Journal of Retailing and Consumer Services 71 (2023) 103191

Contents lists available at ScienceDirect

Journal of Retailing and Consumer Services

journal homepage: www.elsevier.com/locate/jretconser

Exploring the choice between in-store versus online grocery shopping through an application of Semi-Compensatory Independent Availability Logit (SCIAL) model with latent variables

Kaili Wang [a,*], Ya Gao [a,b], Yicong Liu [a], Khandker Nurul Habib [a]

[a] Department of Civil & Mineral Engineering, University of Toronto, 35 St. George St, Toronto, ON, M5S 1A4, Canada
[b] Department of Urban Planning, Tongji University, 1239 Siping Road, Shanghai, 200020, China

ARTICLE INFO

Keywords:
Online shopping
Grocery delivery
Semi-compensatory choice model
Latent variables
Willingness to pay

ABSTRACT

This paper examines individuals' choice of in-store and online grocery shopping channels using stated preference (SP) choice experiments. The study uses 1,391 records from a stated preference choice experiment in the Greater Toronto Area (GTA), Canada. It applies a Semi-Compensatory Independent Availability Logit (SCIAL) Model with latent variables. The methodology accounts for semi-compensatory choice behaviour through probabilistic choice set formation considering effects from socioeconomic and psychological variables. This study demonstrates the advantage of considering probabilistic choice set formation and semi-compensatory behaviour in modelling the adoption of innovative products. Empirical results reveal that shoppers demonstrated similar myopic behaviours once they firmly considered in-store grocery and subscribed free delivery services in their choice sets. They are equally likely to choose both channels without careful comparison to alternative channels once they firmly consider both channels in the choice set. However, considering the latter in choice sets is much costlier than in-store shopping. Therefore, in-store grocery shopping will still dominate the grocery shopping channel unless all home delivery services become free. Moreover, grocery shoppers value same-day delivery service. For typical delivery services charged between \$4 and \$20 in the GTA, Canada, grocery shoppers are willing to pay between \$3.91 and \$8.44 for same-day delivery. The latent variable describing shoppers' perceived pandemic fear significantly contributes to the choice set inclusion probability of in-store grocery pick-up services, but the effect is not significant for other home delivery channels. This highlights heterogeneity in grocery shoppers' choice behaviour within the online channel.

1. Introduction

Online shopping has experienced persistent increases over the past years. Moreover, online grocery shopping has been fueled by the COVID-19 pandemic (McKinsey, 2022a). Since the onset of the pandemic in 2020, grocery stores have started providing online shopping and home delivery services. E-commerce giants like Amazon and transportation network companies (TNCs) like Uber also launch their grocery delivery services. Resultantly, the growth in market share of online grocery shopping became a worldwide phenomenon. In North America, online grocery accounted for 14.3% of the total grocery sales in 2021, compared to 8.3% in 2020 (McKinsey, 2022b). In Europe, the online grocery market shares were led by the United Kingdom, France, and the Netherlands. In 2021, their domestic market sold 12%, 8.6%, and 7.5% of their grocery online (McKinsey, 2022b). By moderate forecast, the online grocery market share will account for 19%, 16%, and 17% of the market share in the United Kingdom, France, and the Netherlands by 2030 (McKinsey, 2022b). In Asia, the growth in online grocery was led by China. The online sale took 20%–25% of the domestic grocery market in 2022, compared to 10% in 2019 (McKinsey, 2022c).

Online grocery shopping services allow shoppers to purchase perishable products, packaged foods, beverages, and other daily necessities online and deliver them to homes. Although the popularity of online groceries was instigated by measures like social distancing during the pandemic, industry experts believe that shoppers who adopted online grocery shopping tend to stay with the service in the post-pandemic

* Corresponding author.
E-mail addresses: jackkaili.wang@mail.utoronto.ca (K. Wang), gaoya@tongji.edu.cn (Y. Gao), nora.liu@mail.utoronto.ca (Y. Liu), khandker.nurulhabib@utoronto.ca (K. Nurul Habib).

https://doi.org/10.1016/j.jretconser.2022.103191
Received 30 July 2022; Received in revised form 19 October 2022; Accepted 6 November 2022
Available online 15 November 2022
0969-6989/

era (Shen et al., 2022). The market penetration of online grocery shopping might continue after the pandemic. McKinsey (2022a) predicts that by 2030, online services could account for 18 to 30 percent of the food-at-home market in leading developed countries. The shifts in shopping behaviours may have long-lasting effects on all stakeholders. Consumers might shift their daily routine and activity-travel patterns (Mokhtarian et al., 2006; Shamshiripour et al., 2020). Retailers are expected to develop omnichannel strategies to maintain competitiveness. Urban freight transport needs to be reconfigured to serve the needs of home delivery (Marcucci et al., 2021). Policymakers need to understand the potential impact to better plan and manage urban areas. Therefore, it is crucial for all stakeholders to fully comprehend the determinants of grocery shoppers' shopping channel choices.

Despite the extensive body of literature on retail shopping behaviours, scientific literature focusing on online grocery shopping channel choice are underdeveloped. Earlier studies in retailing and marketing science commonly focused on shopping location and store choice. These studies commonly treated formative alternatives by brand names and store types (Aaker and Jones, 1971; Bhatnagar and Ratchford, 2004; Reutterer and Teller, 2009; González-Benito, 2002). Later, researchers were interested in channel choice behaviour (e.g., online vs. in-store) for durable goods (Farag et al., 2006a; Hsiao, 2009; Cao, 2012). Until recently, grocery shopping channel choice behaviours started to receive attention. Several studies investigated determinants of grocery channel choices (Beckers et al., 2018; Clarke et al., 2015; Wieland, 2022; Shen et al., 2022; Schmid and Axhausen, 2019; Marcucci et al., 2021; Gatta et al., 2021; Suel and Polak, 2017).

However, there are two notable limitations in existing literature studying grocery shopping channel choice. Firstly, most studies only focused on binary channel choice between online and in-store grocery shopping channels (Beckers et al., 2018; Clarke et al., 2015; Wieland, 2022; Shen et al., 2022; Schmid and Axhausen, 2019). The online grocery shopping ecosystem has introduced a number of new business models, such as ordering online and picking up at the nearest brick-and-mortar stores or offering membership subscriptions to enjoy free home delivery for subscribers (Dias et al., 2020). This highlights the discussion by Suel and Polak (2018) on the risk of channel choice studies focusing on aggregated channel choice and overlooking within-channel heterogeneity. Secondly, previous studies all used discrete choice techniques that assumed full-compensatory choice behaviours. This assumption implies individuals put all shopping channels into their consideration and evaluate them carefully. This assumption might not be appropriate to examine innovative products in their early adoption stage. Many consumers may not consider innovative products as feasible options; therefore, in their choice process, they will hardly evaluate the performance attributes associated with innovative products.

This study contributes to the literature on grocery channel choice by dealing with the two limitations. Firstly, this study considers heterogeneity in online grocery shopping channels. The stated preference (SP) experiment conducted in this study considers three types of online delivery business models. The first online model charged consumers per delivery. The second model provides free delivery to homes with monthly subscription charges. The third model allows consumers to order online and pick up in stores. Examining several online channels simultaneously will shed light on consumers' within-channel behavioural heterogeneity. Secondly, this paper applies a novel discrete choice modelling framework using the Semi-Compensatory Independent Availability Logit (SCIAL) Model with latent variables in the specification. The modelling framework accommodates semi-compensatory choice behaviour by considering probabilistic choice set formation using choice-makers' socioeconomic and psychological variables. The modelling approach captures the choice process when consumers think certain shopping channels are infeasible. The model formulation presented in this paper integrates the Semi-Compensatory Independent Availability Logit (SCIAL) model and the classical integrated Choice and Latent Variable (ICLV) model into one.

The remainder of the paper is organized as follows. **Section 2** presents brief literature reviews and summarizes the research gaps. **Section 3** presents the data used in this study, descriptive analysis, and experimental designs. **Section 4** presents the proposed modelling methodology. Model results are discussed in **Section 5**, followed by conclusions and discussions on future research in **Section 6**.

2. Literature review

Literature has made extensive progress in examining shoppers' choices between in-store and e-shopping. A heated debate was about whether e-shopping had a substitution, complementarity, modification, or neutrality effect on in-store shopping (Farag et al., 2003; Choo et al., 2007; Mokhtarian and Patricia, 1990; Salomon, 1986). Among the four effects, the potential of e-shopping to substitute in-store shopping has important implications for travel demand management and congestion mitigation (Cao, 2009). However, the existing literature did not settle with any specific effect above. They found that neither e-shopping nor in-store shopping uniformly dominates the other in consumer shopping channel choices. Some literature found a certain degree of substitution between online and in-person shopping, but the magnitude of the substitution effect did not significantly reduce the total amount of shopping trips (Handy and Yantis, 1997; Sim and Koi, 2002; Tonn and Hemrick, 2004).

Conversely, several other studies also identified complementary effects between in-store and e-shopping (Cao et al., 2010; Farag et al., 2006a, 2007). Besides substitution and complementarity effects, some studies also found complicated shopping behaviours. Consumers combine different shopping channels during different stages of the consumer buying process to maximize their total shopping utility (Hsiao, 2009; Cao, 2012). In-store buyers might search online to gather information before their in-store purchase. Likewise, e-shoppers might also travel to stores for product trials before their purchase online.

Literature also examined the determinants of consumer e-shopping behaviours. The influential factors found include consumers' personal and household characteristics, vendor and product characteristics, and attitudinal factors. Rotem-Mindali and Salomon (2007) found that online shoppers had higher shopping orientations and more frequent internet usage. Several studies came to similar conclusions that male, young, highly educated, and wealthy individuals were more likely to purchase online (Farag et al., 2006a; Sim and Koi, 2002; Beckers et al., 2018). Some researchers also found a non-linear relationship between age and online buying, with people aged between 26 and 45 being more inclined to buy online (Farag et al., 2005).

While e-shopping was in its early introductory stages, as with any emerging technology, literature investigated its penetration and adoption patterns. Anderson et al. (2003) proposed two hypotheses on possible e-shopping adoption patterns. First, the innovation-diffusion hypothesis viewed e-shopping as technological innovation. Thus, adopting e-shopping will follow a typical early adopter and follower pattern. Second, the efficiency hypothesis treated e-shopping as a substitution for in-person shopping, especially for consumers with low shopping accessibility. Testing the two hypotheses, several studies examined land-use characteristics such as residential location and shopping accessibility. Farag et al. (2003, 2005) found that people living in densely populated urban areas were more likely to adopt online shopping. At the same time, people having low in-store shopping accessibility were also more likely to use online shopping. The earlier finding supported the innovation-diffusion hypothesis. They found that urban residents might be more tech-savvy and have access to convenient internet and supply chain services. These factors lead to more frequent e-shopping purchases. The latter finding also supported the efficiency hypothesis that e-shopping was used to substitute in-person shopping for efficiency.

Beckers et al. (2018) used revealed preference (RP) data and found that most online shoppers reside in densely populated urban areas.

However, they pointed out that socioeconomic status also impacts online shopping demand in densely populated urban regions. For instance, densely populated urban areas primarily occupied by low-income households could process lower e-shopping demand.

Zhen et al. (2018) also used RP data to investigate the effects of spatial attributes on shopping channel choices. Like Farag et al. (2003, 2005), their findings supported both hypotheses. Individuals who living in peripheral areas are more likely to shop in person. In the meantime, individuals with lower accessibility to shopping outlets were more likely to shop online.

All studies summarized above share at least one of the common characteristics, if not both. First, their study periods were relatively early. The market share of e-shopping was only marginal while most works were conducted. Secondly, these studies only focused on durable goods (e.g., books, clothing, and electronic devices). None explicitly consider grocery shopping, which has become one of the latest niches for e-commerce investments.

Recently, several studies have been specifically focused on grocery shopping. A number of studies are concerned with online grocery shopping demands. Suel et al. (2018) used the hazard-based model to investigate the temporal structure of grocery shopping demand. They found that performing e-shopping greatly reduces the probability of in-store grocery shopping. They also found that e-shopping could reduce trips to grocery stores but could not alter the household's overall grocery shopping demand. Again, their finding confirmed the substitution pattern between online and in-store shopping. Effective delivery services might be a key factor influencing consumers' choice of transaction media. The rapid advancement of information communication technology (ICT) and supply chain management allowed same-day delivery of orders placed online. Xi et al. (2020) studied the influences of same-day delivery e-shopping on in-store shopping frequency for five stores (supermarkets, convenience stores, vegetable markets, fruit stores, and restaurant products). They found that same-day delivery services would reduce the demand for in-store shopping trips to all types of local stores.

Several pieces of literature focused on consumers' shopping channel choice and their determinants. Clarke et al. (2015) used revealed preference (RP) shopping data to investigate determinants for online grocery shopping in the UK. They found that males, young individuals and high-income households were more likely to shop online. The findings are similar to previous studies summarized above. They also found that online grocery shoppers had poorer accessibility to grocery retail stores (measured by the average distance to the nearest grocery stores) compared to in-person shoppers. Their findings support the efficiency hypothesis that shoppers used online services to compensate for their lack of accessibility.

