Joint Model of Private Passenger Vehicle Type Ownership and Fuel Consumption in Metro Manila: Analysis and Application of Discrete-Continuous Model

Monorom RITH[a], Jose Bienvenido M. BIONA[a], Alexis FILLONE[b], Kenji DOI[c], Hiroto INOI[d]

[a] *De La Salle University, Department of Mechanical Engineering, 2401 Taft Avenue, Manila City, Metro Manila*
Email: rith_monorom@dlsu.edu.ph
Email: jose.bienvenido.biona@dlsu.edu.ph
[b]*De La Salle University, Department of Civil Engineering, 2401 Taft Avenue, Manila City, Metro Manila*
Email: alexis.fillone@dlsu.edu.ph
[c]*Osaka University, Department of Civil Engineering, Suita, Osaka Prefecture 565-0871*
Email: doi@civil.eng.osaka-u.ac.jp
[d]*University of Toyama, Department of Urban, Toyama city, Toyama Prefecture 930-8555*
Email: inoi@civil.eng.osaka-u.ac.jp

Abstract: This study is aimed to identify micro-level determinants of the joint model for vehicle type ownership-cum-fuel consumption in Metro Manila. The developed model can be used to predict the mentioned output variables, vehicle kilometers traveled, and CO_2 emissions based on various scenarios considered. Frank copula-based joint Multinomial Logit-Linear Regression model is implemented to build the joint model and analyze the sample data. Vehicles are categorized into two alternatives based on engine size: small car (≤2.0L) and large car (>2.0L). As evident from the empirical findings, an increase in gas price is the potential factors in reducing vehicular fuel consumption, vehicle usage, and CO_2 emissions. The empirical results of this study is highly expected to be informative for policy makers in crafting intervention for road transport energy management, vehicular emissions, and the current traffic congestion.

Keywords: Frank copula, Discrete-continuous model, Vehicle ownership, Fuel consumption, Emissions

1. INTRODUCTION

The final energy consumption in Southeast Asia grows about 70% faster than the world average from 2016 to 2040, and the transport sector accounts for 26% of the total energy demand (WEO, 2017). This sector consumes petroleum-based fuels about 60% of the total oil consumed (WEO, 2017). Transportation sector is considered as the main contributor of economic development, while the urban areas are cited as the engines of growth of a country (Fabian and Gota, 2009). However, a rapid economic growth like in the Philippines has provoked Metro Manila (the national capital region) into severe issues including traffic congestion and accidents, urban air quality degradation, noise pollution, and vibration (NEDA and JICA, 2014). The transport sector is the largest share of energy consumption (10.56 Mtoe or 34% in 2015) and GHGs emissions (26.36 $MtCO_2e$ or 39.37% in 2005) (PSY, 2016; Department of Energy, 2009). Road transport typically accounted for about 80% of the transport sector's energy demand (Regidor et al., 2011). NEDA and JICA (2014) indicates that the private car-choice trips account for 31.7% (for person trips) or 71.3% (for vehicle trips) in Metro Manila in 2012 with the per-annum 3.3% growth over the last 16 years. In order to mitigate traffic congestion, fuel consumption, and GHGs emissions from the wealthier people who prefer private vehicles, we need to discourage them from owning and using large fuel-inefficient vehicles. First, we need to understand determinants of vehicle type choice and energy consumption, and then we have to vary the factors' value to predict the output variables. Such a study has been widely studied in developed countries, but it has been less carried out for the developing world, especially in Southeast Asia.

This current paper is aimed to understand the micro-level determinants of individual vehicle type ownership and fuel consumption in Metro Manila. The developed model was implemented to predict the output variables in response to explanatory variable changes. The empirical findings are highly expected to be informative for policy makers in crafting appropriate solutions to manage energy consumption at urban level and reduce circulation of private vehicles in the urban area. Also, the proposed computation procedure is informative for analysts to apply for other case studies to support policy makers.

The remainder of the paper is organized as follows. After a brief description of the background literature, the paper demonstrates the methodological framework with procedures of data modelling. The next section shows the descriptive statistics of output and explanatory variables. The penultimate part interprets the estimated results and apply the developed model for forecasting the output variables, vehicle kilometers travelled (VKT), and CO_2 emissions affected by changes in explanatory variables. The final part is the conclusions of the study and recommendations for future research.

2. BACKGROUND LITERATURE

In the real world, people might make two choices or more at the same time and the two choices might be inter-dependence (called self-selection effect) such as individual vehicle type holding and vehicle mile of travel (Nguyen et al., 2017), mode choice and commuting trip timing (Habib et al., 2009), commuting mode choice coupled with work start time and duration (Habib, 2012), parking type choice and activity scheduling process (Habib et al., 2012), residential neighborhood choice and daily household vehicle usage (Bhat and Eluru, 2009), and vehicle type holding and use (Spissu et al., 2009). Bhat and Eluru (2009) generalized Lee's approach by adopting the "Copula" approach to couple two univariate marginal distributions that can allow non-linear and asymmetric dependencies. In addition, the copula-based joint model can couple any form of marginal distributions and is reliable in applications of the developed model. The copula-based discrete-continuous choice model therefore provides some merits over the previous joint models including the two-step approach of Hay (1980) and Dubin and McFadden (1984) based on Heckman's (1974, 1976) and the full-information maximum likelihood approach proposed by Lee (1983). The copula-based joint model can allow analysts to be familiar with maximum log-likelihood function and simulate the two choice equations simultaneously.

