New insights about the relation between modal split and urban density: the Lisbon Metropolitan Area case study revisited

J. de Abreu e Silva\textsuperscript{1,2} & F. Nunes da Silva\textsuperscript{1}
\textsuperscript{1}CESUR – Centre of Urban and Regional Systems, IST – Technical University of Lisbon, Portugal
\textsuperscript{2}TIS.pt Lisbon, Portugal

Abstract

Urban density is usually considered as one important factor of modal split, although its weight, compared with other variables (like car ownership, income, public transport supply, etc), is not consensual. In a previous paper the impact of urban density in modal split (defined as car share in the total number of trips made by the population of a specific urban area) in the Lisbon Metropolitan Area (LMA) was analyzed. The main conclusion was that urban density was not the most important explicative variable. From these conclusions several questions have emerged, which the present paper addresses, such as:

• What is the importance of urban density in the modal split of trips made by the residents of each zone?
• What is the importance of urban density in the modal split of trips to and from the centre of the LMA?
• What is the importance of urban density in the number of car kilometers traveled per capita?
• What is the importance of urban density in the modal split in terms of car kilometers traveled?

The developed approach uses multivariate regression analysis, considering as independent variables urban density, car availability, income and accessibility indicators. The obtained results are compared and discussed, in order to identify which aspects are conditioning—favouring, or not—the weight of density in modal split.

Keywords: transport and land use, urban density, modal split, metropolitan areas.
1 Introduction

This paper is organised in two main sections. The first one is a general overview of part of the extensive bibliography on this subject – the relation between density and modal split, namely with car use. The second part is the case study analysis.

The case study comprehends four different sections. In the first a brief geographical overview of Lisbon Metropolitan Area is made, followed by general description of the analysed survey. After an explanation of the methodology, a description of the variables construction is done.

Finally, the results are presented and some conclusions are summarised.

2 Overview

The study of the relations between land use patterns and mobility patterns started in mid eighties, and it was fuelled by the study made by Peter Newman and Jeffrey Kenworthy in 1989 [1]. One of the main conclusions was that urban density was an important variable in explaining the consumption of energy in the transport sector in big cities. This study was subjected to several critics, namely that it didn’t take into account variables like income and transport costs [2]. The study was revised later on, in order to correct for these variables. The conclusions were very much the same, the evidence of strong correlation between urban density and energy consumption [2]. Several other studies analyzing the relations between land use and mobility patterns were carried out thru the nineties. For instance, van Wee [2] refers several ones, which results are summarised in the following way: Gordon (published in 1997) concludes that density accounts for about one third of the total variation of the energy spent in transport. Anderson et al (published in 1996) revises several studies and concludes that land use patterns have a significant impact in car use. Hillbers et al (published in 1999) states that the levels of car use in new zones built within an existing city is lower than in new zones built on the edge of existing cities, all else being equal.

Another study made by Ewing and Cervero (published in 2002) [3] points out that:

- Trip frequency is mainly a function of socio-economic characteristics and secondary of the land use patterns;
- As trips lengths are concerned, the relative importance of socio-economic characteristics and land use patterns is reversed;
- Mode choice depends both of land use patterns and socio-economic characteristics, probably more of the later;
- Land use patterns are a more significant predictor of VMT (vehicle miles travelled).

The SESAME research project [4] reveals that in dense cities the modal split favours the public transport. This project also advanced that the offer of public transport was directly related with the density.
Regarding the relations between density and car share, in the SESAME Project [4] the following conclusions were drawn:

- Density is not included in the best fit model of car share;
- Density only has a strong negative correlation with the segment of drivers in car share;

Nevertheless it appears that the true nature of the relations between land use and mobility patterns is not yet fully understood. The debate evolves around the issue of causality [3]. It is possible that people who live in denser areas do it because they prefer to use the car less. In a recent study Bagley and Mokhtarian [5] found that when attitudinal and lifestyle variables are introduced the land use variables cease to have an important role in explaining mobility patterns. This supports the argument that the relation between land use and mobility patterns is due mainly to correlations between the former and attitudinal variables.

