From home to work and vice versa; analyzing the footprint of commuting to work in the greater area of Athens

Abstract

Commuting to work constitutes an important procedure, taking place in the contemporary cities. The present paper, in the basis of the strong relation between commuting and the factors defining the urban fabric, investigates the above relationship in the Attica region. The research focuses on detecting both the geography of commuting flows (Average Commuting Distance) of the municipalities and its correlation with other characteristics (geographic, social, urban and transport). The data were drawn from Hellenic Statistical Authority (ELSTAT), Openstreetmap, Athens Regional Plan, Athens Urban Transport Organisation S.A. (OASA S.A.) and were consequently analyzed in R, Ms Excel and Geographic Information Software (GIS). The analysis procedure includes mapping of the variables, spatial autocorrelation of the dependent variable and multiple linear-regression between the Average Commuting Distance and a set of other characteristics. The findings show that Land Use Mix influences negatively the Average Commuting Distance whilst the Car Property and the Municipality’s Distance from the main metropolitan center influence it positively. The conclusions contribute to a better understanding of the spatial expression of commuting and sustain the urban and regional policymaking. Finally, the methodological steps can be applied to other regions leading to new spatial «revelations».

Key words: Commuting, spatial autocorrelation, multiple linear regression, Attica, GIS.

Introduction

Cities are considered as the appropriate environment where a variety of different activities happens. Undoubtedly, one of the most important and influential of these activities seems to be work. Working as a repeating and mainly an obligatory process shapes and affects in a significant way residents’ everyday life and urban geography in general. Specifically, the examination of the relation between the working place and the place of residence contributes seriously to the understanding of cities functionality and sustainability. This phenomenon is referred as commuting to work. The present paper, in the basis of the strong connection between commuting and the factors defining the urban fabric, investigates the above relationship in the Attica region. The research focuses on detecting both the geography of commuting flows (commuting distance) of the mainland municipalities and its correlation with other characteristics (geographic, social, urban and transport). Furthermore, this investigation contributes to the definition of the proper policies regarding urban and transport issues about the procedure of commuting to work. Therefore, our goal is to arrive at conclusions assigned to the “volume” of commuting distance (to work) performed by the residents of each municipality, the factors that affect this mobility and finally the existence of high or low value spatial clusters.
Regarding the structure, the paper is divided into five chapters, which they are then sub-divided into individual entities. The second chapter contains all the theoretical background of the research. In the third chapter, are described the various characteristics of the study area, the dataset used in the research analysis and the methodology steps that we followed. The next chapter contains the analysis in total. Mapping of commuting flows, spatial autocorrelation and regression analysis regarding the commuting distance and a set of other features are the main factors of the fourth chapter. In the last chapter, we present the conclusions of the above-mentioned analysis as well as some proposals beyond the findings.

**Theoretical background**

The scope of this chapter is to the present the terminology and the description of the features referred in the research. It deals with issues like urban sprawl, land use and transport and commuting to work.

Urban sprawl is a multilayered phenomenon that describes the spread of urban fabric, its functions and suburbs to the periurban area (Dimopoulou et al., 2015). There is a plethora of definitions for this phenomenon. The dominant definition links urban sprawl to urban sprawl in an uncontrolled, scattered way. Specifically, in the Oxford English Dictionary (2001) it is defined as “A large area covered with building, which spreads from the city into the countryside in an ugly way”. In addition, urban sprawl does not refer only to the expansion of the urban fabric into a broader area, but it is also linked to population growth (Barnes et al. 2001). Despite the idyllic promises of suburbanization, these newly created areas present a series of problems, such as long-term work trips, great dependence on car use and inadequate social infrastructure (Cervero & Kockelman, 1997; Newman & Kenworthy, 1989; Dimopoulou et al, 2015). The explanation of urban sprawl and the identification of the factors that affect it, is a rather difficult process (Johnson, 2001) because cause and effect are involved in a procedure of mutual feeding, alternating the role of the other.

