Australasian Transport Research Forum 2011 Proceedings
28 - 30 September 2011, Adelaide, Australia
Publication website: http://www.patrec.org/atrf.aspx

The role of attitudes and public transport service on vehicle ownership in Ho Chi Minh, Vietnam

Chinh Quoc Ho[1], Toshiyuki Yamamoto[2]

[1]Institute of Transport and Logistics Studies, the University of Sydney Business School, Australia

[2]EcoTopia Science Institute, Nagoya University, Japan

quoc.ho@sydney.edu.au

Abstract

Empirical evidence suggests that urban form has a strong influence on travel behaviour. This finding, however, has been challenged by the debate on the confounding effects of subjective dimensions such as individual attitudes and preferences (the self-selection hypothesis). Residential self-selection has been found to significantly influence travel behaviour though the effects of land use remain important. Although people have more options than just residential choice to self-select, researchers have focused on residential self-selection only, mostly in the US. The paper examines the effects of attitudes and public transport service on household multiple vehicle ownership behaviour, controlled for socio-demographic and land use characteristics. This study contributes empirical results of Generalised Nested Logit models using large scale household interview survey data collected in Ho Chi Minh metropolitan area, Vietnam. The study finds that both subjective and objective dimensions of the built environment such as bus coverage, bus operators' attitudes, walking and cycling conditions, and the diversity of land use at residential locations are important to multiple vehicle ownership behaviour while the effects of self-selection are relatively modest.

1. Introduction

With hundreds of studies on the relationship between the built environment and travel behaviour, it is quite clear that there is a correlation between these two factors. Generally, residents of pedestrian-friendly neighbourhoods walk more than those of non-walkable neighbourhoods; and people living in areas with better public transport services tend to use public transport more, compared to those living in areas with poorer quality public transport. However, significant correlation between land use patterns and travel does not necessarily mean that changes in land use patterns will trigger changes in travel behaviour. More specifically, confounding factors such as residential self-selection may cause a spurious relationship between the built environment and travel. For instance, residents who prefer using public transport may intentionally choose to live in neighbourhoods that are well served by public transport. Thus, we observe the association between the built environment and travel behaviour because people who prefer public transport also like to live in public transport oriented developments. This problem is referred to as residential self-selection in the literature.

Self-selection relevant for travel behaviour extends beyond the scope of residential choice. People can theoretically self-select with respect to other choices such as destination locations, travel modes, and levels of exposure to negative traffic externalities (van Wee, 2009). However, due to concerns regarding respondent burden and difficulties in measuring attitudes, most of the existing studies, if controlling for self-selection at all, limit self-selection to residential location only. The research contribution of this analysis lies in the ways it extends current understandings of self-selection effects – mainly focused on residential self-selection – to those with respect to negative traffic externalities on household multiple vehicle ownership behaviour.

The remainder of this paper is organised as follows. Section 2 briefly reviews the literature and establishes background for this study. Section 3 introduces the data source and method of the analysis, followed by estimation results in Section 4. Section 5 puts the developed model into practice. Conclusions and policy implications are drawn in Section 6.

2. Research and policy background

2.1. Effect of the built environment and residential self-selection on travel

There has been much study on the role of land use factors in forming individuals' travel behaviour in general, and vehicle ownership in particular. The research findings may be summarised in the statement that residents living in dense areas with mixed land use and good access to public transport service generate shorter trips, and lower rate of car travel (Cervero and Kockelman, 1997; Boarnet and Crane, 2001; Ewing and Cervero, 2010). This can be explained firstly by the variety of activities that are accessible by alternative modes to the car, and secondly by traffic conditions and restrictions for car travel (such as parking, road pricing) in these areas which encourage people to use public transport or non-motorised modes.

These findings, however, have been challenged by the debate on the confounding effect of residential self-selection. Evidence on residential self-selection is ample (see Cao et al. 2009). Cao et al. (2009) reviewed 38 empirical studies and concluded that residential self-selection exists in all the studies, and that the effect of the built environment on travel behaviour is significant and independent of residential self-selection. Understanding the confounding effect of residential self-selection is very important for transport planners to ascertain whether changes to the built environment are a cost-effective way to change travel behaviour. This is because the observed associations between the built environment and travel behaviour may due in large part to the effect of individuals' preferences and attitudes. Accordingly, one implication is that the effectiveness of using land use measures to change travel behaviour may be mainly limited to the market share of individuals whose attitudes are favourable to sustainable transport and/or location choices.

Most of the studies on self-selection focused geographically on the US with high car use. Further, the literature has not yet drawn the boundary between different types of self-selection. Consequently, some kinds of self-selection which have nothing to do with residential location choice are seen as and therefore combined with *residential* self-selection, though this is just a formal term that says little about the reasons why traffic congestion in the CBD, for instance, should influence travel behaviour. This paper aims to explore the relevance of other aspects of self-selection to better understand the land use – travel relationship. These aspects are discussed in the next section.

2.2. Self-selection with respect to negative traffic externalities

2.2.1. Traffic congestion

Some people may be more averse than others to traffic congestion. There is some evidence that commuters perceived congestion differently (Redmond and Mokhtarian, 2001). With the same level of traffic congestion in the CBD, people who are more averse than others to traffic congestion may not choose the CBD as a location for a family or an individual activity. They instead consider alternative locations (e.g., suburban centres) to which public transport services may not be as good. Consequently, they travel by car. Thus, the effect of traffic congestion on travel mode choice and/or vehicle ownership will be wrongly estimated if individuals' averseness to traffic congestion is ignored. Here, people self-select destinations according to their concerns and judgement on the level of traffic congestion at alternative destinations.

It may also be possible that people who are more concerned about traffic congestion perceive it as more serious than others. Their perception of traffic congestion may in turn influence the relative attractiveness of the private car due to the less significant difference in travel speed between car and other modes. On the other side, due to environmental protection objectives calling for a reduction in private motorised vehicle ownership and use, people who think traffic congestion is a real concern may supportively choose not to use or own private motorised modes. Thus, the effects of the real level of traffic congestion on mode choice or vehicle ownership will be biased if attitudes to traffic congestion are not accounted for.

