A bottom-up study on the relationship between transportation expenditure and socio-demographic variables: evidence from the Italian case study

Giorgio Besagni*, Marco Borgarello

Ricerca sul Sistema Energetico - RSE S.p.A., Power System Development Department, via Rubattino 54, 20134 Milano (Italy)

* Corresponding author: Giorgio Besagni, giorgio.besagni@rse-web.it, Ricerca sul Sistema Energetico - RSE S.p.A., Power System Development Department, via Rubattino 54, 20134 Milano (Italy)

Abstract

A precise understanding of the relationships between the household characteristics and the transportation expenditures is of paramount importance to support bottom-up policies, aiming at defining decarbonisation pathways keeping into account the household budget constraints. Despite the considerable amount of research activities carried out during the last decades, an agreement regarding the factors influencing the transportation expenditures is far from being reached. This paper contributes to the present-day discussion, focusing on the Italian case study, by analyzing the relationships between the private, public and total transportation expenditure and the socio-demographic and geographical dimensions. The impact that the household characteristics have on the transportation expenditures have been explored by coupling (a) the ordinary least squares method, to determine the relationship between the variables, (b) the variance inflation factor, to check for multicollinearity issues, (c) the least absolute shrinkage and selection operator, to select variable. Subsequently, a segmentation of the Italian families is proposed, by using a segmentation-tree approach and the outcomes of the previous analysis. It is found that the geographic area (in terms of the macro-scale as well as the micro-scale geographic locations) as well as income-related variables are likely to be factors influencing the transportation expenditures. These

observations may serve as bottom-layer for the forthcoming studies regarding decarbonisation of the transportation sector, considering also the household budget constraints.

Keywords. Transportation expenditure; Residential sector; Socio-demographics; Lasso regression; Multicollinearity; Household segmentation

1 Introduction

The "*sustainable mobility paradigm*", defined by Banister (Banister, 2008), is a priority in the current research agenda owing to the large share of the primary energy consumption as well as the emissions related to the transportation sector (Anciaes and Jones, 2020; Ben-Salha et al., 2018; Sajid et al., 2019). For example, in 2016 the transportation sector accounted approximately 28% of the total final consumption on a global perspective. In recent years, an increasing number of research activities has been conducted to support the pathways towards the "*decarbonisation of country-scale energy systems*"[1], as mentioned by Sovacool et al. (Sovacool et al., 2018), Tapio et al. (Tapio et al., 2007) and reviewed by Miller et al. (Miller et al., 2016). In addition, Zawieska and Pieriegudb (Zawieska and Pieriegud, 2018) and Ventura et al. (Ventura et al., 2017) mentioned that achieving in the international aims (European Commission, 2011) relies on a profound transformation of the transportation sector. This transformation should be guided by following two perspective: (i) actions on the technology-side and (ii) actions on the policy-side. The former relies on pursuing the electrification pathway, on deploying storage technologies (at different scales) and on improving vehicle efficiencies in general; the latter relies on pursuing and adopting top-bottom policies, whose correct implementation relies on the precise understanding of the bottom implications, i.e., the relationships between the household characteristics and the transportation dimension (Besagni et al., 2019). Indeed, when planning top-bottom policy schemes and when designing decarbonisation pathways, the household-scale constrains should be accounted. For example, it should be considered that the "*decarbonisation pathways*" can be sustained as long as they are able to satisfy the constrains related to the household budget (Dias et al., 2019). For this reason, studying the factors influencing the household energy-related expenditures is of fundamental importance.

In the case of transportation, studying the relationships between the household characteristics and the transportation expenditure belongs to the field of the so-called "*human dimension*" of the energy-intensity

---

[1] It is worth noting that, despite ''*decarbonisation*'' is used, the wording ''*defossilisation*'' is more correct. Indeed, hydrocarbons will be used on the long-terms (i.e., biomass, power-to-gas/fuels, power-to-methane, …). These synthetic fuels do not contribute to the CO_2 concentration in the ecosphere, since in closed loops.

in transportation and its subsequent "*energy metabolism*" (Lowe et al., 2018; Stephenson et al., 2015). This concept was also addressed by Tian et al. (Tian et al., 2016), who pointed out that the energy consumption at the "*household-scale*" determines the carbon emission at the "*country-scale*". In this perspective, Longhi (Longhi, 2015) and Besagni and Borgarello (Besagni and Borgarello, 2018) mentioned that the precise understanding of the factors influencing the energy expenditure serves as basis for policymakers when planning investments aiming at reducing the energy consumption at the "*household-scale*". It is worth mentioning that analysing the travel expenditure has some other advantages, as it will open up new studies regarding "*decarbonisation pathways*" under household budget constraints. On the practical point of view, the prediction of the travel expenditures might be employed in forthcoming activities regarding the decarbonisation of the transportation sector keeping into account the household socio-economic constrains. This paper contributes to the present-day discussion by investigating the transportation sector and the existing relationships between the "*transport expenditure*" (private, public and total) and the "*socio-demographic dimension*" (i.e., household composition, income of the household, qualification of the occupants…) and the "*geographic dimension*" (i.e., the macro-scale and the micro-scale geographic location; see also the recent paper of Ke and McMullen (Ke and McMullen, 2017)). This analysis is interesting also when considering the demographic shift experienced by European countries, thus causing changes in behaviour and attitudes (Torgler et al., 2008). As the population is progressively ageing, this may reflect in the energy metabolism at the different levels: Liddle (Liddle, 2014) and Brand et al. (Brand et al., 2013) mentioned that the ageing of population is likely to increase the residential energy consumption and reduce the transport-related energy use. This topic was further investigated by Bardazzi and Pazienza (Bardazzi and Pazienza, 2018), who studied how the changes in the "*socio-demographic dimension*" and in the "*economic drivers*" would affect the private transport-related fuel demand in Italy. They concluded that, on one hand, fuel consumption is likely to decrease with age and, on the other hand, behaviour and attitude variables are important determinants which should be taken into account when modelling the whole transportation sector (i.e., encompassing all ages and social stratifications). Similar outcomes were also derived by Soltani et al. (Soltani et al., 2018), considering the Iranian case study by Orru et al. (Orru et al., 2019), considering the northern/eastern Europe case study. Okada (Okada, 2012) contributed to this discussion by proposing an inverted *U-shaped* relationship between the share of elderly people and carbon dioxide emissions from transportation. In this perspective, the interested reader may also refer to the study of O’Neill et al. (O'Neill et al., 2012) regarding the relationship between carbon dioxide emissions and population changes over the time.

The above literature survey supports the importance of a precise modelling of the transportation sector, encompassing the household-scale. In the following, a brief literature survey regarding the different works on the factors influencing transportation attitude is proposed to better fit the proposed contribution within the existing body of knowledge. Arbués et al. (Arbués et al., 2016) considered the Spain case study and applied a multilevel multinomial logit model to investigate the relationship between the travel mode and the socio-demographic, economic variables, land-use features and trip attributes. They found that the socio-demographic variable, transportation type and the geographical variables have a significant influence of the above-mentioned relationship. Soltani et al. (Soltani et al., 2018) considered the Iranian case study and, by applying different regression approaches, found that socio-demographic (viz., employment status, household size, car ownership, …) and geographical variables are significant factors influencing to explain the car use. Orru et al. (Orru et al., 2019) considered the northern/eastern Europe case study (viz., Sweden and Estonia) and explored the relationships between socio-demographic variables (also considering the behaviour variable and the income characteristics) and transportation patterns. Besides the differences observed in the two countries (suggesting a higher geographical dimension of the transportation consumption patterns), they found that socio-demographic variables (i.e., social positions, household variables, income-related variables …) have significant relationship with respect to travel intensity (car travel frequencies). Etminani-Ghasrodashti et al. (Etminani-Ghasrodashti et al., 2018) considered the Iranian case study and, in particular, they studied the (sub-)socio-demographical stratification determined by university students. By applying a multinomial logit method, they explored the perceptions of the respondent class with respect to the public/private transportation, travel attitudes and lifestyle. Finally, Abenoza et al. (Abenoza et al., 2017) applied a cluster analysis to a Sweden case study to offer a segmentation of the population based on socio-demographic variables, travel characteristics as well as accessibility.

This paper contributes to the existing discussion, focusing on the Italian case study, by analysing the relationships between the transportation expenditure (private/public/total) and the household variables. This study is intended to clarify the factors influencing transportation intensity and, in a broader perspective, to provide information regarding the share of household budget devoted to transportation. The second outcome is particularly interesting for future researchers aiming at simulating the large scale diffusion of electric vehicles, which should take into account the household budget constrain. When considering the Italian case study, it is worth mentioning that it is peculiar owing to three concepts (Bardazzi and Pazienza, 2018): (i) Italy experiences a high motorcycle car ownership rate compared with other countries; (ii) Italy is experiencing a fast aging of population and (iii) Italy is characterized by a constant increase in life expectancy. In particular, in this paper, the impact that the household characteristics (both the socio-demographic and the geographical dimensions) have on the transportation expenditures have been explored, based on the microdata taken from the Italian Household Budget Survey published by the Italian Statistical Office. The analysis is performed by coupling the ordinary least squares method, to determine the relationship between the variables, the variance inflation factor, to check for multicollinearity issues, and the least absolute shrinkage and selection operator, to select significant predictors. Subsequently, a segmentation of the Italian families is proposed, by using a *CART* approach. In summary, this study is intended as a first step toward a complete and comprehensive definition of the factors influencing the transportation expenditure: the proposed outcome may be coupled, in the future, with behavioural determinants (Acheampong and Cugurullo, 2019), by considering additional datasets obtained by the Italian Statistical Office.

This contribution is organized as follows. Section 2 discusses the dataset and the statistical methods; Section 3 describes the results of the statistical procedure and, finally, Section 4 contains our conclusions.

