URBAN FORM AND BUS RIDERSHIP IN SPANISH CITIES

Rocío Cascajo, TRANSyT-Transport Research Centre, Universidad Politécnica de Madrid, ETSI Caminos, Canales y Puertos, C/ Profesor Aranguren, s/n. 28040 Madrid, Spain

Steven Farber, McMaster Institute for Transport and Logistics, School of Geography and Earth Sciences, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada

Pablo Jordá, TRANSyT-Transport Research Centre, Universidad Politécnica de Madrid, ETSI Caminos, Canales y Puertos, C/ Profesor Aranguren, s/n. 28040 Madrid, Spain

Antonio Páez, McMaster Institute for Transport and Logistics, School of Geography and Earth Sciences, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada

Andrés Monzón, TRANSyT-Transport Research Centre, Universidad Politécnica de Madrid, ETSI Caminos, Canales y Puertos, C/ Profesor Aranguren, s/n. 28040 Madrid, Spain

ABSTRACT

The effect of urban form on mobility is a question with important implications for transportation and land use research and planning. A number of studies have investigated the relationship between mobility and various indicators of urban form while controlling for socio-economic characteristics. A majority of these have been concerned with mobility by private vehicles, although more recently there are also examples of research that explores the relationship between urban form and the use of public transportation. The objective of this paper is to investigate the demand for public transportation in a selection of cities in Spain, from the perspective of urban form (density, metropolitan area, population ratios), and other variables describing the characteristics of public transport supply (density of network and of stops), urban socio-economic profile, and competing and complementary modes (metro, suburban train and auto ownership rates). The paper is based on aggregate bus usage data collected for a number of Spanish cities and metropolitan areas for the years 2003-2007, and the use of autocorrelation models to account for intra-class correlations for multi-year observations of metro areas/cities. The results of the analysis indicate that demand for public transport in Spain is related to the size and population density of the urban area, the supply of bus routes, and the existence of a metro rail alternative. The results also indicate the existence of substantial autocorrelation effects.

Keywords

Public transport; Spain; urban form; density; autocorrelation models

INTRODUCTION

The interaction between urban form and mobility is recognized as a key topic of research that provides a useful point of contact, and sometimes creative tension, between the various measures that planners and policy makers can implement to achieve planning goals, for instance spatial planning measures (e.g. development policies, land use measures) or transportation interventions (e.g. supply of transportation infrastructure, travel demand management). Given the important implications of mobility patterns and urban form for the understanding of urban processes, and their practical implications, there have been numerous studies that attempt to clarify the relationship between these two aspects of urban systems – including research by Giuliano and Small (1993), Cervero and Kockelman (1997), Stead (2001), Maat et al. (2005), Giuliano and Dargay (2006), Van de Coevering and Schwanen (2006), Vance and Hedel (2007), and Estupiñán and Rodriguez (2008), among other authors.

Previous research on the topic has provided valuable insights about the relationship between mobility and urban form. For instance, Cervero and Kockelman (1997) carry out statistical tests using variables that aim to capture three different dimensions of neighbourhood design, namely density, land use diversity, and pedestrian friendliness. The results indicate that higher density, mix of uses, and accommodation of walking as a transport mode significantly correlate with non-auto travel. Van de Coevering and Schawnen (2006) in more recent research, further expand the scope of metropolitan-wide travel research by considering, in addition to urban form and transportation infrastructure, various other socio-demographic, housing, and history-related variables associated with inter-metropolitan differences in travel patterns. These authors find that there are significant variations between different regions, and uncover evidence to suggest that average distance travelled by car depends to a lesser degree on urban form and the socio-demographic profile of cities in the US compared to Europe or Canada. Other researchers also highlight that private motorised mobility is determined by the structure of a city. Cameron (2003), for example, applies a dimensional analysis to a wide-ranging set of possible drivers of urban mobility, including population of the metropolitan area, number of jobs, length of metropolitan road and rail networks, public transport passenger boardings and seat kilometres, average public transport trip length, annual number of walking and cycling trips, and gross domestic product per capita. With this analysis, he finds that the urbanised land area and population of an urban area determine to a large extent the private motorised mobility of the city and in turn the aggregate vehicle kilometres of travel. Vance and Hedel (2007), taking a different approach, study the determinants of the discrete decision to use the car and the continuous decision of distance travelled by estimating econometric models on a panel of travel-diary data, using a two-part model, a procedure involving probit and OLS estimators. The results suggested that urban form has a causative impact on car use, emphasizing the potential for integrating urban design into transportation demand management.

