Transportation (2018) 45:1755–1781
https://doi.org/10.1007/s11116-017-9794-y

Mode choice and travel distance joint models in school trips

Alireza Ermagun[1] · Amir Samimi[2]

Published online: 4 July 2017

Abstract We test a copula-based joint discrete–continuous model to unravel mode choice and travel distance decisions in a joint framework for school trips. This framework explicitly accounts for common unobserved factors that may affect both the mode choice and travel distance. Joint estimation of the models makes a significant difference in the effect of travel distance on willingness to walk to school. The absolute value of the travel distance coefficient in the mode choice model increases by 22% when a joint formulation is adopted instead of the conventional single estimations. We find a significant decrease of 19% in the coefficient of travel safety perception in the joint mode choice model compared to the single model. This underscores the impact of model specification, in terms of the variable effect interpretation and policy assessments. The effect magnitude of several policy-sensitive variables is discussed and compared with previous studies. Particularly, we indicate that the probability of walking is reduced by 0.85% due to a 1% increase in travel distance; accordingly, it propels parents to select non-active modes, particularly school bus. This study also demonstrates how addressing parental concerns about travel safety could double the propensity to walk to school.

Keywords Copula-based model · Mode choice · School siting · School trips · Safety

✉ Alireza Ermagun
alireza.ermagun@northwestern.edu

Amir Samimi
asamimi@sharif.edu

1 McCormick School of Engineering and Applied Science, Northwestern University, 2145 Sheridan Rd, Evanston, IL 60208, USA

2 Department of Civil Engineering, Sharif University of Technology, Azadi Avenue, Tehran, Iran

Introduction

Developed and developing countries have witnessed an increase in the use of private cars in school trips, which has resulted in a diminution in the share of active transportation modes (such as walking and biking). According to the Department for Transport in Great Britain, 55% of all secondary school children walked to school in 1976, while this share dropped to 38% by 2012 (Easton and Ferrari 2015). Surveys conducted in Australia also revealed that the share of children's active and independent travelling to school declined from 61% to 32% between 1991 and 2012 (Schoeppe et al. 2015). Such behavioral shifts have adverse consequences that include: physical inactivity among students (Cooper et al. 2005), increasing costs of obesity (Sepulveda et al. 2010), traffic jams at peak hours (Pedestrian and Bicycle Information Center 2008), and adverse environmental impacts (Wilson et al. 2007). Hence, detecting factors that affect this change is crucial for controlling this trend and promoting active modes of travel.

Among the effective factors, travel distance is found to be very influential, if not the most influential, on the choice of school travel mode (Seraj et al. 2012). All the previous studies found a negative correlation between home-to-school distance and students' propensity to walk to school (Ermagun and Levinson 2017). Therefore, two major research venues, namely school travel mode and home-to-school distance, have been developed to find effective policies that may encourage active modes among children. There are two dominant viewpoints in modeling travel distance and travel mode. First, both decisions are considered to be made jointly by the parents. This school of thought acknowledges when parents are deciding on which school to register their children in, they also consider different travel options for them. For example, parents who are more worried about the safety of their children may choose a school that either has a safe walkway to home or is easily accessible by their private car. Therefore, this point of view strongly supports a joint formulation for travel distance and mode. Second, home-to-school distance is part of very complex choices that could be tightened with residential location choice and other long-term decisions as well. This school of thought mentions when parents attempt to enact a school choice other than their current close school, they might even end up taking on a change in residential location to move into the desired school district. Therefore, travel distance is not actually a choice, but rather a consequence of school choice. According to this viewpoint, travel distance is an exogenous variable in the mode choice model, in which a joint formulation for the modes may be undermined.

This study argues for the significance of a joint formulation for school travel mode and distance, no matter which viewpoint is followed. We test a copula-based joint multinomial discrete–continuous model framework to account explicitly for common unobserved factors that may affect both the mode choice and travel distance decisions. We also compare a joint copula model for travel mode and distance with single distance (linear regression) and travel mode (multinomial logit) models. This analysis is conducted in Tehran, with more than 5300 schools and 1.1 million registered k-12 students (Statistics of Minister of Education 2012). Absence of such studies in the Middle East with arguably different culture, attitudes, and preferences, further motivated this research.

The remainder of this study is structured as follows. First, a critical review of the literature is provided. Then, the modeling method is explained, followed by a description of the empirical data. The estimation results of the joint model is proposed and compared to the single models. Finally, the paper ends with policy recommendations and concluding remarks.