2 Research design and methods

Within this section, the employed dataset (Section 2.1), the dependant variables (Section 2.2), the predictors (Section 2.3) and the statistical methods (Section 2.4) are presented and discussed.

2.1 The dataset

The dataset considered in this research is the "*Household Budget Survey: microdata for research purposes*" (reference year: 2015, (ISTAT, 2017)), which is representative of the whole Italian population and was obtained by Italian National Institute of Statistics. The micro-data (more than 1,264 variables concerning monthly expenditures along with data regarding socio-demographic, dwelling and appliances) were collected in 502 different municipalities from 15,015 households.

2.2 The dependent variables

As this study aims at relating the transportation expenditure (private, public and total) to the household variables, three dependant variables are considered. The first dependent variable is the "*private transport expenditure*" (θ_1), which has been obtained by summing different expenditures, to describe the different patterns of the "*private transportation*": (a) gasoline expenditure, (b) diesel expenditure, (c) "*other fuel*" expenditure, (d) timetable parking expenditure, (e) motorway tolls, (f) tire expenditure, (g) vehicle spare part expenditure, (h) car accessory expenditure, (i) lubricant expenditure, (j) maintenance/repair service expenditure, (k) car/motor expenditure. The second dependent variable is the "*public transport expenditure*" (θ_2), which has been obtained by summing the following contributions: (a) integrated transport (train/bus/coach/metro/tram) tickets/subscriptions, (b) school busses expenditure, (c) other integrated transport ticket or subscriptions. Finally, the third dependent variable is the "*total transport expenditure*" (θ_3), which has been obtained by summing θ_1 and θ_2. Households having private/public/total transport expenditure equal to zero were excluded from the analysis: Figure 1 provide a descriptive overview of the employed variables after this procedure.

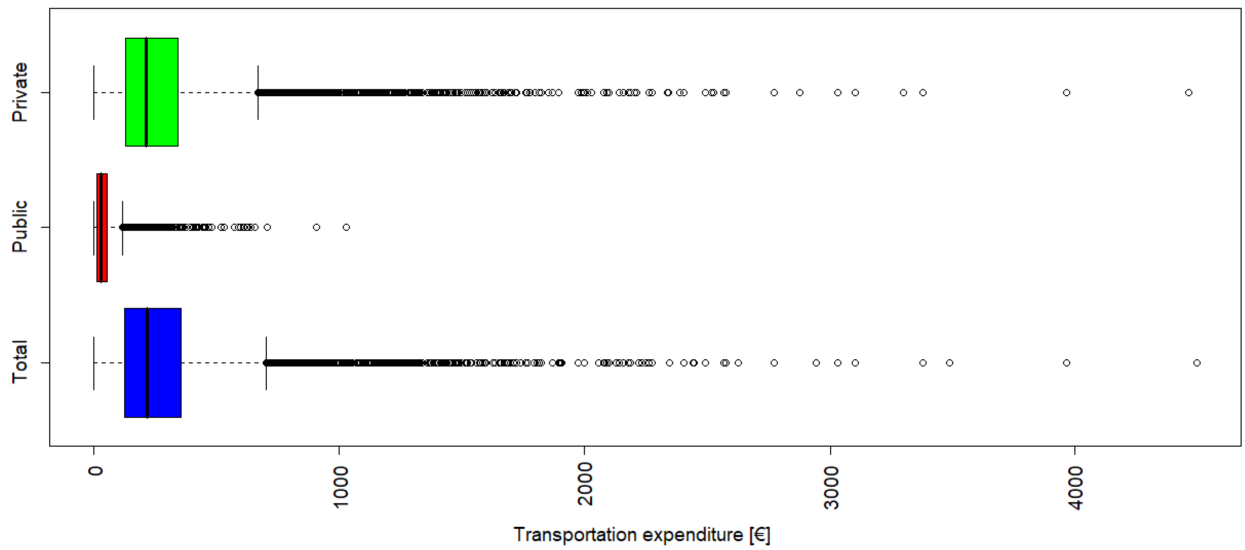
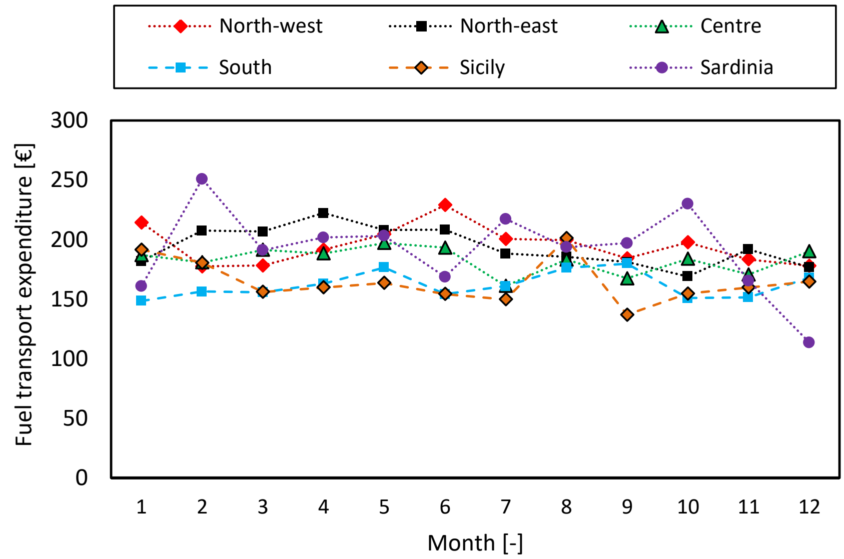
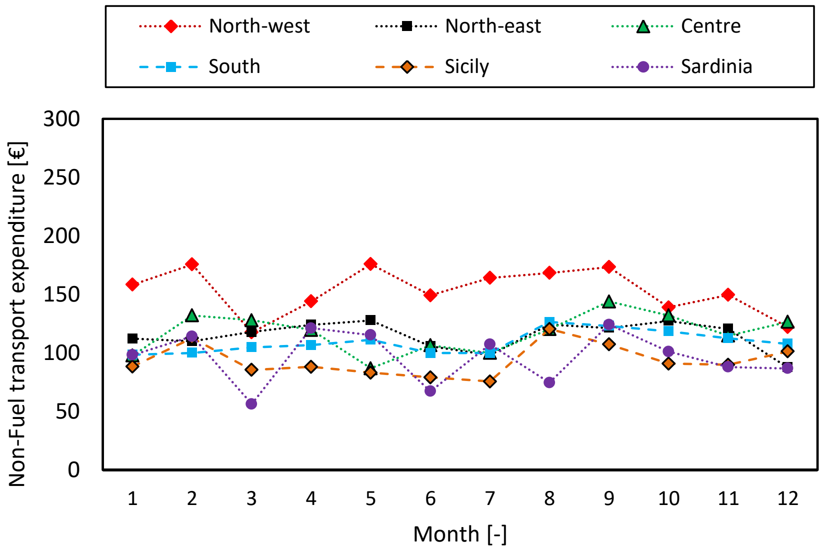


Figure 1: Box plot of the dependant variables.

It is noted that the micro-data included in the dataset concerns monthly expenditures and, in particular, they refer to a precise month. Before going ahead with the analysis, it should be verified whether an annual calibration is needed or not (Besagni and Borgarello, 2018). For the sake of clarity, Figure 2 and Figure 3 displays the relationship between the time variable (month) and the different dimensions of the transportation expenditures. In particular, Figure 2 focuses on the differences between fuel and non-fuel expenditures (to unveil that both of them should be considered), Figure 3 focuses on the differences between private, public and total transportation expenditures. It should be noted that, in September and in October, some areas exhibit higher public transport expenditures: this is likely to be caused, for example, by school opening, which force a change of behavior of the household components. Similarly, the higher transportation expenditures in November and in April might be caused by the tire shifts. The slight changes in shape of the expenditure patterns in the southern areas, in Sicily and in Sardinia in the summer seasons might be accused by two effects: (i) the high-intensity tourism (which also affects the other regions, by cross-migration effects) and (ii) the migration of students and workers from/towards the norther areas. In summary, Figure 2 and Figure 3 summarizes the average values of the private and total transportation expenditures in the different Italian regions, to provide insights within their relative percentage. It is worth noting that the differences between the southern areas and the norther areas are likely to be caused by the different income in these two areas as well as the lack of infrastructures in the southern parts (i.e., for the public transport). To sum up, the dependent variable does not exhibit high dependency with respect to the time variable. Hence, an annual calibration is not needed (as done in ref. (Besagni and Borgarello, 2018)).

