

Bayesian Hierarchical Tobit Models: an application to travel distance analysis

Modelos Tobit Bayesianos Jerárquicos: aplicación al análisis de la distancia de viaje

Jonathan Agüero Valverde
Universidad de Costa Rica. Director, Departamento de Transportes
Escuela de Ingeniería Civil
jonathan.aguero@ucr.ac.cr
<http://orcid.org/0000-0002-9096-9274>

Recibido: 1 de diciembre 2016

Aceptado: 12 de enero 2017

Abstract

The objective of travel distance models is to better understand travel behavior so that policies can be implemented for reducing travel and with that the externalities of transport such as air pollution, congestion, and crashes. Hierarchical Bayesian models offer a flexible framework to analyze travel behavior by allowing the study of short term decisions of the activity and travel choices as well as long term decisions of residential and employment location. Since travel distance is censored at zero for a significant fraction of the observations, parameter estimates obtained by conventional regression methods are biased. Consistent parameter estimates can be obtained by using the Tobit model. The purpose of this paper is to demonstrate the application of fully Bayesian Tobit hierarchical models to the analysis of travel distance, aiming to accommodate the multilevel and censored nature of the data.

Results show that the hierarchical Tobit Model performs significantly better than the non-hierarchical model as measured by the Deviance and Deviance Information Criteria. Further, the highly significant variance at the individual and location levels, demonstrates the importance of using a multilevel approach.

The distance traveled increases with years of study and job qualification. In addition, all the members of the household travel less than the householder and women travel less than men. Industry sectors also show significant differences in travel time: workers in the secondary and tertiary sectors travel farther than workers in the primary sector. Land price is significantly correlated with distance traveled in both residence and employment locations.

Keywords

Residential location, employment location, multilevel models, censored data



Resumen

El objetivo de los modelos de distancia de viaje es entender el comportamiento de viaje de los usuarios, de forma tal que se puedan implementar políticas para reducir la distancia de viaje y, con esto, externalidades del transporte tales como contaminación del aire, congestión y accidentes. Los modelos Bayesianos Jerárquicos ofrecen una metodología flexible para analizar el comportamiento de viaje al permitir el estudio tanto de las decisiones de corto plazo de la actividad y las selecciones de viaje así como las decisiones de largo plazo como la localización de la vivienda y el lugar de trabajo. Como la distancia de viaje está censurada en cero para una proporción importante de los datos, los parámetros obtenidos por medio de regresiones lineales convencionales están sesgados. Estimaciones no sesgadas de los parámetros pueden ser obtenidas usando modelos Tobit. El propósito de este artículo es demostrar la aplicación de modelos Tobit Bayesianos jerárquicos al análisis de la distancia de viaje, considerando la naturaleza multinivel y censurada de los datos.

Los resultados muestran que el modelo Tobit Bayesiano jerárquico tiene un desempeño significativamente mejor que el modelo no jerárquico al medir la bondad de ajuste la Devianza y el Criterio de Información de la Devianza. Más aún, la varianza es estadísticamente muy significativa tanto para el nivel individual como para el nivel de ubicación, lo cual demuestra la importancia de usar una metodología multinivel.

Palabras clave:

Ubicación de la vivienda, ubicación de lugar de trabajo, modelos multinivel, datos censurados

1. INTRODUCTION

The variables that affect travel distance and travel time have been of interest for transportation engineers and economists for several decades (1-11). The objective of travel distance and travel time models is to better understand travel behavior, so that policies can be implemented for reducing travel distances. With that, the externalities of transport such as air pollution, congestion, and crashes can be reduced as well (10).

Transport demand is derived; people travel in order to satisfy a need (study, work, leisure). These activities are distributed over space (12); hence, there is a close relationship between land use and transportation policy. A two way transportation-land use interaction exists, in which transportation is a derived demand from the urban system and the transportation system influences land development and location choice (13).

