Transportation Research Part A 70 (2014) 1–9

Contents lists available at ScienceDirect

Transportation Research Part A

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

Explore the relationship between online shopping and shopping trips: An analysis with the 2009 NHTS data

Yiwei Zhou [a,1], Xiaokun (Cara) Wang [b,*]

[a] Department of Civil and Environmental Engineering, 4027 JEC Building, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590, USA
[b] Department of Civil and Environmental Engineering, 4032 JEC Building, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590, USA

A R T I C L E I N F O

Article history:
Received 16 February 2014
Received in revised form 25 September 2014
Accepted 30 September 2014

Keywords:
Online shopping
E-commerce
Personal trips
NHTS
SEM

A B S T R A C T

The rapid growth of ecommerce brings great changes to the transportation system. However, most existing studies focus on the impact of ecommerce on freight system. Its impact on personal trips is relatively less studied. It is reasonable to argue that online shopping reduces the need of shopping trips by making goods accessible via door-to-door deliveries. On the other hand, online shopping may also create more shopping trips as online shoppers travel to stores to experience, compare or pick up the goods. Understanding the connections between online shopping and shopping trips is critical for transportation planners to prepare for changes that information technology will continue to bring to this nation in the future. Using the 2009 National Household Travel Survey (NHTS) data and a structural equation model (SEM), this paper disentangles the bidirectional connections between online shopping and shopping trips. Results show that online shopping encourages shopping trips while shopping trips tend to suppress the online shopping propensity. Besides, both online shopping and shopping trips are influenced by exogenous factors such as shoppers' demographic features, regional specific factors and household attributes. A closer examination at the state level further confirms model validity while disclosing spatial variation in their relationship.

1. Introduction

The advance of information and communication technology has profoundly changed people's way of life. With the widespread use of internet, more and more people choose to shop online. Online shopping, or more often categorized as e-commerce, has been increasing dramatically during the past decade. In the U.S., total e-commerce sales in 2012 reached about $225.5 million, an increase of 15.8% compared to 2011. It took up 5.2% of total sales and the percentage continues to increase (Census, 2013).

The rapid growth of online shopping has brought great changes to the transportation system. As summarized by Mokhtarian (2004), the potential impacts of online shopping include changes in shopping mode share, changes in the volume of goods purchased, changes in per capita consumption spending, and demographic changes. From the perspective of planners and policy makers, these changes imply impacts on economy, population, land use, freight transportation and passenger transportation. As one can expect, online shopping changes freight trip pattern to and from businesses, creates freight trips

* Corresponding author. Tel.: +1 518 276 2098.
E-mail addresses: zhouy7@rpi.edu (Y. Zhou), wangx18@rpi.edu (X. (Cara) Wang).
[1] Tel.: +1 518 867 6669.

http://dx.doi.org/10.1016/j.tra.2014.09.014
0965-8564/

to residential areas, and influences residents' shopping trips. This paper focuses on the last aspect and examines the relationship between online shopping and personal shopping trips.

There has been a long-lasting debate on the impact of online shopping on shopping trips, whether being substitution or complement (Aguiléra et al., 2012). One hypothesis proposed by some of these studies is that online shopping serves as a substitute of shopping trips because there is no need for people to make the actual shopping trips when they can shop online. For instance, Sim and Koi (2002) found that 12% of online buyers reduced their trips to stores based on a survey on 1500 local consumers in Singapore. Tonn and Hemrick (2004) conducted a web survey among residents in Tennessee. Results showed that about 40% of residents reported less driving with the use of internet. Weltevreden and Rietbergen (2007) studied the impact of e-shopping on in-store shopping based on data of 3074 internet users in the Netherlands. Results indicated that more than 20% online buyers made fewer trips to city center stores.

An alternative hypothesis is that online shopping complements personal shopping trips: As people are able to freely browse and choose among numerous products, their desire of shopping may be stimulated, leading to more shopping trips. Choo et al. (2008) analyzed relations between transportation and communications based on the U.S. Consumer Expenditure 1984–2002. Results indicated that both substitute and complementary effects existed, but the impact of communications on transportation was dominantly complementary. Cao et al. (2012) studied relationship between online searching frequency, online buying frequency and in-store shopping frequency using data of 539 adults from a shopping survey in Minneapolis-St Paul metropolitan area and found that online shopping tended to have a complementary effect on in-store shopping.

