

Measuring Urban Motorized Passenger Mobility: Evidence from Sample Selection Poisson Regression Model in the Case of Istanbul

Enver Cenan İnce*, Hüseyin Murat Çelik

Faculty of Architecture, Istanbul Technical University, Turkey

Copyright©2019 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Unlike any previous researches of urban passenger mobility demand on behalf of the travel demand behaviors of the individuals, this study firstly proposes a measurement focusing specifically on the interrelation between the *travel times of the individuals* and the *number of motorized trips* they exhibit in a day. It is the first time in the literature that the related measurement is focusing on the additional number of daily motorized trips -instead of focusing on measuring vehicle miles traveled- as a result of decrease in daily motorized travel time. In this sense, a single equation *Sample Selection Poisson Regression Model (SSPRM)* seems a good candidate to measure the motorized passenger mobility and to integrate it into a trip generation model. The proposed model has shown that a 26 percent decrease in average motorized travel time per capita in Istanbul will make the motorized trip making increase by 6.5%. Non-linear structure of elasticity estimate of SSPRM may further allow us to estimate the spatial variation of generative impact of induced urban passenger mobility within the trip generation models since it is possible to account for the individual characteristics in the estimation of elasticities as long as we have disaggregate spatial data. By the way, these measures of induced passenger motorized mobility might be used in formulating and in assessing any travel demand management policies so as to minimize the motorized trip makings within the urban spaces.

Keywords Travel Demand Models, Urban Motorized Passenger Mobility, Sample Selection Poisson Regression Model

1. Introduction

After many years of continuing investment in highway network to solve traffic congestion in urban areas, transportation professionals came to the idea that the

efforts to solve the traffic congestion by constructing new highways might be a futile effort since these new highways are re-congested in a relatively short period of time confirming the contention that “you cannot build your way out of traffic congestion”. According to the theory, the newly created capacities to solve an existing congestion problem stimulate the suppressed demands and cause shifts in time, route and mode of daily travels as called “triple divergence” by Downs [1]. Emergence of the suppressed demand accompanying with these convergences as a consequence of the constructing a new highway facility is called as “induced urban passenger mobility” as the empirical framework of urban passenger mobility. Herein, by the notion called induced urban passenger mobility, it is explicitly referred to the additional daily number of motorized trips that is caused by some amount of decrease in daily motorized travel time of an individual.

Induced demand as the specific empirical consideration of urban passenger mobility makes also sense from the economic theory perspective as the price of some goods goes down, the demand for it will go up. The improvement or capacity increase in transportation (highway) infrastructure implicitly or indirectly decreases the price of transportation as lower user costs. Eventually, the amount of travel demand will increase up. Induced urban passenger mobility may be composed of several parts. While some part of the induced urban passenger mobility comes from diversion of the existing demand, some other parts are newly generated traffic. Newly generated traffics have two main forms. One is the release of suppressed demand, and the other comes from new urban development around improved facilities. While divergence and suppressed demand effects are short term immediate effects, developmental effects are generally realized in the long term.

Since traffic congestion is realized on highway network, the literature on induced urban passenger mobility is

concentrated on the measurement of the highway distance with respect to travel time. The obtained elasticity is deemed reflecting the people are willing to travel as travel time decreases. These measurements may have several deficiencies: first of all, these measurements may reflect a partial urban equilibrium when especially done in facility based or corridor based measurement. Area-wide measurements, however, may implicate urban development effect and suffer in isolating individuals' elasticity to travel more or farther. Cervero [2] isolated the urban development effect and revealed that an important part of the induced urban passenger mobility comes from urban development.

These studies have proved enough so far that people are willing to make more vehicle miles in their car trips as travel time decreases, resulted in higher urban passenger mobility. This increase can come from either making longer trips for their existing trips by shifting the origin/destination of the trip, or making additional trips not needed or suppressed previously. That part of the research has not made a clear measurement on this issue yet.

As Cervero stated "until trip generation techniques adequately account for latent trips, the traffic assignment step adequately captures route shifts, and dynamic feedback loops are created to account for land use shifts spurred by new road, the art and science of travel demand forecasting will fall far short of the ideal. "Progress is sorely needed on this front"[2]. The intention of this research is an attempt to measure induced urban passenger mobility and to integrate it into trip generation using total travel times spent in daily trips with their personal characteristics to suggest some progress on this front.

The vitality of motorized traffic in our study is based on the two main points. One is that the main opponent of usage of green modes of transportation is able to be declared as the motorized flows within the urban spaces and the other is that the main source of gas emissions in urban spaces is motorized flows of passengers. On the other hand, the vitality of motorized traffic in relation to the improvement of transportation infrastructure in our study is able to be explained by the fact that the improvements of transportation infrastructure in urban spaces are explicitly able to be indicated as the primary factors that are making the amount of motorized trips enormously increase. In this sense, any improvement of transport infrastructure in an urban space induces new motorized travel demand with refers to the notions in literature called diverted traffic and generated traffic as asserted in the paper. The former is directly related to the diversion of the existing motorized travel demand to the routes of newly improved transportation facilities, while the latter refers to two main forms called release of the suppressed demand and newly generated traffic from new urban development around improved facilities. Herein, the release of suppressed motorized travel demand is directly related to the new motorized trips that are newly exhibited

by the passengers in the short run due to the shorter travel times when compared to their previous travel times belonging to the period, which is before the construction of the new transportation infrastructure.

On the other hand, the newly generated traffic -as a result of new urban development along improved transportation facilities- is a kind of long-run motorized travel demand effect. In other words, the suppressed and diverted motorized traffic are taken into considerations as the short-run effects, while the remaining are long-run effects in urban spaces. Therefore, the effect of newly generated traffic due to the new urban development along new transportation facilities in the long-run would be out of the scope of this paper. Such this scope requires highly different datasets and methodological framework, which would be different from the one of this study. Instead, this study is mainly focusing on the diverted and suppressed demand effect within the urban spaces in the short-run, which are empirically able to be investigated via the effects of average daily travel time changes of the passengers as the outcome of newly improved transportation facilities. Accordingly, the effect of one unit change in average daily motorized travel time of a passenger on the number of daily motorized trips of this passenger would be the fresh empirical research, which would be able to highlight explicitly the amount of induced motorized travel demand of individuals as a result of newly developed transportation facilities.

