The influence of E-shopping on Shopping Travel: Evidence from Twin Cities

Xinyu (Jason) Cao
Humphrey Institute of Public Affairs, University of Minnesota, 301 19th Ave S Minneapolis MN
612-625-5671 cao@umn.edu

Frank Douma
Humphrey Institute of Public Affairs, University of Minnesota, 301 19th Ave S Minneapolis MN
612-626-9946 douma002@umn.edu

Fay Cleaveland
Minnesota Department of Transportation, 395 John Ireland Boulevard, St. Paul, MN 55155
651-366-4194

Abstract

This research aims to reveal travel impacts of e-shopping in the Minneapolis-St Paul metropolitan area. Specifically, using a sample of internet users that drawn from urban, suburban, and exurban neighborhoods, we sought to identify the relationship between e-shopping and in-store shopping. An on-line survey composed of direct and attitudinal questions was used to obtain the data. Ordered probit models were developed to account for the influences of a variety of confounding factors such as shopping attitudes, shopping accessibility, shopping responsibility, and socio-demographics. The preliminary results showed that e-shopping behavior (in terms of online searching and online buying) tends to have a complementarity effect on in-store shopping, controlling for the confounding factors.

Words: 5000

Tables and Figures: 5

Key words: ICT, telecommunication, teleshopping, travel demand management

1. Introduction

E-shopping has the potential to replace traditional in-store shopping. It is well perceived that information and communication technologies (ICTs) have had pervasive impacts on modern society - they are changing how and where we work, shop, and in other ways live our lives. Significant research has been conducted to understand the impact of ICTs on where work is done and how this affects travel. Recently, e-shopping has become a centerpiece of ICTs because of its unprecedented proliferation. In the U.S., online retail spending grew by 19% a year to $136.4 billion in 2007, which accounted for 4% of total retail sales (InternetRetailer.com, 2008). In the Netherlands, online sales have increased from less than €200 million in 1999 to more than €1.6 billion in 2004, as Farag (Farag, 2006) reported. The growth of e-shopping has reshaped consumers' shopping behavior. Online buying could be a substitute for traditional shopping media, and may well dominate the exchange of certain products (e.g., digital assets) in the future (Cao and Mokhtarian, 2005). E-shopping as used here refers to the business-to-consumer (B2C) segment of e-commerce (Mokhtarian, 2004), i.e. product information search (online searching) and product transactions (online buying/purchasing) via internet, unless otherwise indicated.

With the rise of e-shopping and its long-term potential in the retail industry, transportation planners are interested in the changes that it will bring about to transportation systems. In particular, they focus on the effects of e-shopping on individuals' activity-travel patterns and on freight transport. Changes in freight transport can result from the growth of delivery trips to end consumers and the bypass of wholesalers and retailers in the network from manufacturers to consumers (Anderson et al., 2003). The potential of online buying to substitute for traditional in-store shopping and reduce personal shopping travel has important implications for travel demand management and congestion mitigation. According to the 2001 National Household Transportation Survey (NHTS), on average shopping travel accounted for 14.4% of annual VMT per household and 21.1% of annual vehicle trips per household (Hu and Reuscher, 2004). Therefore, the growth of online buying could have the potential to reduce traffic if it does replace physical shopping. On the other hand, if e-shopping induces new shopping trips, it is likely to generate more personal travel to existing transportation systems.

Although many studies have been conducted to model the adoption of e-shopping (Chang et al., 2005), little empirical work has been conducted to relate e-shopping to travel behavior and the existing findings are limited to few geographical areas. Using 591 adult internet users living in the Minneapolis-St. Paul metropolitan area, this study aims to provide further insight into the relationships between e-shopping and in-store shopping. Specifically, a series of ordered probit models were developed to investigate their connections while controlling for a number of confounding factors. This research addresses the following central question: how, and to what extent, does e-shopping impact individuals' physical shopping? The paper is organized as follows. The next section summarizes studies on the relationships between e-shopping and in-store shopping. Section 3 describes the data, variables, and analysis approach. Section 4 presents model results. The last section recapitulates the findings and makes recommendations for future research.

2. Literature Review

Conceptual and empirical studies in the field of ICTs and transportation suggest that e-shopping may interact with travel behavior in four ways: substitution, complementarity, modification, and neutrality (Choo et al., 2007; Mokhtarian, 1990; Pendyala et al., 1991; Salomon, 1986), although the connections can be even more complex (Mokhtarian, 2004). In the context of e-shopping, s*ubstitution* denotes that a physical trip to traditional stores is replaced by an online transaction; *complementarity* means that e-shopping generate new demands for trips to stores, either because of information obtained online or resulting from an online purchase; *modification* denotes that e-shopping does not affect the amount of physical travel to stores but changes the characteristics of trips such as timing and chaining; n*eutrality* means that e-shopping is independent of traditional shopping. Among these effects, substitution carries important positive implications for congestion mitigation. Planners are eager to disentangle the complex connections between e-shopping and travel behavior, and to measure the degree of substitution.

