

# UC Irvine

## UC Irvine Previously Published Works

### Title

E-shopping changes and the state of E-grocery shopping in the US - Evidence from national travel and time use surveys

### Permalink

<https://escholarship.org/uc/item/78g4079g>

### Authors

Saphores, Jean-Daniel  
Xu, Lu

### Publication Date

2021-06-01

### DOI

10.1016/j.retrec.2020.100864

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



# E-shopping changes and the state of E-grocery shopping in the US - Evidence from national travel and time use surveys

Jean-Daniel Saphores<sup>a,\*</sup>, Lu Xu<sup>b</sup>

<sup>a</sup> Departments of Civil and Environmental Engineering, Economics, and Urban Planning & Public Policy, Institute of Transportation Studies, University of California, Irvine, CA, 92617, USA

<sup>b</sup> Department of Civil and Environmental Engineering, Institute of Transportation Studies, University of California, Irvine, CA, 92617, USA

## ARTICLE INFO

### JEL classification:

R0 (General)  
R2 (Household Analysis)  
R4 (Transportation Economics)

### Keywords:

e-grocery  
Online shopping  
National household travel survey  
American Time use survey

## ABSTRACT

In spite of the popularity of e-shopping, only 16% of US adults have ordered groceries online, and 7 out of 10 of those who currently buy groceries online do so at most twice a month. Understanding the determinants of e-grocery shopping is important for grocers, supply chain managers, and urban planners. In this context, we first explore how deliveries from online shopping have been changing over time. From our analysis of the 2009 and 2017 National Household Travel Surveys, we found that online shopping has been embraced by increasingly diverse households, although income, education, and some racial/ethnic differences persist. Our analysis of the 2017 American Time Use Survey shows that Americans are 24 times more likely to shop for groceries in stores than online. Moreover, in-store grocery shoppers are more likely to be female and unemployed, but less likely to belong to younger generations, to have less than a college degree, or to be African American. The gender imbalance in grocery shopping is larger online than in stores, but e-grocery shoppers do not otherwise differ from the general population. Future travel and e-shopping surveys (especially for e-grocery) should combine time use and travel questions with retrospective questions about online purchases.

## 1. Introduction

By expanding the range of products available to consumers, stimulating competition, and enhancing shopping convenience, e-commerce is changing the way people shop. Its popularity is growing. According to the Pew Research Center (Smith & Anderson, 2016), four out of five Americans have purchased items online at least once (up from 22% in 2000). Globally, e-commerce is taking an increasing share of total retail sales, rising from 7.4% in 2015 to 11.9% in 2018 (eMarketer, n. d.). These changes have widespread implications for freight and supply chains management (Perboli & Rosano, 2019), travel (Calderwood & Freathy, 2014; Suel & Polak, 2018), the environment (Cherrett et al., 2017; Dost & Maier, 2018), in-store shopping (Farag, Schwanen, Dijst, & Faber, 2007; Lee, Sener, Mokhtarian, & Handy, 2017), and land use planning (Pettersson, Winslott Hiselius, & Koglin, 2018).

The growth of online shopping (e-shopping) is far from homogeneous from a geographic point of view, however. For example, in 2018 online retail sales were approximately 28.6% of total consumer retail

sales in China (InsideRetail Hong Kong, n. d.), versus less than 10% in the United States (US Census Bureau News, 2019). Moreover, e-shopping in a given sector can differ widely even between countries that are culturally and economically similar. Indeed, despite an average annual growth of 18.7% between 2000 and 2016 (US Census Bureau, 2018a), e-commerce sales of food, beer, and wine<sup>1</sup> in the United States represent currently only 0.35% of total food and beverage purchases (US Census Bureau, 2018b). By comparison, online sales made up 5.3% of total food retail sales in the UK (Office of National Statistics, 2018), which outlines the need to study e-grocery in the United States, even though this topic has already received much attention elsewhere, especially in Europe.

In this context, this empirical study has two purposes. Our first purpose is to understand changes between 2009 and 2017 – two years selected because of data availability - in residential deliveries from online shopping in the United States. We focus on residential deliveries because national data on grocery deliveries from online purchases in the United States are not, as far as we know, publicly available. Understanding residential deliveries from e-shopping is clearly important to

\* Corresponding author.

E-mail addresses: [saphores@uci.edu](mailto:saphores@uci.edu) (J.-D. Saphores), [lux15@uci.edu](mailto:lux15@uci.edu) (L. Xu).

<sup>1</sup> There is no separate category for groceries in the Census data on e-commerce sales. Moreover, data from the Census in Historical Tables 4 and 5 do not distinguish between meals delivered to a home/office and foods bought at a grocery store.

logistics managers (so they can supply warehouses serving residences), to transportation engineers (so they can maintain residential roads and adequately update their design), and to transportation planners (so they can mitigate the externalities of changing freight flows and accommodate new delivery options). Better quantifying the traffic and environmental impacts of residential deliveries from online shopping is particularly of concern because soaring uncoordinated deliveries will increase residential traffic congestion, noise, and air pollution, and exacerbate parking shortages in dense urban areas.

Our second purpose is to characterize US households who are shopping online for groceries, which is salient because of the importance and the challenges of grocery retailing. Since we are not aware of any public national dataset on household deliveries of groceries in the US, an indirect way of analyzing e-grocery deliveries is to couple characterizations of e-grocers with an understanding of deliveries from online purchases, as analyzed in the first part of this paper. Although a number of papers have analyzed online shoppers (e.g., see [Brashear, Kashyap, Musante, & Donthu, 2009](#); [Ganesh, Reynolds, Lockett, & Pomirleanu, 2010](#); [Crocco, Eboli, & Mazzulla, 2013](#); [Bressolles, Durrieu, & Senecal, 2014](#); or [Harris, Dall'Olmo Riley, Riley, & Hand, 2017](#), and references herein), much of the recent literature has focused on Europe, and there is a dearth of academic research on e-grocery in the United States. Profiles of online shoppers generated by consultants for grocers can also be found online, but they typically rely on univariate analyses and none of the profiles we found analyzed datasets representative of the US population.

To the best of our knowledge, our paper is the first to analyze recent changes in residential deliveries from online shopping in the United States and to examine online grocery (e-grocery) shopping using publicly available survey data.

Our point of departure is the latest (2017) National Household Travel Survey (NHTS), which we contrast with its previous edition, the 2009 NHTS. We analyze these two national surveys because they asked participating households how many deliveries from online shopping they received in the 30 days preceding their assigned survey day. To explain the number of these deliveries and understand how they changed between the 2009 and the 2017 NHTS, we estimate similar zero-inflated mixture models on 2009 ( $N = 134,371$ )<sup>2</sup> and 2017 ( $N = 123,148$ ) NHTS data, and test differences in model coefficients for these two years.

Since the 2017 NHTS does not ask about e-grocery shopping, we also analyze data from the 2017 American Time Use Survey (ATUS) to contrast socio-economic characteristics of people who engage in online grocery shopping with conventional grocery shoppers. To do so, we first estimate logit models on a subset ( $N = 2934$ ) of the 2017 ATUS to consider only households likely to have had access to e-grocery shopping in 2017. Since the number of people who shopped for groceries online in the 2017 ATUS is small, we compare the distributions of selected socio-economic characteristics of e-grocery shoppers with those of conventional grocery shoppers using Kruskal-Wallis tests. Finally, we contrast the distributions of shopping start times (i.e., when a customer starts browsing to buy groceries) between these two groups, since proponents of e-grocery shopping highlight the convenience of shopping at any time.

In the next section, we review how e-grocery shopping first emerged, assess what obstacles led to early failures, and explain how they were (at least partly) overcome. We then briefly discuss characteristics of e-shoppers in the United States, and review some potential impacts of e-shopping and e-grocery. In Sections 3 and 4, we respectively present our data and introduce our modeling approach. In Section 5 we discuss our results. In Section 6 we summarize our findings, discuss potential impacts on local deliveries, outline some limitations of our study, and

propose some avenues for future work.

## 2. Background and literature review

### 2.1. The first coming of e-grocery in the US

Alternative channels to traditional retail (e.g., mail order) predate by decades the arrival of online shopping but their market share has always been small, especially for groceries ([White, 1997](#)). The emergence of the internet was a game changer because it considerably expanded consumer choice, sped up deliveries, and provided new ways to learn about products ([Seaman, 1995](#)).

The first big push to develop online grocery shopping took place in the late 1990s when consumers started buying products over the internet. Grocery was an early target because it is the single largest retail sector, most consumers shop for groceries frequently, and many do not particularly enjoy it ([Saunders, 2018](#)). Proponents argued that e-grocery would give consumers the freedom to shop at the time and from the place of their choosing, expand consumer choice, and stimulate competition by facilitating price comparisons. Convenience and time saving were seen as major advantages at a time when women's participation in the labor force was rising ([Morganosky & Cude, 2000](#)).

E-grocery pioneers were technology companies eager to leverage their knowledge of information technologies to take over what they saw as an underachieving sector ([Saunders, 2018](#)). Companies like Webvan (founded in 1996) and [HomeGrocer.com](#) (started in 1997) rode the [dot.com](#) bubble. They attracted large investments to buy warehouses, delivery vans, and marketing campaigns, and built from scratch purely online businesses. However, when consumer demand failed to meet expectations, investments dried up with the burst of the [dot.com](#) bubble and they went bankrupt (Webvan purchased [HomeGrocer.com](#) in September 2000, and it filed for bankruptcy in July 2001) ([Grunert & Ramus, 2005](#); [Saunders, 2018](#)). Partly as a result of these failures, online grocery shopping has been called "the Bermuda Triangle of e-commerce" - a place where investments vanish without leaving a trace ([McDonald, Christensen, Yang, & Hollingsworth, 2014](#)).

What went wrong? US e-grocery pioneers overlooked several key characteristics of the grocery sector, and underestimated the magnitude of the change they wanted to introduce.

First, e-shopping implies that a number of tasks previously undertaken by customers, including picking, packing, and delivering goods, are taken over by the retailer. This adds to the costs of retailers and squeezes their already thin profits.

Second, since groceries are quite diverse and some are perishable, they require more complex logistics ([Murphy, 2003](#)). Moreover, delivering to a customer's residence raises new issues. Indeed, if no one is present to receive an order, coming back at a different time is costly. Conversely, if an order is left on a buyer's doorstep, delivered goods could spoil or be stolen. The problem is especially acute for prepared foods, whose temperature, texture, taste, and appearance can quickly change over time.