Attitudes to traffic congestion can influence travel behaviour both directly and indirectly through location choices. Firstly, people who dislike traffic congestion may avoid it by not generating trips. For example, they can work from home or use technology such as telephone and internet to minimise travelling. Secondly, they can choose to locate and work in less congested areas. Thirdly, when they have to travel for non-work purposes, less congested destinations are prioritised over congested ones, all else equal. Alternatively, they can choose the time of travel to avoid traffic congestion in peak hours. Therefore, the relevance of ignoring self-selection with respect to traffic congestion may be comparable to ignoring residential self-selection.

2.2.2. Traffic safety

Some people may prefer to own and travel by car because they feel safer than other modes. Similarly, safety concerns may prevent some people from owning and using motorcycle and bicycle. People can self-select the type of vehicle to own and use according to their perceptions and acceptance of safety levels. This self-selection may also have impacts on route choice, and therefore travel distance (van Wee, 2009). This type of self-selection is highly correlated with residential self-selection and characterised by ‘safety of car’ variable in the literature (Handy et al., 2006). Ignoring self-selection with respect to traffic safety may result in a fault in estimating mode choice and vehicle ownership behaviour (Cao et al., 2007).

2.2.3. Noise and air pollution

People may choose a residential location based on their noise sensitivity and health-related concerns. A process of self-selection may take place on the basis of exposure to air and noise pollution. People with high sensitivity to noise or with health-related concerns might choose to live in quiet neighbourhoods or regions with low exposure to air pollution, respectively. These areas are normally found in suburban areas where public transport service is also less frequent. This process should be referred to as residential self-selection because individuals’ preferences for noise and air quality influence residential location choice, thereby the associated residential built environments. The only research focusing on a self-selection process due to noise sensitivity carried out by Nijland et al. (2007) found no statistical association between noise sensitivity and residential location in the Netherlands. One of the reasons they argued for this finding was that home buyers may trade off noise levels for other qualities of the dwelling and the neighbourhood, especially in the tight housing market. However, in a dwelling market driven by demand side rather than supply side factors, self-selection due to exposure to noise and air pollution may exist, impacting residential location choice thereby travel mode choice and vehicle ownership.

Preferences for noise and air pollution may also relate to choices other than residential location such as vehicle ownership and use. The level of exposure to air and noise pollution for different modes is clearly different. Thus, people who strongly prefer a lower level of noise and air pollution may choose to own and use a car instead of a motorcycle or a bicycle.

In the empirical analysis presented in sections 3 and 4, these three aspects of self-selection are combined and referred to as self-selection with respect to negative traffic externalities.

3. Method and data source

3.1. Data description

The paper seeks to expand the understanding of choice factors and provide empirical evidence of self-selection in environmental settings of developing countries, using the first general purpose large scale urban household interview survey in Ho Chi Minh metropolitan area, Vietnam (Almec Corporation, 2004). The study area had a population of 7 million, of which 5 million lived in Ho Chi Minh City (HCMC) while the remaining lived in districts of adjacent provinces which are currently forming or will form part of the metropolitan area from the viewpoint of regional development. The survey contained 28,001 households, equivalent to 102,407 persons (sampling rate: 1.45% households and 1.34% individuals), conducted from August 2002 to June 2004. Both socio-economic and attitudinal data for households and individuals, as well as vehicle ownership, travel behaviour, location, mobility restrictions, and overall assessment of basic urban services, were collected. The data collected included:

- Household: size, type of dwelling and tenancy, number of workers, number of children, income, number of owned cars, motorcycles, bicycles, and total vehicle operating cost.
- Individual: age, gender, education, occupation, typical working hours, driving licence, vehicle ownership, income, mobility restriction due to physical/mental condition.
- Travel diaries: daily logs of all trips made by each household member, including trip origin and destination, trip purpose, travel mode, parking location, departure and arrival time, travel mode, reason for mode choice, and reason for not choosing bus.
- Perceived characteristics: individual's overall assessment of key urban issues (e.g., environment protection, security and safety, economic development), urban services (e.g., recreation and sport facilities), transport aspects (e.g., walking and cycling infrastructure, traffic control and traffic safety), bus service, motorcycle mobility, and walking and cycling conditions both in community and in city centre.
- Attitudinal data: Attitudes towards negative traffic externalities, bus services, private travel modes and leisure preferences.

Because vehicle ownership is considered as a medium to long-term decision which normally occurs at the household level, the household is chosen as the analysis unit. Given that the aim of this study is to explore the role of attitudes and public transport service on household vehicle ownership, for a household's observation to be selected, at least one household member provided an overall assessment on urban and transport aspects. The final estimation data set included 22,830 households.

3.2. Variables

Disaggregate choice models have been typically used to describe car (vehicle) ownership behaviour where vehicle ownership has been defined based either on the number of owned vehicles (Train, 1980), or on various vehicle attributes such as vehicle type (Choo and Mokhtarian, 2004; Bhat et al., 2009), engine displacement (Zhang et al., 2009), car size (Hayashi et al., 2001), fuel type and automaker (Koh, 2003). In Vietnam, vehicle type is usually classified into three categories: bicycle, motorcycle and car. Further, it is a common practice that a household owns more than one vehicle of the same or different types. Therefore, this study defines vehicle ownership based on both the number of owned vehicles and vehicle types where the common term vehicle refers to car, motorcycle or bicycle. This classification is not only well known to vehicle users in Vietnam but also directly related to evaluation of emissions, fuel consumption, sustainability and effects of motorized vehicle-related taxation.