Furthermore, the vitality of measuring motorized traffic in relation to the improvement of transportation infrastructure in urban spaces would mostly come from the concern of minimizing (or managing) motorized traffic so as to enhance the level of social benefits of urban societies via stimulating the usage of green modes of transport, mostly referring to the non-motorized trips such as cycling, walking, etc., and minimizing the level of gas emissions within the urban spaces. In this sense, the paper aims to measure the motorized traffic and then to constitute a baseline to monitor any travel demand management policies regarding the motorized traffic flows within urban spaces. In addition, these considerations have been asserted in the study so as to make policy makers see the urgency of discouraging enlargement of new urban roads for motorized traffic, since each new motorized route induces extra motorized travel demand. All these aims are directly related to understanding the mechanism of motorized traffic of passengers on the urban scale and then to manage it. Herein, measuring induced motorized travel demand via the changes in average daily motorized travel times of the passengers would explicitly be able to feed the empirical baseline of designating and monitoring the performances of travel demand management policies in urban spaces. In another words, any new improvement of transportation infrastructure would be able to make the average daily motorized travel time of individuals decrease, which

would cause dramatic increases in the average number of daily motorized trips of these individuals. In this context, the comparison of the potential changes in motorized travel times as the result of developing new optional transportation projects would be added to the cost & benefit analysis of these related projects so as to select the one with optimum social benefit. Herein, it is referred to the related project's contribution to the urban society's benefit with regards to the inter-relationship between daily motorized travel times and the number of daily motorized trips. It is meant that the more the average daily motorized travel time decreases, the more number of motorized trips would be exhibited by each individual due to the fact that travel time is a kind of disutility for passengers.

On the other hand, it is not meant that the worst project with highest travel times would be selected within the process of selecting any transportation project among its options. Instead, it is here meant that the reciprocal relationship between the number of motorized trips and motorized travel times is to be investigated in that the decrease in travel times would induce new motorized trips, while these new induced motorized trips would make the decreased travel times enormously increase again. This reciprocal relationship would be referring to a kind of recursive optimization between these two parameters with refers to the selection of potentially optimum transport project on the urban scale. From the view of traditional four stages travel demand forecasting models, the number of motorized trips produced & attracted are estimated in the first stage called trip generation stage, while the total system travel time- with refers to the total motorized travel time- is estimated in the fourth stage called network assignment. The findings of this study would be able to make this modelling process more dynamic in that after predicting the total system travel time – with refers to the total motorized travel times- via the fourth stage called network assignment, then the first stage called trip generation is to be run again till the optimization between travel times and number of motorized trips generated is grasped. Such this dynamic process would be able to be implemented within the classical travel demand forecasting models and then the selection of transportation project or bundle of projects would be carried out according to the findings of this iterative process, which is able to be integrated to the travel demand models. Such this consideration would be referring to a kind of milestone for the considerations of travel demand models in literature, which would be referring to another contribution of this paper to the related literature.

Besides, such an optimizing concern would be referring to a min-min problem in that the transport project (or a bundle of transport projects) in urban space is to be selected systematically with refers to the minimum travel times and minimum number of motorized trips in a day as much as possible. To put it in a different way, the urgency of spreading non-motorized green modes of trips within

urban spaces would be able to be justified via the findings of this study in that for motorized trips, it is explicitly seen that each new improvement or development of transportation facility would be able to induce new motorized travel demand resulting in a kind of vicious circle for urban transportation policy makers. This study is just aiming to initiate such these considerations in literature. Obviously, other than Istanbul, more measuring is required and the findings of the study should be improved via numerous urban areas with new studies.

On the lights of all these views, the paper has been structure by seven sections. The next section is devoted to the review of approaches in literature. Description of study area and the data is given in the third section called Study Area, Data & Materials. Section four explains the models. The results are discussed in section five, and section six discusses the results. Lastly, section seven concludes the study.

2. Review of Approaches

A significant amount of effort is devoted to empirical measurement of induced urban passenger mobility in the literature since it may overestimate travel time saving in capacity expansions leading to wrong decision in transport policy. Obviously it would be voluminous to cite all here. Thus only a compact summary of these researches will be included in our study. In relatively recent studies, Cervero [2] and Noland & Lem [3] summarized these researches in very well defined frameworks. Cervero [2] summarizes the literature from the perspective of spatial resolution, measurement, specification, and normative significance.

Spatial resolution in these studies varies from facility specific studies to area-wide studies. A few recent of the facility specific studies are Pells [4], Hansen et al. [5], Kroes et al. [6], Luk and Chung [7], and Mokhtarian et al. [8]. Recent examples of area-wide studies on the other hand, include Hansen and Huang [9], Noland and Cowart [10], Fulton et al. [11], Cervero and Hansen [12], Cervero [13], and Ozuysal and Tanyel [14]. In these studies, elasticities of vehicle miles traveled with respect to lane-mile addition or construction are estimated. In facility-specific studies, generally growth-comparison or matched pair analysis are conducted and estimated elasticities are given as per cent changes. Area-wide studies generally employed econometric models (i.e. ordinary least squares, or two-stage/three-stage least squares to eliminate the simultaneity bias), auto-regressive models, and travel demand models. By adding, a time lag variable of lane-mile addition into the regression equations, intermediate and long term elasticities are also estimated. "At the road segment level, estimated elasticities of Vehicle Miles Traveled (VMT) as a function of lane miles were in the following ranges: 0.15 to 0.30 across a four year horizon, 0.30 to 0.40 across a ten year

horizon and 0.40 to 0.60 across a sixteen-year horizon. At the county level, short-term elasticities were higher, in the range of 0.30 to 0.50, and at the metropolitan scale analysis, short-term elasticities edged even higher: 0.54 to 0.61” [2].

There are several techniques to measure induced urban passenger mobility in the literature as aggregate and disaggregate models. The aggregate models are either facility based or area based models while disaggregate models measure individual and household level elasticities. Dependent variable in these studies is almost always vehicle miles travelled. Independent variables in aggregate models are generally lane-miles of highway facilities in the area with several lagged and geographical variables. The aggregate models are generally time series models and dominate the literature. In disaggregated models however, main independent variable is either travel time or travel speed with other individually characterizing variables. The preferred functional forms in these models are generally log-linear specification to be able to use the estimated coefficient as direct elasticity as a convenience.

The estimated elasticities get higher as the size of the area gets larger. The above mentioned studies are aggregate in measurement scale, and they might include a certain level of errors that come from certain resources. As it was stated earlier, induced urban passenger mobility demand has two main components: generative (i.e. unleashed suppressed demand by a faster moving traffic), and redistributive (diverted trips from other routes, modes, and hours) as a result of capacity improvements. However, even if no capacity improvement takes place in an area, there are other factors that may increase travel demand beyond inducement. These factors include development in the area’s economy, increase in personal incomes, labor force participation, car ownership, and changes in land-use patterns.

Not being able to isolate such effects from induced urban passenger mobility in model specification may lead to erroneous measurement, and this error may aggravate as size of the area gets larger. One possible way to isolate these effects would be disaggregate modeling. Dominating part of the literature is aggregate models with VMT as dependent variable. Unfortunately there are not many studies with disaggregate level measurement. Only Barr [15] specified a set of models at household level with cross-sectional US Nationwide Personal Transportation Survey. The dependent variable in this study is again the logarithm of VMT while the independent variables include socio-economic characteristics of the households as well as the logarithm of travel time. The models are stratified by certain criteria, and one of them is the size of the metropolitan area, and the findings confirm that at disaggregate level, the size of a metropolitan area has no discernible effect on induced urban passenger mobility.