To assess the effects of e-shopping on individuals' shopping travel, why not just ask them? Several studies have found consistent substitution effects across many countries. Sim and Koi (Sim and Koi, 2002) stated that 12% of 175 Singapore online buyers had reduced their travel to stores although the remainder did not feel any influence. Tonn and Hemrick (Tonn and Hemrick, 2004) found that some internet users in the Knoxville metropolitan region (TN, U.S.) have

reduced trips to stores although a smaller percent indicated generating new trips. Weltervreden and van Rietbergen (Weltevreden and Van Rietbergen, 2007) found that more than 20% of Dutch respondents reported fewer trips to city center stores. Cairns et al. (Cairns et al., 2004) reported that 80% of 538 U.K. internet users polled by British Telecom have saved at least one car trip because of online buying.

Some studies adopted a different questioning strategy and yielded different results. Wilson et al. (Wilson et al., 2007) asked respondents in three U.S. cities what they would have done if they had not found the last purchase online, and they found that 79% of respondents would have made a special trip to a store for that item; that is, online buying substituted for 79% of shopping trips. On the other hand, 21% would not have happened; that is, these online purchases may not influence store shopping. They also found that 55.5% of respondents generated new trips because of product information obtained online. Corpuz and Peachman (Corpuz and Peachman, 2003) showed that 19% of respondents in Sydney would have made a special trip if online buying were not available and that 14% would have not made the purchase. These two studies yielded mixed results of substitution, complementarity, and independence: online information may generate new trips to retail stores; some online purchases may replace shopping trips; and some may not affect shopping trips at all.

Going beyond descriptive questioning, a few studies have applied multivariate models to reveal the relationships between e-shopping and travel behavior. Using 634 respondents to a shopping survey from Utrecht, The Netherlands, Farag et al. (Farag et al., 2006) employed linear regression to explore the impacts of online buying frequency on trip frequencies of daily in-store shopping and non-daily in-store shopping, controlling for socio-demographics, spatial attributes, internet experience, and attitudes. They found that frequent online buyers were more likely to make shopping trips (for daily and non-daily products). They concluded that “the relationship between online buying and in-store shopping is not one of substitution but of complementarity” (p. 43).

Farag et al. (Farag et al., 2005) applied path analysis to a sample of 826 respondents to the other shopping survey in one urban municipality (Utrecht) and three surrounding suburban

municipalities in the Netherlands. The model contained six endogenous variables: online buying frequency, online searching frequency, in-store shopping frequency, internet use frequency, and two attitudes toward online buying and in-store shopping, respectively. This study assumed a unidirectional influence from e-shopping to in-store shopping. They found that online searching frequency had a positive impact on in-store shopping frequency although the influence of online buying frequency was insignificant. Using the same dataset, Farag et al. (Farag et al., 2007) augmented the path analysis to a Structural Equations Model (SEM), in which traditional shopping and e-shopping were assumed to influence each other. In terms of direct effects, they found again the positive association between online searching and in-store shopping. Further, people who shop often in stores tend to purchase often online (but not vice versa), after controlling for shopping attitudes. Thus, in summary, they concluded a complementary relationship exists between online searching and in-store shopping, but online buying does not have a significant impact on in-store shopping.

Several studies have used activity/travel dairies to explore the implications of online buying for travel. Based on a supplementary survey, Zmud et al. (Zmud et al., 2001) classified respondents to the 1999 household travel survey in Sacramento, CA into two groups: internet shopper and non-internet shopper. They found that internet shoppers had a higher overall trip rates than non-internet shoppers and hence concluded that online buying may generate new trips. However, they did not control for any third-party variables such as income and hence did not rule out the possibility of spuriousness.

Ferrell (Ferrell, 2004) used the 2000 San Francisco Bay Area Travel Survey (BATS 2000) to explore the relationships between home-based teleshopping (by internet, catalog, and television) and travel behavior. He first developed two models to predict out-of-home shopping duration and non-shopping activity duration (with sociodemographics being explanatory variables) and then inserted predicted values into the equations of travel behavior. With respect to shopping travel, Ferrell found that teleshopping *frequency* has positive influences on shopping trip frequency and trip chaining frequency, but its influence on shopping trip distance is insignificant. He concluded a complementary effect. Ferrell (Ferrell, 2005) applied a SEM approach to the BATS 2000 to investigate the connections between teleshopping *duration* and shopping travel.

He found that "people substitute home teleshopping time for shopping travel time and that teleshoppers take fewer shopping trips and travel shorter total distances for shopping purposes" (p. 212). Specifically, every 100-minute teleshopping is associated with a 5-minute reduction in shopping travel time, with a 1-mile reduction in travel distance, and with 0.2-trip reduction in frequency. Ferrell did note that these two studies (Ferrell, 2004, 2005) yielded contradictory relationships between teleshopping and shopping travel, even though they used the same data. Ferrell attributed this critical difference to the limitations of the dataset and different units of analysis in the two studies: household vs. person.

It is becoming clear that the relationships between e-shopping and in-store shopping are more complex than they initially appeared: different survey methods yield different results; and different model specifications produce different results. In particular, the studies using the descriptive questioning approach tend to show both substitution and complementarity effects, and the studies using activity/travel diaries also produce mixed findings of substitution and complementarity. The studies using shopping surveys found complementary and independent effects between online buying and traditional shopping, More importantly, the findings of existing multivariate analyses indicate only association between e-shopping and personal travel but not causality.