While most of the literature on urban form and travel patterns has been concerned with auto travel, there have been some forays into the theme of public transportation as well. In the paper previously cited, Van de Coevering and Schwanen (2006) examine the statistical relationships between land use and total distance travelled by public transport, reaching several conclusions: distance travelled by public transport tends to be longer in metropolitan areas that have more jobs per hectare in their cores, and also in those cities with lower public parking availability in their CBDs. In contrast, kilometres travelled by public transport decrease with increasing gross regional product. Another finding is that transit is more competitive for commuting than for other travel purposes. Finally, regression models for modal split for commuting show that the share of public transport rises in proportion to job density in inner parts of the city, but at the expense of walking and cycling. In separate research also concerned with public transport, Estupiñán and Rodriguez (2008) examine the built environment characteristics related to stop-level ridership for Bogotá's bus rapid transit (BRT) system, and finds that environmental support for walking and barriers to car use are related to higher BRT use. Also recently, Brown and Thompson (2008) use time-series analysis to compare patterns of transit patronage change over time with patterns of growth and decentralisation of population and employment in Atlanta and MARTA (the Metropolitan Atlanta Rapid Transit Authority). The results provide evidence that the decline in transit patronage is to some extent attributable to employment and population decentralisation in metropolitan areas. Based on these findings, Brown and Thompson suggest that said decline could be reduced if the transit system makes employment in non-centralized locations more reachable. In a recent study, Taylor et al. (2009) found that most of the variation in transit ridership among urbanized areas can be explained by factors outside of the control of public transit systems, i.e., regional geography, metropolitan economy, population characteristics, and auto/highway system characteristics.

As the brief review of the literature above indicates, several authors have concluded that urban form is indeed related to travel behaviour, both by private and public transportation. The extent of the relationship, on the other hand, seems to differ depending on the variables examined and also across regions. In this respect, it is important to note that there has been a great variety of approaches in terms of the level of analysis, from the study of neighbourhoods within cities (e.g. Cervero and Kockelman, 1997), macro-measures of urban form for metropolitan areas (e.g. Vance and Hedel, 2007), and intercity comparisons (e.g. van de Coevering and Schwanen, 2006). Some studies have made use of disaggregate data to investigate mobility patterns (e.g. Stead, 2001), while others have relied on aggregate measures of travel such as minimum average commute distance (e.g. Giuliano and Small, 1993). The present paper can be placed within the context of this body of research. More concretely, the objective of this paper is to investigate, at the city/metropolitan level, the relationship between aggregate measures of travel by public transport in Spain, specifically urban bus services, and a selection of urban form, public transport supply, and socio-economic variables, in addition to indicators of complementary and competing modes (i.e. presence of suburban trains and auto ownership rates).

Analysis in this paper is based on data drawn from the Metropolitan Mobility Observatory (MMO), a multi-city, multi-year effort to pool information regarding public transportation usage in Spain. While the main objective of the MMO is to provide information useful to assess the state of public transportation across cities, this is the first time there is an effort to estimate relationships between travel and urban characteristics in Spain. For the purpose of the analysis, spatial autocorrelation models are used to account for intra-class dependencies of substantive and nuisance natures (Anselin, 1988b). Models of this type have been used in the past to study autocorrelated error terms in travel flow models (Bolduc et al, 1992), as well as inter-personal dependencies on mode choice decisions (Goetzke, 2008). In the present case, autocorrelation models are essential for the analysis, since most metropolitan areas in the dataset are observed in multiple years, a violation of the independence assumption common to linear regression models.

CONTEXT FOR STUDY

Data source: Spain's Metropolitan Mobility Observatory

The data used in this study is drawn from the Spanish Metropolitan Mobility Observatory (MMO), an initiative established with the objective of observing and assessing the general mobility trends in the metropolitan areas comprised within the study. The MMO was created in 2003 following the initiative of a Reflection Group set up by the Public Transport Authorities (PTA) of the 6 main metropolitan areas in Spain, in conjunction with the Ministries of the Environment and Development and other national institutions. In 2009, the number of participants has increased to 16 metropolitan areas and their corresponding PTA, which represent the transit providers for about 22 million inhabitants, or in other words, about 49% of the total population in Spain (Monzón et al, 2009). The Observatory achieves its objective of assessing mobility trends by analysing specific indicators such as transport supply and demand, funding and investment and environmental indicators in the different participant areas. It is important to note, however, that this is the first time that these indicators are analyzed in a multivariate statistical framework. The MMO publishes a report every year with the main results of the indicators for one year, and the evolution of some of them from 2002. The Public Transport Authorities and the National Rail Operator (RENFE) are the main suppliers of data. The annual reports for the past five years serve as the data source for this research.

