based classification approaches provide more conservative and precise estimates than classification based on department totals, as shown in Appendix Figure D1. Additionally, the variety differences between Instacart and BM trips are significantly larger when evaluating all trips (see Table 4) rather than just characteristic trips, indicating that our main specification results focused on characteristic trips provide a conservative lower-bound estimate of the differences in basket variety.

Next, we investigate whether the diversity of Instacart baskets varies with a household's experience with Instacart. In particular, we consider a specification that adds an interaction term with the number of Instacart trips (as a proxy for experience with Instacart) to

Equation 2. We find that as households utilize Instacart more frequently, the previously pronounced pattern of lower variety on Instacart becomes somewhat less prominent. One possible explanation is that households are more willing to consider a wider range of products as they become more experienced and familiar with the Instacart platform; however, even the most experienced Instacart users appear to exhibit significantly lower basket variety compared to their BM counterparts (see Appendix Table D4 for more details).

Finally, in Appendix D.3 we examine the heterogeneity in variety differences across observable household characteristics such as income, education, and race. Our findings indicate that the online vs. offline basket variety differences remain largely consistent across various household characteristics.

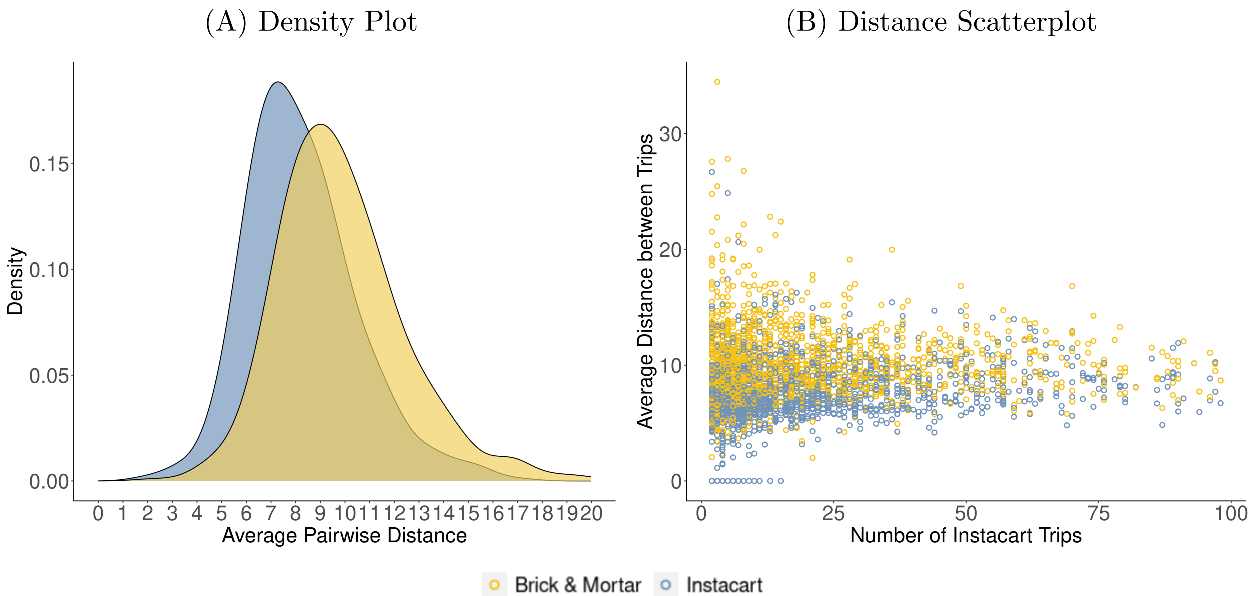
4.3 Characteristic Basket Similarity

Our next set of results pertains to comparing household-level basket similarity between Instacart versus BM shopping trips. To obtain *category-level* similarity measures, we use pairwise distances between different trips. Because all trips are represented in 100-dimensional space, we can calculate the Euclidean distance between any two characteristic trips. Using these distances, we calculate a similarity score for Instacart characteristic trips for each household by taking the average of the within-household pairwise distances between all combinations of such trips. Analogously, we compute a similarity score for characteristic trips made to brick-and-mortar (BM) stores and then compare the two scores. An important nuance of this approach is that there are significantly more BM characteristic trips in the data compared to Instacart characteristic trips. Therefore, differences in similarity scores between the two channels may be an artifact of the disparity in the number of trips. To address this issue, we randomly sample the same number of BM characteristic trips whenever the number of Instacart characteristic trips is lower than the number of BM characteristic trips.[7] Finally, we evaluate mean differences in similarity scores among Instacart and BM baskets.

[7] We also use a method similar to bootstrapping to assess the robustness of our approach. We repeat the random draw process of N BM trips 1000 times. We then compute the average of household-specific pairwise distances from 1000 iterations. The result from Welch's t-test for difference in means of Euclidean Distances is quantitatively similar to the result without the bootstrapping procedure.

Panel A in Figure 6 provides a visualization of the density plots of the average of within-household pairwise Euclidean distances of BM and Instacart characteristic trips for the households which have more than two Instacart characteristic trips. This graph shows that the distribution of Instacart pairwise distances is to the left of the distribution of BM pairwise distances, indicating that within each household, the Instacart trips are significantly more similar to each other than the BM trips. Specifically, we find that the category-level within-household pairwise distances are 27.1% larger for BM shopping trips as compared with Instacart shopping trips (see Table 6).

Figure 6: Similarity Metrics for Instacart and Brick & Mortar Trips



Notes: Panel A shows the distribution of household-specific pairwise Euclidean Distances for BM and Instacart characteristic trips. The sample includes households with more than 2 Instacart characteristic trips. Panel B shows the scatter plot of the average distance between the different types of trips as a function of experience with Instacart (excluding outliers with $N > 100$ of Instacart trips).

Next, we investigate the relationship between a household's experience with Instacart and the similarity of their Instacart baskets. Panel B of Figure 6 displays pairwise Euclidean distances for Instacart and BM trips for each household (Y axis) as a function of the number of Instacart trips a household has made (X axis). We observe that, on average, the pairwise Euclidean distances of Instacart trips are lower than those of BM trips, and this pattern is mostly consistent across the range of experience with Instacart, although the gap appears to narrow. To formally test this, we regress the relative gap in distances *((BM Distance - Instacart Distance)/BM Distance)* on the number of Instacart trips. The results indicate that while the gap slightly narrows with experience with Instacart, Instacart baskets remain significantly more similar to each other than BM baskets, even among the most experienced Instacart users. See Appendix Table D5 for the detailed results.

As an alternative method of assessing the similarity of baskets, we utilize the Jaccard Index to measure the extent of overlap between *items* purchased in any two successive trips to the same retailer. Specifically, we compute the Jaccard Index between any two consecutive characteristic trips at the $retailer \times household \times channel$ level. We then calculate the average Jaccard Index separately for BM and Instacart characteristic trips and perform a t-test. The results are reported in Table 6 and indicate that within-retailer similarity is, on average, twice as high for Instacart shopping trips than for BM shopping trips.

Table 6: Similarity Metrics by Characteristic Trip Type

Distance Measure	Instacart	BM	p-value	Mean difference	95% CI Lower	95% CI Upper
Euclidean Distance (Category-level)	7.531	9.576	<0.0000	2.045	2.035	2.054
Jaccard Similarity (Retailer-Item-level)	0.143	0.070	<0.0000	0.073	0.066	0.079

Notes: Euclidean distance refers to the distance between two different baskets represented in the 100-dimension space using Category-level embeddings. The Jaccard Similarity (Item-level) refers to the similarity measure computed using the unique Item IDs.

Overall, our findings suggest that although Instacart characteristic trips have 9.6% (14.1%) lower variety in categories (items) compared to BM characteristic trips, they exhibit a 27.1% higher degree of similarity in terms of categories and over 100% higher degree of similarity in terms of item overlap across successive trips to the same retailer. These differences are somewhat mitigated by the household's experience with Instacart but persist even for the most experienced users.

One plausible explanation for the differences in basket similarity could be potentially attributed to Instacart's 'Buy It Again' feature, which enables users to add previously purchased items to their current shopping cart. Intuitively, this feature might reduce a household's incentive to explore the full online assortment during each shopping occasion, thereby contributing to the increased similarity of Instacart baskets. However, it's important to note that our data and setting do not permit us to establish a definitive causal relationship between successive online basket similarities and this feature. Furthermore, identifying all mechanisms that may contribute to the observed similarity differences is beyond the scope of this paper. Nevertheless, we offer an informal discussion on other potentially relevant mechanisms in Appendix C.3.

4.4 Characteristic Basket Composition

Next, we examine the systematic differences in basket composition between BM and Instacart shopping trips by analyzing the variation in the number of items purchased within each category. As the Numerator categories are highly disaggregated, to increase the power of our tests, we aggregate similar categories and reduce the number of highly granular categories from 217 to 88. For example, we aggregate categories such as Bacon, Beef, Lamb, Pork, Poultry, Meat Snacks, and Sausage into a more general category called *Meat*. For each resulting aggregated category, we construct a dependent variable in the form of *log(1+Number of Items)* and loop over these 88 categories to estimate Equation 2.

In Table 7, we present the ten categories with the largest negative β estimates (in absolute magnitude) and the only four categories with significant positive β estimates from 88 regressions. Our results reveal that, spanning all categories, the most substantial difference—amounting to 13.6% fewer purchases in Instacart characteristic baskets—is observed in the fresh vegetables category. The next set of categories exhibits differences with noticeably lower magnitudes, and this set includes several items from the impulse purchase categories, such as candy, bakery desserts, and savory snacks such as chips. The observed differences in the fresh vegetable and impulse category purchases correspond to fewer purchases along both dimensions, namely the intensive and extensive margins (see Appendix D.10). Additionally, we observe differences in fresh meat purchases, lower by about 4.3%, although these differences are partially offset by higher frozen meat purchases, as reported in the *More* column of Table 7. The same column shows that all categories with systematically higher purchase incidences in Instacart baskets center around frozen foods. Contrary to the overall variety or similarity patterns in section 4.2 and section 4.3, we find no evidence suggesting that these patterns vary with a household's experience with Instacart in any of the categories where differences in basket composition have been identified.

While it is outside the scope of this paper to document the underlying mechanisms explaining the differences in shopping basket composition, our informal conversations with

Table 7: Categories that Exhibit the Largest Differences in Instacart vs. Brick & Mortar Characteristic Baskets

Less Category Name	Estimate	More Category Name	Estimate
Fresh Vegetables	-0.136*** (0.016)	Frozen Meat	0.025*** (0.006)
Candy	-0.071*** (0.007)	Frozen Vegetables	0.016* (0.006)
Cheese	-0.061*** (0.011)	Frozen Breakfast	0.014* (0.005)
Bakery Desserts	-0.059*** (0.008)	Frozen Fruit	0.008** (0.002)
Savory Snacks	-0.047*** (0.009)		
Fresh Meat	-0.043*** (0.012)		
Bread	-0.029*** (0.008)		
Sweet Snacks	-0.028*** (0.007)		
Deli	-0.027*** (0.007)		
Pasta	-0.026*** (0.005)		

Notes: This table compiles the categories with the largest negative β estimates (in absolute magnitude) from 88 regressions, each for a different category, estimated using Equation 2. For each regression, the dependent variable is *log(1+Number of Items)* from that respective category. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in the parentheses.

Instacart consumers and Instacart data scientists give us some clues as to why we might observe these differences. For instance, the decrease in purchases of fresh vegetables in Instacart baskets could be attributed to consumers wanting to have greater control over the selection of these items, as they may not trust the Instacart shopper to select the freshest options. Instacart is aware of this potential issue as they have specific policies to encourage its shoppers to "look for the freshest possible items and pay special attention to expiration dates, broken seals, and the quality of fresh produce" (Instacart 2021). The lower number of impulse category purchases, on the other hand, can potentially be explained by behavioral mechanisms related to self-control. For instance, it might be easier for shoppers to stick to their shopping lists and avoid impulse purchases when they are not exposed to the visual temptations of items in physical stores (Huyghe et al. 2017).

4.5 Compensation and Adjustment Beyond Characteristic Trips

So far we have demonstrated that there are systematic differences in basket variety, similarity, and composition between BM and Instacart characteristic trips. Next, we investigate the extent to which households adjust their grocery shopping, dining out, and/or food delivery behavior outside the focal characteristic trips. Specifically, we examine whether households compensate for the differences in their characteristic trips by altering their behavior in other shopping trips or dining occasions. For example, do households purchase additional candy during their convenience store runs to make up for not buying candy on Instacart? Similarly, do they modify their dining out or food delivery behavior to compensate for changes in their grocery shopping behavior?

We begin by examining whether there are systematic differences in grocery shopping patterns in the seven-day window surrounding focal BM and Instacart characteristic trips (i.e., three days before, three days after, and other shopping episodes on the same day as the focal trip). We investigate adjustments in both the extensive and intensive margins. The extensive margin is measured by determining whether households make more non-characteristic trips around Instacart characteristic trips than around BM characteristic trips. The intensive margin is measured by calculating the total spending in non-characteristic trips to see whether households spend more to compensate for fewer purchases around Instacart shopping trips than around BM shopping trips. The results, presented in the first column of Table 8, suggest that households do not have a higher frequency or spending of non-characteristic grocery trips.

In addition to observing all grocery receipts, we also observe restaurant and food delivery trips and the total spending on those trips. To explore whether households' shopping behavior towards restaurants/food delivery differs when they shop for groceries online, we analyze the data on the number and spending of these trips. On average, each household in our sample is observed to have 1.12 such trips per week. The results in the second and third columns of Table 8 indicate that there are no detectable differences in either the number or

Table 8: Extensive and Intensive Margin Differences around Instacart Trips

	Non-characteristic trips	Restaurant trips	Delivery trips
	(A) *Extensive Margin: By Number of Trips*		
Instacart	-0.017*	0.001	0.004
	(0.008)	(0.007)	(0.003)
Num.Obs.	434,322	434,322	434,322
R^2	0.678	0.741	0.681
	(B) *Intensive Margin: By Dollar Spend*		
Instacart	-0.001	-0.001	0.009
	(0.011)	(0.012)	(0.008)
Num.Obs.	434,322	434,322	434,322
R^2	0.661	0.727	0.670

Notes: The dependent variable for extensive margin regressions is *log(number of non-characteristic trips)*, and for intensive margin is *log(dollar spent in non-characteristic trips)*. All models use $HH \times Retailer \times Quarter$ and $HH \times Month$ fixed effects. *** p<0.001, ** p<0.01, * p<0.05; standard errors reported in parentheses.

spending across restaurant or delivery trips.