However, simultaneous maximum simulated log-likelihood estimation (MSLE) of such joint model entails some demerits. Such approach is computationally prohibitive in response with an increase in number of observations, explanatory variables, and alternatives.

Furthermore, the estimated percentage share of the discrete choice component is less accurate than those of the joint model of independence assumption and the continuous output is underestimated (see Nguyen et al. (2017). This is the main cause that the previous studies have applied the developed model to predict the percentage changes of the continuous output variables in place of numerical values (see Nguyen et al., 2007; Spissu et al., 2009).

Consequently, this present paper proposes to use the sequential MSLE approach of the copula-based joint model. First, we estimated the parameters of the discrete choice equation and then we replaced the estimated parameters into the joint model to estimate the remained parameters of the joint model. Frank copula is used to link the two univariate marginal distributions of the two choices in this study, as this copula type provides the best fit of the joint model as compared to the other copula types (Nguyen et al., 2017; Spissu et al., 2009; Bhat and Eluru, 2009). Mathematical framework and procedure are detailed in the next section.

3. MODELLING METHODOLOGY

The discrete choice is modelled using the Multinomial Logit (MNL) model, while the linear regression is employed to model the fuel consumption. The two outcome equations are linked together using Frank copula to become a single model bundle of two dimensions, because there may be unobserved factors affecting the two outcomes at the same time.

3.1 Multinomial Logit (MNL) Model

For single discrete choice analysis, MNL is most widely employed because the formula for probability takes the closed form and is readily interpretable (Train, 2003). In 1974, McFadden developed the MNL model, derived from random utility theory (Train, 2003). Each alternative of a choice set has its own utility function composed of observed term and unobserved term. The observed component is a function of explanatory variables and the unobserved component is assumed to be identically and independently distributed (IID) Type I Extreme-Value (Gumbel). The latent utility equation is typically expressed as linear in parameter form (see equation 1). Where β'_t is a vector of alternative specific parameters of vehicle type "t" chosen by an individual "n"; x_{nt} is a column vector of the explanatory variables (including a constant); and ε_{nt} defines the error term.

$$U_{nt} = V_{nt} + \varepsilon_{nt} = \beta'_t x_{nt} + \varepsilon_{nt} \quad (1)$$

Regarding to the properties of Type I Extreme-Value distribution, the maximum over an IID Extreme-Value random variable is also extreme value distributed and the difference of the two IID Extreme-Value random terms is logistically distributed (Pendyala and Bhat, 2004). Therefore, the indicated cumulative distribution of the random error term of a chosen alternative $F(\varepsilon_{nt})$ can be shown as follows (Train, 2003).

$$F(\varepsilon_{nt}) = \Pr(V_{nt} > V_{nT}) = \Pr(t) = \frac{\exp(\beta'_t x_{nt})}{\exp(\beta'_t x_{nt}) + \sum_{T \neq t} \exp(\beta'_T x_{nT})} \quad (2)$$

3.2 Linear Regression Model

As fuel consumption "l" is positive, we assume "l" has a log-normal distribution or ln(l) is normally distributed. The fuel consumption equation of vehicle type "t" chosen by an individual "n" can be written as a function of explanatory variables as seen in equation 3.

Where α is a column vector of generic parameters of all vehicle types; y_{nt} is a column vector of the observed variables (including a constant); and η_{nt} defines the error term. Let $\ln(L_{nt})$ denotes natural logarithm of the fuel consumption outcome by vehicle type "t" chosen by an individual n. The probability density function $f(\eta_{nt})$ and the cumulative distribution function $F(\eta_{nt})$ then can be expressed as equation 4 and 5, respectively (Johnson et al., 1994). Where σ_t is standard deviation of alternative "t". The respective ϕ and Φ are the probability density function and cumulative distribution function of the standard normal distribution. $F(\eta_{nt})$ is the marginal distribution function for fuel consumed by vehicle type "t".

$$\ln(l_{nt}) = \alpha' y_{nt} + \eta_{nt} \tag{3}$$

$$f(\eta_{nt}) = \Pr\big(\ln(L_{nt}) = \ln(l_{nt})\big) = \frac{1}{\sigma_t}\phi\left(\frac{\ln(l_{nt}) - \alpha' y_{nt}}{\sigma_t}\right) \tag{4}$$

$$F(\eta_{nt}) = \Pr\big(\ln(L_{nt}) \leq \ln(l_{nt})\big) = \Phi\left(\frac{\ln(l_{nt}) - \alpha' y_{nt}}{\sigma_t}\right) \tag{5}$$

3.3 Joint Model and Computational Procedure

Supposed that the vehicle type choice and fuel consumption are inter-dependent (correlated), the joint probability of the two choices are demonstrated as equation 6. As a consequence, there is a probability of vehicle type choice conditional on fuel consumption.

$$\Pr\big(m = t, \ln(L_{nt}) = \ln(l_{nt})\big) = \Pr\big(m = t|\ln(L_{nt}) = \ln(l_{nt})\big)\Pr\big(\ln(L_{nt}) = \ln(l_{nt})\big) \tag{6}$$

