Article:

Flavia, A and Choudhury, C orcid.org/0000-0002-8886-8976 (2019) Temporal Transferability of Vehicle Ownership Models in the Developing World: a Case Study of Dhaka, Bangladesh. Transportation Research Record, 2673 (3). pp. 722-732. ISSN 0361-1981

https://doi.org/10.1177/0361198119836760

TEMPORAL TRANSFERABILITY OF VEHICLE OWNERSHIP MODELS IN THE DEVELOPING WORLD: A CASE STUDY OF DHAKA, BANGLADESH

Anyiko Flavia
Institute for Transport Studies, University of Leeds
34-40 University Road, Leeds LS2 9JT, UK
Anyflavie@gmail.com

Charisma Choudhury
Institute for Transport Studies & Choice Modelling Centre, University of Leeds
34-40 University Road, Leeds LS2 9JT, UK
C.F.Choudhury@leeds.ac.uk

Word count: 5,593 words text + 4 tables x 250 words (each) = 6,593 words

Initial Submission Date: 1 August 2018
Re-submission Date: 15 November 2018

ABSTRACT

Temporal transferability of model parameters is a critical issue, especially in the context of developing countries where data and resources for transport model development are extremely limited. This study investigates the temporal transferability of vehicle ownership models with special emphasis on exploring the effect of model structure on temporal transferability. The performance of potential updating methods for making the models more transferable are also compared. The household survey data collected from Dhaka, Bangladesh in 2005 (STP 2006) and 2010 (DHUTS 2011) have been used in this regard. Different forms of Random Utility Based Discrete Choice and Count Regression Models of car, motorcycle and bicycle ownership have been developed using income, household size, and number of workers, children and licensed drivers as explanatory variables. The temporal transferability of each model between the two time-periods has been compared rigorously using statistical tests. Results indicate that Multinomial Logit model has better temporal transferability compared to the Count Regression Models. In terms of model updating, the Combined Transfer Estimation method for model updating is found to perform better than the Bayesian updating. The findings can provide useful guidance during application of a pre-existing model in the context of a developing country.

Keywords: Global south; Car ownership; Bangladesh; Count Regression; Bayesian updating; Combined Transfer Estimation

1. INTRODUCTION

The economic growth of a developing country is very often highly inter-linked with the growth of its transport sector. In recent years, economic growth has facilitated rapid urbanisation in most countries of the developing world followed by exponential increase in ownership and use of private vehicles (1). This, in turn, has increased the demand for transport infrastructure and services and in many cases intensified the negative transport externalities such as air pollution, high energy consumption and loss of lives from accidents (2). These highlight the extreme importance of robust travel demand models that can be used for informing and guiding policy decisions directed at sustainable planning, control and management of transport services and infrastructure. Development of travel demand models however often requires significant resources which are not readily available in developing countries given the financial constraints. Transferability of the models across time, either in the original form or with limited updating, offers an economic solution to this problem.

Several studies have investigated temporal transferability of travel behavior but mostly using data from developed countries (3-8). However, the developing countries typically have very different transport contexts and travel behavior. For example, in 2013, the per capita motorized vehicle ownership was about 0.52 in the UK and 0.058 in China (9). Further, the transport and economic landscape are changing at much faster rates in the developing countries. China, for example, is expected to have a 13–17% per year increase in car ownership till 2020 (10) whereas recent data indicates that many of the European countries have almost reached their 'peak car' levels (11). This warrants the need for detailed research in the context of developing countries on vehicle ownership models, their temporal transferability and measures to improve temporal transferability.

The particular research questions we investigate are as follows:

- Which model structure best explains the vehicle-ownership decision in the context of a developing country where the car-ownership level is very low compared to the total population?
- How do the performance of different model structures compare in terms of temporal transferability?
- Which candidate method has the best performance in improving the temporal transferability?

Disaggregate data collected from Dhaka, the capital of Bangladesh and one of the fastest growing megacities in the world, has been used to investigate these research questions. Dhaka already hosts more than 18 million people and attracts 300,000 to 400,000 new migrants every year from different parts of the country (12). To meet the mobility demands of the rapidly growing population, the number of vehicles is increasing at an alarming rate. According to Bangladesh Road Transport Authority (BRTA), the number of newly registered vehicles in Dhaka in 2004 was 21,471 and 95,743 in 2015 (13). The rapidly changing demography and transport scenario of the city (further detailed in Section 4), makes it an interesting test-bed for conducting the research on temporal transferability. Further, despite the very high growth rate, the vehicle ownership levels in Dhaka are one of the lowest in the world with more than 90% households not owning any vehicle (car, motorcycle or cycle). From a modelling perspective, this (i.e. excessive occurrence of zeros as dependent variables) poses additional challenges and prompts us to investigate the most appropriate model structure for predicting vehicle ownership in the context of very low ownership levels.

In this research, vehicle ownership models are estimated using household survey data from two different time periods 2005 (14) and 2010 (15). The models are estimated as disaggregate at the household level taking into consideration the communal nature of making travel decisions (16). The effect of the excessive zeros in the vehicle ownership data (which is typical in the developing world), different model structures have been estimated and compared in terms of goodness-of-fit and temporal transferability. The potential methods to improve temporal transferability have also been tested. The rest of this paper is organized in the following sequence; data, methodology, results and conclusions.