The relationship between land use and transport has always been one of the most important subjects that researchers associated to spatial issues have to deal with (Cervero, 1988, 1991; Frank & Pivo, 1994; Milakis & Vlastos, 2007). The issues of spatial organization and planning are becoming more and more complex as the reference area gets wider. Taking into account both the factor of social costs (traffic congestion, environmental burden) and purely individual consequences (cost of economic and time travel), many researchers believe in the impact of land uses to reduce time and length of journey (Cervero, 1989; Levine, 1992). In addition, Cervero (1996) highlights the importance of the characteristics of a site on the flows produced. This implies policies to safeguard the balance between jobs and housing at the local level.

Commuting to work is a research issue in both urban and regional development as it is crucial to understanding the facts that take place in a city and the urban economy. “The connection between home and workplace has been, and remains, a central part of theories of urban spatial structure” (Clark et al., 2003). Numerous researchers from different fields of science have attempted or attempt to interpret commuting patterns both theoretically and quantitatively. Thus, the investigation of the patterns of mobility is not a simple matter, because it consists of a complex system of moving people - units, motivated by a set of their own motives and characteristics (Szalai, 1972). For anyone studying such phenomena, there are two options: either to focus on exploring one of the features linked to commuting
(Soltani, 2005; Schwanen et al., 2004), or to treat the phenomenon as multidimensional using the appropriate tools for the corresponding quantitative methods (Giuliano & Small, 1993). In these researches, the initial data are presented based on some general characteristics and according to them they are being classified. Thereafter, these data refer to a large number of household variables by explaining the phenomena being studied (Koppelman & Pas, 1985). The methodologies developed in these researches are extremely useful for understanding and representing the variables that make up the phenomenon, but they are acknowledged to have their limitations and shortcomings.

The table below summarizes the goals of other researches carried out on relevant subjects:

<table>
<thead>
<tr>
<th>Citation</th>
<th>Aim of research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black, 1997</td>
<td>Environment: the longer cars are on roads, the more emissions they expend, and</td>
</tr>
<tr>
<td></td>
<td>additional fuel is wasted.</td>
</tr>
<tr>
<td>Green, Hogarth &amp; Shackleton, 1999</td>
<td>Social concerns: long commutes as practice require additional time away from</td>
</tr>
<tr>
<td></td>
<td>family.</td>
</tr>
<tr>
<td>Koslowsky, Aizer &amp; Krausz, 1996</td>
<td>Social concerns: place additional stress on individuals and limit their ability</td>
</tr>
<tr>
<td></td>
<td>to participate in desired activities outside of work.</td>
</tr>
<tr>
<td>Sultana, 2002</td>
<td>Land uses: impacts of the density and mix of land use on commuting outcomes.</td>
</tr>
<tr>
<td>Giuliano &amp; Small, 1993</td>
<td>Job–housing balance: describe the relative locations of jobs with respect to</td>
</tr>
<tr>
<td></td>
<td>housing in a given area.</td>
</tr>
<tr>
<td>Peng, 1997</td>
<td>Job–housing balance: empirical research on the relationship between commuting</td>
</tr>
<tr>
<td></td>
<td>and jobs–housing balance in Portland.</td>
</tr>
<tr>
<td>Cervero, 1989</td>
<td>Job–housing balance: proving that a well-designed urban area with integrated</td>
</tr>
<tr>
<td></td>
<td>residential and workplace locations (i.e., a balance between jobs and housing)</td>
</tr>
<tr>
<td></td>
<td>should promote less commuting.</td>
</tr>
<tr>
<td></td>
<td>commute lengths at the metropolitan scale are correlated with the level of</td>
</tr>
<tr>
<td></td>
<td>internal jobs–housing balance.</td>
</tr>
</tbody>
</table>

**Data and methods**

The methodological procedure of the present paper is divided into five steps, which are described in the following flow chart.