Third, early e-grocers did not appreciate the difference between buying groceries online and shopping for groceries in a conventional store ([Robinson, Dall'Olmo, Rettie, & Rolls-Willson, 2007](#)). In particular, sensory information (e.g. smell and touch) and interpersonal interactions are lacking online ([Hansen, 2005](#)). This is not an issue for search goods (whose characteristics are easily evaluated before purchase), but it matters for experience goods (which can only be evaluated after a purchase) ([Nelson, 1970](#)). In brick-and-mortar stores, fresh produce can be touched and smelled so they are search goods, but for online shoppers they become experience goods ([Weathers, Sharma, & Wood, 2007](#)).

In spite of its failure, the first wave of e-groceries caught the attention of traditional grocers as online shopping revolutionized shopping in other sectors, such as books or electronics.

<sup>2</sup> In this paper, ( $N =$  number) refers to the size of the sample on which specific models were estimated.

## 2.2. The second coming of e-grocery in the US

As the first wave of internet-only grocers were closing, traditional grocers started experimenting with on-line shopping. Moreover, new start-ups began emerging with innovative solutions to address some of the shortcomings of earlier e-grocery shopping models.

In addition to attractive and easily navigable websites (Freeman & Freeman, 2011), one key to success in e-grocery shopping is low operational costs in order to offer competitive prices and effective delivery services (Anckar, Walden, & Jelassi, 2002; Kamarainen, Smaros, Jaakola, & Holmstrom, 2001). This lesson was learned by both traditional grocers and mega retailers such as Walmart or Target, and Amazon, which acquired Whole Foods in 2017 to boost its physical presence.

To keep costs down, some grocers partnered with start-ups that provide a platform to customers who order from the grocers' websites and have their employees pick, pack, and deliver orders in exchange for payments from both grocers and shoppers (they also make money from customer information). The largest of these start-ups is Instacart, created in 2012 (Lien, 2017). By the end of 2018, Instacart had partnerships with over 300 retailers operating over 15,000 grocery stores. A number of other start-ups have been offering similar services, including Deliv, DoorDash, Postmates, or Shipt (acquired by Target at the end of 2017).

To avoid unsecured deliveries, e-grocers have experimented with different alternatives: 1) Click and pick, where consumers order online but pick up at a store or a warehouse; 2) Bring to a local storage area and deliver when customers are home; 3) Allow the delivery person to leave purchases inside a customer's home. This service is offered by smart lock maker August with delivery partner Deliv and several retailers (Macy's, Best Buy, Bloomingdale's and PetSmart). It is also offered by Amazon for its Prime customers who live in selected cities, and subscribe to Amazon Key. Amazon Key requires buying an Amazon cloud cam and installing a compatible smart lock at home (Wollerton, 2018); and 4) Deliver an order to the customer's car trunk, if he/she is an Amazon Prime subscriber, has an active connected car service plan, drives a GM or a Volvo vehicle from 2015 or newer, and lives in one of 37 US cities (Hawkins, 2018).

Leading e-grocery retailers in 2017 include Wal-Mart, Costco, Sears (which filed for bankruptcy in 2018), Amazon, Kmart, but also Kroger (the largest overall grocer in the US), and Safeway (an Albertsons brand). However, this sector has been changing quickly with Amazon's purchase of Whole Foods and Target's acquisition of Shipt (both in 2017), for example. While in 2017 just over 30% of grocery stores in the US offered home delivery/store pickup of online orders, this percentage had jumped to over 52% by 2019 (Conway, 2020).

## 2.3. E-shoppers and e-groceries

To inform our choices of explanatory variables, we also reviewed paper characterizing people who engage in e-grocery shopping. Early studies reported that online grocery shoppers are typically younger, better educated, and tend to have higher incomes than the general population (Morganosky & Cude, 2000, 2002). They are also more likely to be female because women are typically more involved in grocery shopping (Morganosky & Cude, 2000, 2002). Other studies reported that some seniors and some disabled individuals also shop online for groceries (Anckar et al., 2002; White, 1997).

As expected, convenience is a driving force behind e-groceries, but situational factors (such as a recent baby or a deteriorating health) also matter (Hand, Riley, Harris, Singh, & Rettie, 2009). People comfortable navigating the internet are not necessarily online shoppers, however, and when they shop online, they do not usually discontinue offline shopping (Hand et al., 2009). Furthermore, Kang, Moon, Kim, and Choe (2016) showed that the impact of convenience depends on experience with e-shopping and with the type of product considered. Moreover, although the time requirement to access offline grocery markets has no effect on the adoption of online grocery shopping, it may affect the

amount of groceries purchased online.

A number of papers have inquired about the determinants of consumers' channel choice (e.g., see Melis, Campo, Breugelmanns, & Lamey, 2015; or Wang, Malthouse, & Krishnamurthi, 2015). Melis et al. (2015) found that when consumers start buying groceries online, they tend to select the online store from their preferred offline stores; moreover, the offline environment is important when customers are new to online shopping, although it matters increasingly less as they gain more experience with on-line shopping. The device used for e-grocery shopping also seems to matter: according to Wang et al. (2015), m-shopping (i.e., shopping via smartphones or tablets) increases the rate of orders, especially for low-spending customers.

In spite of high growth rates and enthusiasm for online grocery shopping, a recent Gallup survey (Redman, 2018) showed that 84% of US adults have never ordered groceries online, and that 7 out of 10 of those who buy groceries online do so twice a month or less.

## 2.4. Impacts of e-shopping/e-groceries on land use, retailers, and supply chain planning

As e-shopping and e-grocery become more common in the US, they may have multiple impacts. Here, we briefly consider impacts on land use, retailers, and supply chain planning.

As they become increasingly affordable and ubiquitous, information and communication technologies (ICT) are decoupling activities such as work or shopping from specific times and spaces (Kwan, 2007). By decreasing the cost of exchanging information, ICT may weaken agglomeration forces and promote the emergence of decentralized, smaller urban centers (Ioannides, Overman, Rossi-Hansberg, & Schmidheiny, 2008). However, concrete evidence that internet use (and in particular e-shopping) has impacted urban structure is still lacking (Ioannides et al., 2008), possibly because it takes years to substantially change the structure of an urban area, but also because of the complexity of the changes induced by ICT, and more particularly by e-shopping (Nahiduzzaman, Aldosary, & Mohammed, 2019). While we can expect to see a shift in demand from retail space to warehousing or other types of storage space, the magnitude of that shift is still uncertain. It may be amplified, however, with the widespread adoption of self-driving technologies, which are expected to substantially cut the cost of freight transportation (Andersson & Ivehammar, 2019; Wadud, 2017). This shift may be less important for e-grocery than for e-shopping in general. It will also depend on the dominant form of grocery (e.g., click-and-pick vs. home deliveries) and on the extent to which grocers adopt omni-channel strategies, where the business processes of multiple retail channels are increasingly integrated (Marchet, Melacini, Perotti, Rasini, & Tappia, 2018).

Although online shopping is often invoked to explain high-profile bankruptcies among US retailers (e.g., Sears, Sport Authority, Payless) and store closures by major retailers such as J.C. Penney and Macy's that took place over the last decade, other factors may have contributed just as much to retail store closures in the US, including an excessive number of malls and shifts in consumers' spending habits (Thompson, 2017). Grocery stores have also been affected as they are facing increasing competition from Walmart, Aldi, and Amazon (Meyerson, 2019), but casualties so far have only been small and regional firms that were out of sync with the markets, or could not afford costly investment to expand online (Meyerson, 2020).

In addition to the obstacles associated with e-shopping in general, such as risks associated with the safety of internet connections and the payment system, or the lack of complete information about online orders (Wat, Ngai, & Cheng, 2005), e-grocers need to adopt efficient home delivery solutions that can accommodate the requirements of groceries (Punakivi & Saranen, 2001). This entails ensuring delivery during tight time windows while observing adequate temperature requirements, the ability to promptly respond to demand, and having enough information to avoid failed home deliveries due to a customer's absence (Punakivi &

Saranen, 2001), all of which are issues that delivery service companies or the postal service typically do not face. A number of options have been proposed to manage the demand for grocery deliveries based either on time slot allocations or time slots pricing, either in static (i.e., forecast-based) or dynamic (i.e., real-time) settings (Agatz, Campbell, Fleischmann, Van Nunen, & Savelsbergh, 2013; Klein, Neugebauer, Ratkovitch, and Steinhardt (2019).

More generally, brick-and-mortar retailers aiming to be competitive online need to redefine their logistics networks (Wollenburg, Hübner, Kuhn, & Trautrim, 2018). This involved adapting inventory management, distribution settings (i.e., the number and type of logistics facilities handling online orders), fulfilment strategy, deliveries, and return management policies (Marchet et al., 2018).

### 3. Data

As explained in the introduction, this study relied on data from several publicly available datasets. First, to understand how deliveries from online shopping have been changing over the last few years, we analyzed data from both the 2009 and the 2017 National Household Travel Surveys (NHTS) (FHWA, 2010, 2018). These surveys provide comprehensive national data on households, their members, their vehicles, and daily travel for all purposes and by all modes of transportation. Second, since (to the best of our knowledge) there is no publicly available national dataset on e-grocery in the US, we analyzed data from the 2017 American Time Use Survey (ATUS) to characterize US consumers who shop online for groceries. These datasets and our variable choices are presented in turn.

#### 3.1. Data from the 2009 and 2017 NHTS

Compared to the 2009 NHTS, data for the 2017 NHTS were collected using a new sampling strategy and a new methodology (e.g., data were retrieved using a self-completed web-based survey instead of via an interviewer assisted phone survey) to lower the burden on respondents and improve coverage. In 2009, the sample frame was obtained using random digit dialing of landline phone numbers, but this approach is no longer appropriate because since 2016 over half of American homes have abandoned landlines in favor of cell phones (Blumberg & Luke, 2017). Approximately 45% of households who participated in the 2017 NHTS have no land line, and many include ethnic minorities and younger people (McGuckin & Fucci, 2018).

Overall, the 2017 NHTS collected data from 129,969 households who made 923,572 trips between April 2016 and April 2017. Conversely, the 2009 NHTS collected data from 150,147 households who made 1,167,321 trips between March 2008 and April 2009 (FHWA, 2011).

Both the 2009 and the 2017 NHTS inquired about the number of times each respondent purchased something online and had it delivered in the 30 days preceding their survey day. We note, however, that the 2009 NHTS question specifies deliveries to home, while that 2017 NHTS question we analyze does not. We aggregated individual answers to this question by household to create the dependent variable for our models that explain deliveries from online shopping. We focused on households here because it is not uncommon for one household member to order goods for other household members, especially if they are children.

Both surveys also asked how frequently respondents use the internet. We used this information to exclude households who stated they never use the internet but receive deliveries from online purchases.