Suel et al. (2015) used UK Living Cost and Food Survey data to examine buyers' choice between online and in-store grocery shopping within a 14-day diary period. Substitution effects between online and in-store grocery shopping for both household and individual levels were identified. Their study showed that the types of products purchased would determine the choice of purchase channels. Also, e-shopping basket sizes turned out to be bulkier than in-store shopping basket sizes since the burden of transporting large baskets could be offloaded to delivery services.

Suel and Polak (2017) modelled consumers' store choices considering all individual grocery stores, including online stores, in the study area. Their approach implicitly dealt with channel choice. They spent incredible data collection efforts collecting real purchase data using consumer panels whereby participants could log their daily purchases, especially groceries. The dataset was then augmented with transport-related and store attributes from a database maintained by private companies. Their data collection and preparation efforts are very costly and hard to be replicated. The alternatives in the choice sets were all individual retail stores in the study area and several online stores that provided equivalent services. The final dataset recorded 272 shopping instances. However, only 11 observations were e-shopping indicating the market share of online grocery shopping was still marginal when the study was conducted. The limited observations found that e-shopping alternatives attracted higher-income groups with larger shopping baskets. They also found that shopping for groceries online mainly substituted driving trips instead of walking and transit trips to stores. This finding highlights the possibility of alternated vehicular travel patterns due to the mass adaption of online grocery shopping.

The pandemic fueled the market penetration of online grocery shopping and called for up-to-date empirical evidence describing the state of development in this niche. Shen et al. (2022) collected RP data to investigate individuals' choice of their primary channel of grocery shopping (in-store vs. e-shopping). They found females, vehicle owners, and health-constrained and risk-aversion individuals were more likely to use online grocery shopping.

Most recently, Wieland (2022) also collected RP data to investigate determinants for online shopping on individuals' shopping behaviours in German. He considered four types of goods: groceries, clothing, electronic devices, and furniture. On top of socio-economic variables, he considered latent psychological variables like shopping and pandemic-related attitudes. Like other relevant studies, he found that aged individuals were less likely to shop for groceries online. For psychological factors, he found that risk-aversion to the virus and perceiving wearing a face mask as inconvenient did not increase the likelihood of using online grocery shopping.

All studies discussed above used RP data. The data reflected real grocery purchasing behaviours. However, they all suffered from data availability issues (except for Suel and Polak (2017), who spent an incredible effort to augment the dataset), so their studies did not consider any cost-related explanatory variables such as shopping baskets cost, travel impedance to stores, and monetary cost for home delivery services. These variables are vital to explain shopping channel choice behaviour in a specific choice context.

To overcome the data availability issue in RP data, studies used the stated preference (SP) experiment technique to examine channel choices between in-store and online grocery shopping (Schmid and Axhausen, 2019; Marcucci et al., 2021; Gatta et al., 2021). Schmid and Axhausen (2019) investigated buyers' channel choices for grocery and durable goods shopping. They collected data using the stated preference technique and applied integrated choice and latent variable (ICLV) models. The result indicated that grocery shopping was still mainly conducted in person. E-shoppers were characterized as younger and well-educated male individuals. Furthermore, respondents with positive e-shopping attitudes demonstrated more price sensitivity. However, their choice set only contains two alternatives: online and in-store shopping. This overlooked various business models operating simultaneously in the home grocery delivery sector.

Marcucci et al. (2021) and Gatta et al. (2021) conducted similar SP choice experiments on grocery shopping channel choices in Norway and Shanghai, China, respectively. In Norway, they found that consumers could be categorized into two classes. One class is consisted of hard-core in-store shoppers, whereas the other class represents online shoppers. Individuals in the online shopper class were sensitive to product price, delivery fee, travel time and delivery time (Marcucci et al., 2021). They found that Chinese consumers were also sensitive to similar cost-related attributes (Gatta et al., 2021).

Using stated preference choice experiments, Rossetti et al. (2022) investigated grocery shoppers' trade-off between the perceived inconvenience of shopping and various safety measures introduced because of the pandemic. They found that rigorous health and safety measures attracted shoppers. Grocery shoppers are willing to accept longer queues in exchange for rigorous health and safety measures while shopping in stores. Although Rossetti et al. only focused on in-person store choice, their work demonstrates the necessity to consider psychologic factors in discrete choice models, especially in disruptive contexts.

To sum up, previous studies on grocery shopping channel choice examined the effects of socioeconomic attributes, attitudinal attributes, land-use characteristics, and variables that describe shopping channels' performance (e.g., cost & travel impendence). However, most studies, except Schmid and Axhausen (2019), only tested part of the variables listed above. Still, Schmid and Axhausen (2019) only focused on aggregated channel choice overlooking within-channel heterogeneity. Moreover, all work above assumed full-compensatory choice behaviour. The limitation of using full-compensatory choice models without consideration of choice set formation to study innovative products is discussed in detail in Section 1.0. This study overcomes the limitations summarized above using the SCIAL model, considering a holistic set of explanatory variables.

Fig. 1 presents the conceptual framework of the study. In detail, consumers' grocery shopping channel choices are outcomes of the interaction between the supply and demand sides. The demand side describes consumers' grocery shopping needs and their propensity to consider specific channels as feasible options. Their consideration will be influenced by socioeconomic, attitudinal, and land-use characteristics constraints. Finally, their channel choice will be determined by evaluating the performance of all feasible channels. On the other hand, the supply side describes the performance of each shopping channel. Retailers' management decisions determine channel performance. Retailers' location, operation logistics, and pricing strategy will determine consumers' temporal and monetary costs to use each channel. Meanwhile, retailers' product mix strategy will influence the types of grocery products available in each channel. Retailers could modify the strategies mentioned above to achieve their market share target. Moreover, retailers could also conduct marketing efforts targeting specific consumer groups to influence their choice set formation consideration.

3. Data & experimental design

The following section describes the data used in this study and the detailed experimental design of the online grocery shopping and home delivery stated preference (SP) choice experiment. For simplicity, the choice experiment is referred to as the home delivery SP experiment in the remainder of the paper.

3.1. COVHITS survey & sample statistics

The stated-preference (SP) choice experiment was conducted as part of the **COV**id-19 influenced **H**ouseholds' **I**nterrupted **T**ravel **S**chedules (COVHITS) survey. The COVHITS surveys are a series of online household travel surveys monitoring passenger travel demands in the Greater Toronto Area (GTA), Canada, during the COVID-19 pandemic (Wang et al., 2021). In total, three cycles of COVHITS surveys were conducted in the Fall of 2020, the Summer of 2021, and the Fall of 2021. The home delivery SP choice experiment was conducted as part of the 2021 Summer COVHITS cycle. The 2021 Summer COVHITS survey is a proxy-based household travel survey. Individuals at least 18 years old are allowed to participate in the survey as self-respondents. Other than typical information (e.g., household, personal socioeconomic attributes, and travel diaries for household members over six years old), the survey also conducted home delivery SP experiments on the self-respondents. Moreover, attitudinal questions regarding the pandemic and online shopping were collected following the SP experiment.

The 2021 Summer COVHITS cycle was conducted between July and August 2021. The samples were randomly drawn from panels maintained by market research companies. The final dataset of the 2021 Summer COVHITS survey contains 1,876 households. The final dataset was cleaned in this study to ensure data quality. The lowest and highest 2.5th percentile are removed from the analysis based on respondents' average response time for choice scenarios. Respondents who merely glanced at the SP scenario or idled for hours during the SP section reflect their careless attitudes towards the experiment. In addition, respondents who provided completely conflicting responses to paired attitude questions are also removed from the analysis. After data cleaning, 1,391 valid samples remained for further analysis.

Table 1 presents key descriptive statistics of the samples used in the study. Representativeness of the sample is assured by comparing sample statistics with the 2016 Transportation Tomorrow Survey (TTS). The TTS was a regional household travel survey covering the study area of this study (Data Management Group, 2018). The latest TTS cycle in 2016 is used as the reference for this study. Overall, the sample matches well with the 2016 TTS. All general personal and household attribute trends are similar between the two surveys. However, some discrepancies do exist. On the personal level, the sample over-represents individuals younger than 30 years old compared to the 2016 TTS. The discrepancy is caused by different survey modes between the two surveys. The COVHITS surveys are online surveys. However, the 2016 TTS adopted online and telephone modes. Therefore, the samples collected in the 2021 Summer COVHITS survey contain younger individuals who are more used to online devices.

On the household level, the 2021 Summer COVHITS survey has higher proportions of single and two-person households compared to the 2016 TTS. This is consistent with the discussion above, as individuals between 18 and 30 years old are more likely to reside in smaller households. Also, samples in this study appear to contain more medium ($60,000 - $99,999) and high-income (>$100,000) households in the 2021 Summer COVHITS survey. 65.5% of the samples in the COVHITS survey are from these two categories. However, only 49.1% of the samples in the 2016 TTS are medium and high-income households. The different sampling methods utilized in the two surveys could cause this discrepancy. In the 2016 TTS, samples were drawn randomly from

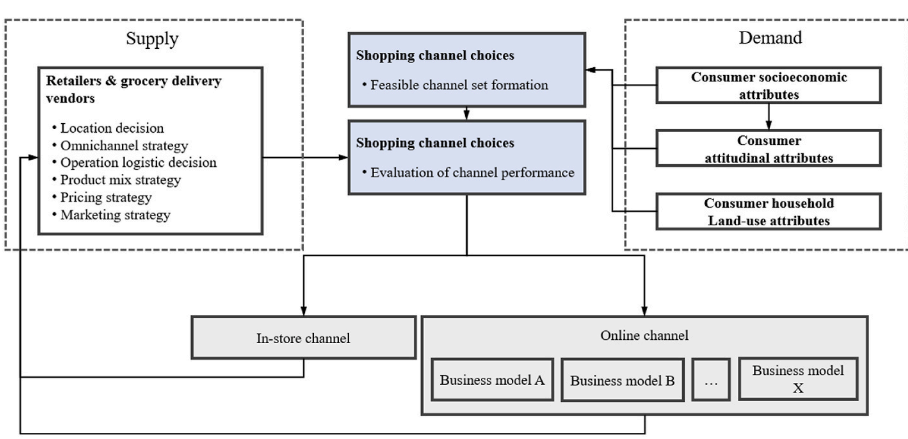

Fig. 1. Conceptual framework of the study.

Table 1
Descriptive statistics of key socioeconomic attributes.

Attributes	2021 Summer COVHITS (N = 1,391)	2016 TTS
Personal attributes of self-respondents		
Gender		
female	55.6%	49.5%
Age		
18-29	21.7%	11.2%
30-39	21.2%	19.4%
40-49	18.0%	21.0%
50-59	15.0%	21.2%
60-64	8.4%	8.0%
≥65	15.7%	19.2%
Employment status		
Full-time	54.1%	58.8%
Part-time	13.8%	9.4%
Not employed	32.1%	31.7%
Household attributes		
Household size		
1	30.2%	24.8%
2	34.8%	28.2%
3	16.9%	17.4%
4+	18.2%	29.7%
Household having dependent children		
Yes	25.2%	
Household dwelling types		
House	59.7%	46.4%
Townhouse/apartment	40.3%	53.6%
Household income		
< $14,999	2.4%	5.0%
$15,000 - $59,999	24.6%	28.2%
$60,000 - $99,999	29.9%	21.4%
> $100,000	35.6%	27.7%
Decline	7.5%	17.7%
Household vehicles		
0	8.8%	17.4%
1	45.8%	40.7%
2	35.9%	31.7%
3+	9.5%	10.3%
Household having online grocery membership		
Yes	21.5%	

regional databases of mailable addresses. (Data Management Group, 2018). On the other hand, panelists from online commercial survey panels were randomly selected to participate in the 2021 Summer COVHTIS survey. The commercial panels might contain a higher proportion of medium- and high-income households compared to the study area. To sum up, the samples used in this study have reasonable repetitiveness to the general population in the study area. The overall trends of all key socioeconomic statistics match closely with the reference regional travel survey. However, the samples in this study might be slightly younger, wealthier, and entitled to higher mobility compared to the general population in the study area. Therefore, caution should still be used while applying this study's results to the study area's general population. Moreover, the readers should note that individuals who are not internet users are not in the sampling frame of the COVHITS survey. The Canadian internet user penetration in 2021 was 97.9% (Statista, 2022). Thus, the sub-group of the population left by the sampling frame should be marginal and will not significantly affect the findings in this study.

3.2. Experimental design

The SP choice experiment examines individuals' trade-offs between in-store and online grocery shopping via different channels. The experiment considers two forms of grocery shopping: online versus in-store shopping. For online grocery shopping, delivery fee and subscription-based services account for different business models available in the market. In total, five distinct alternatives are taken into consideration. **Alternative 1** considers in-store grocery shopping. **Alternative 2** considers online grocery shopping and having the products delivered to homes. In alternative 2, individuals do not have subscriptions with the service provider. So, they are subjected to delivery fees for each home delivery service. **Alternative 3** considers individuals shopping grocery online, having the products delivered to the home, and purchasing service subscriptions simultaneously. Individuals will be entitled to free home delivery services once they subscribe. This alternative is designed to reflect the customer acquisition strategy commonly used by service providers. While customers purchase groceries, service providers will also try to sell their membership subscriptions with incentives such as free deliveries.