Finally, we examine potential category-specific adjustments related to Instacart characteristic trips. Our analysis shows no evidence that households make up for the items they didn't buy via Instacart by purchasing them through other channels. Specifically, our findings indicate that when an average household replaces a BM shopping trip with an Instacart shopping trip, they purchase 13% fewer vegetables and 5-6% fewer impulse purchases in that week.[8] Appendix D.9 details our category-specific analysis and reports the results.

Overall, our results uncover no apparent patterns consistent with major compensatory adjustments at the extensive or intensive margins around Instacart characteristic trips.

4.6 Robustness

In this section, we provide brief details about each of the robustness checks and other analyses we implement, which are detailed in respective Appendices. Although these robustness

[8]Since an average Instacart user employs Instacart for one out of every five regular grocery shopping trips (see Appendix C.1), a back-of-the-envelope estimate suggests that overall vegetable purchases by an average Instacart-adopting household are about 3% lower.

checks help strengthen our main findings and demonstrate that they are not artifacts of potential confounders, we emphasize that establishing robustness does not imply causality. We interpret our main findings as descriptive evidence of systematic differences in basket variety, similarity, and content across online and offline channels.[9]

Specifically, in Appendix C.1, we conduct multiple robustness checks to ensure that our documented patterns are not artifacts of pandemic-induced changes in consumption behavior. In Appendix C.2, we demonstrate that systematic price discrepancies between Instacart and BM channels are unlikely to explain our main findings. Lastly, in Appendix C.3, we provide an informal discussion outlining other potential explanations of our results.

5 Managerial and Policy Implications

Our research provides implications for retail management and sheds light on consumer shopping habits in the context of health and wellness. Furthermore, we highlight the potential of the online grocery channel to amplify the inertia in consumer behavior patterns, an aspect that warrants further exploration and carries significant implications for industry competition and brand management. We discuss each of these implications next.

First, we demonstrate that it is possible to predict which products are less (or more) likely to be purchased when households complete their routine grocery shopping online as opposed to in BM stores. With this information, both BM retailers and online grocery platforms can take steps to counteract the decline in sales in specific categories. For instance, if customers are purchasing fewer fresh produce items online due to concerns about the quality, the platform could provide training to shoppers to pick higher quality produce or select better replacements in the event of stock-outs.

Second, our research carries significant implications for the nutritional content of consumers' purchases as online baskets appear to contain fewer vegetables, but at the same time, fewer unhealthy items, such as junk food. These findings might help provide guidance to food assistance programs such as SNAP, which could consider encouraging beneficiaries

[9]Even in an ideal randomized experiment setting, establishing causal effects may not be possible, as randomizing households to forcibly use one channel over another would not be feasible.

to utilize offline channels to redeem their benefits in order to increase their consumption of fresh produce. Moreover, being aware of whether recipients redeem their SNAP benefits online or offline, can provide policymakers with more insights into the impact of these policies (Hinnosaar 2023). Retailers and online platforms can also play a role by promoting healthier food choices among online customers and enhancing access to nutritious ingredients for households residing in food deserts with limited availability of healthy food options. Furthermore, the emergence of CRISPR-edited produce, which has the capability to enhance the perishability and quality of fresh food items, may lead to increased consumer confidence in purchasing produce through online channels (Deng et al. 2023).

Third, our findings may have implications for how brands can cultivate customer loyalty in online shopping environments. We observe that online grocery baskets exhibit considerable similarity compared to offline baskets. Although further research is needed to explicitly explore the link between basket similarities and consumer inertia across channels, these findings hold important implications for brand management. Our results suggest that brands should prioritize establishing a presence in customers' online baskets early on, as overlooking this aspect may lead to the brand being disregarded in future purchases. In the online shopping environment, traditional in-store promotional tactics like product displays and samples are not applicable. Therefore, brands will need to re-evaluate their approaches to introducing new products. To establish a brand presence in online shopping environments, promotional activities and generating awareness through paid search and premium placements on digital shelves will become increasingly important. We are already witnessing brands investing in these strategic efforts, as evidenced by the significant growth in Instacart's revenue from their internal advertising bidding platform, which nearly doubled from 2019 to 2022 (Anderson 2022).

Finally, the growing reliance on online grocery shopping has the potential to significantly impact competition within the grocery sector. Previous studies in marketing (Guo and Wang 2023, Danaher et al. 2003) have shown that customers tend to display higher brand loyalty in the online channel. Consequently, this heightened brand loyalty may present challenges for new entrants seeking to establish a customer base. This aligns with prior research on barriers to entry in the retail market, indicating that consumer inertia can create difficulties for newcomers, encourage larger incumbents to raise their prices, and favor production within larger firms with higher markups, while simultaneously discouraging the emergence of smaller players (Bornstein 2020, Pozzi 2012).

6 Concluding Remarks

In this paper, we begin by outlining the general patterns of grocery shopping behavior in offline and online channels. Then, to gain insights into shopping patterns when households are likely substituting their regular offline trips with online Instacart trips, we identify characteristic shopping trips for each household and measure the systematic differences in basket variety, similarity, and composition across channels. We find that Instacart trips exhibit a lower basket variety compared to offline trips, and within households, Instacart baskets have higher similarity in food contents than offline baskets. Additionally, we examine systematic basket composition differences and find significant differences in the fresh vegetable and impulse purchase categories. Finally, we highlight that when households substitute some of their characteristic offline restocking trips with Instacart orders, they do not compensate for the reduced purchases in the fresh vegetable and impulse categories through alternative grocery shopping or restaurant trips.

As with every study, ours has limitations that might present opportunities for future research. While we observe household purchases, we do not observe actual consumption. A priori, the fewer purchases of the types of items that we document might have ambiguous effects on consumer welfare. On one hand, research shows that approximately 40% of food is wasted by US households, and this percentage increases to 50% when considering fresh produce (Goldenberg 2016). Therefore, lower purchases of specific categories may help reduce household food waste. On the other hand, our findings draw attention to fewer purchases of categories that are unequivocally healthy (vegetables) and unequivocally unhealthy (impulse purchases). Although evaluating consumer welfare is beyond the scope of this paper, future studies could investigate whether lower purchases contribute to reduced waste or actually alter consumption behavior. Examining the healthiness of foods purchased online versus offline could also offer new substantive insights into the possibility of "online food deserts" and ultimately contribute to the ongoing debate about food deserts in general (Caoui et al. 2022, Kolb 2021).

Another important dimension of consumer welfare that we do not address in this paper is the trade-off between differences in basket variety/composition and time savings associated with online grocery shopping (Chintagunta et al. 2012). Unfortunately, our data does not contain information needed for calculating the distances between household residences and BM stores, which could serve as a useful proxy for time savings. Nonetheless, Huang and Bronnenberg (2022) provide a promising framework for examining these trade-offs in detail, as they demonstrate potential consumer welfare gains with the e-commerce channel in the retail fashion industry.

Data limitations prevent us from formally modeling the differences in consumer consideration set formation, which is likely to vary across channels. Past research in this area has mainly focused on either offline or online channels (Bronnenberg et al. 2023, Chakravarti and Janiszewski 2003), but not both. This may be an important avenue for future research, as our results suggest that online channels may accelerate consumer inertia and progression through the purchase funnel. Additionally, product availability may influence consideration sets, and future research could incorporate inventory-related information or shocks into product demand forecasts (Ergin et al. 2022, Levine and Seiler 2022).

Finally, our study highlights the potential of flexible and easily scalable embeddings-based methods for revealing informative household shopping patterns. Future research can expand upon these machine learning techniques to improve demand estimation (Magnolfi et al. 2022) and forecasting (Doan et al. 2018).

Funding and Competing Interests. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.

References

Anderson G (2022) How will grocers react to Instacart's expanded digital ad network? *RetailWire* URL https://retailwire.com/discussion/how-will-grocers-react-to-instacarts-expanded-digital-ad-network/, (accessed January 20, 2023).

Aparicio D, Metzman Z, Rigobon R (2021) The pricing strategies of online grocery retailers. Technical report, National Bureau of Economic Research.

Bawa K, Shoemaker R (2004) The effects of free sample promotions on incremental brand sales. *Marketing Science* 23(3):345–363.

Bell DR, Gallino S, Moreno A (2018) Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science* 64(4):1629–1651.

Bornstein G (2020) Entry and profits in an aging economy: The role of consumer inertia. *working paper, The Wharton School, University of Pennsylvania* .

Bronnenberg BJ, Dubé JPH, Gentzkow M (2012) The evolution of brand preferences: Evidence from consumer migration. *American Economic Review* 102(6):2472–2508.

Bronnenberg BJ, Klein TJ, Xu Y (2023) Consumer time budgets and grocery shopping behavior. *Management Science* .

Brynjolfsson E, Hu Y, Simester D (2011) Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science* 57(8):1373–1386.

Campbell J (2019) Why does instacart charge more than the store? (or do they?). *The Grocery Store Guy* URL https://thegrocerystoreguy.com/why-does-instacart-charge-more-than-the-store/, (accessed November 16, 2021).

Caoui EH, Hollenbeck B, Osborne M (2022) The impact of dollar store expansion on local market structure and food access. *Available at SSRN 4163102* .

Chakravarti A, Janiszewski C (2003) The influence of macro-level motives on consideration set composition in novel purchase situations. *Journal of Consumer Research* 30(2):244–258.

Chen F, Liu X, Proserpio D, Troncoso I (2022) Product2vec: Understanding product-level competition using representation learning. *working paper, NYU Stern School of Business* .

Chen F, Liu X, Proserpio D, Troncoso I, Xiong F (2020) Studying product competition using representation learning. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1261–1268.

Chintagunta PK (1998) Inertia and variety seeking in a model of brand-purchase timing. *Marketing Science* 17(3):253–270.

Chintagunta PK, Chu J, Cebollada J (2012) Quantifying transaction costs in online/off-line grocery channel choice. *Marketing Science* 31(1):96–114.

Chiou L, Tucker C (2020) Social distancing, internet access and inequality. Technical report, National Bureau of Economic Research.

Choi J, Bell DR (2011) Preference minorities and the internet. *Journal of Marketing Research* 48(4):670–682.

Chu J, Chintagunta P, Cebollada J (2008) Research note—a comparison of within-household price sensitivity across online and offline channels. *Marketing Science* 27(2):283–299.

Cui TH, Ghose A, Halaburda H, Iyengar R, Pauwels K, Sriram S, Tucker C, Venkataraman S (2021) Informational challenges in omnichannel marketing: remedies and future research. *Journal of Marketing* 85(1):103–120.

Damiani J (2020) Instacart Surges Past Walmart In Online Grocery Market. *Forbes* URL www.forbes.com/sites/jessedamiani/2020/06/09/instacart-surges-past-walmart-in-online-grocery-market/, (accessed November 16, 2021).

Danaher PJ, Wilson IW, Davis RA (2003) A comparison of online and offline consumer brand loyalty. *Marketing Science* 22(4):461–476.

Datta H, Knox G, Bronnenberg BJ (2018) Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. *Marketing Science* 37(1):5–21.

de Vaan M, Mumtaz S, Nagaraj A, Srivastava SB (2021) Social learning in the covid-19 pandemic: Community establishments' closure decisions follow those of nearby chain establishments. *Management Science* 67(7):4446–4454.

Degeratu AM, Rangaswamy A, Wu J (2000) Consumer choice behavior in online and traditional supermarkets: The effects of brand name, price, and other search attributes. *International Journal of research in Marketing* 17(1):55–78.

Deng S, Adalja A, Liaukonyte J (2023) Consumer Acceptance of CRISPR-Edited Food Products and Implications for Online Grocery Shopping. *working paper, Cornell SC Johnson College of Business* .

Doan T, Veira N, Keng B (2018) Generating realistic sequences of customer-level transactions for retail datasets. *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, 820–827 (IEEE).

Donnelly R, Kanodia A, Morozov I (2023) Welfare effects of personalized rankings. *Marketing Science* .

Dubé JP, Hitsch GJ, Rossi PE (2010) State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics* 41(3):417–445.

Entezari N, Papalexakis EE, Wang H, Rao S, Prasad SK (2021) Tensor-based complementary product recommendation. *2021 IEEE International Conference on Big Data (Big Data)*, 409–415 (IEEE).

Erdem T, Sun B (2001) Testing for choice dynamics in panel data. *Journal of Business & Economic Statistics* 19(2):142–152.

Ergin E, Gümüş M, Yang N (2022) An empirical analysis of intra-firm product substitutability in fashion retailing. *Production and Operations Management* 31(2):607–621.

Ertekin N, Gumus M, Nikoofal M (2021) Online-exclusive or hybrid? channel merchandising strategies for ship-to-store implementation. *Management Science* forthcoming.

Fleder D, Hosanagar K (2009) Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science* 55(5):697–712.

Freimer M, Horsky D (2012) Periodic advertising pulsing in a competitive market. *Marketing Science* 31(4):637–648.

Gabel S, Guhl D, Klapper D (2019) P2v-map: Mapping market structures for large retail assortments. *Journal of Marketing Research* 56(4):557–580.

Gabel S, Timoshenko A (2021) Product choice with large assortments: A scalable deep-learning model. *Management Science* forthcoming.

Gao S, Rao J, Kang Y, Liang Y, Kruse J, Dopfer D, Sethi AK, Reyes JFM, Yandell BS, Patz JA (2020) Association of mobile phone location data indications of travel and stay-at-home mandates with covid-19 infection rates in the us. *JAMA network open* 3(9):e2020485–e2020485.