3. Methodology

3.1 Data

The data came from a shopping survey administered in the Minneapolis-St. Paul (Twin Cities) seven-county metropolitan area in 2008-2009. Because the purpose of this study is to investigate e-shopping via internet, the population is chosen as adult internet users living in the Twin Cities area. To identify our sample frame, we initially hoped to acquire internet subscriber data from local internet providers. In the Twin Cities, Comcast and Qwest are the major providers of cable and DSL internet connections, respectively. Unfortunately, the data have two main problems. First, the data these companies can provide are inherently incomplete; internet subscribers who use other service providers or connection types are excluded from their lists. Second, neither Comcast nor Quest readily provides subscriber data for research purposes. Therefore, we decided to sample households from our chosen neighborhoods.

Because accurate internet penetration data are not readily available, we used income as a proxy for internet accessibility and drew our sample from households living in certain census block groups with relatively high median household incomes. In particular, we mapped census block groups by median household income in ArcGIS 9.2. Based on our local knowledge, we identified 16 sets of census block groups (totaling 78 block groups). Household median income in each of these block groups is well above the Twin Cities Metropolitan Statistical Area median household income ($54,304), with the median household income in most of the groups being 50 to 100 percent higher.

Previous research has shown that consumers' adoption of e-shopping and the interaction between e-shopping and in-store shopping are likely to be influenced by retail accessibility (Ren and Kwan, 2009; Weltevreden and Van Rietbergen, 2007). Therefore, we selected sample households from three different location types: urban neighborhoods, suburban neighborhoods, and exurban neighborhoods. These locations represent different levels of shopping accessibility. Because internet penetration in exurban neighborhoods tends to be lower than that in the other two types of neighborhoods, we oversampled exurban neighborhoods by choosing four urban, four suburban and eight exurban neighborhoods (16 sets of block groups) to achieve a relatively balanced sample of internet users from all three location types. The geographical distribution of these neighborhoods is shown in Figure 1.

[Insert Figure 1 here]

We purchased a random sample of 500 addresses in each of the 16 neighborhoods from AccuData Integrated Marketing, a commercial data provider (http://www.accudata.com). The survey was administered on Survey Monkey, a web-based survey provider, from December 2008 to January 2009. An invitation letter and two reminder postcards (1 and 4 weeks later) were mailed during the period. Five $100 gift cards were provided as the incentive for the survey. After removing duplications and two non-adult participants, the number of respondents totaled 591. The survey response rate was 7.85% (=591/7,533, based on valid addresses only). However, this response rate is likely to be substantially lower than the true response rate because

the population of this survey is internet users while many (though unknown) households in our sample frame are not internet users.

Table 1 presents the number of respondents in each neighborhood and Table 2 depicts selected demographic characteristics among survey respondents. Generally, respondents are well-educated and well-paid workers with few limitations on their shopping behaviors such as low income or lack of a driver's license or credit card. The sample may not be a good representative of internet users in the Twin Cities, especially when overdrawing high income households. However, since the focus of our study is on explaining the relationships of other variables to shopping behavior rather than on describing shopping behavior per se, this is not expected to materially affect the results of multivariate analysis (Babbie, 2007). For example, we are not trying to replicate the population distribution of income, but rather the impact of income on shopping behavior. Since we have ample diversity of income across the sample and thence ample opportunity to properly capture its effect on shopping behavior, the under- or over-representation of certain income categories is not serious. A similar principle holds with respect to other sources of non-representativeness, provided those sources are exogenous - such a situation is equivalent to exogenously stratified random sampling, and under those circumstances, coefficients of linear regression models are statistically unbiased (Hausman and Wise, 1981). Thus, although we do not claim perfect representativeness for our sample in terms of univariate distributions, the relationships among the variables of interest that are captured by our models (and that are the focus of this study) are generalizable.

[Insert Tables 1 and 2 here]

3.2 Variables

The rationale for variable development is as follows. Does a positive (or negative) association between shopping behaviors mean that e-shopping has a complementarity (or substitution) effect on in-store shopping? The answer is not straightforward, however. The positive association can be attributed to at least three mechanisms (Mokhtarian and Circella, 2007). First, online buying induces the need for additional related goods such as accessories (a direct causal influence). Second, time savings from online buying are reallocated to other shopping activities and travel (an indirect causal influence). Last but not the least, the association may result from factors

antecedent to both shopping behaviors. Antecedent factors can result from various sources. For example, affluent individuals are likely to have a higher demand for online buying and traditional shopping than the poor; similarly, other socio-demographics (such as household size and age) can cause the demand for both shopping activities in the same way. Shopping enjoyment may also increase the demand for shopping through multiple channels. Household shopping responsibility is another important, but mostly ignored, factor. Shopping is motivated by enjoyment but also by household responsibility. Failure to control for the responsibility may produce spurious relationships between online buying and traditional shopping at the individual level (Cao, 2009). Furthermore, shopping accessibility and location attributes may be third-party variables according to the efficiency hypothesis: people with a low shopping accessibility or in exurban neighborhoods may make fewer in-store purchases but conduct more online buying (Anderson et al., 2003).

In this shopping survey, instruments were designed to account for the influences of the antecedent factors on both e-shopping and traditional in-store shopping. The variables in this study were classified into six groups: shopping behavior, shopping attitudes, internet experience, demographics, household responsibility, and shopping accessibility.