3.2.1. Vehicle ownership

Table 1 shows the household distribution by the number of vehicles and vehicle types. The sample was first grouped into seventeen categories presented in the first column of Table 1. Households with one bicycle, two motorcycles and no car, for example, fall into the category *B1MC2*. Similarly, households with three or more bicycles and no motorised vehicle are classified into the category *B3*$^+$.

Based on the observed distributions, the seventeen categories are grouped into ten choice alternatives for modelling. For convenience, the ten alternatives are numbered from 1 through 10 and indicated in the last column of Table 1. In defining choice set alternatives, the objective of representing variation in observed household distributions was balanced against the ability to estimate significant coefficients and the practical needs for computational tractability.

Table 1 Household distribution by the number of vehicles and vehicle types (vehicle bundle)

Vehicle	No. of vehicles by type			No. of households		Modelled vehicle bundle			
bundle[a]	Bicycle	MC	Car	Cases	%	Name	Cases	%	Alt.[b]
No vehicle	0	0	0	134	0.6%	No vehicle	134	0.6%	1
B1	1	0	0	505	2.2%	B1	505	2.2%	2
B2	2	0	0	347	1.5%	B2$^+$	555	2.4%	3
B3$^+$	3$^+$	0	0	208	0.9%				
MC1	0	1	0	3,124	13.7%	MC1	3,124	13.7%	4
MC2	0	2	0	4,634	20.3%	MC2$^+$	7,803	34.2%	5
MC3$^+$	0	3$^+$	0	3,169	13.9%				
B1MC1	1	1	0	3,219	14.1%	B1MC1	3,219	14.1%	6
B1MC2	1	2	0	2,732	12.0%	B1MC2$^+$	4,122	18.1%	7
B1MC3$^+$	1	3$^+$	0	1,390	6.1%				
B2MC1	2	1	0	1,275	5.6%	B2$^+$MC1	1,602	7.0%	8
B3$^+$MC1	3$^+$	1	0	327	1.4%				
B2MC2	2	2	0	851	3.7%	B2$^+$MC2$^+$	1,392	6.1%	9
B3$^+$MC2	3$^+$	2	0	138	0.6%				
B2MC3$^+$	2	3$^+$	0	320	1.4%				
B3$^+$MC3$^+$	3$^+$	3$^+$	0	83	0.4%				
CAR[c]	-	-	1$^+$	374	1.6%	CAR[c]	374	1.6%	10
Total				22,830	100%		22,830	100%	

[a] B = Bicycle, MC = Motorcycle; [b]Alt. = Alternative.

[c] Including all bundles containing car regardless of the number of bicycles and motorcycles.

Similar shaded rows (vehicle bundles) are combined for the modelling purpose.

3.2.2. Explanatory variables

The built environment is measured using perceived transport condition characteristics as well as objective measures of accessibility. For perceived characteristics, respondents were required to evaluate, on a 5-point scale from “very bad” (1) to “very good” (5), how well various aspects of transport are performing. Through principal component factor analysis, these aspects were reduced to three factors shown in Table 2. Perceived public transport services were measured in a similar manner by asking respondents to assess a series of aspects about bus service, reducing them to five factors using principal component analysis (see Table 2).

Respondents were also asked to select the three most critical urban issues, the five most important basic urban services, the three most important transport aspects, the three most important bus service aspects, and the most important feature of motorcycle mode from a list of corresponding items. Based on these selected most important aspects, individual subjective dimensions of travel attitudes and preferences were inferred. For instance, if any item making up the “negative traffic externalities” factor was chosen as one of the three most important transport aspects, negative traffic externalities are one of the individual’s concerns. This was then transformed into a dummy variable representing attitude towards negative traffic externalities.

Table 2 Factors for perceived characteristics of transport supply and travel mode*

Factor	Item	Loading
Perceived transport condition characteristics		
Negative traffic externalities	Traffic safety	.814
	Traffic enforcement	.804
	Air pollution from vehicles	.737
	Traffic control measures	.636
Transport supply conditions	Travel conditions	.790
	Roads and facilities	.778
	Public transport services	.586
Parking/ walking conditions	Parking at destination	.802
	Parking at home	.704
	Walking condition	.614
Perceived public transport services		
Bus coverage	Operating hours	.847
	Routing	.803
	Frequency	.696
Bus operators' attitudes	Drivers' attitude	.884
	Conductors' attitude	.872
Ease of bus use	Waiting condition	.808
	Access to bus stops	.770
	Bus colour/design	.313
	On-board comfort	.420
Bus vehicle quality	Vehicle quality	.802
	Travel speed	.635
	Bus colour/design	.604
Qualitative level of bus service	On-board comfort	.740
	On-board security	.656
	Fare	.641

Factor	Item	Loading
Perceived characteristics of motorcycle		
Convenience	Convenience	.818
	Speed	.702
	Freedom	.569
	Comfort	.349
Safety	Freedom	.408
	Safety	.855
	Comfort	.642
Perceived walking conditions in community		
Walking conditions in community	Safety	.437
	Comfort	.482
	Convenience	.400
Perceived walking conditions in city centre		
Walking conditions in city	Safety	.434
	Comfort	.506
	Convenience	.406
Perceived cycling conditions in community		
Cycling conditions in community	Safety	.422
	Comfort	.498
	Convenience	.435
Perceived cycling conditions in city centre		
Cycling conditions in city	Safety	.418
	Comfort	.477
	Convenience	.433

* Extraction method: principal component analysis; rotation method: Kaiser normalization; factor loadings lower than 0.3 are suppressed.

The other explanatory variables selected based on preliminary analysis include: the number of adults, the number of children, household income, large home size indicator, distance from home to CBD as the crow flies, mixed land use index, population density, and total vehicle operating cost. Table 3 shows variable definitions along with their descriptive statistics, excluding the factors shown in Table 2 which are all standardised to have means of zeros and standard deviations of ones.