To explain deliveries from online purchases, we relied on a wide range of directly observable household characteristics and on land use variables available in both the 2009 and the 2017 NHTS. First, our models include variables describing household composition (see Table 1), and a count of the number of household women over 18, as women are often in charge of shopping. Since the age of household members likely matters, we kept track of the number of children

(household members under 18), and of the number of household members 18 or older from each generation, as defined by the Pew Research Center (2018).

Because income categories differ between the 2009 and the 2017 NHTS, we created approximate quintiles (20% strata). We also added the number of household adults who work part-time and full-time since working decreases the time available for activities such as shopping. To capture the highest educational achievement in the household, we relied on the five common categories shown in Table 1.

In addition, we included in our models race variables, a Hispanic/Latino indicator, and we counted the number of household members born abroad. The purpose of these variables is to help capture cultural differences or uneven access to online shopping.

People with a medical condition that impairs their mobility would likely benefit from online shopping, so we created an indicator variable for them. Home ownership (a wealth indicator) may also play a role here, and so could the number of vehicles per adult.

Finally, we created two population density indicator variables (<300 people per square mile and  $\geq 7000$  people per square mile) and an indicator variable for heavy rail. The low-density variable could reflect that people in rural areas have fewer shopping options and may therefore benefit more from online shopping. Conversely, the high-density indicator (and the heavy rail variable) could capture impacts from enhanced online shopping and delivery options.

After removing households with missing data and with over 2 deliveries per day because the latter are unusual and influential observations (39 and 117 observations for the 2009 and 2017 NHTS respectively), our final samples have 134,371 households for 2009 and 123,148 for 2017. Summary statistics for both are provided in Table 1.

#### 3.2. Data from the 2017 ATUS

To get a profile of US consumers who shop for groceries online, we analyzed data from the 2017 ATUS. The main goal of this survey is to understand how noninstitutionalized US residents who are civilians 15 or older allocate their time. ATUS samples individuals who participated in the Current Population Survey. It is conducted annually by the US Census Bureau by phone, using Computer Assisted Interview software.

Although grocery shopping is a household activity, our basic unit of analysis here is the individual because only one person per household participates in ATUS. For each person in our sample, we gathered a broad range of socio-economic variables. Individual characteristics include marital status, gender, generation (based on Pew Research Center definitions), work status, education level, race, Hispanic/Latino status, and presence of a mobility impairment. Household characteristics include the number of children and annual income.

We considered two dependent variables obtained by combining activity (grocery shopping) and location data. The first dependent variable indicates whether or not a respondent shopped for groceries in a store during their ATUS survey day. The second dependent variable indicates whether or not a respondent shopped for groceries online (i.e., when the respondent shopped for groceries, he/she was neither in a grocery store nor in another store/mall). Out of 10,223 persons in our sample, 1420 (13.9%) shopped for groceries in a store, but only 59 (0.57%) shopped for groceries online. This low percentage is not surprising since by August 2018, of the 16% of US adults who had ever ordered groceries online, 7 out of 10 did so twice a month or less (Redman, 2018). Interestingly, nobody in the ATUS dataset shopped for groceries both in a store and online on their survey day.



**Table 1**  
Summary statistics for data used to explain deliveries from online shopping.

Variable	2017 NHTS (N = 123,148)				2009 NHTS (N = 134,371)			
	Min	Mean	Max	Std. Dev.	Min	Mean	Max	Std. Dev.
Number of deliveries to the household from online shopping in the past month	0	4.709	60	6.700	0	2.105	60	4.194
<i>Household composition</i>								
1 adult without children	0	0.181	1	0.385	0	0.102	1	0.302
2+ adults without children	0	0.214	1	0.410	0	0.213	1	0.410
1 adult with children	0	0.035	1	0.184	0	0.026	1	0.160
1 adult, retired, without children	0	0.140	1	0.348	0	0.130	1	0.336
2+ adults, retired, without children	0	0.242	1	0.428	0	0.276	1	0.447
Count of household members younger than 18	0	0.355	8	0.816	0	0.467	11	0.931
Count of women 18 and over in the household	0	0.948	8	0.511	0	0.956	6	0.476
<i>Age structure (# of household members ≥ 18)</i>								
Generation Z (born after 1997)	0	0.043	6	0.224				
Generation Y (born 1981 to 1996)	0	0.329	8	0.661	0	0.125	5	0.400
Generation X (born 1965 to 1980)	0	0.392	5	0.682	0	0.324	4	0.643
Baby Boomers (born 1946 to 1964)	0	0.734	7	0.809	0	0.730	4	0.815
Silent generation (born before 1946)	0	0.282	4	0.578	0	0.585	5	0.761
<i>Annual household income</i>								
First quintile	0	0.192	1	0.394	0	0.214	1	0.410
Second quintile	0	0.221	1	0.415	0	0.175	1	0.380
Fourth quintile	0	0.236	1	0.424	0	0.208	1	0.406
Fifth quintile	0	0.172	1	0.378	0	0.200	1	0.400
<i>Count of household members who:</i>								
Work full-time	0	0.756	6	0.797	0	0.736	5	0.784
Work part-time	0	0.208	6	0.460	0	0.222	5	0.471
<i>Highest education achieved in the household</i>								
Less than high school	0	0.020	1	0.138	0	0.049	1	0.216
High school graduate or GED	0	0.133	1	0.340	0	0.210	1	0.408
Bachelor's degree	0	0.263	1	0.440	0	0.237	1	0.425
Graduate or professional degree	0	0.302	1	0.459	0	0.221	1	0.415
<i>Ethnicity of the household head</i>								
Black or African American	0	0.076	1	0.265	0	0.062	1	0.241
Asian	0	0.037	1	0.190	0	0.020	1	0.138
More than one ethnicity	0	0.028	1	0.165	0	0.006	1	0.077
Other (single ethnicity)	0	0.024	1	0.153	0	0.046	1	0.209
Household head is Hispanic/Latino	0	0.069	1	0.253	0	0.067	1	0.250
Count of household members born abroad	0	0.196	9	0.583	0	0.169	7	0.513
Count of household members with a medical condition that makes it difficult to travel outside of home	0	0.191	7	0.448	0	0.218	6	0.463
Household owns dwelling	0	0.757	1	0.429	0	0.873	1	0.333
Number of vehicles per household adult	0	1.138	12	0.630	0	1.111	27	0.571
<i>Land use</i>								
Household resides in lower density census tract	0	0.325	1	0.468	0	0.345	1	0.475
Household resides in higher density census tract	0	0.257	1	0.437	0	0.233	1	0.423
Household MSA has heavy rail	0	0.154	1	0.361	0	0.170	1	0.376

Notes.

- 1) The dependent variable (deliver) was truncated at 60.
- 2) MSA stands for Metropolitan Statistical Area.
- 3) Lower density census tracts have less than 300 people per square mile. Higher density census tracts have over 7000 people per square mile.

Since e-grocery shopping was not available everywhere in the United States in 2017, we pared down our sample by keeping only respondents who reside in core-based statistical areas<sup>3</sup> where at least one ATUS respondent shopped for groceries online. This reduced our ATUS sample to 2934 respondents. This sample includes only people who are African American, Asian, or White, and it does not include people with a mobility impairment. As a result, other ethnic variables and the variable indicating the presence of a mobility impairment could not be included in the logit models estimated on that sample.

Summary statistics for our two ATUS samples are presented in Table 2.

<sup>3</sup> A core-based statistical area (CBSA) is a US geographic area that consists of one or more counties (or equivalents) anchored by an urban center of at least 10,000 people plus adjacent counties that are socioeconomically tied to that urban center by commuting. Source: <https://catalog.data.gov/dataset/core-based-statistical-areas-national>.

4. Models

In this section, we first describe the mixture models we relied on to explain changes between 2009 and 2017 in residential deliveries from online shopping in the United States. We then explain our strategy to characterize Americans who shop at brick-and-mortar grocery stores versus Americans who engage in e-grocery. Combining results from both analyses could allow identifying areas with a high potential for the delivery of groceries from online shopping.

4.1. Number of deliveries from online shopping (2017 and 2009 NHTS)

We first explained the number of deliveries from online shopping by household in the last 30 days preceding their assigned survey day, and how it changed between 2009 and 2017. In 2009, 57.8% of households had no deliveries from Internet shopping during the 30 days preceding their survey day; in 2017, this percentage fell to 31.3%. To account for this relatively high percentage of zeros, we estimated mixture models (Greene, 2011), which assume that our samples are composed of two distinct groups of households. Households in the first group never purchase goods online, and therefore get no deliveries from online

**Table 2**  
Summary statistics for ATUS analysis.

Variable	N = 10,223			N = 2934		
	Min	Mean	Max	Min	Mean	Max
<b>Dependent variables</b>						
Conventional grocery shopping indicator	0	0.148	1	0	0.150	1
E-grocery shopping indicator	0	0.006	1	0	0.017	1
<b>Explanatory variables</b>						
<i>Marital status (baseline: married)</i>						
Never married	0	0.240	1	0	0.266	1
No Spouse	0	0.279	1	0	0.243	1
Gender (1 if female)	0	0.546	1	0	0.545	1
<i>Age (baseline: Baby Boomer)</i>						
Generation Z (born after 1997)	0	0.053	1	0	0.057	1
Generation Y (born 1981 to 1996)	0	0.221	1	0	0.222	1
Generation X (born 1965 to 1980)	0	0.272	1	0	0.298	1
Silent generation (born < 1946)	0	0.132	1	0	0.113	1
Number of household children	0	0.745	11	0	0.763	11
<i>Annual household income (baseline: \$60,000 to \$99,999)</i>						
Less than \$30,000	0	0.268	1	0	0.223	1
\$30,000 to \$59,000	0	0.265	1	0	0.258	1
Over \$100,000	0	0.244	1	0	0.306	1
<i>Work status (baseline: full-time job)</i>						
Part-time job	0	0.125	1	0	0.131	1
Not employed	0	0.393	1	0	0.363	1
<i>Education (baseline: college degree)</i>						
Less than high school	0	0.122	1	0	0.123	1
High school	0	0.235	1	0	0.207	1
Some college/associate degree	0	0.271	1	0	0.241	1
Graduate/professional degree	0	0.146	1	0	0.172	1
<i>Ethnicity (baseline: White)</i>						
African American	0	0.147	1	0	0.176	1
Asian	0	0.041	1	0	0.071	1
Other	0	0.021	1	NA	NA	NA
Hispanic/Latino	0	0.155	1	0	0.208	1
Mobility impairment	0	0.037	1	NA	NA	NA

Notes: 1) All variables are binary or count variables so we did not present standard deviations.

