Factors determining Commuting Intensity; an Empirical Analysis in Greece

Maria Stefanouli
Department of Planning and Regional Development
University of Thessaly
mstefanouli@gmail.com

Serafeim Polyzos
Department of Planning and Regional Development
University of Thessaly
spolyzos@uth.gr

Abstract
Commuting –defined as the daily travelling for employment purposes- has gradually grown in importance as in the past decades more and more employees commute daily towards their workplace and naturally this renders it an essential element of any sustainable policy. Greece's case is a peculiar one as there are noticeable differences from one region to another, due to variations in the basic organisation and function of spatial units, related to economic, demographic, social and geographic factors. Given this, this paper aims to analyse the commuting intensity at relatively fine unit scales (Local Administrative Unit - LAU1), and to propose – through its use - a method of evaluation of both out and in-commuting intensity through indicators suited to this spatial scale. Factors used are of different nature in order to present a holistic view of the commuting intensity. For this purpose, a multi-regressive methodological framework is constructed for the analysis of commuting intensity, applying the multinomial logistic regression model. The conclusions reached come in general agreement with the recent literature, whilst are of special interest for Greece as commuting in this country has not been studied yet extensively.

Keywords: Commuting, Out-commuting intensity, In-commuting intensity, Multinomial Logistic Regression, Sustainable policy

JEL classification: C35, R40, R49

Introduction

What is most definitely a complex matter is the development of urban areas in a sustainable manner with the aim of attracting eventually new residents and new employees but not losing their distinctive features at the same time. As far as Wong (1995) is concerned, the basic principles of urban and regional planning sustainable development policies can be distinguished into quantifying needs and opportunities of each region with regard to resource distribution, determining all the necessary conditions for improving an area determining political goals having first identified the opportunities and the problems of each region.

Consequently, it is obvious that commuting flows between municipalities, as a factor of great importance, should be taken into consideration for sustainable development policies. The term commuting is defined as the daily travelling for work beyond the administrative unit (Polyzos, 2015; Stefanouli & Polyzos, 2015; Tsiotas & Polyzos, 2013; Van der Laan & Schalke, 2001).

Until the 19th century, employees' first priority in terms of commuting was to live near their workplace so they spend less time commuting (Polyzos et al., 2014). As the concept of a polycentric city emerged, the more urban cores and the extensive use of road networks altered this prioritisation (Axisa et al., 2012; Clark et al., 2003; Harris et al., 2008). Nowadays, commuting –a daily act of a significant part of employees- is integral to the daily life of most and has acquired multivariate characteristics especially after the technological evolution (Polyzos, 2015; Polyzos et al., 2014).

Against this background, commuting is shaped by economic, social and geopolitical factors and this renders it a multivariate phenomenon. Its draws influences from various subject areas, such as Regional Economics, Sociology, Human Geography, Economic Geography etc and naturally is based on many factors, such as the travel cost, the mode of commuting, the commuting distance etc. (Polyzos et al., 2013; Tsiotas & Polyzos, 2013).

According to Nielsen & Hovgesen (2008), the commuting maps shed light on the human behaviour factor, which is influenced and shaped according to desires, trends, wealth and mobility, which is something Crane (1996, in Clark et al., 2003) also concurs with as to him expectations for future employment and housing opportunities shape the link between home and workplace. It stands to reason then that the key elements of this need to be understood in order to shape an effective and sustainable planning and policy (Polyzos et al., 2013; Tsiotas & Polyzos, 2013).

Commuting Intensity

A wide range of various variables (such as commuters' social and demographic features which have also been proven important (Susilo & Maat, 2007)) have been used -some successfully and some not so much-to better explore and understand commuting behaviour (Axisa et al., 2012).

There is a growing body of research concerned with commuting distance, commuting flows, etc., however, out-commuting intensity and in-commuting intensity remain outside the spotlight and for no good reason as they could prove very helpful.

With out-commuting, which is defined as commuting flows from one municipality -the original host- to another, out-commuters are calculated in accordance with the workforce of the original host, whilst with in-commuting, which is described as the commuting flows received by a municipality -the destination- from another, in-commuters are calculated in accordance with the workforce of the destination (Duquenne & Kaklamani, 2009; Harris et al., 2008).

Therefore, commuting intensity can be defined as the sum of employees who move daily out or in a municipality in comparison with the ones who live and work inside the municipality. Specifically, commuting intensity is estimated by the following relations (1).

$$Out_{Intensity}\,(i) = \frac{Or_i}{Fix_i}$$

$$In_{intensity}\,(i) = \frac{Ir_i}{Fix_i} \tag{1}$$

Where:
Or_i = active population living in municipality i and working in another municipality
Ir_i = active population living in another municipality but working in municipality i
Fix_i = active population living and working in the same municipality i

In the same vein, there is a useful measure - the Jobs to Workers Ratio (JWR). With the computation of this ratio, which traditionally is defined as the number of jobs per resident worker within a geographical unit, the degree of mixed land uses as well as the job accessibility can be apprehended (Antipova et al., 2011).

Hence, the JWR is calculated for the Greek municipalities and its descriptive statistics are shown at table 1. As it can be observed, in both cases the mean is smaller than 1 which indicates that in most municipalities richness in job opportunities does not exist, resulting in a longer commuting trip.