2. DATA

The data used by the study is obtained from two household surveys conducted in the Dhaka Metropolitan Area, Bangladesh in 2005 and 2010. Area definitions in both years were the same though the 2010 sample is much bigger (18084 households) than the 2005 one (655 households). The surveys, originally conducted for developing the Strategic Transport Plan for Dhaka (STP 2006 (14)) and the Dhaka Urban Transport Strategy (DHUTS 2011(15)) respectively used the same questionnaire and identical stratified sampling strategies. The socio demographic data collected in the survey included household income, total number of persons per household and the household composition (e.g. number of children, workers, students and licensed drivers). TABLE 1 below presents a comparison of the household level demographics in the two datasets and more detailed statistical analyses are presented in Figures 1 and 2.

TABLE 1: Comparison of the Two Datasets

		2005	**2010**
Average Income (BDT/month)		22732.82	31785.46
Average Income (USD/month)		270.29	457.18
Income Distributions (%)	Low	39.50	26.20
	Medium	52.40	58.90
	High	8.10	14.90
Average Household Size		4.23	4.00
Average Number of Workers		1.37	1.38
Average Number of Licensed Drivers		0.16	1.10
Average Number of Cars per Household		0.06	0.07
Average Number of Motorcycles per Household		0.04	0.03
Average Number of Bicycles per Household		0.01	0.01
Total Vehicle Ownership per Houshold (%)	0	89.31%	90.30%
	1	9.78%	8.97%
	1+	0.91%	0.74%

*The GDP per capita in Dhaka changed from 485.21 USD to 762.81 USD during this time period (9), the exchange rates were 1 USD = 84.10 BDT (2005) and 69.53 (2010) respectively

As seen in Table I, there are some similarities between the demographics in the two datasets as well as differences. For instance, the average numbers of licensed drivers are significantly higher in the 2010 sample and the average incomes are higher as well. The differences are not unexpected though given the increase in GDP in this time period.

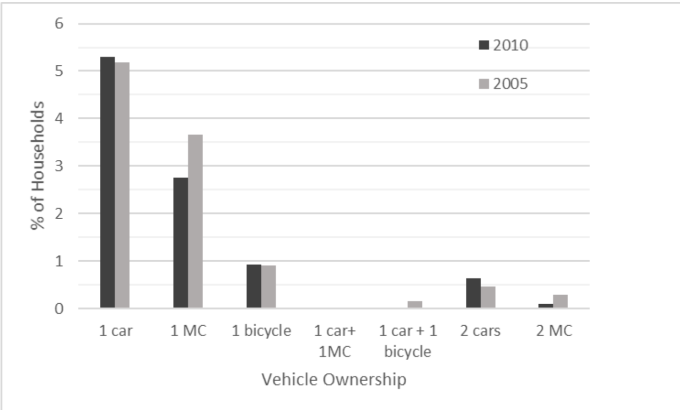
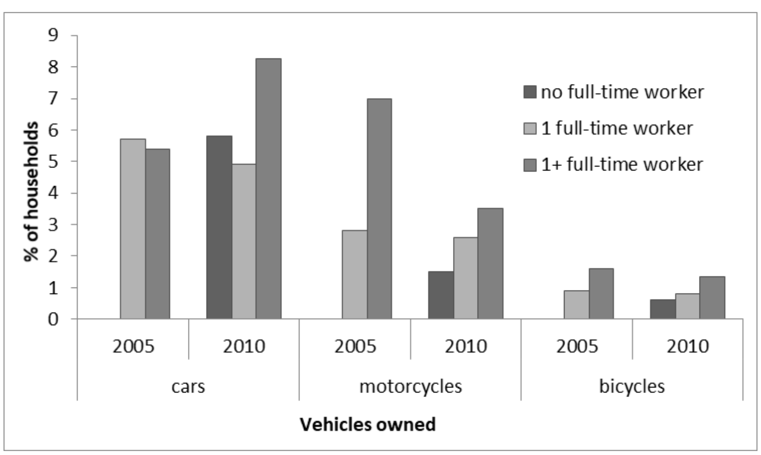


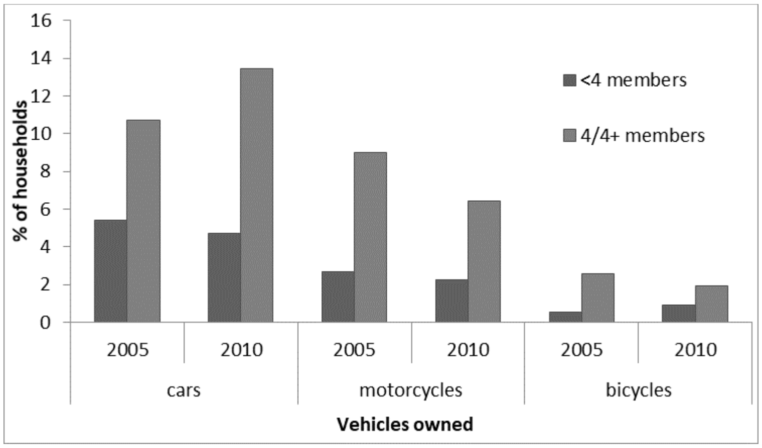
FIGURE 1 Vehicle ownership distribution.

Fig 1 shows the detailed comparison of the type of vehicle ownership in the two datasets. As seen in the Figure, the total percentage of households owning at least one car is very low (around 6%), out of which less than 1% own two or more cars. Bicycle ownership is also very low among the surveyed households with no household owning more than one bicycle. It may be noted that though Dhaka is a flat city, bicycles are not popular because of safety (there are no designated cycle lanes) and security issues (they can easily get stolen in absence of proper bicycle racks). Also, it is considered culturally improper for women to ride bicycles.