Ge Y, Zhao S, Zhou H, Pei C, Sun F, Ou W, Zhang Y (2020) Understanding echo chambers in e-commerce recommender systems. *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval*, 2261–2270.

Gijsenberg MJ (2017) Riding the waves: revealing the impact of intrayear category demand cycles on advertising and pricing effectiveness. *Journal of Marketing Research* 54(2):171–186.

Goldenberg S (2016) Half of all US food produce is thrown away, new research suggests. *The Guardian* URL https://www.theguardian.com/environment/2016/jul/13/us-food-waste-ugly-fruit-vegetables-perfect, (accessed July 19, 2022).

Gordon BR, Goldfarb A, Li Y (2013) Does price elasticity vary with economic growth? a cross-category analysis. *Journal of Marketing Research* 50(1):4–23.

Guo B, Wang D (2023) Will online shopping lead to more brand loyalty than offline shopping? the role of uncertainty avoidance. *Journal of Marketing Research* 0(0):00222437231153075, URL http://dx.doi.org/10.1177/00222437231153075.

Gupta S, Chintagunta PK, Wittink DR (1997) Household heterogeneity and state dependence in a model of purchase strings: Empirical results and managerial implications. *International Journal of Research in Marketing* 14(4):341–357.

Haddon H (2019) Amazon to Whole Foods Online Delivery Customers: We're Out of Celery, How's Kale? *Wall Street Journal* URL https://www.wsj.com/articles/amazon-to-whole-foods-online-delivery-customers-were-out-of-celery-hows-kale-11553425200, (accessed December 8,2021).

Harris-Lagoudakis K (2021) Online shopping and the healthfulness of grocery purchases. *American Journal of Agricultural Economics* forthcoming.

Haws KL, Liu PJ, Redden JP, Silver HJ (2017) Exploring the relationship between varieties of variety and weight loss: When more variety can help people lose weight. *Journal of Marketing Research* 54(4):619–635.

Hinnosaar M (2023) The persistence of healthy behaviors in food purchasing. *Marketing Science* 42(3):521–537.

Holtz D, Carterette B, Chandar P, Nazari Z, Cramer H, Aral S (2020) The engagement-diversity connection: Evidence from a field experiment on spotify. *Proceedings of the 21st ACM Conference on Economics and Computation*, 75–76.

Huang Y, Bronnenberg BJ (2022) Consumer transportation costs and the value of e-commerce: Evidence from the dutch apparel industry. *Marketing Science* .

Hussey E (2021) Browsing In The Aisles Has Been Replaced By Browsing Mobile Apps | Jācapps. URL https://jacapps.com/browsing-in-the-aisles-has-been-replaced-by-browsing-mobile-apps, (accessed May 19, 2021).

Huyghe E, Verstraeten J, Geuens M, Van Kerckhove A (2017) Clicks as a healthy alternative to bricks: How online grocery shopping reduces vice purchases. *Journal of Marketing Research* 54(1):61–74.

Instacart (2021) Instacart Help Center. *White paper* URL https://www.instacart.com/help/section/360007797972/360039569911, (accessed December 8, 2021).

IRI Worldwide (2020) The Changing Shape of the CPG Demand Curve: E-commerce. *White paper* URL https://www.iriworldwide.com/IRI/media/Library/COVID-19-Changing-Shape-of-the-Demand-Curve-Part-6-7-29-20.pdf, (accessed May 19, 2021).

Jindal P, Zhu T, Chintagunta P, Dhar S (2020) Marketing-mix response across retail formats: The role of shopping trip types. *Journal of Marketing* 84(2):114–132.

Kolb KH (2021) *Retail Inequality: Reframing the Food Desert Debate* (Univ of California Press).

Levine J, Seiler S (2022) Identifying state dependence in brand choice: Evidence from hurricanes. *Marketing Science* .

Li KT (2020) Statistical inference for average treatment effects estimated by synthetic control methods. *Journal of the American Statistical Association* 115(532):2068–2083.

Li X, Grahl J, Hinz O (2022) How do recommender systems lead to consumer purchases? a causal mediation analysis of a field experiment. *Information Systems Research* 33(2):620–637.

Luo L, Sun J (2016) New product design under channel acceptance: Brick-and-mortar, online-exclusive, or brick-and-click. *Production and Operations Management* 25(12):2014–2034.

Magnolfi L, McClure J, Sorensen A (2022) Triplet embeddings for demand estimation. *Available at SSRN* .

Manchanda P, Ansari A, Gupta S (1999) The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing Science* 18(2):95–114.

Mantha A, Arora Y, Gupta S, Kanumala P, Liu Z, Guo S, Achan K (2020) A large-scale deep architecture for personalized grocery basket recommendations. *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3807–3811 (IEEE).

McKinsey (2022) The next horizon for grocery e-commerce. https://www.mckinsey.com/industries/retail/our-insights/the-next-horizon-for-grocery-ecommerce-beyond-the-pandemic-bump#/, (Accessed on 05/31/2023).

Mercatus (2020) June 2020 Online Grocery Scorecard: Growth in sales & HH penetration continues. URL https://www.brickmeetsclick.com/june-2020-online-grocery.-scorecard--growth-in-sales---hh-penetration-continues, (accessed May 19, 2021).

Mercatus (2021) Total U.S. Online Grocery Sales for March 2021 Up 43% Versus Year Ago - Mercatus. URL https://www.mercatus.com/newsroom/announcements/total-u-s-online-grocery-sales-for-march-2021-up-43-versus-year-ago, (accessed May 19, 2021).

Meyer KA, Guilkey DK, Ng SW, Duffey KJ, Popkin BM, Kiefe CI, Steffen LM, Shikany JM, Gordon-Larsen P (2014) Sociodemographic differences in fast food price sensitivity. *JAMA internal medicine* 174(3):434–442.

Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J (2013) Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 3111–3119.

Milkman KL, Rogers T, Bazerman MH (2010) I'll have the ice cream soon and the vegetables later: A study of online grocery purchases and order lead time. *Marketing Letters* 21(1):17–35.

Mulhern FJ, Williams JD, Leone RP (1998) Variability of brand price elasticities across retail stores: Ethnic, income, and brand determinants. *Journal of Retailing* 74(3):427–446.

Nagaraj A, Reimers I (2021) Digitization and the demand for physical works: Evidence from the google books project. *Available at SSRN 3339524* .

Narang U, Shankar V (2019) Mobile app introduction and online and offline purchases and product returns. *Marketing Science* 38(5):756–772.

Neslin SA (2022) The omnichannel continuum: Integrating online and offline channels along the customer journey. *Journal of Retailing* 98(1):111–132.

Netzer O, Feldman R, Goldenberg J, Fresko M (2012) Mine your own business: Market-structure surveillance through text mining. *Marketing Science* 31(3):521–543.

Oblander S, McCarthy D (2023) Estimating the long-term impact of major events on consumption patterns: Evidence from covid-19. *Marketing Science* forthcoming.

Oestreicher-Singer G, Sundararajan A (2012) Recommendation networks and the long tail of electronic commerce. *MIS quarterly* 65–83.

Orhun AY, Palazzolo M (2019) Frugality is hard to afford. *Journal of Marketing Research* 56(1):1–17.

Platzer M, Reutterer T (2016) Ticking away the moments: Timing regularity helps to better predict customer activity. *Marketing Science* 35(5):779–799.

Pozzi A (2012) Shopping cost and brand exploration in online grocery. *American Economic Journal: Microeconomics* 4(3):96–120.

Rao S, Zhang L (2021) The algorithms that make instacart roll: How machine learning and other tech tools guide your groceries from store to doorstep. *IEEE Spectrum* 58(3):36–42.

Ratchford B, Soysal G, Zentner A, Gauri DK (2022) Online and offline retailing: What we know and directions for future research. *Journal of Retailing* .

Reutterer T, Platzer M, Schröder N (2021) Leveraging purchase regularity for predicting customer behavior the easy way. *International Journal of Research in Marketing* 38(1):194–215.

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics* 20:53–65.

Ruiz FJ, Athey S, Blei DM (2020) Shopper: A probabilistic model of consumer choice with substitutes and complements. *The Annals of Applied Statistics* 14(1):1–27.

Sangani K (2022) Markups across the income distribution: Measurement and implications. *Available at SSRN 4092068* .

Sorvino C (2021) Instacart Survived Covid Chaos — But Can It Keep Delivering After The Pandemic? *Forbes* URL https://www.forbes.com/sites/chloesorvino/2021/01/27/instacart-survived-covid-chaos---but-can-it-keep-delivering-after-the-pandemic/?sh=3473e423bfa1, (accessed November 16, 2021).

Tucker C, Zhang J (2011) How does popularity information affect choices? a field experiment. *Management Science* 57(5):828–842.

Van Heerde HJ, Gijsenberg MJ, Dekimpe MG, Steenkamp JBE (2013) Price and advertising effectiveness over the business cycle. *Journal of Marketing Research* 50(2):177–193.

Vijjali R, Bhageria D, Tamhane A, TM M, Sathyanarayana J (2022) Foodnet: Simplifying online food ordering with contextual food combos. *5th Joint International Conference on Data Science & Management of Data (9th ACM IKDD CODS and 27th COMAD)*, 178–185.

Wakefield KL, Inman JJ (2003) Situational price sensitivity: the role of consumption occasion, social context and income. *Journal of Retailing* 79(4):199–212.

Wang K, Goldfarb A (2017) Can offline stores drive online sales? *Journal of Marketing Research* 54(5):706–719.

WSJ (2023) Instacart's revenue and profit climb ahead of public listing. https://www.wsj.com/articles/instacart-sees-revenue-profit-boost-ahead-of-public-listing-1d7891d, (Accessed on 05/12/2023).

Zentner A, Smith M, Kaya C (2013) How video rental patterns change as consumers move online. *Management Science* 59(11):2622–2634.

Web Appendix

Table of Contents

A Detailed Data Description **2**

A.1 Numerator Managed Panel . . . 2

A.2 Numerator vs. Earnest Research Instacart Sales data . . . 2

A.3 Demographics of the Numerator Managed Panel . . . 2

B Characteristic Trips and Classification Details **4**

B.1 Identifying Characteristic Trips . . . 4

B.2 Intuition Behind Shopping Basket Embeddings . . . 6

B.3 Classification Approaches . . . 7

B.4 Validation of Classification Approaches . . . 8

B.5 Optimal Number of Clusters . . . 9

C Robustness of Main Findings **11**

C.1 Covid-19 and Reduced Mobility . . . 11

C.2 Price Differences . . . 16

C.3 Other Alternative Explanations . . . 18

D Additional Results **20**

D.1 Variety Results In Table Format . . . 20

D.2 Variety Differences Across Time . . . 20

D.3 Variety Differences and Household Heterogeneity . . . 22

D.4 Variety Differences Across Classification Approaches . . . 22

D.5 Poisson Specification . . . 22

D.6 Basket Variety and Similarity by Experience with Instacart . . . 24

D.7 Stockpiling . . . 25

D.8 Evaluating Substitution of BM and Instacart Characteristic Trips . . . 26

D.9 Category Compensation Results . . . 27

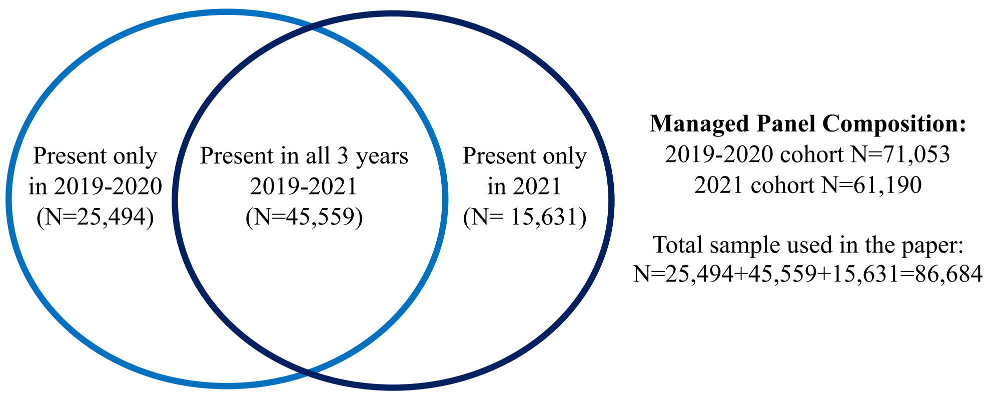
D.10 Decomposing Extensive and Intensive Margins in Category Changes . . . 29

A Detailed Data Description

A.1 Numerator Managed Panel

The data from Numerator captures the shopping behavior of 132,243 unique households (47,321,908 trips) from 2019 to 2021. Not all of these households consistently upload their shopping trip receipts, therefore Numerator does not include such households in their managed panel—a static group of consumers who consistently report shopping data and have verified demographic information. Consequently, 86,884 households are part of this managed panel, which constitutes the primary sample that we use in the paper.[10] This panel is built to be representative of the US population and normalized to account for the expected panelist churn. We had two sequential pulls of the data from the company, one for the 2019-2020 panelist cohort and another one for the 2021. Figure A1 illustrates the composition of our sample that covers the years 2019-2021. We observe Instacart purchases only for the subset of panelists (N=4,388) who gave Numerator permission to access their Instacart shopping data.

Figure A1: Numerator Managed Panel Composition



A.2 Numerator vs. Earnest Research Instacart Sales data

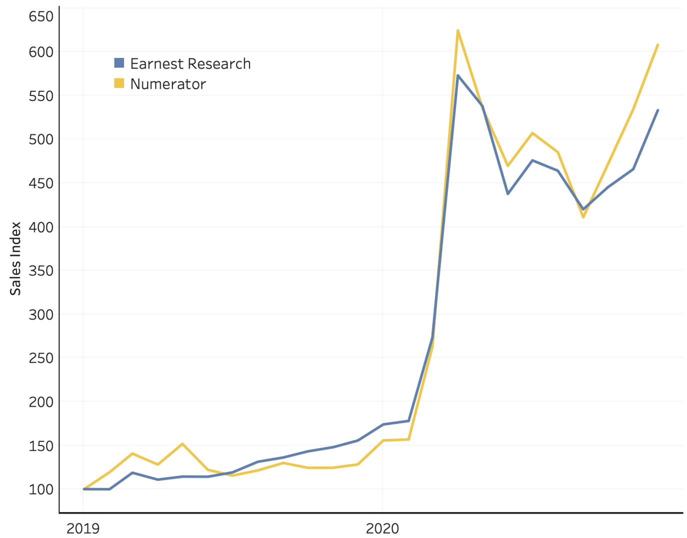
Figure A2 compares the Instacart sales index (normalized to 100 for January 2019) between two data providers: Numerator and Earnest Research. Earnest Research data rely on credit and debit card transaction data from 1.56 million panelists, which are meant to be representative of the US population. Both data sources show similar sales patterns during 2019-2020.