2) The smaller sample (N = 2934) was obtained by keeping only ATUS respondents from core-based statistical areas (CBSA) where at least one other ATUS respondent shopped for groceries online. We removed respondents whose ethnicity is “other” (neither African American, Asian, nor White) and who had a “mobility impairment” because none of them shopped for groceries online so the impact of these characteristics on e-grocery shopping could not be estimated.

shopping. Households in the second group make online purchases from time to time, so the number of their deliveries is modeled via a count model. To model the probability of belonging to either group, we relied on a logit model. The resulting mixture model can be written (Greene, 2011):

$$Pr(N_i = n_i | x_i, z_i) = \begin{cases} \phi(\gamma' z_i) + \{1 - \phi(\gamma' z_i)\}g(0|\beta' x_i), & \text{if } n_i = 0 \\ \{1 - \phi(\gamma' z_i)\}g(n_i|\beta' x_i), & \text{if } n_i > 0 \end{cases} \quad (1)$$

where:

- $N_i$  designates the random variable that generated  $n_i$ , which is the number of packages from online shopping received by household  $i$  in the 30 days up to their survey day;
- $\phi(\gamma' z_i)$  is the probability that household  $i$  belongs to the group that never makes online purchases (the first group);
- $g(n_i|\beta' x_i)$  is the probability that household  $i$  received  $n_i \geq 0$  packages from online purchases over the 30 days up to their survey day;
- $\gamma$  and  $\beta$  are vectors of unknown coefficients that need to be estimated; and
- $z_i$  and  $x_i$  are vectors of explanatory variables, respectively for the logit model that identifies households who never buy goods online (the first group) and for the count model.

If we model counts with a Poisson process, we obtain a zero-inflated Poisson (ZIP) mixture model; using a negative binomial regression model instead gives a zero-inflated negative binomial (ZINB) mixture model (Long, 1997).

To avoid multicollinearity, we checked that the variance inflation factors for all our explanatory variables are below 10 (they are).

To assess if the 2017 estimate of the coefficient for an explanatory variable ( $\hat{\beta}_{2017}$ ) differs from its 2009 value ( $\hat{\beta}_{2009}$ ), we relied on the test statistic:

$$Z = \frac{\hat{\beta}_{2017} - \hat{\beta}_{2009}}{\sqrt{SE^2(\hat{\beta}_{2017}) + SE^2(\hat{\beta}_{2009})}} \quad (2)$$

Under the null hypothesis  $H_0: \beta_{2017} = \beta_{2009}$ , Z has approximately a standard normal distribution as the difference of two (asymptotically) independent normal random variables, assuming that the 2009 and the 2017 NHTS samples are independent (Greene, 2011).

#### 4.2. Characteristics of US e-grocery shoppers (2017 ATUS)

To obtain a baseline profile of Americans who shop at brick-and-mortar grocery stores, we first estimated a logit model (Greene, 2011) on our full ATUS sample (N = 10,223).

We then estimated two more logit models on the sub-sample (N = 2934) of ATUS respondents who live in core-based statistical areas where at least one person shopped for groceries online. As mentioned above, we constructed this sub-sample because e-grocery shopping was not available everywhere in the United States in 2017. The first logit model estimated on our ATUS sub-sample again characterizes ATUS respondents who shopped at a physical grocery store. It allows us to check that grocery shoppers in this sub-sample do not differ substantially from those in the full ATUS dataset. The second logit model characterizes ATUS respondents who shopped for groceries online.

Since only 59 people shopped for groceries online in the 2017 ATUS, it is difficult to fully capture the determinants of online grocery shopping. We therefore also compared the distributions of selected socio-economic characteristics of people who shopped for groceries online with those who shopped in stores using Kruskal-Wallis (KW) tests (Conover, 1999). A KW test assesses whether different samples originate from the same distribution. Finally, we analyzed the distribution of shopping start times (the self-reported time when ATUS respondents started browsing for groceries online) since one argument for e-grocery is the convenience afforded by the ability to shop at any time.

#### 4.3. Interpreting results

As discussed above, we estimated ZINB (a mixture consisting of a negative binomial regression model with a logit) and logit models. For a negative binomial regression model, the coefficient of an explanatory variable represents the difference of the logs of expected counts when that explanatory variable is increased by one unit, holding constant all other explanatory variables (Long, 1997). In a logit model, the coefficient of an explanatory variable represents the change in the logit (the log of the probability of 1 divided by the probability of 0) of the probability associated with a unit change in that explanatory variable holding all other predictors constant.

### 5. Results

Our results were obtained with Stata 15.1. They are presented in Tables 3 and 4, and illustrated on Figs. 1–3. In Table 3, shaded numbers indicate when the 2017 value of a coefficient differs from its 2009 value. To better link our explanations with results presented in Tables 3 and 4, we occasionally report estimated coefficients and their statistical significance (see notes below these tables).

**Table 3**  
Results for Zero-Inflated Negative Binomial models (2009 & 2017 NHTS) explaining deliveries from online shopping.

Variable	Count sub-models (ZINB) (number of deliveries from online shopping)		Logit sub-models (households who never buy goods online)	
	2017	2009	2017	2009
<i>Household composition (baseline: 2 + adults, not retired, with children)</i>				
1 adult without children	-0.368***	-0.069	0.304**	0.304***
2+ adults without children	-0.012	0.035	0.157	0.259***
1 adult with children	-0.204***	-0.094*	-0.14	-0.223*
1 adult, retired, without children	-0.465***	-0.084	0.506***	0.677***
2+ adults, retired, without children	-0.006	-0.002	-0.305**	0.041
Count of household members younger than 18	0.039***	0.032***	0.120***	-0.049
Count of women 18 and over in the household	0.051***	-0.005	-0.427***	-0.167***
<i>Age structure (count of household members ≥ 18)</i>				
Generation Z (born after 1997)	0.154***		-0.185*	
Generation Y (born 1981 to 1996)	0.239***	0.268***	-0.340***	-0.820***
Generation X (born 1965 to 1980)	0.155***	0.338***	-0.006	-0.792***
Baby Boomers (born 1946 to 1964)	0.033**	0.268***	0.371***	-0.544***
Silent generation (born before 1946)	-0.148***	0.133***	1.028***	0.092*
<i>Annual household income (baseline: third quintile)</i>				
First quintile	-0.189***	-0.112***	1.391***	1.083***
Second quintile	-0.124***	-0.091***	0.622***	0.506***
Fourth quintile	0.155***	0.116***	-0.586***	-0.420***
Fifth quintile	0.407***	0.330***	-1.721***	-1.084***
<i>Work Status. Count of household workers who:</i>				
Work full-time	-0.005	-0.049***	-0.211***	-0.147***
Work part-time	0.011	-0.022	-0.361***	-0.275***
<i>Highest education achieved in the household (baseline: some college/associate degree)</i>				
Less than high school	-0.304***	-0.108	1.544***	1.923***
High school graduate or GED	-0.168***	-0.142***	0.773***	0.728***
Bachelor's degree	0.103***	0.114***	-0.490***	-0.471***
Graduate or professional degree	0.172***	0.228***	-0.780***	-0.764***
<i>Ethnicity of the household head (baseline: White)</i>				
Black or African American	-0.336***	-0.208***	0.713***	0.856***
Asian	-0.132***	-0.203***	0.424***	0.251*
More than one ethnicity	0.017	-0.044	0.152	0.270
Other (single ethnicity)	-0.088**	-0.025	0.441***	0.627***
Household head is Hispanic/Latino	-0.136***	-0.108***	0.440***	0.415***
Count of household members born abroad	-0.070***	-0.025*	0.059	0.205***
Count of household members with a medical condition that makes it difficult to travel outside of home	0.109***	0.086***	-0.078**	0.233***
Household owns dwelling	0.000	-0.008	-0.084*	-0.114**
Number of vehicles per household adult	0.048***	0.101***	-0.258***	-0.177***
<b>Land use</b>				
Household resides in lower density census tract	0.031***	-0.002	0.133***	0.145***
Household resides in higher density census tract	0.054***	-0.009	-0.005	-0.081*
Household MSA has heavy rail	0.084***	0.084***	-0.279***	-0.108**
Constant	1.324***	0.418***	-1.689***	0.301**
Ln(α)	-0.217***	0.134***		

Notes.

- 1) The dependent variable is the number of deliveries to the household from online shopping in the past month. It was truncated at 60.
- 2) \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001. To assess statistical significance, we used robust standard errors to mitigate the potential impacts of heteroskedasticity.
- 3) A shaded cell indicates that the 2017 coefficient value for an explanatory variable differs from its 2009 value (p-value ≤ 0.05).
- 4) Count sub-models explain the number of deliveries to each household from online shopping. Logit sub-models explains the characteristics of households who never order goods online.
- 5) Lower density census tracts have less than 300 people per square mile. Higher density census tracts have over 7000 people per square mile.
- 6) MSA: Metropolitan Statistical Area.
- 7) α is the inverse of the scale parameter of the gamma noise variable in the negative binomial count component of a zero-inflated negative binomial model. Stata estimates and reports Ln(α), as well as its statistical significance.
- 8) The sample size is 123,148 for the 2017 NHTS and 134,371 for the 2009 NHTS.

**Table 4**  
Results for logit models characterizing in-store & e-grocery shoppers (2017 ATUS).

Variable	In-store grocery shopping		E-grocery shopping
	N = 10,223	N = 2934	N = 2934
<i>Marital status (baseline: married)</i>			
Never married	0.009	0.017	0.392
No Spouse	0.033	0.065	0.51
Gender (1 if female)	0.390***	0.413***	0.879**
<i>Age (baseline: Baby Boomer)</i>			
Generation Z (born after 1997)	-1.134***	-1.274***	-1.119
Generation Y (born 1981 to 1996)	-0.057	-0.353**	0.783
Generation X (born 1965 to 1980)	0.044	-0.151	0.67
Silent generation (born < 1946)	-0.082	-0.224	0.092
Number of household children	0.01	0.045	0.041
<i>Annual household income (baseline: \$60,000 to \$99,999)</i>			
Less than \$30,000	0.073	0.132	-0.365
\$30,000 to \$59,000	0.106	0.151	0.075
Over \$100,000	0.184**	0.133	-0.225
<i>Work status (baseline: full-time job)</i>			
Part-time job	0.102	0.021	0.493
Not employed	0.303***	0.293**	0.368
<i>Education (baseline: college degree)</i>			
Less than high school	-0.444***	-0.637***	0.295
High school	-0.276***	-0.344**	0.311
Some college/associate degree	-0.177**	-0.315**	-0.212
Graduate/professional degree	0.061	-0.064	-0.266
<i>Ethnicity (baseline: White)</i>			
African American	-0.393***	-0.338**	-0.503
Asian	0.033	0.107	-0.377
Other	0.01	NA	NA
Hispanic/Latino	0.116	0.108	-0.778
Mobility impairment	-1.012***	NA	NA
Constant	-1.971***	-1.819***	-5.187***

Notes: 1) \*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01.