The variables of interest are respondent's e-shopping and in-store shopping behavior regarding non-daily purchases ("products such as books, clothes or electronics"). These variables include frequency of store trips, frequency of internet purchases, frequency of product information search via internet, preferred alternative to internet purchases, and a dummy indicating whether or not the respondent has made a store trip as a result of internet browsing.

Shopping attitudes variables were derived from a factor analysis. In the survey, respondents were asked to indicate whether they agree or disagree with 15 statements on a five-point Likert scale from "strongly disagree" to "strongly agree". The statements reflect fundamental shopping attitudes that are expected to influence the respondents' shopping behavior. Since some of these statements measure similar dimensions of shopping attitudes and are highly correlated, we conducted a factor analysis to identify latent constructs of the attitudes. As shown in Table 3, four factor scores (price consciousness, time consciousness, shopping enjoyment, and impulsive

shopper) were derived using common factor analysis (called principle axis factoring in SPSS) with Varimax rotation.

[Insert Table 3 here]

The procedure for factor analysis is as follows. Before conducting factor analysis, missing value analysis with expectation maximum algorithm was applied to 15 statements due to 20 missing cells (out of 15*588 = 8820 cells) in the data. Because the purpose of the factor analysis is to identify latent constructs of shopping attitudes, we adopted common factor analysis (CFA) rather than principle components analysis (PCA). It is worth noting that CFA tends to produce lower factor loading and percent variance explained by the factor solution than PCA. However, CFA is still superior to PCA because the PCA loadings are more biased estimators than are the CFA loadings (Widaman, 1993). When identifying the number of factors, a number of criteria were used. In particular, both eigenvalue-one rule and scree role point to the four-factor solution shown in the Table. We also tried a five-factor solution. However, the statements underlying the fifth factor are a subset of the statements for the factor score "shopping enjoyment". Thus, we decided to choose the four-factor solution as our final results. Overall, 58.2% of total variation is explained by the factor scores using initial eigenvalues and CFA explains 43.5% of total variation.

The internet experience category included variables describing the respondent's usage of the internet: years of using internet, frequency of personal internet use, and internet connection method (dial-up or broadband).

The demographic category included demographic characteristics that could influence the frequency of respondent's shopping activity, which consist of gender, age, possession of a driver's license, possession of a credit card, household income, number of household vehicles, household size, number of children, education background, student status, employment status, and any conditions preventing the respondent from driving, walking, or taking public transportation.

The household responsibility category included a single variable, the percent of household shopping activities for which the respondent is primarily responsible.

Shopping accessibility variables were derived from ArcGIS analysis. We calculated retail accessibility for each of the 78 block groups included in this study. When computing these variables, we relied on the U.S. Census Bureau's Longitudinal Employer-Household Dynamics data, which contains the number of retail jobs per block group in the Twin Cities area. Using number of retail jobs as a proxy for retail opportunity, we calculated five cumulative shopping accessibility variables for each block group: the number of retail jobs within a pre-specified buffer (1, 2, 5, 10 and 20 miles). Because some block groups may partially fall within the pre-specified buffer, the number of retail jobs in a block group was counted when the centroid of the block group fell within the buffer.

3.3 Analysis approach

In the survey, in-store shopping frequency was measured in an ordinal scale ranging from 0 (less than once a month) to 3 (more than once a week). Specifically, 24.6% of respondents go to a store to purchase non-daily products for themselves and their households less than once a month (one respondent reported never); 47.9% go to a store one to three times per month; 20.3% go to a store once a week, and 7.2% go to a store more than once a week. So we employed ordered-response techniques to model the behavior. In an ordered-response model, individuals' behavior is observed in an ordinal scale (Y = 0, 1, 2,…j,…, J), and we assume that an underlying latent continuous variable, Y*, represents a respondent's propensity to store shopping (Daykin and Moffatt, 2002). Y* is expressed in the following form:

$$Y^* = \beta' X + e, \qquad (1)$$

where X is a vector of explanatory variables, β is a vector of parameters, and e is the unobserved error term. The relationship between the latent Y* and the observed Y is:

$$Y = j \text{ if } \mu_{j-1} < Y^* \leq \mu_j, \text{ j} = 0,1,2,...,\text{J}, \qquad (2)$$

where the μ_js are cut points or threshold parameters, defined as $\mu_{-1} = -\infty$, $\mu_J = +\infty$, and $\mu_{j-1} < \mu_j$ for all j. In the context of the ordered probit model, $e \sim N[0, \sigma_e^2]$, and we have the following probabilities:

$$P(Y = j) = P(\mu_{j-1} < Y^* \leq \mu_j) = \Phi(\frac{\mu_j - \beta' X}{\sigma_e}) - \Phi(\frac{\mu_{j-1} - \beta' X}{\sigma_e}), \quad (3)$$

Where Φ denotes the standard normal CDF. The coefficients are derived from maximum likelihood estimation of a sample.