A.3 Demographics of the Numerator Managed Panel

Table A1 summarizes the demographic variables that we observe in the Numerator data for the full managed panel sample (N=86,684) as well as for the permission sample (N=4,388)

[10]We use *all* the trips from *all* 132,243 households to train our custom Skipgram model.

Figure A2: Instacart Sales Trend 2019-2020



where we observe at least one Instacart purchase. The breakdown of the summary statistics is very similar for the 2019-2020 and 2021 cohorts separately.

Table A1: Summary Statistics of the Sample vs. All Households

		Instacart (N=4,388)		All (N=86,684)	
		N	%	N	%
Education Level	Advanced	914	20.8	15066	17.4
	College	2885	65.7	53857	62.1
	High School	537	12.2	15658	18.1
	Less than high school	52	1.2	2103	2.4
Income Quartile	1(less than $49,999)	1248	28.4	30114	34.7
	2($50,000 - $79,999)	1012	23.1	20564	23.7
	3($80,000 - $124,999)	1139	26.0	989	22.5
	4($125,000 - $250,000+)	989	22.5	15599	18.0
Ethnicity	Asian	288	6.6	5821	6.7
	Black or African American	553	12.6	9030	10.4
	Hispanic/Latino	526	12.0	9381	10.8
	Other	77	1.8	1687	1.9
	White/Caucasian	2944	67.1	60765	70.1
Percentage of trips to:					
Food, Mass, Club retailers		71.02%		75.95%	
Gas, Convenience stores		5.79%		6.05%	
Dollar stores		6.36%		7.46%	
Drug stores		5.24%		4.17%	

B Characteristic Trips and Classification Details

B.1 Identifying Characteristic Trips

To identify characteristic shopping trips using embeddings, we follow two main steps. The first step is to obtain a *representation* of each shopping basket in a lower dimensional vector space. In the second step, we use these vector representations to *classify* trips into characteristic and non-characteristic trips. Our analysis will consider five different pathways for the classification step for the purpose of establishing trip types and demonstrating robustness. These ML-based classification algorithms are cross-checked using human evaluation of the random subset of trips in a validation exercise described later in Appendix B.4.

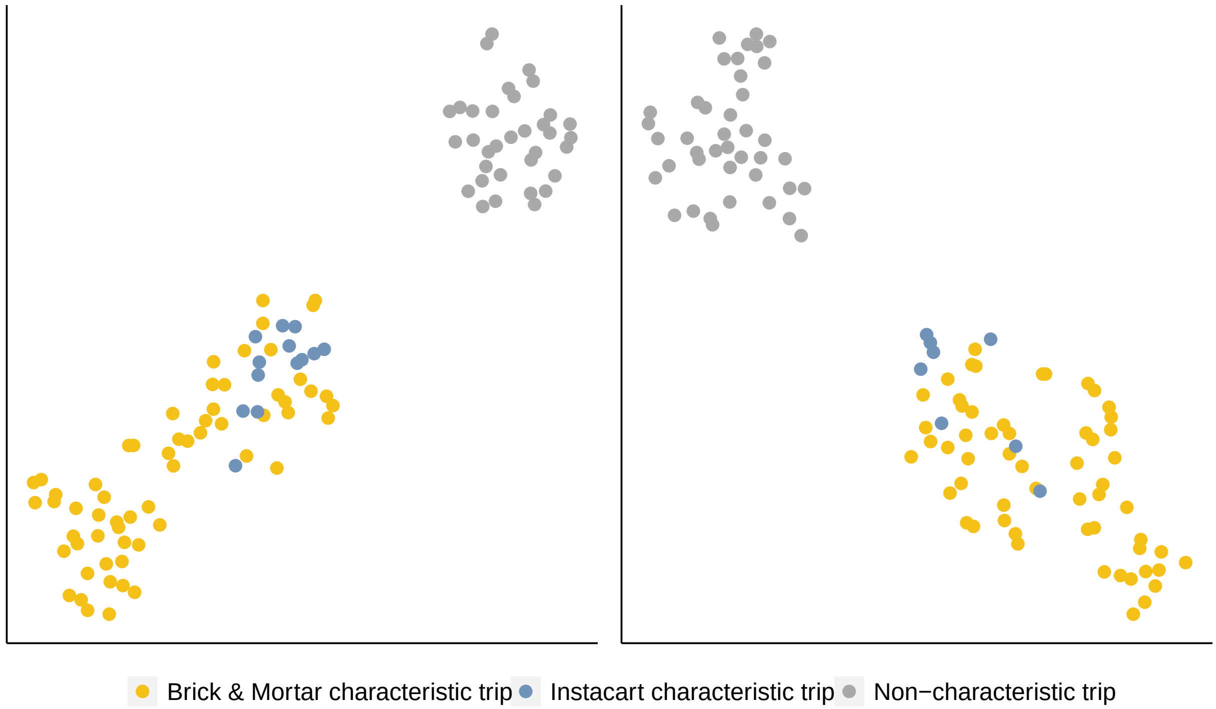
For the *representation* step, we adopt the Skipgram algorithm (Mikolov et al. 2013) originally developed in Natural Language Processing (NLP), which is used to create word embeddings—low-dimensional vector representations of words. These embeddings encode the semantic meaning of words such that words that are semantically similar lie close to each other in the vector space. Word2Vec model is the most well-known application of the Skipgram model, where words are the input to the model, and embeddings capturing the semantic relationship between words are the output. The original Word2Vec model is trained on a Google News dataset of about 100 billion words. Analogously, in our custom application of the Skipgram model, product categories become the input to the model in order to obtain the product category embeddings that capture the contextual relationships between product categories[11]. The resulting product category embeddings are in 100 dimensions. Our Skipgram model is trained using the *gensim* package in Python on the entire shopping trip data (not just the permission sample) consisting of more than 47 million shopping trips (see section A.1 for more details on the entire dataset).

A methodological advantage of the Skipgram-based model is that it uses not only the information about the product categories that are frequently bought together but also takes into account the similarity across the observed shopping contexts that a given category is present within. That is, categories observed within similar shopping contexts are assigned similar vector embeddings. The model learns these latent associations using a shallow two-layer Neural Network. As such, any two product category embeddings observed to be similar (as measured by the distance between those vectors) can be interpreted as categories that are closely related to each other. We then construct the basket-level embeddings by summing up the embeddings of the constituent categories in a particular basket. Using this approach, two baskets with similar constituent product categories will have the respective basket embeddings close to each other in the vector space.

For the *classification* step, we apply K-means clustering on the resulting basket embeddings at a household level to identify the characteristic and non-characteristic trips. We note that the use of clustering-type methods has been commonly used in marketing research in order to better understand co-occurrences in the data (Netzer et al. 2012). The optimal number of clusters for this step are identified using the approach described in Appendix

[11] We did not explicitly use the number of categories as an input in the training or classification procedure, but this outcome measure will be correlated with the probability of observing a certain trip classified as a characteristic trip. However, our classification process is channel-agnostic: The algorithm does not use any information about whether a trip is BM or Instacart.

Figure B1: t-SNE Plots of Grocery Trips for Two Sample Households



Notes: The t-SNE plots show the trips as two-dimensional points after reducing the dimensions from 100 to 2.

section B.5.

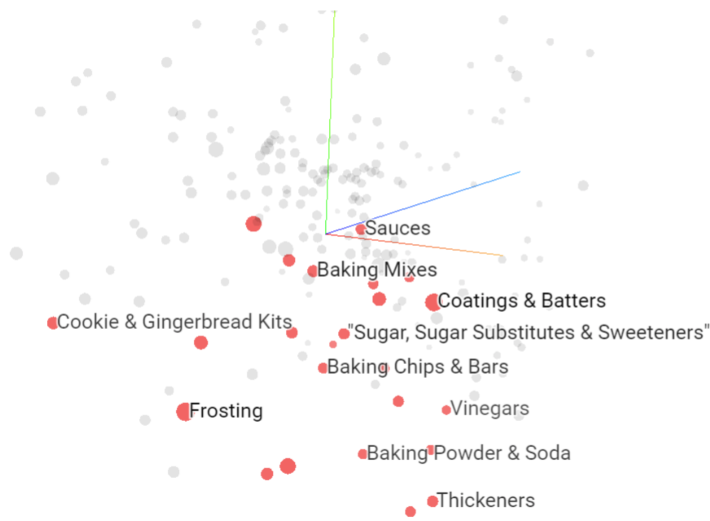
Most importantly, our classification approach is agnostic to the channel type (we do not include information on whether a basket is from BM or Instacart channel) to facilitate the identification of comparable baskets. We inspect the resulting classifications and find that all the trips in one cluster have larger baskets/spend and are more consistent in basket composition within each household, resembling regular restocking trips. Therefore, we label this cluster as containing characteristic trips. On the other hand, the trips in the second cluster are distinctly different, usually containing significantly fewer items that do not reflect household restocking behavior. We label those trips as non-characteristic trips.

As mentioned in the paper, one limitation of our data is that Numerator provided us with a subset of data limited to the grocery sector (i.e., food) items. Consequently, while we can observe all the food items and the overall basket totals, we do not observe non-food items purchased during the same shopping trips. However, considering that food items constitute the majority of the items in the analyzed baskets, we expect that the inclusion of non-food items would likely not have a significant impact on the classification. Nonetheless, we leave the formal evaluation of this matter to future research that may have access to more comprehensive basket data.

B.2 Intuition Behind Shopping Basket Embeddings

Figure B1 illustrates t-SNE plots[12] of basket-level embeddings for two sample households. If two shopping baskets are similar to each other, then these two baskets should appear close to each other on the t-SNE plot. In these representative plots, we see a clear separation between the characteristic and non-characteristic trips. Figure B2 shows the t-SNE plot of product categories primarily belonging to the bakery department. The t-SNE coordinates in the three-dimensional space are mapped from 100-dimensional vectors obtained from our Skipgram model. The similarity between two categories (i.e., any two dots in the plot), is determined by the distance between the two category embedding vectors. This figure highlights select dots in red indicating various categories in the bakery department and shows that they are close to each other in the vector space.

Figure B2: t-SNE Plot in Three Dimensional Space of Baking Department Categories



Furthermore, Figure B3 illustrates two receipts summarized at the category level that are classified as two characteristic trips for a specific household—one BM and one Instacart trip. We observe several patterns that help illustrate the flexibility of our algorithm. First, the receipts need not be identical to be classified as characteristic. Second, there is some exact overlap in the categories indicating that this is a characteristic restocking trip. Third, some categories are very similar. For example, BM trip has a category "Salad Dressings" and Instacart trip has a category "Mayonnaise & Mayonnaise Dressing", recognized by our algorithm as similar categories. The higher the overlap of the exact or similar items, the higher the similarity between the baskets.

[12]t-SNE plot stands for t-distributed stochastic neighbor embedding plot, which is a commonly used statistical method for visualizing high-dimensional data by giving each data point a location in a two or three-dimensional map.

Figure B3: Two Characteristic Trips within one Household

(a) Brick & Mortar Characteristic Trip

MAJOR_CATEGORY	ITEM_TOTAL
Beans	3.2
Beef	16.36
Cheese	8.5
Eggs	2
Frozen Vegetables	2
Fruits	2.59
Herbs, Spices & Seasonings-Single	4.25
Milk, Cream, & Milk Substitutes	1.88
Oil & Shortening	2
Packaged Cookies	4.49
Pasta	2.64
Poultry	8.95
Salad Dressings & Toppings-Ref	4.99
Salad Mixes & Salad Kits	5
Sauces	0.59
Sports & Energy Drinks	2
Stocks & Broths	3.32
Sugar, Sugar Substitutes & Sweeteners	2.5
Vegetables	12.1
Yogurt & Yogurt Drinks	5.99
GRAND TOTAL	95.35

(b) Instacart Characteristic Trip

MAJOR_CATEGORY	ITEM_TOTAL
Beef	8.91
Cheese	3.99
Cheese-Deli	5.7
Chips	5.49
Cold Cuts-Deli	7.25
Frozen Potato Snacks	3.88
Frozen Vegetables	1.25
Fruits	5.67
Mayonnaise & Mayonnaise Dressings	3.77
Oil & Shortening	6
Packaged Rolls & Buns	2.59
Pasta	1
Popcorn	4.83
Poultry	4.69
Salad Mixes & Salad Kits	3.58
Sauces	3.78
Snack Seeds, Nuts & Trail Mixes (Snack)	9.99
Tea	1.49
Vegetables	9.6
Yogurt & Yogurt Drinks	5
GRAND TOTAL	98.46

B.3 Classification Approaches

In this section, we now explore different classification approaches to establish robustness and generalizability of our main findings, as well as suggest our preferred classification approach. For our sensitivity analysis, we consider four different classification approaches based on basket embeddings. For completeness, we also implement the fifth classification approach based on department category totals.