Table 1: JWR Descriptive Statistics

JWR	N	Range	Min	Max	Mean	Std. deviation
2001	1033	3.08	0.30	3.38	.92	.221
2011	325	2.46	0.28	2.74	.91	.254

The next step deals with the computation of commuting intensity of each Greek municipality for both years 2001 and 2011. In an attempt to make the results more comprehensible, they are mapped. Therefore, maps in the figures 1 and 2 show the intensity of commuting flows for the employed people in Greece in years 2001 and 2011. Through this initial estimation of intensity, a diversification between production areas and residence-consumption areas can be observed and different functions of space can be revealed. Moreover, it is noticed, through the differentiation of the colours of the municipalities, that the commuting intensity increased in a decade.

Greek municipalities differ between them in terms of commuting as depicted on the visual representation on maps. Out-commuting intensity is low in most of the islands and in the municipalities of Tripoli, Mani, Grevena, Komotini and others. With regard to in-commuting intensity, it tends to be low in most islands too, as well as in some municipalities, such as Grevena, Florina, Tripoli, Mani etc. At first glance, low commuting intensity in general is located mainly in islands and possibly in remote, mountainous areas like Grevena, Florina, Tripoli and others. It is interesting to note that in Grevena and Komotini the primary sector of the economy is highly developed and naturally are considered rather autonomous, since most positions are occupied by local residents who do not commute.

On the other hand, as it is shown in figure 2, high in-commuting intensity appears in the large urban centres in Athens, in Thessaloniki and in the nearby municipalities, whilst, as expected, high out-commuting intensity (figure 1) is observed in municipalities close to urban centres.

Therefore, differentiation of commuting intensity among municipalities is obvious, as is the polarization of employment in the main urban centres, which in turn means that the issue of jobs available in small and medium sized towns becomes increasingly important.

The remainder of this paper is organised as follows: Section 2 presents the methodological framework used in the analysis and the available data. Section 3 illustrates the results of the analysis and, finally, Section 4 concludes the paper and outlines some future research directions.

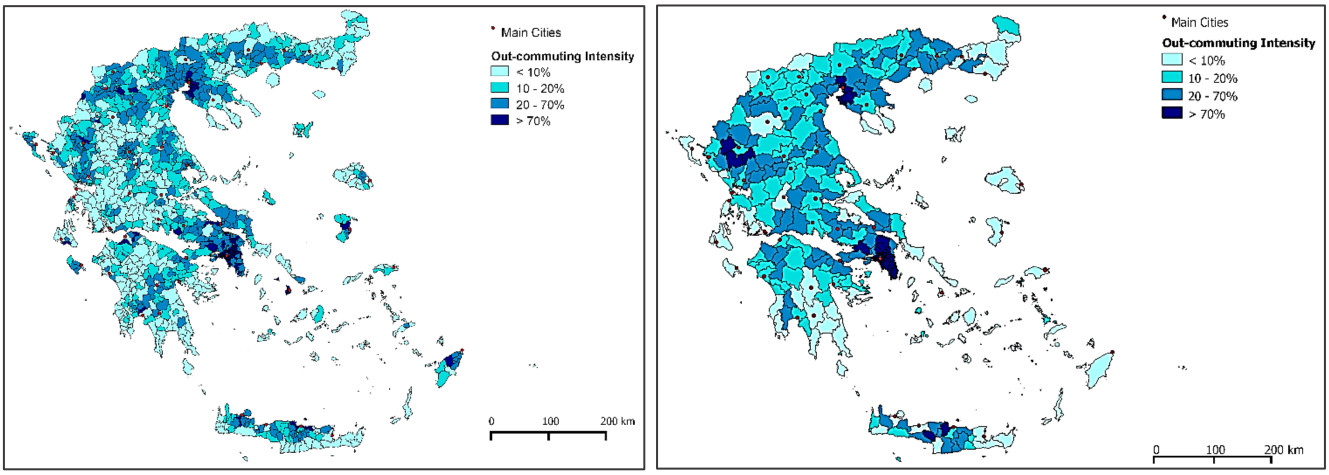

Figure 1: Out-commuting intensity in Greece (2001 left – 2011 right)

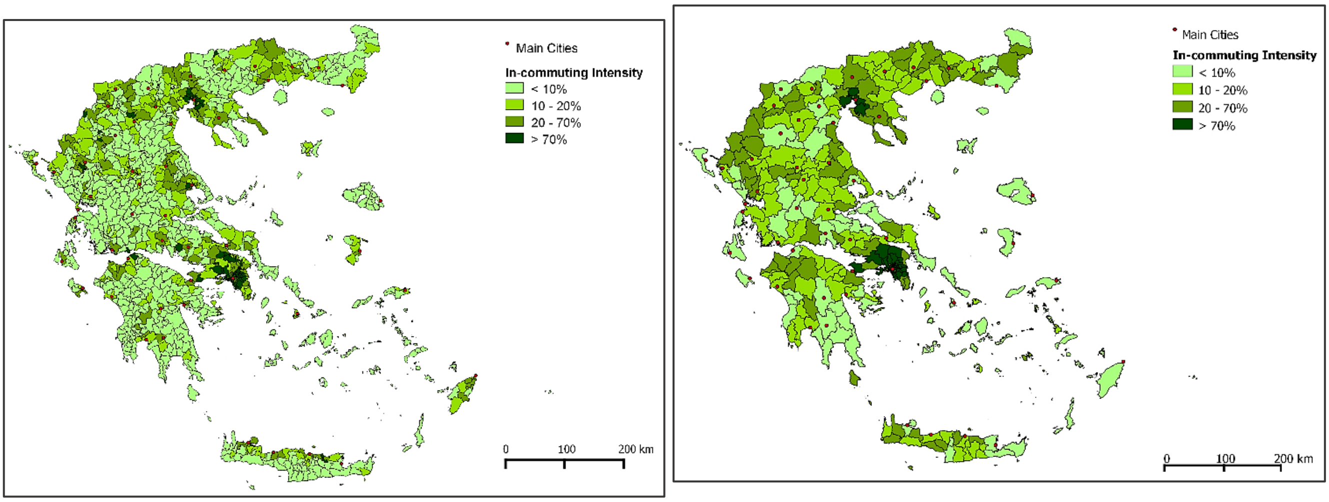

Figure 2: In-commuting intensity in Greece (2001 left – 2011 right)