Within this analysis, only data from 9 metropolitan areas will be used. These areas include some major cities, such as Madrid, Barcelona and Valencia, in addition to a number of medium- and small-size urban areas (see figure 1). More generally, the areas can be classified in 3 different classes regarding their population-size. Four of them are considered major metropolitan areas, with metropolitan populations exceeding one million inhabitants (Madrid, Barcelona, Valencia and Seville). A second medium-size class with populations ranging from 500,000 to one million inhabitants is constituted by one urban area (Bilbao). Finally, there is a group of 4 smaller urban areas with less than 500,000 inhabitants (Alicante, Corunna, Oviedo and Pamplona). These urban locations display a wide range of attributes in terms of size, population and number of municipalities integrated in the metropolitan area. Their structure on the other hand also presents some similarities, since a majority of these urban areas consist of a densely populated urban core (from 1,100 to 15,000 inhabitants per km^2), and a lower density surrounding metropolitan ring. Regarding the evolution of the population, the number of inhabitants in these areas has increased on average at the rate of 1.6% per year in the period covered (between 2002-2007), with much of this growth being absorbed by the outer metropolitan rings as opposed to the urban centres, thus reflecting ongoing dispersion trends. Several of these metropolitan areas have been the object of substantial immigration influxes since the year 2000, and as a result their rates of growth have been more considerable, sometimes exceeding 10% in this period. This is the case of Madrid, Seville and Alicante. The MMO provides information for a selection of metropolitan areas of various sizes, and for different years depending on when the transportation authorities in the cities concerned joined the effort and shared their information for the database. In addition to the data provided by the PTA of the different cities/metropolitan areas, information is provided by the national rail operator (RENFE) on suburban rail networks, by the General Directorate of Traffic (DGT) on accidents, and other information independently retrieved from the National Statistics Institute (INE).

Figure 1 *– Metropolitan areas comprising within the study*

Public Transport in Spain

Public transport is an important element of the transportation system in many cities in Spain, with considerable shares in some cities that moreover have increased in recent years (12.9%

on average of total trips between 2002 and 2007). There is a consensus among transit operators that consistently high quality of service, good coordination between the different service providers involved, and fare integration are key elements for the continued success of the public transport system in Spanish urban areas. At the same time, there are concerns that other factors, which escape the control of public transport providers, may negatively affect their ability to offer an effective and attractive alternative to other modes. In particular, the continued outward growth of cities, changes in density of development, and increased motorization of the population, represented by the rates of auto ownership, are factors that may negatively affect public transportation.

To be sure, public transport patronage in Spain varies considerable across metropolitan areas and by trip purpose. The share of public transport in major cities is higher than in smaller ones for all journeys' purposes (Madrid: 31.6%, Barcelona: 18.6%, Seville: 13.4%, Alicante: 12.3%), and the use of public transport in many places tends to be slightly higher when the purpose is travel to work or to school (Madrid: 40.4%, Barcelona: 23.9%, Seville: 12.1% Alicante: 8.8%). The picture for car trips is similar: private cars are used more for trips to work, reaching shares that exceed 50%. This similarity highlights a notable characteristic of Spanish cities compared to other countries, that is the high share of walking trips, especially for the case of non-compulsory journeys (leisure, shopping), which display shares between 40-60% depending on the city. However, the proportion of walking trips in compulsory journeys (work or studies) decreases in favour of car trips. There is also considerable use of suburban rail services in these areas, with rates around 10% to 20% in relation to the total journeys taken on public transport. These data substantiate international experience which indicates that an adequate supply of rail modes must be available in order to achieve a high quota of public transport patronage (Monzón et al, 2007). The high shares observed for public transport and walking trips are thought to be related to historically high levels of density in Spanish cities, an attribute that is gradually declining with the increased dispersion of metropolitan developments.

Descriptive Statistics

Information collected from the different reports of the Spanish Metropolitan Mobility Observatory covers the period between 2003 and 2007. Although 16 metropolitan areas collaborated in the MMO last report, not all PTAs submitted the data required for the analysis. Therefore, the research reported here is concerned with 9 metropolitan areas, most of which provide multiple year information (see table 1). The variables considered in the analysis are defined in table 2, and their descriptive statistics are shown there as well.

Table 1 *– Urban Areas in Study and Years Available*

	Year				
	2003	2004	2005	2006	2007
Madrid	✓	✓	✓	✓	✓
Barcelona	✓	✓	✓	✓	✓
Valencia	✓	✓	✓	✓	✓
Seville	n.a.	✓	✓	✓	✓

Bilbao	✓	✓	✓	n.a.	n.a.
Alicante	n.a.	✓	✓	n.a.	n.a.
Corunna	n.a.	✓	✓	✓	n.a.
Oviedo	n.a.	n.a.	n.a.	n.a.	✓
Pamplona	✓	n.a.	n.a.	n.a.	n.a.

✓: Data available
n.a.: Data not available

ANALYSIS

Methods

Spatial regression modelling is a statistical technique widely used in the analysis of cross-sectional data since it allows for the determination of autocorrelated dependent observations, in the case of the lag model, or an autocorrelated covariance structure in the case of the error model (Anselin, 1988a). These techniques are commonly applied to spatial economic analyses such as hedonic house price modelling (Can and Megbolugbe, 1997; Basu and Thibodeau 1998) and models of spatial spillovers (Fischer and Varga, 2003), but given the spatial structure of transportation data, applications of spatial regression techniques also exist in the transportation literature (Bolduc et al., 1992; Goetzke, 2008). Similarly, analysis of the clustered structure of the MMO dataset into city groups lends itself to spatial regression techniques. In particular, given the stability of transportation behaviour over time, it is likely that observed public transport usage in a given city and year is dependent on observations for the same city at different times. This substantive form of autocorrelation can be modelled with a lagged dependent specification to ensure estimation consistency and efficiency. In addition to this substantive form of autocorrelation, nuisance autocorrelation, arising from a structured pattern of missing information and resulting in an autocorrelated residual vector, must also be investigated in order to insure efficient parameter estimates.