Table B1: Summary of Classification Approaches

	Classification Type	Classification Level
Classification 1	Basket embeddings	Household
Classification 2	Basket embeddings	Household×Retailer
Classification 3	Basket embeddings	Household×Quarter
Classification 4	Basket embeddings	Household×Retailer×Quarter
Classification 5	Department totals	Household

The five classification approaches are summarized in Table B1. Classification 1 identifies characteristic trips across all trips for a given household. Classification 2 allows for the characteristic trips to differ across different retailers. Since we are interested in a household's representative restocking shopping baskets, we focus on the primary retailers for this classification. Classification 3 is a more flexible version of Classification 1, as it allows for the characteristic trips to change over every quarter. For example, if a household becomes income-constrained and their shopping patterns change, this classification approach would be flexible enough to allow for those systematic changes. Similarly, Classification 4 is a more flexible version of Classification 2, as it allows for both retailer- and time-specific shopping pattern changes. Finally, Classification 5 no longer uses basket-embeddings; instead, it uses only department expenditure sub-totals, which is a classification approach similar to the one implemented in Jindal et al. (2020). We note that department is a significantly coarser level classification than product categories, as there are 26 different departments in our data (e.g., Frozen foods, Meat, and Produce). Table B2 summarizes shopping basket outcomes of interest across these five approaches.

Table B2: Summary Statistics by Trip Type using Different Classification Approaches

	Characteristic trip		Non-characteristic trip	
	mean	sd	mean	sd
Classification 1: *k-means clustering using basket embeddings*				
Grocery amount per trip (in $)	76.37	52.75	17.80	22.42
Number of items bought per trip	25.16	17.12	4.98	5.98
Number of categories bought per trip	13.11	6.26	2.85	2.18
Number of trips	433,791		1,534,601	
Overlap with human classification	87.68%			
Classification 2: *k-means clustering using basket embeddings by retailer*				
Grocery amount per trip (in $)	79.49	56.13	23.00	25.74
Number of items bought per trip	25.67	17.99	6.92	7.83
Number of categories bought per trip	13.31	6.85	3.89	3.17
Number of trips	298,632		662,130	
Overlap with human classification	79.33%			
Classification 3: *k-means clustering using basket embeddings by quarter*				
Grocery amount per trip (in $)	72.64	52.69	17.64	22.54
Number of items bought per trip	23.93	17.05	4.92	6.05
Number of categories bought per trip	12.50	6.39	2.81	2.20
Number of trips	467,543		1,500,849	
Overlap with human classification	88.92%			
Classification 4: *k-means clustering using basket embeddings by retailer and quarter*				
Grocery amount per trip (in $)	73.39	55.33	24.46	29.50
Number of items bought per trip	23.81	18.01	7.32	8.75
Number of categories bought per trip	12.41	7.01	4.07	3.67
Number of trips	316,102		644,660	
Overlap with human classification	75.42%			
Classification 5: *k-means clustering using department totals*				
Grocery amount per trip (in $)	84.74	65.19	23.90	29.08
Number of items bought per trip	23.48	21.05	7.66	9.95
Number of categories bought per trip	11.30	8.34	4.33	4.47
Number of trips	220,278		1,748,114	
Overlap with human classification	61.16%			

Notes: Total number of trips: 1,968,392; Total number of primary retailer trips: 960,762

B.4 Validation of Classification Approaches

To assess the validation of our classification approach, we conduct a Qualtrics survey on a test pool of $N = 290$ student participants. We constructed a randomly drawn sub-sample of 1,800 shopping trip receipts from 300 randomly drawn households that contain both characteristic and non-characteristic trips (from both BM and Instacart channels). Each participant was asked to review and classify six receipts each from five different households; so a total of 30 receipts per participant. On average, each shopping trip receipt was evaluated by 4.8 participants.

Note that unlike an exercise to demonstrate external validity, in this survey, we are imposing the trip classification where our objective is to simply validate that humans will label the trips similarly when provided the classification our analysis relies upon.

Before participants evaluated the receipts, the survey introduced the intuition behind a *regular* shopping trip, which was explained to be a household-specific regular restocking trip of commonly used household grocery items. Subjects were then asked first to review all six receipts for a given household and then to classify each receipt as a regular or a non-regular (intermittent) shopping trip for that household.

We report the percentage overlap across human-classified and machine-classified trips in Table B2. The results of the survey suggest that Classifications 1 and 3 have similar and higher overlap rates (88% and 89%, respectively) with human classification relative to Classifications 2 and 4. Furthermore, Classification 5, which uses information about only the department totals and *does not* use basket embeddings (i.e., there's no representation step), has noticeably the lowest overlap rate of around 61%. When manually investigating differences in classification approaches, we find that Classification 5 fails to capture the relationships between the frequently co-purchased categories and departments and instead interprets them as discrete silos. On the other hand, our embeddings-based classifications appear to represent these relationships well and use the information that some categories/departments are more related than others in classifying the trips. In summary, Classification 3 yields the highest overlap rate with human classification, and thus we indicate the results using Classification 3 as our preferred classification (though we note that our findings are robust across all classification approaches).

B.5 Optimal Number of Clusters

The silhouette coefficients (Rousseeuw 1987) incorporate intra-cluster (a measure of tightness) and inter-cluster (a measure of separation) distances, and reward specifications that excel at these dimensions. Since we classify the trips at household level, we calculate the silhouette scores for number of clusters varying from 2 to 10 for each household. The mean silhouette scores are presented in Table B3. K-means clustering with 2 clusters ($K = 2$) yields the highest mean value of silhouette scores. Therefore, we assess the optimal number of clusters is 2.

Table B3: Mean silhouette scores for K-means clustering

Number of Clusters	Classification 1	Classification 2	Classification 3	Classification 4	Classification 5
2	**0.511**	**0.386**	**0.505**	**0.399**	**0.489**
3	0.359	0.271	0.341	0.271	0.383
4	0.251	0.223	0.246	0.215	0.331
5	0.186	0.204	0.201	0.191	0.298
6	0.153	0.197	0.180	0.178	0.276
7	0.136	0.195	0.169	0.168	0.257
8	0.127	0.195	0.164	0.161	0.240
9	0.123	0.195	0.161	0.154	0.232
10	0.120	0.195	0.160	0.146	0.219

C Robustness of Main Findings

We establish robustness of our main empirical findings. In particular, we demonstrate that our findings are not driven by the reduced mobility and increased stay-at-home incidence due to Covid-19. Moreover, we provide evidence that pricing differences between the two channels and other potential alternative mechanisms are unlikely to explain our main findings. Note that even though these specifications demonstrate robustness of the main empirical patterns, we emphasize that these tests do not provide grounds for a causal interpretation.

C.1 Covid-19 and Reduced Mobility

In order to examine whether or not our main results are identified primarily from households switching to Instacart due to Covid-19 concerns and avoidance of in-person shopping, we consider another robustness test that uses household mobility rates. The underlying assumption is that reduced mobility rates are correlated with the localized level of concern due to Covid-19 risk and thus the propensity to use Instacart might be higher in order to avoid in-person shopping. If our results are driven by such Instacart cases, then we will be unlikely to make a generalizable claim that such effects would persist outside of pandemic conditions that might subside over time.

For this robustness check, we use the Social Distancing Metrics dataset from SafeGraph, a company that specializes in collecting the location-based data of over 20 million anonymous mobile devices in the US and Canada. The Social Distancing Metrics are generated using a panel of GPS pings from anonymous mobile devices to determine the geographic movement of a given device in a given time period. The dataset also has information about the total number of devices in a zipcode. We use the ratio of the number of devices categorized as completely at home and the total number of devices in a given zipcode to compute the *stay at home rate*, for each zipcode-quarter combination. These stay-at-home metrics from the SafeGraph data have also been widely adopted in recent research to study pandemic-related questions (Gao et al. 2020, de Vaan et al. 2021, Chiou and Tucker 2020).

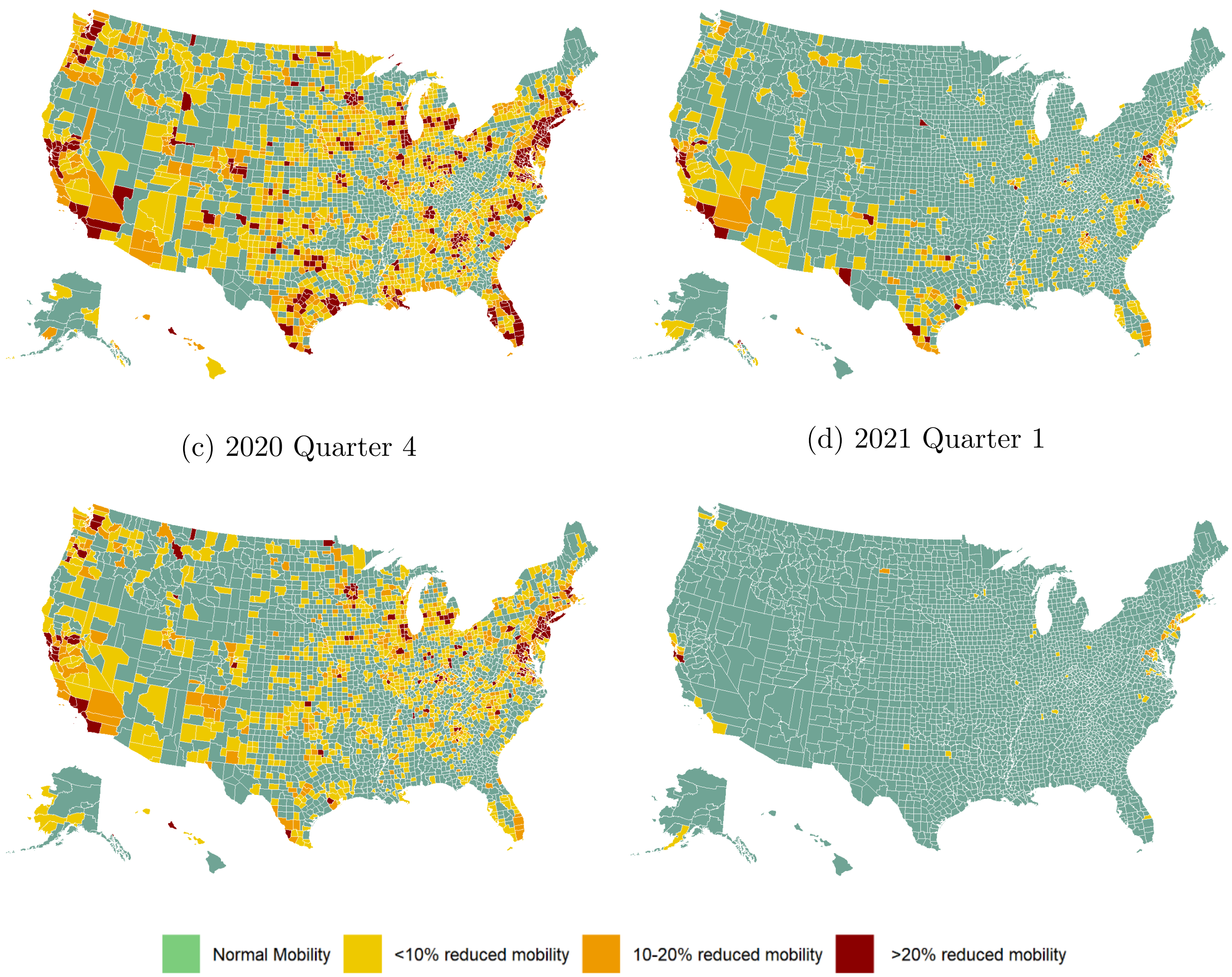
Our analysis evaluates mobility rates for each zipcode for nine quarters: four quarters in 2019, four quarters in 2020, and one quarter in 2021. Safegraph discontinued the social distancing dataset availability after April 2021, therefore we are unable to include data for the last three quarters in 2021. By construction, since we are studying mobility constraints due to the pandemic, all zipcode-quarter combinations in 2019 are considered to have normal mobility rates. The rest of the zipcode-quarter combinations are evaluated relative to their respective baselines in 2019: For example, the relative mobility level for the first quarter in 2020 is calculated as the stay-at-home rate in that quarter in 2020 divided by the stay-at-home rate in the same quarter in 2019 for the same zipcode. If the mobility level is less than or equal to 1 (indicating that the mobility levels have not changed or increased from the baseline quarter in 2019), we categorize that zipcode-quarter as exhibiting normal mobility rates.

Next, we keep the subset of zipcode-quarter combinations for which the mobility rates are classified as normal and exclude the observations in zipcode-quarter combinations with decreased mobility levels. Figure C1 illustrates the locations with normal and reduced mobility

Figure C1: Mobility Heatmaps by Quarter

(a) 2020 Quarter 2 (b) 2020 Quarter 3

(c) 2020 Quarter 4 (d) 2021 Quarter 1



Notes: These maps represent relative county-level mobility rates by quarter: For each county-quarter combination, we calculate the % decrease in mobility relative to the baseline mobility in 2019 for the same quarter and county. The corresponding figure for 2020 Q1 is overwhelmingly green since it is a pre-pandemic quarter. We use stay-at-home rates calculated from SafeGraph data.

rates by quarter.[13] As such, our robustness test estimates Equation 2 using only observations from normal mobility zipcode-quarter combinations. Table C1 reports the results for basket variety comparisons and Table C2 reports the results for basket composition comparisons. We find that the main basket variety and composition results are robust, albeit slightly lower in magnitude,[14] indicating that our baseline basket variety and composition differences appear not to be driven by Covid-19.

Furthermore, Table C2 displays the basket composition results, which complement the

[13]Note that the figure represents county-level mobility rates, while in estimation we use more granular zipcode level mobility rates

[14]While the point estimates are slightly lower than in our baseline specification, they are not statistically significantly different due to overlapping confidence intervals.

Table C1: Variety Differences (Only Observations Without Reduced Mobility due to Covid-19)

	(A) Normal Mobility Subset All Nine Quarters	**(B) Normal Mobility Subset All Quarters excluding 2019**
Instacart	-0.071***	-0.079***
	(0.011)	(0.013)
Num.Obs.	189,344	79,749
R^2	0.665	0.635

Notes: This table reports the basket variety results where we keep the subset of zipcode-quarter combinations for which the mobility rates are classified as normal and exclude the combinations with decreased mobility levels. Outcome variable is *log(Number of Categories)*. All models use $HH \times Retailer \times Quarter$ and $HH \times Month$ fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.

basket variety results reported in Table C1 for the Covid-19 mobility test.