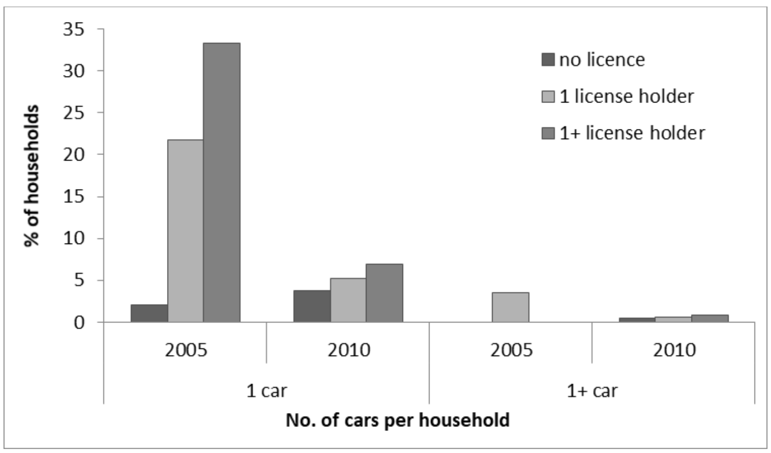
In terms of vehicle ownership among different demographic groups, in most cases, car and motorcycle ownership is highest among households with two or more full-time working members (with the exception of car ownership in case of 2005) (Figure 2a). This is expected as an increase in workers is reflective of an increase in the household income. However, though vehicle ownership is low for households with no full-time working members, in the 2010 dataset, some of such households do report that they own vehicles. It is suspected that these household may have one/more members working part-time or have members who have retired from their jobs. Vehicle ownership rates are higher for larger households (Figure 2b). This is expected as mobility needs are expected to increase as the number of people in a household increases. The relationship between the number of license holders and the number of cars in a household (Figure 2c) demonstrates a weak correlation with car ownership. This is not unexpected given that most of the cars in Dhaka are chauffeur driven and it is common to own a car without having a driving license or to have a driving license but not own a car (i.e. work as a chauffeur by profession which is a low income profession).



a) Vehicle ownership and number of workers



b) Vehicle ownership and household size



c) Vehicle ownership and licensed drivers

FIGURE 2 Demographic distribution and vehicle ownership.

3. METHODOLOGY

The candidate model structures, methods for testing transferability and updating the model parameters (using limited data) are discussed in this section. For all cases, the state-of-the-art is presented first followed by details of the selected methods.

3.1 Model Structures

Vehicle ownership is typically modelled using Ordered and Unordered Discrete Choice Models or Count Regression Models.

Among the unordered models, Multinomial Logit (MNL) and Nested Logit (NL) models have been widely used for their ease of analysis and availability of estimation software, both in medium and long term (see 17 for further details). In these models, following the principles of utility maximisation, the decision maker chooses the alternative that provides the greatest satisfaction. Therefore, for a given set of alternatives, the probability of household n choosing alternative I, given choice set C_n , can be expressed as follows:

$$P_n(i) = P\left(V_{ni} + \epsilon_{ni} \geq V_{nj} + \epsilon_{nj}\right) \forall C_n, j \neq i \quad (1)$$

Where, V_{ni} and ε_{ni} represent the observed and random components of utility of alternative i. The distribution of the unobserved error term ε_i indicates if the model is MNL or NL.

The MNL structure assumes that the error term is independently and identically (Gumbel) distributed across households. The probability of household n selecting vehicle ownership alternative i, is therefore expressed as follows:

$$P_n(i) = \frac{e^{V_{ni}}}{\sum_{j \in C_n} e^{V_{nj}}} \tag{2}$$

Ordered models are based on the assumption that a latent intangible variable represents a household's propensity to own vehicles and the probabilities of owning certain number of vehicles are then obtained by matching specific ranges of the values of the latent variable to the corresponding numbers:

$$y_n^* = \beta' X_n + \varepsilon_n$$

$$y_n = \begin{cases} 0 & if\; y_n^* \leq \mu_0 \\ 1 & if\; \mu_0 < y_n^* \leq \mu_1 \\ 2+ & if\; \mu_1 < y_n^* \end{cases} \tag{3}$$

Where; y_n^* is the car trip generation propensity for household n; y_n is the car trip generation for household n; X_n is the vector of explanatory variables; ε_n is the random error term (normally distributed for Ordered Probit and logistically distributed for Ordered Logit); μ_0 and μ_1 are the threshold parameters; and β is the vector of model coefficients.

Among the Count Regression Models, the Poisson and the Negative Binomial regression models are some of the most commonly used to estimate and analyse count data, though their applications have been primarily in the context of crash frequency (see 19 for a comprehensive review) , trip-generation (e.g. 20 - 23 etc.), etc. with a few applications for vehicle ownership decisions (e.g. 24, 25).

In Poisson regression, it is assumed that the number of occurrences (k) of the dependent variable y has a Poisson distribution given the independent variables X_1, X_2, …, X_n:

$$P(y_n = k | X_1, X_2, \ldots, X_n) = \frac{e^{-\mu}\mu^k}{k!}, k = 0,1,2, \ldots \tag{4}$$

It assumes $\ln(\mu)$ is a linear function of independent variables, denoted as:

$$\ln(\mu) = \ln(N) + \alpha_c + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{5}$$

Where, y_n is the number of vehicles owned by household n and X_1, X_2,…, X_n are household characteristics in the context of vehicle ownership, N is the sample size, α_c is a constant, and $\beta_1 \ldots \beta_n$ are coefficients of the household characteristics. The Pearson's goodness of fit test is used to check model appropriateness to the data distribution. P-values less than 0.05 mean the data is significantly different from a Poisson distribution.