To model the interaction between transport and land use in travel distance models one can use a hierarchical or multilevel approach with at least two levels of characteristics: the individual characteristics and the location characteristics (14). Hierarchical models are appropriate to structure some dependence into the parameters at the higher levels (i.e. the location level), thus avoiding problems of overfitting or variability underestimation (15). This is due to the fact that individuals living and traveling to the same locations are clearly correlated but single level models are not capable of accommodating this correlation. Further, travel choices can be seen as embedded in a wider set of choices (16). According to Salomon and Ben-Akiva the short term decisions of the activity and travel choices are dependent on long term decisions of residential location, employment location, automobile ownership and mode of travel to work. The life-style choice is added at the top of the hierarchy. The life-style choice can be defined as the longer term decision on the type or patterns of activities one aspires to engage on (16). Figure 1 shows the hierarchy of analysis proposed for this study. Note that even though this paper makes no attempt to model the third level of the hierarchy proposed by Salomon and Ben-Akiva (i.e. life-style), the proposed approach can easily accommodate it; providing that one has a rich enough dataset to model life-style choices. Tobit models (17) have been used to model travel distance (4, 18-20) due to the censored nature of the data. Since travel distance and travel time are censored at zero for a significant fraction of the observations, parameter estimates obtained by conventional regression methods (e.g. ordinary least squares) are biased (21). Consistent parameter estimates can be obtained by using Tobit models which are a special case of the more general censored regression model.

Now, in order to accommodate the hierarchy proposed in a Tobit model a Full Bayes hierarchical approach is needed, since traditional frequentist approaches only allow for repeated measurements or panel data models (22). Fully Bayesian methods offer a comprehensive and robust approach to model estimation for multilevel models with nested random effects (23). In transportation, fully Bayesian Tobit approaches have been used mainly for estimating discrete-continuous models of vehicle choice and usage (10, 24, 25); however, the hierarchical nature of travel choice have not been explored in the context of travel distance.

The purpose of this paper is to demonstrate the application of fully Bayesian Tobit

hierarchical models to the analysis of travel distance; this with the goal of accommodating the multilevel and censored nature of the data. This paper is organized as follows: the next section describes the analytical methods used; then the dataset is described, followed by the discussion of results, conclusions and recommendations for future research.

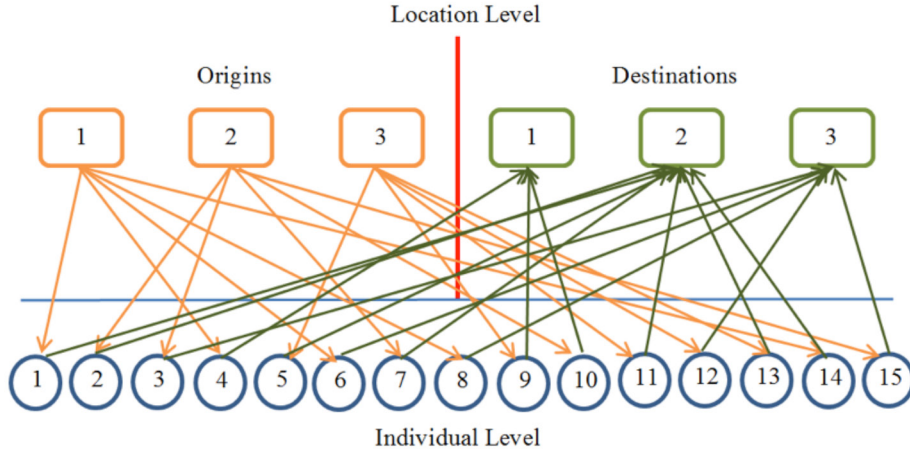


Figure 1. Hierarchy of analysis for travel distance modeling.

2. METHODOLOGY

Full Bayes hierarchical approach is used to estimate the models. Two levels of hierarchy are analyzed in this work: the person level and the location level. At the location level, residential location (origins) as well as employment locations (destinations) are analyzed.

The response variable is the distance traveled by each person from the household to the work place. The distance traveled follows a Tobit model (17):

$$\begin{aligned}
 y_{ijk}^* &= \boldsymbol{\beta}' \mathbf{x}_i + \epsilon_j + \omega_k, \\
 y_{ijk} &= 0 \text{ if } y_{ijk}^* \leq 0, \\
 y_{ijk} &= y_{ijk}^* \text{ if } y_{ijk}^* > 0.
 \end{aligned}
 \tag{1}$$

where y_{ijk} are the observed traveled distance for person i from location j to location k , $\boldsymbol{\beta}$ is the estimated coefficient vector for the individual characteristics, \mathbf{x}_i are the covariates or individual characteristics for person i , ϵ_j are the random effects for origin j , and ω_k are the random effects for destination k .

