CHEMICAL ENGINEERING TRANSACTIONS

VOL. 76, 2019

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Petar S. Varbanov, Timothy G. Walmsley, Jiří J. Klemeš, Panos Seferlis

ISBN 978-88-95608-73-0; **ISSN** 2283-9216

DOI: 10.3303/CET1976180

On the Relationship between Private Transportation Expenditure and Socio-Demographic Variables

Giorgio Besagni*, Corine Nsangwe Businge, Marco Borgarello

Ricerca sul Sistema Energetico - RSE S.p.A., Power System Development Department, via Rubattino 54, 20134 Milano, Italy
giorgio.besagni@rse-web.it

A precise understanding of the relationships between household characteristics and transportation expenditures is of paramount importance to support bottom-up policies. Despite the considerable amount of research activities carried out in the last decades, an agreement regarding the determinants of the transportation expenditure is far from being achieved. This paper contributes to the existing discussion, focusing on the Italian case study, by analysing the relationships between the private transportation expenditure and the socio-demographic variables. In particular, the impact that the different household characteristics have on the private transportation expenditures have been examined. The analysis is performed by coupling three statistical methods: the ordinary least squares method, the variance inflation factor, the least absolute shrinkage and selection operator. It is found that the geographic area (in terms of the macro-scale and the micro-scale geographic locations) as well as income-related variables are likely to be determinants for the private transportation expenditure.

1. Introduction

The so-called “sustainable mobility paradigm”, as defined by Banister (2008), is a cutting-edge research topic and a priority in the international agenda. Owing to the large share of primary energy consumption and emissions related to the transportation sector, a considerable amount of research activities is conducted to support pathways towards emission reduction and “decarbonisation” (viz., reduction in greenhouse gas emissions). Out of the many studies, Van Fan et al. (2018) discussed sustainable alternatives for freight transportation and Promjittiphong et al. (2018) analyses the energy consumption of the tourist-sector. To reach these targets, effective policies are needed: their success relies on the precise understanding of the relationships between the household characteristics and the transportation expenditure/fuel consumption. The study of these relationships belongs to the field of the so-called “human dimension” of the energy-intensity in transportation along with the subsequent “energy metabolism” (Stephenson et al, 2015). This concept was addressed by 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 (2015) mentioned that the precise understanding of the determinants of the energy consumption/expenditure serves as basis for policymakers when planning investments aiming to reduce the energy consumption at the bottom-level (Longhi, 2015). This paper contributes to the present-day discussion, by investigating the transportation sector and the existing relationships between the “private transport expenditure” and both 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 locations). This context is interesting if considering the demographic shift experienced by European countries. As the population is progressively ageing, this may reflect in the energy metabolism at the different levels: for example, both Liddle (2014) and Brand et al. (2013) mentioned that the ageing of population is likely to increase the residential energy consumption and to reduce the transport-related energy use. This topic was further investigated by 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. Unfortunately, despite the considerable amount of research activities carried out in the last decades, an agreement is far from being achieved on both the methodological point of view and the relationships

Paper Received: 22/01/2019; Revised: 28/03/2019; Accepted: 28/03/2019
Please cite this article as: Besagni G., Businge C.N., Borgarello M., 2019, On the Relationship between Private Transportation Expenditure and Socio-Demographic Variables, Chemical Engineering Transactions, 76, 1075-1080 DOI:10.3303/CET1976180

between the transportation expenditure and the many variables describing households. This paper contributes to the existing discussion, focusing on the Italian case study, by analysing the relationships between the private transportation expenditure and the household variables. In particular, the impact that the different household characteristics have on the transportation expenditures have been examined, based on the microdata taken from the Italian Household Budget Survey published by the Italian Statistical Office. The analysis is performed by coupling three methods: the ordinary least squares method, the variance inflation factor, the least absolute shrinkage and selection operator, to select relevant variable. It is found that the geographic area as well as income-related variables are likely to be determinants for the private and public transportation expenditures.
The paper 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

In this section, the dataset (Section 2.1), the dependant variable (Section 2.2), the predictors (Section 2.3) and the statistical methods (Section 2.4) are discussed and commented.

2.1 The dataset

The dataset used in this study 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 were collected from 15,015 households, in 502 different municipalities; for each household more than 1264 variables are available, regarding socio-demographic information, dwelling characteristics, appliances and monthly expenditures.