As mentioned early, Frank copula approach is used to capture the linkage structure of the joint discrete-continuous choices. Based on Sklar's theorem, two univariate distributions are joint by a bivariate distribution function represented by a copula C (Trivedi and Zimmer, 2005; Nelsen, 2006; Yan, 2007). Then, the bivariate distribution of vehicle type choice and fuel consumption can be shown as equation 7. The parameter θ represents the linkage between the two univariate distributions. In the conditional statement, the conditional joint distribution function is derived from a partial derivation of the copula, called a conditional copula which determines the conditional probability (Trivedi and Zimmer, 2005; Nelsen, 2006). The probability of vehicle type t conditional on fuel consumption ln(L)=ln(l) is shown as equation 8. Using the conditional copula, the equation 6 of the joint probability of vehicle type choice and fuel consumption can be derived into equation 9.

$$F(\varepsilon_{nt}, \eta_{nt}) = C_\theta(F(\varepsilon_{nt}), F(\eta_{nt})) \tag{7}$$

$$\Pr\big(m = t|\ln(L_{nt}) = \ln(l_{nt})\big) = \frac{\partial C_\theta(F(\varepsilon_{nt}), F(\eta_{nt}))}{\partial F(\eta_{nt})} \tag{8}$$

$$\Pr\big(m = t, \ln(L_{nt}) = \ln(l_{nt})\big) = \left(\frac{\partial C_\theta(F(\varepsilon_{nt}), F(\eta_{nt}))}{\partial F(\eta_{nt})}\right)\big(f(\eta_{nt})\big) \tag{9}$$

The log-likelihood function of the Frank copula-based joint model can be written as equation 10. R_{nm} $[R_{nm} = 1]$ defines the dummy variable of the chosen vehicle type "m" (m=1,2,3…M) made by an individual n.

$$LL = \sum_{n=1}^{N}\sum_{m=1}^{M} R_{nm}\left[\ln\left(\frac{\partial C_\theta(F(\varepsilon_{nm}), F(\eta_{nm}))}{\partial F(\eta_{nm})}\right) + \ln\left(f(\eta_{nm})\right)\right] \tag{10}$$

Supposed $u_1 = F(\varepsilon_{nm})$ and $u_2 = F(\eta_{nm})$, thus the partial derivative of copula can be written as in Table 1.

Table 1. Partial derivative of Frank copula (Bhat and Eluru, 2009)

Copula	Partial derivate of copula conditional on u_2
Frank	$1 - e^{\theta u_2}\left(e^{\theta u_1} - e^{\theta}\right)\left[e^{\theta u_1}e^{\theta u_2} + e^{\theta}\left(1 - e^{\theta u_1} - e^{\theta u_2}\right)\right]^{-1}$
Range of θ	$-\infty \leq \theta \leq \infty$
Kendall's tau	$1 - \frac{4}{\theta}\left[1 - \frac{1}{\theta}\int_{t=0}^{\theta}\frac{t}{e^{t}-1}dt\right]$

The code language written in R programming was employed to estimate the maximum log-likelihood function by maximizing the above log-likelihood function via applying Newton Raphson type optimization routine. The "maxLik" package developed by Henningsen and Toomet (2011) was applied to simulate the log-likelihood function, while the package "spcopula" developed by Graeler (2014) was implemented to calculate the partial derivative of Frank copula. Firstly, maximum likelihood function of the MNL model was estimated to ascertain the estimated parameters of the discrete choice component, and then the estimated parameters were replaced in the joint model (equation 10) to compute parameters of the continuous choice equation and dependency parameters between unobserved variables of the discrete and continuous choice equations.

4. SAMPLE DATA

2,300 households in Metro Manila were randomly selected based on simple random sampling technique to participate face-to-face interviews which was conducted in April and May, 2017. The questionnaires comprised three main sections: household information (residential location, socio-economic and demographic), household vehicles' characteristics (make, model, vintage year, year of purchase, fuel type, purchase cost, mode of purchase, new car or used car when purchased, expenditure on fuel per week, impact of gas price fluctuation on travel demand, which household member holds vehicle type for regular trip), and daily travel behavior of household members (destination, days of travel per week, number of commuters in vehicle). Such information from the questionnaire is useful to merge the secondary data with, in order to achieve the study objectives. Any questionnaire forms not having complete answers were removed and only the household members who regularly use four-wheelers for their daily trips are encoded. Vehicles purchased before year 2007 and vehicle for rents or used as public transport mode were not deliberately included for data analysis to avoid data consistency problem. After cleaning the data, there were only 1039 available observations. Vehicles are classified into two categories based on engine size (see Table 2). In average, the small car consumed fuel 87.23liters/month, which is less than the fuel consumed by the large car (119.51 liters/month). Summary of basic features of the explanatory variables for empirical analysis are tabulated in Table 3. Majority of observations own a house; are married, male and aged above 40 years; served as an employee; and earned bachelor degree or higher.