Another possible source of error in measurements could come from simultaneity bias in a sense that the growth in

VMT could be induced by lane-mile additions while lane-mile additions could be aggravated by the growth in VMT. Noland and Cowart [10] used instrumental variables, and Cervero and Hansen [12] used a two-stage least squares simultaneous equation model to eliminate this bias. To sort out simultaneity and causality effect deeper, Cervero [13] launched another four simultaneous equation model to isolate the urban development, lane-miles growth, increases in VMT and travel speed. It is concluded that “Overall, models that have sought to account for two-way causality have yielded lower elasticities (in absolute terms) than those based on single-equation analyses” [2].

Furthermore, the effects of certain other attributes on induced urban passenger mobility were also addressed in the literature. One of them is whether level of congestion creates a variance on induced travel demand (and so on urban passenger mobility). Even though intuition suggests that under heavy congestion the induced urban passenger mobility effect would be higher the intuition is not confirmed by the empirical research [10]. The other issues tested empirically in the literature were that the effects of urban settings of an area, the size of the metropolitan area, and type of the facility on induced urban passenger mobility demand. The findings on these issues are not unanimous [2, 3].

The main motivation above cited literature is to confirm the existence and to measure the magnitude of the induced urban passenger mobility demand at mostly aggregate level. The existence is well confirmed while the magnitude depends on various factors such as the definition of induced urban passenger mobility demand, the model specification, selection of the model and variables, area size and type; and type of the facility. However, none of these studies define a procedure to include induced urban passenger mobility effect in the conventional four-step travel demand models, especially in the trip generation.

In this sense, the present study defines a methodology to include the “generative” part of the induced urban passenger mobility into the travel demand modeling using count data models with individual level disaggregate data. The trip generation is the best defined part of the conventional modeling. It can be specified at either aggregate or disaggregate or both level. The details of these models can be found in any transportation planning textbook [16, 17]. Only Litman [18] explains the qualitative aspects and implications of trip generation and induced travel. Using count data models in trip generation is not a new approach as Jang [19] and Gurmu and Trivedi [20] used count models in predicting disaggregate level of trip generation. However, it is the first time that the present study uses them to account the effect of induced urban passenger mobility in trip generation modeling.

3. Study Area, Data & Materials

The population of Istanbul in 2005 was around 10.500.000, and the administrative borders of the city covers 5.400 km². Istanbul was the capital of the Byzantine and the Ottoman Empires. Even though the capital was moved to Ankara after the Republic of Turkey was established in 1923, the city has sustained her economical supremacy over the country. After the 1950, when the high rate of urbanization was started in the country, Istanbul was the main destination of internal migration. Today, the city carries the 17% of national population while the administrative area of the city is only seven per ten thousand of the country. Furthermore, the city includes approximately 34% of national

manufacturing and 35% of national financial employments, and approximately 44% of foreign trade of the country was done from Istanbul.

Istanbul is located around the Bosphorus; the water strait separating the Asia and Europe. The historical core which is the central business district is circled area in figure 1. Urban fabric starts from the south coasts and expands into the upper north forest areas and valuable agricultural lands. These forest areas include eight potable water dams, host many environmentally sensitive areas, and establish life support systems of the metropolitan area. With this fragile geography, a sustainable population of the city is estimated around to be 16 million while the trends tend to 22 million implying very serious future environmental challenges.