There are also studies suggesting neutral effect of online shopping on shopping trips. The e-commerce had only a modest impact on consumer travel patterns as moderated by other exogenous factors (Calderwood and Freathy, 2014). In general, depending on survey samples, conclusions on the relationship between online shopping and shopping trips differ significantly. It can be substitute, complement, or neutral.

A few studies tried to explore such inconsistency by examining the spatial variation of the shopping behavior. For example, using a sample of 740 adults from Seattle, Kansas City and Pittsburgh, Krizek et al. (2005) found that people living in suburbs or away from central business district (CBD) were more likely to make online shopping, but the effects are insignificant. Farag et al. (2006) studied shopping behavior of 2190 individuals in the Netherlands. Results suggested that "people are more likely to adopt e-shopping when their accessibility to shops is relatively low".

2. Methodology

In terms of methodology investigating the relationship between online shopping and shopping trips, various methods can be used depending on the characteristics of available data. Among them, the most widely adopted approaches are path analysis, binary logit model, multinomial logit (MNL), joint decision models and structural equation modeling (SEM).

Unlike traditional regression models, SEM allows reciprocal influence among variables. In fact, as stated by Pearl (2013), linear SEM serves as a "microscope" for causal analysis. It helps quantify the causal assumptions in the model and assess the impact of a particular phenomenon. Another advantage of SEM is that the framework accommodates the use of latent variables, which can be used to conveniently incorporate unobservable variables such as attitude into the model. SEM is widely used in many fields such as agriculture, meteorology, economy, psychology, sociology and engineering. Farag et al. (2007) studied the interactions between online searching frequency, online shopping and shopping trips using SEM. In addition to disclosing a complementary effect between online shopping and shopping trips, the study also found that urban residents shopped online more frequently than suburban residents. Cao et al. (2012) used SEM to examine the relationship between online searching frequency, online buying frequency and in-store shopping frequency. Positive mutual influence between each pair of the dependent variables was identified. Household income, age, education, and working status were also found significantly influential to individual shopping behavior. Existing methods to study impacts of online shopping are summarized in Table 1.

Table 1
Existing methods summary.

Literatures	Methods	Key variables	Conclusions
Farag et al. (2005)	Path analysis	Socio-demographic, land use, behavioral, and attitudinal variables	Complementary effect between online shopping and shopping trips
Ren and Kwan (2009)	Binary logit	Travel pattern, internet diary and accessibility data	People with low accessibility to local shopping and white people are more likely to adapt to online shopping
Weltevreden and Rietbergen (2007)	MNL	Shopping enjoyment, internet access, education, accessibility and travel mode	Complementary effects in short run but substitution in long run
Ferrell (2005)	SEM	Time use variables of work, maintenance and shopping activities	Small substitution effect of online shopping on shopping trips
Ferrell (2004)	2SLS	Time use variables and trip activity data	Complementary effect of online shopping on shopping trips
Cao et al. (2012)	SEM	Shopping and internet use frequency, social-demographics	Complementary effect of online shopping on shopping trips
Farag et al. (2007)	SEM	Life style, shopping attitudes, land use, shopping behavior	Complementary effect of online shopping on shopping trips

This paper investigates the relationship between online shopping and shopping trips using the 2009 National Household Travel Survey (NHTS) dataset with the SEM model. Analysis is conducted at both national and state levels. Results from the national level will be great addition to existing studies, which mostly focus on specific regions. The state level analysis will provide further insights into the regional differences and the effects of various influential factors.

In the next section, data sources and variable characteristics are described, followed by explanation of model specification and analysis of results. The paper concludes with key findings and potential future work.

3. Data description

The data used in this study is derived from the 2009 NHTS data (FHWA, 2012). Carried out every 5–10 years, NHTS collects detailed household travel data to assist transportation planning and research. Previous NHTS surveys were carried out in 1969, 1977, 1990, 1995 and 2001. The recently released 2009 NHTS dataset includes latest information on household trip purpose, mode, time, frequency, together with information on household income, family size, education level, and location, etc. The survey was conducted through interviews on 150,147 households, including 308,901 individuals in the U.S.

Besides rich household travel information, the 2009 NHTS also collects information for frequency of online shopping, frequency of web use and number of shopping trips. This is the first time that NHTS collects variables related to online shopping. The large sample and the comprehensive information create great opportunity to conduct in-depth analysis on shopping behavior. As discussed above, empirical data in previous studies was usually collected from surveys in specific regions, containing hundreds, at most thousands, of records with limited number of variables. In contrast, the NHTS dataset covers residents in the entire national, with around 300,000 individual records and more than 200 variables. NHTS data was collected and post-processed carefully to ensure data validity. In short, the 2009 NHTS provides unprecedentedly comprehensive, accurate and timely dataset for the research on online shopping and shopping trips.