Now, assuming that the random effects for origins and destinations are normally distributed; the following prior distributions are proposed at the second level of the hierarchy:

$$\begin{aligned} \epsilon_j &\sim N(\boldsymbol{\gamma}'\mathbf{z}_j, \tau_\epsilon^{-1}), \\ \omega_k &\sim N(\boldsymbol{\delta}'\mathbf{w}_k, \tau_\omega^{-1}) \end{aligned} \tag{2}$$

where $\boldsymbol{\gamma}$ is the estimated coefficient vector for the location characteristics at the origins, \mathbf{z}_j are the covariates or location characteristics for origin j , and are the random effects for origin j . On the other hand $\boldsymbol{\delta}$ is the estimated coefficient vector for the location characteristics at the destinations, are the covariates or location characteristics for destination k , τ_ϵ are the random effects for destination k ω_k is the inverse of the variance for the origins, also known as precision and τ_ω is the precision for the destinations.

The hyperpriors for the coefficients, including the intercepts, are modeled using very non-informative Normal priors:

$$\begin{aligned} \beta &\sim N(0,1000) \\ \gamma &\sim N(0,1000) \\ \delta &\sim N(0,1000) \end{aligned} \tag{3}$$

The precisions have a gamma prior:

$$\begin{aligned} \tau_\epsilon &\sim \text{Gamma}(0.01, 0.001) \\ \tau_\omega &\sim \text{Gamma}(0.01, 0.001) \end{aligned} \tag{4}$$

with a mean of 10 and a variance of 10000, which makes this a very non-informative prior. In the case of the single level Tobit model equation 1 is replaced by

$$y_i^* = \boldsymbol{\beta}' \mathbf{x}_i + \psi_i \tag{5}$$

where ψ_i are the random effects for each person i :

$$\psi_i \sim N(0, \tau(-1)), \tag{6}$$

here τ has a gamma prior as defined by equation [4].

Two different goodness-of-fit measures are frequently used for model comparison in a fully Bayesian approach: posterior mean deviance (\bar{D}) and Deviance Information

Criterion (DIC). The posterior mean deviance can be taught as a Bayesian measure of fit or ‘adequacy’. To account for model complexity the Deviance Information Criterion was proposed by Spiegelhalter *et al.* (26).

$$DIC = D(\bar{\theta}) + 2p_D = \bar{D} + p_D \quad [7]$$

The DIC is considered the Bayesian equivalent of the Akaike Information Criterion (AIC). DIC is defined as an estimate of fit plus twice the effective number of parameters:

where $D(\bar{\theta})$ is the deviance evaluated at $\bar{\theta}$, the posterior means of the parameters of interest, p_D is the effective number of parameters in the model, and \bar{D} is the posterior mean of the deviance statistic $D(\theta)$. As with AIC, models with lower DIC values are preferred.

3. DATA DESCRIPTION

In contrast with most travel distance or travel time models that are based on travel survey data, this study is based on the data from the latest Census in Costa Rica, which was completed in 2011. For the first time in a Census in Costa Rica, the location of the job was recorded, which makes it possible to analyze the journey to work while also exploring the rich dataset from the census.

The spatial units used in the census are cantons, which are equivalent to counties or municipalities. The area of study was defined as the Greater Metropolitan Area in Costa Rica, which is composed of 31 cantons.

Since the census is not explicitly created to analyze travel data, the distance traveled for each person is not directly measured; therefore, the distance was estimated as the travel distance between the canton of residence and the canton where the job is located. This introduces censoring in the data, since each person that lives and works in the same canton has a ‘‘censored’’ travel distance to work of zero. In addition, the travel distance between cantons is estimated as the shortest path between locations which is not the path that all the persons follow; which potentially introduces a measurement error in the variable.

The descriptive statistics of the data are shown in Table 1. A sample of 10% of the population censused in 2011 is used in the analysis and only trips completely inside the area of study are included (78012 individuals). The Area of Study as well as a graphical representation of the Origen-Destination Matrix are shown in Fig. 2.