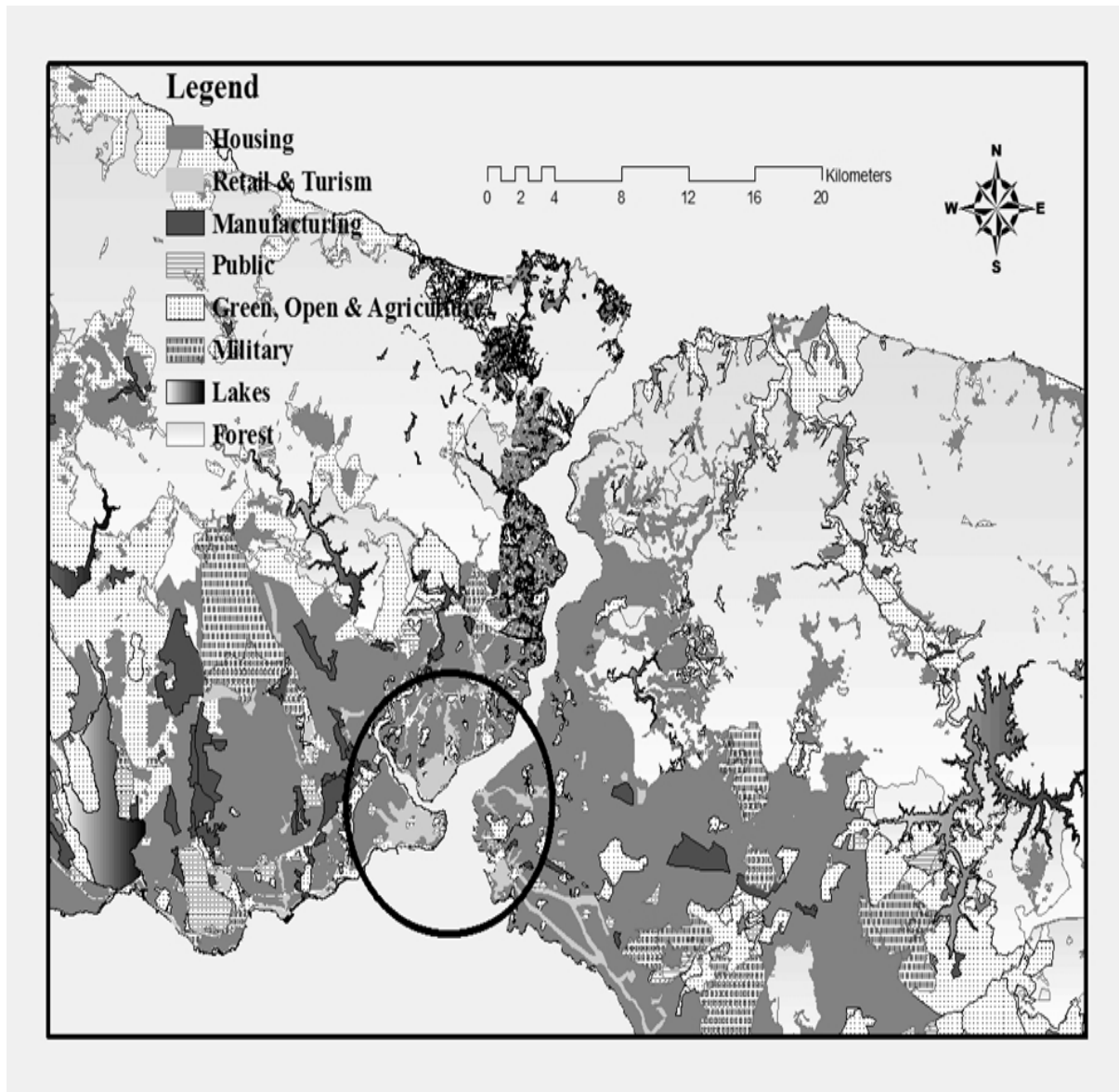


Figure 1. Land Use in the Istanbul Metropolitan Area

The city is connected with two suspension bridges, and the main destination of morning commuting is towards the CBD, and to the European rim since the European rim accommodates 65% of population and 72% of all employment. This unbalanced distribution of population and employment would have been caused by the lack of an extensive rail transit network aggravate traffic congestion in the city especially for the continent crossing. There are 21 million daily trips in the metropolitan area, and half of them are motorized as figure 2 reveals. 6% of all trips make continent crossing. Approximately, 300.000 of them were carried by the ferries while 1.000.000 trips use the bridges.

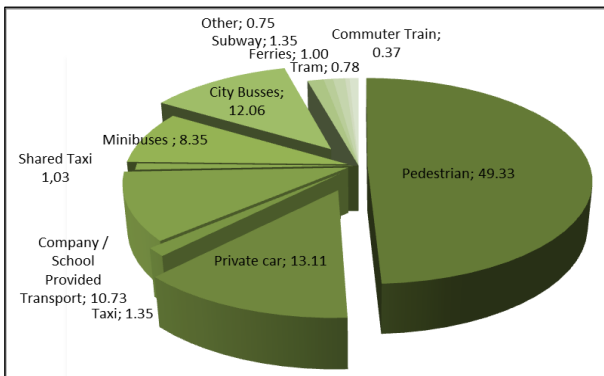


Figure 2. Modal Split of Istanbul

The share of rail transit was only around 3%, and marine transit around 1%. Approximately 15% of the whole trips were made by private car, and the remaining 32% was carried by rubber-tired public transit. Even though the share of private car and the trip rate are lower than the most western countries, the level of congestion in Istanbul is very high. There are two important reasons for such level of congestion: (i) the city is very densely occupied especially in and around the CBD, (ii) the public transit system of the city heavily depends on rubber-tired public transit worsening the traffic congestion and air-pollution, and (iii) feeder arterials of the bridges are also used for intra-rim traffic and when they are congested by the continent crossing traffic, the intra-rim traffic also gets congested in the both sides of the city.

There are some 147 km of rail network with different capacities within the city at the moment. When ongoing rail projects are completed, there will be approximately 270 km of core rail network in Istanbul. Even though all is completed without any delay, the rail network will be lower than those of world's comparable sized metropolitan areas. The late transportation network plan states that the rail network can be extended to 550 km if enough level of resources is raised.

Data used in this study comes from the 2006 Household Origin-Destination (O/D) Survey of Istanbul Metropolitan Area. The survey was conducted by the Department of Transportation of the Metropolitan Municipality of the Istanbul. The sample size in this survey was 90.000

households, and sampling rate was 3%. Samples were chosen by two-stage random cluster sampling. At the first stage, approximately 4000 primary sampling units (PSU) were randomly drawn from the latest building list of the Metropolitan Municipality. In the second stage, 30 households out of the 90 recorded households around the PSU were surveyed systematically as the secondary sampling units in 450 traffic analysis zones. Once the samples were selected, no sample replacement was allowed to avoid non-sampling errors. In this way, 80% unit response rate was achieved. In 72.000 households, a total of 264.000 persons were interviewed face to face, and a total of 356.000 daily trips were recorded between 451 Origin-Destination pairs defining 203.401 distinct movements. The Zone system of Istanbul is given in figure 4.

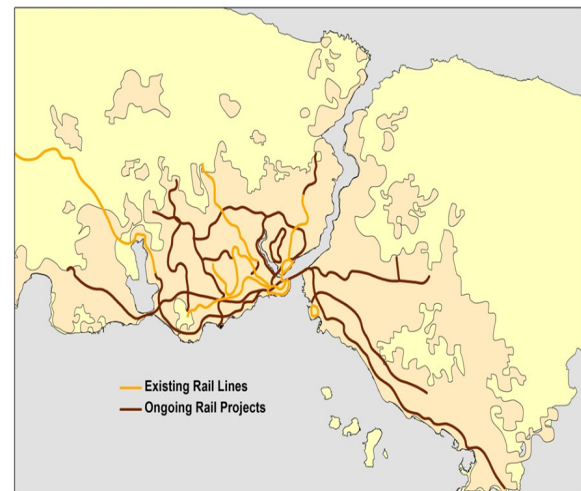


Figure 3. Rail Network of Istanbul



Figure 4. Traffic Analysis Zones of Istanbul Metropolitan Area

All people older than 16 are included in this study and the sample size included 194.000 people. The means and the standard deviations of all variables used in the models are presented in table 1. Total motorized number of daily trips made by an individual (motor_trip) is the dependent variable in the models. The frequency distributions of motorized trips are given in figure 5. Zero trips have the highest frequency in motorized trips. Number of the trips was obtained from the 2006 household survey while trip

distances are the system produced shortest paths between the O/D pairs.

Table 1. Descriptive Statistics of Variables

	mean	st. dv.
motor_trip	0,79	1,12
motor_time	37,85	66,53
male_d	0,5	0,5
hh_head_d	0,36	0,48
hh_income	1,057	1,28
auto_n	0,39	0,58
hh_size	3,92	1,08
age	37,83	15,45
worked	0,34	0,47
student	0,05	0,23
yr_schooling	7,35	3,87

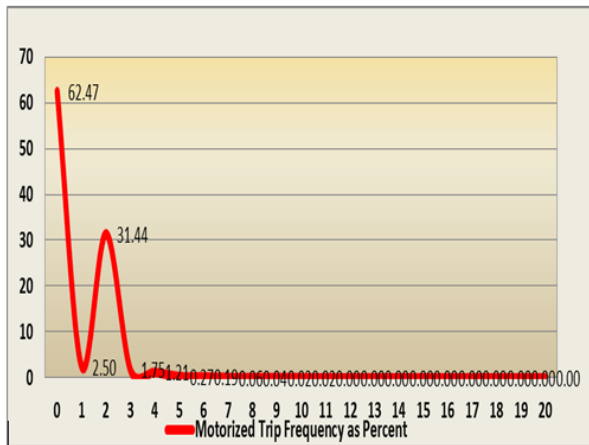


Figure 5. Motorized Trip Frequencies

Set of explanatory variables included the negative of total motorized travel times (motor_time) spent by an individual in daily motorized trips, and personal & family characteristics. Travel time variables are the key variables, and their elasticities with respect to total number of motorized trips are assumed to reflect the willingness to travel more depending on a reduction in daily motorized travel times. The other explanatory variables are dummy variable specifying sex (male_d), dummy variable specifying if the individual is household head (hh_head_d), household disposable income in thousand Turkish liras (hh_income), number of the private cars in the family (auto_n), size of the household (hh_size), age of individual (age), dummy variables if individual is working (worked) or student (student) respectively, and year of schooling the individual attended (yr_schooling).