Data and Methodology

The commuting data of the research derives from the official Census conducted by the National Statistical Service in 2001 and in 2011 and refers spatially to the Greek municipalities. This administrative unit is not the smallest one and it is obvious that the complexity of the commuting network increases as the scale gets smaller, since more daily trips to work take place in many different directions. Thus, the choice of the spatial scale used in the analysis is essential, influencing the analysis of intensity but also the spatial variability of the phenomenon.

The validity of this study has to do with the approved applicability and generality of the techniques which are used in the methodological framework. Regarding the quantitative approach of this study, in an empirical analysis a statistical model of generalized linear regression model is constructed and applied, which basically forms a multivariate mathematic function between a dependent variable and a set of independent variables. In particular, the present approach applies the Multinomial Logistic Regression version (MLR) of the Generalised linear models –the generalization of the General linear models-, which considers as dependent the variable having as values the predefined intensity categories and as predictor variables the numerical attributes characterising the municipalities and their population.

As regards the chosen method, it is common practice in several studies to consider a continuous variable as ordinal or nominal, in case the researchers wish to set certain thresholds, which essentially categorize the examined variable, but nonetheless the Simple Linear Regression models are very sensitive to the categorization of a dependent variable resulting in general failure to reveal the relation with the independent variables and production of unrealistic predictions (Sentas et al., 2005).

On the other hand, the logistic regression is defined as the technique used to determine which predictor variables are most strongly and significantly associated with the probability of a particular category in the criterion variable occurring. Because events either occur or do not occur, logistic regression assumes that the relationship between the predictors and the criterion is S-shaped or sigmoidal rather than linear. The change in a predictor is expressed as a logit or the natural logarithm of the odds (Cramer & Howitt, 2004).

The starting point of the methodological framework is the selection of the variables to be used, which is paramount to the effectiveness of the model. In this paper the variables, which together with a short description are shown in table 2, were selected on the basis of existing literature, experience of researchers and of course available data.

Table 2: Definition of variables and data sources

Variable	Symbol	Description	Measure	Primary data source and year
Out-commuting Intensity	Y1	Daily out-commuters of a municipality in comparison with the fixed employees	Percentage	EL.STAT., (2001,2011)
In-commuting Intensity	Y2	Daily in-commuters in a municipality in comparison with the fixed employees	Percentage	EL.STAT., (2001,2011)
Population density	X1	Population density of the municipality	Number of citizens / Area (km2)	EL.STAT., (2001,2011)

Dwelling density	X2	Dwelling density of the municipality	Population / Dwellings	EL.STAT., (2001,2010)
GDP per capita	X3	Gross Domestic Product per capita of municipality	Euros	EL.STAT., (2001,2011)
Participation of the secondary sector at the GDP	X4	Participation of the secondary sector to each prefecture's GDP	Percentage	EL.STAT., (2001,2011)
Educational level	X5	Number of people of the municipality with bachelor's degree or above	Number of people	EL.STAT., (2001,2011)
Immigrant density	X6	Immigrant density of the municipality	Number of immigrants / Area (km2)	EL.STAT., (2001,2011)
Unemployment	X7	Number of registered unemployed of the municipality	Number of people	EL.STAT., (2001,2011)

Brief descriptions of the variables, as well as hypotheses regarding the relation between independent and dependent variables, are laid down, having first selected them accordingly, as described.

At first, the density of the population, which can be seen as a measure of how close geographically it is or how urbanised the municipality is, may be highly related to commuting intensity and to elaborate this further - areas with high degrees of population and urbanisation tend to experience commuting within the spatial unit and so few move outside the city of residence and hence commuting tends to be shorter (Antipova et al., 2011; Polyzos et al., 2014; Susilo & Maat, 2007).

Dwelling density is also explored, following on from land use, which is used extensively in related papers, as it could serve as an indication of the commercial/office or the residential land type of a municipality, which would play a role in the commuting intensity as explained above (i.e., municipalities with an equilibrium of job positions and dwellings would depict lower commuting intensity).

With regard to GDP per capita, to Östh & Lindgren (2012), changes in GDP play a significant role in determining commuting distances. More specifically, urban commuting distances increase as an immediate response to GDP growth, while rural commuting increases eventually. Moreover, as far as Polyzos et al. (2013) are concerned, a high welfare of municipality is expected to reduce the out-commuting flows, keeping the employees within the municipality boundaries. Against this background, the in-commuting flows should be intensified.

To Polyzos et al. (2013) the higher the participation of the secondary sector to the GDP of the city is, the lower the potential of long distance commutes. On the other hand, according to Polyzos et al. (2014), the abovementioned variable is insignificant as regards commuting distance. In the same vein, the participation of the tertiary sector at the GDP of the city would also be used, but after running the statistical test on the assumption of no multicollinearity, multicollinearity was detected between these two variables, resulting in abandoning the last mentioned one.

As far as educational level is concerned, it seems to have a positive relation with the commuting distance and time. However, Antipove et al. (2011) deduced the non-significance of educational attainment in commuting behavior (Polyzos et al., 2013; Shoag & Muehlegger, 2015).