For unchosen alternatives (bundles) in the choice set, it is better to ask the households to report those attributes in the survey. However, questions aiming to identify competing alternatives as well as their specific attributes were not included in the household interview survey. Thus, it is necessary to impute these attributes. At least two ways could be used to impute uncollected information: one is to use relevant information from external sources and the other is to use information from the survey. As mentioned, this was the first urban household survey in the study area; therefore, external data sources are not available for the desired attributes. Thus, the alternative specific attributes of unchosen bundles have been computed using relevant average values of similar households. This imputation assumes that households with some similar socio-demographic characteristics have the same choice set and share the same average alternative specific attributes. A problem with this imputation method is that it arbitrarily increases similarities in attributes among households, thereby reducing the inherent variations of these attributes which in turn may result in lower model goodness-of-fit (Zhang et al., 2009). In this study, the average values of vehicle operating cost of unchosen vehicle bundles were computed with respect to different criteria including household size, the number of children, household income, and residential location. It is found that vehicle operating cost imputed with respect to household income results in the highest model fit. Consequently, similar household income is adopted as the criterion to impute vehicle operating costs of unchosen alternatives in this study. A by-product of *partial* household-specific choice set has resulted from this imputation where *partial* means households with similar income have the same choice set. Effects of this imputation of vehicle operating cost as well as household-specific choice set on model accuracy are discussed in the next section.

Table 3 Explanatory variables and their descriptive statistics (excluding the standardised factors shown in Table 2)

Variable definition	Mean	Std.Dev.
Demographic and land use characteristics		
Number of adults[a]	3.773	1.375
Number of children less than or equal to 6 years of age	0.265	0.578
Annual household income (1k US$)	2.149	1.232
Owner of home with area larger than 50m^2 (0/1)	0.378	0.485
Distance from CBD in 10 km	0.855	0.843
Mixed land use index, calculated as unnormalised entropy[b]	1.046	0.100
Population density at residential zone (10,000/km^2)	2.103	1.807
Annual vehicle operating cost (100 US$)	1.325	0.901
Travel attitudes and preferences		
Attitude to environmental protection (0/1)	0.425	0.494
Attitude towards negative traffic externalities (0/1)	0.621	0.485
Leisure preferences (0/1)	0.195	0.397
Attitude towards bus operator's attitudes (0/1)	0.219	0.413
Attitude towards bus onboard services (0/1)	0.462	0.499
Motorcycle convenience preference (0/1)	0.175	0.380
Motorcycle speed preference (0/1)	0.160	0.366
Motorcycle freedom preference (0/1)	0.172	0.377
Motorcycle safety preference (0/1)	0.440	0.496
Motorcycle comfort preference (0/1)	0.055	0.227

[a] Number of household members over 6 years old.
[b] Unnormalised entropy = $-\sum[p_{it}\ln(p_{it})]$, where p_{it} is the relative frequency of the trips with purpose t (work, school, return home, and shopping) in the total number of trips attracted to each destination zone i (see, e.g., Cervero, 2002; Bodea et al., 2008).

3.3. Model specification

Given the way in which a choice set is defined, some bundled alternatives share a common element. Consequently, some alternatives may have correlations in terms of unobserved components. Also, three types of vehicle may have their own correlations. As a result, the standard logit model and conventional Nested Logit model are inappropriate for this issue. The Generalised Nested Logit (GNL) model (Wen and Koppelman, 2001), which allows alternatives to be allocated in more than one nest, is capable of gauging these correlations and is therefore adopted in this study.

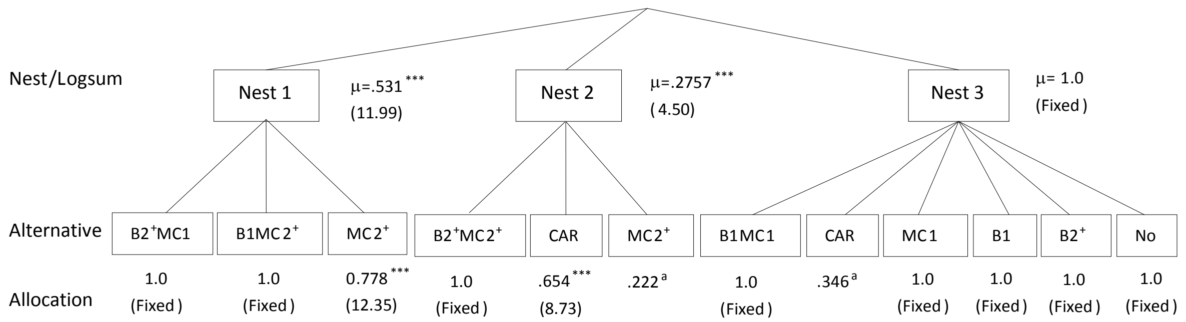
In the model presented in Figure 1 and Table 4, Alternative 1 (owning no vehicle) was treated as the reference option, meaning coefficients on the utility function should be interpreted in reference to the owning no vehicle alternative. Also, all parameters (except for vehicle operating cost) were first specified as alternative-specific coefficients under the assumption that socio-demographics and spatial factors have different effects on household vehicle ownership behaviour. In the interest of parsimony, parameters were constrained to be equal across bundled alternatives when their alternative-specific parameters appeared to be statistically indifferent.

4. Estimation results

4.1. Tree structure and model fit statistics

Figure 1 shows the resultant structure of a 2-level GNL model with three nests based on results of several dozen alternative Nested Logit models, and a comprehensive GNL structure specified with a maximum of four alternatives per nest and estimated with a quarter of the full dataset. This is due to the computationally expensive cost of estimation of GNL models. The adopted GNL model, however, was estimated with the full dataset. It has two *similarity nests* and one *dissimilarity nest*. The former nests (i.e., Nests 1 and 2) include portions of alternatives grouped together to represent different similarity relationships while the latter (i.e., Nest 3) includes portions of alternatives that are not allocated to any of the *similarity nests*. Figure 1 indicates that *MC2*$^+$ alternative is allocated in two nests. It is estimated that the probability of the *MC2*$^+$ alternative being allocated to nest 1 is 0.778. Similarly, the estimated probability of the *CAR* alternative being allocated to nest 2 is 0.654. The estimated logsum parameters (denoted by μ and shown next to each nest in Figure 1) represent the dissimilarity between alternatives within each nest. Together with allocation parameters, they provide the level of substitutions between alternatives within and across nests. The logsum parameters are significantly different from one another, implying the complex substitution patterns between vehicle bundles. The upper level suggests household choices for affordable, luxurious, and meagre bundles. The lower level shows household choices for the number of vehicles and vehicle types in each bundle.