While the Poisson model assumes the mean and variance are equal, the Negative Binomial model takes into account the possibility of over-dispersion in the data due to large differences between the observed mean and variance (see 26 for details). The Negative Binomial regression model takes the following form:

$$P(y_n) = \left(\frac{1}{1+\alpha\mu_n}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha\mu_n}{1+\alpha\mu_n}\right)^{y_n} \quad (6)$$

μ_n is the mean and α is the over dispersion parameter which is greater than zero.

To investigate model appropriateness (i.e. to test if Negative Binomial model, which has one additional parameter, is superior to the Poisson model), a Likelihood Ratio Test is performed and the likelihood ration (LR) is compared with the chi square distribution:

$$-2\big(LL(P) - LL(NB)\big) \sim \chi_k^2 \quad (7)$$

Where, LL(P) and LL(NB) are the log-likelihoods of the Poisson model and the Negative Binomial model respectively and use of the Negative Binomial model is justified if LR is greater than the critical chi square value at k degree of confidence .

A variant of these models are Zero Inflated Poisson and Zero Inflated Negative Binomial models which addresses the issues associated with excessive zeroes in the data and are expressed as follows respectively:

$$P(y_n) = \frac{e^{-\mu_n}\mu_n^{y_n}}{y_n!} \quad (8)$$

$$P(y_n) = \frac{\Gamma(\alpha y_n+1)}{y_n!\Gamma 1} \left(\frac{1}{1+\alpha\mu_n}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha\mu_n}{1+\alpha\mu_n}\right)^{y_n} \quad (9)$$

The Vuong test (27) is carried out to compare the models to the simpler variants.

3.2 Transferability

Temporal transferability of the individual parameters is checked by testing whether or not there is a significant difference between the parameter estimates of equivalent variables in the two cities (28). Minimum and maximum t-ratio values of -1.96 and 1.96 corresponding to the 95% confidence interval are taken as the critical values.

$$t_{diff,k} = \frac{\hat{\beta}_{trans,k} - \hat{\beta}_{appl,k}}{\sqrt{\left(\frac{\hat{\beta}_{trans,k}}{t_{trans,k}}\right)^2 + \left(\frac{\hat{\beta}_{appl,k}}{t_{appl,k}}\right)^2}} \quad (10)$$

Where; $\beta_{trans,k}$ and $\beta_{appl,k}$ are the estimates for the k-th parameter in the transferred and application areas; $t_{trans,k}$ and $t_{appl,k}$ are the respective t-ratios of the parameter estimates; and $t_{diff,k}$ is the t-ratio for the difference between parameters.

Global measures of model transferability are also obtained using the transferability test statistic (TTS) (28, 29).

$$TTS_{appl}(\hat{\beta}_{trans}) = -2 \times \left(LL_{appl}(\hat{\beta}_{trans}) - LL_{appl}(\hat{\beta}_{appl})\right) \quad (11)$$

Where; $LL_{appl}(\hat{\beta}_{trans})$ is the log-likelihood on the application context data with transferred context parameters; $LL_{appl}(\hat{\beta}_{appl})$ is the log-likelihood on the application context data with application context parameters; and $TTS_{appl}(\hat{\beta}_{trans})$ is the transferability test statistic of the transferred model in application context.

The transferability test statistic follows a chi-squared distribution with degrees of freedom equal to the number of parameters estimated and its value should be less than the critical chi-square value at the chosen level of significance for good transferability.

3.3. Model Updating

Findings from previous studies indicate that temporal transferability of a model is improved by updating the model parameters with some information from the application context (e.g. 30). Several updating approaches have been suggested in literature. Among these, the two most widely used are explained below:

Bayesian updating

The Bayesian updating process follows the Bayes theorem in which prior information about the model is combined with a random sample from the application context to get updated information that is important in reducing doubt during prediction (33). The parameters estimated with the data from the first location can be used as the prior information in this case and the following formula can be used (Equation 12):

$$\beta_{updated} = \left(\frac{\beta_{trans}}{\sigma^2_{trans}} + \frac{\beta_{appl}}{\sigma^2_{appl}}\right)\left(\frac{1}{\sigma^2_{trans}} + \frac{1}{\sigma^2_{appl}}\right)^{-1} \tag{12}$$

Where β_{trans} and β_{appl} are the vectors of parameters of the originally estimated model and the application context model respectively and σ_{trans} and σ_{appl} are corresponding vectors of standard deviations.

Combined Transfer Estimation

The combined transfer estimation method (34) acknowledges the variations between parameters due to long time gaps and other differences between the estimation and application contexts such that the updated parameters are estimated as (Equation 14):

$$\beta_{updated} = \left(\frac{\beta_{trans}}{\sigma^2_{trans}+\alpha\alpha^{-1}} + \frac{\beta_{appl}}{\sigma^2_{appl}}\right)\left(\frac{1}{\sigma^2_{trans}+\alpha\alpha^{-1}} + \frac{1}{\sigma^2_{appl}}\right)^{-1} \tag{13}$$

Where: $\alpha = \beta_{\text{trans}}-\beta_{\text{appl}}$ and $\alpha' = \beta_{\text{appl}}-\beta_{\text{trans}}$

It may be noted that though there are simpler methods like updating only the constants of the model or using scaling of the model parameters, based on the exploratory analysis results, they were not deemed to be appropriate in these cases and have not been tested rigorously.