The spatial lag and error models are both simple extensions of the linear regression model. Given a vector of observations on the dependent variable, Y, a matrix of exogenous factors, X, and a matrix W representing the contiguity structure of the analysis units, the spatial lag model is formulated as:

$$Y = \rho WY + XB + \varepsilon,$$

where ρ and B are parameters estimated via maximum likelihood methods, and ε is a vector of random normal errors with mean 0 and estimated variance. Simply put, this model supposes that the observed variable is a function of other measures of the observed value for neighbouring units, as well as an additive linear function of exogenous attributes of the units.

***Table 2** – Variable Definitions and Descriptive Statistics*

Type of variable	*Variable name*	*Description*	*Min*	*Max*	*Mean*	*Stand. Dev.*
Demand for public transport	Trips-km	Number of trips-km in urban bus network per main city inhabitants	204.5	596.8	386.66	123.01
Urban Form	Metro Population	Metropolitan population (inhabitants)	216607	6081689	2471883.5	2194342.0
	Main Population	Main city population (inhabitants)	190937	3155359	1347623.9	1204062.1
	Metro Surface	Metropolitan area (km^2)	36.8	8030.1	2554.0	2752.6
	Urban Surface	Main city area (km^2)	25.0	606.0	234.1	198.0
	Population Density	Number of residents/km^2 for the main city	1158.3	8609.7	5986.5	2094.4
	Density Difference	Percentage difference between main city density and the density of the ring	0.56	1.00	0.890	0.106
	City/Metro population ratio	Ratio between population in the main city and population in the whole metropolitan area (%)	0.31	1.00	0.60	0.19
	Large City	City population greater than 500kinhabitants (1 when the city has more than 500kinhabitants, 0 otherwise)	0	1	0.7	0.5
Transport Supply	Bus Length	Length of bus lines per main city inhabitants	5.75	23.11	9.209	3.403
	Bus Stops	Number of stops within the bus network per 1,000 main city inhabitants	1.94	4.96	2.78	0.73
Socio-economic characteristics	GDP	Per capita gross domestic product (€)	15524	30419	21852.0	4100.2
Competing and complementary modes	Motorization	Number of cars per 1,000 inhabitants	389	546	461.4	38.5
	Metro trips-km	Number of trips-km in metro network (million)	0	4807.0	1202.55	1615.12
	Light rail trips-km	Number of trips-km in light rail network (million)	0	120.8	18.99	32.75
	Suburban railway trips-km	Number of trips-km in suburban railway network (million)	0	4123.5	1420.07	1633.67
	Metro	Existence of metro network (1 when metro network present, and 0 otherwise)	0	1	0.62	0.49
	Light Rail	Existence of light rail network (1 when light rail network present, and 0 otherwise)	0	1	0.45	0.51
	Suburban Rail	Existence of suburban rail network (1 when suburban rail network present, and 0 otherwise)	0	1	0.86	0.35

Similar to the lag model, the error autocorrelation model can be written as:

$$Y = XB + (In - \lambda W)^{-1} \bullet \varepsilon$$

where *In* is the nxn identity matrix, λ is the strength of error association to be estimated, and all other variables are defined as above. Both cases reduce to ordinary least squares regression in the absence of significant autocorrelation (i.e. if ρ, λ = 0). This results in some fairly straightforward test formulations based on the model likelihood functions (Anselin, 1988b; Anselin et al, 1996), namely the likelihood ratio and Lagrange multiplier (LM) tests.

In the application currently under investigation, the matrix of contiguity, W, is simply defined as a binary matrix with W_{ij} equal to 1 if observations i and j are on the same city, and 0 otherwise. In this application, the W matrix is row-normalized such that $(WY)_i$ can be interpreted as a neighbourhood average for observation i, and $(In - \lambda W)^{-1} \bullet \varepsilon$ is an autocorrelated error vector. Given that each city needs at least 1 neighbour in W for the accurate calculation of LM diagnostics in GeoDA (Anselin et al., 2006), the two cities that were only observed once, Pamplona and Oviedo, were joined together as neighbours in W. Following some experimentation with joining these cities to existing city groups, which always resulted in less precise estimates, the cities were joined with each other in order to contain the potential misspecification error.