Figure 1 Tree structure for Generalised Nested Logit model



***Significant at the 1% level; t-stat (in parenthesis) is calculated with respect to 1 for logsum and 0 for allocation parameters

[a] Allocation parameter = 1 - allocation parameter of the corresponding alternative in the other nest.

Table 4 presents the parameter estimation results for the above GNL model. Insignificant coefficients were discarded from the model, except for those having important policy implications and expected signs. McFadden's adjusted Rho-squared is 0.302 which indicates a relatively good fit to the data. Based on the result of the likelihood ratio test provided in Table 4, the GNL model soundly rejects the MNL specification with similar explanatory variables. Also, most of the parameter estimates are significant with logically expected signs. These results together suggest that the GNL model is appropriate to describe household multiple vehicle ownership behaviour, given that the alternatives were defined as bundles of vehicles in this study.

Table 4 Estimation results for the GNL model of household multiple vehicle ownership

Variable[a]	Coefficient[*]	Std. Err
Socio-economic attributes		
Annual household income in 1k US$ (2,3)	1.02980***	0.1296
Annual household income in 1k US$ (5,7,9)	2.03779***	0.1234
Annual household income in 1k US$ (4,6,8)	1.63549***	0.1232
Annual household income in 1k US$ (10)	1.50891***	0.1259
Number of adults (2,3,4,6)	.21149***	0.0661
Number of adults (5,7,8,9)	.81633***	0.0660
Number of adults (10)	.66676***	0.0720
Number of children (2-9)	.70208***	0.1923
Number of children (10)	.77486***	0.2067
Owner of home with area larger than $50m^2$, dummy (3,4,6,8)	.37778***	0.1064
Owner of home with area larger than $50m^2$, dummy (5,7,9)	.67332***	0.1076
Owner of home with area larger than $50m^2$, dummy (10)	1.32919***	0.1518
Annual vehicle operating cost in 100 US$ (all)	-7.37150***	0.0744
Annual vehicle operating cost squared (all)	1.71776***	0.0179
Land use characteristics		
Population density at residential zone in 10,000/km^2 (3,8)	-.17434***	0.0109
Population density at residential zone in 10,000/km^2 (6,7,9)	-.05951***	0.0060
Mixed land use index (5,10)	-.28505*	0.1529
Mixed land use index (6,7,9)	-.44928***	0.1516
Distance from CBD in 10 km (10)	0.09037	0.0679
Negative traffic externalities, perceived (2-9)	-.14328***	0.0404
Bus coverage, perceived (5,7)	-0.01425	0.0111
Bus coverage, perceived (10)	-.09801**	0.0450
Ease of bus use, perceived (2,3,4,6,8)	-.06004***	0.0145
Ease of bus use, perceived (7,9)	-.02513**	0.0115
Bus operator's attitudes, perceived (5,7,9,10)	-.04235***	0.0137
Walking conditions in CBD, perceived (all)	-.17988*	0.0999
Walking conditions in community, perceived (5,7)	-.03830***	0.0138
Walking conditions in community, perceived (6,8)	-.05469***	0.0180
Cycling conditions in community, perceived (2,3,6-10)	.02017*	0.0103
Safety of motorcycle mode, perceived (5)	.02494**	0.0104
Travel attitudes/preferences		
Motorcycle freedom preference, dummy (6)	.12861**	0.0557
Motorcycle comfort preference, dummy (6,8)	.16482***	0.0594
Motorcycle convenience preference, dummy(9)	.09089**	0.0461
Leisure preference, dummy (2)	-.32709**	0.1299
Attitude to negative traffic externalities, dummy (2-9)	-0.21076	0.1922
Attitude to negative traffic externalities, dummy (10)	-0.29073	0.2191
Constant term (2)	2.04823***	0.2634
Constant term (3)	3.54837***	0.2698
Constant term (4)	5.13598***	0.2610
Constant term (5)	5.77630***	0.3100
Constant term (6)	6.32715***	0.3082
Constant term (7)	5.71004***	0.3110
Constant term (8)	4.41875***	0.2641
Constant term (9)	4.64693***	0.3121
Constant term (10)	3.99308***	0.3664
Summary statistics		
Sample size	22830	
Log likelihood function value at convergence	-29469	
Log likelihood function value at market share	-42269	
Goodness-of-fit (McFadden)	ρ^2=0.303, $\rho^2_{(adj)}$=0.302	
Log likelihood function value of MNL model	-30030	
2[LL(GNL) - LL(MNL)] > $\chi 2\ (0.01, 4)$	1122.9 > 13.3	

[*]Significant at 10% level; [**]Significant at 5% level, [***]Significant at 1% level.

[a] Numbers in parentheses after variable names in the first column indicate bundled alternatives, described in Table 1, associated with the variable.

4.2. Effect of household socio-demographics

All household socio-demographic variables are significant at 99% level of confidence, implying an important role for socio-demographic characteristics in explaining household vehicle ownership behaviour. More specifically, the effect of annual household income indicates that high income households are more likely to own vehicle, and that the income effect is highest on bundles containing two or more motorcycles (alternatives 5, 7 & 9), followed by bundles containing one motorcycle (alternatives 4, 6 & 8), bundles containing car(s), and those containing bicycle(s) only (alternatives 2 and 3).