The monthly family income includes primary income and other receipts (pensions, grants, etc) by all family members. The household income was categorized as dummy variable in the questionnaire forms to encourage responses and simplify the respondent's task. The dummy variable can be transformed into the continuous variable as a function of household demographic variables using Order Probit model. Such an approach performs much better in imputation of a continuous variable of income from a dummy variable, compared with the Naïve or the Midpoint approaches (Bhat, 1994). Furthermore, the Order Probit model can clearly recognizes ordinality of the dependent variables, avoid arbitrary assumption, allow for analysis of the statistical framework, and possibly test the assumption based on probability distribution (Winship and Mare, 1984). The eighth and ninth rows of Table 3 show the descriptive statistics of household income dummy variable and continuous household income variable converted from the dummy variables. The fuel cost/estimated household income ratio and the vehicle purchase cost/estimated household income ratio were employed to investigate the impact of household income, fuel price, and vehicle purchase price on vehicle type choice and fuel consumption and predict the output variables under changes in the mentioned input variables.

Table 2. Classification of vehicle types and the corresponding per-vehicle fuel consumption

Vehicle type	Small car	Large car
Engine size	≤2.0V Liter	> 2.0 Liter
Frequency (% share)	708 (65.74)	369 (34.26)
Per-vehicle fuel consumption (liters/month)	88.79	135.45

The procedure of the Order Probit model was detailed by Winship and Mare (1984) and Bhat (1994). Equation (11) and (12) were applied to calculate the estimated household income. In our study, supposed the household income has a log-normal distribution with disturbance in term of mean zero and variance ($\omega' w_n$), which can vary regarding to household characteristics of individual "n". The threshold value $a_{n,j}$ is known as the pre-specified intervals j=$\{1, 2, 3, \ldots, 12\}$, w_n denotes a column vector of explanatory variables (including a constant), γ and ω are column vectors of the corresponding coefficients for mean of log-income and standard deviation of log-income, respectively. I_n refers the estimated household income of individual "n". Maximum simulated likelihood estimation approach was applied to estimate the results. The respective ϕ and Φ are the probability density function and cumulative distribution function of the standard normal distribution. Table 4 demonstrates the estimated parameters of household income as a function of household demographic variables. The monthly household income was found higher among families with large family size, more working adults, and presence of people working overseas but lower among the households with pre-schoolers and children go to school. Evident from the last six rows of the table, presence of family member working oversea was found non-significant on the variance of disturbance, while the family size variable increases and the other variables reduce the variance of disturbance. Such demographic variables were not included into the joint model as those variables were significantly contributed to the estimated household income.

$$\Pr(\ln(I_n) = j) = \Phi\left(\frac{\ln(a_{n,j}) - \gamma' w_n}{\omega' w_n}\right) - \Phi\left(\frac{\ln(a_{n,j-1}) - \gamma' w_n}{\omega' w_n}\right) \quad (11)$$

$$I_n = \exp\left[\gamma' w_n + (\omega' w_n) \times \frac{\phi\left(\frac{\ln(a_{n,j-1}) - \gamma' w_n}{\omega}\right) - \phi\left(\frac{\ln(a_{n,j}) - \gamma' w_n}{\omega' w_n}\right)}{\Phi\left(\frac{\ln(a_{n,j}) - \gamma' w_i}{\omega' w_n}\right) - \Phi\left(\frac{\ln(a_{n,j-1}) - \gamma' w_n}{\omega' w_n}\right)}\right] \quad (12)$$

Table 3. Descriptive statistics of the explanatory variables

Explanatory variables	Min	Median	Mean	Max
Family size	1	3	3.227	9
Employed individuals	0	2	1.906	6
Number of preschoolers	0	0	0.178	3
Number of children go to primary and high school	0	0	0.475	5
Number of children go to university	0	0	0.241	3
Number of overseas Filipino workers	0	0	0.139	3
Household income level per month (dummy variables) [a]	3	8	8.590	12
Estimated household income in Php per month $[10^4]$	1.336	6.994	9.692	40.730
Population density km^{-2} $[10^4]$ (city level)	0.124	2.801	3.320	7.330
Housing type				
Own house = 1	0	1	0.775	1
Rental house/share-house = 0 (reference)	-	-	-	-
Marital status				
Married = 1	0	1	0.878	2
Single/divorcee = 0 (reference)	-	-	-	-
Sex				
Male = 1	0	1	0.811	1
Female = 0 (reference)	-	-	-	-
Age				
≤40 years = 1	0	0	0.417	1
>40 years = 0 (reference)	-	-	-	-
Occupation				
Employee = 1 (if employee)	0	1	0.758	1
Self-employed = 1 (if self-employed)	0	0	0.207	1
Non-working (reference)	-	-	-	-
Education level				
Bachelor degree or higher = 1	0	1	0.921	1
Lower than bachelor degree = 0 (reference)	-	-	-	-
Number of commuters in vehicle	1	1	1.255	5
Monthly household income/(monthly fuel cost [× 10])	0.333	1.962	2.482	24.373
Annual household income-vehicle cost ratio	0.211	1.335	2.044	15.302

[a] Household income level: 1<Php5,000; 2=Php5,000-9,999; 3=Php10,000-14,999; 4=Php15,000-19,999; 5=Php20,000-29,999; 6=Php30,000-39,999; 7=Php40,000-59,999; 8=Php60,000-79,999; 9=Php80,000-99,999; 10=Php100,000-149,999; 11=Php150,000-299,999; 12≥Php300,000