2.2 The dependent variables

The dependent variable “private transport expenditure” (coded as ϑ, in the followings) was obtained by summing different monthly 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. As mentioned above, the data are available in the dataset in terms of monthly expenditures and refer to a precise month during the year. Before going ahead with the analysis, it should be verified whether an annual calibration is needed or not (Besagni and Borgarello, 2018). To this end, Figure 1 displays the relationship between the time variable (month) and the “private transport expenditure”: the dependent variable does not exhibit high dependency with respect to the time variable. Hence, an annual calibration is not needed. As the dependent variable has left-skewed distribution, it is implemented as log-transformed. Another advantage of using logs is that the regression coefficients refer to the relative changes rather than the absolute changes in per-capita energy expenditures. This concept was described by Longhi (2015) and by Besagni and Borgarello (2018). Households having “private transport expenditure” equal to zero were excluded from the analysis.

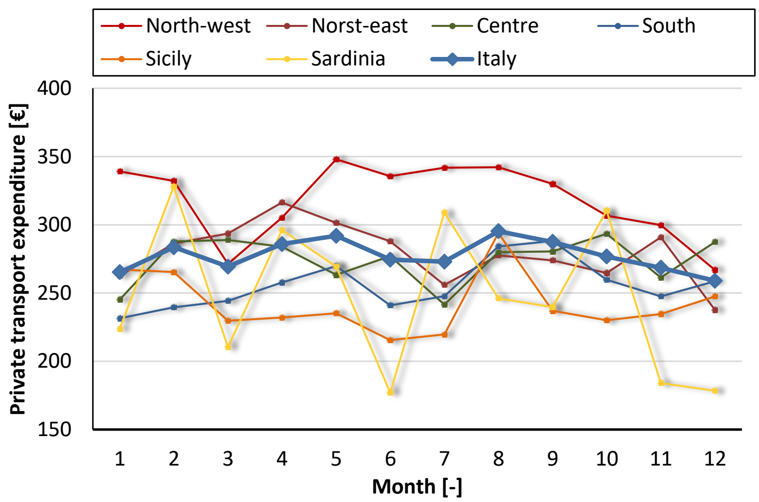

Figure 1: Relationship between time of the year and “private transport expenditure

2.3 The predictors

Table 1 lists the variables used as predictors in the regression analysis, with their frequencies and summary statistics (viz., mean and standard deviation for the continuous variables - the category printed in bold indicates the reference category). In Table 1, HRP represents the Household Representative Person, which is the individual that is taken to represent the household. In this study, HRP describes the highest income earner in the household Summary statistics evaluated on the whole data-set.

2.4 The statistical methods

The statistical approach couples three methods: (a) the ordinary least squares method (OLS), (b) the variance inflator factor (VIF), (c) the Least absolute shrinkage and selection operator (LASSO). A similar procedure has been implemented by Besagni and Borgarello (2018). The procedure con consists of four phases, as discussed in the following:

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

$$y_i = \ln\ (\vartheta) = \beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\ldots + \beta_k x_{ik}+ \varepsilon_i = \beta_0+ \sum_{j=1}^{w} \beta_j x_{ij} + \varepsilon_i \quad (1)$$

  In Eq. (1), y_i is the logarithm of the dependent variable 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. The model fit is evaluated on the well-known adjusted coefficient of determination (R^2_{adj}). After the OLS analysis, owing to possibility of multicollinearity issues, the variance-inflation factors (VIF_j) have been inspected for every β_j:

$$VIF_j = \frac{1}{1-R_j^2} \quad (2)$$

  If the predictors are uncorrelated, VIF = 1; conversely, if VIF > 1, a possible correlation exist. At present, there is no agreement on the cut-off point for critical values of VIF (VIF_{max}). In this study, VIF_{max} = 3, has been selected, accordingly with Besagni and Borgarello (2018).
- Phase#b. If multicollinearity has been detected in phase#a, LASSO regression, a variable shrinkage based on a penalty, is applied to select relevant predictors by using a constrained optimization problem Once the significant variables have been selected, the OLS is repeated and another R^2_{adj} is obtained; subsequently, VIF are inspected again.
- Phase#c. If multicollinearity is detected after LASSO procedure, a progressive selection of predictors is conducted, by excluding, one by one, the least significant predictors and, at each step, the changes in R^2_{adj} are inspected and VIF is monitored (a reduction in R^2_{adj} is considered acceptance only if within 0.5 % and, in this phase VIF_{max} = 10).
- Phase#d. If no multicollinearity has been detected after LASSO, a selection of the predictors is obtained thorough the recursive procedure described in phase#c (a reduction of R^2_{adj} is considered acceptable if within 0.2 %).