Costa Rica’s Greater Metropolitan Area includes the Capital City of San José as well as the other three major cities in the country: Alajuela, Heredia and Cartago. As shown in Fig. 1, these cities concentrate the majority of the jobs in the area and attract trips from cantons that can be considered suburbs of the cities such as Desamparados, Alajuelita and La Unión.

Table 1. Summary Statistics of the Data by Individual and Location Level

| | Variable | min | median | mean | max | S.D. |
|--|---|-----------|--------|--------|-------|-------|
| | Distance traveled (km) | 0 | 2.79 | 5.81 | 63.70 | 7.45 |
| | age | 12 | 35 | 37.21 | 98 | 12.44 |
| | number of dependent children | 0 | 0 | 0.57 | 19 | 1.17 |
| | years of study | 0 | 11 | 10.60 | 20 | 4.55 |
| | Categorical Variables | | | | | |
| | | frequency | % | | | |
| | Sex (female) | 29984 | 38.0% | | | |
| | Household Relationship | | | | | |
| | Householder | 36705 | 47.1% | | | |
| | Spouse or unmarried partner | 12741 | 16.3% | | | |
| | Son or daughter | 21959 | 28.1% | | | |
| | Other | 6607 | 8.5% | | | |
| | Job Qualification | | | | | |
| | Non-qualified | 11194 | 14.3% | | | |
| | qualified (technical) | 40769 | 52.3% | | | |
| | qualified (professional degree) | 26049 | 33.4% | | | |
| | Employment Status | | | | | |
| | self-employed | 15485 | 19.8% | | | |
| | Employee | 62527 | 80.2% | | | |
| | Industry Sector | | | | | |
| | Primary sector | 2735 | 3.5% | | | |
| | Secondary Sector | 15636 | 20.0% | | | |
| | Accommodation, Food Services and Trade | 23421 | 30.0% | | | |
| | Real Estate and Finance | 3704 | 4.7% | | | |
| | Health Care and Education | 15658 | 20.1% | | | |
| | Professional Services | 10388 | 13.3% | | | |
| | Other | 6470 | 8.3% | | | |
| | Transportation | | | | | |
| | Car ownership | 39105 | 50.1% | | | |
| | | min | median | mean | max | S.D. |
| | Land Price (thousands of Colones/m ²) | 0.792 | 18.453 | 31.061 | | 4.022 |
| | Density (thousands of persons/km ²) | 0.140 | 0.783 | 1.689 | 7.956 | 1.886 |

variables are included at the location level, land price and population density; however, only land price resulted significant in the model. As a result, population density was removed from the final model.

In terms of the Goodness-of-fit metrics the hierarchical model is far superior to the base model. With the inclusion of the location level variables and random effects the hierarchical model decreases the deviance from 529 408 to 504 245 or a 4.8%. The DIC also presents a significant reduction going from 529 427 to 504 324 or a 4.7%. As expected, the effective number of parameters (p_D) increases for the hierarchical model but not at the pace of the added degrees of freedom; p_D increased by about 60 while the degrees of freedom increased by 66.

The random effects also show a significant change when the hierarchical structure is introduced. The standard deviation (σ) for the individual random effects decreases from 7.2 to 6.1 but standard deviations at the location level are highly significant and of similar magnitude to the individual level: 5.5 and 6.0 for origins (σ) and destinations (σ) respectively. This clearly indicates that by introducing a second level in the analysis much more of the variability presented in the data is explained by the model.

In terms of the coefficients for the individual level variables, the magnitudes of all of them were reduced in the hierarchical Tobit model in comparison to the Base model. Some of the reductions were mild (i.e. about 20%) but some were more significant, reducing the coefficients values to less than half of the initial Base Tobit model. This result highlights the importance of including the hierarchical structure in the analysis of travel distance, since the coefficients are overestimated in the absence of location level covariates and random effects.

Since the hierarchical model is clearly superior to the single level Base Tobit model, the discussion about covariate parameters is based on the former model. All the covariates are significant for the individual level except for age, which is significant for the Base Model but non-significant for the Hierarchical Model. For the number of dependent children the coefficient is negative which is consistent with the findings from Maat and Timmermans (28). This result was expected since more dependent children in the household means more household responsibilities and therefore less time for traveling even for the journey to work.