4. RESULTS

4.1 Model Coefficients

Review of precious studies on vehicle ownership (car ownership in particular) have shown that the private ownership decisions are affected by both the socio-demographics and the urban forms. For example, it has been reported household's decision when purchasing the first car is primarily based on socio-economic factors (income, age of household members, value of time, etc.), while the decision for purchasing a second car (or more) is largely based on traffic network, efficiency, and transit level-of-service parameters (35,36). In the context of developing countries it has been also reported that good transit services decrease the tendency of households to own more motorcycles (37).
However, in this study, between the two waves, there were no significant changes in the transport and urban landscape. For instance, there were no significant improvements or investments in the public transport. Nor were any new major roads constructed within the city. Rather, the economic landscape has undergone major improvements. This motivated us to focus on the socio-demographic variables.The model parameters of all models are estimated using maximum likelihood technique. The model parameters are retained based on their statistical significance. However, in case of some variables, the coefficients are found to be statistically significant in any of the models is retained for consistency.
The results of the MNL model are presented in Table 2. The estimated results indicate that the vehicle

ownership decisions are significantly affected by the income and the number of licence holders in the household. These are intuitive and in agreement with literature. For example, previous studies in developing countries also indicate that household income is one of the major determinants of car and motorcycle ownership (39). This indirectly implies that vehicle ownership will continue to increase as fast as per capita income growth in developing countries until saturation is reached (9**Error! Reference source not found.**). Among other household characteristics, number of workers is found to positively impacts vehicle ownership, but the parameter is statistically not significantly different from zero in the 2005 data. In the 2010 data, it is however significantly different from zero at 95% level of confidence. The increase in number of workers per household has been reported to be positively correlated to vehicle ownership in previous studies as well. For example, in Chennai city in India, ownership of two wheeled vehicles increased with the increase of young workers by 25% and indirectly encouraged car ownership as well since it translated to increased income (38). Increase in household size, however, is found to have a negative impact on vehicle ownership. This agrees with the suggestion of Zegras and Gakenheimer (39) that a decrease in household size could encourage vehicle ownership. This is because smaller households are likely to have fewer dependants (and therefore less expenses and more savings) to facilitate vehicle ownership. It may be noted that the coefficients of household size and number of workers have been found to be statistically different from zero only for 2010. The effect of number of children has not been found to be significantly different from zero in either year and not included in the final model.

TABLE 2: MNL Model Estimation Results

	2005		**2010**		
	Coefficient	**Robust t-stat**	**Coefficient**	**Robust t-stat**	**t-test diff**
Alternative Specific Constant - 1 car	-4.91	-8.94	-4.25	-4.06	-1.17
Alternative Specific Constant 1+ car	-7.30	-9.73	-6.37	-3.11	-1.22
Alternative Specific Constant - Motorcycle	-4.10	-7.86	-4.54	-6.57	0.82
Alternative Specific Constant - Bicycle	-5.42	-8.19	-5.67	-3.81	0.37
High income	2.47	4.41	2.48	3.15	-0.01
Middle income	1.03	3.21	1.03	2.74	0.00
Household size	0.06	-0.17	-0.08	-2.46	0.43
No. of workers	0.00	-0.02	0.17	4.64	-1.11
1 driving license holder dummy	2.65	6.86	0.27	3.14	**6.01**
1+ driving license holder dummy	2.87	3.98	0.09	6.57	**3.85**
Adjusted Rho square	0.751		0.762		
Final log likelihood, LL	-251.921		-6925.139		
LL (applied)					-9281.99
TTS					**4713.70**

It may be noted that market-segmentation tests have also been performed, but the coefficients were not found different for different segments of income, household size and number of workers. Estimation of the Poisson model (separate for number of cars and number of motorcycles/bicycles) with 2005 data in Table III shows similar trends as the MNL model in terms of effects of increase in income, but interestingly, the effect was found to be statistically insignificant in case of motorcycles/bicycles in the 2005 dataset. The number of driving license holders was significant for the car-ownership model for both years. The workers and household size were found to be statistically significant in the 2010 dataset only (both for car and motorcycles/bicycles). Interestingly, the Pearson's goodness of fit test showed that the 2005 data is not significantly different from a Poisson distribution while 2010 data is different from the

Poisson distribution.

The Negative Binomial model results were very similar to the Poisson model, both in terms of magnitude of the coefficients and statistical significance and similar to the Poisson model. The goodness-of-fit measures indicated slight improvement over the Poisson model though the very small values of alpha ruled out over-dispersion for the car-ownership model for both years and motorcycle/bicycle ownership model in 2005. However, both the likelihood ratio test result and the alpha estimate indicated that for the 2010 motorcycle/bicycle ownership data, there is significant over-dispersion. The model is hence retained for the transferability analysis.

The results of the Zero Inflated Binomial Models (ZINB) were substantially different from the Poisson and the Negative Binomial models. The Vuong test for appropriateness of the ZINB model indicates that the model is indeed better suited for 2010 data and 2005 car ownership than the Negative Binomial model. This is evident in the statistically significant z-values at 95% level of confidence as shown by results in Table 3. It may be noted that the coefficients in this case predict the occurrence of zeros and have opposite interpretation of the signs to the previous two models. Though for the sake of conformity, all variables significant in other models were retained, other than income, no variables were found to be statistically significant.