Specification

Specification of the final model concerns both the selection of the most appropriate form of regression (i.e. OLS, Lag, or Error model), as well as a parsimonious set of regression variables. The current state of practice in the spatial modelling literature suggests constructing a best fitting OLS regression with the regressors available, followed by a series of diagnostic tests on the residuals of the model in order to determine which (if any) spatial specification is the most appropriate. In this case, the process of variable selection was quite tedious due to the high degree of multicollinearity within the assortment of regression variables available for analysis. The final selection of variables (seen in Table 3) contains representative vectors from the following dimensions of explanatory variables: urban form, transport supply, and competing mode-use (see Table 2), but we could not use the economic indicator, GDP per capita, without introducing severe levels of variance inflation due to multicollinearity. These variables explain a very large proportion of the variance of urban bus use per capita as illustrated by an extremely high R-Square and a highly significant likelihood ratio score with respect to the constant-only model. In addition, none of the variables have variance inflation factors above 2.5, suggesting a high degree of linear independence between them and that their confidence bounds are fairly unaffected by multicollinearity. (Fox, 2002). Despite this high level of fit, and the presence of well-signed, independent, and significant regression coefficients, the diagnostic tests for residual and substantive dependence in Table 4 suggest that the OLS model is mis-specified. As is often the case with autocorrelated data, the LM tests suggest that both the LAG and ERR models are significant improvements over the OLS specification, but LM(ERR)>LM(LAG), and the robust

LM(ERR) test is far more significant than the robust LM(LAG) test, evidence to suggest that the correlated error process is most suitable (Florax et al., 2003).

In addition to the diagnostic LM tests, the assumption of error autocorrelation is far more palatable than the lag assumption. Suggesting that there is a structural pattern to missing information and measurement errors that follows city groupings defined by W is extremely reasonable, whereas the assumption that bus usage in one year actually depends on bus usage in another year for the same city is slightly more difficult to comprehend. Moreover, as is seen in Table 3, the error model produces results with the highest level of fit (indicated by the R-Squared, Schwartz Criterion, and Log Likelihood), the lowest standard error of regression, the most significant improvement over the OLS model (measured by the likelihood ratio test), and it is the only model to produce random errors that do not contain significant levels of autocorrelation within city groupings (according to Moran's I). Thus, from an interpretative point of view, over and above the similarly poignant diagnostic tests, the error model is preferred to the lag model.

Table 3 *– Results of model estimation*

	Model 1 (OLS)		Model 2 (ERR)		Model 3 (LAG)	
Variable	estimate	p-value	estimate	p-value	estimate	p-value
Constant	373.085	0.0000	338.22	0.0000	95.831	0.0212
λ	-	-	0.673	0.0000	-	-
ρ	-	-	-	-	0.6073	0.0000
Urban Surface	0.315	0.0000	0.331	0.0000	0.137	0.0010
Population Density	-0.033	0.0000	-0.029	0.0000	-0.008	0.0423
Bus Length	16.865	0.0000	18.359	0.0000	9.654	0.0000
Large City	39.998	0.0370	42.900	0.1361	21.736	0.0619
Metro	-73.881	0.0023	-88.217	0.0045	-46.405	0.0016
R-Squared	0.913		0.975		0.966	
S.E. of Regression	36.23		19.19		22.24	
Schwartz Criterion	303.99		282.43		292.21	
Log Likelihood	-141.892		-131.114		-134.319	
	Stat	p-value	Stat	p-value	Stat	p-value
Likelihood Ratio [a]						
vs Model 0	16.496	0.0210	-	-	-	-
vs Model 1	-	-	21.56	0.0000	15.14	0.0001
Moran's I (Residuals)	0.6707	0.0002	0.0565	0.2715	0.2331	0.0585

[a] Model 0 is the constant only OLS model

***Table 4** – Diagnostics of OLS residuals for dependence*

Test	Stat	p-value
Lagrange Multiplier (LAG)	13.0228	0.0003
Robust LM (LAG)	0.7450	0.3881
Lagrange Multiplier (ERR)	15.6553	0.0001
Robust LM (ERR)	3.3774	0.0661

Results and Discussion

In this section, the coefficients of the error model (found in the middle section of Table 3) are interpreted. Comparing the coefficient of the constant with the mean value of the dependent (in Table 2), we observe that the constant takes a reasonable (i.e. non-negative) value quite close to the sample-mean of the dependent variable. This intercept is shifted upward for cities with high populations (Madrid, Barcelona, Valencia and Seville), indicating that all else being equal, in these cities the bus is used to a greater extent in comparison to the lesser populated cities.

The two coefficients for urban form variables, urban surface area and population density, indicate that bus kilometres travelled per person increases with increased surface area, and decrease with urban density. While the first finding is clearly a result of needing to travel longer distances in larger cities, the second is somewhat more controversial given that urban density is often toted as a key factor influencing people to use more public transit (Cervero, 1998). However, this model is predicting bus kilometres travelled, not the mode choice decision, and holding all else constant, bus trips in more compact cities such as Barcelona, Bilbao and Corunna apparently are shorter than those in more sprawling metropolises such as Alicante, Seville and Madrid. It is also related to the particularity of Spanish cities, whose high density incites people to walk, instead of take the bus.