The number of adults has a substantial effect on vehicle ownership and its effect also differs across the alternatives. The influencing trend is being driven by the number of vehicles within each bundle. In particular, the number of adults has a greater effect on bundles with more vehicles. Having children less than or equal to 6 years of age also induces households to own vehicles of any type, all else being equal.

House area is also found to be a factor affecting a household's demand for vehicles presumably because large house area facilitates inside-house parking which is a very common practice in Vietnam. It is expected that households with large house area are more inclined to own bundles with larger size vehicles. This expectation is confirmed by the decreasing estimates of the *large home size* indicator for car bundle, bundles with two or more motorcycles (alternatives 5, 7 & 9), and those containing one or less motorcycle. The parameter is not significant for alternative 2. This would appear logical because not much room is required for parking one bicycle. These results extend previous findings by Ho and Yamamoto (2009) and Osara et al. (2009) in which not only tenancy status but also house area significantly influence household vehicle ownership.

4.3. Effect of vehicle operating cost

The only alternative-specific attribute in the analysis is the total operating cost of all vehicles in each bundle. The highly significant estimates of the linear and quadratic terms of this variable indicate that the effects of vehicle operating cost on household propensity for owning vehicle of any type are not linear. *Ceteris paribus*, higher operating cost decreases household vehicle ownership, with a declining marginal effect reflected in the significant positive coefficient on the annual operating cost squared variable.

The operating cost of all vehicles in the chosen bundle was reported by each household while values for unchosen alternatives were imputed using information from the collected samples. The effectiveness of imputing and using vehicle operating cost as a predictor is explored by comparing the two model specifications: one with and one without vehicle operating cost attribute. The model goodness-of-fit in terms of adjusted McFadden pseudo Rho-squared significantly improves from 0.086 to 0.302. There are three main reasons for this remarkable improvement in model fit. Firstly, the household-specific choice set has been defined as a by-product of an operating cost imputing process. Compared to the assumption of a universal choice set (i.e., all bundled alternatives are available for each household) the established household-specific choice set assumption is more realistic. This is reflected in a substantially higher Rho-squared (0.266) of the model with household-specific choice set and without vehicle operating cost. Secondly, operating cost is probably the important variable in explaining household vehicle ownership. Thirdly, it may be that the imputed vehicle operating costs are highly correlated with vehicle bundles' prices – an important determinant of vehicle ownership but not collected in this study. Consequently, a part of vehicle operating cost influence is probably attributed to vehicle price. However, since household income was used as a matching criterion to impute vehicle operating costs of unchosen alternatives, the improvement in the model fit may partly represent effects of income on choice. It is therefore necessary for future research to further disentangle the effect of this imputation method on model performance.

4.4. Effect of the built environment and transport supply conditions

Several built environment variables were found to be significant in alternative bundles' utility functions. The negative coefficients of population density confirm previous research findings that households residing in highly populated neighbourhoods are less likely to own vehicles (see, e.g., Bhat and Sen, 2006). However, the effect of population density at residential zone on multiple motorcycle and car bundles was not significant though its parameters have an expected negative sign.

Table 4 shows land use diversity also mattered. Generally, mixed land use settings at the residential zone tended to work against motorised vehicle ownership. This finding is consistent with research showing mixed use developments induce people to use alternative modes to the private car (e.g., Ewing and Cervero, 2010) thereby relieving the need for owning motorised modes.

On the subjective side of the built environment, perceived factors of transport supply conditions and the motorcycle mode appear to influence household vehicle ownership to some degree. Specifically, the perceived factor characterising negative traffic externalities was highly significant, with negative effects on motorcycle and bicycle ownership.

As for transport supply conditions, several composite factors characterising public transport services and walking and cycling infrastructure were found to significantly work against motorised vehicle ownership and in favour of bicycle alternatives. As expected, perceived bus coverage reduced car ownership and multiple motorcycle bundles (alternatives 5 & 7) though the effect on multiple motorcycle alternatives was not statistically significant. Also, ease of bus use was found to be inversely correlated with the likelihood of vehicle holding. More interestingly, bus operators' attitudes significantly influenced household vehicle ownership, ostensibly reflecting the fact that good bus operators encouraged people to use public transport more thereby reducing the need for owning private vehicles. This is a very important finding. It is known that improving public transport services by means of a widened public transport network and/or increased service frequency is the common practice to increase public transport use. However, an increase in market share for public transport due to such incentive policies normally comes with high costs. Although the magnitude of changes in travel behaviour resulting from better bus operators' attitudes required further exploration, this finding suggests a supplementary/alternative way to reduce private vehicle ownership thereby vehicle use. This finding, however, might result from the cognitive dissonance of bus captive users. They have to use bus even if the bus operators' attitude is not good, so they might adjust their evaluation on operators' attitudes to justify their bus use (see Festinger (1957) for detailed discussion of cognitive dissonance). Thus, more in-depth investigation on this endogeneity issue is required and reserved for future study.

Walking conditions in the CBD and community both influenced the propensity of owning vehicles. In particular, a good walking condition in CBD tended to reduce the likelihood of owning a vehicle of any type while its counterpart in community significantly eased the need of owning motorcycle(s). Interestingly, households residing in neighbourhoods with good cycling conditions indicated a strong inclination toward bundles containing bicycle (alternatives 2, 3, 6-10). Similarly, highly perceived safety of motorcycle mode was correlated with higher propensity of multiple motorcycles owning.

4.5. Effect of self-selection variables

Among the set of variables aiming to capture the effects of self-selection, only some variables were found to be significant in a few alternative bundles. Of the significant variables, preferences towards motorcycle mode appear to be most important ones. Preferences for convenience, freedom and comfort were found to be factors influencing motorcycle ownership (similar findings for the case of car mode can be found in Cao et al., 2007). Also, leisure preference was significant, with a negative effect on the propensity of owning one bicycle. As Table 4 showed, self-selection with respect to negative traffic externalities was not a factor affecting household vehicle ownership behaviour, although the variable charactering this type of self-selection had its estimated parameters with an expected sign.