As an additional robustness test for confounds related to the pandemic, we design a placebo test to show that we do not find the documented Instacart variety and composition differences where we do not expect to find them. This placebo test is particularly suited to demonstrate that the Instacart basket composition differences are not correlated with the systematic changes in BM baskets that might be induced by the factors related to the pandemic.[15] For example, is it possible that the fewer purchases of the vegetables that we find in Instacart baskets is correlated with the possibility that consumers consume fewer vegetables during the pandemic or that fewer vegetables are available on store shelves?[16] This placebo test explores whether or not this confound is an important concern.

Among our sample households, 71.78% of them begin using Instacart after the first Covid-19 lockdown (March 19th, 2020). We refer to these households as the *pandemic Instacart adopters*. Pandemic Instacart adopters use Instacart in one out of every five regular grocery shopping trips. We proceed by constructing a set of *placebo Instacart adopters* by using a 1:1 nearest neighbor matching procedure between pandemic Instacart adopters and the households who are never observed to use Instacart. In other words, for every pandemic Instacart adopter, we identify a household that is very similar in all observable dimensions *except* for having adopted Instacart. Since the Covid-19 based restrictions were usually state-specific, we require for these two households to be located in the same state. We match the households on the demographic characteristics such as gender, age, marital status, family size, employment, and income level. In addition to these demographic features, we also add matching criteria based on the household's shopping patterns: the number of characteristic trips, the number of characteristic trips post the first lockdown, average number of days between characteristic trips, average spend per characteristic trip, average number of categories purchased in a characteristic trip, and the individual quantities of

[15] An important caveat of this analysis is that we cannot confirm with complete certainty that the households that we observe without any Instacart purchases are the ones that did not use Instacart, as there may be data omission resulting from data permission constraints.

[16] Note that our specification in Equation 2 already controls for that potential confound to a certain extent as our effect is identified from the variation in basket composition *within* household-retailer-quarter combination after controlling for seasonality.

Table C2: Basket Composition Differences (Only Observations Without Reduced Mobility due to Covid-19)

Less Category Name	Estimate	More Category Name	Estimate
Fresh Vegetables	-0.127***	Frozen Meat	0.036***
	(0.024)		(0.010)
Candy	-0.082***	Frozen Vegetables	0.006
	(0.011)		(0.010)
Cheese	-0.031	Frozen Breakfast	0.013
	(0.017)		(0.010)
Bakery Desserts	-0.035**	Frozen Fruit	0.012**
	(0.013)		(0.004)
Savory Snacks	-0.039**		
	(0.014)		
Fresh Meat	-0.041*		
	(0.020)		
Bread	-0.011		
	(0.014)		
Sweet Snacks	-0.024*		
	(0.010)		
Deli	-0.029*		
	(0.012)		
Pasta	-0.019		
	(0.010)		

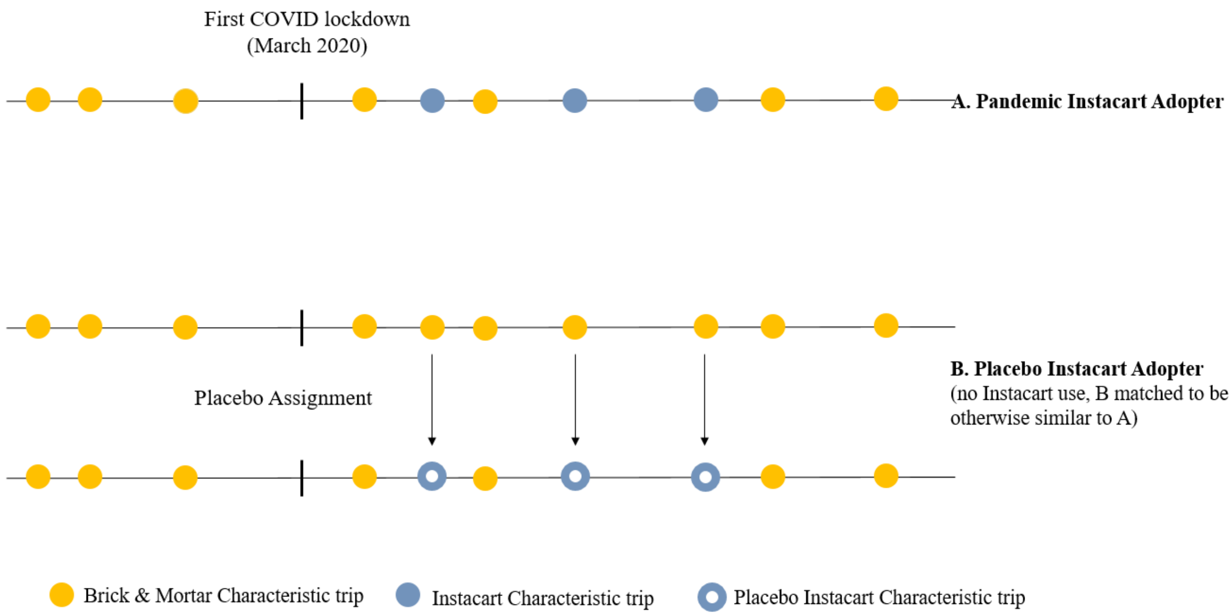
Notes: This table mirrors the categories from Table 7, but for the households with normal mobility rates. For each regression, the dependent variable is log(1+Number of Items) from that respective category. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in the parentheses. Cheese and Pasta categories are significant at p<0.10 level.

each of the top ten categories for which we estimated systematic differences in Instacart characteristic trips (see Table 7).

The placebo trip assignment procedure is illustrated in Figure C2. In this illustrative example, let Household A (adopts Instacart after Covid-19), and Household B (never observed to use Instacart) be the matched households. We count the number of Instacart characteristic trips of Household A, which is three, in this example. We then randomly assign three placebo Instacart characteristic trips to Household B in the same week (i.e., we mark three BM characteristic trips of Household B as placebo Instacart characteristic trips).[17] Next, we

[17] 116 placebo Instacart adopter households could not be assigned placebo Instacart characteristic trips due to the same week criterion. We exclude those placebo Instacart adopter households and the corresponding pandemic

Figure C2: Placebo Test Illustration



estimate the same specifications as in section 4.2 to evaluate basket variety and composition effects for actual and placebo Instacart adopters separately. Table C3 reports the results for basket variety and Table C4 reports the results for basket composition. These results confirm, as expected, that the basket variety and composition effects that are present among the actual Instacart adopters are no longer present among the placebo Instacart adopters. Therefore, we conclude that our main results are not an artifact of systematic changes in basket variety or composition due to the pandemic.

Table C3: Robustness tests: Placebo Results

	(A) Pandemic Instacart Adopters	**(B) Placebo Instacart Adopters**
Instacart	-0.131***	0.002
	(0.011)	(0.006)
Num.HH.	3,034	3,034
Num.Obs.	301,795	327,639
R^2	0.855	0.863

Notes: Pandemic Instacart Adopters are households that are observed to use Instacart for the first time after Covid-19 onset. Placebo Instacart Adopters are matched households who never adopted Instacart. Outcome variable is *log(Number of Categories)*. All models use $HH \times Retailer \times Quarter$ and $HH \times Month$ fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.

Furthermore, Table C4 reports the basket composition results, which complement the basket variety results reported in Table C3 for the placebo test.

Instacart adopter households from the estimation.

Table C4: Basket Composition Test Results - Pandemic Instacart Adopters and Placebo Instacart Adopters

(A) Pandemic Instacart Adopters		(B)Placebo Instacart Adopters	
Category Name	**Estimate**	**Category Name**	**Estimate**
Fresh Vegetables	-0.170***	Fresh Vegetables	0.007
	(0.024)		(0.012)
Candy	-0.067***	Candy	0.008
	(0.009)		(0.008)
Cheese	-0.101***	Cheese	-0.001
	(0.019)		(0.010)
Bakery Desserts	-0.083***	Bakery Desserts	-0.001
	(0.012)		(0.009)
Savory Snacks	-0.058***	Savory Snacks	0.001
	(0.014)		(0.009)
Fresh Meat	-0.068***	Fresh Meat	-0.001
	(0.018)		(0.011)
Bread	-0.028*	Bread	-0.004
	(0.012)		(0.008)
Sweet Snacks	-0.039***	Sweet Snacks	-0.010
	(0.009)		(0.007)
Deli	-0.037***	Deli	0.010
	(0.010)		(0.007)
Pasta	-0.028**	Pasta	0.009
	(0.008)		(0.007)

Notes: This table mirrors the categories from Table 7, but for the placebo test. For each regression, the dependent variable is log(1+Number of Items) from that respective category. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in the parentheses.

C.2 Price Differences

One candidate alternative mechanism that could potentially explain Instacart's lower variety and composition might have to do with price differences between the two channels: Consumers might strategically reduce their consumption of items when they face higher prices for some categories over others. In general, consumers are aware and expect that prices are typically higher when purchased via Instacart as opposed to BM (Campbell 2019), but the extent to which online grocery channels mark up their prices can vary across products, categories, and time (Aparicio et al. 2021). A direct test of the pricing mechanism is challenging given that price information is only available for items that are purchased. If an item is purchased at a BM grocery store, we do not know the counterfactual price of that item on Instacart at the time of purchase and vice versa. Similarly, if an item is not purchased due to its price, we do not observe the price in those instances either. Furthermore, as we discuss in section 2.1, the data limitations prevent us from linking item-level data in BM and Instacart baskets to be able to compare prices. In light of these data-specific limitations that prevent us from analyzing price information directly, we develop a series of indirect tests to explore the potential role of pricing.

To develop these indirect tests, we turn to insights from past research that uncover linkages between price sensitivity with income (Sangani 2022, Meyer et al. 2014, Mulhern et al. 1998, Orhun and Palazzolo 2019, Wakefield and Inman 2003) and general macroeconomic conditions (Gijsenberg 2017, Gordon et al. 2013, Van Heerde et al. 2013). These studies have shown that consumers with lower income or those who face unfavorable macroeconomic conditions (e.g., negative income shocks or unemployment) are more price elastic. As household self-reported income is directly observable in our data, we investigate the heterogeneity of our main results for basket variety and composition across different income groups. Intuitively, households who are most price sensitive might be the ones most impacted by higher prices in the online channel and thus display the largest basket variety and composition changes. Therefore, if price differences are a driving factor in online versus offline grocery shopping, we should observe noticeable differences across the income groups. On the other hand, if we find similar basket variety and composition effects across different income groups, we can conclude that price differences are not the driving mechanism behind our main results.

Table C5: Variety Differences by Income Quartiles

	Estimate
Instacart x First Quartile	-0.087***
	(0.016)
Instacart x Second Quartile	-0.085***
	(0.015)
Instacart x Third Quartile	-0.107***
	(0.011)
Instacart x Fourth Quartile	-0.101***
	(0.014)
Num.Obs.	434,322
R^2	0.658

Notes: This table reports variety differences across household income quartiles. Outcome variable is *log(Number of Categories)*. All models use $HH \times Retailer \times Quarter$ and $HH \times Month$ fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses. Appendix Table A1 reports the income quartile cutoffs.

Table C5 and Table C6 report the results by income quartile for basket variety and composition effects, respectively. These results reveal that the main effects are indeed very similar across the income quartiles. While individual point estimates across the four income quartiles are not statistically significantly different from each other, the pattern of the relative magnitudes of these point estimates is actually in the opposite direction of what we would expect if prices were indeed the primary reason for basket variety and composition adjustment (we would expect variety differences to be higher among the lower-income households relative to the higher-income households). Therefore, we conclude that price differences are an unlikely driving mechanism behind our main results.

Furthermore, Table C6 reports the basket composition results, which complement the basket variety results reported in Table C5 for the Income quartiles.

Table C6: Categories with Fewer Purchases in Instacart (as Compared with BM) by Income Quartile

Category Name	**Instacart x I Quartile**	**Instacart x II Quartile**	**Instacart x III Quartile**	**Instacart x IV Quartile**
Fresh Vegetables	-0.146***	-0.137***	-0.155***	-0.111***
	(0.033)	(0.035)	(0.031)	(0.033)
Candy	-0.053**	-0.056*	-0.078***	-0.085***
	(0.016)	(0.013)	(0.010)	(0.016)
Cheese	-0.057**	-0.061**	-0.050*	-0.074**
	(0.022)	(0.021)	(0.023)	(0.022)
Bakery Desserts	-0.050**	-0.067**	-0.074***	-0.046*
	(0.019)	(0.018)	(0.015)	(0.015)
Savory Snacks	-0.014	-0.045**	-0.071***	-0.048**
	(0.021)	(0.016)	(0.016)	(0.018)
Fresh Meat	-0.029	-0.068**	-0.037	-0.040
	(0.027)	(0.024)	(0.023)	(0.023)
Bread	-0.037	-0.017	-0.023	-0.039*
	(0.019)	(0.016)	(0.014)	(0.018)
Sweet Snacks	-0.020	-0.020	-0.024*	-0.045**
	(0.014)	(0.013)	(0.011)	(0.016)
Deli	-0.021	-0.023	-0.014	-0.046**
	(0.014)	(0.014)	(0.014)	(0.016)
Pasta	-0.036*	-0.025*	-0.021*	-0.023*
	(0.015)	(0.011)	(0.009)	(0.009)

Notes: This table mirrors the categories from Table 7 for different Income Quartiles. For each regression, the dependent variable is *log(1+Number of Items)* from that respective category. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in the parentheses.