Secondly, the indicator of bus transport supply, length of bus network per capita, obtains a very large and positive coefficient that suggests that increasing bus network length by one kilometre per person, results in 18 additional kilometres travelled per person, a very large response to an increase in bus supply.

Finally, the coefficient for the competing and complementary modes variable, Metro, indicates that the distance travelled per person by bus decreased with the existence of a metro system in the city. This finding is reasonable since there is a shift of passengers from bus to metro in those cities having this rail mode. There are also other reasons for Metro presence reducing bus trip length. For instance, metro trips should be longer than bus trips, especially in city centres where congestion of roads makes metro trips faster.

In order to further assess the goodness-of-fit of the model, the coefficients obtained from estimation are used to project values of the dependent variable. The diagnostic plots in figures 2 and 3 illustrate the relative fits of the three models. Since all the models have very high R-Squares, we are not surprised by the overall excellent fit observed in Figure 2, but do draw your attention to the poor performance of the Lag model for Oviedo and Pamplona, no doubt a function of their being forced into a common city group for technical reasons as noted above. Besides these two observations, the models fit extremely well, with no systematic difference in prediction error size and direction along the range of observed values. The real advantage of the error model is best observed in Figure 3, a plot of the model residuals (ε) across the cities on the horizontal axis. Here we see that the random components of the spatial models are more frequently closer to zero than for the OLS specification, and seldom do we find that the residual of the lag model is smaller than the error model. This graphical evidence supports our suggestion that the best specification is in fact the correlated error model.