3 Case study

3.1 Geographical overview

Lisbon Metropolitan Area, comprehends 17 municipalities, involving the city of Lisbon (the capital of Portugal). The Tagus River estuary, which crosses through the region, with a width ranging from 2 Km to 12 Km, is a considerable barrier by splitting this region in two sub-regions, each one in the banks of the estuary.

In 1991 Lisbon Metropolitan Area had 2,540 inhabitants, representing 26% of Portugal’s total population. Between 1991 and 2001 the total population of this region has increased in about 140 thousand inhabitants. The demographic growth rate of Lisbon Metropolitan area (between 1991 and 2001) was 0,65% higher than the country’s growth rate. This area is also the richest region of Portugal and its most important economical driving region.

3.2 Survey

The mobility survey analysed was made seven years ago, between 1993 and 1994 [6]. The survey was made both by telephone and personal interviews. In every interview the social economical characteristics of each household member were established, and one of them was randomly chosen in order to find out his mobility patterns. The universe comprehended all the residents in the Metropolitan Area of Lisbon with an age above ten years old (age of children above elementary school). For that survey, 30680 interviews were made, characterising 101 337 persons (with an average of about 3,3 persons per household).

3.3 Methodology

An earlier study tried to find the relations between urban density and car share (% of car trips in the total number of trips). The results were not conclusive. It seemed that density had some explainable power in car use although it was not significant. So a new approach was tried. New variables to describe car use were then tested. Those were:
• Weight of car use in total kilometres travelled by the residents of each zone in all the trips which have at least one extreme in the zone (variable name \textit{PTIDIST});
• Weight of car use in total kilometres travelled by the residents of each zone between the zone and the city of Lisbon (variable name \textit{PTIDISTLX});
• Weight of car use in total kilometres travelled by the residents of each zone, independent of its origin and destination (variable name \textit{PTIDHAB});
• Number of kilometres per capita of mobile residents in each zone (variable name \textit{DISTIHAB})

The study focused on the suburban civil parishes of northern bank of Tagus River. In these civil parishes, the link to Lisbon and between them is not disturbed by the important barrier made by the river. The developed models considered only the weekday trips.

The developed methodology consisted in an exploratory line of work, in which stepwise multivariate linear regression was utilized. Every time, the density variable didn’t fit in the model, it was inserted in a first block, and in the second block the stepwise method was run, with all the other variables.

Figure 1: Suburban train lines in Lisbon Metropolitan Area.

In order to be possible to use this approach it was necessary to cover all potential aspects capable of influencing mobility patterns. So a comprehensive group of variables was built. This group included the following categories of variables: land use, socio-economical, accessibility, motorization levels, public transport supply and road supply.
The total number of considered zones (civil parishes) was 32, which were at the time served by suburban trains. The lines considered are: Cascais Line – in yellow; Sintra Line– in green; Azambuja Line – in red. Sado Line – in blue, was not considered due to the fact that the river introduces a gap between the terminal of this line and central Lisbon. Their geographical location is shown in figure 1.

3.3.1 Variables construction
In this section we will briefly discuss the variable construction, sources of information, estimation processes and possible errors.

Land use variables, this group of variables includes: global density (considering inhabitants, students and employees) transformed by logarithm – \( LNDENS\_GL \), index of compactness - \( COMP \), index of entropy (which measures the mix of land uses) - \( ENTROP \), mix of functions (percentage of students and employees in the global number of people) - \( MIX \), distance from the CBD - \( DISTCBD \) and percentage of each zone’s area occupied by urban functions - \( PAURB \).

Socio-economical variables, which includes: percentage of residents not able to read - \( TXANALF \), percentage of residents with college degrees - \( TXUNIV \), percentage of residents working in agriculture – \( TX\_SECT\_I \), percentage of residents working in the service sector – \( TX\_SECT\_III \), index of power of purchase - \( IPC \), percentage of employed residents working in Lisbon – \( DEP\_EMP\_LX \), percentage of families with 0 – \( PFAM0EMP \), 1 – \( PFAM1EMP \), 2 – \( PFAM2EMP \) or more than 2 employed members – \( PFAMM2EMP \), percentage of families with 0 – \( PFAM0LIB \), 1 - \( PFAM1LIB \) or more than one liberal professionals – \( PFAMM1LIB \) and percentage of families with children below ten years old - \( PFAMCRIANC \).