In parallel with the above, immigration, a very current topic in the whole Europe, is also a factor in the concept, as immigrants often face discrimination in finding a job, which together with their marginalisation in specific parts of the city, it affects where and how far they can commute (Antipova et al., 2011; Östh & Lindgren, 2012).

Finally, the unemployed are a special social group, as they are willing to travel longer distances in order to work, generating distant commuting flows and so this is issue is deemed noteworthy(Östh & Lindgren, 2012; Polyzos et al., 2013).

At the second step of the methodology, a multinomial model is chosen for application among the logistic regression models. In multinomial logistic regression, the criterion has more than two categories. The general form of the model for G outcome levels is the following (Kleinbaum & Klein, 2010):

$$\ln\left[\frac{P(D=g|X)}{P(D=0|X)}\right] = a_g + \sum_{i=1}^{k} \beta_{gi} X_i \tag{1}$$

Where g= 1, 2,…., G-1

Note that, an ordinal logistic regression model was applied firstly but the proportional odds assumptions were not met. In particular odds were not equal across all levels of intensity. It is clear to the reader of the literature that it is exceedingly rare that the parallel line assumptions in particular are met. Therefore, without considering the ordering of the data structure, a multinomial logistic regression model was used in order to relax the proportionality assumptions and to also offer an additional perspective as it provides a better fit to the data than the ordinal logistic regression, even if the categories of the dependent variable are ordered, according to Spitznagel (2008).

Therefore, in order to meet the standards for applying the multinomial logistic regression model, the scale response variables were transformed into ordinal (or nominal) ones, by dividing the range of their values into only three categories (shown in table 3) – no more, no less – as more categories did not provide valid models, given that in some categories, and especially in the last ones (high intensity), the number of municipalities was very small. Consequently, there was no need for more than three categories, which were set using as bounds approximately the tertiles of their empirical distribution (e.g. low, medium and high). This leads to almost equally probable categories.

Table 3: Descriptive statistics of the multinomial logistic regression model

Variable	Category code	Category spacing	Degrees of freedom (N) (2001)	Percentage of category (2001)	Degrees of freedom (N) (2011)	Percentage of category (2011)
Out-commuting intensity	[1] Low	Y1 ≤ 10%	383	37,1%	89	27,4%
	[2] Medium	10 <Y1≤ 40%	439	42,5%	144	44,3%
	[3] High	Y1 > 40%	210	20,3%	92	28,3%
In-commuting intensity	[1] Low	Y2 ≤ 10%	587	56,9%	95	29,2%
	[2] Medium	10 <Y2≤ 40%	318	30,8%	148	45,5%
	[3] High	Y2 > 40%	127	12,3%	82	25,2%

Results and Discussion

Before conducting the multinomial logistic regression analysis, the fitness ability of the models is checked, by using the likelihood ratio and the statistics of Pearson and Deviance. The calculations of these statistics of both models and for both years are presented in tables 4, 5, that prove that the full models statistically significantly predict the dependent variables better than the corresponding intercept only models (i.e., models with no predictors) alone. In other words, the values of the statistical significance of the final models chi-square prove statistically the presence of a relation between the dependent variable and the combination of the independent variables. Besides, the values of the statistics Pearson and Deviance higher than .05 (tables 4, 5) also evidence that the models fit the data well.

In the case of multinomial logistic regression, where the response variable is categorical, three indicators describe the model’s ability of determination, which constitute a generalization of the coefficient of determination R^2 (Nagelkerke, 1991; Polyzos et al., 2013). In particular, these indicators are the statistics, whose calculations are shown in tables 4, 5. The Cox and Snell R-square achieves a maximum of less than 1 for discrete models, in contrast with the Nagelkerke R-square whose upper bound is 1 (Nagelkerke, 1991). The values of the statistics of both models (tables 4, 5) are considered acceptable, bearing in mind that many researchers find these indicators only of marginal interest. In addition to this, to Polyzos et al. (2013), the pseudo-coefficients present lower values than the corresponding R^2 of a linear regression case.

Table 4: Goodness of fit statistics (Out-commuting intensity models)

2001						
Model	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	2.182	2.192	2.178			
Final	1.656	1.764	1.612	566,44	20	,000

Goodness-of-Fit				Pseudo R-Square		
	Chi-Square	df	Sig.	Cox and Snell	,422	
Pearson	1933,85	2042	,957	Nagelkerke	,481	
Deviance	1612,01	2042	1,000	McFadden	,260	
2011						
Model	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	701,19	708,76	697,19			
Final	478,38	538,92	446,38	250,82	14	,000
Goodness-of-Fit				**Pseudo R-Square**		
	Chi-Square	df	Sig.	Cox and Snell	,538	
Pearson	580,23	634	,938	Nagelkerke	,609	
Deviance	446,38	634	1,000	McFadden	,360	

Table 5: Goodness of fit statistics (In-commuting intensity models)

2001						
Model	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	2.267	2.267	2.267			
Final	1.377	1.506	1.352	941,93	26	,000
Goodness-of-Fit				**Pseudo R-Square**		
	Chi-Square	df	Sig.	Cox and Snell	,599	
Pearson	2046,34	2038	,444	Nagelkerke	,673	
Deviance	1325,60	2038	1,000	McFadden	,415	
2011						
Model	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	696,37	703,94	692,37			
Final	446,27	506,81	414,27	278,107	14	,000
Goodness-of-Fit				**Pseudo R-Square**		
	Chi-Square	df	Sig.	Cox and Snell	,575	
Pearson	544,69	634	,996	Nagelkerke	,653	
Deviance	414,27	634	1,000	McFadden	,402	

Afterwards, in the tables titled "Likelihood Ratio Tests" (tables 6, 7) the overall statistical significance of the independent variables is checked. In particular, in case of the out-commuting intensity model, all control variables are statistically significant. On the other hand, in case of the in-commuting intensity model (2011), the predictors X6, X7, which correspondingly refer to the immigrant density and the unemployment of the city destination, do not seem to be statistically significant.