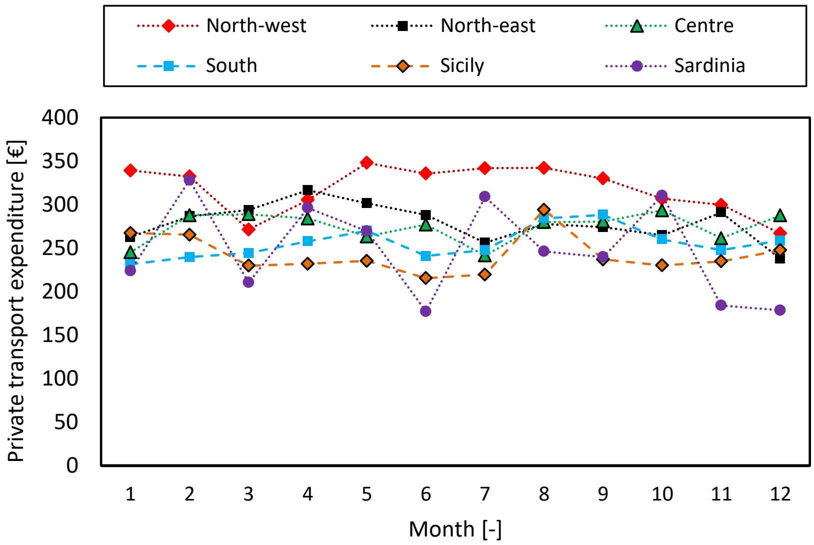


(a) Fuel expenditure

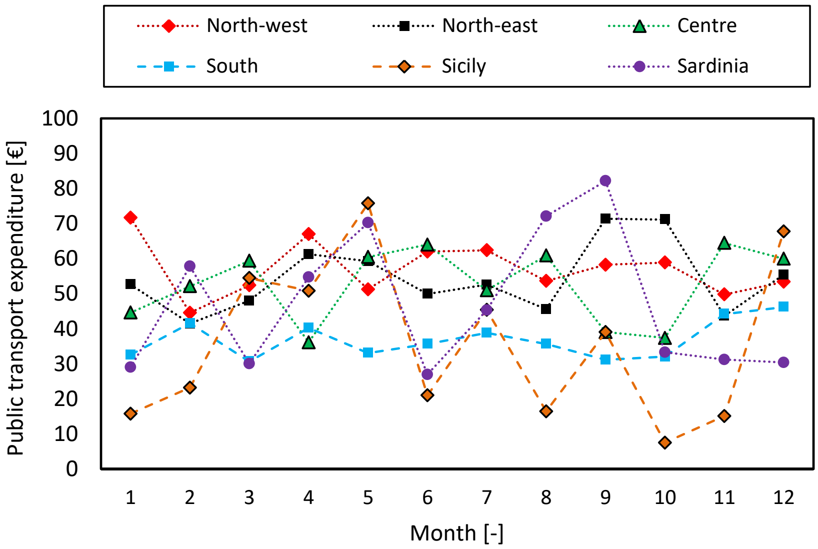


(b) Non-fuel expenditure

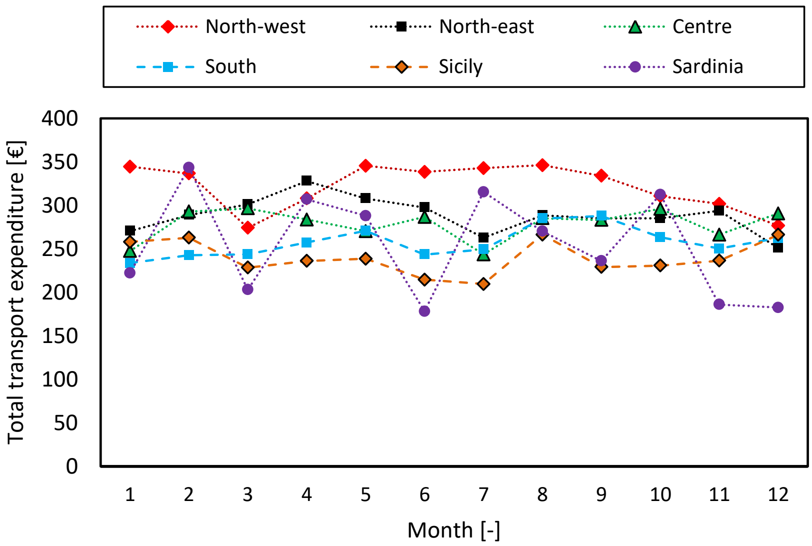
Figure 2: Relationship between time of the year and "*private transport expenditure*": differences between fuel and non-fuel expenditures.



(a) Private transport expenditure



(a) Public transport expenditure



(c) total transport expenditure

Figure 3: Relationship between time of the year and "*private transport expenditure*": differences among private, public and total expenditures.

As discussed by Longhi (Longhi, 2015), an advantage of using logs is that the regression coefficients refer to the relative changes rather than the absolute changes in the transportation expenditures.

2.3 The predictors

Table 1 lists the variables used as predictors in the regression analysis, along with their frequencies and summary statistics. In the following, *HRP* is the Household Representative Person, which is the individual who represents the household (viz., the highest income earner in the household). Categorical variables were dummy-coded prior to the statistical analyses.

Table 1: Socio-demographic variables with their frequencies. (* = reference category) - Summary statistics are computed on the entire data-set.

Variable	**Summary statistics**
Sex of the *HRP*	(a) Male [10,193]*, (b) Female [4,820]
Current economic resources	(a) Optimal [279], (b) Adequate [7,912]*, (c) Scarce [5,651], (d) Insufficient [1,171]
Changes in economic resources compared to the previous year	(a) Much improved [30], (b) A little bit improved [512], (c) More or less the same [8,488]*, (d) A little worsened [4,626], (e) Much worsened [1,357]
Absolute poverty	(a) Yes [834], (b) No [14,179]*
Birth place of the household components	(a) Only born in Italy [13,456]*, (b) At least one born abroad [973], (c) Only born abroad [584]
Citizenship of the household components	(a) Only Italian citizens [14,176]*, (b) At least one foreign citizens [257], (c) Only foreign citizens [580]
Marital status of the HRP	(a) Unmarried [2,551], (b) Married or cohabitant [8,252]*, (c) Married but not cohabitant [355], (d) Legally separated [625], (e)) Divorced [698], (f) Widow or widower [2,532]
Qualification of the occupants	(a) No member has a qualification [377], (b) At least one member with elementary school [1,978], (c) At least one member with junior high school [3108], (d) At least one member with high school [6,483]*, (e) At least one member with a degree [3,067]
Work contract of the occupants	(a) There is neither temporary job nor permanent job [7,536]*, (b) At least one temporary job [1,125], (c) At least one permanent job [6,352]
Source of income of the occupants	(a) There is no income [83], (b) At least one maintained [413], (c) At least one pension [4,911], (d) At least one income [9,606]*
Enrolment in study courses	(a) No members enrolled in a course [10,930]*, (b) At least one in no title school [419], (c At least one in elementary school [747], (d) At least one in junior high school [584], (e) At least one in high school [1,244], (f) At least one in a degree or post-degree course [1,089]
Expenditure for elderly or	(a) Yes [100], (b) No [14,913]*

disabled people	
Household structure	(a) Single person 18-34 years [391], (b) Single person 35-64 years [1,817], (c) Single person 65 years and more [2,240], (d) Couple without children with HRP 18-34 years [178], (e) Couple without children with HRP 35-64 years [1,350], (f) Couple without children with HRP 65 years and more [2,164], (g) Couple with 1 child [2,276]*, (h) Couple with 2 children [2,184], (i) Couple with 3 children or more [495], (l) Mono parent family [1033], (m) Others [885]
Number of workers in the primary sector	(a) No one [13,622]*, (b) One [1,100], (c) More than one [291]
Number of workers in the secondary sector	(a) No one [9,766]*, (b) One [4,098], (c) More than one [1,149]
Number of workers in the tertiary sector	(a) No one [4,577], (b) One [6,195]*, (c) More than one [4,241]
Number of managers and employees	(a) No one [8,227]*, (b) One [4,739], (c) More than one [2,047]
Workers and similar (employers)	(a) No one [8,166]*, (b) One [4,741], (c) More than one [2,106]
Entrepreneurs and freelancer workers	(a) No one [13,696]*, (b) One [1,172], (c) More than one [145]
Self-employed workers	(a) No one [11,876]*, (b) One [2,583], (c) More than one [554]
Age of the *HRP*	(a) Up to 34 years [995], (b) From 25 to 44 years [2,343], (c) From 45 to 54 years [3,059]*, (d) From 55 to 64 years [2,934], (e) From 65 to 74 years [2841], (f) From 75 years [2,841]
Type of municipalities	(a) Centre of metropolitan area [1,889], (b) Periphery of metropolitan area and municipalities with 50.001 inhabitants and more [4,032], (c) Other municipalities until 50.000 inhabitants [9,092]*
Geographic location	(a) North-west [3,284], (b) North-east [3,382], (c) Centre [2,791]*, (d) South [4,385], (e) Sicily [753], (f) Sardinia [418]
Number of cars	(a) no one [2,761]*, (b) one [7,324], (c) two [4,226], (d) three or more [702]
Free time expenditures[2]	Continuous variable [Mean = 22.68 / Variance = 3027]

[2] This variable is following expenditures: sport events/activity/subscriptions, cultural and naturalistic visits ticket/subscriptions, cinema/theatre ticket/subscriptions, recreational activities ticket/subscriptions, training courses, …

2.4 The statistical methods

The statistical approach couples four methods: (a) the ordinary least squares method (OLS), (b) the variance inflator factor (VIF), (c) the Least absolute shrinkage and selection operator (LASSO), (d) a *CART* approach to perform the tree-segmentation procedure.

2.4.1 Regression and selection of variables

A similar procedure has been implemented by Besagni and Borgarello (Besagni and Borgarello, 2018). The procedure consists of the following phases:

- Phase#a. *OLS* relates the dependent variable with the predictors listed in Table 1:

$$y_i = \ln(\vartheta_{1,2,3}) = \beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\ldots+\beta_k x_{ik}+\varepsilon_i = \beta_0+\sum_{j=1}^{w}\beta_j x_{ij} + \varepsilon_i \qquad (1)$$

  In Eq(1), y_i is the logarithm of the selected dependent variable ($\vartheta_{1,2,3}$) for the i-household; x_{ij} is the i-predictor for the j-household out of w-predictors; β_0 is the constant term (viz. the intercept); β_j is the j-coefficients for the x_{ij} variable; ε_i is the error having null mean and constant variance. As the dependent variable has a left-skewed distribution, it is implemented after a log-transformation. The performance of the model as a whole is estimated based on the adjusted coefficient of determination (R^2_{adj}). Once the *OLS* analysis is completed, multicollinearity is checked by inspecting the variance-inflation factors for every β_j, as follows:

$$VIF_j = \frac{1}{1-R_j^2} = \begin{cases} if\ \mathrm{VIF} = 1 & \text{predictors are uncorrelated} \\ if\ \mathrm{VIF} > 1 & \text{predictors may be correlated} \end{cases} \qquad (2)$$

  To date, there is no widely-accepted agreement on the cut-off point for VIF (VIF_{max}); based on the outcomes of (Besagni and Borgarello, 2018), VIF_{max} = 3 has been selected.