C.3 Other Alternative Explanations

While it is beyond the scope of this paper to evaluate the exhaustive list of potential mechanisms and alternative explanations, in this section, we provide an informal discussion outlining other potential explanations and their likelihood of explaining the totality of our results. Some of these assertions are informed by helpful discussions with Instacart economists and data scientists themselves.[18]

We start by evaluating whether the main variety and composition effects that we document can be explained by the differences in item availability across the shopping channels. When we restrict the analysis to the categories that are available in both BM and Instacart channels, we obtain quantitatively similar results. While we have no data to accurately evaluate the inventory differences across Instacart and the respective BM stores at any given time, we believe that this mechanism is unlikely to explain our results for several reasons. First, since items from shopping receipts are classified into shopping categories by Numerator algorithms (and are not the categories assigned by retailers or delivery providers), we observe items assigned into a standardized set of categories across all retailers and channels. The categories where we find most noticeable basket composition effects—vegetables and impulse purchase categories—are well represented on Instacart (we confirm this with our data). Second, Instacart and its respective BM store inventory are designed to be the same—in many cases, BM stores share their inventory databases with Instacart, and this information is updated daily (or even minute-by-minute) based on real-time data directly sent to Instacart from the BM stores (Rao and Zhang 2021). Most importantly, our Instacart basket variety and composition effects are identified from the variation of basket composition *within* a household, retailer, and quarter (or month) combination. Thus, inventory stock-outs in one channel would likely be mirrored by stock-outs in the other channel at a given retailer and time period, thus canceling out in our estimation. Finally, online grocery apps tend to substitute items when the requested item is missing (Haddon 2019), thereby increasing the likelihood that consumers purchase items they may not have initially planned on purchasing. We would then expect the basket variety to remain the same or increase in this case[19]. To summarize, while our data limitations prevent us from evaluating the differences in inventory directly, several institutional details suggest that inventory differences are not likely to explain the documented basket variety and composition differences.

[18]While Instacart was not involved in our study, they showed interest in discussing this study's findings.

[19]Our Instacart transactions provide data on the products received by the household. As a result, there is potential for measurement error since the delivered products may not match precisely what the households ordered, particularly in cases where Instacart shoppers had to substitute out-of-stock items. However, we anticipate that any such measurement error would result in underestimating the actual variety, as forced substitutions are likely to increase variety.

D Additional Results

D.1 Variety Results In Table Format

Table D1 complements Figure 5 in the paper and reports the variety results for all 7 specifications that correspond to the point estimates depicted in specification chart in Figure 5. There could be a potential concern that the time-varying changes in Instacart's firm-level advertising strategy, interface, and recommendation system design might be correlated with the observed Instacart basket variety and composition changes. To investigate this potential mechanism, we estimate time-varying basket variety and composition effects for each six-month period in our sample. Though we find some variation in the estimated effects, they are quantitatively and qualitatively similar. Therefore, we conclude that advertising, recommendation systems, and interface design changes are unlikely alternative explanations.

Table D1: Results for the Variety Measure (Classification 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(A) Categories			
β	-0.128***	-0.128***	-0.101***	-0.099***	-0.100***	-0.093***	-0.096***
	(0.007)	(0.007)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)
Num.Obs.	434,322	434,322	434,322	434,322	434,322	434,322	434,322
R^2	0.553	0.553	0.772	0.820	0.810	0.875	0.855
				(B) Items			
β	-0.175***	-0.175***	-0.146***	-0.141***	-0.147***	-0.137***	-0.141***
	(0.009)	(0.009)	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)
Num.Obs.	434,322	434,322	434,322	434,322	434,322	434,322	434,322
R^2	0.518	0.519	0.688	0.688	0.698	0.701	0.619
HH f.e.	yes	yes	no	no	no	no	no
Retailer f.e.	yes	yes	no	no	no	no	no
Month f.e.	no	yes	no	no	no	no	no
Quarter f.e.	yes	no	no	no	no	no	no
HH×*Retailer* f.e.	no	no	yes	yes	no	no	no
HH×*Month* f.e.	no	no	no	yes	no	no	yes
HH×*Quarter* f.e.	no	no	yes	no	no	no	no
HH×*Retailer* × *Quarter* f.e.	no	no	no	no	yes	no	yes
HH×*Retailer* × *Month* f.e.	no	no	no	no	no	yes	no

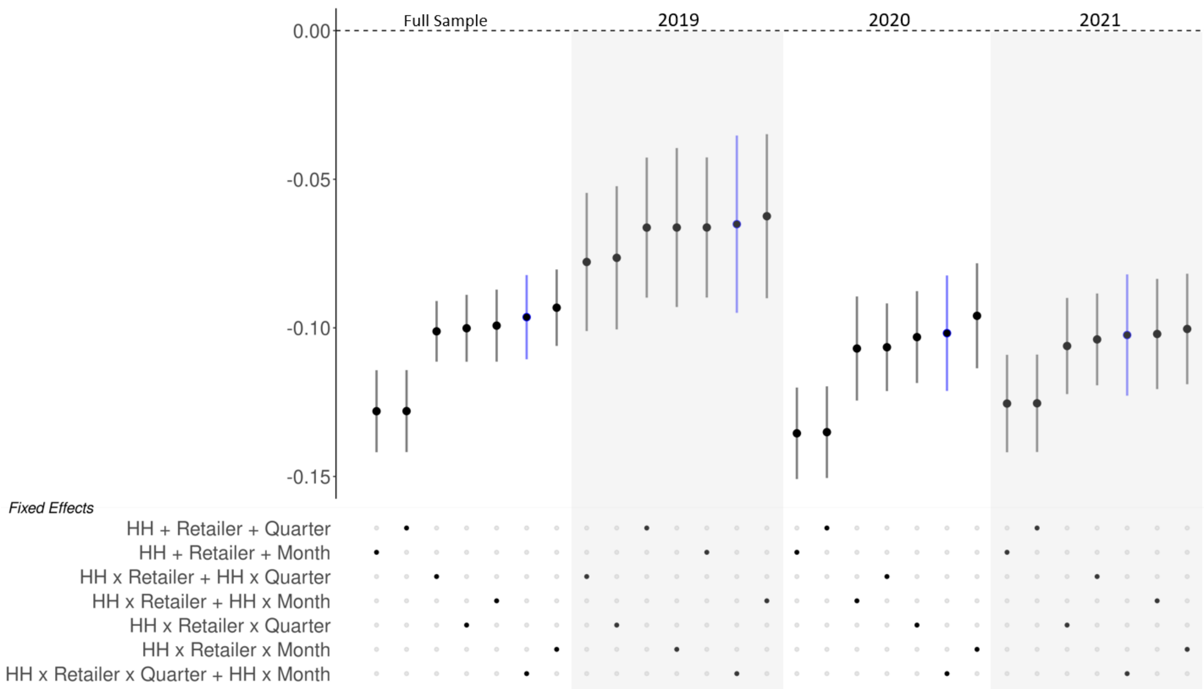
Notes: Outcome variable is *log(Number of Categories)* for Panel A and it is *log(Number of Items)* for Panel B. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.

D.2 Variety Differences Across Time

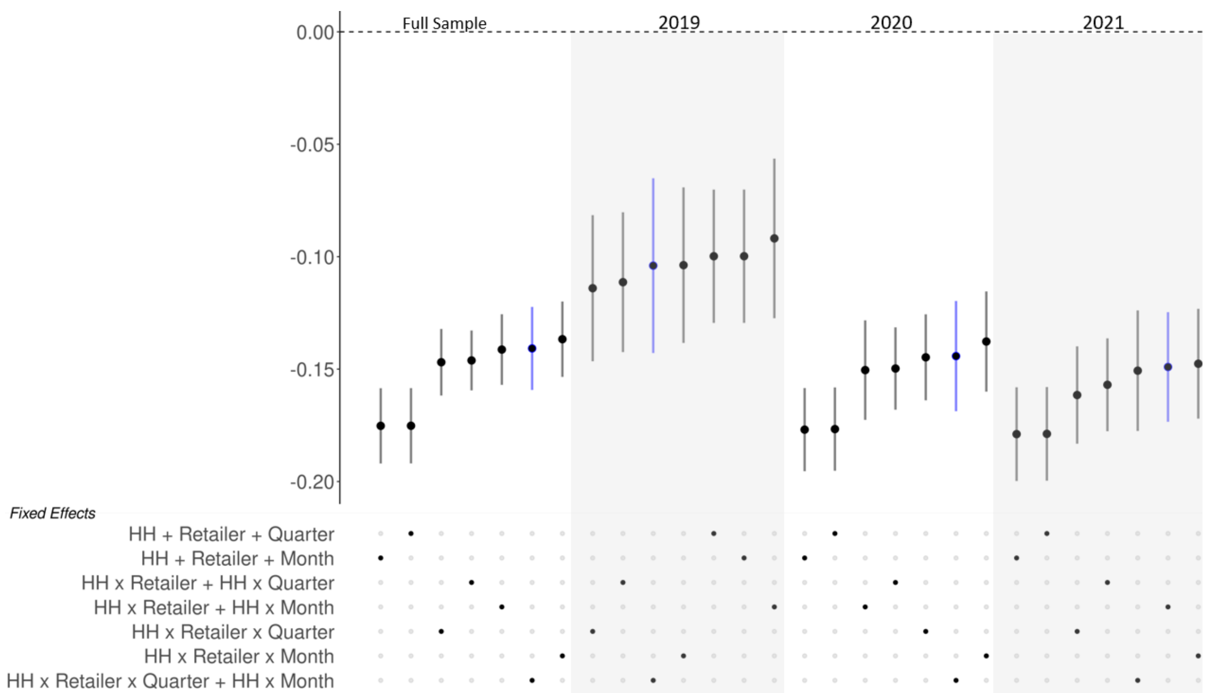
Figure D1 shows the specifications charts for variety differences for the entire sample, and for different years - 2019, 2020, and 2021.

Figure D1: Specification Chart for Variety Differences

(A) Categories



(B) Items



Notes: The figures above present the specification chart for 28 different specifications with different sets of fixed effects. Outcome variable in all specifications is *log(Number of Categories)* for Panel A and *log(Number of Items)* for Panel B. The baseline specifications as specified in Equation 2 are highlighted in blue. Error bars represent 95% CI using clustered standard errors.

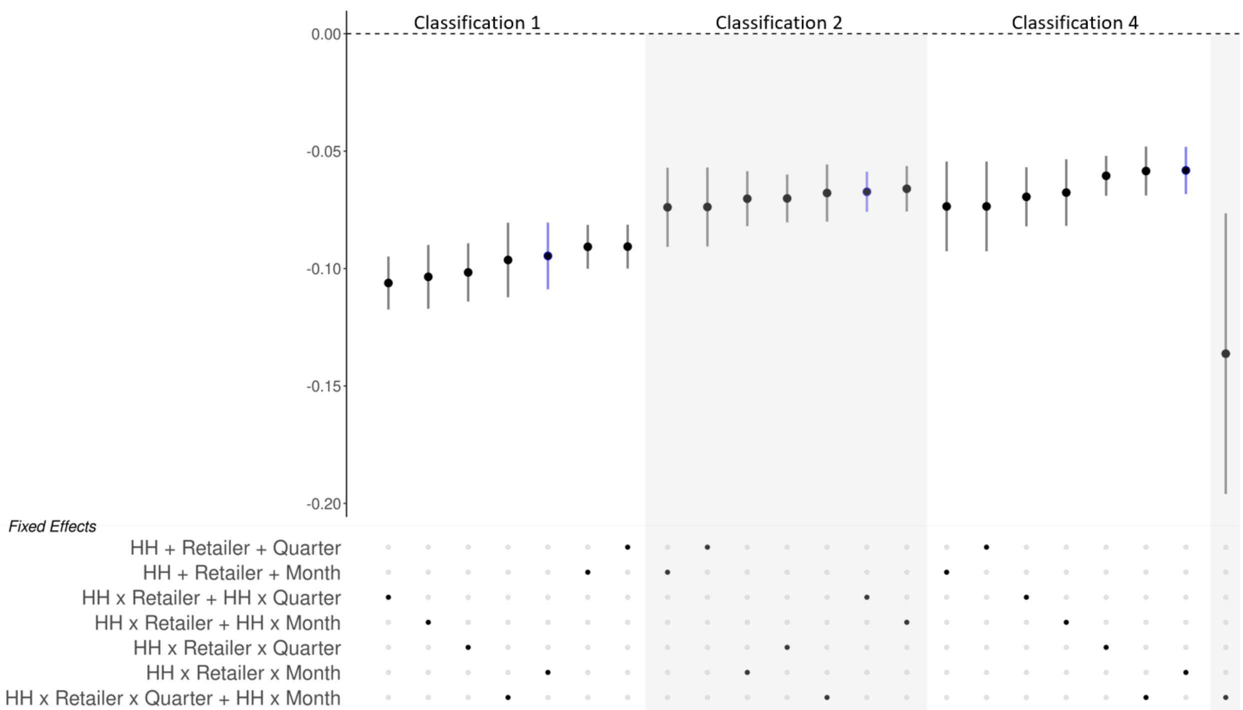
D.3 Variety Differences and Household Heterogeneity

Table D2 reports the results of basket variety by education, race, and income. It shows that Instacart baskets show a significantly lower variety among households with less than a high school education. However, for households with other education levels, the estimates do not depart much from the baseline results. Regarding household race, the impact on variety is lower for Asian households compared to other races, where the estimates remain similar to the baseline. Lastly, the effect on variety remains consistent across different income levels.

D.4 Variety Differences Across Classification Approaches

Figure D2 complements Figure 5 in the paper and provides the specification chart of the variety differences across alternative trip classifications described in section B.3.

Figure D2: Specification chart for Variety Differences across Classification Approaches



Notes: This figure presents specification chart for 22 different specifications with different sets of fixed effects for classifications 1, 2, 4, and 5 (the last point estimate). The outcome variable in all specifications is *log(Number of Categories)*. The baseline specifications with preferred sets of fixed effects as specified in Equation 2 are highlighted in blue. Error bars represent 95% CI using clustered standard errors.

D.5 Poisson Specification

Table D3 compares the baseline specification results, where the dependent variable is *log(number of categories)*, with the specification that uses count data and implements Poisson regression.