Table 6: Likelihood Ratio Tests (Out-commuting intensity models)

2001						
Effect	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	1727,740	1826,525	1687,740	75,729	2	,000
X1	1683,155	1781,940	1643,155	31,144	2	,000
X2	1686,043	1784,828	1646,043	34,031	2	,000
X3	1660,199	1758,984	1620,199	8,188	2	,017
X4	1679,778	1778,563	1639,778	27,767	2	,000
X5	1678,074	1776,860	1638,074	26,063	2	,000
X6	1676,872	1775,657	1636,872	24,861	2	,000
X7	1676,719	1775,504	1636,719	24,708	2	,000
X2*X5	1659,405	1758,191	1619,405	7,394	2	,025
X3*X5	1667,141	1765,926	1627,141	15,130	2	,001
X5*X6	1673,645	1772,430	1633,645	21,634	2	,000
2011						
Effect	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	511,155	564,129	483,155	36,779	2	,000
X1	504,204	557,177	476,204	29,828	2	,000
X2	494,824	547,798	466,824	20,448	2	,000
X3	515,663	568,636	487,663	41,287	2	,000
X4	494,134	547,107	466,134	19,758	2	,000
X5	486,830	539,803	458,830	12,454	2	,002
X6	483,083	536,056	455,083	8,707	2	,013
X7	480,781	533,754	452,781	6,405	2	,041

Table 7: Likelihood Ratio Tests (In-commuting intensity models)

2001						
Effect	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	1379,689	1498,231	1331,689	6,087	2	,048
X1	1424,524	1543,066	1376,524	50,922	2	,000
X2	1441,274	1559,816	1393,274	67,672	2	,000
X3	1384,049	1502,591	1336,049	10,447	2	,005
X4	1389,065	1507,607	1341,065	15,463	2	,000
X5	1395,265	1513,807	1347,265	21,663	2	,000
X6	1380,394	1498,936	1332,394	6,792	2	,034
X7	1385,031	1503,573	1337,031	11,429	2	,003
X1*X2	1434,403	1552,945	1386,403	60,801	2	,000
X2*X3	1397,318	1515,860	1349,318	23,716	2	,000
X2*X6	1382,087	1500,630	1334,087	8,485	2	,014

X4*X5	1384,437	1502,980	1336,437	10,835	2	,004
X4*X7	1380,689	1499,231	1332,689	7,087	2	,029
X5*X7	1379,689	1498,231	1331,689	6,087	2	,048
2011						
Effect	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	475,413	528,387	447,413	33,147	2	,000
X1	452,403	505,376	424,403	10,137	2	,006
X2	456,908	509,881	428,908	14,642	2	,001
X3	458,583	511,557	430,583	16,317	2	,000
X4	463,628	516,602	435,628	21,362	2	,000
X5	448,930	501,903	420,930	6,664	2	,036
X6	443,088	496,061	415,088	,822	2	,663
X7	443,060	496,034	415,060	,794	2	,672

Before the last step of the model's parameters analysis, separately for each model, the classification tables 8, 10 and specifically the classification accuracy rates have to be checked. In both models the criteria for classification accuracy are satisfied, supporting in this way the utility of the model, since the classification accuracy rates are greater than the corresponding increased by 25% proportional by chance accuracy criteria (tables 9, 11).

Table 8: Classification accuracy (Out-commuting intensity models)

2001				
Observed	**Predicted**			
	Low	Medium	High	Percent Correct
Low	222	151	10	58,0%
Medium	131	283	25	64,5%
High	16	62	132	62,9%
Overall Percentage	35,8%	48,1%	16,2%	61,7%
2011				
Observed	**Predicted**			
	Low	Medium	High	Percent Correct
Low	58	30	1	65,2%
Medium	22	119	3	82,6%
High	6	28	58	63,0%
Overall Percentage	26,5%	54,5%	19,1%	72,3%

Table 9: Case Processing Summary (Out-commuting intensity models)

2001			
		N	**Marginal Percentage**
OUTCOM	Low	383	37,1%
	Medium	439	42,5%
	High	210	20,3%

Valid	1032	100,0%
Missing	2	
Total	1034	
Subpopulation	1032	
Chance rate		44,9%
2011		
	N	**Marginal Percentage**
OUTCOM Low	89	27,4%
OUTCOM Medium	144	44,3%
OUTCOM High	92	28,3%
Valid	325	100,0%
Missing	1	
Total	326	
Subpopulation	325	
Chance rate		43,9%

Table 10: Classification accuracy (In-commuting intensity models)

2001				
Observed	**Predicted**			
	Low	Medium	High	Percent Correct
Low	551	34	2	93,9%
Medium	183	113	22	35,5%
High	20	24	83	65,4%
Overall Percentage	73,1%	16,6%	10,4%	72,4%
2011				
Observed	**Predicted**			
	Low	Medium	High	Percent Correct
Low	50	43	2	52,6%
Medium	24	121	3	81,8%
High	4	16	62	75,6%
Overall Percentage	24,0%	55,4%	20,6%	71,7%