TABLE 3: Poisson and Negative Binomial Regression Model Estimation Results

	Poisson regression model					Negative binomial regression model					Zero Inflated negative binomial model				
	2005		2010		Comparison	2005		2010		Comparison	2005		2010		Comparison
	Coefficient	z-value	Coefficient	z-value	z-test diff	Coefficient	z-value	Coefficient	z-value	z-test diff	Coefficient	z-value	Coefficient	z-value	z-test diff
CAR															
Constant	-4.99	-7.31	-4.75	-4.94	0.21	-4.99	-6.26	-4.85	-4.81	0.12	0.68	-0.92	-1.41	-1.01	1.32
High income	3.29	5.16	3.48	3.69	0.17	3.29	4.95	3.48	3.98	0.17	-4.07	-2.89	-8.56	-5.35	**2.11**
Middle income	1.70	2.87	1.68	2.79	0.02	1.70	2.73	1.68	4.43	0.03	-1.03	-2.12	-3.92	-1.94	1.39
Household size	0.02	0.18	-0.16	-4.05	1.36	0.02	0.14	-0.16	-4.24	1.08	-0.04	-0.24	-0.05	-1.61	0.06
No. of workers	-0.37	-1.34	0.23	5.56	**2.16**	-0.37	-1.41	0.23	5.61	**2.27**	-0.10	-0.43	0.12	0.01	0.02
1 driving license holder dummy	2.34	6.47	0.25	3.06	**5.63**	2.34	6.21	0.25	3.09	**5.42**	-0.04	-0.40	-0.43	-0.61	0.54
1+ driving license holder dummy	2.35	4.37	0.62	6.77	**3.16**	2.35	3.47	0.62	6.80	**2.52**	-1.67	-1.89	-0.19	-1.79	1.66
Inflate constant											1.35	3.50	1.50	18.25	0.38
Final log likelihood, LL	-101.10		-3595.00			-101.10		-3595.00			-128.55		-3614.09		
Pearson's value	1.00		0.89												
Alpha						0.0031		0.0045							
Vuong test z value											4.47		7.88		
LL (applied)					-7185.48					-6505.41					-3692.54
TTS					**7180.96**					**5820.82**					**156.91**
MOTORCYCLE & BICYCLE															
Constant	-4.18	-7.05	-4.23	-5.82	0.05	-4.17	-7.02	-4.22	-2.08	0.02	-3.54	**-3.79**	-3.38	-20.37	0.16
High income	-0.58	-0.57	0.89	8.52	1.45	-0.57	-0.54	0.89	8.49	1.37	1.58	0.82	-12.37	-0.02	0.02
Middle income	0.34	0.93	0.60	6.24	0.69	0.34	0.88	0.60	6.28	0.66	0.84	0.72	1.50	4.51	0.55
Household size	0.20	1.46	0.10	2.00	0.73	0.20	1.37	0.09	2.14	0.70	0.11	0.76	0.10	2.70	0.07
No. of workers	0.31	1.72	0.10	2.00	1.13	0.31	1.52	0.10	1.89	1.01	0.45	**2.27**	0.10	2.05	1.69
Inflate constant											-0.79	-0.37	-1.23	-3.13	0.20
Final log likelihood, LL	-135.20		-3008.00			-135.20		-3001.76			-136.11		-3003.31		
Pearson's value	0.40		0.10												
Alpha						0.0966		1.026							
Vuong test z value											1.07		4.30		
LL (applied)					-3142.82					-3139.99					-3375.17
TTS					**269.64**					**276.46**					**743.74**

4.2 Assessing temporal transferability

Going by the results of the t-statistic for the difference in parameters test shown in Table 2, most of the coefficients of the MNL model are found to be transferable between 2005 and 2010. The only exception is the number of licensed driver dummy variables. However, despite most variables proving transferable, the model itself is not transferable between 2005 and 2010 as indicated by the transferability test statistic 4713.70 which is much greater than the critical chi-square value ($\chi^2_{0.05,11}$=19.675). This is in line the findings of the exploratory analyses where the relationship between car-ownership and possession of driving license were found to have very different patterns - with 2005 having the higher proportion of households without licensed drivers yet at the same time having 45% higher car ownership among license holders compared to 2010.

The findings of the Poisson model for car ownership were somewhat similar to the MNL with the coefficient of the number of licensed driver dummy being significantly different in the two years. In addition, the coefficient of the number of workers was also found to have statistically significant differences originating from the t-stat being significant in 2010 and insignificant in 2005. The model as a whole was however not temporally transferable as the TTS values 7180.97 is much greater than the critical chi-square value ($\chi^2_{0.05,7}$=14.076). The same phenomenon was observed with the Negative Binomial regression model as well.

For the Poisson and Negative Binomial distribution for motorcycles/bicycles, the difference between the parameters were all found to be statistically insignificant, but the TTS values , 269.6 and 276.46 respectively are much greater than the critical chi-square value ($\chi^2_{0.05,5}$=11.07) .

For the ZINB model, all parameters of the car-ownership model except high income dummy are found to be transferable across time as all values of the z-statistic are below 1.96. For the motorcycle/bicycle ownership model, the high income dummy was however not fund to be transferable. The models are not transferable as a whole as well as the TTS values are significantly larger than the critical chi square value ($\chi^2_{0.05,8}$=15.507 and $\chi^2_{0.05,6}$= 12.592).

4.3 Improving temporal transferability

As mentioned in Section 3.3, two methods of updating have been tested: Bayesian Updating and Combined Transfer Estimation.

For the Bayesian method, three small samples of 3616 households were drawn randomly from the application data, i.e., the 2010 data. The sample size was one fifth of the entire 2010 sample a size recommended by Koppelman, Kuah and Wilmot (38) in Santoso and Tsunokawa (39) as suitable for updating procedures. Three random samples were used to eliminate any bias and check for consistency in the data. The models were run using each sample and the resulting parameter estimates used to calculate updated estimates by equation. The updated model for each sample was then tested for transferability. Using another set of random samples measuring one third of the entire 2010 sample, the updating procedure was repeated to check for the effect of bigger sample size on the resulting model transferability.

Similarly, the models were again examined for improved temporal transferability following updating of parameters by the combined transfer estimation method. Modified parameter estimates were calculated using Equation (13).