To further disentangle the possible confounding effects of self-selection from the built environment on household vehicle ownership behaviour, we discarded all attitudinal variables to determine self-selection from the model, re-estimated and compared the results with the original one. It is found that dropping these attitudinal variables from the model, while downgrading the overall model fit, hardly changes the built environment effects as well as socio-economic coefficients. This finding is consistent with Chatman (2009).

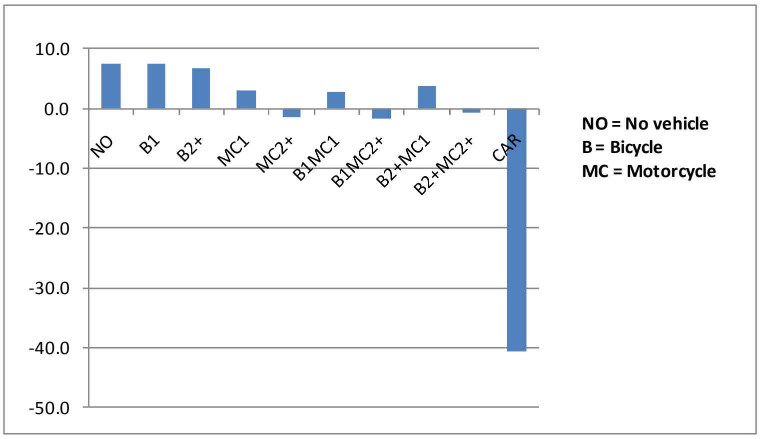
Clearly, given the preliminary nature of the studies focusing on other aspects of self-selection taken to date, these results are suggestive, rather than conclusive. However, it is argued that the findings are somewhat encouraging as tentative indicators of the relevance of other aspects of self-selection to understanding of travel behaviour and external effects of transport.

5. Model application

With the rapid increase in GDP in Vietnam recently, the rate of private motorised use and ownership has also increased at the same fast rate or even more, causing many significant negative consequences to the network and the environment. For the study area, the recent report by HCMC Department of Transport (2007) revealed that GDP of the city annually increased at the rate of 12% while motorcycle and car ownership increased by 10.6% and 14% per year, respectively. Consequently, some transport policies have been considered to reduce private motorised vehicle ownership and use. The model developed in this study can be used to evaluate the change in household vehicle ownership resulting from government intervention. The latest proposal of HCMC People's Committee which proposed an annual holding fee of VND500k (approx. US$ 24 in 2011) per motorcycle and VND10m (approx. US$ 484 in 2011) per car is applied.

The developed model provides two options to evaluate the impact of this intervention: (1) increase vehicle operating cost, and (2) reduce annual household income. As discussed earlier, it is probably that the imputed vehicle operating cost is also a proxy for vehicle price. Consequently, adjusting vehicle operating cost to reflect the policy is unrealistic and undesirable because this works as though the owned vehicles' price also increases annually. Thus, annual household incomes were modified to reflect an additional annual cost of US$ 24 for each motorcycle and US$ 484 for each car. Revised expected aggregate shares were then computed to obtain a percentage change from the baseline estimate. Figure 2 shows the simulated results.

Figure 2 Estimated percentage changes in vehicle ownership in response to annual using fee



With the proposed policy, it is found that private motorised-vehicle ownership, especially car ownership, can be reduced to some extent while bicycle alternatives and bundles containing bicycle(s) and one motorcycle can be increased. On average, car ownership is cut by more than 40% and multiple motorcycle ownership by about 1%. Conversely, bicycle ownership is increased by 7%, and bicycle(s) and one motorcycle bundles ownership by 3%.

6. Conclusions and policy implications

This paper has sought to enhance our understanding of confounding effects of self-selection on the association between the built environment and travel behaviour in terms of multiple vehicle ownership. This study provides new evidence that other aspects of self-selection such as those with respect to negative traffic externalities are also relevant to travel behaviour. The findings partly support the self-selection hypothesis in that failing to account for subjective dimensions of attitudes and preferences leads to misestimating of the influence of the built environment on travel behaviour. However, the measured bias due to self-selection is relatively minor in the environmental settings of the HCMC study area. A possible explanation is that people in developing countries do not base their choices of residential location, destinations and travel modes on their attitudes but mainly on income or other factors. If so, transport policies and planning practices aiming to bring about better built environment (e.g., diversity of land use, good access to local facilities, high quality infrastructure for walking and cycling), and better public transport services have more impact in Vietnam and developing countries than in developed countries like US.

In the empirical study, principal component analysis and GNL model have been employed to examine how spatial factors and public transport services affect vehicle ownership behaviour amongst households in Ho Chi Minh metropolitan area while controlling for the confounding effects of self-selection and socio-demographic characteristics. The study has shown that both subjective and objective measures of the built environment have a strong influence on household vehicle ownership behaviour. In particular, public transport services in terms of bus coverage and ease of use have strong negative influences on the propensity of owning vehicles, especially car and motorcycle. More importantly, bus operators' attitudes were found to be inversely correlated with vehicle ownership, suggesting an alternative or supplementary way to reduce private motorised vehicle ownership and use. While improving public transport services with a widened bus network and increased bus frequency is necessary, especially in areas with currently poor public transport services like HCMC, building up good bus operators' attitudes also helps promote public transport use by reducing motorised vehicle ownership and use.

Other findings related to local and central walking/cycling conditions also provide some encouragement that changes to the built environment that improve walking and cycling conditions may in fact result in a lower level of motorised vehicle ownership and a higher level of bicycle owning. However, translating these results to planning practice and policy implementation requires further research to confirm and expand aspects of the built environment that are most important for easing the need of motorised vehicle ownership and use. Our analysis indicates that comfort is the most important element for bringing about better walking/cycling conditions which in turn reduces vehicle ownership. Safety and convenience also play a comparably important role. Yet, since these qualities often contain several physical factors of the built environment, further research is needed to decompose these aspects into better policy-supportive measures.