Motorization variables, which included: the percentage of families with more than one car – \( FAM>1TI \) and the number of cars per 1000 inhabitants - \( TXMOT \).

Public transport supply variables, this group includes: variables like percentage of the global population (residents, employees and students) within a radius of 400 m of distance from each bus stop – \( PPGLOB400BUS \) and the same type of variable for each rail station \( PPGLOB400TCSP \).

Road supply variables, which includes the capitation of road kilometres (considering residents, employees and students) – \( KMVIA/PESSGL \) and the percentage of people within a radius of 1000 metres form each freeway junction – \( PPESSGL1000 \).

Accessibility variables, in this class of variables only one variable was considered, which was the logarithm of the quotient between the public and car accessibility index – \( LNACTC\_TI \).

The accessibility indexes were built using a gravitational approach method, which can be represented by the following expression:

\[
A_i = \sum_{j=1}^{n} D_j \cdot f(c_{ij})
\]
where:

$A_i$ is the accessibility index do zone $i$

$D_j$ is the number of opportunities (residents, employees and students) of trips within a 400 m radius of the destination coordinates;

$f(c_{ij})$ is a function of cost (considering the travel time) calibrated with the results from the mobility survey.

Several variables were built considering the survey results, while others were built using 1991 census data. As stated before, the survey was made between 1993 and 1994, so the population coefficients utilized were the ones derived from the 1991 national census.

### 3.4 Results

For the model considering the variable $PTIDIST$, the results obtained by the linear regression are presented in the next table.

The first model tested, didn’t include the variable $LNDENS\_GL$, so in a second model (presented above) that variable was included in a first block. The results show that this variable is not significant at a 95% level. The significance level of $LNDENS\_GL$ decreases abruptly when the variable of accessibility is included in the model. That might imply a resemblance of effects. It is expected that with an increase in urban density at least the levels of public transport accessibility should also increase (the correlation between the indicator of accessibility in public transport and $LNDENS\_GL$ is high, about 0.68).

All the other variables included in the model have the expected signals.

#### Table 1: Linear regression model summary - dependent variable: $PTIDIST$.

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R$ Square</th>
<th>Adjusted $R$ Square</th>
<th>Std. Error of the Estimate</th>
<th>Change Statistics</th>
<th>df1</th>
<th>df2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5448</td>
<td>0.2969</td>
<td>0.2734</td>
<td>0.0933</td>
<td>0.2969</td>
<td>12.6655</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.6770</td>
<td>0.4584</td>
<td>0.4219</td>
<td>0.0833</td>
<td>0.1615</td>
<td>6.6496</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.8318</td>
<td>0.6918</td>
<td>0.6588</td>
<td>0.0639</td>
<td>0.2334</td>
<td>21.2104</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.8624</td>
<td>0.7437</td>
<td>0.7057</td>
<td>0.0594</td>
<td>0.0519</td>
<td>5.4647</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.8943</td>
<td>0.7998</td>
<td>0.7613</td>
<td>0.0535</td>
<td>0.0561</td>
<td>7.2902</td>
<td>1</td>
</tr>
</tbody>
</table>

1 - Predictors: (Constant), $LNDENS\_GL$
2 - Predictors: (Constant), $LNDENS\_GL$, $TX\_SECT\_III$
3 - Predictors: (Constant), $LNDENS\_GL$, $TX\_SECT\_III$, $KMVIA/PESSGL$
4 - Predictors: (Constant), $LNDENS\_GL$, $TX\_SECT\_III$, $KMVIA/PESSGL$, $LNACTC\_TI$
5 - Predictors: (Constant), $LNDENS\_GL$, $TX\_SECT\_III$, $KMVIA/PESSGL$, $LNACTC\_TI$, $PFAM1EMP$

The results from the second variable tested, $PTIDISTLX$, are presented in Table 2. In this case, density was not included in the first stepwise model, so it
was inserted in a separated block the regression model. In his case the variable density is significant, and its effect is the expected one. The other variables included in the model reflect both the presence of activities to which the private transport is more useful (agriculture) and also the levels of income. Usually people working in the service sector tend to have higher income levels. The effect of the last variable is contrary to what was expected. Normally families with children tend to use more the car (for instance to take their children to school).