On the other hand, **alternative 4** is for individuals who have already subscribed to a grocery delivery service. Alternative 4 considers online grocery shopping and free delivery to homes because of free already subscripted memberships. Finally, **alternative 5** is shopping for groceries online and picking them up in the stores. The alternative eliminates in-store shopping time for grocery shoppers. Alternatives 1, 2, and 5 are always available for all respondents. Availability of alternatives 3 and 4 are determined by respondents' reported subscription status for home delivery services. Respondents who are not subscribed to any online grocery delivery service will have availability for alternative 3 but not alternative 4, and vice versa.

Appendix A summarizes the characteristics of online grocery shopping services already operated in the GTA. Overall, there are two types of service providers in the market. Third-party online grocery companies serve multiple brands with their personnel and resources. At the same time, grocery brands themselves will offer their online grocery and delivery service. Some service providers deliver all types of products in the stores. However, certain service providers only limited their delivery service to non-perishable products for quality assurance. There are two business models for price schemes. Service providers could charge delivery fees to customers on each purchase.

Moreover, some service providers might have a policy to waive delivery fees once the basket price exceeds specific amounts. On the other hand, customers could also purchase providers' membership subscriptions. Once subscripted, customers will be entitled to fee delivery services under conditions such as their orders exceeding certain thresholds. Meanwhile, the waiting period for groceries to arrive at homes ranges from 1 h to three weeks, depending on providers. Most service providers offer same-day or next-day delivery services if they have enough capacity.

Attributes and levels are selected based on the characteristics of online grocery shopping services summarized above and reference from the literature on shopping channel choice (Chintagunta et al., 2012; Schmid and Axhausen, 2019; Marcucci et al., 2021). Table 2 presents attributes and levels for all five alternatives in the choice experiment. In total, four types of attributes are considered. The first set of attributes describes the characteristics of goods purchased. Basket price defines the total amount consumers spend on shopping carts. The price will range from $25, $50, $75, $100, $150 to $300 (all currency is in Canadian dollars). The basket characteristics consider different proportions of perishable and standardized products in the shopping carts. Perishable products are goods such as fruit and meat, which are likely to decay with time and have higher variance in terms of quality. On the other hand, standardized products are goods such as canned food, beverages, and sanitary products that are unlikely to decay and have little to no quality variance. Then, the attribute of delivery service provider describes whether the service is provided by a third-party company, or the grocery store itself.

The third sets of attributes consider time and travel impedance associated with each alternative. Shopping time in stores accounts for time spent in grocery stores for alternative 1. The attribute ranges from 15, 30, 45, to 60 min. Travel modes to stores consider cars, transit, and active modes (e.g., walking). Meanwhile, travel time to stores ranges from 10, 15, 30–45 min. Delivery time and pick-up time characterize the waiting period between the time of online purchasing and items arriving home or ready for pick-up. They share similar levels ranging from 2 to 4

Table 2
Attributes and levels for the online grocery shopping and home delivery SP choice experiment.

Attributes	Applicable alternatives	Levels
Characteristics of goods		
Basket price	1,2,3,4,5	\$25, \$50, \$75, \$100, \$150, \$300
Basket characteristics	1,2,3,4,5	context 1: only perishable products context 2: only standardized products context 3: mostly standardized products, half standardized & half perishable products, mostly perishable products
Characteristics of delivery service provider		
Service provider types	2,3,4,5	brand operated, third-party operated
Time and travel		
Shopping time in stores	1	15 min, 30 min, 45 min, 60 min
Travel modes to stores	1,5	car, transit, active modes
Travel time to stores	1,5	10 min, 15 min, 30 min, 45 min
Delivery time	2,3,4	2–4 h, same day, next day, a week or later
Pick-up time	5	2–4 h, same day, next day, a week or later
Delivery cost		
Minimum order for fee delivery	2	not applicable, order over \$99
Delivery fee	2	\$4, \$8, \$10, \$12
Subscription fee for unlimited free delivery	3	Starting to pay \$5, \$7, \$10 or \$12 per month now
Minimum order for fee delivery	4	free for any order, order over \$20, order over \$40, order over \$70
Delivery fee if under minimum order	4	\$2, \$4, \$6
Pick-up fee	5	\$0, \$2, \$4, \$6

h, same day, next day to a week or later.

Finally, the fourth set of attributes defines the cost of delivery or pick-up service associated with alternatives 2, 3, 4, and 5. In alternative 2, respondents will be charged \$4, \$8, \$10, or \$12 for groceries delivered to their homes. However, their fee might be exempted in some scenarios if their basket price exceeds \$99. In alternative 3, respondents will face the choice of paying monthly subscription fees of \$5, \$7, \$10, or \$12 for the privilege of free grocery delivery service. In alternative 4, respondents will be charged \$4, \$8, \$10, or \$12 for their groceries to be delivered, if their basket price is lower than the minimum order required by the service provider. Otherwise, their grocery delivery will be free. Finally, in alternative 5, respondents will be charged \$0, \$2, \$4, or \$6, so they can skip the line and pick up their groceries directly in the stores.

Three contexts of choice scenarios are designed. Choice scenarios are contextualized by the characteristics of grocery products in the shopping carts. The first context account for only perishable products. The second context accounts for only scandalized products. Finally, the third context accounts for a mixture of perishable and scandalized products. The number of choice scenarios in the experimental design will be governed by the attribute levels (Kocur et al., 1981). Therefore, twelve scenarios are designed for each context. The D-efficient design was utilized to ensure sufficient variable variations in choice scenarios (ChoiceMetrics, 2018). During the survey, for each context, 2 out of 12 scenarios are randomly presented to each respondent to avoid respondent fatigue. Thus, each respondent will face six scenarios in total.

4. Methodology

This section presents the Semi-Compensatory Independent Availability (SCIAL) model with latent variables. Fig. 2 presents the modelling workflow. The methodology accommodates semi-compensatory choice behaviour by considering choice set formation with choice-makers' psychological factors. The modelling framework reaps the features of the fully compensatory Independent Availability Logit (IAL) model, Constrained Multinomial Logit (CML) model, and classical Integrated Choice and Latent Variable (ICLV) model. The remainder of the section will describe the model formulation in detail.

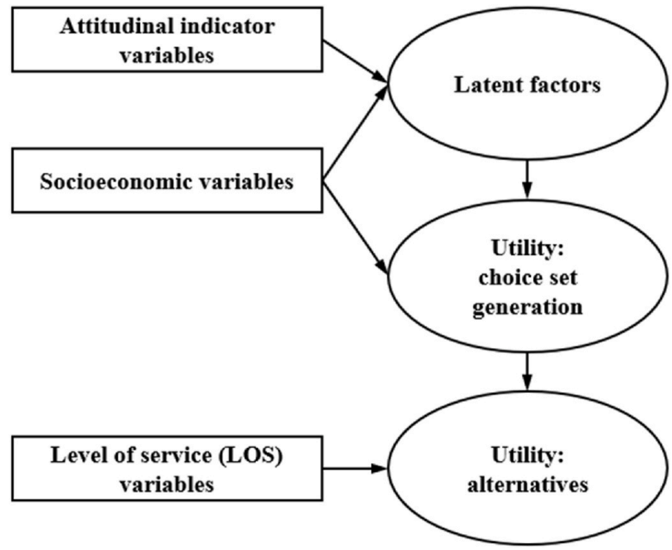

Fig. 2. Schematic diagram of the Semi-Compensatory Independent Availability Latent Variable (SIALV) Logit model.

4.1. The choice model component

The total utility U_j of a choice alternative under the Random Utility Maximization (RUM) framework follows.

$$U_j = V_j + \frac{1}{\mu}\ln(A_j) + \varepsilon_j \ , j \in C_m \tag{1}$$

where V_j is the linear-in-parameter systematic utility. A_j is the arbitrary penalty function to account for the availability of alternative j. ε_j is the random error term following independent and identically distributed (IID) Type I Value distribution. μ is the scale parameter of distribution of random error term. C_m is the master choice set containing all alternatives.

Following the logit formulation, the choice probabilities (P_j) of observing choosing alternatives j in a particular choice scenario is defined as:

$$Pr(j) = \frac{\exp(\mu(V_j) + \ln(A_j))}{\sum_{j\in C_m}\exp(\mu(V_j) + \ln(A_j))} = \frac{A_j \exp(\mu(V_j))}{\sum_{j\in C_m} A_j \exp(\mu(V_j))} \tag{2}$$

4.2. Probabilistic choice set generation

The formulation of $Pr(j)$ above assumes deterministic choice sets generation. While considering probability choice set formation, the unconditional probability (Manski, 1977):

$$Pr(j) = Pr(j \mid C_m)\Pr(C_m) \tag{3}$$

where C_m denote the master choice set that contains all alternatives.

Swait and Ben-Akiva (1987) postulated the probability that some collection of alternatives in C_k is individual i's choice set as:

$$Pr(C_k | not\ all\ A_j = 0, j \in C_m) = \frac{\Pr(A_i = 1, \forall i \in C_k)\ and\ \Pr(A_l = 0, \forall l \in S_{m-c})}{Pr(not\ all\ A_j = 0, j \in C_m)} \tag{4}$$

where C_k denotes the choice set for individual k. Finally, S_{m-c} denotes the complement set of $C_m \cap C_k$.

Assuming independent availability, Swait and Ben-Akiva (1987) further specified the choice set formation probability in Equation (4) as:

$$Pr(C_k|C_m) = \frac{\prod_{i \in C_k} \Pr(A_i = 1) \prod_{j \in S_{m-c}} \Pr(A_j = 0)}{1 - \prod_{j \in C_m} \Pr(A_j = 0)} \tag{5}$$

where A_j is the binary availability indicator for alternative j, $A_j \in \{1, 0\}$.

Combining equations (3) and (5), Swait and Ben-Akiva (1987) postulated a fully compensatory Independent Availability Logit (IAL) model as follows:

$$P(j \mid C_k) = \frac{A_j^{'} \exp(\mu(V_j))}{\sum_{j \in C_m} A_j^{'} \exp(\mu(V_j))} \tag{6}$$

$$A_j^{'} = 1 \ if \ A_j^{'} > 0, A_j^{'} = 0 \ otherwise \quad for \ j \in C_k \tag{7}$$

Martínez et al. (2009) proposed Constrained Multinomial Logit (CML) model, which relaxed the binary condition of $A_j^{'}$ stated in equation (7). Later, Habib (2019) proposed a Semi-Compensatory Independent Availability Logit (SCIAL) model by combining IAL and CML. Under SCIAL model, the $P(j \mid C_k)$ is as follows:

$$P(j \mid C_k) = \frac{A_j \exp(\mu(V_j))}{\sum_{j \in C_m} A_j \exp(\mu(V_j))} \tag{8}$$

A_j is the probabilistic choice set formation function that follows:

$$A_j = \frac{1}{1 + \exp\left(-\sum \beta_j x_k\right)} \quad for \ j \in C_k \tag{9}$$

where $\beta_j x_k$ is the liner-in-parameter systematic utility as a function of socioeconomic variables for individual k. With equation (9), equation (5) can be written as:

$$Pr(C_k|C_m) = \frac{\prod_{j \in C_k} \Pr(A_j) \prod_{j \in S_{m-c}} \Pr(A_j)}{1 - \prod_{j \in C_m} \Pr(A_j)} \tag{10}$$

4.3. Latent variable with the choice set formation

The systematic utility of a choice alternative j in typical integrated choice latent variable (ICLV) formation follows (Bhat and Dubey, 2014; Ben-Akiva and Boccara, 1995; Habib and Zaman, 2012; Vij and Walker, 2016):

$$V_j = \beta_j x_k + \gamma_j q_j + \alpha_j h_k + \varepsilon_j \tag{11}$$

where x_k is the vector of socioeconomic variables for individual k, q_j is the level-of-service variables, h_k is the unobserved latent variables, β_j, γ_j and α_j are vectors of estimated coefficients, and ε_j is the error term following Type I Extreme Value distribution.

The formation above should be further improved to capture more realistic choice marking behaviours. In the modelling framework of this paper, socioeconomic and latent variables are postulated to affect choice set formation. Socioeconomic and psychological variables describe the choice maker's medium to long-term status and will govern their choice set formation. While the choice set is probabilistically formed, the choice among alternatives within the available choice set is picked by evaluating level-of-service attributes specific to each choice scenario. This approach is believed to be a more realistic reflection of reality. For example, in travel mode choice, individuals might consider private vehicles in their choice set by purchasing private cars based on their socioeconomic status or latent factors such as preferred lifestyles. However, when choosing the mode of travel for specific trips, individuals will evaluate the level-of-service associated with each alternative in their available choice set and select the best available alternative.

Therefore, in the formulation of the SCIAL model considering latent variables (LV), equation (11) will be re-specified as:

$$V_j = \gamma_j q_j + \varepsilon_j \tag{12}$$

The choice set inclusion probability in Equation (9) will be re-specified as:

$$A_j = \frac{1}{1 + \exp\left(-\sum \beta_j x_k + \alpha_j h_k\right)} \quad for \ j \in C_k \tag{13}$$

The conditional choice probability, $P(j \mid C_k)$, and choice set formation probability, $Pr(C_k|C_m)$ still follow Equations (8) and (10) in the original formulation of SCIAL model.