Background

Many studies (Yarlagadda and Srinivasan 2008; McDonald and Aalborg 2009; Ermagun et al. 2015a) have analyzed factors that influence travel behavior in school trips. Table 1 shows some of the past studies with a wide diversity in geographic contexts, time spans, and analysis methods. Past studies typically used statistical description (McDonald and Aalborg 2009) or a binary logit model specification (McMillan 2007) to model travel mode. Further, a few studies (Wilson et al. 2010; McDonald 2008) have used multinomial logit formulation to examine a broader range of modes. Independent from irrelevant alternatives (IIA), a restrictive property of multinomial logit formulation on choices (Train 2009) is remedied, to an extent, in very few studies (Ermagun and Samimi 2015).

Factors affecting the choice of travel mode in school trips can be categorized in four classes: (1) student characteristics, (2) household demographics, (3) urban design and built environment, and (4) parental reservations. As shown in Table 2, some parameters are found to have contradictory effects on school trip modal selection. This difference could be caused by non-uniformity of samples and difference of surrounding conditions. For example, age of students could change certain travel behaviors, and thus, paradoxical conclusions may be reached in studies that target dissimilar age groups. Student gender is another characteristic with a determining role in mode choice. Many studies (Nelson et al. 2008; McDonald 2008) found a higher tendency to use active modes for boys than girls. However, a broad study (Leslie et al. 2010) among students aged 10–14 in Australia showed a stronger tendency for girls to choose walking more than boys.

Among the factors related to household characteristics, income and vehicle ownership have a significant influence on the choice of travel modes. All studies (Ermagun and Levinson 2016; Wen et al. 2008) found that families with private cars are more willing to use this mode to transport their children. Also, children in high income families are more likely to use private car compared to walking or taking transit (Wilson et al. 2010).

Distance from home to school is among the built environment factors that affect travel mode in school trips. All the previous studies (Wilson et al. 2007; Mitra et al. 2010; Larsen et al. 2009) showed a negative relationship between distance and tendency to use active modes of travel. The significance of this variable is such that some studies tried to determine a tolerable walking distance. In particular, Ermagun et al. (2016) developed a two-step framework encompassing a random utility method and a hazard-based model. They showed 85% of students fail to walk when their travel distance exceeds 1.9 km. Nelson et al. (2008) found 82% of Irish students from 15 to 17 years that live closer than 2.4 km from school tend to walk. Schlossberg et al. (2006) mentioned that 52% of students who live in a 1.6 km buffer around their schools either walk or bike to school.

Long-term effects of school siting in urban development policies are undeniable. For example, in 1958 American policymakers argued in favor of establishing larger schools, and argued that it would improve learning opportunities for students (Walberg 1992). Between 1940 and 2001, the U.S. population increased 70% while the number of public schools fell by 69% (Ewing et al. 2004; Walberg 1992). Many students did not live near their school, which meant fewer opportunities for walking or biking. Understanding the effects of travel distance on the mode choice of students, the U.S. government started to build a greater number of neighborhood schools (Council of Educational Facility Planners, International 2004). Hence, recognizing the factors that influence school trip distance and the reasons for such decisions by the parents is essential for long-term solutions toward promoting active travel modes to school.

Table 1 Summary of school trips studies

S. No	References	Country	Age	Mode				Analysis method	Distance analysis
				Active (walk/bike)	Automobile	Public transit	School bus		
1.	Ermagun and Levinson (2017)	Iran	12–17	×	×	×	×	Cross-nested logit	Variable
2.	Li and Zhao (2015)	China	13–15	×	×	×		Multinomial logit	Split model
3.	Broberg and Sarjala (2015)	Finland	11–14	×	×			Multinomial logit	Variable
4.	Mitra and Buliung (2015)	Canada	11, 14-15	×	×	×	×	Multinomial logit	Variable
5.	Ermagun and Samimi (2015)	Iran	12–17	×	×	×	×	Three level nested logit	Variable
6.	Samimi and Ermagun (2012)	Iran	12–17	×	×	×	×	Binary logit	0/1Variable
7.	D'Haese et al. (2011)	Belgium	11–12	×	×			Two level bivariate regression	Descriptive
8.	Sidharthan et al. (2011)	U.S	5–15	×	×		×	Spatial lag	
9.	Alemu and Tsutsumi (2011)	Japan	15–18	×	×	×		Multinomial logit	–
10.	Leslie et al. (2010)	Australia	10–14	×	×	×		Binary logistic regression	–
11.	Wilson et al. (2010)	U.S.	7–12	×	×	×		Multinomial logit	–
12.	Van Dyck et al. (2010)	Belgian	17–18	×	×	×		Logistic multi level	Descriptive
13.	Mitra et al. (2010)	Canada	11–13	×	×			Binomial logit	–
14.	Larsen et al. (2009)	Canada	11–13	×	×	×	×	Logistic regression	–
15.	McDonald and Aalborg (2009)	U.S.	10–14	×	×	×	×	Descriptive	–
16.	Nelson et al. (2008)	Ireland	15–17	×	×	×		Logistic regression	Descriptive
17.	Yarlagadda and Srinivasan (2008)	U.S.	<18	×	×	×	×	Multinomial logit	–
18.	Wen et al. (2008)	Australia	9–11	×	×	×		Logistic regression	–
19.	McDonald (2008)	U.S.	7–14	×	×	×		Multinomial logit	0/1Variable
20.	McMillan (2007)	U.S.	9–11	×	×			Binomial logit	0/1Variable
21.	Martin et al. (2007)	U.S.	9–15	×				Logistic regression	–