Table 2: Linear regression model summary - dependent variable: PTIDISTLX.

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
<th>Change Statistics</th>
<th>df1</th>
<th>df2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4881</td>
<td>0.2382</td>
<td>0.2128</td>
<td>0.1300</td>
<td>0.2382</td>
<td>9.3819</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.7047</td>
<td>0.4966</td>
<td>0.4619</td>
<td>0.1075</td>
<td>0.2584</td>
<td>14.8886</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.7789</td>
<td>0.6067</td>
<td>0.5646</td>
<td>0.0967</td>
<td>0.1101</td>
<td>7.8347</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.8223</td>
<td>0.6761</td>
<td>0.6281</td>
<td>0.0894</td>
<td>0.0694</td>
<td>5.7864</td>
<td>1</td>
</tr>
</tbody>
</table>

In the following model (dependent variable PTIDHAB) the density variable enters when the capitation of roads is taken from the model. That fact is not surprisingly, because those two variables are highly correlated (the correlation coefficient is -0.756). The other selected variables, generally have the expected influence, with the exception of PFAMCRIANC and IPC. In the first of these two variables, the expected coefficient should be positive, due to the reasons explained earlier. In the second case, the IPC, which is a variable that describes income, should have a positive effect on the levels of car use. Although this variable is strongly correlated with another variable included in the model, which can also act as a proxy of income, TX_SECTIII (correlation coefficient 0.793). It is also positively correlated with the levels of motorization, although not in such strong way (correlation coefficient 0.42). So it is possible that the negative coefficient of IPC could be only in order to adjust the effects of both other referred variables. In the following table the model results are presented.

The last of the studied models, has as dependent variable the total average travelled distance in car by mobile resident (DISTTIHAB). In this model the variable density (LNDENS_GL) entered directly in the first tested stepwise regression. Although, due to collinearity problems, some variables had to be rejected, and in the final, and selected model, LNDENS_GL was inserted in a first block, with a second block using a stepwise method.
# Table 3: Linear regression model summary - dependent variable: PTIDHAB.

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
<th>Change Statistics</th>
<th>df1</th>
<th>df2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R Square Change</td>
<td>F Change</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.6534</td>
<td>0.4269</td>
<td>0.4078</td>
<td>0.0837</td>
<td>0.4269</td>
<td>22.3444</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.7674</td>
<td>0.5888</td>
<td>0.5605</td>
<td>0.0721</td>
<td>0.1620</td>
<td>11.4238</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.8047</td>
<td>0.6475</td>
<td>0.6097</td>
<td>0.0679</td>
<td>0.0586</td>
<td>4.6566</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.8268</td>
<td>0.6835</td>
<td>0.6366</td>
<td>0.0655</td>
<td>0.0361</td>
<td>3.0757</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.8466</td>
<td>0.7167</td>
<td>0.6622</td>
<td>0.0632</td>
<td>0.0332</td>
<td>3.0428</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0.8738</td>
<td>0.7635</td>
<td>0.7068</td>
<td>0.0589</td>
<td>0.0469</td>
<td>4.9557</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0.9024</td>
<td>0.8143</td>
<td>0.7601</td>
<td>0.0533</td>
<td>0.0507</td>
<td>6.5552</td>
<td>1</td>
</tr>
</tbody>
</table>