In terms of level of education, the coefficient for years of study is statistically significant and positive. This is inconsistent with the results from Wilson (29) who found a negative relationship between education and travel distance but consistent with Feng *et al* (30) and LaMondia *et al* (31) who found a positive relationship for seniors and long commuters respectively.

Sex differences are also significant for travel distance. Women generally feel more responsible for caring for children, while men traditionally feel more responsible for providing the family income (32) which normally results in shorter work trips for women. For the hierarchical Tobit model the coefficient for female is negative which is consistent with many previous studies (11, 29, 33).

Among household characteristics the relationship with the householder is explored

Table 2. Estimates for Base Tobit and Hierarchical Tobit Models

| Variable | Base Tobit | | Hierarchical Tobit | |
|---|------------|--------|--------------------|--------|
| | mean | s.d. | mean | s.d. |
| constant | 0.660 | 0.206 | | |
| age | -0.008 | 0.003 | 0.001 | 0.002 |
| number of dependent children | -0.115 | 0.031 | -0.072 | 0.027 |
| years of study | 0.157 | 0.008 | 0.139 | 0.007 |
| Sex (female) | -0.728 | 0.075 | -0.571 | 0.064 |
| Household Relationship | | | | |
| Householder | | | | |
| Spouse or unmarried partner | -0.318 | 0.084 | -0.151 | 0.071 |
| Son or daughter | -0.263 | 0.072 | -0.232 | 0.061 |
| Other | -0.552 | 0.100 | -0.188 | 0.085 |
| Job Qualification | | | | |
| Non-qualified | | | | |
| qualified (technical) | 0.662 | 0.088 | 0.472 | 0.073 |
| qualified (professional degree) | 1.852 | 0.110 | 1.512 | 0.093 |
| Employment Status | | | | |
| self-employed | | | | |
| Employee | 2.227 | 0.069 | 1.349 | 0.059 |
| Industry Sector | | | | |
| Primary sector | | | | |
| Secondary Sector | 2.577 | 0.154 | 1.640 | 0.135 |
| Accommodation, Food Services and Trade | 1.840 | 0.151 | 0.902 | 0.133 |
| Real Estate and Finance | 2.584 | 0.190 | 0.817 | 0.167 |
| Health Care and Education | 1.049 | 0.161 | 0.333 | 0.141 |
| Professional Services | 2.685 | 0.160 | 1.305 | 0.143 |
| Other | 1.072 | 0.168 | 0.322 | 0.147 |
| Transportation | | | | |
| Car ownership | -0.591 | 0.058 | -0.343 | 0.050 |
| τ | 0.0193 | 0.0001 | 0.0266 | 0.0001 |
| $\sigma = \sqrt{\tau^{-1}}$ | 7.201 | 0.018 | 6.128 | 0.016 |
| Origin | | | | |
| Land Price | | | -0.107 | 0.024 |
| τ_{ϵ} | | | 0.034 | 0.009 |
| $\sigma_{\epsilon} = \sqrt{\tau_{\epsilon}^{-1}}$ | | | 5.531 | 0.748 |
| Destination | | | | |
| Land Price | | | 0.115 | 0.026 |
| τ_{ω} | | | 0.029 | 0.007 |
| $\sigma_{\omega} = \sqrt{\tau_{\omega}^{-1}}$ | | | 6.037 | 0.814 |
| GOF | | | | |
| \bar{D} | 529408 | | 504245 | |
| DIC | 529427 | | 504324 | |
| P_D | 19.1 | | 79.0 | |

Note: Statistically significant variables (95%) are shown in gray.

in the models. Among all the household members the head of the house have the longest journey to work since spouses, sons and daughters, and the other working persons in the household have negative coefficients.

Job qualifications are also important covariates in travel distance. Technical workers travel 0.66 km more than their non-qualified counterparts and professional workers 1.85 km. This indicates that the more qualified and possibly better paid the job is, the longer the workers are willing to commute.