![Methodological steps](Figure 1: Methodological steps)
The present paper focuses on the commuting distance of the mainland part of the Attica region. This specific part is divided into four regional entities, the first two and the most significant are Athens (Central, North, South and West sectors) and Piraeus entity respectively, the third one is East Attica and the last one is East Attica. Moreover, the study area consists of 59 municipalities and its whole geographic area equals to 2932 km² (77% of Attica region). Regarding its population, mainland Attica has 3.753.783 residents, constituting a percentage of 98, 05% of the whole region. The entity with the biggest population is Athens entity that has 3090508 residents and the fewest Piraeus with no more that 490.000 (448.997) residents. The other ones, West and East Attica have 592490 and 502348 residents respectively. Furthermore, in the study area, are located 1.483.516 households, with mean size 2,53 residents per household. Also, in mainland Attica live 1.744.316 economically active residents, with an amount of 1.430.554 residents being on a working state and the rest being unemployed (18%). In the matter of employment, the vast majority of the employees (82%) work in the third sector (services), 17% in the second sector (industry, handicraft) and only the 1% of the employees of Attica works in the first sector activities.

Regarding the urban features, the “mixed residence” is a dominant element of the central areas. This fact contributes positively to land use mix and subsequently enhances the public transport use as well as walking. On the contrary, suburban, outer urban and mainly north and south areas of Athens are at “net residence” and thus residents tend to prefer the car use. Despite the size of the study area, Attica functions as standard single-core city and not as an organized polycentric metropolis (Milakis, 2006). Municipality of Athens undoubtedly constitutes the center of the region in general as it provides an important amount of jobs, most of them, belonging to the third sector of production. Therefore, the majority of the population is forced to commute to and from this center.

We used a set of various data that concern the house residence and the work place of the Athenian employees as well as other variables relevant with the character of the study area. The choice of these parameters is based on:

- Other similar researches
- Availability of associated data
- Their expected utility in the greek region

The spatial analysis level is that of the municipality. Similar researches work with the same spatial level (Cervero, 1989, 1996, Milakis & Vlastos, 2007) and therefore we chose it as though. Moreover, using a wide area of analysis leads to deterioration of the results and incapability of proper findings (Handy, 1996; Badoe & Miller, 2000). Specifically, the home and work place data, as well as the rest social and some urban elements were extracted from the database of the Elstat (2011). Moreover, the transportation data were drawn from Oasa (public transport) and open street map (road network). As far as the cartographic materials are concerned; we also used ELSTAT data in combination with features from the geodata.gov online platform and the NTUA. The environment of ArcGis and QGis, Microsoft Excel and the statistical package of R were used for the processing of the above variables and features. Subsequently, after the collection of the aforementioned data, we proceeded to their categorization in different subgroups (geographical, urban, social, transport) in order to ensure that we have covered the majority of the aspects that affect the Average Commuting Distance. These categories are described in the next pages.

In order to compute the value of average commuting distance (ACD) for each municipality, we applied the following steps. Firstly, we identified the centroid coordinates of each
municipality polygon and subsequently for each of the pairs we calculated the Euclidean distance. This procedure though, does not take into consideration the inbound commuting flows. Therefore, in this case we set the Euclidean distance as the approximate radius of the municipality polygon. The complete formula of the ACD is defined as follows:

$$ACD_i = \frac{\sum_{j=1}^{n} d_{ij} \times CW_{ij}}{\sum_{j=1}^{n} CW_{ij}}$$

Where, ACD constitutes the average commuting distance to work, CW is the commuting employees and d is the Euclidean distance between the polygons’ centroids. The measurement unit is kilometers (km).

In order to achieve a better outline of the features and their formulas, we created a table summarizing all the above. This table is shown directly below.