Improvements that were sought for were reductions in the TTS values.

The results are summarized in Table 4. As observed in the table, it is evident that model updating improves temporal transferability as the TTS values in the updated models are much less. It may be note though even after updating none of the models resulted TTS values smaller than the critical chi square values. The Combined Transfer Estimation approach shows better performance in improving the TTS in all model forms. This is expected because the transfer bias in combining parameters is taken into consideration in this method.

TABLE 4: Comparison of Model Performance

Models and Transfer tests	Bayesian updating				Combined transfer estimation	Critical Chi-Square
	Sample 1	Sample 2	Sample 3	Average		
MNL	175.22	108.24	139.37	140.94	105.84	19.68
Poisson - Car ownership	138.30	145.10	144.10	142.50	129.90	14.08
Poisson - Motorcycles and bicycles	112.40	166.70	164.20	147.77	138.40	11.07
Negative Binomial - Car ownership	154.10	153.10	156.90	154.70	130.70	14.08
Negative Binomial - Motorcycles and bicycles	148.00	159.00	129.50	145.50	134.70	11.07
ZINB - Car ownership	108.50	96.80	153.94	119.75	113.02	15.51
ZINB - Motorcycles and bicycles	108.70	237.80	184.20	176.90	174.10	12.59

Among different model forms, the MNL model shows better improvement than all other models.

5. CONCLUSIONS

Different forms of RUM and Count Regression based models have been rigorously tested in this paper in the context of vehicle-ownership in Dhaka. The aim was to contrast the values of the coefficients across different structures, investigate which models are more temporally transferable and assess the performances of candidate parameter updating methods.

The key findings are listed below:

- The model form results in some, but not substantial, differences in sensitivities towards different influencing factors. For example, in almost all models, the income levels are found to be statistically significant.
- In terms of transferability, most coefficients are individually transferable, but the models are not transferable as a whole as the TTS values were above the critical chi-square values.
- Updating methods result in reductions in TTS values, but they are still above the critical chi-square values
- Of the model structures explored by this study, the MNL model was found to be more temporally transferable, both before and after updating.
- Among Bayesian Updating and Combined Transfer Estimation, the latter results larger improvement in increasing temporal transferability.

It may be noted that 2005 and 2010 are not too far apart and there have not been any significant changes in the urban or transport landscape in this period. Our results are therefore more on the conservative side. But the results serve as a proof of concept that updating of estimated models for temporal transferability is indeed a practical way for developing countries to make better travel demand forecasts without the encumbrance of extensive new data collection and model estimation. The findings are expected to be of utmost practical importance to transport planners working in developing countries where very often it is not possible to collect detailed data on a frequent basis due to resource constraints.

For future research, we recommend further investigation into the performance of other updating methods on temporal transferability and testing temporal transferability of more advanced model structures such (as the mixed logit). Future studies on temporal transferability in the context of developing countries could also examine the use of the more flexible predictive tests such as model elasticity to check the sensitivity of the model to variations in input variables and the relative error measure to compare parameter values between the estimation and transfer context (6) to complement statistical tests of transferability as used by this study. This is relevant since models found statistically not transferable may still prove useful in forecasting with a reliable degree of practical accuracy.

Contributions

Flavia Anyiko conducted the main data analyses and model development. Dr Charisma Choudhury proposed the idea and guided the analyses and model development. Both authors contributed to writing the paper.

Acknowledgement

Flavia Anyiko is funded by the UK Commonwealth Scholarship Commission.

References

1. Sanko, N., Dissanayake, D., Kurauchi, S., Maesoba, H., Yamamoto, T. and Morikawa, T. 2014. Household car and motorcycle ownership in Bangkok and Kuala Lumpur in comparison with Nagoya. Transportmetrica A: Transport Science. **10**(3), pp.187-213.
2. Dargay, J., Gately, D. and Sommer, M. 2007. Vehicle Ownership and Income Growth, Worldwide: 1960-2030. The Energy Journal. **28**(4), pp.143-170.
3. Karasmaa, N. and Pursula, M. 1997. Empirical Studies of Transferability of Helsinki Metropolitan Area Travel Forecasting Models. Transportation Research Record: Journal of the Transportation Research Board. **1607**(-1), pp.38-44.
4. Hadayeghi, A., Shalaby, A.S., Persaud, B.N. and Cheung, C. 2006. Temporal transferability and updating of zonal level accident prediction models. Accident Analysis & Prevention. **38**(3), pp.579-589.
5. Forsey, D., Nurul Habib, K., Miller, E.J. and Shalaby, A., 2014. Temporal transferability of work trip mode choice models in an expanding suburban area: the case of York Region, Ontario. Transportmetrica A: Transport Science, 10(6), pp.469-482.
6. Fox, J. 2015. Temporal transferability of mode-destination choice models. Doctor of Philosophy thesis, University of Leeds.
7. Han, Y. and Zegras, C., 2016. Exploring Model and Behavior Uncertainty: Temporal Transferability Assessment of Vehicle Ownership Models for Boston, Massachusetts, Metropolitan Area. Transportation Research Record: Journal of the Transportation Research Board, (2563), pp.122-133.
8. Sanko, N. 2016. Temporal transferability: trade-off between data newness and the number of observations for forecasting travel demand. Transportation, pp.1-18.
9. The World Bank. 2014. [Online database], https://data.worldbank.org/indicator, accessed 2017-02-09.
10. Wang, Y., Teter, J. and Sperling, D., 2011. China's soaring vehicle population: Even greater than forecasted?. Energy Policy, 39(6), pp.3296-3306.
11. Van Wee, B., 2015. Peak car: The first signs of a shift towards ICT-based activities replacing travel? A discussion paper. Transport Policy, 42, pp.1-3.
12. The World Bank. 2018. Towards Great Dhaka: A New Urban Development Paradigm Eastward. Directions in Development;. Washington, DC: World Bank.
13. Siddique, M.A.B. and Choudhury, C.F., 2017. Modelling the Behavioural Response to Congestion Pricing in Dhaka, Bangladesh. Transportation in Developing Economies, 3(2), p.23.
14. STP. 2006. Strategic Transport Plan for Dhaka, Prepared by Louis Berger Group and Bangladesh Consultant Ltd.
15. DHUTS. 2011. Dhaka Urban Transport Network Development Study, Draft Final Report. Prepared by Katahiraand Engineers International, Oriental Consultants Co. Ltd., and Mitsubishi Research Institute, Inc.
16. Rajagopalan, B. and Srinivasan, K., 2008. Integrating household-level mode choice and modal expenditure decisions in a developing country: Multiple discrete-continuous extreme value model. Transportation Research Record: Journal of the Transportation Research Board, (2076), pp.41-51.
17. de Jong, G., Fox, J., Daly, A., Pieters, M. and Smit, R., 2004. Comparison of car ownership models. Transport Reviews, 24(4), pp.379-408.
18. Train, K.E. 2009. Discrete choice methods with simulation. Cambridge university press.