At the location level, land price is significantly correlated with both origins and destinations but the effects are opposite. For the origins the coefficient is negative, which means that the lower the residential land price is the farther the workers are willing to travel. On the other hand, the land price coefficient for job location is positive, which is consistent with the location of better jobs in more expensive but also more convenient places such as the Central Business District. As mentioned before, residential and employment locations are long term decisions; therefore, the effect of variables such as land price and population density is not easy to capture in a cross-sectional analysis such as this one. Nevertheless, the current land price might be highly correlated with the past land price, which seems to be the case for the area of study.

The random effects at the location level give us the opportunity to explore the differences among locations. Table 3 presents the estimated random effects for origins and destinations. For identifiability purposes the constant is removed in the Hierarchical Tobit Model; hence, the random effects are not centered at zero. In any case, the differences among locations are more important than the absolute value.

In terms of the random effects by origin, cantons such as Belén, San José and Montes de Oca have the lowest values. For example, San José, which is the Capital City of Costa Rica, has the highest number of workers (116 754), the highest number of jobs (240 747) and most of its workers work in the canton (79 295 or 68.1%); which explains why workers in this location travel shorter distances to work. Besides, as shown in Fig. 2, these locations are positioned near the center of the metropolitan region which makes the work trips shorter since cantons are closer together. On the other hand, workers from cantons around the periphery of the region have some of the longer trips as shown also in Fig. 2. Atenas, Alvarado and Paraiso are some of these locations and the arrows depicting the most important pairs of the Origin-Destination matrix show significantly long trips.

For the destination random effects, the locations that attract jobs from farther distances include San José, Escazú, Santa Ana and Belén. Not surprisingly, San José, the Capital City and the business and political center of the country is among the locations that attract jobs with longer commuting trips. Escazú, Santa Ana and Belén are relatively new attractors based on office centers and high-tech industry, whose high-end jobs attract workers from longer distances.

On the other side of the spectrum, Atenas, Alvarado, Paraíso, El Guarco and Oreamuno are among the locations that attract jobs with shorter commutes. These towns are small and located in the periphery of the region. In addition a high percentage of attracted

Table 3. Random effects by Canton for the Hierarchical Tobit Models

| Canton | Origin | | Destination | |
|---------------------|---------|-------|-------------|-------|
| | mean | sd | mean | sd |
| San José | -10.205 | 0.724 | 10.539 | 0.721 |
| Escazú | -9.968 | 0.734 | 12.674 | 0.729 |
| Desamparados | -5.819 | 0.726 | 4.991 | 0.729 |
| Aserrí | -2.223 | 0.743 | 1.779 | 0.768 |
| Mora | -1.006 | 0.784 | 2.375 | 0.830 |
| Goicoechea | -8.079 | 0.729 | 7.240 | 0.730 |
| Santa Ana | -9.983 | 0.737 | 13.844 | 0.732 |
| Alajuelita | -5.378 | 0.732 | 4.408 | 0.751 |
| Vázquez de Coronado | -3.838 | 0.735 | 3.728 | 0.748 |
| Tibás | -7.874 | .732 | 8.732 | 0.733 |
| Moravia | -7.288 | 0.734 | 6.623 | 0.740 |
| Montes de Oca | -9.859 | 0.740 | 8.929 | 0.727 |
| Curridabat | -7.176 | 0.733 | 7.681 | 0.733 |
| Alajuela | -6.496 | 0.726 | 7.248 | 0.724 |
| Atenas | 11.705 | 0.804 | -10.439 | 0.840 |
| Poás | 2.729 | 0.774 | -2.692 | 0.800 |
| Cartago | 2.294 | 0.731 | -1.992 | 0.731 |
| Paraíso | 7.303 | 0.749 | -7.496 | 0.761 |
| La Unión | -2.397 | 0.731 | 2.734 | 0.739 |
| Alvarado | 15.158 | 0.850 | -14.872 | 0.885 |
| Oreamuno | 4.729 | 0.751 | -5.852 | 0.767 |
| El Guarco | 3.381 | 0.752 | -3.504 | 0.757 |
| Heredia | -8.275 | 0.730 | 9.511 | 0.726 |
| Barva | -5.444 | 0.741 | 5.188 | 0.762 |
| Santo Domingo | -7.036 | 0.743 | 8.107 | 0.744 |
| Santa Bárbara | -3.298 | 0.748 | 4.302 | 0.779 |
| San Rafael | -5.078 | 0.743 | 4.204 | 0.769 |
| San Isidro | -2.627 | 0.764 | 4.051 | 0.801 |
| Belén | -10.264 | 0.754 | 14.303 | 0.733 |
| Flores | -7.167 | 0.756 | 9.207 | 0.757 |
| San Pablo | -5.991 | 0.749 | 5.467 | 0.801 |

Note: Statistically significant variables (95%) are shown in gray.

jobs are occupied by persons from the same town reducing the commuting distances; these percentages range from 67.7% for El Guarco to 94.1% for Atenas.