Table 4. Estimated results of the order probit model for household income

Variables	Coefficients (t-value)
Mean of log (household income)	
Intercept	1.364 (26.296)**
Family size	0.166 (5.253) **
Household employed members	0.158 (5.222) **
Number of pre-schoolers	-0.156 (-3.670) **
Number of children go to primary and high school	-0.208 (-6.028) **
Number of children go to university	-0.131 (-3.605) **
Number of overseas Filipino workers	0.539 (10.197) **
Standard Deviation of log (household income)	
Intercept	0.402 (11.532) **
Family size	0.068 (3.483) **
Household employed members	-0.064 (-3.615) **
Number of pre-schoolers	-0.065 (-2.384) *
Number of children go to primary and high school	-0.049 (-2.357) *
Number of children go to university	-0.118 (-4.656) **
Number of overseas Filipino workers	-

LL at convergence = –1794.69
LL at null hypothesis = –2017.19
* significance at 95% interval; ** significance at 99% interval

5. DATA ANALYSIS

5.1 Model Estimation Results

This section reports a description of model estimation results for Frank copula-based discrete-continuous joint model for vehicle type ownership and fuel consumption. The estimated percentage share of the discrete choice and per-vehicle fuel consumption are demonstrated in Table 5. The estimated results were found very close to the observed data for both discrete and continuous choice components. The sequential MSLE approach is superior to the simultaneous MSLE approach in term of estimated percentage share of discrete choice and estimated mean continuous output variable. The estimated per-vehicle fuel consumption of small car type was slightly higher than the corresponding actual data.

Table 5. Comparisons of estimated output variables with the observed data

Choice equation	Vehicle type ownership		Mean fuel consumption	
	Small car	Large car	Small car	Large car
Actual	65.74 %	34.26 %	88.79 (liters/month)	135.45 (liters/month)
Estimated	65.74 %	34.26 %	92.74 (liters/month)	135.08 (liters/month)

Table 6 tabulates the estimated parameters of the two choice equations. The second column of Table 6 has no data, as the small car type is used as the reference category for discrete choice data analysis. The third column lists the estimated parameters for discrete choice, while the fourth and fifth columns list the estimated parameters of the continuous choice components for small car and large car respectively. The values in the parentheses are the t-values employed to identify the significant status and magnitude of the factors. The non-zero dependence parameters define that there are unobserved factors affecting both vehicle type choice and fuel consumption for each vehicle type at the same time (see the last row of Table 6). The negative sign of the dependency parameters implies positive dependence between discrete choice and continuous choice, and vice versa (Spissu et al., 2009; Habib et al.,2009). This is associated with that unobserved factor increasing propensity to own vehicle type "t" also increase propensity of fuel consumption. The last row of table 6 presents that the values of dependence parameters for small car type and large car type were 4.081 (or Kendall's tau = 0.394) and 1.38 (or Kendall's tau =0.149), respectively. These figures suggest that people who are more likely to own vehicles are declined to consume fuel. Those who are likely to own small car are more declined to consume fuel than the others.

The intercept parameter of the discrete choice component suggests that people in Metro Manila are less likely to acquire large cars relative to small cars in the ten-year period prior to 2017. People living in higher density urban area are less likely to own large cars as compared with those residing in lower density urban area. It is probably due to that higher density area has narrow road space, small parking space, and slow traffic flow. Small cars therefore are more convenient in circulating around high density urban area. Similar result was reported by Bhat et al. (2009) that people living in rural area are more likely to own SUVs and pickups than those living in urban area. Those owning houses have a higher baseline preference for large cars relative to small cars. Young people group (aged $\leq$ 40 years old) are less prone than older people group to own large cars. People serving as self-employed have higher propensity to hold large cars than small cars, as large cars have large seating and luggage space capacity which are more convenient for business trips. Those travelling with commuters are more inclined to own large cars. The ratio of household income to gas price implies that an increase in gas price is associated with a decrease in preference to hold large cars. The ratio of household income to vehicle cost shows that an increase in vehicle cost is consistent on increase in propensity to own large cars, presumably due to availability of used large car with cheap price relative to new small cars. It is interesting to see that the rest variables (marital status, gender, educational level and job as employee) did not show significant impact on vehicle type ownership.

The constant coefficient of larges cars (5.026) higher than that of small cars (4.800) indicates that larges car consumes more fuel than small cars. People living in higher density area are less likely to consume fuel than those who reside in lower density area. The married, male and young aged ($\leq$ 40 years) people group who own small cars have lower likelihood to consume fuel than the other people group. It is surprising that married people consume small car-related fuel less than single people; males are declined to consume small car-related fuel than females; and younger people group are less prone than older people group to consume small car-related fuel. Working adults have propensity to use more fuel than non-working adults. Those who hold bachelor or higher degree are more prone than those holding lower educational degree to consume fuel. Drivers who have regular trips with commuters are likely to consume fuel more than those who have no commuter. The household income-gas price ratio and the household income-vehicle cost ratio indicate that an increase in gas price and vehicle cost can force people to consume less fuel. Finally, the significant scale parameters postulate that there are considerable unobserved factors affecting fuel consumption patterns for all vehicle types. Large car has higher scale parameter than small car, indicating that variability among those who are inclined to acquire large car is higher in fuel consumption than the others.