The Veall-Zimmermann R^2 was chosen as a goodness-of-fit measure for the models because it is better than McFadden R^2 when the number of ordinal categories exceeds three (Veall and Zimmermann, 1996). The R^2_{vz} is calculated as follows:

$$R^2_{vz} = \frac{[LL(\beta) - LL(C)][N - 2LL(C)]}{-LL(C)[N + 2LL(\beta) - 2LL(C)]}, \quad (4)$$

where LL(c) is the log-likelihood for the model where all coefficients but J-1 thresholds are restricted to zero (equivalent to a constant-only model); LL(β) is the log-likelihood at convergence (i.e., the final model); N is the number of observations in the model.

4. Results

In this section, the results of descriptive questioning are presented before proceeding to multivariate analysis results. Similar to Wilson et al. (Wilson et al., 2007), this survey directly asked respondents what they would have done if their last purchase over the internet had not been available over the internet. Consistent with previous research, this question produces mixed results on the interactions between e-shopping and in-store shopping. In particular, about 29% of respondents would have made a special trip to the store to find this item or a similar item. This means that 29% of online purchases may substitute for in-store shopping. About 18% of respondents would have not made the purchase, which suggests that 18% of online purchases are likely to be induced demand. Further, 14% would have tried to purchase the item or similar item through catalog or other media, a possible substitution for other teleshopping channels. Finally, 39% of respondents would have waited until their next trip to the store. Note that although the last three scenarios may not have generated additional trips to the store, the fact that the respondent did purchase the item on-line likely resulted in some form of home delivery. In

another question, 49.3% of respondents indicated that they have made a special trip to the store because of something they have seen online, which provides evidence for complementarity effect.

Using the SPSS 17.0 software package, we developed five ordered probit models to explore the correlates of store shopping frequency. The first model illustrated the relationships between store shopping frequency and demographic components of the sample and those between store shopping behavior and internet experience. The second incorporated data from variables regarding respondents' internet use for online shopping and online searching. The third added neighborhood dummy variables and shopping accessibility variables. The fourth incorporated four factor scores of shopping attitudes. In the last specification, the variable indicating respondents' shopping responsibility in the household was incorporated in the model. The interest of this model is to show the relationship between in-store shopping and e-shopping behavior, with other variables being controls. When developing these models, we dropped variables whose p-values were greater than 0.2.

The results of these models are shown in Table 4. After we dropped some variables, Model 1 shows that household income is positively associated with store shopping frequency, although the association is not significant at the 0.05 level. The number of children (12-15 years old) has a positive association with store shopping frequency, indicating the larger demand induced by those children. Among students, full-time workers, and retirees, students were least likely to shop at a store frequently, followed by full-time workers, whereas retirees were most likely to shop at a store frequently. Further, respondents who had a broadband internet connection tend to go shopping more frequently. Number of years using internet is negatively associated with store shopping frequency, but the association is insignificant at the level of 0.05.

[Insert Table 4 here]

After controlling for demographics and internet experience, we found that two more variables, how often respondents consulted the internet for information about non-daily products and how often they actually bought those products over the internet, are positively associated with store

shopping frequency (Model 2). These two positive influences suggest that e-shopping (online searching and online buying) is likely to have a complementary effect on store shopping.

In Model 3, we further incorporated neighborhood location dummies and shopping accessibility variables. The suburban dummy is significant at the 0.05 level. In particular, suburban respondents were more likely to shop in stores. However, none of the remaining location variables are insignificant at the 0.2 level. In this model, e-shopping behavior variables remain statistically significant.

Our fourth model added the shopping attitudes identified in the factor analysis. This model showed that three attitudes had positive impacts on in-store shopping frequency although the cost consciousness variable is not significant even at the 0.1 level. As expected, those who intrinsically enjoyed shopping as well as impulsive shoppers are more likely to shop in stores. When comparing the coefficients of online shopping frequency across the models, interestingly, we found that the coefficient in Model 4 is considerably larger than that in Model 3. In other words, once we control for shopping attitudes, the complementary connection between online shopping and store shopping becomes stronger. This is contrary to our expectation. If an antecedent variable affects two outcome variables in the same way, the association between the latter two variables tends to diminish once the antecedent variable is controlled for. By contrast, if the antecedent variable influences the two outcome variables in opposite ways, the association between outcomes becomes stronger when we account for the influence of the antecedent variable on both outcomes[i]. Through a further exploration, we found that the shopping enjoyment attitude is positively associated with store shopping frequency, but negatively associated with online shopping frequency. Therefore, it is reasonable that the coefficient of online shopping frequency in Model 4 is larger than that in Model 3. This result suggests that if we do not control for shopping attitudes, we tend to underestimate the complementarity between online shopping and store shopping.

We also included the shopping responsibility variable in the model. However, it is not significant at the 0.2 level. Thus, the model is not presented in the table. A bivariate correlation analysis confirmed that the variable is not significantly associated with either online shopping frequency

or store shopping frequency. Therefore, this study failed to support the alternative hypothesis that shopping responsibility in the household is an antecedent factor of online and store shopping frequencies.

Veal-Zimmerman R^2 increased from 0.104 of Model 1 to 0.240 of Model 4. The large increase results from e-shopping behavior and shopping attitudes. Based on these models, we conclude that both online searching and online buying tend to have a complementary, rather than substitution, effect on in-store shopping trips. Although it appears that those who have used the Internet longest may have adopted behaviors that allowed them to reduce in-store shopping, increased access to the internet in the form of a broadband connection, gathering information about a product via internet, and online shopping allowed individuals to increase their shopping activities in-stores.