- 2) For in-store grocery shopping, the dependent variable equals 1 if the ATUS respondent shopped in a brick-and-mortar grocery store on the survey day and zero otherwise. For e-grocery shopping, the dependent variable equals 1 if the ATUS respondent shopped online for groceries on the survey day and zero otherwise.
- 3) Our reduced sample (N = 2934) was obtained by keeping only ATUS respondents from core-based statistical areas (CBSA) where at least one other ATUS respondent shopped for groceries online. We estimated e-grocer characteristics using this sample to avoid the bias that would result from analyzing people who had no access to e-grocery shopping. We removed respondents whose ethnicity is "other" (neither African American, Asian, nor White) and who had a "mobility impairment" because none of them shopped for groceries online so the impact of these characteristics on e-grocery shopping could not be estimated.



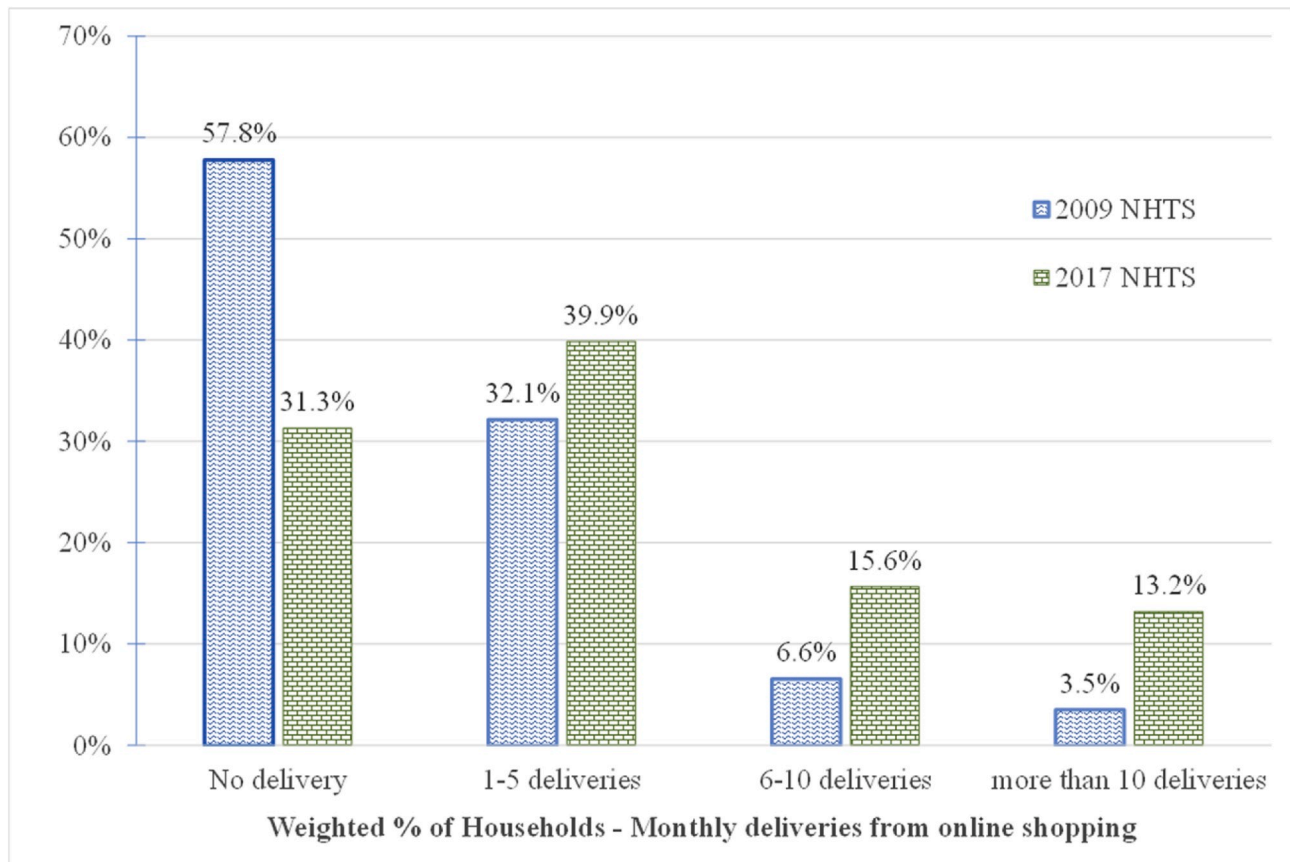


Fig. 1. Change in monthly household package deliveries from online shopping. Source: 2009 and 2017 National Household Travel Surveys.

5.1. Number of deliveries from online shopping (2017 and 2009 NHTS)

Before discussing results from our count models, it is instructive to contrast monthly household package deliveries from online shopping for 2009 and 2017 (Fig. 1). Results are weighted to be representative of the US population. As expected, Fig. 1 shows a sharp reduction in the percentage of households who do not get any deliveries (it drops from 57.8% to 31.3%) together with an increase in the percentage of households who received packages. This increase is especially marked for the 6–10 deliveries category (from 6.6% to 15.6%) and for more than 10 deliveries (from 3.5% to 13.2%).

For robustness, we estimated Poisson, negative binomial, ZIP, and ZINB models for our 2009 and 2017 NHTS samples using maximum likelihood. The ZINB models presented in Table 3 have the best (lowest) AIC and BIC values.

From Table 3, we first see that the impact of household composition on deliveries from internet shopping has been changing over time. Overall, the likelihood to never order online has decreased for most household types, and especially for households with 2 or more retired adults (−0.305\*\*). These households are now less unlikely than baseline households to shop online, possibly because older adults with a spouse are more likely to use modern information technologies (Vroman, Arthanat, & Lysack, 2015). However, differences in deliveries have sharpened (household composition made little difference in 2009): compared to our baseline households (2+ adults with children), 1-adult households (with and without children, retired or not) received fewer deliveries in 2017.

The impact of the number of children (household members under 18) has also been evolving and it is mixed. In 2017, having more children increased the likelihood of never shopping online, but it also slightly increased the number of deliveries from online shopping.

As reported by Ferrell (2005), having more females in the household

matters. In 2009, it decreased the likelihood of never shopping online (−0.167\*\*\*), without impacting the number of deliveries. In 2017, it both decreased the likelihood of never shopping online (−0.427\*\*\*), and increased deliveries (0.051\*\*\*).

The importance of the generational structure of households has also been changing, suggesting a broader adoption of online shopping. In 2009, an increase in the number of household members from younger generations (Gen X and Gen Y, and to a lesser extent Baby Boomers) sharply reduced the likelihood of never shopping online. Furthermore, increasing the number of household adults raised the number of deliveries. By contrast, in 2017, the magnitude of model coefficients for Gen X and Gen Y decreased. As Baby Boomers aged, however, they became more likely to never order goods online (0.371\*\*\*), although not as much as members of the Silent generation (1.028\*\*\*). As Parment (2013) explained, unlike Gen Y members, Baby Boomers often prefer to start a purchase with a retailer they trust, before eventually committing to a purchase, either online or in a store.

The impact of household income is monotonic. As their income rises, households are less likely to never order goods online and they tend to get more deliveries (as in Wang & Zhou, 2015). This effect increased across the board between 2009 and 2017. Likewise, as they gain workers, households are less likely to never order goods online, although the impact on deliveries is insignificant in 2017 (and in 2009 for part-time workers).

Similarly, as their level of education increases, households are less likely to never order goods online, and their deliveries increase. Except for households with less than a high school education in 2009, this relationship is monotonic. This result echoes the reported correlation between education and computer proficiency (Burroughs & Sabherwal, 2002).

As reported by Ren and Kwan (2009), race matters. Compared to White households (our baseline), African Americans (0.856\*\*\* for

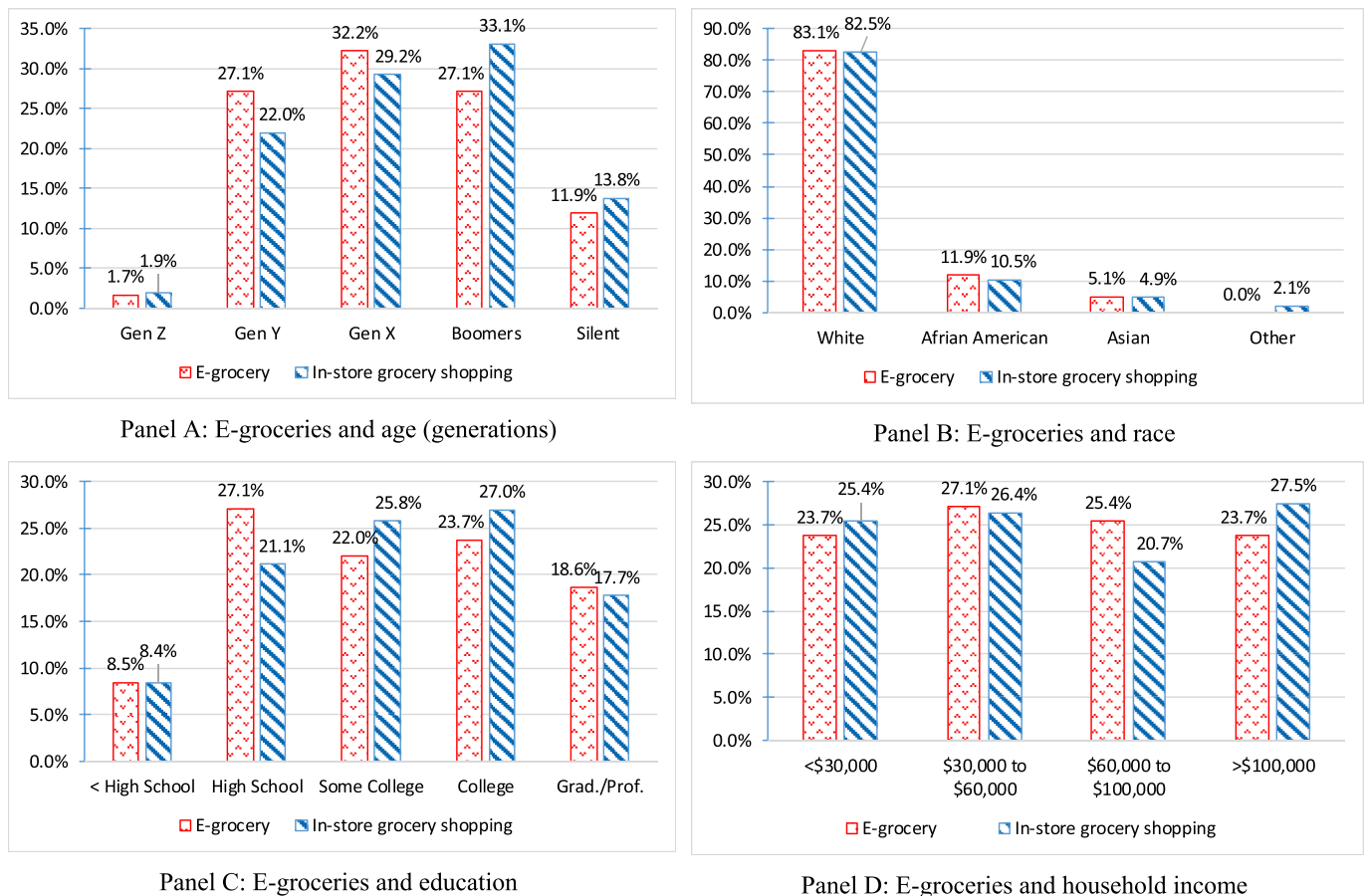


Fig. 2. Comparison of characteristics between online and conventional grocery shoppers. Source: 2017 American Time Use Survey.