The exploration of this latest 2009 NHTS dataset has just started and literature has not been widely seen yet. This paper is by far the first attempt to use 2009 NHTS to rigorously assess the relationship between online shopping and shopping trips. Results from this new dataset will enrich existing literature on this topic and supplement the group of soon-to-emerge travel behavior studies based on the 2009 NHTS dataset.

The NHTS 2009 dataset consists of four data files: household file, person file, vehicle file and trip file. As their names imply, each file contains different sets of variables. These files were merged and processed to generate personal-specific variables. Variables representing daily travel pattern are first derived based on the original trip information by calculating summation, average, or percentage of specific types of trips. 85,663 records remain after removing invalid records. The final dataset used for this study contains five types of variables: internet related, regional-specific, household-related, personal-related, and travel pattern-related. Their descriptions and key statistics are listed in Table 2.

Removing invalid observations inevitably reduces the sample size (from the original 308,901 to the current 85,663 records) and potentially changes the sample distribution. As summarized in Table 3, male, higher income families, frequent online users and individuals with bachelor's degrees tend to be overrepresented in the final sample. As the NHTS involves sophisticated sample bias correction process, this study does not make more statistical adjustments to the sample in order to avoid further complication. However, limitation of the sample is recognized and the analysis results will be interpreted with care.

4. Model specification

As previously discussed, online shopping may influence in-store shopping and vice versa, yet some previous studies only focus on one side of the bidirectional influence. An advantage of SEM is that it enables the assessment of bidirectional relations between multiple endogenous variables, providing a flexible framework to explore interactions between online shopping and shopping trips.

Another advantage of SEM is its convenient accommodation of latent variables. It can be noticed that the NHTS 2009 asked respondents to provide their web use and online purchase frequency in the *past month*, while the number of shopping trips is only reported for the *specific travel day*. However, because online shopping and shopping trips are rare events for their respective given time frames, these two variables are not sufficient to accurately reflect the true frequency of a person's online shopping and shopping trips. The true frequency, or people's propensity to shop online (or make shopping trips), is unobserved. Therefore, two latent variables are created: "prop_purchase" representing an individual's propensity of online shopping and "prop_shoptrp" indicating an individual's propensity of making shopping trips. In addition to the mutual influence, these two variables are also influenced by a set of exogenous factors including the various regional specific, household-related, personal-related and travel pattern related variables. Retrospective variables such as web use and online purchase frequency in the past month could be used as measurement variables for "prop_purchase." Similarly, total number of trips and total number of shopping trips on the survey day can be used as measurement variables for "prop_shoptrp." It can be expected that the total number of shopping trips on a specific day is also significantly influenced by whether that day is weekend.

The structural relationship between latent variables, measurement variables and exogenous variables can be illustrated in Fig. 1.

Table 2
Variable description and key statistics.

Variable names	Description	Mean	Std. dev	Min	Max
Internet related					
purchase	Number of times purchasing via internet in past month	2.277	4.823	0	250
webuse	Internet use frequency in past month (Base case: once a week or less)				
webuse1	1 if almost every day; 0 otherwise	0.815	0.388	0	1
webuse2	1 if several times a week; 0 otherwise	0.117	0.322	0	1
Regional specific					
urban	1 if in an urban area; 0 otherwise	0.629	0.483	0	1
hbhtnrnt	Percentage of renter-occupier at block group level	22.44	20.67	0	95
gdp	GDP per capita by state ($1000)	49.49	8.78	32.96	174.5
popden	Population density by state	85.30	160.0	0.46	3999
Household related					
hhfaminc1	1 if HH income<$50,000; 0 otherwise	0.224	0.417	0	1
hhsize	Count of household members	2.909	1.277	1	14
hhvehcnt	Count of household vehicles	2.593	1.176	0	27
Person related					
r_age	Respondent age	47.85	12.52	18	92
gender	1 if female; 0 otherwise	0.468	0.499	0	1
educ	Education level (Base case: graduate or professional degree)				
educ1	1 if high school or less; 0 otherwise	0.202	0.402	0	1
educ2	1 if bachelor's degree; 0 otherwise	0.576	0.494	0	1
wkftpt	1 if work full time; 0 otherwise	0.807	0.395	0	1
Travel pattern related					
shoptrp	Number of shopping trips on travel day	0.824	1.176	0	11
cnttdtr	Total number of trips on travel day	4.844	2.565	2	27
trvlmin	Average travel time per trip (minute)	21.78	21.41	1	615
dweltime	Total dwell time on travel day (hour)	7.793	3.688	0	25.32
strttime_pct	Percentage of trips starting after 6 pm	0.171	0.218	0	1
trptrans	Percentage of transportation mode used (Base case: all by transit or other)				
mode1_pct	Percentage of trips by motor vehicles	0.915	0.206	0	1
mode2_pct	Percentage of trips by bicycle and walk	0.070	0.178	0	1
tdwknd	If travel day is weekend (indicator)	0.298	0.440	0	1
gasprice	Gas price (dollar)	2.859	0.943	1.50	4.46