Table 1 continued

S. No	References	Country	Age	Mode				Analysis method	Distance analysis
				Active (walk/bike)	Automobile	Public transit	School bus		
22.	Mota et al. (2007)	Portugal	12–16	×	×	×		Logistic regression	–
23.	Kerr et al. (2007)	U.S.	5–18	×				Logistic regression	–
24.	Kerr et al. (2006)	U.S.	5–18	×				Logistic regression	–
25.	Merom et al. (2006)	Australia	5–12	×	×	×		Logistic regression	Descriptive
26.	Timperio et al. (2006)	Australia	5–6, 10–12	×				Logistic regression	–
27.	Schlossberg et al. (2006)	U.S.	12–15	×	×	×		Logistic regression	Descriptive
28.	Schlossberg et al. (2005)	U.S.	12–14	×	×	×		Descriptive	–
29.	Ewing et al. (2004)	U.S.	7–18	×	×		×	Multinomial logit	–
30.	Evenson et al. (2003)	U.S.	12–15	×				Logistic regression	–

Table 2 Summary of influential parameters on mode choice in previous studies

Characteristic	Parameter	Private		Public		School bus		Walk	
		+	–	+	–	+	–	+	–
Students	Girl	(19)*	–	–	(19)	–	–	(6), (10)	(16), (19)
	Age	–	–	(17), (11), (19)	–	(17)	–	(11), (13), (19)	(21)
Household demographics	Education	–	–	–	–	–	–	–	(21)
	Income	(11), (19)	–	–	(19)	–	–	–	(13), (14), (19), (29)
	Full time work	(18)	–	–	–	–		–	–
	No. car	(5), (18)	–	–	(17), (19)	–	(17)	–	(19), (29)
	No. license	–	–	–	(29)	–	–	–	(13)
	No. children	(17)	(19)	(11), (19)	–		(17)	(11), (19)	(13)
Built environment	Distance	(18)	–	–	(11)	–	–		(11), (13), (14), (29)
	Population density	–	(19)	–	–	–	–	(13), (16), (19)	–
Parental Reservation	Safety	–	–	–	–	–	–	–	(20), (24)
	Convenience	–	–	–	–	–	–	–	(6), (20)
	Travel time	–	(19)	–	–	–	–	–	(19)

* Numbers in parentheses are based on the numbers of the studies in Table 1

This study is an effort to augment the current literature of school trips by developing a joint formulation for school travel mode and distance. This model is presented and compared with the conventional single models to highlight the significance of model specification in terms of policy assessment. Further, lack of school travel behavior studies in the Middle East propelled this research to shed light on the cultural impacts of travel decisions. We account for parents' perception of safety, convenience, and reliability when deciding on school trip mode and distance. These factors have received little, if any, attention in previous studies.

Method

Joint models are not necessarily adopted for simultaneous decisions. Indeed, when two variables are modeled jointly, interdependencies of the error terms are accounted for. There are several studies that jointly modeled decisions that are not made simultaneously in practice, but, have common unobserved factors. Spissu et al. (2009) developed a joint model of vehicle type choice and utilization using the 2000 San Francisco Bay Area Travel Survey. They underlined that the introduced joint discrete–continuous model system explicitly accounts for common unobserved factors that may affect the choice and utilization of a certain vehicle type. Habib (2012) also tested a joint trivariate discrete–continuous–continuous model for commuters' mode choice, work start time, and work duration to capture correlations among random components influencing these decisions. In a more recent study, Habib (2015) developed a joint model to investigate not only the factors affecting mode choice and travel distance of older people, but also their interrelationship. Likewise, school travel mode and distance could have several common unobserved variables, such as neighborhood walkability, safety and security perceptions, cultural preferences, and subject well-being that favor a joint formulation whether or not these decisions are made jointly in practice.

Several studies investigated a variety of methods to model interdependent decisions in a joint structure in recent decades. A simple approach is to combine various choices of different decisions (bundling), make new choices, and then calibrate a random utility model to predict choice behaviors (Ermagun et al. 2015b; Yarlagadda and Srinivasan 2008; De Jong and Ben-Akiva 2007). Although the use of a multinomial logit specification is open to question, one may address such shortcomings with a generalized extreme value model specification. Another approach employs a copula function that is mostly celebrated for its closed-form formulation and efficient estimations (Trivedi and Zimmer 2007). We test a joint copula model of multinomial logit and linear regression, respectively, for travel mode and distance in school trips. In the following, the modeling methods of the mode choice and travel distance are discussed, which is followed by a discussion of the joint copula method.