Finally, the unconditional choice probability of the SCIAL model considering latent variables in choice set formation follows:

$$Pr(j) = \sum_{C_k \in C_m} Pr(j \mid C_k) \Pr(C_k) \tag{14}$$

The Multiple Indicator Multiple Cause (MIMIC) model is used to make use of latent variables.

The structural equations are specified as:

$$h_k = \alpha \bullet x_k + \xi_k \tag{15}$$

where h_k is the vector of continuous latent variables, x_k is the vector of socioeconomic variables of choice maker k, α is the vector of estimated coefficients, and ξ_k is normally distributed errors with zero means and unit variance for estimation ease.

The measurement equations are specified as:

$$y_k = \gamma \bullet h_k + v_k \tag{16}$$

where y_k is the vector of the observed indicator variable, h_k is the vector of the latent variable, γ is the vector of estimated coefficients, and v_i is normally distributed errors with zero means and unit variance.

If considered y_k as normally distributed continuous variables:

$$\Pr(y_k) = \frac{1}{\sigma_k} \Phi\left(\frac{y_k - \gamma \bullet h_k}{\sigma_k}\right) \tag{17}$$

4.4. Likelihood function and model estimation

The likelihood function is open formed with multi-dimensional integration. Therefore, the simulation likelihood approach is used to estimate parameters for the SCIAL model with latent variables. For latent variables, the likelihood of observing a particular response to a specific set of N indicators for an individual k is:

$$LI_{k,n} = \prod_{n=1}^{N} \Pr(y_{k,n}) \tag{18}$$

for stated-preference choice experiments, the simulated log-likelihood function for the SIALV Logit model is as follows:

$$LL = \sum_{k=1}^{K} ln\left(\frac{1}{R} \sum_{r=1}^{R} \left[\prod_{i=1}^{I} Pr(j)\right]\right) \tag{19}$$

where K is the total number of records in the sample, R is the total number of simulated draws, I is the total number of stated preference choice scenarios observed from each record, and N is the total number of observed latent variable indicators from each record.

The likelihood function of SCIAL and MIMIC components can be estimated using sequential and simultaneous approaches. Both approaches will result in unbiased estimators (Raveau et al., 2010). The study applies a sequential estimation approach due to the feasibility of computation time. Indeed, the simultaneous approach is more efficient because it jointly uses all information available. However, the complexity of the model formulation makes the simultaneous approach impractical with commonly available computation resources. Calculating unconditional choice probability in Equation (14) requires exclusive consideration of all possible choice set combinations. With n alternatives, the consideration has $2^n - 1$ cases. In addition, estimation of the MIMIC model must use a simulated likelihood function which typically requires more than 1,000 random draws to ensure results reliability. The simultaneous estimation approach must fulfill these requirements, resulting in infeasibly long computation time. As a result, this paper uses a sequential estimation approach to overcome the above-stated issue. The likelihood functions can be estimated in GAUSS using the classical BFGS gradient search algorithm (Aptech Systems, 2014).

4.5. Explanatory variables

This study examines the influence of four sets of explanatory variables on grocery shopping channel choice. The first set of variables describes the level-of-service (LOS) associated with each shopping channel. The LOS variables define the performance of each channel. The LOS variables are included to examine characteristics of shopping channels and various transaction costs (e.g., temporal, monetary and transportation-related costs) on channel choice. Their attribute levels are determined in the SP experimental design. Readers could refer to Table 2 and the discussion in **Section 3.2** for a detailed description of each variable.

The second set of variables considers the influence of socioeconomic variables on grocery shoppers' choice set inclusion probability. All previous studies on shopping channel choice studied the impact of socioeconomic characteristics (Farag et al., 2006a; Beckers et al., 2018; Clarke et al., 2015; Zhen et al., 2018; Schmid and Axhausen, 2019; Wieland, 2022; Marcucci et al., 2021; Gatta et al., 2021). Five socioeconomic variables are included in this study. First, age is included as a continuous variable. It is expected to be inversely associated with the likelihood of using online channels. One dummy variable indicates if the individual has male gender. Third, household size is included as a continuous variable. A dummy variable for household income is included if the household earns more than $100,000 yearly. It is expected that high-income households are more likely to consider online channels. Lastly, the number of vehicles owned by the household is included as a continuous variable. It is a proxy measurement for the level of mobility processed by the household.

The third sets of variables reflect the influence of land-use characteristics on choice set inclusion probability. Two variables are considered. The first variable describes the population density in the travel analysis zone (TAZ) where households locate. TAZ is the unit of geography used in transportation studies (Miller, 2021). The urban region is divided into mutually exclusive and exhaustive zones. The second variable is a continuous variable describing the number of retail facilities in the TAZ where households locate. This variable is a proxy of households' grocery retail accessibility. The land-use variables are included to test the innovation-diffusion and efficiency hypotheses (Anderson et al., 2003). Detailed descriptions of the two hypotheses can be found in **Section 2.0**. If the innovation-diffusion hypothesis is supported, residents in densely populated areas are more likely to choose home delivery channels. If the efficiency hypothesis is supported, individuals with higher impedance to access grocery outlets are more likely to consider online channels.

The fourth sets of variables are latent variables that might affect shopping channel choices. The first latent variable concerns individuals' shopping channel choices due to the pandemic.

The latent variable describes individuals' risk aversion toward the COVID-19 pandemic. Rossetti et al. (2022) found that risk-aversion attitudes toward the pandemic would influence grocery shopping behaviours. It is expected that attitudes concerning the risk of infection would lead to a higher likelihood of considering online shopping as a feasible channel. The second latent variable describes respondents' perceived level of convenience in receiving deliveries at home. None of the prior studies tested this latent variable. It is expected that higher perceived convenience in receiving deliveries contributes to a higher likelihood of considering online channels as a feasible channel. Lastly, the third latent variable describes respondents' degree of pickiness towards the quality of perishable grocery products. Chintagunta et al. (2012) found the inability to perform quality verification induced additional transaction costs in online grocery shopping channels. Therefore, it is expected that individuals with a higher level of quality pickiness are more likely to consider the in-store channel.

5. Modelling results

This section presents results from the empirical investigation. The section has two parts. The first part presents results of psychometric modelling that identifies and validates latent factors from choice-makers using Confirmatory Factor Analysis (CFA). After identifying latent factors, the second part applies the SCIAL model with latent variable (LV) to investigate individuals' determinants of online grocery shopping via different channels.

5.1. Psychometric modelling

5.1.1. Factor analysis

Factor analysis was used to determine the number and nature of the latent factors. The 2021 Summer COVHITS survey collected 22 attitudinal questions related to seven potential latent factors. These factors are informed by earlier literature and the judgment of the authors (Farag et al., 2005; Mokhtarian et al., 2006; Xi et al., 2020). All attitudinal questions have scaled from 1 to 5 (from strongly disagree to strongly agree). All psychometric models are estimated using the *lavaan* package in R (Rosseel, 2012). The confirmatory factor analysis (CFA) identifies three latent factors: pandemic fear, delivery convenience, and product quality. Pandemic fear describes individuals' levels of concern toward the COVID-19 pandemic and their beliefs about the risks associated with conducting out-of-home activities during the pandemic. (Mashrur et al., 2022). Delivery convenience refers to the perceived levels of easiness for individuals to receive deliveries at home. Product quality is the degree to which individuals value the quality of perishable groceries they purchase.

Measurements of latent factors and their standardized factor loadings are reported in Table 3. The results indicate excellent levels of reliability and validity of latent and indicator variables. All the factor loadings are higher than the 0.40 standard limits and statistically significant at the 95th percentile confidence level (Stevens, 2002). The reliability and validity of indicators are measured using Cronbach's alpha coefficients and composite reliability (CR) scores (Tavakol and Dennick, 2011). Cronbach alpha and CR scores for all indicators are greater than the threshold of 0.7. The omega (ω) coefficients, first mentioned by McDonald, are recommended to report in recent research (e.g., Tahlyan et al., 2022). Finally, the omega (ω) coefficients are found to be 0.86, 0.84, and 0.79 for the three latent factors, indicating sufficient reliability (McDonald, 1999; Tahlyan et al., 2022).

Additionally, convergent and discriminant validity tests are also conducted on all three latent factors. Table 3 also presents the results of both tests. For all latent factors, the average variance extracted (AVE) values for all factors are above the threshold value of 0.5 (Hair et al., 2006). For discriminant validity, the square root values of AVE for each factor exceeded the intercorrelations between all the possible pairs of factors. All results above confirm the validity of the latent factors identified in this study.

Table 3
Factor structures of identified latent factors and results of validity tests.

Latent factors & Observed indicators	Mean	SD.	Factor loading (λ)	Cronbach's alpha (α)	Composite reliability	Omega (ω)
Pandemic fear						
It is important to practice social distancing	4.22	0.97	0.879	0.850	0.835	0.861
I believe mandatory face covering is important to protect shoppers	4.32	0.98	0.87			
I followed the directives to avoid non-essential out of home activity	4.28	0.88	0.686			
Delivery convenience						
The place where I am living can receive home delivery conveniently, especially when I am not at home	3.87	1.02	0.907	0.829	0.861	0.835
Receive home delivery is convenient for my residence	3.65	1.01	0.78			
Product quality						
The high quality of perishable products is very important	4.42	0.74	0.875	0.783	0.786	0.786
I will always pick and choose the best quality products when shopping for perishable products	4.25	0.77	0.736			

Latent constructs	No. of items	AVE	Discriminant validity		
Pandemic fear	3	0.680	**0.824**		
Delivery convenience	2	0.717	0.182	**0.847**	
Product quality	2	0.649	0.552	0.294	**0.806**

Notes: Notes: SD = standard deviations; λ = standardized factor loadings from the confirmatory factor analysis; AVE = average variance extracted; square root values of AVE are bolded.

5.2. The SCIAL model with latent variables

5.2.1. The latent variable (LV) component

The latent variable component is estimated sequentially to identify the causal effects of choice-makers' socioeconomic variables on latent variables. The results of the latent variable component are shown in Table 4. All estimated variables are statistically significant at the 95 percent confidence interval (t-statistics of 1.64 for the one-tailed test). For the causal effect of socioeconomic variables, younger individuals and males are found to contribute to the less perceived fear of the pandemic and vice versa. Results also suggest that the perceived home delivery convenience are higher for females, individuals having dependent children, and households with more than one. Furthermore, elderly individuals, female, and individuals from high-income households are found to have higher perceived values over the quality of their perishable groceries.

Table 4
Estimated parameters of the MIMIC component.

		Est.	t-stat
Structural model			
Pandemic fear	constant	3.51	13.12
	logarithm of age	0.34	4.78
	gender as male	−0.20	−3.21
Delivery convenience	constant	3.78	70.43
	gender as male	−0.12	−2.51
	having at least one dependent child	0.15	2.49
	single-person household	−0.27	−5.41
Product quality	constant	2.30	6.82
	logarithm of age	0.43	4.79
	gender as male	−0.22	−2.87
	Household income > $100,000	0.13	1.70
Measurement model			
Pandemic fear	social distancing	0.97	58.50
	face covering	0.99	58.38
	avoid non-essential travel	1.00	–
Delivery convenience	ease of receiving delivery when not home	1.06	52.91
	overall ease of receiving delivery	1.00	–
Product quality	high quality perishable products is important	1.01	47.33
	always search for high quality products	1.00	–
	log-likelihood	−13,457	
	log-likelihood - null model	−21,306	
	McFadden's Rho- square	0.37	

5.2.2. The choice model component

The SCIAL model is then sequentially estimated with latent variables identified in the previous step. The final specification and goodness-of-fit of the SCIAL Logit model with the latent variable are presented in Table 5. A multinomial logit (MNL) model is also presented in Table 5 for the purpose of reference. The SCIAL model has an Akaike Information Criterion (AIC) value of 17,738 and a Bayesian Information Criterion (BIC) value of 17,658. The MNL model has an AIC value of 18,389 and a BIC value of 18,355. The model with lower AIC and BIC values should have superior specification parsimony and dataset fitting. The SCIAL model outperforms the MNL model formulation. McFadden's Rho-square value for the SCIAL model, is 0.24, indicating a good model fit, especially given the complexity of the model formulation. Most parameters are statistically significant at the 95% confidence interval based on the one-tailed *t*-test. The final specification also keeps some parameters with t-statistics less than 1.64 since they represent expected behavioural effects.

The choice component examines the effects of level-of-service variables on consumers' shopping channel choices. Modelling results show that shoppers prefer to only shop for standardized items online and have them delivered to their homes. This finding is consistent with empirical findings reported by Chintagunta et al. (2012). Chintagunta et al. quantified and compared various consumer transaction costs between in-store and online grocery shopping channels. They found that consumers treated purchasing perishable items online as costlier (in terms of transaction cost) than purchasing in-store. They attributed the higher transaction cost in the online channel for perishable items to the inability to verify the product quality of perishable items. In this study, the modelling results also indicate that when ordering online and picking up groceries in-store, shoppers positively accepted purchasing some perishable items in their shopping baskets. The acceptance can be explained by the fact that in-store pick-up allows shoppers to perform quality checks on items purchased. It is easier to swap items with unsatisfied quality on-site. Although home delivery vendors might provide similar customer services. However, item swapping, returning, and disputing might take longer and costlier in online channels. The findings discussed above suggest a possible product mix strategy that online service might emphasize on standardized items instead of perishable items.