Note: A travel day is defined as the 24-h period starting from 4:00 AM of the day assigned until 3:59 AM of the following day. On weekends the travel day starts from Friday 6:00 PM and ends on Sunday at midnight (FHWA, 2014).

Table 3
Key variable distributions in original and final datasets.

Variable	Description	Percentage in original dataset (%)	Percentage in final dataset (%)
Gender	Female = 1	53.7	46.8
	Male = 0	46.3	53.2
hhfaminc1*	<$50,000 = 1	39.4	22.4
	>$50,000 = 0	53.9	77.6
webuse1*	Everyday = 1	49.5	81.5
	Otherwise = 0	37.9	18.5
educ2*	Bachelor = 1	41.6	57.6
	Otherwise = 0	42.9	42.4

Note: "*" indicates that invalid records exist in the original file, leading to a total less than 100%.

The structural relationship depicted in Fig. 1 can be generally expressed as:

$$Y = BY + \Gamma X + \zeta \tag{1}$$

where Y represents variables of interest, including "prop_purchase" and "prop_shoptrp"; X is the vector of exogenous variables and measurement variables, including the internet-related, regional specific, household-related, personal-related and travel pattern related variables; B and Γ are estimable coefficient matrices and ζ represents errors.

Regarding the estimation methods for SEM, as summarized by Fan and Wang (1998), there were generally three types of estimation approaches: (1) Generalized Least Square (GLS); (2) Maximum Likelihood (ML); and (3) the Asymptotically Distribution Free (ADF) estimator. GLS and ML are often used for normally distributed data with independent factors and errors. ADF estimator can be used with large (over 2500) sample size for non-normally distributed data. This paper uses the ML approach for estimation.

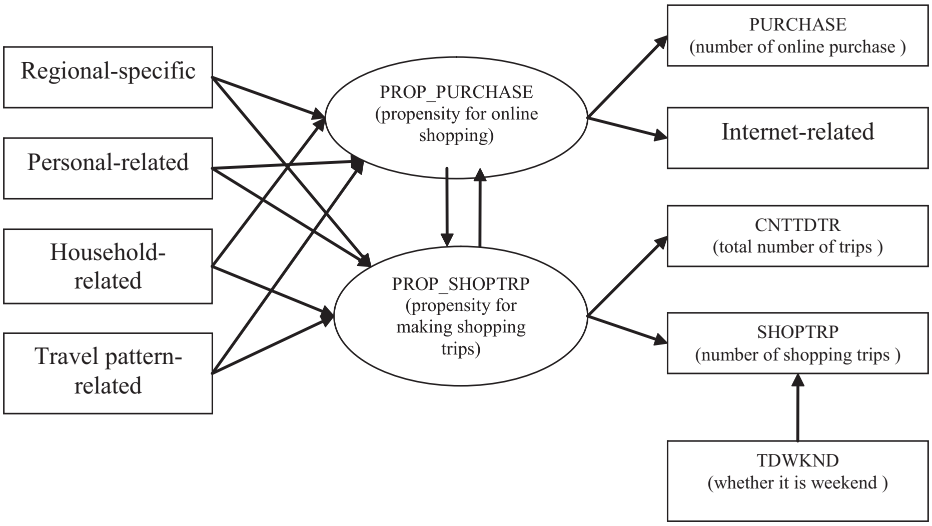

Fig. 1. SEM model structure.

Table 4
Direct effects of national level results.