The developed model can be used to evaluate the change in private vehicle ownership due to intervening policy such as one proposed by HCMC People's Committee recently. The model application found that the policy of an annual vehicle holding fee can be used as a means of reducing motorised vehicle ownership, especially car, while encouraging more sustainable modes such as bicycle.

Acknowledgements

The authors are grateful to William Greene and David Hensher for their works related to the coding of additional options into NLOGIT version 5.0 procedures to assist the estimates of Generalised Nested Logit models presented in this paper. We also thank Rhonda Daniels, Corinne Mulley and John Rose of ITLS for their helpful comments, suggestions, and language support.

References

Almec Corporation (2004) The Study on Urban Transport Master Plan and Feasibility Study in Ho Chi Minh Metropolitan Area (HOUTRANS), Vol. 1, 2, 3.

Bhat, C. R. & Sen, S. (2006) Household Vehicle Type Holdings and Usage: An Application of the Multiple Discrete-Continuous Extreme Value (MDCEV) Model. *Transportation Research Part B,* 40, 35-53.

Bhat, C. R., Sen, S. & Eluru, N. (2009) The Impact of Demographics, Built Environment Attributes, Vehicle Characteristics, and Gasoline Prices on Household Vehicle Holdings and Use. *Transportation Research Part B,* 43, 1-18.

Boarnet, M. G. & Crane, R. (2001) *Travel by Design: The Influence of Urban Form on Travel.* Oxford; New York, Oxford University Press.

Bodea, T., L. Garrow, et al. (2008) Explaining Obesity with Urban Form: A Cautionary Tale. *Transportation,* 35, 179-199.

Cao, X., Mokhtarian, P. L. & Handy, S. L. (2007) Cross-Sectional and Quasi-Panel Explorations of the Connection between the Built Environment and Auto Ownership. *Environment and Planning A,* 39, 830-847.

Cao, X., Mokhtarian, P. L. & Handy, S. L. (2009) Examining the Impacts of Residential Self-Selection on Travel Behaviour: A Focus on Empirical Findings. *Transport Reviews,* 29, 359 - 395.

Cervero, R. & Kockelman, K. (1997) Travel Demand and the 3Ds: Density, Diversity, and Design. *Transportation Research Part D,* 2, 199-219.

Cervero, R. (2002) Built Environments and Mode Choice: Toward a Normative Framework. *Transportation Research Part D,* 7, 265-284.

Chatman, D. G. (2009) Residential Choice, the Built Environment, and Nonwork Travel: Evidence Using New Data and Methods. *Environment and Planning A,* 41, 1072-1089.

Choo, S. & Mokhtarian, P. L. (2004) What Type of Vehicle Do People Drive? The Role of Attitude and Lifestyle in Influencing Vehicle Type Choice. *Transportation Research Part A,* 38, 201-222.

Ewing, R. & Cervero, R. (2010) Travel and the Built Environment -- a Meta-Analysis. *Journal of the American Planning Association,* 76, 265-294.

Festinger, L. (1957) *A Theory of Cognitive Dissonance.* Stanford University Press.

Handy, S. L., Cao, X. & Mokhtarian, P. L. (2006) Self-Selection in the Relationship between the Built Environment and Walking: Empirical Evidence from Northern California. *Journal of the American Planning Association,* 72, 55 - 74.

Hayashi, Y., Kato, H. & Teodoro, R. V. R. (2001) A Model System for the Assessment of the Effects of Car and Fuel Green Taxes on CO_2 Emission. *Transportation Research Part D,* 6, 123-139.

HCMC Department of Transport (2007) Plan for Mitigation of Traffic Congestion and Traffic Accidents in HCMC from 2007 to 2010. *A Proposal Submitted to Ho Chi Minh City People Committee in May 2007* (in Vietnamese).

Ho, C. Q. & Yamamoto, T. (2009) Comparative Analysis on Determinant Factors of Household Vehicle Ownership between Hochiminh and Hanoi Metropolitans, Vietnam. *The 8th International Conference of Eastern Asia Society for Transportation Studies*, Surabaya, Indonesia, 132-147.

Koh, W. T. H. (2003) Control of Vehicle Ownership and Market Competition: Theory and Singapore's Experience with the Vehicle Quota System. *Transportation Research Part A,* 37, 749-770.

Nijland, H. A., Hartemink, S., van Kamp, I. & van Wee, B. (2007) The Influence of Sensitivity for Road Traffic Noise on Residential Location: Does It Trigger a Process of Spatial Selection? *The Journal of the Acoustical Society of America,* 122, 1595-1601.

Osara, T., Nakamura, F., Okamura, T. & Wang, R. (2009) A Study on Ownership and Usage of Car and Motorcycle in Hanoi City, Vietnam. *Proceedings of Infrastructure Planning,* 40, CD-ROM (in Japanese).

Redmond, L. S. & Mokhtarian, P. L. (2001) The Positive Utility of the Commute: Modeling Ideal Commute Time and Relative Desired Commute Amount. *Transportation,* 28, 179-205.

Train, K. (1980) A Structured Logit Model of Auto Ownership and Mode Choice. *The Review of Economic Studies,* 47, 357-370.

van Wee, B. (2009) Self-Selection: A Key to a Better Understanding of Location Choices, Travel Behaviour and Transport Externalities? *Transport Reviews,* 29, 279-292.

Wen, C.-H. & Koppelman, F. S. (2001) The Generalized Nested Logit Model. *Transportation Research Part B,* 35, 627-641.

Zhang, J., Kuwano, M., Lee, B. & Fujiwara, A. (2009) Modeling Household Discrete Choice Behavior Incorporating Heterogeneous Group Decision-Making Mechanisms. *Transportation Research Part B,* 43, 230-250.