- Phase#b. If multicollinearity is detected by Eq. (2), a LASSO regression procedure will be applied. LASSO procedure is a variable shrinkage based on a penalty. This approach solves a constrained optimization problem. The procedure proposed in ref. (Besagni and Borgarello, 2018) has been applied here too. Based on this approach, significant variables are selected and, subsequently, a regression procedure (Eq. (1)) is repeated again. Finally, R^2_{adj} is computed and VIF are inspected again.

- Phase#c. If multicollinearity is detected again, after phase#b, a selection of predictors will be conducted by a progressive exclusion of the predictors. Predictors are progressively excluded from the least significant: at each step, the changes in R^2_{adj} are analyzed and VIFs are inspected (R^2_{adj} may reduce at maximum of 0.5 %). The detailed procedure follows ref. (Besagni and Borgarello, 2018).
- Phase#d. If no multicollinearity is detected following the LASSO procedure, a selection of predictors will be obtained by the above recursive procedure (R^2_{adj} may reduce at maximum of 0.2 %).

2.4.2 Household segmentation

After above-mentioned regression procedure is completed, the household segmentation is employed aiming at identifying the classes of households having homogeneous characteristics with respect to the transportation expenditure (in terms of the significant predictors obtained in the above-mentioned procedure). To this end, the *CART*[3] approach, introduced by Breiman et al. (Breiman et al., 1984), has been applied (using as input the variables found significant after above-procedure): it is based on a binary and recursive partitioning of the dataset and it uses a flowchart-like tree structure to segregate the complete dataset into various classes. In the present case, the tree is expended up to reaching its asymptotic region (where additional splitting does not let a considerable improvement in the results (as done in ref. (Besagni and Borgarello, 2018). the *CART* method is based on a binary and recursive partitioning of the dataset and it uses a flowchart-like tree structure to segregate the complete dataset into various classes. In the inverted-tree structure, three types of nodes can be observed: root nodes, internal nodes, and leaf nodes, which represents the outcomes of the classification; a non-terminal (or parent) node is a node that splits into two daughter nodes (see the Fig. 1 in ref.(Yu et al., 2010)). The splitting criterion is based on the selection of the independent variable which allows the largest reduction in Eq. (3):

[3] The *CART* method includes both classification and regression procedures; in the case of a categorical variable the procedure takes the name of classification; conversely, if continuous, like the energy expenditure (viz. the present case), variable the procedure takes the name of regression trees.

$$SS_P - (SS_L + SS_R) \quad (3)$$

Where $SS_P = \sum(z - \bar{z})^2$ is the sum of the squares of the parent node and SS_L, SS_R are the sum of the squares of the left and right children nodes, respectively. In addition, the splitting is regulated by the complexity parameter (*cp*): at every split R^2_{CART} should increase of, at least, *cp*, R^2_{CART} is defined as follows:

$$R^2_{\mathrm{CART}} = 1 - \underbrace{\frac{\sum_{t \in \tilde{T}} \sum_{i \in t} [z_i - \bar{z}(t)]^2}{\sum_{i=1}^{n} [z_i - \bar{z}]^2}}_{relative\ error} \quad (4)$$

In Eq. (4), $\bar{z}$ is the predicted transportation expenditure, for the terminal node *t,* defined as the mean of the expenditure at that node; the sum at the numerator in Eq. (4) covers all the terminal nodes $t \in \tilde{T}$ (in each of them, all households included are considered). As this term monotonically reduces by increasing the number of slitting, a cross-validation procedure is needed. On the practical point of view, the *CART* method is implemented in three phases, as described in the following and displayed in Figure 8a:

1. **Phase#0.** The dependent variable is selected and the corresponding set of independent variables is selected;
2. **Phase#1.** The *overgrow* tree is generated, by setting the *complexity parameter* (*cp*) at a low value;
3. **Phase#2.** A preliminary pruning procedure is applied to the *overgrow* tree, by selecting the *cp* value corresponding to the minimum cross-validation error (the *V-fold* approach has been applied);
4. **Phase#3.** The tree obtained as the phase#2 is progressively modified and splitting are allowed till R^2_{CART} increase lass than 0.5 % (above this value, there is no further benefit from additional splitting). Also, the minimum number of households in every terminal node should not fall below 100 (< 1% of the data set).

It has been observed that using the same procedure of Besagni and Borgarello (Besagni and Borgarello, 2018) (viz., selecting the regression tree corresponding to the minimum relative error (as obtained in the cross-validation procedure plus the standard error) was not affective in this case, as it would lead towards an *overgrow* tree.

3 Results and discussions

As stated in the introduction, the goal of this paper is to assess the "*socio-demographic and geographical dimensions*" of the transportation expenditure in Italy. To this end, this section mainly answers to the following question: is the "*transport expenditure*" (private, public and total) mainly related to the socio-demographic dimension or to the geographic one? The answer is found in Section 3.1 (when discussing the outcomes of the regression approach) and in Section 3.2 (when discussing the outcomes of the household segmentation). In particular, within Section 2.1, the answer to above-questions are unveiled in Table 2 (private transportation expenditure), Table 3 (public transportation expenditure) and Table 4 (total transportation expenditure).

In these tables, for every variable, the values of the coefficients in Eq. (1), the standard error, the t-test value, the p-value (indicated by Pr(>|t|)), the level of significance and the VIF values are presented; the first row displays the value of the intercept, $\widehat{\beta}_0$, whereas in the subsequent rows, the other coefficients, $\widehat{\beta}_j$, are listed (Eq. (1)). When interpreting these results, it should be noted that, as log-transformed dependent variable are used, interpreting the value of the coefficients is quite straightforward: if we change a certain coefficient (i.e., β_1) by unit, we would expect ϑ to change by $100 \cdot \beta_1$ percent (Longhi, 2015). It is worth noting that all the predictors are characterized by VIF < 3, thus suggesting that the OLS-VIF-LASSO procedure eliminated all the multicollinearity issues.

3.1 The regression approach

In this section, the outcomes of the regression approach are discussed. Firstly, the coefficients of determinations are presented, to provide a global view of the model outcomes. Secondly, the details of the different regression models are resented and commented.

3.1.1 Coefficient of determinations

The coefficient of determination of the three regression models are summarized in Figure 4: (a) private transportation expenditure, R^2_{adj}=28.99 %; (b) public transportation expenditure, R^2_{adj}=9.29 %; (c) total transportation expenditure, R^2_{adj}=36.69 %.These observations suggest that the proposed models are able to explain a small portion of $\theta_{1,2,3}$ variance; nevertheless, the reader should consider that R^2_{adj} represents the proportion of the variance (of the dependent variable) explained by the selected predictors under the linear modeling approach expressed in Eq. (1). For this reason, a low value of R^2_{adj}, along with an high significance of the statistical model, can imply that there is an high relationship between the dependent variable and the predictor, but the dependence is non-linear and/or additional variables should be include

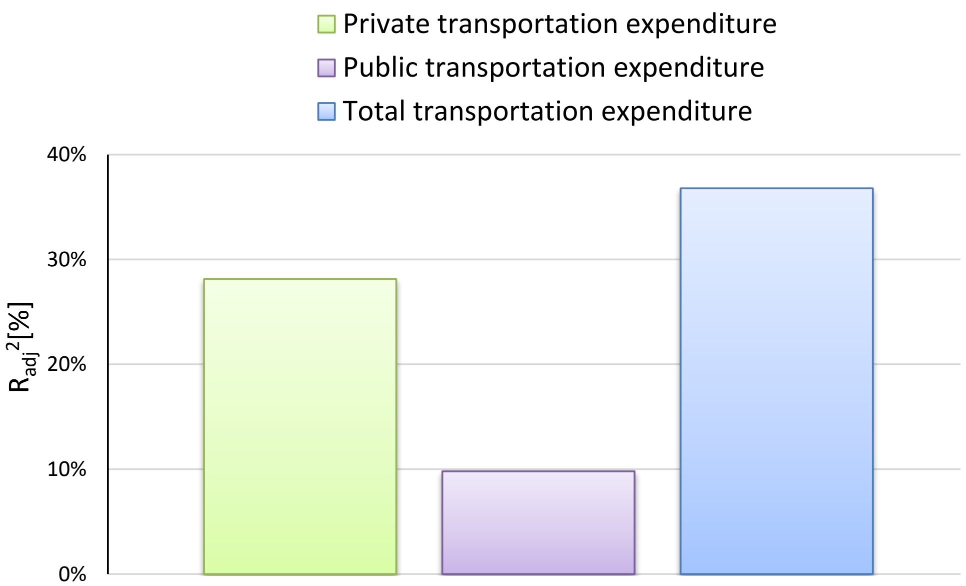


Figure 4. Coefficient of determinations of the proposed regression models (Table 2, Table 3 and Table 4).