4. The Models

The count models are specified in this study, which is the main point of the paper. The purpose of these models

is to define a procedure to account the effect of induced urban passenger mobility into trip generation modeling. This specification is intended to provide comparable parametric elasticity estimates for similar studies as another contribution into the literature.

Trip generation is modeled either at aggregate zone level using the linear regression models (LRM) or household data is used to produce cross classification tables which can be considered as a discrete approximation to LRM at disaggregate level [16]. Daily trip making is discrete and non-negative in nature, its occurrence is relatively small, and it may have very small variance. For these reasons, the LRM do not always produce consistent, efficient and unbiased results [21]. The LRM can produce continuous dependent variable ranging from negative infinity to positive infinity, and underlying distribution is generally assumed as gaussian normal. However, frequency distribution of trip generation is hardly gaussian, trip generation is strictly non-negative, and may include excess number of zeros [19].

“In principle, we could analyze these data using multiple regressions. But, the preponderance of zeros and the small values and clearly discrete nature of the dependent variable suggest that we could improve on least squares and the linear model with a specification that accounts for these characteristics” [22]. The family of the models to achieve such improvements is called the count data models, and they are extensively studied and used in the literature. The Poisson Regression Model (PRM) and Negative Binomial Regression Model (NBRM) are the most basic forms of the count models. It is needless to say that there many variations of such models to improve the estimation efficiency [21, 23, 24]. Application of these models takes place in many diversified areas: crime analysis, disease occurrence, doctor visits, occupational injuries, software faults, accident analysis and prevention, manufacturing defects to name the few.

The PRM is the most basic form of the count models. According to the Poisson Regression Model, the random variable, y_i shows a poisson distribution, and mean of this distribution is λ_i .

$$Prob(Y_i = y_i) = \frac{e^{-\lambda_i} \times \lambda_i^{y_i}}{y_i!}, y_i = 0, 1, 2, \dots \quad (1)$$

The mean of the distribution, λ_i , is explained by a set of variables, x_i . the formulation to estimating model parameters is the log-linear model.

$$\ln(\lambda_i) = \beta' \times x_i \quad (2)$$

The basic assumption of this model is the equidispersion, meaning that conditional mean and conditional variance are equal.

$$E(y_i | x_i) = Var(y_i | x_i) = \lambda_i = e^{\beta' \times x_i} \quad (3)$$

The elasticity with respect to any given variable is nonlinear, and can either be estimated at the variable means or as the mean of individual elasticities in the sample.

$$\frac{\partial E(y_i | x_i)}{\partial x_i} = \lambda_i \times \beta = \lambda_i \times e^{\beta' x_i} \quad (4)$$

The PRM is nonlinear and maximum likelihood can be used for parameter estimation as a mathematical simplicity.

$$\ln(L) = \sum_{i=1}^{i=n} [-\lambda_i + (y_i \times \beta' x_i) - \ln(y_i!)] \quad (5)$$

Equidispersion implicitly assumes that “the formula for the probability of an occurrence is a deterministic function of the explanatory variables –it is not allowed to differ between otherwise- identical individuals” [25]. However, this assumption is relaxed by introducing an “unobserved heterogeneity” effect into the conditional mean, “called scale variable”. This leads to a different model, the NBRM in which the conditional variance is larger than conditional mean.

$$\text{Prob}(y_i | x_i) = \frac{r(\theta + y_i)}{r(\theta) \times r(y_i + 1)} \times r_i^y \times (1 - r_i)^\theta, \quad (6)$$

$$\text{where } r_i = \frac{\lambda_i}{(\lambda_i + \theta)}$$

Conditional mean of this distribution is λ_i and conditional variance $\lambda_i (1 + (1/\theta) \lambda_i)$. The elasticities of the NBRM are still estimated as in “(4)” [23].

This fact implies that “it is important to test for overdispersion if you use the PRM. Even with the correct specification of the mean structure, estimates from the PRM when there is overdispersion is inefficient with standard errors that are biased downwards” [21]. Several tests are suggested for overdispersion [22, 23] without estimating a NBRM. Since the PRM and the NBRM are nested, the log-likelihood of the NBRM needs to be improved over the PRM in case overdispersion is present, and this can be checked by a Log-Likelihood Ratio (LR) test:

$$LR = 2 \times [\ln(L_{NBRM}) - \ln(L_{PRM})] \quad (7)$$

LR shows a chi-square distribution, and any value larger than critical threshold with two degrees of freedom favors the NBRM.

Modeling issues explained so far have only dealt with the discrete nature of our dependent variable (i.e. daily trip making frequency). However, our model may have another important specification problem probably causing *endogeneity bias*: the dependent variable and the key independent variable (i.e. total of reported motorized travel times in minutes spent in these travels) may have

causal relationship. The dependent variable is determined by an explanatory variable in a way as the explanatory variable is also determined by the dependent variable in turn. In such situations, since the error term is correlated with the dependent variable(s), the conventional methods produce biased parameter estimates and standard errors leading to invalidation of the core results.

In principle, endogeneity bias is a form of omitted variables bias, and Mokhtarian and Cao [26] summarizes seven different techniques to deal with endogeneity problems: (i) *direct questioning*, (ii) *statistical control*, (iii) *instrumental variables model*, (iv) *sample selection model*, (v) *joint discrete choice model* (vi) *cross-sectional structural equations*, and (vii) *longitudinal models*. Concerning our model, only three of them, namely, instrumental variables; sample selection model; and structural equation models seem meaningful.

Under ideal conditions, we would have chosen a structural equations model by adding a second equation explaining total travel times exclusively and estimated the system simultaneously so that the bias would be eliminated. However, the second equation explaining the total travel time would have needed the specific information about network situation (congestion, geometric and physical quality etc.) for each trip beyond trip making frequencies as independent variables. However, this is not possible to quantify such network information since the models use combined travel data over individuals. Thus, we had to estimate a single equation model. Time standing instrumental variables estimator would be an alternative to eliminate the bias coming from omitted variables, and these omitted variables would be the network information in our case. However, the discrete nature of our dependent variable, and difficulty of obtaining a good set of instrumental variables have made the sample selection model a better alternative to deal with endogeneity problem. Besides, the structure of our data, selection of trip making or not-making, makes Sample Selection Model (SSM) a viable alternative.