Table 6. Estimated parameters of the joint model

Variable	MNL		Regression	
	Small car (t-value)	Large car (t-value)	Small car (t-value)	Large car (t-value)
Intercept	-	-1.416 (-5.03)*	4.800 (47.72)*	5.026 (24.56)*
Population density/km^2 [$\times 10^4$]	-	-0.117 (-3.45)*	-0.014 (-2.49)*	-0.026 (-2.30)*
Housing type				
Own house	-	0.344 (2.00)*	0.027 (0.93)	0.098 (1.57)
Rent house/ share-house (reference)				
Marital status				
Married	-	-	-0.176 (-5.27)*	-0.074 (-1.05)
Single/ divorced (reference)				
Sex				
Male	-	-	-0.110 (-3.61)*	0.009 (0.15)
Female (reference)				
Age				
<= 40 years	-	-0.383 (-2.67)*	-0.116 (-4.40)*	0.032 (0.55)
> 40 years (reference)				
Occupation				
Employee	-	-	-0.023 (-0.30)	-0.179 (-1.52)
Self-employed	-	1.068 (6.61)*	0.187 (2.37)*	-0.082 (-0.66)
Non-working adult (reference)				
Education level				
Bachelor degree or higher	-	-	0.113 (2.21)*	0.150 (1.78)
Lower than bachelor degree (reference)				
Number of commuters for regular trip	-	0.457 (3.90)*	0.137 (5.43)*	0.102 (2.81)*
Monthly household income/ fuel cost [$\times 10^1$]	-	0.148 (3.68)*	-0.133 (-15.42)*	-0.097 (-7.89)*
Annually household income/ vehicle purchase cost	-	-0.087 (2.24)*	0.027 (3.79)*	0.020 (1.53)
Scale parameter	-	-	0.346 (31.99)*	0.461 (23.28)*
Dependency parameter	-	-	4.081 (7.78)*	1.378 (1.91)

Small car: reference category
* Significance at 95%

5.2 Validation and Applications of the Developed Model

For validation purpose, 70% of the available observations in the whole sample data was randomly selected to validate the developed model for individual vehicle ownership and fuel consumption. Comparison of the actual data with the estimated results are tabulated in Table 7. The estimated results look very close to the observed data, which implies that the developed model perform well in prediction for both discrete choice (see second and third columns) and continuous choice components (see the last two columns). For the discrete choice part, the developed model slightly under-predicted the small car ownership, but over-predicted the large car ownership.

Table 7. Validation of the developed model

	Vehicle type ownership		Per-vehicle fuel consumption	
	Small car (%)	Large car (%)	Small car (liters/month)	Large car (liters/month)
Actual	67.51	32.49	88.66	138.24
Estimated	66.00	34.00	91.83	136.32

This part demonstrates applications of the developed model for policy analysis as examples. We suppose two main scenarios associated with variables of interest:

- First scenario: we change gas price to +50%, +25%, -25%, and -50% (positive and negative singes mean increase and decrease respectively) as compared to the retail gas price in April 2017 in Metro Manila. The retail pump prices were around 47Php/liter (for gasoline RON97) and 30Php/liter (for diesel), according to Department of Energy (2017).
- Second scenario: it consists of vehicle purchase cost changes in +50%, +25%, -25%, and -50%.

Table 7 shows the impact of changes in gas price and vehicle cost on percentage share and per-vehicle fuel consumption of each vehicle type based on the mentioned scenarios. For the continuous choice component, an increase in gas price is equivalent to reduction of an amount of fuel for the same initial price. For instance, the gasoline price is Php47/liter. If the gasoline price is increased 25%, it means that we can buy gasoline 0.75liters (less than 25%) for the same payment (Php47). We used this concept to calculate fuel consumption in term of changes in gas price. As apparent from Table 7, gas price has great impact on changes in percentage share and reduction of per-vehicle fuel consumption for all the vehicle types. An increase in gas price is consistent on increase in percentage share of small car and decrease in percentage share of large car (see columns 3 and 4). It is surprising to see that an increase in vehicle cost is associated with a slight increase in percentage share of large cars, presumably due to the availability of the used car with cheaper price as compared to the new small car in Metro Manila. Changes of vehicle cost have no significant impact on changes in per-vehicle fuel consumption of small car and any car type, but have great effect on fuel consumption of large cars. Therefore, an increase in gas price is a potential solution to reduce private vehicular energy demand and increase in percentage share of small-size engine cars (engine size $\leq$ 2.0L).