Table 2: Details of the final regression model (Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1) – Code names of the variables in Table 1. - Private transportation expenditure

	Estimate	Std. Error	t value	Pr(>|t|)	Significance	VIF
(Intercept) - β_0	4.7258	0.0395	119.7480	< 2e-16	***	1.07
Sex of the *HRP* (b)	-0.1577	0.0151	-10.4640	< 2e-16	***	
Absolute poverty (a)	-0.7241	0.0342	-21.1710	< 2e-16	***	1.04
Qualification of the occupants (a)	-0.3296	0.0862	-3.8230	0.0001	***	1.65
Qualification of the occupants (b)	-0.2251	0.0275	-8.1720	0.0000	***	
Qualification of the occupants (c)	-0.1099	0.0177	-6.2000	0.0000	***	
Qualification of the occupants (e)	0.0196	0.0168	1.1660	0.2435		
Source of income of the occupants (a)	-0.3752	0.0918	-4.0870	0.0000	***	1.39
Source of income of the occupants (b)	-0.2403	0.0478	-5.0310	0.0000	***	
Source of income of the occupants (c)	-0.2437	0.0167	-14.5960	< 2e-16	***	
Workers and similar (b)	-0.0182	0.0152	-1.1960	0.2317		1.22
Workers and similar (c)	0.0684	0.0196	3.4940	0.0005	***	
Type of municipalities (a)	-0.0761	0.0210	-3.6280	0.0003	***	1.11
Type of municipalities (b)	-0.0359	0.0150	-2.4010	0.0164	*	
Geographic location (a)	0.0835	0.0201	4.1490	0.0000	***	1.12
Geographic location (b)	-0.0913	0.0199	-4.5860	0.0000	***	
Geographic location (d)	0.0425	0.0192	2.2070	0.0273	*	
Geographic location (e)	0.0429	0.0331	1.2970	0.1946		
Geographic location (f)	-0.0556	0.0405	-1.3730	0.1699		
Number of cars (b)	0.5428	0.0344	15.7610	< 2e-16	***	1.30
Number of cars (c)	0.9838	0.0363	27.1330	< 2e-16	***	
Number of cars (d)	1.2762	0.0445	28.6680	< 2e-16	***	
Free time expenditures	0.0017	0.0001	15.2000	< 2e-16	***	1.11

Table 3: Details of the final regression model (Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1) – Code names of the variables in Table 1 – Public transportation expenditure

	Estimate	Std. Error	t value	Pr(>|t|)	Significance	VIF
(Intercept) - β_0	3.3183	0.064	51.577	< 2e-16	***	
Absolute poverty (a)	-0.3152	0.098	-3.218	0.001	**	1.07
Enrolment in study courses (b)	-0.2418	0.114	-2.128	0.033	*	1.32
Enrolment in study courses (c)	-0.1982	0.077	-2.590	0.010	**	
Enrolment in study courses (d)	-0.2383	0.074	-3.239	0.001	**	
Enrolment in study courses (e)	0.1162	0.049	2.393	0.017	*	
Enrolment in study courses (f)	0.3626	0.052	6.972	0.000	***	
Type of municipalities (a)	-0.2134	0.044	-4.830	0.000	***	1.19
Type of municipalities (b)	-0.1563	0.042	-3.755	0.000	***	
Geographic location (a)	0.1361	0.052	2.608	0.009	**	1.12
Geographic location (b)	-0.0159	0.052	-0.303	0.762		
Geographic location (d)	-0.3068	0.051	-6.035	0.000	***	
Geographic location (e)	-0.3867	0.122	-3.172	0.002	**	
Geographic location (f)	0.0624	0.132	0.473	0.636		
Number of cars (b)	0.0557	0.052	1.063	0.288		1.41
Number of cars (c)	0.1455	0.059	2.450	0.014	*	
Number of cars (d)	0.2609	0.089	2.922	0.004	**	
Free time expenditures	0.0012	0.000	5.307	0.000	***	1.10

Table 4: Details of the final regression model (Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1) – Code names of the variables in Table 1. – Total transportation expenditure

	Estimate	Std. Error	t value	Pr(>|t|)	Significance	VIF
(Intercept) - β_0	4.2535	0.031	136.866	< 2e-16	***	
Sex of the *HRP* (b)	-0.1788	0.015	-11.655	< 2e-16	***	1.08
Absolute poverty (a)	-0.7146	0.034	-21.137	< 2e-16	***	1.04
Qualification of the occupants (a)	-0.2585	0.084	-3.071	0.002	**	1.92
Qualification of the occupants (b)	-0.2404	0.028	-8.444	< 2e-16	***	
Qualification of the occupants (c)	-0.1004	0.019	-5.400	0.000	***	
Qualification of the occupants (e)	-0.0112	0.018	-0.629	0.529		
Source of income of the occupants (a)	-0.3180	0.094	-3.375	0.001	***	1.41
Source of income of the occupants (b)	-0.2890	0.046	-6.336	0.000	***	
Source of income of the occupants (c)	-0.2708	0.017	-15.854	< 2e-16	***	
Number of managers and employees (b)	0.0421	0.016	2.612	0.009	**	1.55
Number of managers and employees (c)	0.1284	0.022	5.735	0.000	***	
Workers and similar (b)	0.0920	0.024	3.770	0.000	***	1.14
Workers and similar (c)	0.1454	0.065	2.250	0.024	*	
Type of municipalities (a)	-0.0926	0.021	-4.472	0.000	***	1.09
Type of municipalities (b)	-0.0540	0.015	-3.523	0.000	***	
Number of cars (b)	1.0639	0.028	38.473	< 2e-16	***	1.40
Number of cars (c)	1.4630	0.030	48.158	< 2e-16	***	
Number of cars (d)	1.7343	0.041	42.810	< 2e-16	***	
Free time expenditures	0.0019	0.000	16.017	< 2e-16	***	1.10

3.1.2 **Private transportation expenditure**

Table 2 displays the final regression model of the private transportation expenditure (F(22, 12530 = 222.8, p < 2.2e-16), R^2_{adj}=28.99 %). As expected, θ_1 is related to the free time expenditures: an increase in the "*free time expenditure*" equal to 1 € determines an increase ϑ equal to 0.17 %: higher "*free time expenditures*" is likely to determine higher travel intensity, owing to the behavioral/attitude determinants (Acheampong and Cugurullo, 2019; Etminani-Ghasrodashti et al., 2018), Unfortunately, in the employed dataset, there is a lack of attitude patterns and, thus, further insights in behavioral profiles can not be derived. More importantly, the geographic location is significant both in terms of the macro-geographic location and in terms of the type of municipalities, which support the geographical dimension of the "private transport expenditure", in agreement with refs. (Abenoza et al., 2017; Ke and McMullen, 2017). Concerning the influence of municipalities, θ_1 decreases when passing from small municipalities toward the center of metropolitan cities (-7.61 %), possibly owing to the higher availability of public transport system in metropolitan cities; conversely, when passing from small municipalities toward municipalities with more than 50.001 inhabitants (-3.59 %), but less significant. Concerning the influence of macro-geographical location, compared with a household located in the center of Italy, θ_1 increases in the south (+4.25 %) and in the north-west (+8.35 %), whereas it increases in the and in the north-east (-9.13 %); on the other hand, the increases in Sicily and in Sardinia are not significant. It would be interesting to couple the present dataset with more insight regarding economic data in the different regions of Italy to clarify the relationship between transportation expenditures and working conditions/income source by sectors, to clarify the transport-poverty dimension (Ahern et al., 2016; Grieco, 2015). Concerning the socio-demographic variables, it is found that the household structure is not a significant variable, which is in disagreement with previous research activities concerning the relationship between transportation fuel consumption and household variables, as observed in ref. (Büchs and Schnepf, 2013; Clayton et al., 2014; Edwards et al., 2016; King and Scott-Parker, 2016). In this sense, it is worth mentioning that some of the previous literature supported that elderly people tend to use private transportation rather than the public one, which was not observed here (Newbold and Scott, 2017). Conversely, it is has been observed that, in the case the *HRP* person is female, θ_1 decreases by 15.77 %; the importance of the gender on the transportation intensity has been observed also by Arbués et al. (Arbués et al., 2016). In this perspective, some authors stated that men are more likely to travel further compared with women (Basarić et al., 2016; Kawgan-Kagan, 2015; Mahadevia and Advani, 2016; Zheng et al., 2016); conversely, other authors stated that women are more careful towards pro-environmental/pro-sustainability values (Fatma, 2002; O'Connor et al., 1999). Also, the qualification of the occupants is a significant variable: the more the household is "*qualified*" in terms of degree and instruction level, the higher is the "*private transport expenditure*". This result is somehow in agreement with the cluster analysis of of Abenoza et al. (Abenoza et al., 2017) and with the outcomes of Orru et al. (Orru et al., 2019). Considering the previous literature, Pachauri and Jiang (Pachauri and Jiang, 2008) found a relationship between the educational level and the energy consumption. Also, Baiocchi et al. (Baiocchi et al., 2010) found a positive correlation between education level and carbon emissions. It is worth noting that, when considering the relationship between education and the "*private transport expenditure*", there might be a difference regarding the type of education, reflecting in the behavior spectra (i.e., classical courses, engineer courses, …); however, such level of details is missing in the employed dataset and, thus, it is a matter of future studies. In addition, households with poor incomes of with low number of sources of income are more likely to have lower "*private transport expenditures*". This outcome is in agreement with Büchs et al. (Büchs and Schnepf, 2013), who observed that unemployed people tend to have higher public transport emissions compared with households with employed occupants, and with Abenoza et al. (Abenoza et al., 2017), who observed that unemployment people tend to reduce the private transportation use. For example, absolute poverty condition results in a significant decreases of θ_1 (-72.4 %), as this represents of the household income conditions. It has not escaped out notice that this variable can be used as a proxy of the transport poverty condition. Finally, compared with a household with no car, increasing their number increases, as expected, the private transportation expenditure (one car leads +54.38 %, two cars +98.38 %, three or more cars +127.62 %).