	Structural part		Measurement part			
	prop_purchase	prop_shoptrp	purchase	webuse1	cnttdtr	shoptrp
prop_shoptrp	−0.125*** (−4.30)	–			0.849*** (36.75)	0.594*** (38.51)
prop_purchase	–	0.214*** (8.72)	0.280*** (14.39)	0.571*** (22.69)		
cons			0.677*** (13.40)	2.467*** (28.77)	2.028*** (28.34)	0.672*** (13.23)
Regional specific						
urban	0.071*** (4.44)	−0.012 (−1.11)				
hbhtnrnt	0.017 (1.04)	0.024(2.02)				
gdp	0.028 (1.48)	−0.014 (−1.13)				
popden	0.018 (1.11)	−0.011 (−0.98)				
Household related						
hhfaminc1	−0.256*** (−13.28)					
hhsize	−0.048** (−2.74)	0.044*** (3.27)				
Person related						
r_age	−0.118*** (−6.65)	0.128*** (9.94)				
educ						
educ1	−0.441*** (−19.49)					
educ2	−0.189*** (−11.90)					
wkftpt	0.032 (1.91)	−0.071*** (−6.38)				
Travel pattern related						
trvlmin	−0.021 (−1.34)	−0.261*** (−16.67)				
tdwknd						0.166*** (19.87)
N			85,663			
overall R^2			0.3046			

Note: t statistics in parentheses.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

The same model form is used to analyze shopping behavior at both national and state levels. The following section first presents the findings at the national level, and then summarizes estimates at state level.

5. Results analysis

5.1. National level results

As the variables of interest are propensities to shop online and making shopping trips, which do not have to be defined by certain units, the estimation results are analyzed based on standardized coefficients. Table 4 presents the standardized coefficients of the full model including both the structural part and the measurement part. Similar to other conventional econometric models, these coefficients measure the direct effects between variables. As SEM allows the assessment of bidirectional structural relationship, the simultaneity in the system will also cause indirect effects. The summation of the direct effect and the indirect effect is called the total effect, which measures the overall effect of one variable on the endogenous dependent variable after accounting for all the simultaneous effects. Table 5 provides the estimates of the total effects.

According to Table 4, all variables in the measurement part are statistically significant with high *t*-statistics, confirming that these variables measure latent variables well, and people indeed tend to make more shopping trips during weekends.

For the structural part, the results disclose the interactive relations between online shopping and shopping trips. First, people making more shopping trips tend to shop online less frequently, as indicated by the direct effect −0.125 and the total effect −0.122. The results suggest that if the frequency of shopping trips increases by one standard deviation, the frequency of online shopping will eventually decrease by 0.122 standard deviation. It seems that the in-store shopping experience suppresses the desire of online shopping. Hence people making more shopping trips are less likely to shop online.

On the other hand, the direct effect of "prop_purchase" on "prop_shoptrp" is 0.214, and the total effect is 0.208. Such a positive effect is also found by Ferrell (2004), who suggested that teleshopping had positive influence on shopping trip frequency. Cao et al. (2012) also identified a positive relationship between online searching frequency and in-store shopping frequency. Apparently, frequent online shoppers tend to make more shopping trips at the same time, probably because they need to feel and experience some of the commodities in a mortar and brick store.

In contrast to most previous studies that found either substitution or complementary effects between frequency of shopping trip and online shopping, this study suggests that their relationship is neither pure substitution nor pure complement. In-store shopping suppresses the demand of online shopping while online shopping seems to generate more shopping trips. The interaction is asymmetric. Such an asymmetric relationship has also been suggested by previous studies. For example, Farag et al. (2005) indicated a positive relation between online shopping and shopping trips. However, shopping trips had a stronger impact on online shopping (Farag et al., 2007) than the other way round.

Important findings are also discovered between exogenous variables and the shopping propensities. A person living in urban area tends to make more online shopping than their non-urban counterpart. This is expected as people living in urban areas are normally prone to adopting new technologies and have higher propensity for online shopping. The effect of urban location on physical shopping trips, however, is statistically insignificant in both direct effect and total effect. Such insignificance may be a result of the neutralization between various conflicting factors: As Ferrell (2005) suggested, people living near retail opportunities make shopping trips more often. On the other hand, congestion and tight daily schedule may reduce the shopping trip frequency for urban residents. High percentage of renters is positively related to both shopping propensities although the effects are statistically insignificant. One potential explanation is that renters tend to be frequent movers compared to house owners, and therefore have the need to shop more frequently. Higher state-level GDP and population density seem to be related to higher online shopping frequency and lower in-store shopping frequency although the effects are statistically insignificant. The findings suggest that wealthier and highly populated areas tend to have more frequent

Table 5
Total effects of national level results.