Table 11: Case Processing Summary (In-commuting intensity models)

2001		
	N	**Marginal Percentage**
INCOM Low	587	56,9%
INCOM Medium	318	30,8%
INCOM High	127	12,3%
Valid	1032	100,0%
Missing	2	
Total	1034	
Subpopulation	1032	
Chance rate		54,2%
2011		

		N	Marginal Percentage
	Low	95	29,2%
INCOM	Medium	148	45,5%
	High	82	25,2%
Valid		325	100,0%
Missing		1	
Total		326	
Subpopulation		325	
Chance rate			44,5%

The next part of the analysis is the estimation and interpretation of the regression coefficients, taking into consideration that the coefficients express the effects of the predictors on the log odds of being in one category versus the reference one. The results for each one of the four models are shown in tables 12, 13. In both models the last category "High" had been designated as the reference category and consequently each of the other two levels is compared with this one resulting in two sets of logistic regression coefficients for each model.

As regards the out-commuting intensity models, first of all, it is noticed that not all of the predictors are statistically significant for each category (Low – Medium), which means that these variables do not differentiate their groups from the baseline category (High). In addition, many of the significant predictors have B=0 or equivalently exp(B)=1, which corresponds to no change in the odds after a change of the predictor in relation to the reference category. Taking for example for the year 2001 the coefficients of the predictors in the category "Low", since the difference in relation to the baseline category would be more comprehensible, it is noted that for each unit increase in X2 the odds of being in the group "Low" decreases by 76,4% rather than "High" group. This high percent is justified due to value range of the predictor X2, because of which a unit increase is a relative huge increase. Besides this, high values of X2 are noticed in data mainly in densely built up urban areas with intense residential character, which could cause high out-commuting. In the same vein, for each unit increase in X6 the odds of being in the group "Low" increases by 13,4% versus the odds of being in the "High" category. This was expected because immigrants usually do not commute long distances. A similar pattern is found also in the corresponding coefficients of the predictors for the year 2011, reinforcing by this way the results about the variables' influence. It should be stressed that interaction effects (X2*X5, X3*X5, X5*X6) participate in the model of the year 2001, since they do not only improve the goodness of fit, but also are statistically significant. This significance of the interaction terms means that the impact for example of the variable X2 on the out-commuting intensity is not the same for all values of the variable X5.

As regards the in-commuting intensity models, like in the previous one not all of the independent variables are statistically significant for each category and in parallel with this, also in this model many of the significant predictors have exp(B)=1. With regard to coefficients of group "Low" for the year 2001, for each unit increase in X4 the odds of being in the group "Low" decreases by 9,7% rather than "High" group, which is logical sequence since a developed secondary sector leads to many job positions and thus many in-commuters. Similarly, for each unit increase in X5 the odds of being in the group "Low" decreases by 11,5% versus the odds of being in the "High" category, because perhaps a higher educational attainment of a municipality is related to a general better level of living resulting in pulling in-commuters. It should be noted that interaction effects participate again in the model of the year 2001, since they improve the goodness of fit and are also statistically significant. On the other hand, the effects of the most of variables for the year 2011 do not seem so significant, as Exp(B)of the most of them is close to 1.

Table 12: Parameter Estimates (Out-commuting intensity models)

2001									
OUTCOM		**B**	**Std. Error**	**Wald**	**df**	**Sig.**	**Exp(B)**	**95% Confidence Interval for Exp(B)**	
								Lower Bound	**Upper Bound**
Low	Intercept	8,945	1,081	68,448	1	,000			
	X1	-,007	,002	15,543	1	,000	,993	,989	,996
	X2	-1,442	,287	25,202	1	,000	,236	,135	,415
	X3	,000	,000	7,584	1	,006	1,000	1,000	1,000
	X4	-,058	,015	15,504	1	,000	,943	,916	,971
	X5	-,167	,031	28,281	1	,000	,846	,796	,900
	X6	,126	,032	15,657	1	,000	1,134	1,065	1,207
	X7	,001	,000	15,333	1	,000	1,001	1,001	1,002
	X2*X5	,017	,006	7,852	1	,005	1,017	1,005	1,030
	X3*X5	,000	,000	14,141	1	,000	1,000	1,000	1,000
	X5*X6	-,002	,001	10,660	1	,001	,998	,997	,999
Medium	Intercept	4,856	,945	26,418	1	,000			
	X1	-,002	,001	4,901	1	,027	,998	,996	1,000
	X2	-,330	,235	1,974	1	,160	,719	,454	1,139
	X3	,000	,000	2,661	1	,103	1,000	1,000	1,000
	X4	-,016	,014	1,353	1	,245	,984	,958	1,011
	X5	-,064	,028	5,072	1	,024	,938	,887	,992
	X6	,031	,023	1,881	1	,170	1,032	,987	1,079
	X7	,001	,000	6,136	1	,013	1,001	1,000	1,001
	X2*X5	,000	,006	,019	1	,891	,999	,988	1,010
	X3*X5	,000	,000	1,230	1	,267	1,000	1,000	1,000
	X5*X6	,000	,000	1,062	1	,303	1,000	,999	1,000