5. SUMMARY AND DISCUSSION

This research develops a Fully Bayesian Hierarchical Tobit model of distance traveled from residence to employment locations in the Great Metropolitan Area of Costa Rica, for year 2011. Two levels of hierarchy are analyzed in this work: the person level and the location level. At the location level, residential location (origins) as well as employment locations (destinations) are analyzed.

Two models are estimated: a Base Tobit model and a Hierarchical Tobit model. The Hierarchical model is significantly better than the Base model in terms of both Deviance and DIC with reductions of about 5%. Besides, the variance of the random effects at the location level are highly significant and of similar magnitude to the variance at individual level. This clearly indicates that by introducing a second level in the analysis much more of the variability presented in the data is explained by the model.

The estimates for the coefficients also change with the inclusion of the second level variables and random effects in the model. The magnitudes of all of them are reduced in the hierarchical Tobit model in comparison to the Base model. This result stresses the importance of the hierarchical structure in the analysis, because the coefficients are overestimated in the absence of location level covariates and random effects.

According to the model, the distance traveled increases with years of study and job qualification. In addition, all the members of the household travel less than the householder. Industry sectors also show significant differences in travel time among workers in the models. Workers in the secondary and tertiary sectors travel farther than workers in the primary sector. Also consistent with previous studies, it is found that women travel less than men.

Land price is significantly correlated with both residence and employment locations. For the origins the coefficient is negative, which means that the lower the residential land price is the farther the workers are willing to travel. On the contrary, the land price coefficient for the destinations is positive, which is consistent with the location of better jobs in more expensive but also more convenient places.

The random effects at the location level show interesting results. The magnitude of the random effects by origin is lower for the cantons of Belén, San José and Montes de Oca. These cantons are located near the center of the metropolitan region which makes the work trips shorter. On the other hand, workers from cantons around the periphery of the region have some of the longer trips, including Atenas, Alvarado and Paraiso.

The destination random effects show that the locations that attract jobs from farther distances include San José, Escazú, Santa Ana and Belén; which can be explained by the high-end jobs that they offer. Contrastingly, Atenas, Alvarado, Paraiso, El Guarco

and Oreamuno are among the locations that attract jobs with shorter commutes. These towns are small and located in the periphery of the region.

Hierarchical Bayesian models offer a flexible framework to analyze travel behavior by allowing the study of short term decisions of the activity and travel choices as well as long term decisions of residential and employment location. As shown in this study, the hierarchical structure of those decisions can be easily accommodated in a fully Bayesian approach. Further research is needed to test many other variables that influence residential and employment location, particularly land use policies.