	Latent variables	
	prop_purchase	prop_shoptrp
prop_shoptrp	−0.122***	−0.026
prop_purchase	−0.026	0.208***
Regional specific		
urban	0.071***	0.003
hbhtnrnt	0.014	0.027*
gdp	0.029	−0.008
popden	0.019	−0.007
Household related		
hhfaminc1	−0.249***	−0.053***
hhsize	−0.052**	0.033**
Person related		
r_age	−0.131***	0.100***
educ		
educ1	−0.429***	−0.092***
educ2	−0.184***	−0.039***
wkftpt	0.040*	−0.063***
Travel pattern related		
trvlmin	0.011	−0.260***

Note:
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

online shopping activities. Households with lower income (<$50,000) tend to shop less frequently, both online and in-store. This result is consistent with the findings by Srinivasan and Bhat (2005) who concluded that people with higher income tend to make more shopping trips than low income people. Larger households, compared to smaller ones, shop online less frequently but will go to stores more frequently. A relevant finding in Srinivasan and Bhat (2005)'s study also suggested that household with children tend to make more shopping trips. Older people shop online less frequently but make more shopping trips. Other empirical studies also found that older people tend to make more shopping trips than young people (Yun and O'Kelly, 1997). Compared to people with graduate or professional degrees, shoppers with high school or bachelor's degrees shop online less frequently. Psychological studies (Swinyard and Smith, 2003) also confirmed that highly educated people are more likely to shop online. As for the working status, full-time worker tend to make more online shopping and less shopping trips. Cao et al. (2012) found the same impact of being full time worker. As expected, people spending more time on daily travel tend to make less shopping trips. The direct effect of travel time on online shopping frequency is also negative but statistically insignificant. The total effect turns to positive. Apparently, the direct effect is overruled by the indirect effect. This has a rather interesting interpretation: shopping demand of people who have to spend a lot of time traveling is suppressed, both online and in-store. To compensate the lack of in-store shopping, they shop online, which requires less time and physical activities compared to shopping in stores. The overall effect is that people who spend more time traveling shop online more frequently. This interesting insight has not been discovered in previous studies. This finding sheds light on transportation condition and its impact on online shopping and shopping trips.

All findings above provide important reference to study online shopping and its impact on transportation. Professionals may take inspiration from these findings and make better decisions in practice.

5.2. State level results

The above analysis discloses general relationship between online shopping, shopping trips and various socio-economic features in the U.S. However, the insignificance of many intuitively influential factors implies that the national average based on pooled data may have obscured some factors' impacts because the spatial variation is not recognized. As also indicated by previous studies, shopping behavior could vary substantially across different regions. It is thus helpful to further explore the spatial variation by examining shopping behavior at the state level, which is expected to provide further insights into the regional differences and the effects of various factors.

Using the same model (state-level GDP and population density are dropped from the analysis), shopping behavior is analyzed for each state separately. Out of the 50 states and District of Columbia (DC), 24 have non-convergent results due to small sample size or lack of variable variation. Mean, standard deviation, minimum and maximum of the standardized coefficients for the remaining 27 states are summarized in Table 6. For state level analysis, the confidence level is set to be 90% considering the smaller sample sizes.

The variation across the states is clearly indicated by the large standard deviation and the wide range of the coefficient estimates. The impact of a variable could be statistically significant in some states while insignificant in others, and the directions of the impacts also differ. For example, frequency of shopping trips is estimated to significantly impact online shopping in 4 states: In Iowa (IA), New York (NY) and Wisconsin (WI), the effects are negative and in Mississippi (MS) it is positive.

Table 6
Direct effects of state level analysis.

	Latent variables									
	prop_purchase					prop_shoptrp				
	Mean	SD	Min	Max	No. of significant cases	Mean	SD	Min	Max	No. of significant cases
prop_shoptrp	−0.186	0.472	−2.162	0.761	4	–	–	–	–	–
prop_purchase	–	–	–	–	–	0.275	0.349	−0.220	1.696	13
Regional specific										
urban	0.035	0.156	−0.356	0.435	9	−0.010	0.149	−0.302	0.516	5
hbhtnrnt	0.041	0.168	−0.343	0.639	5	0.029	0.106	−0.213	0.282	4
Household related										
hhfaminc1	−0.248	0.168	−0.591	0.109	20	–	–	–	–	–
hhsize	−0.036	0.144	−0.423	0.282	5	0.053	0.082	−0.104	0.284	2
Person related										
r_age	−0.089	0.180	−0.670	0.332	8	0.156	0.098	−0.051	0.439	17
educ										
educ1	−0.498	0.277	−1.745	−0.133	19	–	–	–	–	–
educ2	−0.247	0.272	−1.457	0.106	15	–	–	–	–	–
wkftpt	−0.012	0.143	−0.366	0.201	4	−0.081	0.084	−0.239	0.169	7
Travel pattern related										
trvlmin	0.018	0.104	−0.315	0.215	1	−0.269	0.128	−0.570	0.072	22
N	Mean: 2791		SD: 4014		Min: 106		Max: 12836			

Table 7
Results comparison between different studies.