5. Conclusions

This research aims to reveal the interactions between e-shopping and in-store shopping using a sample of adult internet users in the Minneapolis-St Paul metropolitan area. Ordered probit models were developed to account for the influences of a variety of confounding factors such as shopping attitudes, shopping accessibility, shopping responsibility, and socio-demographics. The preliminary results showed that online searching and online buying tend to have a complementary effect on in-store shopping, controlling for the confounding factors.

However, more sophisticated modeling techniques could be insightful. The data showed that shopping enjoyment is negatively associated with online shopping frequency and positively associated with in-store shopping frequency. In our ordered probit models, we controlled for shopping enjoyment. However, the model did not establish the connection between shopping enjoyment and online shopping frequency. This type of disconnection may be true for other variables in the model. Therefore, a structural equations model may be useful to capture various interactions among variables.

Acknowledgements

This research is funded by the Tech Plan Program of the ITS Institute at University of Minnesota.

Table 1. Number of Respondents by Neighborhood

Response Rates per Neighborhood (N=591)			
Neighborhood		Location	Respondents
Urban	1	Linden Hills Minneapolis	42
	2	Lynhurst, Minneapolis	46
	3	Macaletser-Groveland, St. Paul	51
	4	Cedar-Isles, Minneapolis	44
			183
Suburban	5	Vadnais Heights	53
	6	Eden Prairie	33
	7	Woodbury	31
	8	Edina	36
			153
Exurban	9	Anoka	30
	10	Rogers	26
	11	Chaska	38
	12	Waconia	39
	13	Lakeville	33
	14	Prior Lake	27
	15	New Market	32
	16	Stillwater	30
			255

Table 2. Sample Characteristics

Variables	Mean/Number	Percent
Age (n=566)	48.39	
Number of vehicles per household (n=570)	2.30	
Household size (n=569)	2.81	
Gender (n=572)		
Male	243	42.5
Female	329	57.5
Driver's license (n=570)		
Yes	567	99.5
No	3	0.5
Credit card (n=570)		
Yes	557	97.7
No	13	2.3
Household income (n=546)		
Less than $30,000	20	3.7
$30,000 - $44,999	32	5.9
$45,000 - $59,999	50	9.2
$60,000 - $74,999	50	9.2
$75,000 - $99,000	97	17.8
$100,000 - $124,999	93	17.0
$125,000 - $149,999	63	11.5
$150,000 or more	141	25.8
Educational background (n=570)		
High school or less	32	5.6
Some college or technical school	86	15.1
Two-year associate's degree	29	5.1
Four-year college/technical school degree	202	35.6
Some graduate school	56	9.9
Complete graduate school	163	28.7
Occupation (n=566)		
Student	11	1.9
Sales	38	6.7
Service/repair	15	2.7
Manager/administrator	120	21.2
Production/construction/crafts	23	4.1
Professional/technical	288	50.9
Clerical/administrative support	28	4.9
Homemaker	43	7.6
Employment status (n=568)		
Work full time	376	66.2
Work part time	73	12.9
Not working	49	8.6
Retired	70	12.3