1 - Predictors: (Constant), FAM>1TI
2 - Predictors: (Constant), FAM>1TI, TX_SECT_I
3 - Predictors: (Constant), FAM>1TI, TX_SECT_I, PFAMCRIANC
4 - Predictors: (Constant), FAM>1TI, TX_SECT_I, PFAMCRIANC, TX_SECT_III
5 - Predictors: (Constant), FAM>1TI, TX_SECT_I, PFAMCRIANC, TX_SECT_III, LNDENS_GL
6 - Predictors: (Constant), FAM>1TI, TX_SECT_I, PFAMCRIANC, TX_SECT_III, LNDENS_GL, PPASSE
7 - Predictors: (Constant), FAM>1TI, TX_SECT_I, PFAMCRIANC, TX_SECT_III, LNDENS_GL, PPASSE, IPC

In this model three of the selected variables are related with land use patterns. Besides LNDENS_GL, LNACTC_TI and DISTCBD (the distance between each zone and the Central Business District) are included. The first of these is also linked with land use patterns. In denser areas, usually more people live or work near transit facilities, and also the supply levels of public transport tend to be higher in those areas. The distance from CBD is usually considered as an indicator of car use. The levels of car use tend to increase when the distance from CBD increases.

## 4 Conclusions

The first conclusion that can be drawn is that there are at least some important levels of correlation between land use patterns and levels of car use in the suburban region of Lisbon Metropolitan Area. Of all the considered variables built to describe land use patterns, urban density appears to be the most important variable that can explain car use levels. In only one of the tested models (dependent variable PTIDIST) urban density has been revealed as non-significant at 95% level. Nevertheless, urban density wasn’t usually selected in the first stepwise models tested. It had to be inserted in a previous block of the regression models.
Table 4: Linear regression model summary - dependent variable: DISTTIHAB.

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
<th>Change Statistics</th>
<th>df1</th>
<th>df2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5756</td>
<td>0.3313</td>
<td>0.3090</td>
<td>2.8276</td>
<td>0.3313</td>
<td>14.8643</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.7435</td>
<td>0.5528</td>
<td>0.5219</td>
<td>2.3519</td>
<td>0.2215</td>
<td>14.3619</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.7991</td>
<td>0.6227</td>
<td>0.5623</td>
<td>2.1965</td>
<td>0.0699</td>
<td>5.1872</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.8771</td>
<td>0.7690</td>
<td>0.7352</td>
<td>1.7506</td>
<td>0.1466</td>
<td>17.1637</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.8946</td>
<td>0.8004</td>
<td>0.7620</td>
<td>1.6595</td>
<td>0.0311</td>
<td>4.0457</td>
<td>1</td>
</tr>
</tbody>
</table>

1 - Predictors: (Constant), LNDENS_GL
2 - Predictors: (Constant), LNDENS_GL, LNACTC_TI
3 - Predictors: (Constant), LNDENS_GL, LNACTC_TI, DISTCBD
4 - Predictors: (Constant), LNDENS_GL, LNACTC_TI, DISTCBD, TX_SECT_III
5 - Predictors: (Constant), LNDENS_GL, LNACTC_TI, DISTCBD, TX_SECT_III, PFAMCRIANC

The other territorial variables included in the tested models were LNACTC_TI (the logarithm of the quotient between the measure of accessibility in public transport and in car) and DISTCBD (the distance from CBD). The first of these two variables is, as explained before, linked with land use patterns, particularly with urban density. The second variable is also correlated with land use patterns and with urban density. Usually the density decreases with the distance from the centre of a metropolitan area.

Other important finding is that the relation between urban density and car use is not linear. So the absolute growth of the elasticity of car use regarding density tends to decrease as density increases.

Finally, although the built models included socio-economical variables, they didn’t include attitudinal variables. Also the models built were defined at an aggregated level. So it’s not possible, with this set of data, to verify without any doubt the true nature of the relations between land use and levels of car use. Are they of causal type, or are mainly due to correlations between land use variables and attitudinal variables?

Nevertheless the results obtained are in accordance with several other studies that stated the importance of urban density as one of the variables that should be considered (and not the only one) in order to explain and to decrease the levels of car use. And also they are in accordance with the theoretical mechanisms built to describe the relations between land use patterns and mobility patterns.

References