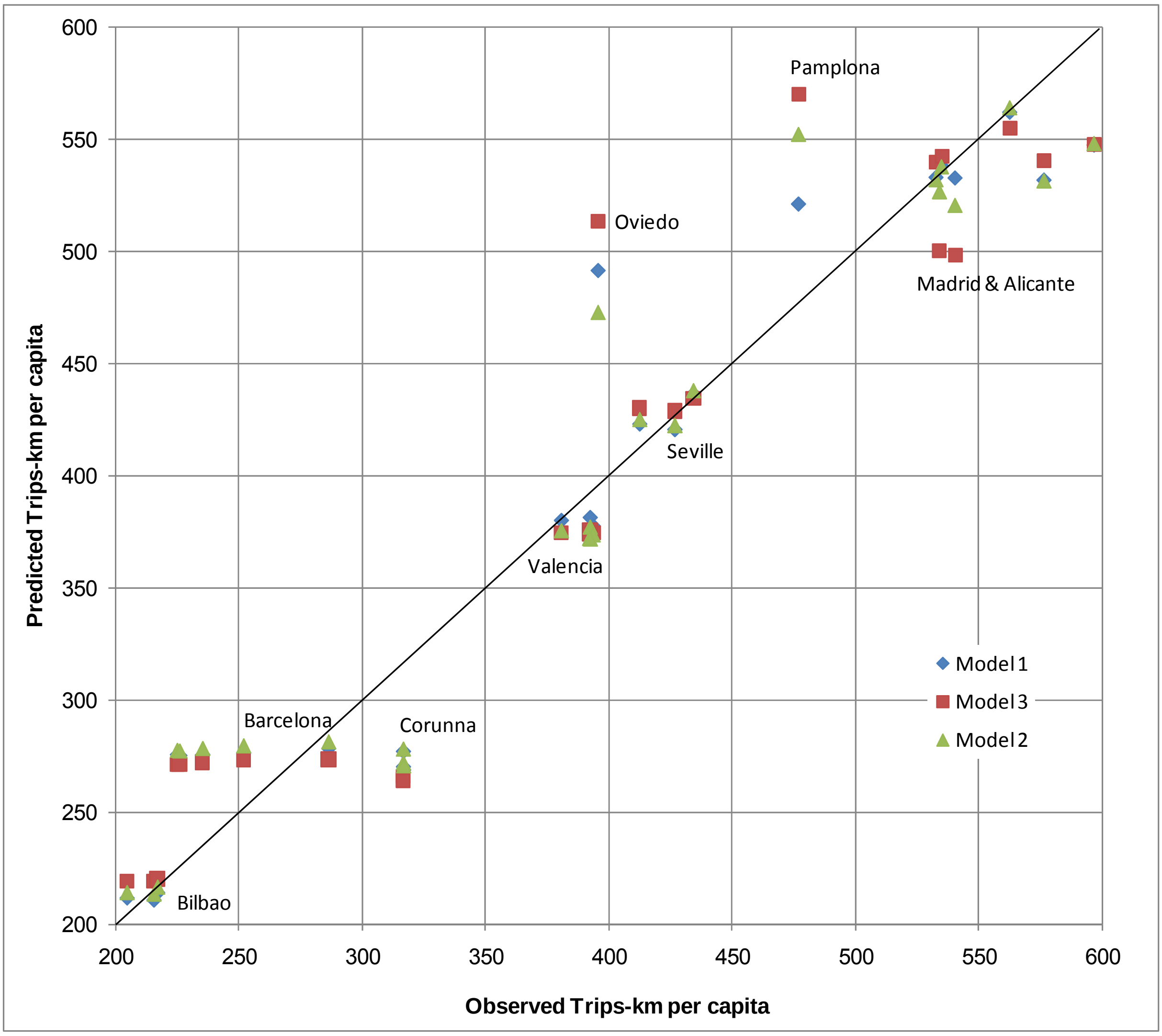

***Figure 2** – Observed and Estimated Values*

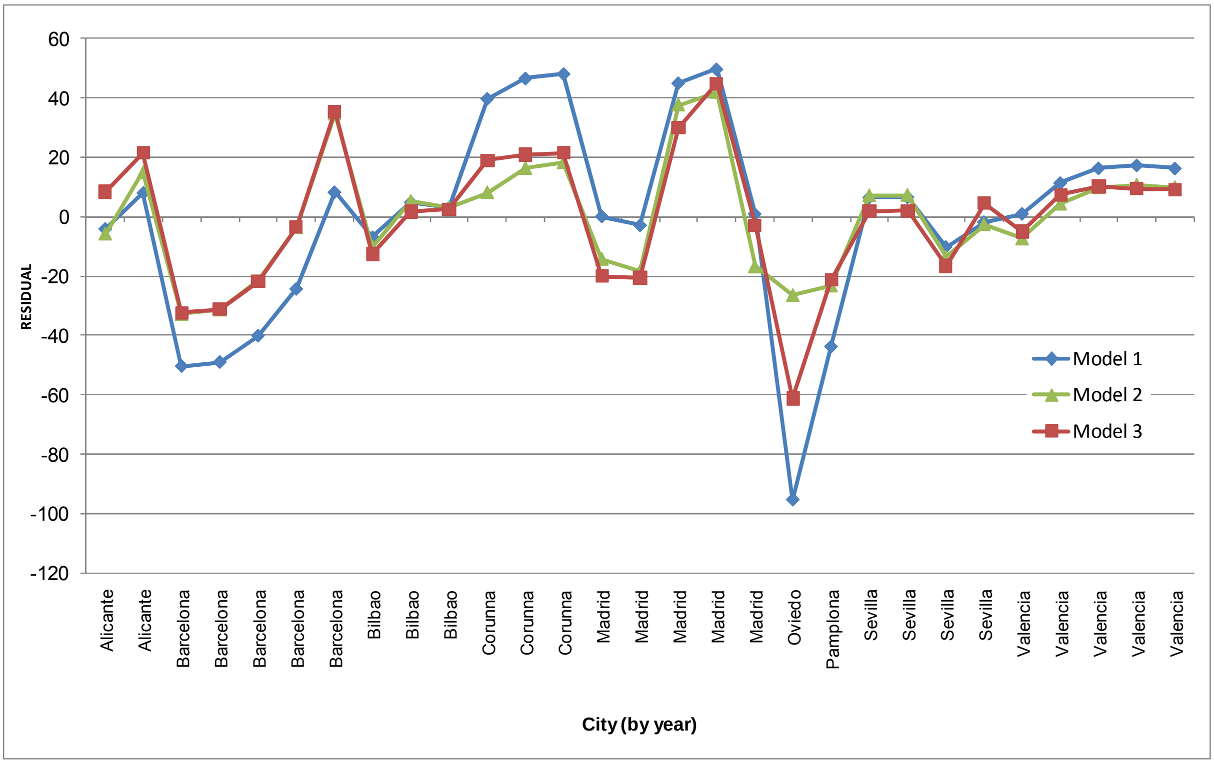

Figure 3 *– Model Residuals*

SUMMARY AND CONCLUSIONS

The objective of this paper has been to investigate the variables that affect the demand for public transport in a selection of Spanish metropolitan areas, with a particular focus on urban form variables. The set of variables available for the analysis also allow for comparisons between public transport supply characteristics, socio-economic conditions, and indicators of competing and complementary modes.

In order to examine the relationship between bus ridership and the variables mentioned above, autocorrelation models were estimated. Autocorrelation modelling techniques are essential in this case in order to account for intra-class correlation, given the multi-year nature of observations for some urban areas. A stepwise approach to variable selection was followed by the estimation and comparison of three model specifications. A spatial error model, with intra-class residual correlation was selected as the most appropriate, and in the end, the final model obtained included three urban form variables, one indicating transit supply, and another for the existence of a competing and complementary mode, namely an urban metro system.

Some conclusions can be gleaned from the analysis presented in this paper. First, the urban form factors, city size and population density, significantly affect the level of per capita urban bus use. Interestingly, while people seem to travel longer distances by bus in larger cities, increasing density seems to reduce the demand for bus transport. Increased density at the local scale is generally understood to be a transit supportive urban form, but given the high degree of land-use mixing typically found in Spanish cities (in comparison to their North American counterparts), increasing density may shift transit riders towards walking (mainly) and cycling, and the negative coefficient on density may be capturing this effect. This result is somehow supported by Cervero (1996), because he stated that the presence of mixed land uses is associated with relatively low vehicle ownership rates and short commuting distances among residents. In terms of transit supply, the model confirms that while holding all else constant, cities with more comprehensive bus network coverage enjoy higher levels of bus trip kilometres per person. The direction of causality in this relationship is not necessarily evident from this model, but this variable serves as a valuable controlling factor so that the relationship with the urban form variables can be interpreted more directly. Similarly, the existence of a metro system makes that the length of the bus trips are shorter than in those cities with no presence of metro.