2011									
OUTCOM		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Low	Intercept	6,526	1,476	19,542	1	,000			
	X1	-,024	,007	11,053	1	,001	,976	,963	,990
	X2	-2,540	,597	18,069	1	,000	,079	,024	,254
	X3	,000	,000	2,003	1	,157	1,000	1,000	1,000
	X4	-,036	,022	2,672	1	,102	,965	,924	1,007
	X5	,000	,000	11,314	1	,001	1,000	1,000	1,001
	X6	,099	,030	10,972	1	,001	1,104	1,041	1,170
	X7	,000	,000	2,783	1	,095	,999	,998	1,000
Medium	Intercept	5,905	1,125	27,565	1	,000			
	X1	-,001	,001	3,968	1	,046	,999	,998	1,000
	X2	-1,112	,436	6,508	1	,011	,329	,140	,773
	X3	,000	,000	31,277	1	,000	1,000	1,000	1,000
	X4	,033	,015	4,738	1	,029	1,034	1,003	1,065
	X5	,000	,000	,223	1	,637	1,000	1,000	1,000
	X6	,009	,006	2,175	1	,140	1,009	,997	1,020
	X7	,000	,000	1,227	1	,268	1,000	1,000	1,001

The reference category is: High.

Table 13: Parameter Estimates (In-commuting intensity models)

2001									
INCOM		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Low	X1	,020	,009	4,552	1	,033	1,020	1,002	1,039
	X2	2,359	,353	44,682	1	,000	10,578	5,297	21,123
	X3	,001	,000	42,554	1	,000	1,001	1,000	1,001
	X4	-,102	,041	6,304	1	,012	,903	,834	,978
	X5	-,122	,038	10,258	1	,001	,885	,821	,954
	X6	-,360	,082	19,014	1	,000	,698	,594	,820
	X7	,003	,001	5,014	1	,025	1,003	1,000	1,006
	X1*X2	-,013	,005	7,809	1	,005	,987	,978	,996
	X2*X3	,000	,000	64,225	1	,000	1,000	1,000	1,000
	X2*X6	,159	,034	21,368	1	,000	1,172	1,096	1,253
	X4*X5	,004	,002	4,350	1	,037	1,004	1,000	1,007
	X4*X7	,000	,000	6,634	1	,010	1,000	1,000	1,000
	X5*X7	,000	,000	4,449	1	,035	1,000	1,000	1,000
Medium	X1	-,002	,002	,925	1	,336	,998	,994	1,002
	X2	1,233	,324	14,513	1	,000	3,432	1,820	6,474
	X3	,000	,000	13,553	1	,000	1,000	1,000	1,001

	X4	-,033	,038	,742	1	,389	,968	,898	1,043
	X5	-,033	,034	,930	1	,335	,968	,905	1,034
	X6	,017	,028	,389	1	,533	1,018	,963	1,075
	X7	,001	,001	1,419	1	,234	1,001	,999	1,004
	X1*X2	,001	,001	1,480	1	,224	1,001	,999	1,003
	X2*X3	,000	,000	29,199	1	,000	1,000	1,000	1,000
	X2*X6	-,011	,012	,819	1	,366	,989	,966	1,013
	X4*X5	,000	,001	,003	1	,957	1,000	,997	1,003
	X4*X7	,000	,000	,158	1	,691	1,000	1,000	1,000
	X5*X7	,000	,000	2,755	1	,097	1,000	1,000	1,000

2011

INCOM		**B**	**Std. Error**	**Wald**	**df**	**Sig.**	**Exp(B)**	**95% Confidence Interval for Exp(B)**	
								Lower Bound	**Upper Bound**
Low	Intercept	9,266	1,885	24,156	1	,000			
	X1	-,020	,007	7,918	1	,005	,980	,967	,994
	X2	-2,378	,678	12,288	1	,000	,093	,025	,351
	X3	,000	,000	6,359	1	,012	1,000	1,000	1,000
	X4	-,071	,025	8,293	1	,004	,931	,887	,978
	X5	,000	,000	4,300	1	,038	1,000	1,000	1,001
	X6	,031	,045	,466	1	,495	1,031	,944	1,126
	X7	,000	,000	,018	1	,894	1,000	,999	1,001
Medium	Intercept	7,958	1,682	22,382	1	,000			
	X1	-,010	,005	3,147	1	,076	,991	,980	1,001
	X2	-1,212	,593	4,175	1	,041	,298	,093	,952
	X3	,000	,000	17,316	1	,000	1,000	1,000	1,000
	X4	,005	,017	,084	1	,773	1,005	,972	1,039
	X5	,000	,000	,193	1	,660	1,000	1,000	1,000
	X6	,035	,022	2,527	1	,112	1,036	,992	1,082
	X7	,000	,000	,290	1	,590	1,000	,999	1,001

The reference category is: High.

Conclusions

Studies on the influence of various factors on commuting intensity are few and far between. Against this background, this paper has conducted a literature approach of commuting and constructed and applied a multi-regressive methodological framework for the determination of the factors which affect the out-commuting intensity and the in-commuting intensity respectively. The rationale of this study is that aspects of commuting, such as intensity, are affected directly by the socioeconomic framework and others. In particular, the use of MLR was proposed, in order to produce category estimates according to estimated probabilities. Category estimates may or actually should be useful for policy and decision making.

Regarding the findings in this study, which examined the influence of the factors at two points in time: 2001 and 2011, some of them are consistent with the literature, for example the low out-commuting of immigrants and the positive relation between the developed secondary sector of a municipality and the commuting intensity. However, many of the predictors were proven insignificant in some levels of the independent variables or corresponded to no change in the odds versus the reference category. Therefore, the use of different explanatory variables instead of the applied here could probably interpret the phenomenon of commuting intensity more completely.