Table D2: Variety Differences by Education, Race, and Income

	Estimate
(A) *Education*	
Instacart x High School	-0.100***
	(0.017)
Instacart x College	-0.097***
	(0.009)
Instacart x Advanced	-0.088***
	(0.015)
Instacart x Less than High School	-0.207***
	(0.062)
(B) *Race*	
Instacart x White/Caucasian	-0.099***
	(0.008)
Instacart x Black/African American	-0.100***
	(0.017)
Instacart x Hispanic/Latino	-0.083**
	(0.026)
Instacart x Asian	-0.055
	(0.029)
(C) *Income*	
Instacart x Income less than $50,000	-0.087***
	(0.016)
Instacart x Income $50,000 - $99,999	-0.100***
	(0.011)
Instacart x Income $100,000 - $199,999	-0.094***
	(0.013)
Instacart x Income greater than $200,000	-0.107***
	(0.021)
Num.Obs.	434,322
R^2	0.658

Notes: Panel A reports variety differences by education level, Panel B reports the variety differences by race, and Panel C reports the variety differences by Income. We exclude the race - Other from Panel B as it consists of only 1.7% of the households. Outcome variable is *log(Number of Categories)*. All models use $HH \times Retailer \times Quarter$ and $HH \times Month$ fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.

Table D3: Results for the Variety Measure with Poisson Specification

	Poisson	**Logarithmic**
Instacart	-0.104***	-0.096***
	(0.008)	(0.007)
Num.Obs	434,322	434,322
R^2	0.190	0.658

Notes: Outcome variable for Poisson regression is *Number of Categories*, and for the other regression that replicates the main results, it is *log(Number of Categories)*. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses. Both the estimations use the baseline specifications.

D.6 Basket Variety and Similarity by Experience with Instacart

In this section, we investigate whether the variety differences exhibit variation with a household's experience with Instacart. In particular, we consider a specification that adds an interaction term with the number of Instacart trips (as a proxy to experience with Instacart) to Equation 2. These results are reported in Table D4 and show that as households increase their usage of Instacart, the differences in variety from shopping on Instacart becomes weakly less pronounced.

In Table D5 we formally investigate whether the gap between the similarity of BM and Instacart baskets diminishes. The dependent variable is the relative gap in distances *((BM Distance - Instacart Distance)/BM Distance)*, and the independent variable is the number of Instacart trips. The results show that while the gap diminishes with experience with Instacart, Instacart baskets are significantly more similar to each other relative to BM baskets, even among the most experienced Instacart users.

Table D4: Variety Differences as a Function of a Household's Experience with Instacart

	Baseline	**Interaction**
Instacart (β)	-0.088***	-0.1172***
	(0.006)	(0.019)
Instacart $\times$ Number of Instacart trips		0.0004*
		(0.0001)
Num.Obs	289,811	289,811
R^2	0.649	0.660

Notes: Outcome variable is *log(Number of Categories)*. The sample includes only households with two or more Instacart trips. All models use $HH \times Retailer \times Quarter$ and $HH \times Month$ fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.

Table D5: Difference in Trip Distance Differences as a Function of a Household's Experience with Instacart

	Estimate
Intercept	0.2074***
	(0.0011)
Number of Instacart trips	-0.0010***
	(0.0002)
Num.Obs	1,586
R^2	0.006

Notes: Outcome variable is the ratio of (Average trip distance of BM trips - Average trip distance of Instacart trips) and Average trip distance of BM trips. The sample includes only households with two or more Instacart characteristic trips. *** p<0.001, ** p<0.01, * p<0.05; Standard errors reported in parentheses.

D.7 Stockpiling

To address the question of bulk purchases/stockpiling, we need to determine if the frequency of these types of trips differs between the online and BM channels. To conduct this test, we examined the distribution of the number of items purchased within a specific category across all trips in the sample (not just characteristic ones). We then flagged trips that had more items within any category than the 90th percentile, which proxies for bulk purchases or stockpiling trips.

Our findings in Table D6 show that Instacart trips are 5% more likely to involve bulk purchases compared to BM trips to the same retailer, and when we expand the analysis to include all online trips, we find that all online trips are 4.6% more likely to have bulk purchases. However, when we restrict the analysis to characteristic trips only, we find that Instacart characteristic baskets are *not* more likely to contain bulk purchases than BM characteristic baskets. This result is expected since we find that Instacart and BM characteristic trips are substitutes.

Table D6: Stockpiling trips - Instacart and Online channels

All Trips		**All Trips**		**Characteristic Trips**	
Instacart	0.050***	Online	0.046***	Instacart	-0.000
	(0.002)		(0.002)		(0.001)
Num.Obs.	1,821,213		1,967,578		434,322
R^2	0.162		0.179		0.127

Notes: The outcome variable is a dummy variable for whether a trip is stockpiling trip or not. The sample includes all trips for the first two columns and only characteristic trips for the last column. All models use $HH \times Retailer \times Quarter$ and HH$\times$*Month* fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.

D.8 Evaluating Substitution of BM and Instacart Characteristic Trips

Evaluating the degree of substitution between Instacart and BM characteristic trips may be challenging as it difficult to disentangle the effects of Covid from those attributable solely to Instacart adoption. Specifically, it can be difficult to determine how much of the change in total regular shopping trips after Instacart adoption is due to Covid's impact on overall shopping patterns. To address this potential confounding factor, we implement a Difference-in-Differences estimation.

Our analysis is centered around households that started using Instacart after the pandemic began (treated households), as it enables us to make a distinct before-and-after comparison. To ensure that our findings are not affected by changes in food consumption or shopping frequency behavior during the pandemic, we needed to construct a set of control households. Using a 1:1 nearest neighbor matching procedure, we construct a group of "control Instacart adopters" by pairing pandemic Instacart adopters with households that were never observed to use Instacart. That is, for each pandemic Instacart adopter, we identified a comparable household that was similar in all observable dimensions except for Instacart adoption. Since the Covid-19 based restrictions were usually state-specific, we require for these two households to be located in the same state. We also exactly match on the number of BM characteristic trips in the pre-period. We match the households on the demographic characteristics such as gender, age, marital status, family size, employment, and income level. In addition to these demographic features, we also add matching criteria based on the household's shopping patterns: the number of characteristic trips, the number of characteristic trips post the first lockdown, average number of days between characteristic trips, average spend per characteristic trip, average number of categories purchased in a characteristic trip. Then we created a balanced $HH \times Month$ panel for both Instacart adopters (treatment) and control non-adopters. Note that by construction treatment and control HH are very similar in the pre-period consumption patterns (and we confirm that pre-period frequency of BM characteristic trips is the same across the two cohorts).[20] We then run the following two-way-fixed-effects (TWFE) DiD specification:

$$Y_{ikt} = \delta[I_k \times Post_t] + \gamma_i + \lambda_t + \varepsilon_{ikt} \tag{3}$$

where k denotes the cohort (Treatment or Control), i denotes household, and t denotes month. Y_{ikt} is the outcome variable. We focus on three types of outcomes: (i) count of characteristic BM trips in a month, (ii) count of characteristic Instacart trips in a month, and (iii) a combo of (i) and (ii) – count of overall characteristic trips in a month. I_k is an indicator variable that takes a value of one for the *Treated* cohort and zero otherwise, $Post_t$ is a post-treatment indicator that equals one for months on or after March 2020, γ_i and λ_t

[20] We acknowledge two caveats relating to this analysis. First, we cannot confirm with complete certainty that the households that we observe without any Instacart purchases are the ones that did not use Instacart, as there may be data omission resulting from data permission constraints. Second, despite matching the control and treatment cohorts on all observable characteristics, there may still be unobserved differences that could be correlated with the treatment effect. Despite these limitations, we believe that this exercise is still valuable in determining the degree of plausible substitutability between BM and Instacart trips.

are household, and month fixed effects, respectively.

The results are presented in Table D7. Our findings indicate that, on average, per month, households that adopt Instacart replace 0.152 BM trips with 0.157 Instacart trips, suggesting that characteristic Instacart and BM trips are almost perfect substitutes. This is reinforced by the third column, which reveals that combining Instacart and BM trips results in no change in the total characteristic trips. Note that this supplementary analysis provides further confirmation of our characteristic trip construct's validation, as our goal was to identify trips that would be substitutes for one another.

Table D7: Evaluating Substitution of BM and Instacart Characteristic trips

	Number of BM Characteristic Trips	Number of Instacart Characteristic Trips	Number of All Characteristic Trips
δ	-0.152**	0.157***	0.006
	(0.056)	(0.007)	(0.056)
Household FE	yes	yes	yes
Month FE	yes	yes	yes
Num.Obs.	224,964	224,964	224,964
R^2	0.307	0.230	0.306

Notes: The above regressions estimate δ from equation 1. The outcome variable in the first column is *Number of BM characteristic trips made by a household in a month.* For the second column, the outcome variable is *Number of Instacart characteristic trips made by a household in a month.* For the third column, the outcome variable is *Number of All characteristic trips made by a household in a month.* The *treated* cohort is the group of households that start using Instacart after the pandemic. The *control* cohort is created using a 1:1 nearest neighbor matching procedure. All models use *HH* and *Month* fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.

D.9 Category Compensation Results

In this section, we evaluate the *total* number of items bought in each seven-day period in each of the categories where we observe significant differences in Instacart baskets (see results in Table 7). In other words, we investigate whether or not households buy the same amount of items across *all* of their trips, including the focal Instacart and BM characteristic trips as well as all non-characteristic trips, in certain categories in a week with an Instacart characteristic trip compared to a week with a BM characteristic trip.[21]

The overall category consumption results are reported in the second column of Table D8. For convenience, the first column reproduces the results from Table 7. We find that the total quantity purchased in a week when the focal characteristic trip was conducted via Instacart is less than when it was carried out through the BM channel. We see that the point estimates across the two columns in all but one of the categories are not statistically significantly different from each other. This means that consumers do not compensate for

[21]The majority of the characteristic trips that we observe are once in a 7-day window or less frequently, making this calculation straightforward. However, in some cases, we observe more than one characteristic trip in a seven day window. To avoid the problem with over counting in such cases, we divide the outcome variables in non-focal trips equally among all the characteristic trips in that seven day window. Our results are robust if we restrict our sample to trips where this division is not necessary.

items that they didn't buy via Instacart and that there is a statistically significant difference in the number of items bought during the weeks with Instacart characteristic trips. A similar Table D9 for a higher consumption shows that there is a statistically significant difference in the quantity bought during the weeks with Instacart characteristic trips.

Table D8: Overall change in categories - Lower purchases

Category Name	Characteristic Trips Estimate	Overall Trips Estimate
Fresh Vegetables	-0.136***	-0.130***
	(0.016)	(0.015)
Candy	-0.071***	-0.060***
	(0.007)	(0.009)
Cheese	-0.061***	-0.061***
	(0.011)	(0.011)
Bakery Desserts	-0.059***	-0.058***
	(0.008)	(0.008)
Savory Snacks	-0.047***	-0.044***
	(0.009)	(0.009)
Fresh Meat	-0.043***	-0.040***
	(0.012)	(0.012)
Bread	-0.029***	-0.030***
	(0.008)	(0.008)
Sweet Snacks	-0.028***	-0.032***
	(0.007)	(0.007)
Deli	-0.027***	-0.024**
	(0.007)	(0.007)
Pasta	-0.026***	-0.084***
	(0.005)	(0.024)

Notes: The first column reproduces the results from Table 7. The second column compiles the βs for the same categories as in Table 7, but for total number of items across *all* trips. For each regression, the dependent variable is *log(1+Number of Items)* from that respective category across *all* characteristic and non-characteristic trips in a 7-day window around any given characteristic trip. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in the parentheses.

Table D9: Overall change in categories - Higher purchases

Category Name	Characteristic Trips Estimate	Overall Trips Estimate
Frozen Meat	0.025***	0.025***
	(0.006)	(0.006)
Frozen Vegetables	0.016*	0.017**
	(0.006)	(0.006)
Frozen Breakfast	0.014*	0.013*
	(0.005)	(0.005)
Frozen Fruit	0.008*	0.008**
	(0.002)	(0.002)

Notes: The first column shows the results for characteristic trips. The second column compiles the βs for the same categories but for total number of items across *all* trips. For each regression, the dependent variable is *log(1+Number of Items)* from that respective category across *all* characteristic and non-characteristic trips in a 7-day window around any given characteristic trip. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in the parentheses.

D.10 Decomposing Extensive and Intensive Margins in Category Changes

Here we evaluate the extensive margin—the incidence of observing a characteristic trip with zero fresh vegetables or zero impulse purchases. We create a dummy variable to indicate whether a trip contains zero vegetables or impulse purchases and use it as the outcome variable in our baseline specification. The results reported in Table D10 suggest that Instacart characteristic trips are more likely to contain baskets without fresh vegetables and impulse categories when compared to BM characteristic trips.

Next, we assess the intensive margin by examining a subset of trips with non-zero fresh vegetable and impulse category purchases. Our analysis using the baseline specification shows that Instacart baskets have fewer fresh vegetables compared to BM baskets (Table D10). The result is similar for the impulse categories. Therefore, we conclude that the observed differences in the decline of fresh vegetable and impulse category purchases can be attributed to both the intensive and extensive margins.

We also examine the level of substitution between fresh and frozen items. We combine the quantities of fresh and frozen vegetables purchased and use the baseline specification to estimate the overall change in vegetable consumption. If there is any substitution from fresh vegetables to frozen vegetables, we would anticipate the overall change to be lower than the baseline estimate of 13.6% in fresh vegetable consumption alone. Our results indicate that vegetable consumption is lower by 11.9%. This implies that although there is some weak evidence of substitution between fresh and frozen vegetables, the degree of substitution is small.

Table D10: Extensive and Intensive margins: Vegetables and Impulse Categories

	Fresh Vegetables	**Impulse Categories**
	(A) *Extensive Margin*	
Instacart	0.062***	0.072***
	(0.009)	(0.008)
Num.Obs.	433,304	433,304
R^2	0.123	0.001
	(B) *Intensive Margin*	
Instacart	-0.094***	-0.089***
	(0.018)	(0.011)
Num.Obs.	294,719	302,985
R^2	0.000	0.000

Notes: In Panel A, the dependent variable is whether a trip contains zero fresh vegetables or zero impulse categories. In Panel B, the dependent variable is *log(1+Number of Items)* for vegetables and impulse categories. All models use *HH*×*Retailer*×*Quarter* and *HH*×*Month* fixed effects. *** p<0.001, ** p<0.01, * p<0.05; Clustered standard errors reported in parentheses.