SMM is a two-stage model. In the first stage, a selection equation is specified as a binary outcome. That was trip making or not making in our case with respect to socio demographic variables:

$$z_i^* = (w_i' \alpha) + e_i, \quad (8)$$

$$z_i = 0 \text{ if } z_i^* \leq 0 \text{ and } z_i = 1 \text{ if } z_i^* > 0$$

This selection equation is generally estimated as a binary probit model. In the second stage, an outcome equation is estimated. Independent variables included were total of travel times in daily trips in addition to socio demographic variables:

$$y_i^* = x_i' + \mu_i, \quad (9)$$

$$y_i = y_i^* \text{ if } z_i = 1; y_i \text{ not observed if } z_i = 0$$

In this formulation two errors, u and e “are assumed to have correlation ρ ...the two sets of explanatory variables, w and x , need not be disjoint, and, indeed, in some empirical applications, they are identical” [27]. As Breen [27] and Mokhtarian & Cao [26] stated, the correlation of these two error terms is attributed to the omission of explanatory variables. Heckman’s [28] estimator is recommended as an unbiased alternative for parameter estimation in SSM.

Even though Sample Selection Model can eliminate endogeneity bias, the bias that come from a discrete dependent variable still may exist in our models. However, Terza [29] elaborated on count data models with Sample Selection Poisson Regression Model (SSPRM) that may eliminate a significant portion of above mentioned estimator and specification bias. Furthermore, Terza used the derived model to estimate a model’s parameters of the daily trip frequency (the same dependent variable as ours) versus a set of explanatory variables. Just like our model, the model had an endogeneity problem with one of the explanatory variables, namely, number of cars a person owns. The model proved a successful estimator under the given constraints. For formal treatments of all these models and their marginal effects can be found extensively in Maddala [30], Breen [27], Terza [29] and Green [22, 23].

To provide a comparative evaluation opportunity the results of above mentioned models are presented in this paper. In this way, number of motorized daily trips of an individual has been modelled in this study. In this sense, in addition to the Linear Regression Model (LRM) as a benchmarking modelling structure, Poisson Regression

Model (PRM), Negative Binomial Regression Model (NBRM), Sample Selection Model (SSM), and Sample Selection Poisson Regression Model (SSPRM) are estimated. STATA-15 software package was used for the estimation of all models. STATA-15 was also included built-in estimators to calculate the marginal effects of all variables at variables’ means.

In the count models, dependent variable is the total number of motorized daily trips exhibited by an individual within 24 hours. The main explanatory variable is the negative of total motorized travel time spent in minutes for these daily trips since travel time defines a disutility. The remaining explanatory variables are the personal and family characteristics as it is explained in the data section.

5. Results

As it was stated earlier, the count models are estimated for motorized trips in this paper. The results of the count models for motorized daily trips are given in table 2 below. First of all, a linear model is estimated as a benchmark model. Later, a PRM, a NBRM, a SSM and a SSPRM are estimated. As it can be seen from table 2, likelihood ratio as specified in (7), are 50 and 5926. They are well above the critical chi-squared table value of 5,99 indicating overdispersion and favor to the NBRM for both models. Furthermore, the estimated distribution (scale) variable, θ , is significant in both models confirming an overdispersion. Concerning SSM and SSPRM, log likelihood ratios are well above 5,99 indicating all models outperforms their restricted forms.

Table 2. Count Models’ Parameters for Motorized Daily Trips

	LRM	t	PRM	z	NBRM	z	SSM	z	SSPRM	z
intercept	0,032	4,00	-1,462	-98,13	-1,579	-92,36	1,6503	103,7	0,534	34,1
motor_time	-0,01	-421,5	-0,005	-284,2	-0,008	-810,00	-0,003	-93,3	-0,0014	-45,4
male_d	0,013	3,20	0,093	13,24	0,08	10,08	0,0365	5,21	0,018	2,59
hh_head_d	0,090	19,52	0,14	18,55	0,135	15,89	0,1087	14,3	0,0513	6,85
hh_income	0,013	10,40	0,007	5,23	0,009	5,00	7,37e-06	4,57	2,69e-06	2,04
auto_n	0,148	53,69	0,167	40,61	0,185	38,54	0,1525	35,6	0,0695	16,9
hh_size	-0,01	-11,06	-0,017	-10,12	-0,019	-10,47	0,003	1,75	0,0016	0,92
age	-0,0003	-2,79	-0,002	-8,50	-0,002	-6,00	-0,002	-7,77	-0,001	-3,66
yr_schooling	0,014	29,61	0,031	44,57	0,027	30,44	0,0069	9,06	0,0036	4,85
worked	0,493	122,51	0,962	139,42	0,90	120,29	-0,0367	-5,44	-0,015	-2,25
student	0,361	49,45	0,854	79,04	0,795	64,14	-0,073	-6,58	-0,032	-2,91
Dispersion					0,177	51,94				
Lambda							-0,16			
MIs Ratio										
Log L			-1201		-11710		-78580,5		-193614,7	
R ²	0,63									
Wild chi2(10)							10942,16		2591,8	
LR			50		5926					

Initial observation for the results of the parameters is that all variables have significant relationship with the total daily motorized trip frequencies, except only for household size. The results partially confirmed the initial expectations in LRM, PRM, and NBRM. However, when the model is corrected for endogeneity, the parameters converged to initial expectations. When we look at the result by SSPRM showed that all variables are positively contributes to total daily trip making. However, being student, being worked, daily motorized travel time, and age are the significant variables affecting daily number of motorized trips negatively. It should be noted that travel time is the most significant (negative as expected) variable among all. According to the results, all models significantly justify the fact that travel time refers to a kind of explicit disutility for individuals' propensities to mobilize in a day. In this sense, the marginal effect of travel time pronouncedly decreases as the model is corrected for endogeneity. This result is not surprising due to the fact that there is a kind of reciprocal relationship between daily travel time and number of daily trips of individuals.

Beyond travel time, age, being a student, and being a worked show negative correlations with daily number of motorized trips, and this result is consistent since students' ages are relatively younger. In addition, being worked would be referring purely to the obligatory trips called home based work trips, which would be able to minimize the other types of motorized passenger mobility such as recreation, shopping, and non-home based others in the case of Istanbul. The precursor factors -affecting daily motorized mobility of the individuals negatively- are being student, being worked, travel time, and age with coefficients 0.032, 0.015, 0.0014, and 0.001, respectively according to the results of SSPRM. On the other hand, the prominent factors inducing the urban motorized passenger mobility positively -with refers to the proxy variable called number of daily motorized trips of an individual-

are number of automobiles owned by the household and the dummy variable asserting that whether the individual is household head or not.