The interested reader may refer to our previous study for additional details regarding the employed methods and modelling approaches (Besagni and Borgarello, 2018).

3. Results and discussions

As stated in the introduction, the main goal of this section is to assess the “socio-demographic and geographical dimensions” of “private transport expenditure” (ϑ); in particular this section tries to answer to the following question: is the “private transport expenditure” mainly related to the socio-demographic aspects or to the geographic ones? In order to answer this question, Table 2 displays the final regression model (F(22, 12530 = 222.8, p < 2.2e-16), R^2_{adj}=28.99 %). In Table 2, 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|)) and the level of significance are presented; the first row displays the value of the intercept, $\hat{\beta}_0$, whereas in the subsequent rows, the other coefficients, $\hat{\beta}_j$, 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. Table 2 also presents the results of the VIF analysis: all the predictors are characterized by VIF < 3, thus suggesting that the OLS-VIF-LASSO procedure eliminated multicollinearity issues.

Table 1: Socio-demographic variables with their frequencies. (= reference category).*

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 citizen [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 disabled people	(a) Yes [100], (b) No [14,913]*
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	(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	Continuous variable [Mean = 22.68 / Variance = 3027]

*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.*

	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	***	
Qualification of the occupants (b)	-0.2251	0.0275	-8.1720	0.0000	***	1.65
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	***	
Source of income of the occupants (b)	-0.2403	0.0478	-5.0310	0.0000	***	1.39
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	***	
Geographic location (b)	-0.0913	0.0199	-4.5860	0.0000	***	
Geographic location (d)	0.0425	0.0192	2.2070	0.0273	*	1.12
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	***	
Number of cars (c)	0.9838	0.0363	27.1330	< 2e-16	***	1.30
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

The proposed model exhibits R^2_{adj}=28.99 %; this result suggests that the proposed model is able to explain quite a small portion of ϑ 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 a 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. Looking at the results in Table 2, it is observed that ϑ 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 %; this result is somehow expected, as higher "free time expenditures" is likely to determine higher travel intensity. More importantly, it is observed that 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 "private transport expenditure". Concerning the influence of municipalities, ϑ 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 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 (see the discussion proposed in the introduction). 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". Considering the previous literature, Pachauri and Jiang (2008) found a relationship between the educational level and the energy consumption. Similarly, households with poor incomes of with low number of sources of income, are more likely to have lower "private transport expenditures". It is also observed that, in the case the HRP person is female, ϑ decreases by 15.77 %. As expected the absolute poverty condition results in a significant decreases of ϑ (-72.4 %), as this is a representation of the household income conditions. In summary, the proposed results, support that both a "socio-demographic dimension" and a "geographical dimensions" of the "private transport expenditure" exist, even if the "socio-demographic dimension" is determined by income-related variables rather than the household-composition variables.

4. Conclusions

Using a nationally representative sample of Italian households, this paper contributes to the existing discussion regarding "human dimension" of the energy-intensity in transportation, 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 "private transport expenditure". 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, thus supporting the geographical dimension of "private transport expenditure". On one hand, it is found that the "socio-demographic dimension" is determined by income-related variables (which are also related to the main occupation of the household components) rather than the household-composition variables; the household structure is not a significant predictor, in disagreement with the previous literature. These results might support, even if following a different path, the discussion of Bardazzi and Pazienza (2018): aging of the population will result in a decrease of household fuel use in Italy in the long-run. This consideration may be of key interest to policymakers when planning investments aiming to reduce the primary energy consumption at the household level (in the view of pathways towards emission reduction and "decarbonisation"), 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. In particular, these results will help forecasting the practical response of the household towards macro transportation policies.

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.

References

Banister D., 2008, The sustainable mobility paradigm, Transport Policy, 15(2), 73-80.

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

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

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.

Fan Y.V., Klemes J.J., Perry S., Lee C.T., 2018, An Emissions Analysis for Environmentally Sustainable Freight Transportation Modes: Distance and Capacity, Chemical Engineering Transactions, 70, 505-510.

ISTAT, 2017, Household Budget Survey: Microdata for Research Purposes (Reference Year: 2015); Rome, Italy.

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(3), 286-304.

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

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

Promjittiphong C., Junead J., Hanpattanakit P., 2018, Greenhouse Gas Emission and Mitigation from Sports Tourism in Benja Burapha Cycling Rally, Sa Kaeo, Thailand. Chemical Engineering Transactions, 63, 397-402.

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 and Social Science, 7 117-123.

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.