2009) and Asians (0.251\* for 2009) are more likely to never order goods online and they tend to receive fewer deliveries. The same holds for Hispanic/Latino households. Conversely, whereas people born abroad were more likely to never order goods online in 2009, it is no longer the case in 2017. Having more foreign-born household members slightly decreases the number of deliveries, however.

People with a medical condition that hinders their mobility appear to take better advantage of the convenience of online shopping. Whereas in 2009 they were more likely to never order goods online (0.233\*\*\*), this effect disappeared in 2017 and those who shopped online received more deliveries in 2017 (0.109\*\*) than in 2009 (0.086\*\*\*).

Home ownership decreases the likelihood of never ordering goods online but it does not impact deliveries from online purchases. Likewise, households who have more vehicles per adult are less likely to never order goods online and they tend to receive more packages from online shopping. This result agrees with Zhou and Wang (2014), who reported that online shopping stimulates shopping trips and found that the number of household vehicles is positively correlated with the level of online shopping.

As expected, population density also plays a role here, although its impact is relatively small. In lower density areas (<300 people per square mile), households are more likely to never order goods online, but those who do received slightly more deliveries in 2017 (0.031\*\*\*). This result illustrates that e-shopping has the potential to increase the range of products available to households currently under-served by brick-and-mortar stores. Conversely, in denser areas (>7000 people per square mile), households are slightly less likely to never order goods online and they receive more deliveries (in 2017, not 2009). The same holds for households who reside in a Metropolitan Statistical Area (MSA) with heavy rail.

To capture the endogeneity of the land use and vehicle ownership variables, we could have estimated generalized structural equation models (GSEM; Kline, 2010) but we opted for ZINB models instead for two main reasons. First, these endogenous effects are small, as we verified by estimating ZINB models without land use and vehicle ownership variables. Second, GSEM models with non-continuous endogenous variables (population density is categorical and the presence of heavy rail is binary) are more difficult to interpret.

### 5.2. Characteristics of e-grocery shoppers (2017 ATUS)

Results from our analysis of the characteristics of US grocery shoppers are shown in Table 4. In this table, a positive coefficient for a variable indicates that the probability to shop for groceries (or e-groceries for the last column of Table 4) increases with that variable.

Starting with the logit model that characterizes people who shopped at brick-and-mortar grocery stores on their ASUS survey day (Column 2 in Table 4), we see that women are more likely to shop for groceries than men (0.390\*\*\*), which is well-known (e.g., see Morganosky & Cude, 2002; or Li, Sun, Zhang, & Hu, 2018). Apart for members of the Z generation who are less likely to shop for groceries, there are no generational effects. Likewise, the number of household children does not matter, and neither does household income, except for members of the highest income group (0.184\*). As expected, grocery shoppers are more likely to be “unemployed” (0.303\*\*\*), likely because they take care of the household while other household adults are at work and homemakers (still usually women) are consider unemployed. Interestingly, people with less education are less likely to shop for groceries. Race matters but only for African Americans, who appear to shop for groceries less frequently than other groups, which agrees with previous

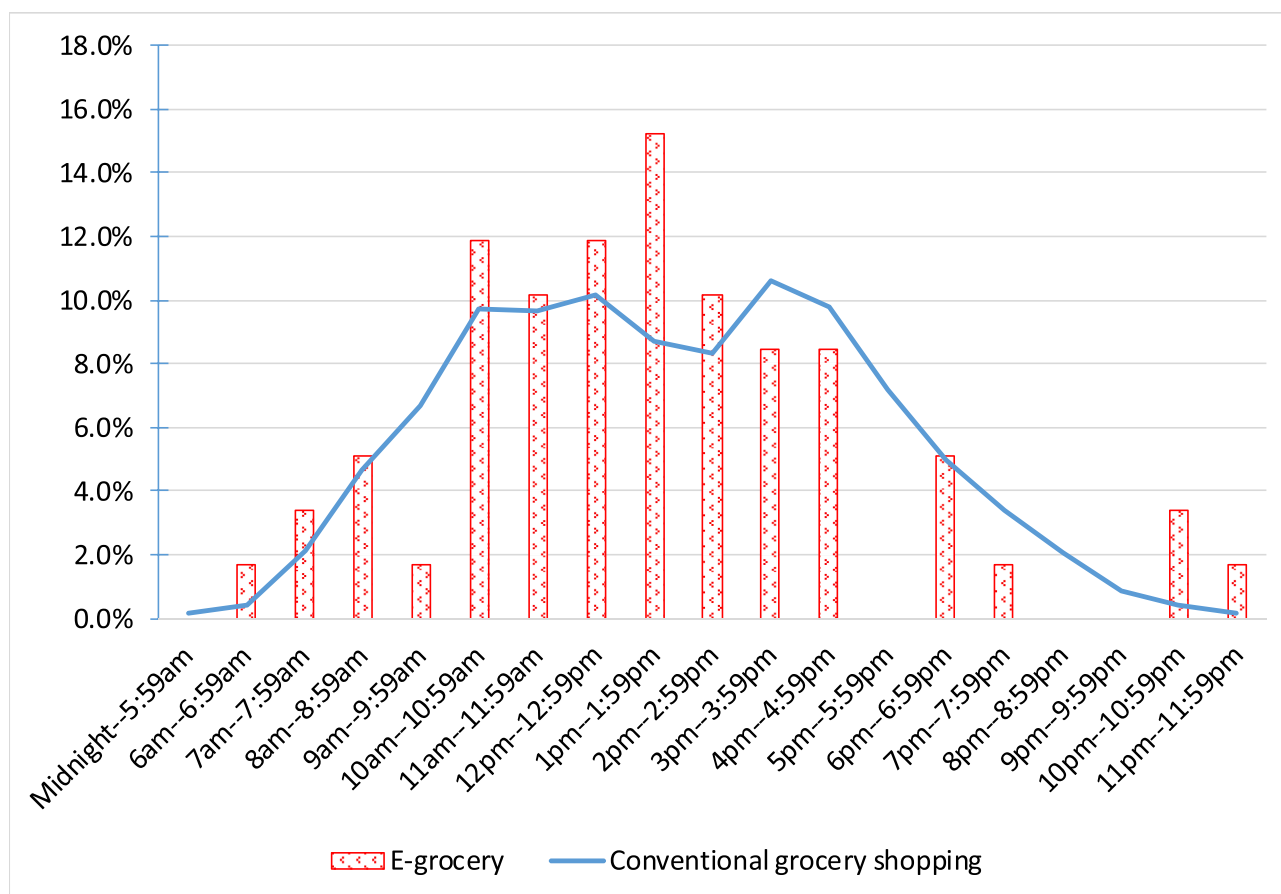


Fig. 3. Comparison of shopping start times. Notes. 1) Source: 2017 American Time Use Survey. 2) The shopping start time is the time when an ATUS respondent starts looking online for groceries.

findings that African Americans have less access to supermarkets (Morland, Wing, Roux, & Poole, 2002; Beaulac, Kristjansson, & Cummins, 2009). To compensate, poor people in predominantly black neighborhoods or in poor rural areas tend to shop at dollar stores<sup>4</sup> (Whalen, 2018), and rely more on fast food outlets (James, Arcaya, Parker, Tucker-Seeley, & Subramanian, 2014). Lastly, respondents with a medical condition that impairs their mobility are less likely to shop for groceries (−1.012\*\*).

When we focus on core-based statistical areas where at least one respondent in our sample shopped for groceries online (Column 3 in Table 4), we observe two main differences. First, members of the Y generation are less likely to shop for groceries. Second, all income groups become equally likely to shop for groceries.

By contrast, the only socio-economic characteristics that is statistically significant for people who shopped online for groceries is gender: women are more likely to shop online for groceries (0.879\*\*) than men, and this gender gap appears wider than for in-store shopping. This result echoes the findings of Morganosky and Cude (2002) in their study of 10 US markets based on three datasets collected between 1998 and 2001. However, when interpreting these results, it is important to keep in mind that the underlying dataset (N = 2,934) only includes 49 people who shopped for groceries online (we lost 10 respondents from our initial sample because their location is unknown).

To take full advantage of our relatively small sample of only grocery shoppers, we also explored differences in the distribution of selected

socio-economic characteristics of people who shop for groceries in stores and online using Kruskal-Wallis (KW) tests. Fig. 2 shows the empirical distributions of selected characteristics of the 1,420 people in the 2017 ATUS dataset who shopped for groceries in stores and the 59 who shopped online (so 24 times more people shopped for grocery in stores than online on a given day in 2017). Only the KW test for gender (not shown) was significant and it indicated that the gender difference for grocery shopping online is more marked than for grocery shopping in stores. Crosstabulation analyses using  $\chi^2$  tests gave similar results.

Fig. 3 contrasts the distributions of the time when people start grocery shopping online and in stores. It shows that online shopping activity picks up early afternoon (especially between noon and 2 pm) at a time when conventional grocery shopping tends to subside, and between 10 pm and midnight, when in-store grocery shopping tapers off. These differences highlight the added flexibility and convenience of online grocery shopping.