Shoppers are indifferent regarding types of service providers when grocery delivery is without subscriptions. While searching for the final specification, none of the parameters indicating service provider types were statistically significant for online & delivery alternatives with unsubscribed providers. However, while purchasing subscriptions with

Table 5
Estimated parameters of the choice and choice set generation component.

	SCIAL		MNL	
Choice model				
	Est.	t-stat	Est.	t-stat
Alternative specific constant (ASC)				
in-store shopping	–	–	–	–
online & delivery - unsubscribed provider	−1.93	−7.72	−1.46	−14.49
online & delivery - begin subscription	−2.52	−9.93	−2.22	−18.48
online & delivery - subscribed provider	1.39	1.51	−1.28	−13.16
online & pick-up	−0.82	−1.62	−2.14	−31.46
Basket characters				
only standardized items				
online & delivery - unsubscribed provider	0.35	3.77	0.24	3.54
online & delivery - begin subscription	0.29	2.20	0.19	1.72
online & pick-up	0.60	4.11	0.41	4.07
half standardized & half perishable				
online & pick-up	0.69	3.87	0.47	3.85
mostly perishable items				
online & delivery - subscribed provider	−0.62	−1.21	−0.31	−1.55
Third-part service provider				
online & delivery - begin subscription	−0.28	−2.73	−0.19	−2.17
Travel modes to stores				
transit				
in-store shopping	−0.59	−7.73	−0.42	−8.94
car				
online & pick-up	0.80	6.30	0.58	7.46
Delivery time				
same day	0.31	3.79	0.31	5.75
online & delivery - unsubscribed provider				
online & delivery - subscribed provider				
a week or later	−0.51	−8.14	−0.42	−9.09
online & delivery - unsubscribed provider				
online & delivery - begin subscription				
online & pick-up				
Logarithm of delivery/pick-up cost	−0.11	−3.34	−0.09	−3.74
online & delivery - unsubscribed provider				
online & delivery - subscribed provider				
online & delivery - begin subscription				
online & pick-up				
Logarithm of shopping basket price	−0.36	−11.04	−0.27	−13.33
in-store shopping				
online & delivery - unsubscribed provider				
online & delivery - subscribed provider				
online & delivery - begin subscription				
online & pick-up				
Logarithm of travel time to store	−0.25	−5.59	−0.18	−6.01
in-store shopping				
online & pick-up				
Choice set generation model				
	Est.	t-stat		
Alternative specific constant (ASC)				
in-store shopping	−2.32	−4.75		
online & delivery - unsubscribed provider	22.47	11.15		
online & delivery - begin subscription	20.00	7.58		
online & delivery - subscribed provider	−0.48	−1.65		
online & pick-up	−3.02	−5.62		
Logarithm of respondents' age				
in-store shopping	1.08	6.39		
online & delivery - unsubscribed provider	−5.29	−11.15		
online & delivery - begin subscription	−4.93	−7.85		
Respondents' gender as male				
in-store shopping	−0.55	−4.81		
online & delivery - unsubscribed provider	−1.07	−6.87		
online & delivery - begin subscription	−0.72	−2.96		
online & delivery - subscribed provider	−0.34	−4.26		
Logarithm of household size				
in-store shopping	0.52	5.17		
online & delivery - begin subscription	1.01	4.54		
Household income > $100,000				
online & delivery - unsubscribed provider	0.32	2.49		
online & delivery - begin subscription	−0.32	−1.69		
online & delivery - subscribed provider	0.38	4.78		
online & pick-up	0.15	2.50		
Logarithm of number of retail facilities in the TAZ where households locate				
online & delivery - unsubscribed provider	0.19	3.65		
online & delivery - begin subscription	0.18	2.11		
Population density in the TAZ where households locate				
online & delivery - subscribed provider	0.01	2.34		
Logarithm of number of vehicles				
online & delivery - subscribed provider	−0.30	−2.55		
Pandemic fear - latent variable				
online & pick-up	1.64	4.17		
log-likelihood		−8,829		−9,177
McFadden's Rho- square		0.24		0.21
AIC		17,738		18,389
BIC		17,658		18,355

Notes: AIC is $-2k + 2LL$, BIC is $-2LL + 2\log(n)k$, where n is the number of observations and k is the number of parameters.

entitlement to free delivery, shoppers prefer brand-owned vendors over thirty-party vendors. The consideration might include multiple reasons. First, shoppers might view brand-owned vendors as more reliable options than third-party vendors. The tangible assets (e.g., stores and inventories) and intangible but long-established brand reputations of grocery brands might be important contributing factors. Moreover, brand-owned home delivery vendors might be able to provide more efficient customer service. Irrespective of who delivered the groceries, the grocery stores where the items are sold must be involved in all consumer services, such as item swapping, returning, and disputing. Therefore, from the perspective of grocery shoppers, subscribing to direct services from stores is the most efficient and reliable option.

Shoppers prefer to use private vehicles if they need to travel to stores. The modelling results indicate that consumers dislike using transit for in-store grocery shopping and favor using private cars to pick up their grocery orders. Such travel mode preference indicates home delivery services could reduce vehicular trips to grocery stores from the consumers' end. Previous studies on grocery shopping channel choice often overlook consumers' travel modes to stores primarily due to data availability. Only Suel and Polak (2017) specifically investigated the effects of travel modes on grocery channel choices. They found that online grocery delivery mostly drew from driving trips instead of walking and transit trips. This study's results are consistent with Suel and Polak (2017). Moreover, logistically planned freight delivery could consolidate up to 60% of the distance travelled on roads, compared to one-to-one trips from stores to consumers' homes (Marcucci et al., 2021). This shows considerable potential for home delivery to reduce road congestion.

For delivery time, grocery shoppers value same-day delivery service and process heavy disutility if the wait time exceeds a week. The result is consistent with the literature that same-day delivery attracted grocery shoppers to use online channels (Xi et al., 2020). Interestingly, existing literature in shopping channel choice using SP choice experiments often presented delivery time to survey respondents as continuous values (e.g., hours or days) (Hsiao, 2009; Schmid and Axhausen, 2019; Marcucci et al., 2021). Their modelling results found respondents displayed disutility towards delivery time for groceries and other durable goods. However, this study demonstrates that differential delivery time can be branded as a premium. Same-day delivery can be introduced as a premium by differentiating it from delivery that takes longer than 24 h. This finding implies an important operation and marketing strategy. First, online grocery vendors should aim to provide same-day delivery service; otherwise, their business might lose attraction. Meanwhile, their marketing efforts should emphasize the term "same-day delivery" instead of the continuous values of time (e.g., hours), which might induce unwanted consumer disutility in their shopping channel choices, even if their online order could be fulfilled within 24 h.

Furthermore, this study also estimates grocery shoppers' Willingness to pay for same-day delivery services. The Willingness to pay under various delivery costs can be read in Fig. 3. Typical grocery delivery cost ranges from $4 to $20 in the GTA (see **Appendix A**). Within this range, the estimated Willingness to pay for same-day delivery varies from $3.91 to $8.44. More specifically, shoppers are willing to pay $6.49 for same-day delivery if their delivery charge is $10. The frontier of travel time saved by same-day delivery is also calculated and presented in Fig. 3. The calculation assumes same-day delivery and in-store grocery shopping are equivalent and interchangeable shopping channels. Also, shoppers will perform dedicated two-trip grocery shopping tours for in-store shopping. Hasnine and Habib (2019) estimated the values of travel time (VOT) in the GTA were $15.47 per hour for two-trip tours. The calculation of the delivery cost and travel time-saving frontier will stay with their reported VOT. The frontier indicates the market penetration potential and attractiveness of same-day grocery delivery services under various pricing schemes. For example, the same-day grocery delivery service charged $10 could potentially attract grocery shoppers with at least 25.2 min of dedicated two-trip grocery shopping tours.

All cost-related parameters, such as shopping basket price, delivery/pick-up cost, and travel time to stores, have expected negative signs. Modelling results show that consumers have different disutility toward basket price and delivery cost. Similar observations can be found in the research of Frischmann et al. (2012) and Marcucci et al. (2021). Resultantly, they suggested a "free shipping" pricing strategy exploiting the biased perception by merging delivery cost into basket price and offering consumers so-called free delivery. Marcucci et al. (2021) tested the "free shopping" pricing strategy and found that the market share of the home delivery alternative increased considerably from 10.9% to 24.7%. The "free shipping" strategy is also tested in this study. By merging delivery cost into basket price, the market share of the "online & delivery - unsubscribed provider" option increased from 23.0% to 24.8%. In contrast, the market share of the "in-store" option decreased from 55.4% to 54.1%. Although the market share gain for the "online & delivery - unsubscribed provider" option is less substantial than the results from Marcucci et al. (2021), the validity of the "free shipping" strategy is confirmed again in this study.

Direct marginal effects of delivery/pick-up cost charged by service providers and travel time to stores are calculated. Fig. 4 presents the direct marginal effects of delivery/pick-up cost. The results reflect considerable behavioural heterogeneity within the online grocery shopping channel. Grocery shoppers have greater price sensitivity when purchasing subscriptions. If the monthly payment doubles, their probability of buying subscriptions will be decreased by around 18%.

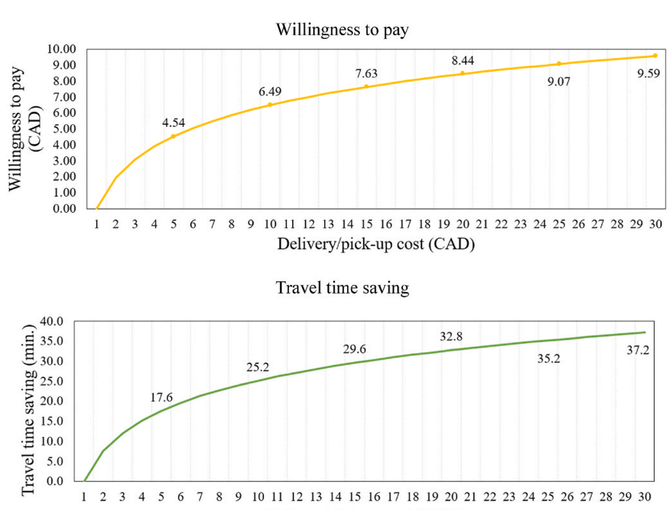

Fig. 3. Estimated Willingness to pay and travel time saving frontier for same day grocery delivery service.

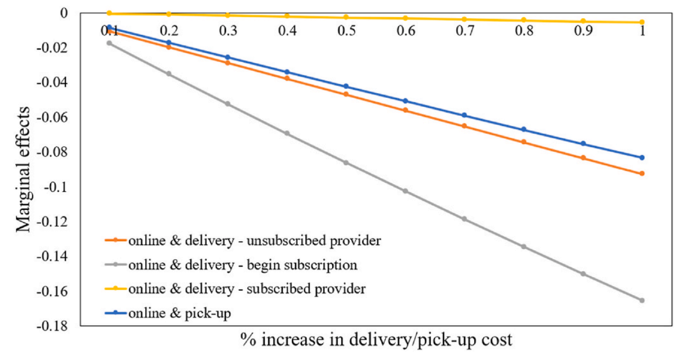

Fig. 4. Direct marginal effects of delivery/pick-up cost.

Shoppers demonstrate moderate price sensitivity when the service is charged per delivery. Suppose the one-time delivery fee doubles, and their probability of using grocery delivery service will be decreased by around 8%. However, for shoppers entitled to free delivery (sometimes under conditions like minimum order for free delivery), their probability of using grocery delivery service is almost irrelevant with changes in delivery cost.

Fig. 5 presents the direct marginal effects of travel time to store. In-store grocery shoppers process less travel time sensitivity compared to online & pick-up shoppers. When travel time doubles, the probability of choosing in-store shopping decreases by around 20%, compared to a 40% decrease for in-store and pick-up. Similar results were reported by Marcucci et al. (2021). They reported that online & pick-up shoppers had greater negative travel time elasticity than in-store shoppers. This fits the expectation since online & pick-up shoppers might order in advance and pick up in-store to save shopping time and bypass checkout lines. Therefore, they are more likely to be more time-sensitive than in-store grocery shoppers.

5.2.3. *The choice set generation component*

The SCIAL model effectively addresses the concept of probabilistic choice set formation by accommodating semi-compensatory choice-making behaviour. Fig. 6 presents the general effects of choice set inclusion probability on the choice probability.