Table 2: Variables of the research

<table>
<thead>
<tr>
<th>m/n</th>
<th>Category</th>
<th>Symbol</th>
<th>Features</th>
<th>Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Geographical</td>
<td>Distance</td>
<td>Distance from center municipality (km)</td>
<td>$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residential</td>
<td>Residential density (res/ha)</td>
<td>$$\theta_{pop} = \frac{N_{residents}}{A_{built}}$$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land use mix</td>
<td>Land use mix (Entropy index)</td>
<td>$$E = \sum_{i} p_i \cdot \ln \left( \frac{p_i}{\theta_{ij}} \right)$$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range and</td>
<td>Range and significance</td>
<td>1st category: Metropolitan, 2nd category: Inter-municipal, 3rd category: Local</td>
</tr>
<tr>
<td></td>
<td></td>
<td>significance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Urban</td>
<td>Car property</td>
<td>Car Property rate (%)</td>
<td>$$\theta_{CPR} = \frac{N_{households \times cars}}{N_{households}}$$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unemployment</td>
<td>Unemployment rate (%)</td>
<td>$$\theta_{UR} = \frac{N_{unemployed \times residents}}{N_{economically \ active \ residents}}$$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Men employed</td>
<td>Men employed rate (%)</td>
<td>$$\theta_{MER} = \frac{N_{employed \times residents}}{N_{economically \ active \ residents}}$$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Upper-class</td>
<td>Upper-class citizens rate (%)</td>
<td>$$\theta_{UCCR} = \frac{N_{upper-class \ citizens}}{N_{economically \ active \ residents}}$$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Public</td>
<td>Urban public transportation area</td>
<td>True or False</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transportation</td>
<td>Public Transportation Accessibility (Bus/Stop/Week)</td>
<td>$$\theta_{PTA} = \frac{N_{bus/stop/week}}{N_{bus/stop}}$$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Road network</td>
<td>Road network service (m/res)</td>
<td>$$\theta_{RNS} = \frac{N_{basic \ road \ network}}{N_{residents}}$$</td>
</tr>
</tbody>
</table>

The basic methods implicated in the present research are spatial autocorrelation and multiple linear regression analysis. According to Griffith (2003), spatial autocorrelation is the correlation between the values of a variable, which derives strictly from the proximity of
these values in the geographic space. This approach is contradicting the general assumption of independent observation in the domain of classic statistics. Spatial autocorrelation can be used to describe and compare the spatial structure of the variable (Overmars et al, 2003). Regarding multiple linear regression analysis, this particular method allows to identify and estimate the main variables that influence the average commuting distance. Over three decades it has major role in researches associated with urban, transportation and geography in general (Hecht, 1974; Cevero, 1989; Levinson, 1997; Muniz & Galindo, 2005).

Findings

This chapter contains the mapping of the average commuting distance, a few descriptive analysis measures and the creation of a multiple linear regression model.

The average commuting distance to work of every municipality in the study area is calculated by the first formula and provides useful findings about the relationship between home and work place. These findings can be exploited in various ways (urban planning, transportation planning, environmental issues etc.). The geographical image of each municipality is illustrated at the next map.

Through the observation of the above map, we can identify a unique geography of the commuting flows in the mainland Attica region. This geography could be characterized as “geography of contrast”, due to its discrete difference that occurs in the study area between central and suburban or outer-urban areas. This given situation illuminates the crucial need of the latter’s residents for travelling many kilometers in order to reach their work place. In addition, these trips probably promote car use, thus leading to negative impacts on the sociality, the urban and natural environment and the transportation system of the
municipalities as well as of the wider area in general. Moreover, we distinguish that among the municipalities, greater distances from the center signify higher values of average commuting distance. However, in the cases of certain distant municipalities (Salamina, Eleusis, Aspropyrgos) differ from this pattern probably because of the opportunity of local employment facilities. The municipality of Athens has the smallest average commuting distance (5.58 km) while the municipality of Saronicos has the biggest, with corresponding value of 18.51 km. The average commuting distance of the sample extracted from the study area is 9.5 km while the standard error that shows the deviation from the actual population mean is 0.37 km.