## 6. Conclusions

In this paper, we first analyzed data from the 2009 and the 2017 National Household Travel Surveys to understand changes in deliveries from online shopping in the US. Results from our zero-inflated negative binomial models show that online shopping in the US has been embraced by a much larger percentage of the US population, and that e-shoppers are more varied in 2017 compared to 2009, although differences in the number of deliveries resulting from online shopping are sharper than in 2009. In particular, households with more adult female members receive significantly more deliveries, and so do households with higher incomes and higher educational achievements. Even after controlling for other socio-economic characteristics, we found that

<sup>4</sup> Dollar stores are discount retailers that sell a wide range of products at low prices. Dollar Tree and Dollar General are the largest dollar stores in the United States, with over 14,000 locations each.

minority households are less likely to buy goods online, a disappointing finding that requires further investigations (one possible reason may be more limited access to credit cards). Finally, households with mobility impaired members rely more in 2017 than in 2009 on online shopping to satisfy their needs.

Understanding the determinants of e-shopping is important for supply chain managers so they can adapt the facilities handling online orders, adjust deliveries, and plan product returns. It is also of interest for planners and policymakers concerned with the externalities generated by deliveries of online orders (congestion, traffic accidents, air pollution, and noise) so they can consider appropriate incentives (e.g., subsidies for electric delivery vehicles in denser areas) and regulations (e.g., on the power train technology of delivery vehicles, hours of deliveries, or maximum noise levels).

Since US national household travel surveys do not track what e-shoppers purchase, we also analyzed grocery shopping data from the 2017 ATUS using logit models and non-parametric tests. Consistent with the literature, our results show that in-store grocery shoppers are more likely to be female and unemployed (because homemakers, who are still often female, are considered unemployed), but less likely to belong to younger generations, to have less than a college education, or to be African American, because poor people in predominantly black neighborhoods or in poor rural areas in the US tend to shop at dollar stores and rely more on fast food outlets. By contrast, the only significant socioeconomic variable for online grocery shoppers is gender: again, women are more likely to shop for groceries, and the gender gap is larger than for in-store grocery shopping. While this result may be partly due to the small number (only 59) of online grocery shoppers in our sample ( $N = 2,934$ ), the small number of people for groceries online reflects that e-grocery is not common currently in the US: on any given day, people are 24 times more likely to shop for groceries in stores compared to online.

Combining the profile of people who order groceries online with information about households who receive many online orders and data on local stores offering e-grocery could help understand where e-grocery is likely to succeed in the US. Policymakers concerned with access to fresh foods in underserved neighborhoods may then consider subsidizing delivery costs and the creation of grocery packing and delivery jobs. Emergency programs may also be put in place to deliver groceries to groups who cannot go grocery shopping, such as the elderly with impaired mobility or persons at-risk during contagious epidemic diseases (such as COVID-19).

As online shopping becomes ever more popular, the need for residential freight deliveries will keep on increasing. However, the magnitude of that increase and its impacts on traffic and the environment depend crucially on how last mile deliveries are organized. If packages are delivered to people's doorsteps with little coordination, soaring residential freight deliveries will increase congestion, noise, and air pollution, not to mention exacerbate parking shortages in denser urban areas. If, however, last mile deliveries are coordinated, performed as part of existing daily deliveries (e.g., via the US Postal Service), done by bicycle or electric vehicles, and/or go to lockers or neighborhood convenience stores (as is commonly done in Taiwan<sup>5</sup>, for example), their external costs could be much reduced (e.g., see Moore, 2019).

The last mile delivery problem is particularly acute for perishable groceries. If local demand is sufficiently high, preferred pricing could foster coordinated local deliveries during specific time windows, which would reduce the need for local freight trips. The widespread adoption of smart home lock systems (see Section 2.2) that allow deliveries to people's fridges when they are not home may also help, although less intrusive solutions such as click-and-pick at grocery stores or deliveries to local convenience stores (where available) may be cheaper and easier to implement.

One limitation of this work is the small number of people who shopped for groceries online in the 2017 ATUS, even though this survey gathered data from a representative sample of over 10,000 Americans (unfortunately only over a single day each), which reflects the current lack of popularity of online grocery shopping in the United States. A second limitation is that our e-grocery dependent variable may have missed some click-and-pick orders.

In future work, it would therefore be of interest to survey US households to gather detailed data about time use and travel, with retrospective questions about online purchases, in order to better capture the links between in-store and online purchases. Until e-grocery becomes more popular in the US, researchers could analyze stated preferences of US consumers for e-grocery to understand their preferences for various delivery and price options, and potential obstacles to e-grocery. To better understand the impact of last mile deliveries, it would also be of interest to monitor local freight deliveries in a wide range of neighborhoods. Lastly, we suggest analyzing international experiences to learn from creative solutions implemented elsewhere to reduce the external costs of e-grocery (and more generally e-shopping).

### CRediT authorship contribution statement

**Jean-Daniel Saphores:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing. **Lu Xu:** Writing - review & editing, Formal analysis, Writing - original draft.

### Acknowledgments

Helpful comments from participants of the Nectar Cluster 2 workshop that took place at Molde University College in Molde, Norway, on October 11 and 12, 2018, are gratefully acknowledged. We also thank the editor-in-chief, Professor Edoardo Marcucci, and anonymous referees for helpful comments that helped improve this paper substantially. As usual, we are responsible for any remaining errors.

### References

- Agatz, N., Campbell, A. M., Fleischmann, M., Van Nunen, J., & Savelsbergh, M. (2013). Revenue management opportunities for Internet retailers. *Journal of Revenue and Pricing Management*, 12(2), 128–138.
- Andersson, P., & Ivehamm, P. (2019). Benefits and costs of autonomous trucks and cars. *Journal of Transportation Technologies*, 9(2), 121–145.
- Anckar, B., Walden, P., & Jelassi, T. (2002). Creating customer value in online grocery shopping. *International Journal of Retail & Distribution Management*, 30(4), 211–220.
- Beaulac, J., Kristjansson, E., & Cummins, S. (2009). A systematic review of food deserts, 1966–2007. *Preventing Chronic Disease*, 6(3).
- Blumberg, S. J., & Luke, J. V. (2017). *Wireless substitution: Early release of estimates from the national health interview survey*. July–December 2016. National Center for Health Statistics. Available from: <https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201705.pdf>.
- Brashear, T. G., Kashyap, V., Musante, M. D., & Donthu, N. (2009). A profile of the Internet shopper: Evidence from six countries. *Journal of Marketing Theory and Practice*, 17(3), 267–282.
- Bressolles, G., Durrieu, F., & Senecal, S. (2014). A consumer typology based on e-service quality and e-satisfaction. *Journal of Retailing and Consumer Services*, 21(6), 889–896.
- Burroughs, R. E., & Sabherwal, R. (2002). Determinants of retail electronic purchasing: A multi-period investigation. *INFOR: Information Systems and Operational Research*, 40(1), 35–56.
- Calderwood, E., & Freathy, P. (2014). Consumer mobility in the Scottish isles: The impact of internet adoption upon retail travel patterns. *Transportation Research Part A: Policy and Practice*, 59, 192–203.
- Cherrett, T., Dickinson, J., McLeod, F., Sit, J., Bailey, G., & Whittle, G. (2017). Logistics impacts of student online shopping—Evaluating delivery consolidation to halls of residence. *Transportation Research Part C: Emerging Technologies*, 78, 111–128.
- Conover, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). Hoboken, NJ: Wiley.
- Conway, J. (2020). Share of grocery stores that offered home delivery/store pickup of online orders in the United States from 2017 to 2019. Available from: <https://www.statista.com/statistics/762746/store-share-online-grocery-delivery-store-pickup-services/#statisticContainer>.
- Crocco, F., Eboli, L., & Mazzulla, G. (2013). Individual attitudes and shopping mode characteristics affecting the use of e-shopping and related travel. *Transport and Telecommunication*, 14(1), 45–56.
- Dost, F., & Maier, E. (2018). E-Commerce effects on energy consumption: A multi-year ecosystem-level assessment. *Journal of Industrial Ecology*, 22(4), 799–812.

<sup>5</sup> Personal communication from Professor May Tsai, National Chung Hsing University, May 19, 2019.