The results indicate that grocery shoppers have the most myopic behaviour once they consider using subscribed free grocery delivery. If they firmly consider subscribing free grocery delivery as a feasible option (choice set inclusion probability = 100%), there is a 76% chance they will choose it. In-store shopping also demonstrates interesting behavioural insights. As the most popular grocery shopping channel, in-store shopping is at least 40% likely to be considered a feasible channel. Once shoppers firmly consider it in their choice set, there is a 75% chance they will shop in-store. The above-discussed findings suggest that shoppers have similar choice probability once they firmly consider in-store grocery shopping and free home grocery delivery in their choice

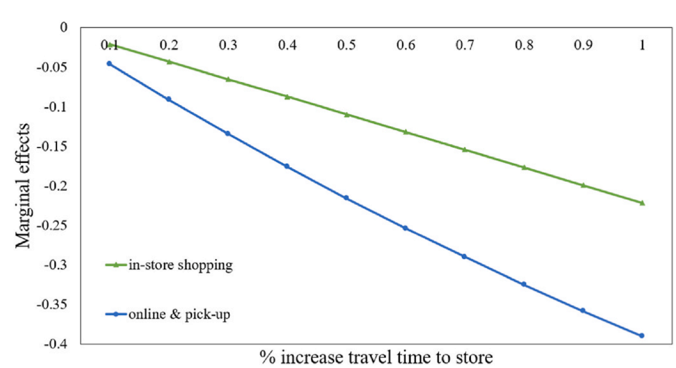

Fig. 5. Direct marginal effects of travel time to store.

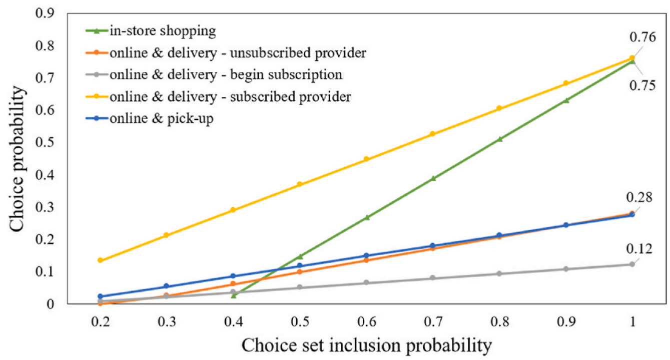

Fig. 6. General effects of choice set generation on choice probability.

set. However, considering subscribed free grocery delivery in the choice set is unavoidably embedded with extra cost. Grocery shoppers must pay a subscription fee. Discussion in **Section 5.2.2** shows that consumers demonstrate considerable sensitivity towards their monthly subscription payment (see estimated direct marginal effects in Fig. 4). On the other hand, in-store grocery shopping is the typical way of grocery shopping, free from any upfront channel selection cost. Any grocery shoppers could consider it as a feasible option. Therefore, the above findings indicate that in-store grocery shopping will still dominate. Unless all online shoppers could enjoy unlimited free grocery delivery service (not the "free shipping" pricing strategy discussed in **Section 5.2.2**). However, unconditional free grocery delivery is impossible without heavy financial subsidies, so the sustainability of the business model is questionable.

Estimated parameters in choice set generation components also shed light on the effects of explanatory variables on choice set inclusion probability. Results indicate that grocery shoppers' socioeconomic, land-use characteristics, and individuals' attitudes contribute to their choice set inclusion probability. Age is found to have a significant contribution. As age increases, grocery shoppers are more likely to consider in-store shopping in their choice set. Conversely, younger grocery shoppers are more likely to consider using home delivery services or purchasing subscriptions. Similar effects of age have been reported many times in the literature (Farag et al., 2006a; Beckers et al., 2018; Clarke et al., 2015; Zhen et al., 2018; Schmid and Axhausen, 2019; Wieland, 2022).

The modelling results propose a plausible explanation for the mixed effects of gender in grocery channel choices in the literature. Farag et al. (2006b) and Shen et al. (2022) reported that females were more likely to be online grocery shoppers. Conversely, Marcucci et al. (2021) reported that female was the explanatory variable attributing individuals' class membership probability to the in-store shopper class. Schmid and Axhausen (2019) found that gender had no significant effect on choosing online grocery shopping; even males were more likely to be pro-online. In this study, the modelling results from the choice set generation components reflect that male shopper generally have less choice set inclusion probability for all shopping channels than their female counterparts. Additionally, Fig. 7 compares relationships between choice set inclusion and the choice probability between males and females. It shows that for all shopping channels, the curves representing male shoppers are above those for female shoppers. This indicates that males are generally more myopic than females. In other words, given the same choice set inclusion probability, males are more likely to choose a particular shopping channel without comparison to other alternative channels. Therefore, the competition between the two mechanisms determines the mixed effects of gender in the grocery channel choice literature.

Larger households are more likely to consider in-store shopping and purchasing subscriptions simultaneously. The two options seem to be at the two ends of the choice set spectrum. However, it is reasonable to expect households with a larger size to consider them simultaneously. Grocery shopping demand should be positively related to household size. Therefore, large households could stick with the in-store channel, which is subjected to fewer additional transaction costs in monetary value. In this case, households use transportation costs (travel time & cost) and in-store shopping time to save delivery costs. On the other hand, purchasing services with unlimited free delivery is also an economically viable option. After the purchase, the households could fulfill all their grocery shopping needs with a fixed monthly payment. Moreover, they could save transportation costs and in-store shopping time for other activities.

Households with income over $100,000 are more likely to consider online grocery with home delivery. This is consistent with the literature that wealthy households were more likely to use online channels Clarke et al. (2015); Suel and Polak (2018); Shen et al. (2022). The negative parameter sign for the alternative to purchase subscriptions is that many high-income households have already purchased delivery subscriptions. The survey design does not allow existing subscribers to purchase dual subscriptions.

This study finds that both innovation-diffusion and efficiency hypotheses drive the adoption of online grocery shopping in the study area. Higher population density and retail accessibility contribute to the higher probability of considering online grocery alternatives. This supports the innovation-diffusion hypothesis (Anderson et al., 2003). As a relatively new channel, online grocery services are more likely to be considered by residents in the densely urban core, which also has greater accessibility to retail activities. Interestingly, higher vehicle accessibility inversely influences the probability of considering using subscribed home delivery services. This finding could also support the efficiency hypothesis (Anderson et al., 2003). Households with lower vehicle access are more likely to consider unlimited free grocery delivery services to complement their relatively lower mobility. This is also consistent with the discussion in **Section 5.2.2** that consumers in the study area prefer vehicular modes to travel for grocery shopping.

The readers should notice that there is a slight deviation from the original efficiency hypothesis. Anderson et al. (2003) stated the efficiency hypothesis that individuals living in suburban areas with low shopping accessibility could use online shopping to overcome spatial barriers. However, in this study, the inefficiency is reflected by low mobility caused by a lack of vehicle accessibility instead of land-use characteristics. Therefore, future research could also consider mobility-related attributes when testing the efficiency hypothesis.

Latent variables describing shoppers' perceived level of delivery convenience and pickiness of product quality are found statistically insignificant with choice inclusion probability. Therefore, they are irrelevant to grocery channel choice. However, perceived pandemic fear is found significantly contributes to the choice set inclusion probability of in-store grocery pick-up services. Grocery shoppers concerned with the pandemic are more likely to consider ordering their groceries online in advance and picking them up in-store. Doing so could avoid unwanted physical contact with individuals in the store. However, perceived pandemic fear is irrelevant with all home delivery alternatives' choice set inclusion probability. This finding reflects individuals' heterogeneous choice behaviour within the online shopping channel. Also, this finding indicates promising growth potential for grocery delivery services after the pandemic. Because their choice set considerations is irrelevant to pandemic fear, which will decay as time progresses to the post-pandemic era.

The following section proposes a characteristics-based investigation of the heterogeneity in the effects of choice set inclusion probability on the choice probability. The samples are separated by gender, age, population density, and retail accessibility, and their relationships between choice set inclusion and the choice probability are compared. For gender, males and females are divided into separate groups. For age, samples are divided by the threshold of 45 years old. Therefore, Millennials and Generation Z are put into one group (age younger than 45),

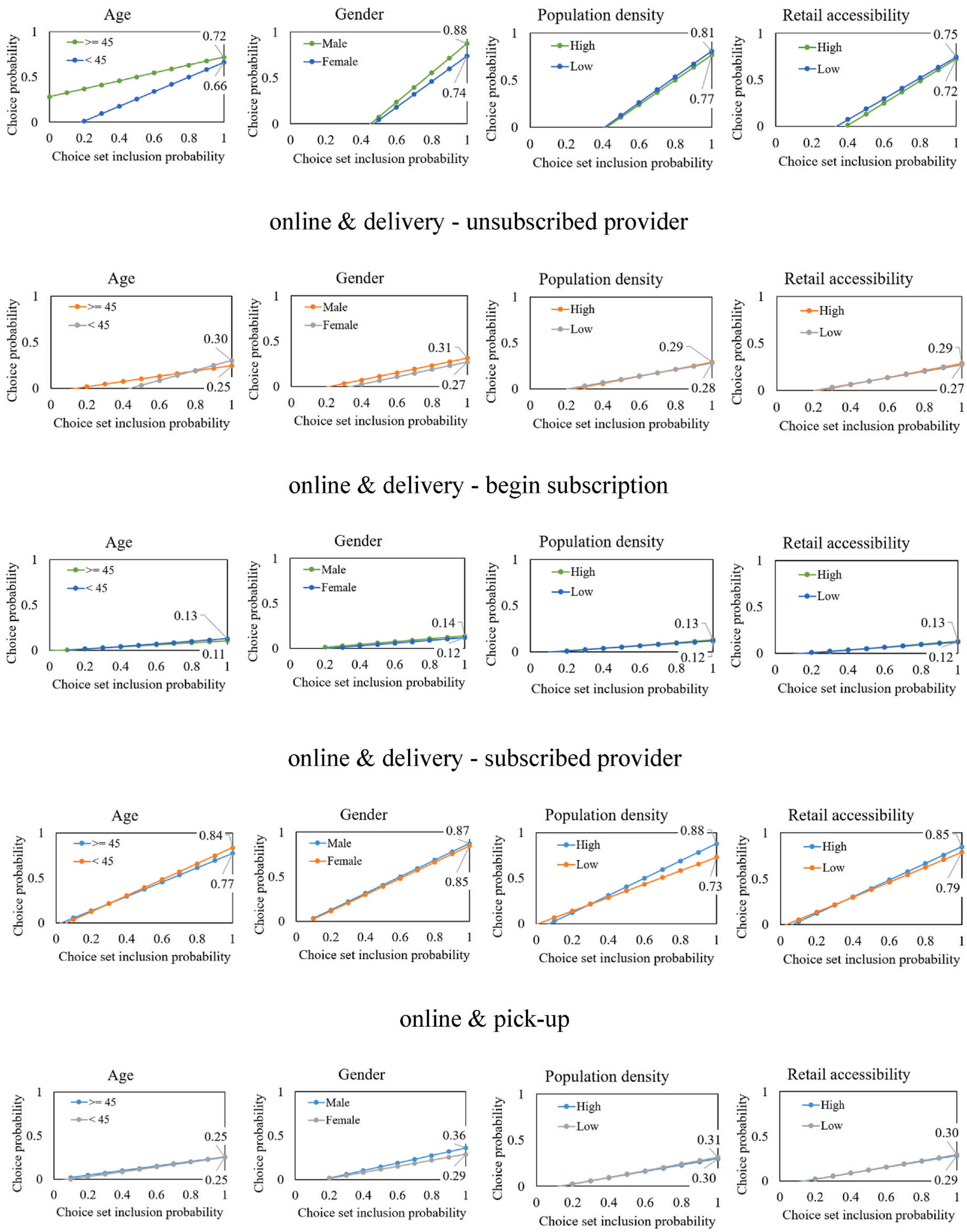

Fig. 7. Effects of choice set generation on choice probability by attributes.

and Generation X and Baby Boomers are put into another group (age older than 45). The effect of population density is compared by samples from the 10th and 90th percentile[1] in terms of population density in the households' TAZ. Similarly, retail accessibility is also compared by samples from the 10th and 90th percentile[2] of the number of retail facilities in the households' TAZ of households. Interestingly, all samples in the 10th percentile have zero retail facilities in their TAZs, meaning they must travel to other zones for grocery shopping. Fig. 7 presents the

[1] The 10th percentile has an average of 625 people per square kilometer. The 90th percentile has an average of 24,781 people per square kilometer.

[2] The 10th percentile has zero retail facility per TAZ. The 90th percentile has an average of 120.8 retail facilities per TAZ.

comparison. The curve with the steeper slope indicates that individuals are more likely to choose the channel without carefully evaluating the level of service attributes against those of alternative channels. Namely, they reflect the sub-groups with myopic choice behaviours. Results could guide marketing strategy for online channels on the target of sub-groups. Marketing campaigns would like to target myopic groups to attract more channel users, supposing marketing efforts could increase people's choice-set inclusion probability by creating channel awareness. Generally, online channels should aim at males, individuals younger than 45 years old, and individuals residing in densely urban areas.

After firmly considering a channel, males are more likely to choose all shopping channels than females. Individuals younger than 45 years old shared similar behaviours as their senior counterparts for the option to begin subscribed home delivery and online & pick-up. However, they are more likely to choose both unsubscribed and subscribed home delivery services than their senior counterparts. Controlling for the choice set inclusion probability, individuals living in high-population-density areas are more likely to use subscribed home delivery than individuals living in low-density areas. Similarly, individuals living in TAZs with a higher number of retail facilities are more likely to use subscribed home delivery than individuals living in TAZs with zero retail facilities.