Table 3. Rotated Factor Matrix of the Factor Analysis

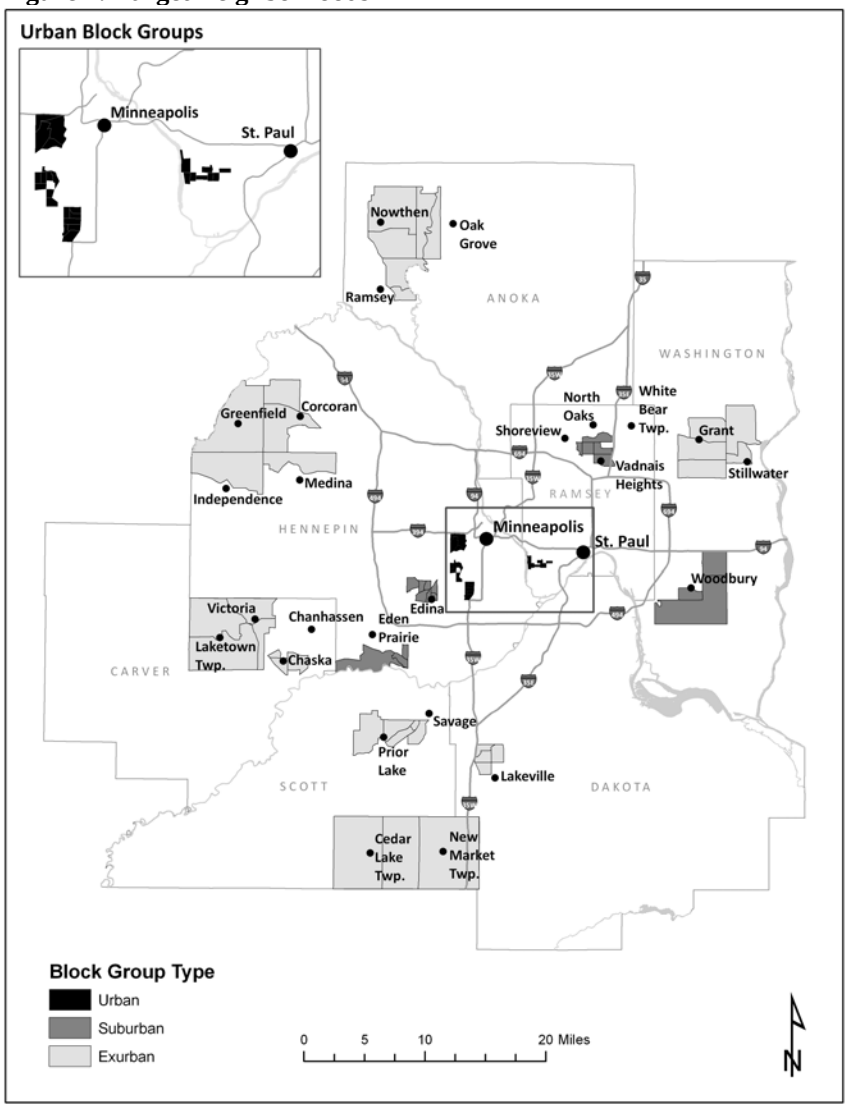
Statements	Shopping enjoyment	Cost consciousness	Impulsive shopper	Time consciousness
I often make unplanned purchases.			0.662	
Before buying something, I generally take some time to think it over.		0.454	-0.352	
I generally stick to my shopping lists.			-0.552	
If I really want something, I'll often buy it even if it costs more than it should.		-0.354	0.428	
I generally compare prices before buying in order to get the best value.		0.751		
It's important to me to get the lowest prices when I buy things.		0.667		
I'm too busy to shop as often or as long as I'd like.				0.637
Being a smart shopper is worth the extra time it takes.		0.661		
Saving time when I shop is important to me.	-0.376			0.224
Under the right circumstances, shopping is fun.	0.608		0.239	0.227
Shopping is usually a chore for me.	-0.859			
For me, shopping can be an important leisure activity.	0.640		0.371	
Sometimes for me, shopping is mostly an excuse to get out of the house or workplace.	0.328		0.379	
Getting to where I usually shop is a hassle.	-0.372			0.221
Shopping at stores is too tiring to be enjoyable.	-0.765			

Note: the analysis is principle axis factoring with Varimax rotation and loadings smaller than 0.2 were suppressed.

Table 4. Ordered Probit Models for In-store Shopping frequency

	Model 1		Model 2		Model 3		Model 4	
	Estimate	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**
Mu_0	-0.666	0.059	-0.215	0.558	-0.116	0.753	-0.007	0.984
Mu_1	0.670	0.057	1.137	0.002	1.243	0.001	1.423	0.000
Mu_2	1.540	0.000	2.040	0.000	2.154	0.000	2.387	0.000
Demographics and internet experience								
Household income	0.036	0.146						
# of children (12-15 years old)	0.179	0.085	0.165	0.107	0.170	0.097	0.202	0.051
Student	-0.626	0.040	-0.581	0.059	-0.551	0.073	-0.703	0.025
Full time worker	-0.217	0.065	-0.187	0.104	-0.181	0.116	-0.169	0.146
Being retired	0.345	0.038	0.377	0.022	0.363	0.027	0.356	0.032
Years using internet	-0.097	0.177	-0.170	0.017	-0.164	0.020	-0.110	0.127
Broadband connection	0.393	0.021	0.280	0.100	0.289	0.089	0.315	0.070
E-shopping behavior								
Frequency of product information search			0.191	0.001	0.187	0.001	0.138	0.021
Frequency of online shopping			0.098	0.011	0.103	0.008	0.121	0.002
Location								
Suburb neighborhoods					0.243	0.021	0.215	0.044
Attitudes								
Shopping enjoyment							0.238	0.000
Impulsive shopper							0.297	0.000
Cost consciousness							0.078	0.163
N	544		567		567		567	
Veall-Zimmermann R-square	0.104		0.158		0.163		0.240	