5. REFERENCES

1. Voorhees AM, Bellomo SJ, Schofer JL, Cleveland DE. Factors in Work Trip Lengths. *Highway Res Rec.* 1966;141:24-46.
2. Gruen AC. Travel time and transportation policy. *J Urban Econ.* 1980;8(2):264-71.
3. Izraeli O, McCarthy TR. Variations in Travel Distance, Travel Time and Model Choice among SMSAs. *Journal of Transport Economics and Policy.* 1985;19(2):139-60.
4. Golob TF, Wissen LV. A joint household travel distance generation and car ownership model. *Transp Res Part B Method.* 1989;23(6):471-91
5. Gordon P, Kumar A, Richardson HW. Gender differences in metropolitan travel behavior. *Reg Stud.* 1989;23(6):499-510.
6. Brännäs K, Laitila T. Modelling and prediction of travel distance by car. *Transportation planning and technology.* 1991;16(2):129-43.
7. Johnston-Anumwo I. The Influence of Household Type on Gender Differences in Work Trip Distance. *Prof Geogr.* 1992;44(2):161-9.
8. Rouwendal J, Rietveld P. Changes in commuting distances of Dutch households. *Urban Stud.* 1994;31(9):1545-57.
9. Aronsson T, Brännäs K. Household work travel time. *Reg Stud.* 1996;30(6):541-8.
10. Fang HA. A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transp Res Part B Method.* 2008;42(9):736-58.
11. Manaugh K, Miranda-Moreno LF, El-Geneidy AM. The effect of neighborhood characteristics, accessibility, home-work location, and demographics on commuting distances. *Transportation.* 2010;37(4):627-46.
12. Ortuzar J, Willumsen LG. *Modelling transport.* 3rd ed. John Wiley & Sons; 2011.
13. Miller EJ. Land use transportation modeling. In: Goulias KG, editor. *Transportation Systems Planning, Methods and Applications.* Boca Raton: CRC Press; 2003.
14. McQuaid RW. A model of the travel to work limits of parents. *Research in Transportation Economics.* 2009;25(1):19-28.
15. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian data analysis.* 2nd ed. London: Chapman & Hall/CRC; 2003.
16. Salomon I, Ben-Akiva M. The use of the life-style concept in travel demand models. *Environ Plan A.* 1983;15(5):623-38.
17. Tobin J. Estimation of relationships for limited dependent variables. *Econometrica.* 1958;26(1):24-36.
18. Redmond LS, Mokhtarian PL. The positive utility of the commute: modeling ideal commute time and relative desired commute amount. *Transportation.* 2001;28(2):179-205.

19. Golob TF. The dynamics of household travel time expenditures and car ownership decisions. *Transp Res Part A General*. 1990;24(6):443-63.
20. Schwanen T, Mokhtarian PL. What if you live in the wrong neighborhood? The impact of residential neighborhood type dissonance on distance traveled. *Transp Res D Transp Environ*. 2005;10(2):127-51.
21. Greene WH. *Econometric Analysis*. 7th ed. Pearson; 2012.
22. Aguiar-Moya JP, Prozzi JA. Accounting for Censoring and Unobserved Heterogeneity in Pavement Cracking. *Journal of Infrastructure Systems*. 2014;21(2): Available from: DOI: [http://dx.doi.org/10.1061/\(ASCE\)IS.1943-555X.0000233](http://dx.doi.org/10.1061/(ASCE)IS.1943-555X.0000233)
23. Congdon P. *Bayesian statistical modelling*. John Wiley & Sons; 2001.
24. Yang J, Miwa T, Morikawa T, Yamamoto T. Forecasting the Demand of Electric Vehicle Ownership and Usage in the Chukyo Region in Japan. *Fourth International Conference on Transportation Engineering*, American Society of Civil Engineers. 2013:245-51
25. Brownstone D, Fang HA. A vehicle ownership and utilization choice model with endogenous residential density. *J Transp Land Use*. 2014;7(2):135-51
26. Spiegelhalter D, Best N, Carlin BP, Linde A. Bayesian Measures of Model Complexity and Fit. *J R Stat Soc Series B Stat Methodol*. 2002;64(4):583-639.
27. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique, and future directions, *Stat Med*. 2009;28:3049-67.
28. Maat K, Timmermans H. A causal model relating urban form with daily travel distance through activity/travel decisions. *Transportation Planning and Technology*. 2009;32(2):115-34
29. Wilson FD. *Journey to Work: Metropolitan-nonmetropolitan comparisons*. Institute for research on Poverty. University of Wisconsin-Madison, 1976.
30. Feng J, Dijst M, Wissink B, Prillwitz J. The impacts of household structure on the travel behavior of seniors and young parents in China. *J Transp Geogr*. 2013;30:117-26.
31. LaMondia J, Aultman-Hall L, Greene E. Long-Distance Work and Leisure Travel Frequencies: Ordered Probit Analysis Across Non-Distance-Based Definitions. *Transp Res Rec*. 2014;2413:1-12.
32. Jorritsma P, Schaap NTW. Families on the run: How do Dutch households with young children organize their travel behavior? In *5th International Conference on Women's Issues in Transportation*. 2014:97-108.
33. Vance C, Hedel R. The Impact of Urban Form on Automobile Travel: Disentangling Causation from Correlation. *Transportation*. 2007;34:575-88.