6. Conclusion & future study

This paper presents an empirical investigation of grocery shoppers' shopping channel choices. The study uses stated preference (SP) experiments conducted with the 2021 Summer COVHITS survey in the GTA, Canada. The following theoretical implications should be considered by researchers and practitioners investigating individual choice behaviours. First, this study demonstrates the advantage of considering probabilistic choice set formation and semi-compensatory behaviour in modelling discrete choices. The Semi-Compensatory Independent Availability Logit (SCIAL) model used in this study allows investigation of the adoption process of novel grocery shopping channels. Like the adoption of many innovative products, the inertia of innovation diffusion could arise because many consumers do not consider novel products as feasible options. Classic full-compensatory discrete choice models could overlook this aspect of choice behaviour.

Moreover, this study proposes a systematic arrangement of socioeconomic, latent psychological and level-of-service (LOS) variables in the specification of systematic utility function for discrete choice models. In typical hybrid choice models, all types of explanatory variables were placed into one systematic utility function (Bhat and Dubey, 2014; Ben-Akiva and Boccara, 1995; Habib and Zaman, 2012; Vij and Walker, 2016). This approach risks overestimating the effects of level-of-service attributes in choice-making behaviour. Logically, choice makers will never evaluate the LOS attributes for alternatives not in their choice set. The above-mentioned approach overlooked it and assumed LOS attributes of all alternatives (deemed to be available by modellers) would be thoroughly evaluated by choice makers. Conversely, the modelling methodology presented in this paper assumes a systemic decision-making process to account for this issue. The methodology postulates that while making choices, individuals first decide their choice set, then evaluate and choose the best alternative they consider feasible. Resultantly, the modelling framework postulates socioeconomic and latent psychological variables affecting choice set formation. Then, LOS attributes of alternatives considered in the choice set determine the final choice. In the SCIAL model, the choice set formation component will penalize the systematic utility function of infeasible alternatives through semi-compensatory behaviour.

The empirical model reveals determinants of the grocery shopping channel choice. Modelling results indicate that shoppers prefer to only shop for standardized items online and have them delivered to their homes. Regarding types of grocery delivery providers, shoppers are indifferent between brand-owned and third-party providers when using grocery delivery services without subscriptions. However, shoppers prefer brand-owned over thirty-party vendors when purchasing subscribed free delivery services. Furthermore, shoppers prefer to use private vehicles if they need to visit stores. This indicates the potential for home delivery services to reduce driving trips to stores.

Grocery shoppers value same-day grocery delivery services. Previous literature in shopping channel choice often reported delivery time as disutility when the message is presented to survey respondents as continuous values (e.g., hours or days) (Hsiao, 2009; Schmid and Axhausen, 2019; Marcucci et al., 2021). However, this study finds that same-day delivery can be introduced as a premium by differentiating it from delivery that takes longer than 24 h. Moreover, this study estimates the Willingness to pay for same-day delivery service. Shoppers are willing to pay between \$3.91 and \$8.44 for same-day delivery, for typical delivery services charged between \$4 and \$20 in the Greater Toronto Area (GTA). From the perspective of travel behaviours, this study also calculates the relationships between delivery fees charged for same-day grocery delivery service and travel time saved to stores. If shoppers are willing to pay \$10 for their same-day grocery delivery service, they are expected to save at least 25.2 min of travel time from their dedicated shopping trips.

The modelling results reveal behavioural heterogeneity within the online grocery shopping channel. Grocery shoppers have greater price sensitivity when purchasing subscription-based grocery delivery services. Shoppers demonstrate moderate price sensitivity when the service is charged per delivery. However, for shoppers already entitled to free delivery, their probability of using grocery delivery service is almost irrelevant with changes in delivery cost. This highlights the discussion by Suel and Polak (2018) on the risk of channel choice study focus on aggregated channel choice and overlooks within channel heterogeneity. Future choice channel studies should consider alternatives within each channel instead of merely focusing on binary choice between online and offline channels.

For choice set inclusion probability, modelling results also indicate that grocery shoppers' age, gender, household size, income, vehicle accessibility, land-use characteristics, and perceived pandemic fear contribute to grocery shoppers' choice set formation. Aged shoppers are more likely to consider in-store shopping in their choice set. This study finds that male shoppers generally have less choice set inclusion probability than females, regardless of shopping channels. However, males are also generally more myopic than females in all shopping channels. In other words, males are more likely to choose a particular shopping channel without comparing it to other channels. The competition of the two effects might lead to a mixed effect of gender in the literature of shopping channel choice (Farag et al., 2006b; Shen et al., 2022; Marcucci et al., 2021; Schmid and Axhausen, 2019). Larger households are more likely to consider in-store shopping and purchasing subscriptions with unlimited free delivery simultaneously.

The results of this study show that the adoption of online grocery shopping is driven by both the innovation-diffusion and efficiency hypotheses (Anderson et al., 2003). Higher retail accessibility and population density contribute to higher choice set inclusion probability for grocery delivery services. This supports the innovation-diffusion hypothesis that online grocery shopping would diffuse from the densely populated urban area first. On the other hand, households with lower vehicle access have the tendency to consider unlimited free grocery delivery services to compensate for their relatively low mobility. This finding also supports the efficiency hypothesis. However, in this case, the inefficiency is measured by mobility instead of land use characteristics.

Besides socioeconomic variables, this study considers three latent factors: individuals' perceived pandemic fear, level of delivery convenience, and their pickiness of the quality of grocery products. The study finds that shoppers' perceived level of delivery convenience and pickiness of product quality are statistically insignificant to their probability of considering grocery delivery services. However, shoppers with higher perceived pandemic fear are more likely to consider online ordering and in-store pick-up services as viable options.

This study also investigates the relationship between the choice set inclusion probability and the choice probability. The results show that grocery shoppers exhibit myopic behaviours once they consider using subscribed free grocery delivery as a viable option. If the choice set inclusion probability is 100% (firmly consider subscribed free grocery delivery in the choice set), there is a 76% chance they will choose this option. On the other hand, once shoppers firmly consider in-store shopping, there is a 75% chance they will shop in-store. However, effectively considering subscribed free grocery delivery as a feasible option is costly. Instead, in-store shopping could be considered by any individual. Therefore, it is safe to conclude that in-store grocery shopping will still be the dominant grocery shopping channel because firmly consider subscribed free grocery delivery service is much costlier (having to pay monthly subscription fees) than in-store shopping.

As with any research, there are limitations of this study. A notable limitation of this study comes from the general nature of SP experiments. The choice variation reflected from SP data is unavoidably different from revealed preference (RP) data (Ortúzar and Willumsen, 2011). The technique of RP-SP modelling could be used to correct variance in SP data. However, as discussed in **Section 1.0**, it is incredibly challenging to obtain RP data, especially the LOS attributes for unchosen channels, in shopping channel choices. In fact, this is the primary reason to use the SP technique to investigate shopping channel choice. Also, in SP experiments, respondents might make decisions based on factors not shown in the hypothetical scenario. Such hidden factors might not be known to researchers, so relevant information is hard to collect. Another limitation is that the dataset was collected before the world economy head into stagnation in 2022 (Gilchrist, 2022). Compared to May 2021, grocery prices rose 9.7% in May 2022 (Statistics Canada, 2022b). This was the largest annual increase since 1983. Significant inflation might affect shoppers' behaviours, such as the Willingness to pay for same-day delivery. Grocery shoppers might become increasingly sensitive to spending, considering the rising price and their limited budget. Nonetheless, significant inflation should be temporary, and general trends and behavioural interpretations reported in this study should still be consistent. Similar research should be conducted in the nearest future to capture the influence of stagnation, and this study could serve as a reference.

Data availability

The data that has been used is confidential.

Acknowledgment

The study was funded by an NSERC Discovery Grant and Percy Edward Hart Professorship Grant. The authors bear the sole responsibility for all results, interpretations, and comments made in the paper.

Appendix A Summary of online grocery delivery services in the Greater Toronto Area

	Instacart	BUGGY	PC Express	Cornershop - Uber	Walmart	Longo's	Costco	Local stores
Home delivery service	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes for all
In-store pick-up	Yes	No	Yes	No	Yes	Yes	No	Varies by stores
Service type	Type 1	Type 1	Type 2	Type 1	Type 2	Type 2	Type 2	Type 2
Product type	Product 1	Product 1&3	Product 1	Product 1	Product 1	Product 1	Product 2	All stores serve product 2
Minimum order for delivery	$10	N/A	$30	N/A	$35	$50	N/A	$0-$99
Minimum delivery fee	$3.99	$9.99	$3.00	$6.90	$7.95	$9.99	$3 per unit	$2.99-$13.99
Maximum delivery fee	$7.99	$19.98	$5.00	$9.90	&14.97	$14.95	$3 per unit	$2.99-$17.99
Membership subscription fee	$99 annual fee.	N/A	$119 annual fee.	$99 annual fee.	$98 annual fee.	$139.99 annual fee.	$60 - $120 annual fee. Membership is mandatory to use the service.	No membership for all stores. However, most of them provide free delivery for order over certain amount ($49-$99).
Incentives with subscription	Free delivery for order over $35	N/A	Free in store pick-up service with priority time slots.	Free delivery on all orders over $40. $4.90 on all orders under $40.	Free same day grocery delivery.	Free unlimited delivery	Free delivery for orders over $75.	
Minimum wait window	2 h	1 h	4 h	1.5 h	2 h	1 day	2 days	1 h – next day
Maximum wait window	5 days	16 days	14 days	2 days	8 days	4 days	10 days	Next day – 21 days
Source	https://www.howtosavemoney.ca/reviews/pc-express	https://www.inabuggy.com/FAQ	https://www.pcoptimum.ca/insiders/en/	https://cornershopapp.com/en-ca/faq	https://www.walmart.com/cp/express-delivery/3696472	https://www.grocerygateway.com/store/membership	https://www.costco.ca/CanadaGroceryDeliveryRedirect	

Notes: (1) Type 1 means the service provider is a third-party company serving multiple brands. (2) Type 2 means the service is operated by the brand itself serving its stores only. (3) Product 1 means the service deliver all types of groceries including perishable products. (4) Product 2 means service deliver non-perishable products only with selective items marked by the store. (5) Product 3 means service deliver alcohol to qualified customers. (6) The currency is Canadian dollar.

References

Aaker, D.A., Jones, J.M., 1971. Modeling store choice behavior. J. Market. Res. 8 (1), 38–42. https://doi.org/10.1177/002224377100800104.

AptechSystems, 2014. Gauss user's manual. http://www.aptech.com.

Anderson, W.P., Chatterjee, L., Lakshmanan, T., 2003. E-Commerce, transportation, and economic geography. Growth Change 34 (4), 415–432. https://doi.org/10.1046/j.0017-4815.2003.00228.x.

Beckers, J., Cárdenas, I., Verhetsel, A., 2018. Identifying the geography of online shopping adoption in Belgium. J. Retailing Consum. Serv. 45, 33–41. https://doi.org/10.1016/j.jretconser.2018.08.006.

Ben-Akiva, M., Boccara, B., 1995. Discrete choice models with latent choice sets. Int. J. Res. Market. 12 (1), 9–24. https://doi.org/10.1016/0167-8116(95)00002-J.

Bhat, C.R., Dubey, S.K., 2014. A new estimation approach to integrate latent psychological constructs in choice modeling. Transp. Res. Part B Methodol. 67, 68–85. https://doi.org/10.1016/j.trb.2014.04.011.

Bhatnagar, A., Ratchford, B.T., 2004. A model of retail format competition for non-durable goods. Int. J. Res. Market. 21 (1), 39–59. https://doi.org/10.1016/j.ijresmar.2003.05.002.

Cao, X., 2009. E-shopping, spatial attributes, and personal travel: a review of empirical studies. Transport. Res. Rec. 2135, 160–169. https://doi.org/10.3141/2135-19.

Cao, X., 2012. The relationships between e-shopping and store shopping in the shopping process of search goods. Transport. Res. Pol. Pract. 46 (7), 993–1002. https://doi.org/10.1016/j.tra.2012.04.007.

Cao, X., Douma, F., Cleaveland, F., Cao, X., Douma, F., 2010. Influence of E-shopping on shopping travel evidence from Minnesota's Twin Cities. Transport. Res. Rec.: J. Transport. Res. Board 147–154. https://doi.org/10.3141/2157-18.

Chintagunta, P.K., Chu, J., Cebollada, J., 2012. Quantifying transaction costs in online/off-line grocery channel choice. January Market. Sci. 31 (1), 96–1*14*. https://doi.org/10.1287/mksc.1110.0678.

Clarke, G., Thompson, C., Birkin, M., 2015. The emerging geography of e-commerce in British retailing. January Regional Studies, Regional Sci*ence* 2 (1), 371–391. https://doi.org/10.1080/21681376.2015.1054420.

ChoiceMetrics, 2018. Ngene 1.2 user manual & reference guide. http://www.choice-metrics.com/NgeneManual120.pdf.