19. Lord, D. and Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice, 44(5), pp.291-305.
20. Jang, T. Y. (2005). Count Data Models for Trip Generation. Journal of Transportation Engineering, 444-450.
21. Schmöcker, J.D., Quddus, M., Noland, R. and Bell, M., 2005. Estimating trip generation of elderly and disabled people: analysis of London data. Transportation Research Record: Journal of the Transportation Research Board, (1924), pp.9-18
22. Noland, R.B., Smart, M.J. and Guo, Z., 2016. Bikeshare trip generation in New York city. Transportation Research Part A: Policy and Practice, 94, pp.164-181.
23. Verhoeven, H., Simons, D., Van Dyck, D., Van Cauwenberg, J., Clarys, P., De Bourdeaudhuij, I., de Geus, B., Vandelanotte, C. and Deforche, B., 2016. Psychosocial and environmental correlates of walking, cycling, public transport and passive transport to various destinations in Flemish older adolescents. PloS one, 11(1), p.e0147128.
24. Zhao, Y. and Kockelman, K.M., 2002, January. Household vehicle ownership by vehicle type: application of a multivariate negative binomial model. In 81st Annual Meeting of the Transportation Research Board. Washington, DC.
25. Hsu, T.P., Tsai, C.C. and Lin, Y.J., 2007. Comparative analysis of household car and motorcycle ownership characteristics. Journal of the Eastern Asia Society for Transportation Studies, 7, pp.105-115.
26. Long, J. S., and J. Freese. 2006. Regression Models for Categorical Dependent Variables Using Stata. Rev. ed. College Station, TX: Stata Press.
27. Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica: Journal of the Econometric Society, pp.307-333.
28. Galbraith, R.A. and Hensher, D.A., 1982. Intra-metropolitan transferability of mode choice models. Journal of Transport Economics and Policy, pp.7-29.
29. Atherton, T.J. and Ben-Akiva, M.E. 1976. Transferability and updating of disaggregate travel demand models.
30. Santoso, D.S. and Tsunokawa, K. 2005. Transferability and updating analyssis of mode choice model between two developing countries - predictive performance point of view . Journal of the Eastern Asia Society for Transportation Studies. **6**, pp.173-185.
31. Koppelman, F. S., & Wilmot, C. G. (1982). Transferability analysis of disaggregate choice models. Transportation Research Record, (No. 895), pp. 18-24.
32. Badoe, D. A., & Miller, E. J. (1995). Comparison of alternative methods for updating disaggregate logit mode choice models. Transportation Research Record, (1493).
33. Dey, S. S., & Fricker, J. D. (1994). Bayesian updating of trip generation data: combining national trip generation rates with local data. Transportation, 21(4), pp. 393-403.
34. Ben-Akiva, M. and Bolduc, D. 1987. Approaches to model transferability and updating: the combined transfer estimator.
35. Karlaftis, M., Golias, J., 2002. Automobile ownership, households with automobiles, and urban traffic parameters: are they related?Transportation Research Record: Journal of the Transportation Research Board 1792, 29–35.
36. Guerra, E., 2015. The geography of car ownership in Mexico City: a joint model of households' residential location and car ownership decisions. Journal of Transport Geography, 43, pp.171-180.
37. Wen, C.H., Chiou, Y.C. and Huang, W.L., 2012. A dynamic analysis of motorcycle ownership and usage: A panel data modeling approach. Accident Analysis & Prevention, 49, pp.193-202.
38. Srinivasan, K., Bhargav, P., Ramadurai, G., Muthuram, V. and Srinivasan, S. 2007. Determinants of changes in mobility and travel patterns in developing countries: case study of Chennai, India. Transportation Research Record: Journal of the Transportation Research Board. (2038), pp.42-52.
39. Zegras, C. and Gakenheimer, R., 2006. Driving Forces in Developing Cities' Transportation Systems: Insights from Selected Cases. Massachusetts Institute of Technology: Cambridge.
40. Santoso, S. D. and K. Tsunokawa, 2005. Spatial transferability and updating analysis of mode choice models in developing countries. Transportation planning and technology, 28(5), pp.341-358.
41. Koppleman, F., G. Kuah and C. Wilmot. 1985. Transfer model updating with disaggregate data. Transportation Research Record, (1037).