In addition to those variables that were significant and therefore included in the final model specification, of interest are those variables that were not. In particular, neither Gross Domestic Product nor the motorization level is a significant predictor of bus use. So, inter-city variation in prosperity does not seem to account for metro use in Spain.

REFERENCES

Anselin, L. (1988a). Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. Geographical Analysis 20, 1–17.

Anselin, L. (1988b). Spatial Econometrics: Methods and Models. Kluwer Academic: Boston.

Anselin, L.; Bera, A.; Florax, R. J. G. M.; Yoon, M. (1996). Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics 26, 77-104.

Anselin, L.; Syabri, I.; Youngihn, K. (2006). GeoDa: An Introduction to Spatial Data Analysis. Geographical Analysis 38, 5-22.

Basu, S.; Thibodeau, T.G. (1998). Analysis of spatial autocorrelation in house prices. The Journal of Real Estate Finance and Economics, Vol. 17, No. 1, pp. 61-85.

Bolduc, D.; Laferrièrre, R.; Santarossa, G. (1992). Spatial autoregressive error components in travel flow models. Regional Science and Urban Economics, Vol. 22, No. 3, pp. 371-385.

Brown, J.R.; Thompson, G.L. (2008). The relationship between transit ridership and urban decentralisation: Insights from Atlanta. Urban Studies, Vol. 45, No. 5-6, pp. 1119-1139.

Cameron, I.; Kenworthy, J.R.; Lyons, T.J. (2003). Understanding and predicting private motorised urban mobility. Transportation Research Part D-Transport and Environment, Vol. 8, No. 4, pp. 267-283.

Can, A.; Megbolugbe, I. (1997). Spatial Dependence and House Price Index Construction. The Journal of Real Estate Finance and Economics, Vol. 14, No. 1-2, pp. 203-222.

Cervero, R. (1996). Mixed land-uses and commuting: Evidence from the American Housing Survey, Transportation Research Part A: Policy and Practice, 30(5), 361-377.

Cervero, R.; Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. Transportation Research Part D-Transport and Environment, Vol. 2, No. 3, pp. 199-219.

Cervero, R. (1998). The Transit Metropolis. Island Press, Washington DC.

Estupinan, N.; Rodriguez, D.A. (2008). The relationship between urban form and station boardings for Bogota's BRT. Transportation Research Part A-Policy and Practice, Vol. 42, No. 2, pp. 296-306.

Fisher, M.M.; Varga, A. (2003). Spatial knowledge spillovers and university research: Evidence from Austria. The Annals of Regional Science, Vol. 37, No. 2, pp. 303-322.

Florax, R.J.G.M.; Folmer, H.; Rey, S. (2003). Specification searches in spatial econometrics: the relevance of Hendry's methodology. Regional Science and Urban Economics 33, 557-579.

Fox, J. (2002). An R and S-Plus companion to applied regression. Sage Publication: Thousand Oaks, CA.

Giuliano, G.; Small, K.A. (1993). Is the Journey to Work Explained by Urban Structure. Urban Studies, Vol. 30, No. 9, pp. 1485-1500.

Giuliano, G.; Dargay, J. (2006). Car ownership, travel and land use: a comparison of the US and Great Britain. Transportation Research Part A-Policy and Practice, Vol. 40, No. 2, pp. 106-124.

Goetzke, F. (2008). Network Effects in Urban Transit Use: Evidence from a Spatially Varying Model Choice Model for New York. Urban Studies, Vol. 45, No. 2, pp. 407-417.

Monzón, A.; Cascajo, R.; Jordá, P. (2007). Summary of the 2005 Report of the Metropolitan Mobility Observatory, Ministerio de Medio Ambiente y Ministerio de Fomento.

Monzón, A.; Cascajo, R.; Jordá, P. (2009). Informe 2007 del Observatorio de la Movilidad Metropolitana. Ministerio de Medio Ambiente y Medio Rural y Marino, y Ministerio de Fomento.

Taylor, B.D.; Miller, D.; Iseki, H.; Fink, C. (2009). Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. Transportation Research Part A-Policy and Practice, Vol. 43, No. 1, pp. 60-77.

van de Coevering, P.; Schwanen, T. (2006). Re-evaluating the impact of urban form on travel patterns in Europe and North-America. Transport Policy, Vol. 13, No. 3, pp. 229-239.

Vance, C.; Hedel, R. (2007). The impact of urban form on automobile travel: disentangling causation from correlation. Transportation, Vol. 34, No. 5, pp. 575-588.