3.1.3 Public transportation expenditure

Table 3 displays the final regression model of the public transportation expenditure (F(17, 3560 = 22.8, p < 2.2e-16), R^2_{adj}=9.29 %). As previously observed, θ_2 is related to the free time expenditures: an increase in the "free time expenditure" equal to 1 € determines an increase θ_2 equal to 0.12 %, which is slight lower compared with the outcome for the private public transport expenditure. As stated above, this results is expected, as higher "*free time expenditures*" increases travel intensity. This result also suggests an insight on the behavioral point of view: "free time expenditures" and travel patterns are more likely to be satisfied by a transport expenditure. As mentioned in Section 3.2, also in this case, the geographic location is significant both in terms of the macro-geographic location and in terms of the type of municipalities, thus proving the geographical dimension of transport expenditure, also for the public transport. Concerning the influence of municipalities, θ_2 decreases (with high significance) when passing from small municipalities toward the center of metropolitan cities (-21.34 %) and municipalities with more than 50.001 inhabitants (-15.63 %). It is worth noting that the differences regarding the public transport expenditure observed in large cities might be related to work-related needs (i.e., rural areas might be less connected by public transport), higher congestion of metropolitan areas as well as income-constrains. Indeed, in large cities a considerable share of the household budget might be related to dwelling expenditures, thus leaving a smaller share to afford several cars; conversely, households in a rural areas, where real estate is cheaper, could afford more cars with the same overall budget. Concerning the influence of macro-geographical location, compared with a household located in the center of Italy, θ_2 increases in the north-west (+13.61 %) and in Sardinia (+6.24 %, but no significant), whereas it decreases in the other locations (up to -28.67 %). This outcome is highly related to the availability of infrastructures and additional studies should be conducted by coupling the present dataset with additional data regarding the different regions in Italy. As observed for the private transportation expenditure, the household structure is not significant also for the public transportation expenditure. On the other hand, the enrolment in study courses is a significant variable: the more household components are enrolled in study courses at higher level, the higher is the "*public transport expenditure*" (up to +36.26 %). This results is particular interesting as it suggest that incentives schemes for the public transportation expenditure should be further proposed based on the enrolment in study courses. As expected the absolute poverty condition results in a significant decreases of θ_2 (-31.52 %); however, this value is lower compared with the private transportation expenditure. This observation is in agreement with Arbués et al. (Arbués et al., 2016) stating that higher income households are more likely to use private transportation. In this sense, it is worth mentioning that Büchs et al. (Büchs and Schnepf, 2013) observed that unemployed people tend to have higher public transport emissions compared with households with employed occupants. The number of cars is a significant variables also in this case, but with a lower effect on the transportation expenditure,

3.1.4 Total transportation expenditure

Table 4 displays the final regression model of the total transportation expenditure (F(19, 12967 = 397.2, p < 2.2e-16), R^2_{adj}=36.69 %). Following the previous considerations, also for the total transportation expenditure θ_3 is related to the free time expenditures: an increase in the “free time expenditure” equal to 1 € determines an increase θ_3 equal to 0.197 %. The geographic location is significant in terms of the type of municipalities, but not on the macro-geographical location. θ_3 decreases (with high significance) when passing from small municipalities toward the center of metropolitan cities (-9.3 %) and municipalities with more than 50.001 inhabitants (-5.4 %). Regarding the socio-demographic variables, the qualification of the occupants is a significant variable: the more the household is “*qualified*” in terms of degree and instruction level, the higher is the transport expenditure, in agreement with Orru et al. (Orru et al., 2019). In addition, in the case the HRP person is female, θ_3 decreases by 17.88 %. As expected, the income-related variables are significant. First, the absolute poverty condition results in a significant decreases of θ_3 (-71.46 %), as this is a representation of the household income conditions. As mentioned previously, this variable can be used as a proxy for a transport poverty condition. Second, the higher the number of workers and managers or the number of employers, the higher is the transportation expenditure. Finally, households with poor incomes of with low number of sources of income, are more likely to have lower transport expenditures (i.e., when there is no source of in income, the transportation expenditure is in the range of -17/-31 %). In addition, regarding the household equipment, increasing their number increases, as expected, the private transportation expenditure (one car leads +106.39 %, two cars +146.30 %, three or more cars +173.43 %).

3.2 Household segmentations

The results of the household segmentation, implemented using as input variables the predictors derived in Section 3.1, are presented in Figure 5 (13 clusters - θ_1), in Figure 6 (13 clusters - θ_2) and in Figure 7 (14 clusters - θ_3). It is worth noting that, in the household segmentations, both socio-demographic and geographical variables have been applied. Thus, these segmentations clarify the combination of factors that influence the expenditures patterns in the Italian households and they reveal the hidden information that cannot be observed by looking at the regression study alone (i.e., some non-linear dependencies). As for θ_1, the first splitting concern the number of cars owned and the second splitting concerns the free time expenditure, thus supporting the importance of the behavioral spectrum. It is observed that, in households with a lower number of cars (zero or one), income-related variables (i.e., absolute poverty or the source of income) define the prevailing household segmentation. Conversely, in households with a higher number of cars (two or more), the behavior spectra (i.e., free time expenditure) and the geographic locations define the prevailing household segmentation. These results suggest that policy-schemes deployed for the private transportation needs to be de-coupled based on the number of cars own, as this variable conceals two different dimension: the former related to income and the latter related to the behavioral and demographic dimensions. This outcome is of practical relevance in forthcoming studies devoted to the large-scale deployment of electrical vehicles and their use in different households. As for θ_2, the first splitting concerns the enrollment in study courses and, secondly, the geographical location and the free time expenditure. In general this segmentation reveals that the public transportation expenditure mainly depends on the mobility requirement to reach study locations; subsequently, the type of transportation is based on the available infrastructures (i.e., in terms of macro-region and micro-locations). In particular, when household components are involved in higher education, the corresponding expenditure is higher (owing to higher mobility requirement); on the other hand, if household components are not involved in study courses, the corresponding expenditure is lower. Finally, the total energy expenditure segmentation is the consequence of the two above-mentioned segmentation trees and can be interpreted as the results of the coupling between transportation demand/availability and household behavior attitudes.

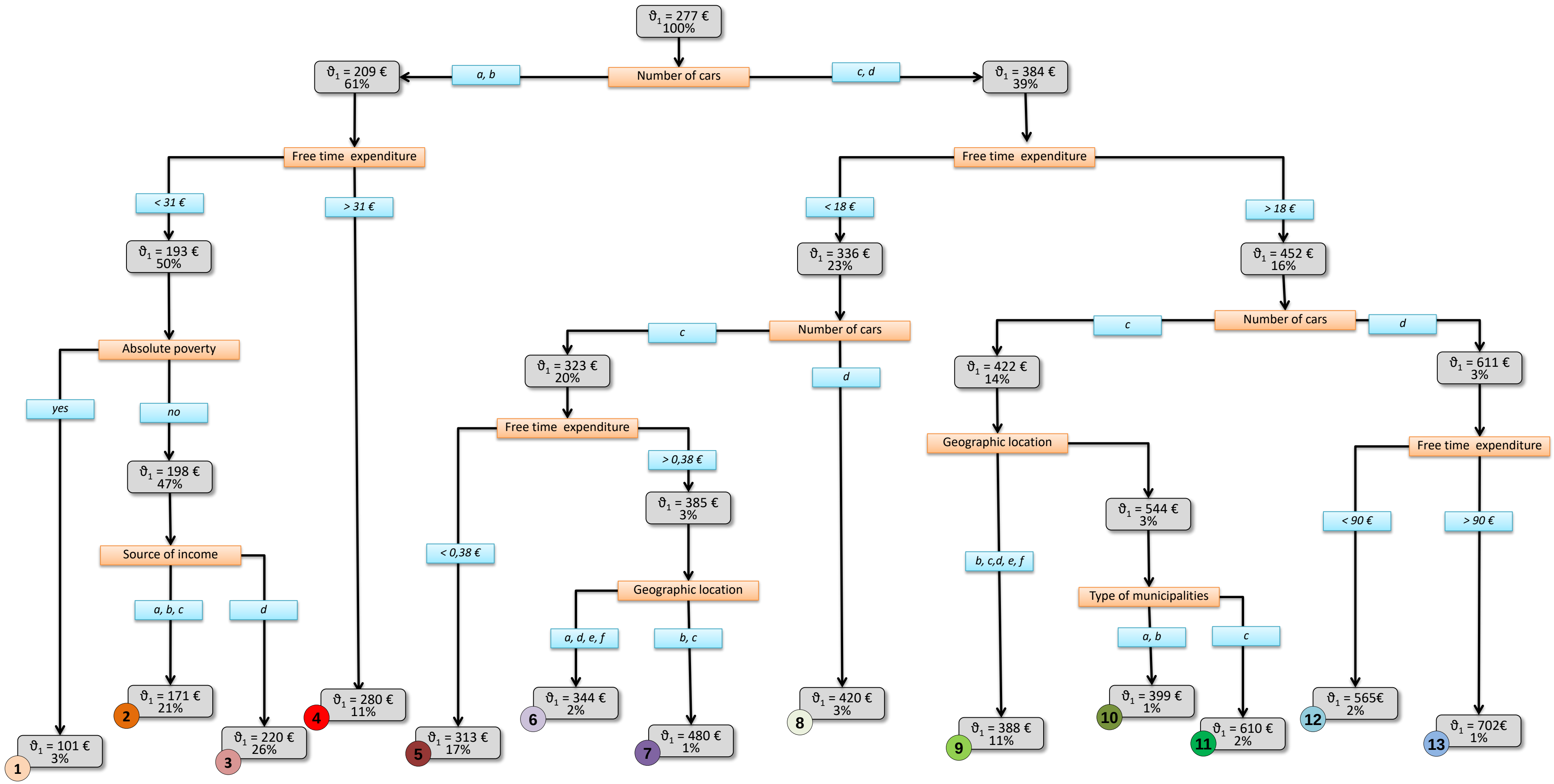


Figure 5. Household segmentation - θ_1.