In any case, daily employee commuting is an important geographical phenomenon and studying commuting is a main part of geographical research, as well as regional and spatial planning. The analysis generally verified the multidimensional nature of commuting and, particularly, that it constitutes a socioeconomic, geographic and political phenomenon, proving in this way the utility of quantitative spatial and socioeconomic analysis to the sustainable urban planning and policy. The interesting outcomes produced by the present multinomial logistic regression analysis or a similar one should suggest consultative material for the strategic planning of the development of a better transportation interregional and hinterland framework, as well as for the policies that target to upgrade the standards of living of the population and finally for the environmental issues.

The present analysis may provide a reference for future comparisons in this study area by applying the methodology with necessary modifications to other data sets. Studies on the driving factors of commuting in Greece are necessary because there is still lack of information. Moreover, some of the above ambiguous results, like insignificant predictors, call for more empirical studies, as well as more convincing theories to untangle the complex interaction between a range of factors and commuting outcomes, because although these findings have been verified in the Greek municipalities context, further research may enhance our understanding of how commuters choose or do not change their employment and residence location. Beside this, further research should be focused on lower spatial hierarchical units, where commuting flows may change evidently and the present results cannot be simply transferred to the commuting models for the lower spatial levels. Concluding, commuting analysis is of big importance in the decision making on the local levels and commuting intensity is widely conditioned by the chosen spatial scale.

References

Axisa, J., Scott, D. and Newbold, B., 2012, "Factors influencing commute distance: a case study of Toronto's commuter shed," *Journal of Transport Geography*, 24, 123–129.

Antipova, A., Wang, F. and Wilmot, C., 2011, "Urban land uses, socio-demographic attributes and commuting: A multilevel modelling approach," *Applied Geography*, 31, 1010–1018.

Clark, W., Huang, Y. and Withers, S., 2003, "Does commuting distance matter? Commuting tolerance and residential change," *Regional Science and Urban Economics*, 33, 199–221.

Cramer, D. and Howitt, D., 2004, *The SAGE Dictionary of Statistics a practical resource for students in the social sciences*, London, SAGE Publications Ltd.

Duquenne, M.N. and Kaklamani, S., 2009, "The daily mobility of employees in Greece: a first attempt to capture the phenomenon intensity," *DemoNews*, [Online]. 6, available at: http://www.demography-lab.prd.uth.gr [Accessed at: 17 August 2015] [in Greek].

Harris, S., Alasia, A. and Bollman, R., 2008, "Rural Commuting: Its Relevance to Rural and Urban Labour Markets," *Rural and Small Town Canada Analysis Bulletin*, 7 (6), 1-23.

Kleinbaum, D. and Klein, M., 2010, *Logistic Regression - A Self-Learning Text*, 3rd ed., New York, Springer.

Nagelkerke, D., 1991, "A note on the general definition of the coefficient of determination," *Biometrika*, 78 (3), 691-692.

Nielsen, T. and Hovgesen, H., 2008, "Explanatory mapping of commuter flows in England and Wales," *Journal of Transport Geography*, 16, 90-99.

Östh, J. and Lindgren, U., 2012, "Do changes in GDP influence commuting distances? A study of Swedish commuting patterns between 1990 and 2006," *Tijdschrift voor Economische en Sociale Geografie*, 103(4), 443-456.

Polyzos, S., 2011, Regional Development, Athens, Kritiki Publications [in Greek].

Polyzos, S., 2015, *Urban Development,* Athens, Kritiki Publications [in Greek].

Polyzos, S., Tsiotas, D. and Minetos, D., 2013, "Determining the Driving Factors of Commuting: An Empirical Analysis from Greece," *Journal of Engineering Science and Technology Review*, 6(3), 46–55.

Polyzos, S., Tsiotas, D. and Papagiannis, K., 2014 "Determining the changes in commuting after the Ionian Motorway's construction", in: MIBES Proceedings 2014, *9th MIBES International Conference*. Thessaloniki, Greece, May 30- June 01, 2014, 187-205.

Sentas, P., Angelis, L., Stamelos, I. and Bleris, G., 2005, "Software productivity and effort prediction with ordinal regression," *Information and Software Technology*, 47, 17-29.

Shoag, D. and Muehlegger, E., 2015, "Commuting times and land use regulations," *Procedia Engineering*, 107, 488-493.

Spitznagel, E., 2008, "Logistic Regression," in: C.R. Rao, J.P. Miller & D.C. Rao, eds. 2008, *Handbook of statistics 27*, Amsterdam, Elsevier B.V., Ch. 6.

Stefanouli, M. and Polyzos, S., 2015, "Determination of Functional Urban Areas in Greece on the basis of commuting", in: *4th Hellenic Conference of Planning and Regional Development*. Volos, Greece, September, 24 – 27, 2015 [in Greek].

Susilo, Y. and Maat, K., 2007, "The influence of built environment to the trends in commuting journeys in the Netherlands," *Transportation*, 34, 589-609.

Tsiotas, D. and Polyzos, S., 2013, "Interregional commuting under the Network Theoretical perspective: an empirical analysis from Greece," in: Academic Conferences Association & Czech Technical University, *Multidisciplinary Academic Conference on Transport, Logistics and Information Technology*. Prague, Czech Republic, May 09-11 2013.

Van der Laan, L. and Schalke, R., 2001, "Reality versus Policy: the Delineation and Testing of Local Labour Market and Spatial Policy Areas," *European Planning Studies*, 9(2), 201-221.

Wong, C., 1995, "Developing Quantitative Indicators for Urban and Regional Policy Analysis", in Hambleton R, Thomas H (eds) *Urban Policy evaluation: challenge and change*, Cardiff, Paul Chapman Publishing Ltd.