Besides the simple description of the results, in order to achieve a more sophisticated approach of the problem, we conducted spatial autocorrelation method (Moran’s Index). The spatial autocorrelation as it was mentioned before contributes to the identification of potential spatial neighborhoods in the study area. The next map depicts the results of the procedure.

The spatial autocorrelation computes results that agree with the findings stated in the previous paragraph. Precisely, we identify two discrete neighborhood categories. The first neighborhood is located at the north and south part of Attica outer-urban area and forms a high-high cluster. This means that these specific areas are characterized by high values of average travelling distance to work and they should be given special attention by the planning procedure. The proper management of the commuting flows in these particular areas is considered a difficult but essential task in order to become more livable, sustainable and functional. On the contrary, the second spatial cluster or neighborhood was identified in the central core of mainland Attica region and contains significant municipalities with metropolitan range like Athens, Piraeus and other important areas (e.g. Peristeri, Kallithea, Vyronas etc). This neighborhood forms a low-low cluster, which represents low values of average travelling distance to work. It is assumed that the municipalities belonging to the second neighborhood, because of their characteristics, prefer mainly to use public

Map 2: Spatial autocorrelation of Commuting Distance (Moran’s)
transportation and walking. The transportation by car option still exists, however is not the dominant one. A further research will surely enlighten this assumption. Nevertheless, despite of the means of transport being used, the low values of commuting distance by itself mean that residents are not obligated to spend a considerable amount of their time just to access their workplace. At last, it has to be mentioned that the z-score of Moran’s Index equals to 9.16, signifying that there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

In order to examine which parameters affect the Commuting Distance regarding the municipalities of the Attica Region, we run a multiple regression model using as dependent variable the logarithm of the Average Commuting Distance and as independent variables those described in the previous section. The estimation of the coefficients derived from the application of ordinary least square method (OLS) in R statistical language, a standard approach in the regression analysis. The results of our analysis are depicted in the table below:

Table 3: Regression model for predicting average commuting distance, simultaneous model estimate

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-Statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_1)</td>
<td>-0.525</td>
<td>0.163</td>
<td>-3.218</td>
<td>0.00229</td>
</tr>
<tr>
<td>(X_2)</td>
<td>0.835</td>
<td>0.101</td>
<td>8.293</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>(X_3)</td>
<td>0.000023</td>
<td>1.4 \times 10^{-6}</td>
<td>17.288</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>Constant</td>
<td>2.217</td>
<td>0.009</td>
<td>248.579</td>
<td>&lt;0.00001</td>
</tr>
</tbody>
</table>

Summary Statistics:

\[R^2 = 0.93\]

Residual Standard error: 0.06

\[F = 248.1, \text{ probability}<0.001\]

No of cases = 58

Variable Definitions

\[\log(\text{Average Commuting Distance}) = \text{is the logarithm of the Average Commuting Distance in km.}\]

\(X_1\) = The variable land use mix described before, subtracted by its expected value.

\(X_2\) = The variable car property subtracted by its expected value.

\(X_3\) = The variable distance from the city center in km, subtracted by its expected value.

The above Table summarizes the stepwise results from estimating the logarithm of the variable Average Commuting Distance that refers to the municipalities of mainland Attica region. At this point we should note that the upper multiple regression model meets all the assumptions of normality, linearity, homoscedasticity, no-multicollinearity and no-autocorrelation of its residuals. This fact in combination with the high value of the multiple \(R^2\) as well as the small residual standard error makes the model suitable for inference and prediction. Three variables that can be characterized as urban (land use mixed), social (car property) and geographical (distance from the city center), entered the model. The reason that we subtracted the expected values from each one of them (centered variables) is to impart a natural and more realistic interpretation to the model. Therefore, as is apparent...
from the estimations that have been retained, the results reveal as follows. The expected value of the dependent variable is equal to 2.22 km when the variables Land Use Mix, Car Property and Distance from the Center are equal to their expected mean values. Otherwise, the geometric mean of the Average Commuting Distance when the variables Land Use Mix, Car Property and Distance from the Center are equal to their expected mean values is equal to 9.2 km. Moreover, we observe that the increase in Land Use Mix measures means the reducing of the Commuting Distance. Holding the centered variables $X_2$ and $X_3$ constant and increasing the variable $X_1$ by one unit the dependent variable reduces by 0.52 km. More clearly, the Average Commuting Distance reduces by 40.5% for one unit increase in $X_1$ while holding the other predictors constant. On the contrary, an increase in Car Property and respectively in the Distance from the city center, leads to the growth of the kilometers in the Average Commuting Distance. Specifically, the Average Commuting Distance increases by 131.6 % for a one unit increase in $X_2$ while holding the other predictors constant while same results arise regarding the centered variable $X_3$. In sum, in regions with high land use mix rate the mobility to work is significantly less than others with lower rates while the factor of the car property as well as the distance from the center make an important contribution to the augmentation of commuting distance.