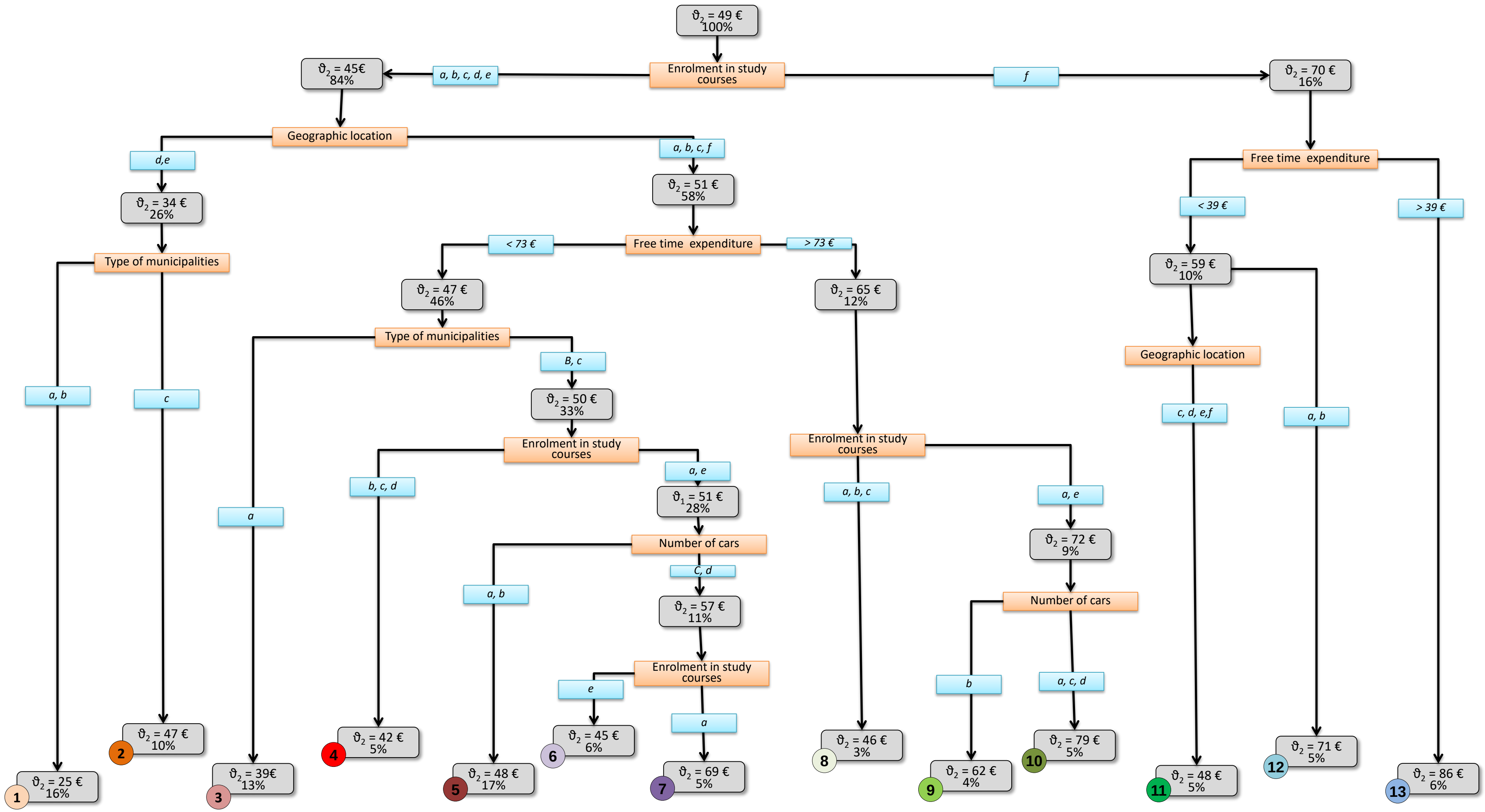


Figure 6. Household segmentation – θ_2.

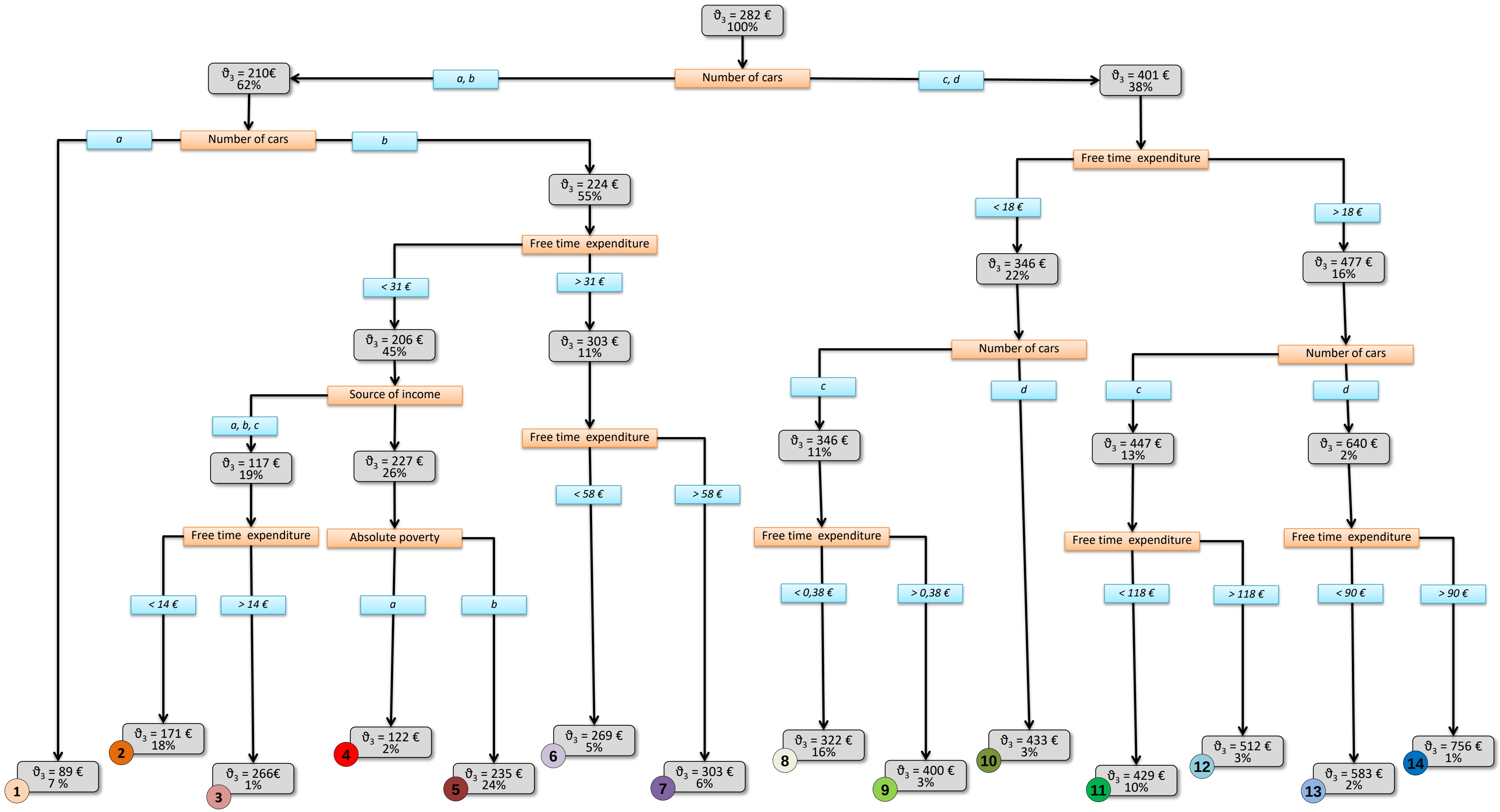


Figure 7. Household segmentation - θ_3.

4 Conclusions, outcomes and outlooks

This paper contributed to the existing discussion regarding the factors influencing the transportation expenditure, thus contributing to the "*human dimension*" of the energy-intensity in transportation, in order to provide a rational basis to evaluate the subsequent energy metabolism. In particular, this paper focuses on the Italian case studies and it evaluates the "*socio-demographic and geographical dimensions*" of the transportation expenditures. It is found that the geographic location is significant both in terms of the macro-geographic location and in terms of the type of municipalities for the private and public transport expenditure. This result is of practical importance in forecasting model for the transportation sector, which should consider also cross-migration effects within the country and from/towards the cities. Conversely, it is found that the "*socio-demographic dimension*" is determined by income-related and behaviour-related variables (which are also related to the main occupation of the household components) rather than the household-composition variables. This result is of practical importance in forecasting model for the transportation sector, which should consider also evolution concerning working types and conditions within the socio-demographic layer. As a consequence of the occupation variables, the obtained results support that aging of the population will result in a decrease of the household fuel use in Italy. In addition, it is found that the private transportation patterns can be differentiated based on the number of cars, as this variable conceals two different dimension: the former related to income and the latter related to the behavioral and demographic dimensions. Instead, the public transportation patterns depend on the requirement to reach study locations and, secondly, on the available infrastructures (i.e., in terms of macro-region and micro-locations).

All above considerations (i.e., the relationship between the variables and the household segmentation) are of key interest to policymakers when planning investments aiming at reducing the primary energy consumption at the household level (in the view of pathways towards emission reduction and "*decarbonisation*" (Sovacool et al., 2018)), by forecasting the transportation energy consumption. The outcomes of this study are also of practical interest in terms of environment policies, as they will provide a statistical basis to estimate the impact of "*country-scale*" transportation policies at the household-scale. Future studies may be as follows:

- the obtained relationships and observations may serve as bottom-layer for studies regarding forecasting of the transportation expenditure at the country-scale as well as a rational basis to explore the transport poverty in Italy (Grieco, 2015);
- the regression models can be coupled with forecast of changes in socio-demographic variables, to predict the primary energy consumption of the transportation sector in the long-term within the so-called decarbonisation pathways (Sofia et al., 2019);
- relate the transportation patterns and health conditions, as a follow-up of the study proposed by Singleton (Singleton, 2018).
- preform measures of the energy intensity of the transportation sector in the different households identified in the segmentation procedure, in order to provide a further insight in the behavior of the different household segmentation;
- couple the present dataset with additional datasets concerning climatic data (i.e., temperature, humidity, heating and cooling degree days, …), to better describe the boundary conditions of the demand side of the transportation;
- elaborate on the "*psychological variables*" and their role in transportation patterns (Abrahamse and Steg, 2009);
- couple the present results with simplified lumped parameter model of the transportation technologies in order to provide a comprehensive model of the transportation sector, to improve the discussion of the decarburization pathways);

5 References

Abenoza, R.F., Cats, O., Susilo, Y.O., 2017. Travel satisfaction with public transport: Determinants, user classes, regional disparities and their evolution. Transportation Research Part A: Policy and Practice 95, 64-84.