Even though we modeled the characteristics of trip making, the core of our paper is the magnitudes of marginal effect of these variables on total number of daily motorized trip making as a proxy for the measurement of induced urban motorized passenger mobility, and these estimations are presented in table 3 below. It should be noted that the key variable measuring induced urban passenger mobility is systematically over-estimated in LRM, PRM, NBRM being the highest in the LRM in both models. When the bias coming from discrete dependent variable was eliminated by the count models, the elasticity of travel time lowers to around 3 per thousand. Elimination of endogeneity, further, decreases the core elasticity around to 3,3 per thousand in SSM and 2,6 in SSPRM for motorized trips.

If we should take SSPRM as reference, actually we should since it eliminates all possible biases; it is possible to say that one percent decrease in travel time increase the trip making by 0,27 percent in motorized travel with refers to the fact that people are more sensitive to the motorized travel time changes. In motorized travel, car ownership and dummy variable of being household head have the highest inducement effects with the marginal elasticity coefficients that are equal to 0,13, and 0,096, respectively. That is to say, one more automobile owned by the household induces 0,13 more motorized trips for each member of that household. In addition, household head has a propensity to exhibit 0,096 more motorized trips in a day when compared to others in a household. Another important finding is that sex also has significantly high inducement effects on trips with the marginal elasticity coefficient that is equal to 0,034. Herein, it is able to be stated that males exhibit 0.034 more daily motorized trips when compared to the females in the case of Istanbul.

Table 3. Count Models' elasticity estimations for motorized daily trips

	LRM	PRM	NBRM	SSM	SSPRM
motor_time	-0,0109	-0,0029	-0,0040	-0,0033057	-0,0026353
male_d	0,0134	0,0502	0,0398	0,0365087	0,0339419
hh_head_d	0,0904	0,0753	0,0675	0,1087435	0,0968455
hh_income	0,0131	0,0037	0,0042	7,37e-06	5,08e-06
auto_n	0,1482	0,0901	0,0924	0,1524823	0,131163
hh_size	-0,0101	-0,0093	-0,0099	0,0029819	0,0029253
age	-0,0003	-0,0009	-0,0009	-0,0019955	-0,0017546
yr_schooling	0,0137	0,0169	0,0137	0,0068702	0,0068009
worked	0,4925	0,5206	0,4507	-0,0366814	-0,028163
student	0,3605	0,4620	0,3973	-0,0729874	-0,06028

Even though our findings confirm the existence of significant induced travel demand effect, as confirmed by the other studies in the literature of urban passenger mobility, the magnitude of it cannot be said overwhelming the magnitude of other trip determinants. In any case, the magnitude of induced urban passenger mobility demand effect remains modest in comparison to the combined effects of individual socio-economical characteristics. For example, the combined effects of number of auto and household income have higher generative effect than the generative impact of travel time saving

These travel time elasticities are higher than those of estimated by Barr [15], which was -0,44 on average. A reasonable explanation for such a result could be that the Turkish cities are more compact than the western cities. Another possible reason could be that the transportation cost is higher since Turkey heavily taxes petroleum. In either case, people would be more sensitive to travel. Furthermore, there are several differences between Barr's study and our approach. Barr used household level car trip data while our data include all trips and motorized trips at individual level. The difference may be attributed to this approach and data difference.

6. Discussions

According to the model results of the paper, to begin with, the findings for the case study of Istanbul have revealed that the induced motorized travel demand is still well and alive and obviously more measuring is required with many urban areas via many new studies.

Secondly, the findings of this paper is based on the case of Istanbul, which is a part of Turkey as a developing country and such empirical researches taking the cities of developing countries into considerations are highly limited. In this context, new empirical researches for different cities of different developing countries would be welcome and the potential findings of such studies would be compared by the findings of our study. In addition, it is to be highlighted that the findings of this study for the case of Istanbul supports the findings of a vast number of studies from western countries in literature asserting that induced travel demand is still alive everywhere and it is to be tackled with immediately after measuring it. The methodological discussions and propositions of this paper would be used via related potential future studies with different case cities other than Istanbul.

Thirdly, it should be highlighted that the findings of this study would be generalized to other cities in formulating and assessing any travel demand management policies. In this sense, the primary factors generating motorized trips would be able to be verified via different cases with different studies. Afterwards, the designation of travel demand management policies would be able to be realized according to these primary factors. To illustrate, in this study, according to the findings of

Sample Selection Poisson Regression Model (SSPRM), the primary factors generating motorized traffic are the number of automobiles owned by household, dummy variable asserting that whether the passenger is household head or not, and the dummy variable asserting that whether the individual is male or not with the marginal elasticities equal to 0,13, 0,1, and 0,034, respectively (see Table 3). That is to say, one more automobile owned by each household is able to make daily amount of motorized trips increase by 0,13 in addition to that a household head and a male has respectively 0,1 and 0,034 more propensity to realize a motorized trip in a day when compared to others in the case of Istanbul. These findings would be indicating that the travel demand management policies for the case of Istanbul as a whole might be formulated by the parameters, namely average number of automobiles per household, per cent of household head in total population, and per cent of men among total population. After verifying these findings by other cities in Turkey, the cities would be categorized according to the automobile ownership, percent of household head and men population among whole and then the top cities according to these parameters would be prioritized in formulating and realizing travel demand management policies. Such this approach might be enlarged by other cases from different countries and any potential similarities and/or differences from our findings would be discussed in detail.

Lastly, such these primary factors would be able to be used as performance indicators of any related travel demand management policies. For example, if the highest effect on inducing motorized trip comes from the number of automobiles owned by the household, then discouraging automobile ownership by additional taxations would be adopted by policy makers. Afterwards, the potential per cent decrease in the average number of automobiles owned by the households would be referring to the performance measure of success in implementing such a policy. Of course, the extensive discussions on policy formulations and on their implementation are able to be exhibited, but such these considerations would be out of the scope of this paper. Instead, we aim to highlight the potential factors inducing motorized trip making as an empirical baseline for designating travel demand management policies with refers to the intention of minimizing motorized trip making in a day. Future studies would be focusing more on formulation and implementation side of such these travel demand management policies with follows to new cases. Once more again, more measuring with much more case study areas (different cities from different countries) is needed on this front.

7. Concluding Remarks

The findings of this paper have explicitly confirmed the existence of induced urban passenger mobility demand of

individuals not only in terms of traveling farther but also in terms of traveling more. It is the first time in the literature that the measurement of induced urban passenger mobility demand is incorporated within a trip generation model via a series of regression models each eliminating a different level of bias in the parameter estimates for the considerations of urban passenger mobility. This variation in the core parameter estimate proves that selection of correct model with correct specification and estimator is essential. In this sense, a single equation *Sample Selection Poisson Regression Model (SSPRM)* seems a good candidate to measure the induced travel demand of urban motorized passengers and to integrate these findings into the related trip generation models.