- eMarketer, & Website (retailtechnewscom). (n.d.). E-commerce share of total global retail sales from 2015 to 2021. In Statista - the statistics portal. Retrieved May 9, 2019, from <https://www.statista.com/statistics/534123/e-commerce-share-of-retail-sales-worldwide/>.
- Farag, S., Schwanen, T., Dijst, M., & Faber, J. (2007). Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping. *Transportation Research Part A: Policy and Practice*, 41(2), 125–141.
- Federal Highway Administration (FHWA). (2011, October). 2009 national household travel survey user's guide. Version 2. Available from: <https://nhts.ornl.gov/2009/public/UsersGuideV2.pdf>.
- Ferrell, C. (2005). Home-based teleshopping and shopping travel: Where do people find the time? *Transportation Research Record: Journal of the Transportation Research Board*, 1926, 212–223.
- Freeman, M., & Freeman, A. (2011). Online grocery systems design through task analysis. *Journal of Enterprise Information Management*, 24, 440–454.
- Ganesh, J., Reynolds, K. E., Luckett, M., & Pomirleanu, N. (2010). Online shopper motivations, and e-store attributes: An examination of online patronage behavior and shopper typologies. *Journal of Retailing*, 86(1), 106–115.
- Greene, W. H. (2011). *Econometric analysis* (7th ed.). Boston, London: Pearson.
- Grunert, K. G., & Ramus, K. (2005). Consumers' willingness to buy food through the internet: A review of the literature and a model for future research. *British Food Journal*, 107(6), 381–403.
- Hand, C., Riley, F. D., Harris, P., Singh, J., & Rettie, R. (2009). Online grocery shopping: The influence of situational factors. *European Journal of Marketing*, 43(9/10), 1205–1219.
- Hansen, T. (2005). Consumer adoption of online grocery buying: A discriminant analysis. *International Journal of Retail & Distribution Management*, 33(2), 101–121.
- Harris, P., Dall'Olmo Riley, F., Riley, D., & Hand, C. (2017). Online and store patronage: A typology of grocery shoppers. *International Journal of Retail & Distribution Management*, 45(4), 419–445.
- Hawkins, A. J. (2018). Amazon will now deliver packages to the trunk of your car. The Verge. Available from: <https://www.theverge.com/2018/4/24/17261744/amazon-package-delivery-car-trunk-gm-volvo>. April 24.
- InsideRetail Hong Kong. (n.d.). E-commerce share of total retail sales in China from 2014 to 2019. In Statista - the statistics portal. Retrieved May 9, 2019, from <https://www.statista.com/statistics/379087/e-commerce-share-of-retail-sales-in-china/>.
- Ioannides, Y. M., Overman, H. G., Rossi-Hansberg, E., & Schmidheiny, K. (2008). The effect of information and communication technologies on urban structure. *Economic Policy*, 23(54), 202–242.
- James, P., Arcaya, M. C., Parker, D. M., Tucker-Seeley, R. D., & Subramanian, S. V. (2014). Do minority and poor neighborhoods have higher access to fast-food restaurants in the United States? *Health & Place*, 29, 10–17.
- Kamarainen, V., Smaros, J., Jaakola, T., & Holmstrom, J. (2001). Cost-effectiveness in the e-grocery business. *International Journal of Retail & Distribution Management*, 29, 41–48.
- Kang, C., Moon, J., Kim, T., & Choe, Y. (2016, January). Why consumers go to online grocery: Comparing vegetables with grains. In *2016 49th Hawaii international Conference on system sciences (HICSS)* (pp. 3604–3613). IEEE.
- Klein, R., Neugebauer, M., Ratkovich, D., & Steinhart, C. (2019). Differentiated time slot pricing under routing considerations in attended home delivery. *Transportation Science*, 53(1), 236–255.
- Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). The Guilford Press.
- Kwan, M. P. (2007). Mobile communications, social networks, and urban travel: Hypertext as a new metaphor for conceptualizing spatial interaction. *The Professional Geographer*, 59(4), 434–446.
- Lee, R. J., Sener, I. N., Mokhtarian, P. L., & Handy, S. L. (2017). Relationships between the online and in-store shopping frequency of Davis, California residents. *Transportation Research Part A: Policy and Practice*, 100, 40–52.
- Li, Y., Sun, Y., Zhang, Y., & Hu, X. (2018). Online grocery retailing for fresh products with order cancellation and refund options. *Journal of Supply Chain and Operations Management*, 16(1), 1–16.
- Lien, T. (2017). Apoorva mehta had 20 failed start-ups before Instacart. Los angeles times (january 27). Available from: <https://www.latimes.com/business/technology/la-fi-himi-apoorva-mehta-20170105-story.html>.
- Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Sage Publications, Inc.
- Marchet, G., Melacini, M., Perotti, S., Rasini, M., & Tappia, E. (2018). Business logistics models in omni-channel: A classification framework and empirical analysis. *International Journal of Physical Distribution & Logistics Management*, 48(4), 439–464.
- McDonald, R., Christensen, C., Yang, R., & Hollingsworth, T. (2014). *Amazon fresh: Rekindling the online grocery market*. Harvard Business School Case, 615-013.
- McGuckin, N., & Fucci, A. (2018). Summary of travel trends: 2017 national household travel survey. FHWA report No. FHWA-PL-18-019. Available from: [https://nhts.ornl.gov/assets/2017\\_nhts\\_summary\\_travel\\_trends.pdf](https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf).
- Melis, K., Campo, K., Breugelmans, E., & Lamey, L. (2015). The impact of the multi-channel retail mix on online store choice: Does online experience matter? *Journal of Retailing*, 91(2), 272–288.
- Meyerson, N. (2019). How Wegmans and H-E-B survived Amazon's onslaught. CNN Business, November 13. Available from: <https://www.cnn.com/2019/11/13/business/heb-wegmans-publix-hy-vee-regional-grocery/index.html>.
- Meyerson, N. (2020). This grocery store wanted to compete with Whole Foods. Now it's going bankrupt. CNN Business, January 28. Available from: <https://www.cnn.com/2020/01/28/business/luckys-market-bankruptcy-aldi-publix-kroger-trnd/index.html>.
- Moore, A. M. (2019). *Innovative scenarios for modeling intra-city freight delivery*. Transportation Research Interdisciplinary Perspectives. Available from: <https://www.sciencedirect.com/science/article/pii/S2590198219300247>.
- Morganosky, M. A., & Cude, B. J. (2000). Consumer response to online grocery shopping. *International Journal of Retail & Distribution Management*, 28(1), 17–26.
- Morganosky, M. A., & Cude, B. J. (2002). Consumer demand for online food retailing: Is it really a supply side issue? *International Journal of Retail & Distribution Management*, 30(10), 451–458.
- Morland, K., Wing, S., Roux, A. D., & Poole, C. (2002). Neighborhood characteristics associated with the location of food stores and food service places. *American Journal of Preventive Medicine*, 22(1), 23–29.
- Murphy, A. J. (2003). Resolving space and time: Fulfillment issues in online grocery retailing. *Environment & Planning A*, 35(7), 1173–1200.
- Nahiduzzaman, K. M., Aldosary, A. S., & Mohammed, I. (2019). Framework analysis of E-commerce induced shift in the spatial structure of a city. *Journal of Urban Planning and Development*, 145(3), 04019006.
- Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2), 311–329.
- Office for National Statistics. (2018). Internet sales as a percentage of total retail sales. Available from: <https://www.ons.gov.uk/businessindustryandtrade/retailindustry/timeseries/j4mc/drs1>.
- Parment, A. (2013). Generation Y vs. Baby Boomers: Shopping behavior, buyer involvement and implications for retailing. *Journal of Retailing and Consumer Services*, 20(2), 189–199.
- Perboli, G., & Rosano, M. (2019). Parcel delivery in urban areas: Opportunities and threats for the mix of traditional and green business models. *Transportation Research Part C: Emerging Technologies*, 99, 19–36.
- Pettersson, F., Winslott Hiselius, L., & Koglin, T. (2018). E-commerce and urban planning—comparing knowledge claims in research and planning practice. *Urban, Planning and Transport Research*, 6(1), 1–21.
- Pew Research Center. (2018). Defining generations: Where Millennials end and post-Millennials begin. Available from: <http://www.pewresearch.org/fact-tank/2018/03/01/defining-generations-where-millennials-end-and-post-millennials-begin/>.
- Punakivi, M., & Saranen, J. (2001). Identifying the success factors in e-grocery home delivery. *International Journal of Retail & Distribution Management*, 29(4), 156–163.
- Redman, R. (2018). *Gallup: Most Americans just aren't into online grocery*. Supermarket News. August 13, Available from: <https://www.supermarketnews.com/consumer-trends/gallup-most-americans-just-arent-online-grocery>.
- Ren, F., & Kwan, M. P. (2009). The impact of geographic context on e-shopping behavior. *Environment and Planning B: Planning and Design*, 36(2), 262–278.
- Robinson, H., Dall'Olmo, R., Rettie, R., & Rolls-Willson, G. (2007). The role of situational variables in online grocery shopping in the UK. *The Marketing Review*, 7(1), 89–106.
- Saunders, N. (2018, October 17). Online grocery: Lessons from history [blog post]. Retrieved from <https://www.onespace.com/blog/2018/10/online-grocery-lessons-history/>.
- Seaman, C. E. A. (1995). Food on the internet. *Nutrition & Food Science*, 95, 26–29.
- Smith, A., & Anderson, M. (2016, December). Online shopping and E-commerce. Pew research center. Available from: <http://www.pewinternet.org/2016/12/19/online-shopping-and-e-commerce/>.
- Suel, E., & Polak, J. W. (2018). Incorporating online shopping into travel demand modelling: Challenges, progress, and opportunities. *Transport Reviews*, 38(5), 576–601.
- Thompson, D. (2017). What in the world is causing the retail meltdown of 2017? The Atlantic. April 10. Available from: <https://www.theatlantic.com/business/archive/2017/04/retail-meltdown-of-2017/522384/>.
- US Census Bureau News. (2019, March 13). Quarterly retail E-commerce sales, 4th quarter 2018. Available from: [https://www.census.gov/retail/mrts/www/data/pdf/ec\\_current.pdf](https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf).
- United States Census Bureau. (2018). Historical table 5. Estimated annual sales for US electronic shopping and mail-order houses (NAICS 54111) - total and E-commerce sales by merchandise line: 1999-2016. Available from: <https://www.census.gov/data/tables/2016/econ/e-stats/2016-e-stats.html>.
- United States Census Bureau. (2018). Historical table 4. Estimated annual US retail trade sales - total and E-commerce: 1998-2016. Available from: <https://www.census.gov/data/tables/2016/econ/e-stats/2016-e-stats.html>.
- Vroman, K. G., Arthanat, S., & Lysack, C. (2015). Who over 65 is online? Older adults' dispositions toward information communication technology. *Computers in Human Behavior*, 43, 156–166.
- Wadud, Z. (2017). Fully automated vehicles: A cost of ownership analysis to inform early adoption. *Transportation Research Part A: Policy and Practice*, 101, 163–176.
- Wang, R. J. H., Malthouse, E. C., & Krishnamurthi, L. (2015). On the go: How mobile shopping affects customer purchase behavior. *Journal of Retailing*, 91(2), 217–234.
- Wang, X. C., & Zhou, Y. (2015). Deliveries to residential units: A rising form of freight transportation in the US. *Transportation Research Part C: Emerging Technologies*, 58, 46–55.
- Wat, F. K. T., Ngai, E. W. T., & Cheng, T. C. E. (2005). Potential risks to e-commerce development using exploratory factor analysis. *International Journal of Services Technology and Management*, 6(1), 55–71.
- Weathers, D., Sharma, S. L., & Wood, S. (2007). Effects of online communication practices on consumer perceptions of performance uncertainty for search and experience goods. *Journal of Retailing*, 83(4), 393–401.
- Whalen, A. (2018). Dollar stores planning for permanent American underclass, sell more groceries than Whole Foods. Newsweek, 12/07/18. Available from: <https://www.newsweek.com/dollar-stores-general-tree-america-economy-poor-people-food-pov-erty-lower-1250186>.



- White, G. K. (1997). Using the WWW for retail food sales: Discussion and demonstration. *Journal of Food Distribution Research*, 28, 26–29.
- Wollenburg, J., Hübner, A., Kuhn, H., & Trautrim, A. (2018). From bricks-and-mortar to bricks-and-clicks. *International Journal of Physical Distribution & Logistics Management*, 48(4), 415–438, 2018.
- Wollerton, M. (2018). Amazon Key took over my door for 3 months. It wasn't as creepy as I expected. CNET. Available from: <https://www.cnet.com/news/amazon-key-too-k-over-my-door-for-3-months-it-wasnt-as-creepy-as-i-expected/>. May 25.
- Zhou, Y., & Wang, X. C. (2014). Explore the relationship between online shopping and shopping trips: An analysis with the 2009 NHTS data. *Transportation Research Part A: Policy and Practice*, 70, 1–9.