Table 8. Prediction of the output variables based on various scenarios

Scenarios		Vehicle type ownership		Per-vehicle fuel consumption		
		Small car (%)	Large car (%)	Small car (liters/month)	Large car (liters/month)	Any car (liters/month)
Actual		65.74	34.26	88.79	135.45	104.76
Gas price	+50%	68.25	31.75	68.41	116.02	83.53
	+25%	67.25	32.75	82.53	130.67	98.30
	-25%	63.17	36.83	101.80	131.87	112.88
	-50%	58.01	41.99	115.18	119.51	117.00
Vehicle cost	+50%	64.54	35.46	92.54	128.83	105.41
	+25%	65.02	34.98	92.50	131.27	106.07
	-25%	66.90	33.10	92.53	141.67	108.79
	-50%	69.08	30.92	92.99	155.77	112.40

"+" sign means increase, and vice versa

The estimated values of per-vehicle fuel consumption by vehicle type in response to changes in gas price and vehicle cost are possibly used to predict the corresponding per-vehicle vehicle kilometers traveled (VKT) and CO_2 emissions. An exact determination of such output variables is widely known as impossible, consequently it is imperative to make assumption to predict the approximate values. From the sample data, most of small car run on gasoline, while about 50% of large car run on diesel. Consequently, we assume that all the small cars are cars running on the gasoline, while all the large cars are utility vehicles with 50% share of utility vehicles fueled with the diesel. Fuel consumption rates and CO_2 emission factors by vehicle type are tabulated in Table 9. Such critical parameters are used to calculate changes in VKT and CO_2 emissions based on the scenarios.

Table 9. Parameters of fuel consumption rate and CO_2 emissions

Scenarios	Per-vehicle fuel consumption	
	Small car (liters/month)	Large car (liters/month)
Fuel consumption rate (liters/100km)	9[a]	14[b]
CO_2 emission factor (kg/liter)	2.39[c]	2.57[d]

[a] Bose, 1996
[b] Average value of gasoline and diesel fuel consumption rates (Bose, 1996)
[c] Emission factor (CO_2) = Carbone intensity × 44/12 (IPCC, 2006)
[d] Average value of gasoline and diesel CO_2 emission factors

Results in Table 10 demonstrates that gas price and vehicle cost have impact on reduction of vehicle usage and vehicular CO_2 emissions. Evident from the estimated results, it suggests that an increase in gas price is the best solution to mitigate vehicle usage and transport-related CO_2 emissions. Consequently, an increase in gas price is significantly contributed to reduction of traffic congestion, urban air quality degradation, and transport-related energy demand.

Table 10. Predicted vehicle kilometers travelled and CO_2 emissions

Scenarios		Per-vehicle VKT (km/month)			Per-vehicle CO_2 emission (kg/month)		
		Small car	Large car	Any car	Small car	Large car	Any car
Actual		987	967	980	212	348	259
Gas price	+50%	760	828	781	163	298	206
	+25%	917	933	922	197	335	242
	-25%	1131	941	1061	243	338	278
	-50%	1279	854	1101	275	307	289
Vehicle cost	+50%	1028	920	990	221	331	260
	+25%	1028	937	996	221	337	262
	-25%	1028	1012	1023	221	364	268
	-50%	1033	1113	1057	222	400	277

VKT: Vehicle kilometers travelled

Furthermore, the developed model for vehicle type ownership and fuel consumption is possibly utilized to predict the percentage share and fuel consumption in the future. Regarding PSY (2016), the household income in average increased from 356,000Php in 2009 to 379,000Php in 2012 in Metro Manila. These figures are equivalent to household income growth rate of 2.15% per annum. The yearly growth rate of population density in Metro Manila is 1.58% from 2010 to 2015 (PSY, 2016). These growth rates of household income and population density were used to predict the phenomenon in the next ten years (2027). Table 11 demonstrate the predicted results in 2027 compared with the actual data in 2017. There are no noticeable changes in percentage share and fuel consumption from 2017 to 2027, even though there is no any intervention from the governmental policy.

Table 11. Percentage share and fuel consumption by vehicle type in 2027

Scenarios	Vehicle type ownership		Per-vehicle fuel consumption		
	Small car (%)	Large car (%)	Small car (liters/month)	Large car (liters/month)	Any car (liters/month)
Actual in 2017	65.74	34.26	88.79	135.45	104.76
Predicted in 2027	66.00	34.00	91.83	136.32	106.96

6. CONCLUSIONS

This present study identifies the determinants and their magnitude on vehicle type ownership and fuel consumption. It highlights that gas price and population density are significantly contributed to reduction of large car percentage share and fuel consumption. The developed model was also applied for policy analysis under changes in gas price and vehicle cost to estimate vehicle type percentage shares and fuel consumptions. The estimated output variables were used to calculate vehicle kilometers travelled and CO_2 emissions. An increase in gas price is associated with reduction in vehicle usage and CO_2 emissions. Furthermore, the developed mode is possibly applied to predict the vehicle type percentage share and fuel consumption in the future. There is no significant changes in percentage share and per-vehicle fuel consumption from 2017 to 2027, in spite of no governmental policy intervention.

Evident from the empirical findings, building compact city and increasing gas price are the proper solutions in reduction of percentage share of large fuel-inefficient vehicles, private vehicle kilometers travelled, and greenhouse gases emissions. Reduction in circulation of private vehicles is speculated to mitigate traffic congestion and combat global climate change in the present-day situation.

However, an increase in gas price may degrade quality of life, slowdowns economic activity, and affects other sectors (industry, household, agrilculture, etc.). Future research would study the impact of increasing gas price on the other factor using input-output table.

ACKNOWLEDGEMENTS

This study is mainly funded by Newton Project of the University of Oxford, London and Japan International Cooperation Agency (JICA) under AUN/SEED-Net project for Ph.D Sandwich program.