The proposed model has shown that a ten minute decrease per capita in Istanbul Metro Area, which makes a 26 percent decrease in average motorized travel time, will increase the motorized trip making 6,5%. In such case, 10,45 million motorized trips will increase to 11,13 million trips in a day (0,68 million more motorized daily trips) in the case of Istanbul.

Another important finding of this study is that even though induced urban passenger mobility demand is *well and alive*, it is not the only and leading evil among other trip generating factors. For example, high elasticity of being employed or being a student implies that participation of women into labor force or increase in girls schooling might easily generate more trips than a transportation network improvement. Obviously more measuring is needed.

Non-linear structure of elasticity estimate of SSPRM may further allow us to estimate the spatial variation of generative impact of induced demand within the trip generation models since it is possible to account for the individual characteristics in the estimation of elasticities as long as we have disaggregate spatial data. The proposed methodology can be used iteratively to estimate the generative impacts of different investment scenarios from the trip assignment phase of travel demand modeling.

Also, the measurement of the generative impact of induced urban passenger mobility demand needs to be tested in different urban settings with the proposed model. Furthermore, it is clear that the functional specification has impact on the magnitude of the elasticities. As it was mentioned in the model section, there are many different specifications of count models especially to eliminate the heterogeneity and overdispersion in the data such as zero-inflated poisson, zero-inflated negative binomial, hurdle, or truncated poisson models obviously with more complex elasticity structures. Measurement of the elasticities with these alternative models may indicate a future research direction in the area of measuring the urban motorized passenger mobility so as to minimize (or manage) it.

REFERENCES

- [1] Downs, A. (1992). Stuck in Traffic: Coping with Peak-Hour Traffic Congestion. Brookings Institute.
- [2] Cervero, R., 2002. Induced travel demand: Research design, empirical evidence, and normative policies, *Journal of Planning Literature* 17(1): 3-20
- [3] Noland, R. B. and Lem, L.L., 2002, A review of the evidence for induced travel and changes in transportation and environment policy in the US and the UK, *Transportation Research Part D*, 7: 1-26.
- [4] Pells, S., 1989, User response to new road capacity: A review of published evidence, Working Paper 283, University of Leeds, Institute of Transport Studies, Yorkshire, UK.
- [5] Hansen, M., Gillen, D., Dobbins, A., Huang, Y., and Puvathingal, M., 1993, The air quality impacts of urban highway capacity expansion: Traffic Generation and land use change, Berkley: University of California, Institute of Transportation Studies.
- [6] Kroes, E., Daily, A., Gunn, H., and Van der Hoorn, T., 1996, The opening of the Amsterdam ring road: A case study of short term impacts on removing bottleneck, *Transportation*, 23: 71-82
- [7] Luk, J., and Chung, E., 1997, Induced demand and road investment: An initial appraisal, Research Report, ARR 299, Vermont South, Australia: ARRB Transportation Research.
- [8] Mokhtarian, P., Samaniego, F., Shumway R., Willits, N., and Azari, R., 2000, Analyzing induced traffic from capacity enhancements using matched pairs: A California Study, Davis: University of California, Institute of Transport Studies.
- [9] Hansen, M. and Huang, Y., 1997, Road supply and traffic in California urban areas, *Transportation Research Part A*, 31: 205-18.
- [10] Noland, R. B. and Cowart W. A., 2000. Analysis of metropolitan highway capacity and the growth in vehicle miles of travel, *Transportation* 27(4): 363-390.
- [11] Fulton, L., Meszler, D., Noland, R. and Thomas J., 2000, A statistical analysis of induced travel effects in the U.S. mid-Atlantic region, *Journal of Transportation and Statistics*, 3(1): 1-14.
- [12] Cervero, R. and Hansen, M., 2002, Induced travel demand and induced road investment: A simultaneous-equation analysis. *Journal of Transport Economics and Policy*, 36:469-490.
- [13] Cervero, R., 2003, Road expansion, urban growth, and induced travel: A path analysis, *Journal of the American Planning Association*, 69(2): 145-163.
- [14] Ozuysal, M. and S. Tanyel, 2008, Induced travel demand in developing countries: Study on state highways in Turkey, *Journal of Urban Planning and Development-ASCE* 134(2): 78-87.

- [15] Barr, L. C., 2000, Testing for the significance of induced highway travel demand in metropolitan areas, *Transportation Research Record: Transportation Planning, Public Participation, and Telecommuting*, (1706): 1-8.
- [16] Ortuzar J. D. and Willumsen L. G., 2001, *Modeling Transport*, 3rd edition, John Wiley & Sons, New York, N.Y.
- [17] Dickey J. W., 1983, *Metropolitan Transportation Planning*, 2nd Edition, Taylor and Francis, Bristol P.A.
- [18] Litman T., 2001, Generated traffic: Implications for transport planning, *ITE Journal*, 71(4): 38-47.
- [19] Jang, T. Y., 2005, Count data models for trip generation, *Journal of Transportation Engineering-ASCE*, 131(6): 444-450.
- [20] Gurmu, S. and Trivedi P., 1996. Excess zeros in count models for recreational trips, *Journal of Business & Economic Statistics*, 14(4): 469-477.
- [21] Long, J. S., 1997, *Regression models for categorical and limited dependent variables*, Sage Publications, Thousand Oaks, London, New Delhi.
- [22] Green, W. H., 2003, *Econometric Analysis*, 5th edition, Pearson, New Jersey.
- [23] Green, W. H., 2007, *Limdep version 9.0: Econometric modeling guide*, v.2, Econometric Software Inc., Plainview, New York.
- [24] Winkelmann, R., 2008, *Econometric analysis of count data*, 5th edition, Springer, Verlag, Berlin, Heidelberg.
- [25] Kennedy, P. 1998, *A guide to econometrics*, The MIT Press, Cambridge, Massachusetts.
- [26] Mokhtarian, P. and Cao, X., 2008, Examining the impacts of residential self-selection on travel behavior: A focus on methodologies, *Transportation Research Part B*, 42: 204-228.
- [27] Breen, R., 1996, *Regression models censored sample selected or truncated data*, Sage Publications, Thousand Oaks, California.
- [28] Heckman, J., 1979, Sample selection bias as a specification error, *Econometrica*, 47(1): 153-61.
- [29] Terza, J. V., 1998, Estimating count data models with endogenous switching: sample selection and endogenous treatment, *Journal of Econometrics*, 84:129-154.
- [30] Maddala, G. S., 1983, *Limited dependent and qualitative variables in econometrics*, Cambridge University Press, New York.