Conclusion

The present research contributes significantly to the discovery of critical issues about the relationship between home and workplace. The main findings concern the geography of commuting flows and the factors affecting it, in the mainland Attica region. It was found that central areas function completely differently in the matter of commuting than suburban or outer-urban ones. The former are characterized by low values of average commuting distance and on the contrary, the latter seem to form high values clusters. This geography could be named as “geography of contrast” signifying the functional and social diversity inside the Attica region. This identified geography is undoubtedly affected by many factors of the city. The factors that were taken into consideration in this particular study were geographical, urban, social and transportation. This choice was made in order to account the multiple dimensions of the city. Through the regression analysis, it was discovered that the factors with greater influence to the average commuting distance were Land Use Mix, Car Property and Municipality’s Distance from the main metropolitan center. Precisely, Land Use Mix influences negatively the Average Commuting Distance whilst the Car Property and the Municipality’s Distance influence it positively.

A study related to commuting patterns has certainly major significance in contemporary urban and regional planning. Commuting consumes notable proportion in the day of an employee, thus affecting strongly his/her daily life habits and behavior. This situation probably leads to free time limitation and therefore social isolation. Furthermore, commuting geography affects crucially the city’s functionality in general. Areas with low values of average commuting distance have greater chances of livability and social interaction; while high values spatial clusters constitute mainly “urban dormitories”. Also, a city where its residents travel great distances far from their home and at the same time adequate public transport services do not exist, would probably encourage car use. The excessive use of private vehicles can cause many negative impacts on the urban environment (air pollution, global warming, anti-social public space, low quality of life etc).
The findings of the research contribute strongly to the mapping and understanding of commuting geography of the city. The discovery of the most significant affecting factors is a crucial spatial as well as social revelation, highlighting the basic roots of the problem. Therefore, the findings should be used in a planning procedure, which takes of course into consideration both the urban and the transportation dimension of the city. Such a planning strategy would be fruitful for the making of sustainable and functional cities.

Regarding the methodological process, the selection of the specific study area, which constitutes the most significant region of Greece with complex characteristics, strengthens the research findings. A study dedicated to a region of great range, can produce various conclusions concerning today’s urban and regional planning issues. It is also a fact, that the number of similar researches in Greece and Athens in particular is limited, so this research represents a notable addition. Trough the elaboration of the research, data collection was a matter of high difficulty. The access constrictions or the absence of essential data undoubtedly affected the research path that we followed. As a result, this study was based mainly on previous literature in combination with the available spatial or other data. Finally, the methods of analysis used (spatial autocorrelation, regression analysis) are not the only one suit for this scientific field and it’s strongly encouraged new methods to be used in similar future researches (e.g. Geographically Weighted Regression).

The research objective as well as the above-mentioned absence of a larger variety of data, makes the present study as a “launchpad” for further research. Specifically, new researches should investigate commuting either in a different level of analysis or correlating it with extra variables. Also, the findings can be used as an input in other studies of similar or other scientific field. Finally, it is recommended the methodological steps used should be applied to other study areas, in order to broaden the results about commuting.

References