Abrahamse, W., Steg, L., 2009. How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings? Journal of Economic Psychology 30, 711-720.

Acheampong, R.A., Cugurullo, F., 2019. Capturing the behavioural determinants behind the adoption of autonomous vehicles: Conceptual frameworks and measurement models to predict public transport, sharing and ownership trends of self-driving cars. Transportation Research Part F: Traffic Psychology and Behaviour 62, 349-375.

Ahern, A., Vega, A., Caulfield, B., 2016. Deprivation and access to work in Dublin City: The impact of transport disadvantage. Research in Transportation Economics 57, 44-52.

Anciaes, P., Jones, P., 2020. Transport policy for liveability – Valuing the impacts on movement, place, and society. Transportation Research Part A: Policy and Practice 132, 157-173.

Arbués, P., Baños, J.F., Mayor, M., Suárez, P., 2016. Determinants of ground transport modal choice in long-distance trips in Spain. Transportation Research Part A: Policy and Practice 84, 131-143.

Baiocchi, G., Minx, J., Hubacek, K., 2010. The Impact of Social Factors and Consumer Behavior on Carbon Dioxide Emissions in the United Kingdom. Journal of Industrial Ecology 14, 50-72.

Banister, D., 2008. The sustainable mobility paradigm. Transport Policy 15, 73-80.

Bardazzi, R., Pazienza, M.G., 2018. Ageing and private transport fuel expenditure: Do generations matter? Energy Policy 117, 396-405.

Basarić, V., Vujičić, A., Simić, J.M., Bogdanović, V., Saulić, N., 2016. Gender and Age Differences in the Travel Behavior – A Novi Sad Case Study. Transportation Research Procedia 14, 4324-4333.

Ben-Salha, O., Hkiri, B., Aloui, C., 2018. Sectoral energy consumption by source and output in the U.S.: New evidence from wavelet-based approach. Energy Economics 72, 75-96.

Besagni, G., Borgarello, M., 2018. The determinants of residential energy expenditure in Italy. Energy 165, 369-386.

Besagni, G., Inzoli, F., Zieghenein, T., Lucas, D., 2019. Experimental study of liquid velocity profiles in large-scale bubble columns with particle tracking velocimetry. Journal of Physics: Conference Series 1224, 012036.

Brand, C., Goodman, A., Rutter, H., Song, Y., Ogilvie, D., 2013. Associations of individual, household and environmental characteristics with carbon dioxide emissions from motorised passenger travel. Applied Energy 104, 158-169.

Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. Classification and regression trees. CRC press.

Büchs, M., Schnepf, S.V., 2013. Who emits most? Associations between socio-economic factors and UK households' home energy, transport, indirect and total CO2 emissions. Ecological Economics 90, 114-123.

Clayton, W., Ben-Elia, E., Parkhurst, G., Ricci, M., 2014. Where to park? A behavioural comparison of bus Park and Ride and city centre car park usage in Bath, UK. Journal of Transport Geography 36, 124-133.

Dias, L.P., Simões, S., Gouveia, J.P., Seixas, J., 2019. City energy modelling - Optimising local low carbon transitions with household budget constraints. Energy Strategy Reviews 26, 100387.

Edwards, S.J., Emmerson, C., Namdeo, A., Blythe, P.T., Guo, W., 2016. Optimising landmark-based route guidance for older drivers. Transportation Research Part F: Traffic Psychology and Behaviour 43, 225-237.

Etminani-Ghasrodashti, R., Paydar, M., Hamidi, S., 2018. University-related travel behavior: Young adults' decision-making in Iran. Sustainable Cities and Society 43, 495-508.

European Commission, S.-G., 2011. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A Roadmap for Moving to a Competitive Low Carbon Economy in 2050 COM(2011) 112 final.

Fatma, D., 2002. Climate Change Vulnerability, Impacts, and Adaptation: Why Does Gender Matter? Gender and Development 10, 10-20.

Grieco, M., 2015. Poverty mapping and sustainable transport: A neglected dimension. Research in Transportation Economics 51, 3-9.

ISTAT, 2017. Household Budget Survey: microdata for research purposes, reference year: 2015, Rome, Italy.

Kawgan-Kagan, I., 2015. Early adopters of carsharing with and without BEVs with respect to gender preferences. European Transport Research Review 7, 33.

Ke, Y., McMullen, B.S., 2017. Regional differences in the determinants of Oregon VMT. Research in Transportation Economics 62, 2-10.

King, M.J., Scott-Parker, B.J., 2016. Older male and female drivers in car-dependent settings: how much do they use other modes, and do they compensate for reduced driving to maintain mobility? Ageing and Society 37, 1249-1267.

Liddle, B., 2014. Impact of population, age structure, and urbanization on carbon emissions/energy consumption: evidence from macro-level, cross-country analyses. Population and Environment 35, 286-304.

Longhi, S., 2015. Residential energy expenditures and the relevance of changes in household circumstances. Energy Economics 49, 440-450.

Lowe, C., Stanley, J., Stanley, J., 2018. A broader perspective on social outcomes in transport. Research in Transportation Economics 69, 482-488.

Mahadevia, D., Advani, D., 2016. Gender differentials in travel pattern – The case of a mid-sized city, Rajkot, India. Transportation Research Part D: Transport and Environment 44, 292-302.

Miller, P., de Barros, A.G., Kattan, L., Wirasinghe, S.C., 2016. Public transportation and sustainability: A review. KSCE Journal of Civil Engineering 20, 1076-1083.

Newbold, K.B., Scott, D.M., 2017. Driving over the life course: The automobility of Canada's Millennial, Generation X, Baby Boomer and Greatest Generations. Travel Behaviour and Society 6, 57-63.

O'Connor, R.E., Bard, R.J., Fisher, A., 1999. Risk Perceptions, General Environmental Beliefs, and Willingness to Address Climate Change. Risk Analysis 19, 461-471.

O'Neill, B.C., Liddle, B., Jiang, L., Smith, K.R., Pachauri, S., Dalton, M., Fuchs, R., 2012. Demographic change and carbon dioxide emissions. The Lancet 380, 157-164.

Okada, A., 2012. Is an increased elderly population related to decreased CO2 emissions from road transportation? Energy Policy 45, 286-292.

Orru, K., Poom, A., Nordlund, A., 2019. Socio-structural and psychological factors behind car use: Comparing Northern and Eastern Europe. Transportation Research Part A: Policy and Practice 119, 313-325.

Pachauri, S., Jiang, L., 2008. The household energy transition in India and China. Energy Policy 36, 4022-4035.

Sajid, M.J., Cao, Q., Kang, W., 2019. Transport sector carbon linkages of EU's top seven emitters. Transport Policy 80, 24-38.

Singleton, P.A., 2018. Walking (and cycling) to well-being: Modal and other determinants of subjective well-being during the commute. Travel Behaviour and Society.

Sofia, D., Gioiella, F., Lotrecchiano, N., Giuliano, A., 2019. Cost-benefit analysis to support decarbonization scenario for 2030: A case study in Italy. Energy Policy, 111137.

Soltani, A., Pojani, D., Askari, S., Masoumi, H.E., 2018. Socio-demographic and built environment determinants of car use among older adults in Iran. Journal of Transport Geography 68, 109-117.

Sovacool, B.K., Kester, J., Noel, L., de Rubens, G.Z., 2018. The demographics of decarbonizing transport: The influence of gender, education, occupation, age, and household size on electric mobility preferences in the Nordic region. Global Environmental Change 52, 86-100.

Stephenson, J., Barton, B., Carrington, G., Doering, A., Ford, R., Hopkins, D., Lawson, R., McCarthy, A., Rees, D., Scott, M., Thorsnes, P., Walton, S., Williams, J., Wooliscroft, B., 2015. The energy cultures framework: Exploring the role of norms, practices and material culture in shaping energy behaviour in New Zealand. Energy Research & Social Science 7, 117-123.

Tapio, P., Banister, D., Luukkanen, J., Vehmas, J., Willamo, R., 2007. Energy and transport in comparison: Immaterialisation, dematerialisation and decarbonisation in the EU15 between 1970 and 2000. Energy Policy 35, 433-451.

Tian, X., Geng, Y., Dai, H., Fujita, T., Wu, R., Liu, Z., Masui, T., Yang, X., 2016. The effects of household consumption pattern on regional development: A case study of Shanghai. Energy 103, 49-60.

Torgler, B., Garcia-Valiñas, M.A., Macintyre, A., 2008. Differences in preferences towards the environment: The impact of a gender, age and parental effect. FEEM Working Paper No. 18.2008.

Ventura, J.A., Kweon, S.J., Hwang, S.W., Tormay, M., Li, C., 2017. Energy policy considerations in the design of an alternative-fuel refueling infrastructure to reduce GHG emissions on a transportation network. Energy Policy 111, 427-439.

Yu, Z., Haghighat, F., Fung, B.C.M., Yoshino, H., 2010. A decision tree method for building energy demand modeling. Energy and Buildings 42, 1637-1646.

Zawieska, J., Pieriegud, J., 2018. Smart city as a tool for sustainable mobility and transport decarbonisation. Transport Policy 63, 39-50.

Zheng, Z., Washington, S., Hyland, P., Sloan, K., Liu, Y., 2016. Preference heterogeneity in mode choice based on a nationwide survey with a focus on urban rail. Transportation Research Part A: Policy and Practice 91, 178-194.

Acknowledgments

This work has been financed by the Research Fund for the Italian Electrical System in compliance with the Decree of Minister of Economic Development April 16, 2018.