Studies	Key variables and impacts on online shopping	Data and region	Conclusions
Ferrell (2005)	Household income: + Shop accessibility: + Employment: +* Time-starved: +*	San Francisco Bay Area Travel Survey 2000 (BATS 2000)	Small substitution effect Home teleshopping time→shopping travel time (−)
Cao et al. (2012)	Household income: +** Education: +* Urban location: +** Full time worker: +** Shopping attitude: −	539 adults survey in Minneapolis-St. Paul seven-county metropolitan area	Complementary effect: online shopping→ shopping trips (+)
Farag et al. (2007)	Urban location: +* Shop accessibility: + Female: − Income: −* Age: − Shopping attitude: +	826 respondents questionnaire in four municipalities in Netherlands	Complementary: shopping trips→online shopping (+) online shopping→ shopping trips(+)
Weltevreden and Rietbergen (2007)	Education: + Male: − Internet access: +*** Shopping enjoyment: −	3074 respondents online survey in 8 cities in Netherlands	Complementary effects in short run but substitution in long run
This paper	Household income: +*** Education: +*** Urban location: +** Full time worker: +*** Shopping attitude: −* Travel time: +	NHTS data at both national and state levels	Asymmetric impacts: online shopping→ shopping trips (+) shopping trips→online shopping (−)

Note: "+" indicates positive impact and "−" indicates negative impact.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Similarly, being a full working is found to have significant impacts on online shopping frequency in 4 states. The effect is negative in Colorado (CO) and Iowa (IA) and positive in Illinois (IL) and Wisconsin (WI). It is meaningful to ask why some states are being statistically significant while others are not, and why the same variable's effect varies so much across states. The publicly available NHTS data has removed geospatial information due to concerns over confidentiality, making it impossible to acquire detailed local built environment information. Therefore, this study only indicates the existence of regional variation. Further investigation needs to rely on the micro-level, geocoded NHTS add-on data maintained by states and MPOs. The add-on data will allow the integration with information related to retail employment density, transportation accessibility and household locations, all important for the shopping behavior analysis. States and MPOs can conveniently adopt the models developed in this research and use the enhanced micro-data to conduct in-depth investigation on the shopping behavior in their respective regions.

6. Conclusions

The growth of ecommerce has great implication for both freight and passenger transportation systems. Understanding the relationship between online shopping and in-store shopping is the critical step to assess the impact of ecommerce on transportation. This paper investigates the relationship between online shopping and shopping trips using NHTS 2009 data. SEM model is used for analysis. Results indicate the interaction is not pure substitution or pure complement. Online shopping stimulates shopping trips while shopping trips tend to suppress online shopping. Results also show that both variables are affected by regional specific, household related, person related and travel pattern related variables. State level results confirm the model validity and indicate that spatial variation indeed exists. A comparison of this paper with a few highly cited previous studies is summarized in Table 7.

Building on the nuanced findings of the previous studies, this paper makes important contributions in three aspects that will benefit future transportation planning and policy making. First, this paper reveals the complex interactions between online shopping and shopping trips. By allowing the assessment of bidirectional impacts, an asymmetric relationship is identified between online shopping and shopping trips.

Second, results are presented at both national level and state level. Previous studies, partially due to limitation of data, mainly focus on regional level. This study provides the big picture at a macro level and also discloses state variation, which should be further explored in future studies as detailed state-level data becomes available.

Third, NHTS 2009 provides newest and most comprehensive dataset compared to other studies, which are limited in sample size and contents due to time and budget. The dataset used in this paper contains a large sample size of 85,663 records with detailed online shopping, regional specific, household person, social demographics and travel information. It greatly ensures the validity, accuracy and promptness of the results.