Figure 1: Target Neighborhoods



References

Anderson W P, Chatterjee L, Lakshmanan T R, 2003, "E-commerce, transportation, and economic geography" *Growth and Change* **34** 415-432
Babbie E R, 2007 *The practice of social research* (Thomson Wadsworth, Belmont, CA)
Cairns S, Sloman L, Newson C, Anable J, Kirkbride A, Goodwin P, 2004, "Smarter Choices – Changing the Way We Travel",
Cao X, 2009, "E-Shopping, Spatial Attributes, and Personal Travel: A Review of Empirical Studies" *Transportation Research Record: Journal of the Transportation Research Board*
Cao X, Mokhtarian P, 2005, "The Intended and Actual Adoption of Online Purchasing: A Brief Review of Recent Literature", (Institute of Transportation Studies, University of California, Davis)
Chang M K, Cheung W M, Lai V S, 2005, "Literature derived reference models for the adoption of online shopping" *Information & Management* **42** 543-559
Choo S, Lee T, Mokhtarian P, 2007, "Do Transportation and Communications Tend to Be Substitutes, Complements, or Neither?: U.S. Consumer Expenditures Perspective, 1984-2002" *Transportation Research Record: Journal of the Transportation Research Board* **2010** 121-132
Corpuz G, Peachman J, 2003, "Measuring the impacts of internet usage on travel behaviour in the Sydney Household Travel Survey", in *Abstract for 2003 ATRF conference*
Daykin A R, Moffatt P G, 2002, "Analyzing ordered responses: A review of the ordered probit model" *Understanding Statistics* **1** 157-166
Farag S, 2006 *E-Shopping and its Interactions with In-Store Shopping*, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands
Farag S, Krizek K J, Dijst M, 2006, "E-shopping and its relationship with in-store shopping: Empirical evidence from the Netherlands and the USA" *Transport Reviews* **26** 43-61
Farag S, Schwanen T, Dijst M, 2005, "Empirical Investigation of Online Searching and Buying and Their Relationship to Shopping Trips" *Transportation Research Record: Journal of the Transportation Research Board* **1926** 242-251
Farag S, Schwanen T, Dijst M, Faber J, 2007, "Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping" *Transportation Research Part a-Policy and Practice* **41** 125-141
Ferrell C, 2004, "Home-Based Teleshoppers and Shopping Travel: Do Teleshoppers Travel Less?" *Transportation Research Record: Journal of the Transportation Research Board* **1894** 241-248
Ferrell C, 2005, "Home-Based Teleshopping and Shopping Travel: Where Do People Find the Time?" *Transportation Research Record: Journal of the Transportation Research Board* **1926** 212-223
Hausman J, Wise D, 1981, "Stratification on Endogenous Variables and Estimation: The Gary Income Maintenance Experiment", in *Structural Analysis of Discrete Data with Econometric Applications* Eds C Manski, D McFadden (The MIT Press, Cambridge, MA) pp 365-391
Hu P S, Reuscher T R, 2004, "*Summary of Travel Trends: 2001 National Household Travel Survey*", (FHWA, USDOT)
InternetRetailer.com, 2008, "E-commerce sales rise 19% last year in U.S."
Mokhtarian P L, 1990, "A Typology of Relationships between Telecommunications and Transportation" *Transportation Research Part a-Policy and Practice* **24** 231-242

Mokhtarian P L, 2004, "A conceptual analysis of the transportation impacts of B2C e-commerce" *Transportation* **31** 257-284
Mokhtarian P L, Circella G, 2007, "The role of social factors in store and internet purchase frequencies of Clothing/Shoes", in *The international workshop on Frontiers in Transportation: Social Interactions* (Amsterdam, The Netherlands)
Pendyala R M, Goulias K G, Kitamura R, 1991, "Impact of Telecommuting on Spatial and Temporal Patterns of Household Travel" *Transportation* **18** 383-409
Ren F, Kwan M P, 2009, "The impact of geographic context on e-shopping behavior" *Environment and Planning B: Planning and Design* **36** 262-278
Salomon I, 1986, "Telecommunications and Travel Relationships - a Review" *Transportation Research Part a-Policy and Practice* **20** 223-238
Sim L L, Koi S M, 2002, "Singapore's Internet shoppers and their impact on traditional shopping patterns" *Journal of Retailing and Consumer Services* **9** 115-124
Tonn B E, Hemrick A, 2004, "Impacts of the Use of E-Mail and the Internet on Personal Trip-Making Behavior" *Social Science Computer Review* **22** 270-280
Veall M R, Zimmermann K F, 1996, "Pseudo-R2 Measures for Some Common Limited Dependent Variable Models" *Sonderforschungsbereich* **386**
Weltevreden J W J, Van Rietbergen T, 2007, "E-shopping versus city centre shopping: The role of perceived city centre attractiveness" *Tijdschrift Voor Economische En Sociale Geografie* **98** 68-85
Widaman K F, 1993, "Common factor analysis versus principal component analysis: differential bias in representing model parameters?" *Multivariate Behavioral Research* **28** 263-311
Wilson R, Krizek K, Handy S, 2007, "Trends in Out-of-Home and At-Home Activities: Evidence from Repeat Cross-Sectional Surveys" *Transportation Research Record: Journal of the Transportation Research Board* **2014** 76-84
Zmud J, Bricka S, Casas J, 2001, "Impact of shopping via internet on travel for shopping purposes", in *The 80th Annual Meeting of the Transportation Research Board* (Washinton, DC)

---

[i] This is equivalent to the concept of partial correlation, which measures the correlation between two variables while removing the effects of a third variable. Simply put, if we want to measure the correlation between online shopping frequency (X) and store shopping frequency (Y) while taking away the effects of shopping enjoyment (Z), the partial correlation between shopping frequencies can be derived using the following equation:
$r_{XY\cdot Z} = \frac{r_{XY} - r_{XZ}r_{YZ}}{\sqrt{(1 - r_{XZ}^2)((1 - r_{YZ}^2)}}$. If the Pearson correlation between the frequencies r_{XY} is positive and both r_{XZ} and r_{YZ} are positive (or negative), the partial correlation $r_{XY\cdot Z}$ may be smaller than the Pearson correlation r_{XY}. If r_{XY} is positive, and one of r_{XZ} and r_{YZ} is positive and the other is negative, the partial correlation $r_{XY\cdot Z}$ is larger than the Pearson correlation r_{XY}.