REFERENCES

Bhat, C. R. (1994). Imputing a continuous income variable from grouped and missing income observations. *Economics Letters, 46*(1994), 311–319.

Bhat, C. R., & Eluru, N. (2009). A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research Part B, 43*(2009), 749–765. doi:10.1016/j.trb.2009.02.001

Bose, R.K. (1996). Energy demand and environmental implications in urban transport case of Delhi. Atmos. Environ. 30 (3), 403–412. https://doi.org/10.1016/1352-2310(95)00111-5

Department of Energy. (2009). *Energy and Environment*: https://www.doe.gov.ph/node/1080

Dubin, J. A., & McFadden, D. L. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica, 52*(2), 345–362. http://www.jstor.org/stable/1911493

Department of Energy (2017). *Retail pump prices-Metro Manila*: https://www.doe.gov.ph/retail-pump-prices-metro-manila?title=&page=4

Fabien, H., & Gota, S. (2009). CO2 Emissions from the Land Transport Sector in the Philippines: Estimates and Policy Implications. *Proceedings of the 17th Annual Conference of the Transportation Science Society of the Philippines (2009).* http://tssp.tripod.com/documents/17th_tssp/fabian.pdf

Graeler, B. (2014). *Modelling Skewed Spatial Random Fields through the Spatial Vine Copula. Spatial Statistics.* Elsevier. doi:10.1016/j.spasta.2014.01.001

Habib, K. N. (2012). Modeling commuting mode choice jointly with work start time and work duration. *Transportation Research Part A, 46*(2012), 33–47. doi:10.1016/j.tra.2011.09.012

Habib, K. N., Day, N., & Miller, E. J. (2009). An investigation of commuting trip timing and mode choice in the Greater Toronto Area: Application of a joint discrete-continuous model. *Transportation Research Part A, 43*(2009), 639–653. doi:10.1016/j.tra.2009.05.001

Habib, K. N., Morency, C., & Trépanier, M. (2012). Integrating parking behaviour in activity-based travel demand modelling: Investigation of the relationship between parking type

choice and activity scheduling process. *Transportation Research Part A, 46*(2012), 154–166. doi:10.1016/j.tra.2011.09.014
Hay, J. W. (1980). *Occupational choice and occupational earnings: selectivity bias in a simultaneous logit-OLS model.* Rockville, Md.: National Technical Information Service.
Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica, 42*(4), 679–694. http://www.jstor.org/stable/1913937
Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement, 5*(4), 475–492. http://www.nber.org/chapters/c10491
Henningsen, A., & Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics, 26*(3), 443-458. doi:10.1007/s00180-010-0217-1
Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions.* New York: John Wiley & Sons.
Lee, L.-F. (1983). Notes and comments: generalized econometric models with selectivity. *Econometrica, 51*(2), 507–512. http://www.jstor.org/stable/1912003
NEDA, & JICA. (2014). *Roadmap for transport infrastructure development for Metro Manila and its surrounding areas (Region III and Region IV-A).* Technical Report No. 2. http://www.neda.gov.ph/wp-content/uploads/2015/03/FR-TR1-ENVI-AND-RISK-12149621.pdf
Nelsen, R. B. (2006). *An introduction to copulas.* Springer. http://www.springer.com/978-0-387-28659-4
Nguyen, N. T., Miwa, T., & Morikawa, T. (2017). Vehicle type choice, usage, and CO2 emissions in Ho Chi Minh city: analysis and simulation using a discrete-continuous model. *Asian Transport Studies, 4*(3), 499–517. https://www.jstage.jst.go.jp/article/eastsats/4/3/4_499/_pdf
Pendyala, R. M., & Bhat, C. R. (2004). An exploration of the relationship between timing and duration of maintenance activities. *Transportation, 31*, 429–456.
PSY. (2016). *Philippine statistical yearbook.* Quezon: Philippine Statistics Authority. https://psa.gov.ph/content/2016-philippine-statistical-yearbook
Regidor, J. R., Montalbo, C. M., Napalang, M. G., Palmiano, H. O., Sigua, R. G., Tiglao, N. C., & Vergel, K. N. (2011). *Formulation of a national environmentally sustainable transport strategy for the Philippines.* Quezon: Design Plus. http://ncts.upd.edu.ph/est/
Spissu, E., Pinjari, A. R., Pendyala, R. M., & Bhat, C. R. (2009). A copula-based joint multinomial discrete–continuous model of vehicle type choice and miles of travel. *Transportation, 36*(2009), 403–422. doi:DOI 10.1007/s11116-009-9208-x
Train, K. (2003). *Discrete choice methods with simulation.* New York, NY.: Cambridge University Press.
Trivedi, P. K., & Zimmer, D. M. (2005). *Copula modeling: an introduction for practitioners.* Boston, MA: now Publishers.
WEO. (2017). *Southeast asia energy outlook 2017.* International Energy Agency. www.iea.org/t&c/
Winship, C., & Mare, R. D. (1984). Regression models with ordinal variables. *American Sociological Review, 49*, 512–525.
Yan, J. (2007). Enjoy the Joy of Copulas: With a Package copula. *Journal of Statistical Software, 21*(4). http://www.jstatsoft.org/