In summary, this paper provides new insights into the connections between online shopping and the transportation system building on a statistically rigorous model and a comprehensive dataset. Future work should collect and integrate other data sources to further improve the model specification. For example, the integration of spatial information with NHTS and the consideration of commodity types may be great supplements for this study, but special attention should be paid for possible bias due to the skew of the cleaned data used in this study. In addition, the SEM model could be further refined, more variables could be added and other methods such as joint decision models could be tried.

Technologies are changing the world and people's way of living. The internet brings people much benefit by shopping online. However, assessing its impact on transportation system, including both freight and personal transportation, in both short run and long run, remains a big challenge for transportation researchers and practitioners. By quantifying the relationship between online shopping propensity and shopping trip frequency, this paper makes an important dent in analyzing the impacts of ecommerce on transportation. The modeling approach and the findings provide important reference for transportation planners and policy makers.

References

Aguiléra, A., Guillot, C., Rallet, A., 2012. Mobile ICTs and physical mobility: review and research agenda. Transport. Res. A Policy Pract. 46 (4), 664–672.

Calderwood, E., Freathy, P., 2014. Consumer mobility in the Scottish isles: the impact of internet adoption upon retail travel patterns. Transport. Res. A Policy Pract. 59, 192–203.

Cao, X., Xu, Z., Douma, F., 2012. The interactions between e-shopping and traditional in-store shopping: an application of structural equations model. Transportation 39 (5), 957–974.

Choo, S., Lee, T., Mokhtarian, P.L., 2008. Do transportation and communications tend to be substitutes, complements, or neither? U.S. consumer expenditures perspective, 1984–2002. Transp. Res. Rec. 2010, 121–132.

Fan, X., Wang, L., 1998. Effects of potential confounding factors on fit indices and parameter estimates for true and misspecified SEM models. Educ. Psychol. Measur. 58, 701–735.

Farag, S., Schwanen, T., Dijst, M., 2005. Empirical investigation of online searching and buying and their relationship to shopping trips. Transp. Res. Rec. 1926, 242–251.

Farag, S., Weltevreden, J., van Rietbergen, T., Dijst, M., van Oort, F., 2006. E-shopping in the Netherlands: does geography matter? Environ. Plann B Plann. Des. 33 (1), 59–74.

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. A Policy Pract. 41 (2), 125–141.

Federal Highway Administration (FHWA), 2012. 2009 National Household Travel Survey. <http://nhts.ornl.gov/download.shtml>> (December 26, 2012).

Federal Highway Administration (FHWA), 2014. 2009 National Household Travel Survey. <http://nhts.ornl.gov/2009/pub/UsersGuideV2.pdf> (September 25, 2014).

Ferrell, C., 2004. Home-based teleshoppers and shopping travel: do teleshoppers travel less? Transp. Res. Rec. 1894, 241–248.

Ferrell, C., 2005. Home-based teleshoppers and shopping travel: where do people find the time? Transp. Res. Rec. 1926, 212–223.

Krizek, K., Li, Y., Handy, S., 2005. Spatial attributes and patterns of use in household-related information and communications technology activity. Transp. Res. Rec. 1926, 252–259.

Mokhtarian, P.L., 2004. A conceptual analysis of the transportation impacts of B2C e-commerce. Transportation 31 (3), 257–284.

Pearl, J., 2013. Linear models: a useful, "microscope" for causal analysis. J. Causal Inference 1 (1), 155–170.

Ren, F., Kwan, M.P., 2009. The impact of geographic context on e-shopping behavior. Environ. Plann. B Plann. Des. 36 (2), 262–278.

Sim, L.L., Koi, S.M., 2002. Singapore's internet shoppers and their impact on traditional shopping patterns. J. Retail. Consum. Serv. 9, 115–124.

Srinivasan, S., Bhat, C.R., 2005. Modeling household interactions in daily in-home and out-of-home maintenance activity participation. Transportation 32 (5), 523–544.

Swinyard, W.R., Smith, S.M., 2003. Why people (don't) shop online: a lifestyle study of the internet consumer. Psychol. Mark. 20 (7), 567–597.

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, 270–280.

U.S. Census Bureau, 2013. The 4th Quarter 2012 Retail E-Commerce Sales Report. <http://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf> (March 13, 2013).

Weltevreden, J.W.J., Rietbergen, T.V., 2007. E-shopping versus city centre shopping: the role of perceived city centre attractiveness. Tijdschrift voor economische en sociale geografie 98 (1), 68–85.

Yun, D.S., O'Kelly, M.E., 1997. Modeling the day-of-the-week shopping activity and travel patterns. Socio-Economic Plann. Sci. 31, 307–319.