Choo, S., Lee, T., Mokhtarian, P.L., 2007. Do transportation and communications tend to be substitutes, complements, or neither? U.S. consumer expenditures perspective, 1984-2002. Transport. Res. Rec. 2010, 121–132. https://doi.org/10.3141/2010–14.

Data Management Group, 2018. 2016 TTS: data guide (issue February). http://dmg.utoronto.ca/pdf/tts/2016/2016TTS_DataGuide.pdf.

Dias, F.F., Lavieri, P.S., Sharda, S., Khoeini, S., Bhat, C.R., Pendyala, R.M., Pinjari, A.R., Ramadurai, G., Srinivasan, K.K., 2020. A comparison of online and in-person activity engagement: the case of shopping and eating meals. Transport. Res. C Emerg. Technol. 114 (March), 643–656. https://doi.org/10.1016/j.trc.2020.02.023.

Farag, S., Dijst, M., Lanzendorf, M., 2003. Exploring the use of E-shopping and its impact on personal travel behavior in The Netherlands. Transport. Res. Rec. 1858, 47–54. https://doi.org/10.3141/1858-07.

Farag, S., Schwanen, T., Dijst, M., 2005. Empirical investigation of online searching and buying and their relationship to shopping trips. Transport. Res. Rec. 1926, 242–251. https://doi.org/10.3141/1926-28.

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. Transport. Res. Pol. Pract. 41 (2), 125–141. https://doi.org/10.1016/j.tra.2006.02.003.

Farag, S., Weltevreden, J., van Rietbergen, T., Dijst, M., van Oort, F., 2006a. E-shopping in The Netherlands: does geography matter? Environ. Plann. Plann. Des. 33 (1), 59–74. https://doi.org/10.1068/b31083.

Farag, S., Krizek, K.J., Dijst, M., 2006b. E-Shopping and its relationship with in-store shopping: empirical evidence from The Netherlands and the USA. Transport Rev. 26 (1), 43–61. https://doi.org/10.1080/01441640500158496.

Frischmann, T., Hinz, O., Skiera, B., 2012. Retailers' use of shipping cost strategies: free shipping or partitioned prices? April Int. J. Electron. Commer. 16 (3), 65–88. https://doi.org/10.2753/jec1086-4415160303.

Gatta, V., Marcucci, E., Maltese, I., Iannaccone, G., Fan, J., 2021. E-Groceries: a Channel choice analysis in Shanghai. MDPI AG Sustainability 13 (7), 3625. https://doi.org/10.3390/su13073625. Retrieved from.

Gilchrist, K., 2022. World Bank Slashes Global Growth Forecast to 2.9%, Warns of 1970s-Style Stagflation. CNBC. https://www.cnbc.com/2022/06/07/world-bank-cuts-global-growth-outlook-and-warns-of-70s-stagflation.html.

González-Benito, Ó., 2002. Overcoming data limitations for store choice modelling. J. Retailing Consum. Serv. 9 (5), 259–268. https://doi.org/10.1016/s0969-6989(01)00035-2.

Habib, K.M.N., Zaman, M.H., 2012. Effects of incorporating latent and attitudinal information in mode choice models. Transport. Plann. Technol. 35 (5), 561–576. https://doi.org/10.1080/03081060.2012.701815.

Habib, K.M.N., 2019. Mode choice modelling for hailable rides: an investigation of the competition of Uber with other modes by using an integrated non-compensatory choice model with probabilistic choice set formation. Transport. Res. Pol. Pract. 129 (August), 205–216. https://doi.org/10.1016/j.tra.2019.08.014.

Hair, J., Black, W., Babin, B., Anderson, R., Tatham, R., 2006. Multivariate Data Analysis, 6th. Pearson Prentice Hall.

Handy, S.L., Yantis, T., 1997. The Impacts of Telecommunications Technologies on Nonwork Travel Behavior (No. SWUTC/97/721927-1). Center for Transportation Research The University of Texas at Austin. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.248.3797&rep=rep1&type=pdf.

Hasnine, M.S., Habib, K.N., 2019. A Dynamic Discrete Choice Model for Tour-Based Mode Choices. Transportation Research Board 98th Annual Meeting, Washington, DC.

Hsiao, M.H., 2009. Shopping mode choice: physical store shopping versus e-shopping. Transport. Res. E Logist. Transport. Rev. 45 (1), 86–95. https://doi.org/10.1016/j.tre.2008.06.002.

Kocur, G., Adler, T., Hyman, W., Aunet, B., 1981. Guide to Forecasting Travel with Direct Utility Assessment (No. UMTA-NH-11-0001-82-1Res. Rpt). Urban Mass Transportation Administration, Washington, DC.

Marcucci, E., Gatta, V., Le Pira, M., Chao, T., Li, S., 2021. Bricks or clicks? Consumer channel choice and its transport and environmental implications for the grocery market in Norway. Cities 110, 103046. https://doi.org/10.1016/j.cities.2020.103046.

Manski, C.F., 1977. The structure of random utility models. Theor. Decis. 8 (3), 229. Fscholarly-journals%2Fstructure-random-utility-models%2Fdocview%2F1303217712%2Fse-2%3Faccountid%3D14771.

Martínez, F., Aguila, F., Hurtubia, R., 2009. The constrained multinomial logit: a semi-compensatory choice model. Transp. Res. Part B Methodol. 43 (3), 365–377. https://doi.org/10.1016/j.trb.2008.06.006.

Mashrur, S., Wang, K., Loa, P., Hossain, S., 2022. Application of Protection Motivation Theory to Quantify the Impact of Pandemic Fear on Anticipated Postpandemic Transit Usage. https://doi.org/10.1177/03611981211065439.

McDonald, R.P., 1999. Test Theory: A Unified Treatment. Taylor & Francis, New York.

McKinsey, 2022a. The State of Grocery Retail 2022 Europe. https://www.mckinsey.com/~/media/mckinsey/industries/retail/our insights/state of grocery europe 2022/navigating-the-market-headwinds-the-state-of-grocery-retail-2022-europe.pdf?shouldIndex=false.

McKinsey, 2022b. The State of Grocery in North America. https://www.mckinsey.com/industries/retail/our-insights/the-state-of-grocery-in-north-america-2022.

McKinsey, 2022c. Disruption and Uncertainty: the State of Grocery 2021 China. https://www.mckinsey.com/~/media/mckinsey/industries/retail/our%20insights/disruption%20and%20uncertainty%20the%20state%20of%20grocery%202022%20china/disruption-and-uncertainty-the-state-of-grocery-2022-china.pdf.

Miller, E.J., 2021. Traffic Analysis Zone Definition: Issues & Guidance. Travel Modelling Group. https://tmg.utoronto.ca/files/Reports/Traffic-Zone-Guidance_March-2021_Final.pd.

Mokhtarian, Patricia, L., 1990. A typology of relationships between telecommunications and transportation. Transport. Res. Gen. 24 (3), 231–242. https://doi.org/10.1016/0191-2607(90)90060-J.

Mokhtarian, P.L., Salomon, I., Handy, S.L., 2006. The impacts of ict on leisure activities and travel: a conceptual exploration. Transportation 33 (3), 263–289. https://doi.org/10.1007/s11116-005-2305-6.

Ortúzar, J. de D., Willumsen, L.G., 2011. Modelling Transport, fourth ed. Wiley-Blackwell.

Raveau, S., Álvarez-Daziano, R., Yáñez, M.F., Bolduc, D., de Dios Ortúzar, J., 2010. Sequential and simultaneous estimation of hybrid discrete choice models: some new findings. Transport. Res. Rec. 2156 (1), 131–139. https://doi.org/10.3141/2156-15.

Reutterer, T., Teller, C., 2009. Store format choice and shopping trip types. Int. J. Retail Distrib. Manag. 37 (8), 695–710. https://doi.org/10.1108/09590550910966196.

Rotem-Mindali, O., Salomon, I., 2007. The impacts of E-retail on the choice of shopping trips and delivery: some preliminary findings. Transport. Res. Pol. Pract. 41 (2), 176–189. https://doi.org/10.1016/j.tra.2006.02.007.

Rosseel, Y, 2012. lavaan: an R package for structural equation modeling. J. Stat. Software. https://doi.org/10.18637/jss.v048.i02.

Rossetti, T., Yoon, S., Daziano, R.A., 2022. Social distancing and store choice in times of a pandemic. J. Retailing Consum. Serv. 65, 102860 https://doi.org/10.1016/j.jretconser.2021.102860.

Salomon, I., 1986. Telecommunications and travel relationships: a review. Transport. Res. Gen. 20 (3), 223–238. https://doi.org/10.1016/0191-2607(86)90096-8.

Schmid, B., Axhausen, K.W., 2019. In-store or online shopping of search and experience goods: a hybrid choice approach. Journal of Choice Modelling 31 (January 2018), 156–180. https://doi.org/10.1016/j.jocm.2018.03.001.

Shamshiripour, A., Rahimi, E., Shabanpour, R., Mohammadian, A., Kouros), 2020. How is COVID-19 reshaping activity-travel behavior? Evidence from a comprehensive survey in Chicago. Transp. Res. Interdiscip. Perspect. 7, 100216 https://doi.org/10.1016/j.trip.2020.100216.

Shen, H., Namdarpour, F., Lin, J., 2022. Investigation of online grocery shopping and delivery preference before, during, and after COVID-19. Transp. Res. Interdiscip. Perspect. 14 (November 2021), 100580 https://doi.org/10.1016/j.trip.2022.100580.

Sim, L.L., Koi, S.M., 2002. Singapore's internet shoppers and their impact on traditional shopping patterns. J. Retailing Consum. Serv. 9 (2), 115–124. https://doi.org/10.1016/S0969-6989(01)00029-7.

Statista, 2022. Canada: Online Usage Penetration 2018-2027 from. https://www.statista.com/statistics/373936/canada-online-penetration/. (Accessed 18 October 2022).

Suel, E., Le Vine, S., Polak, J., 2015. Empirical application of expenditure diary instrument to quantify relationships between in-store and online grocery shopping: case study of greater London. Transport. Res. Rec. 2496, 45–54. https://doi.org/10.3141/2496-06.

Suel, E., Polak, J.W., 2017. Development of joint models for channel, store, and travel mode choice: grocery shopping in London. Transport. Res. Pol. Pract. 99, 147–162. https://doi.org/10.1016/j.tra.2017.03.009.

Suel, E., Polak, J.W., 2018. Incorporating online shopping into travel demand modelling: challenges, progress, and opportunities. Transport Rev. 38 (5), 576–601. https://doi.org/10.1080/01441647.2017.1381864.

Stevens, J., 2002. Applied Multivariate Statistics for the Social Sciences, 4th. Lawrence Erlbaum Associates.
Suel, E., Daina, N., Polak, J.W., 2018. A hazard-based approach to modelling the effects of online shopping on intershopping duration. Transportation 45 (2), 415–428. https://doi.org/10.1007/s11116-017-9838-3.
Swait, J., Ben-Akiva, M., 1987. Incorporating random constraints in discrete models of choice set generation. Transp. Res. Part B Methodol. 21 (2), 91–102. https://doi.org/10.1016/0191-2615(87)90009-9.
Tahlyan, D., Said, M., Mahmassani, H., Stathopoulos, A., Walker, J., Shaheen, S., 2022. For whom did telework not work during the Pandemic? understanding the factors impacting telework satisfaction in the US using a multiple indicator multiple cause (MIMIC) model. Transport. Res. Pol. Pract. https://doi.org/10.1016/j.tra.2021.11.025.
Tavakol, M., Dennick, R., 2011. Making sense of Cronbach's alpha. Int. J. Med. Educ. https://doi.org/10.5116/ijme.4dfb.8dfd.
Tonn, B.E., Hemrick, A., 2004. Impacts of the use of E-mail and the internet on personal trip-making behavior. Soc. Sci. Comput. Rev. 22 (2), 270–280. https://doi.org/10.1177/0894439303262581.
Vij, A., Walker, J.L., 2016. How, when and why integrated choice and latent variable models are latently useful. Transp. Res. Part B Methodol. 90, 192–217. https://doi.org/10.1016/j.trb.2016.04.021.
Wang, K., Liu, Y., Mashrur, S.M., Loa, P., Habib, K.N., 2021. COVid-19 influenced households' Interrupted Travel Schedules (COVHITS) survey: lessons from the fall 2020 cycle. Transport Pol. 112, 43–62. https://doi.org/10.1016/j.tranpol.2021.08.009.
Wieland, T., 2022. Pandemic shopping behavior: did voluntary behavioral changes during the COVID-19 pandemic increase the competition between online retailers and physical retail locations? Papers in Applied Geography 1–19. https://doi.org/10.1080/23754931.2022.2106581.
Xi, G., Cao, X., Zhen, F., 2020. The impacts of same day delivery online shopping on local store shopping in Nanjing, China. Transport. Res. Pol. Pract. 136, 35–47. https://doi.org/10.1016/j.tra.2020.03.030.
Zhen, F., Du, X., Cao, J., Mokhtarian, P.L., 2018. The association between spatial attributes and e-shopping in the shopping process for search goods and experience goods: evidence from Nanjing. J. Transport Geogr. 66, 291–299. https://doi.org/10.1016/j.jtrangeo.2017.11.007.