Mode choice model

A multinomial logit model is used to explain travel mode selection among four alternatives, namely automobile, school bus, public transit, and walk. Utility of mode i. for household q. is formulated as Eq. 1, where x_{qi}. is an observed attribute and ε_{qi} stands for the unobserved part of the utility or a random error term (Greene 2003).

$$u_{qi} = \beta_i x_{qi} + \varepsilon_{qi}. \tag{1}$$

Pursuant to the random utility maximization theory, household q chooses mode i. that comes with the highest utility (Eq. 2).

$$V_{qi} = \max_{j \neq i} u_{qj} - \varepsilon_{qi} \text{ and } \beta_i x_{qi} > V_{qi}. \tag{2}$$

Assuming independent and identical type-1 extreme value distribution for the error, the closed form probability for selecting mode i. by household q. (F_{qi}.) is shown in Eq. 3 (Greene 2003).

$$F_{qi} = \frac{\exp\left(\beta_i x_{qi}\right)}{\sum \exp\left(\beta_j x_{qj}\right)}. \tag{3}$$

Travel distance model

A regression model is used to model home-to-school travel distance for each mode of transportation. Travel distance is modeled using Eq. 4, where m_{qi}. is travel distance for household q. with mode i, γ_i. is the vector of regression coefficients, z_{qi} is the vector of observed attributes τ_{qi} is the error term.

$$m_{qi} = \gamma_i z_{qi} + \tau_{qi}. \tag{4}$$

The joint model: a copula-based approach

The error terms of the logit and regression models are assumed to be type-1 extreme value and normal, respectively. Sklar (1959) showed that a joint distribution for two marginal distributions can be approximated by the copulas, as shown in Eq. 5. In this equation, $F_{\varepsilon_{qi}}(.)$ and $F_{\tau_{qi}}(.)$ are marginal distributions, $C_\theta(.,.)$ is the copula function to generate the joint distribution, and θ is the copula parameter. We considered four types of copulas: Frank, Clayton, FGM, and Joe; and reported the one with the best fit. Bhat and Eluru (2009) suggested estimating Kendall's parameter instead of θ for a better interpretation of dependency between the marginal distributions. This is discussed later in more details.

$$F_{\varepsilon_{qi},\tau_{qi}}(y_1, y_2) = C_\theta\left(F_{\varepsilon_{qi}}(y_1), F_{\tau_{qi}}(y_2)\right). \tag{5}$$

Estimation method

The overall Maximum Likelihood function is written as in Eq. 6 (Bhat and Eluru 2009), where $f_{\tau_i}(.)$ is the probability density function of ε_{qi} and u_{i1} and u_{i2} are the cumulative probability distributions of the two error terms. In other words, $u_{i2} = F_{\tau_i}\left(\frac{m_{qi} - \gamma_i z_{qi}}{\sigma_{\tau_i}}\right)$ and $u_{i1} = F_{Vi}\left(\beta_i x_{qi}\right) = \frac{e^{\beta_i x_{qi}}}{\sum e^{\beta_j x_{qj}}}$; and $\frac{\partial C_{\theta i}(u_{i1}, u_{i2})}{\partial u_{i2}}$ is the partial derivative of the copula in respect to u_{i2}. The parameters γ_i, β_i, θ_i, and σ_{τ_i} are estimated using the maximum likelihood estimation (MLE) method, for which a code was developed in R software (R Core Team 2013) environment.

$$L = \prod_q \left[P\left(m_{qi} | \beta_i x_{qi} > V_{qi}\right) * P\left(\beta_i x_{qi} > V_{qi}\right) \right]$$
$$= \prod_q \left[\frac{1}{\sigma_{\tau_i}} * f_{\tau_i}\left(\frac{m_{qi} - \gamma_i z_{qi}}{\sigma_{\tau_i}} \right) * \frac{\partial C_{\theta i}(u_{i1}, u_{i2})}{\partial u_{i2}} \right]. \quad (6)$$

Data

Tehran is ranked the 16th most densely populated city in the world (City Mayors 2013), with a population of about 8 million. Over 17 million motorized trips are made daily in this city, 27% of which are educational (Tehran Comprehensive Traffic and Transportation Study Co. 2013). About 5300 schools are spread across the city with an area of over 700 square km. Fifty-four percent of schools are private, and they arguably have higher quality. Private schools often accept their students as freshman, and transfers are barely accepted by top-ranked private schools. Although public schools are more affordable, one needs to live in the school district to be eligible for the enrollment much like school districts in the United States. Private schools do not have the same criteria for enrollment, and the students only need to be able to afford the tuition costs and typically pass the entrance exam to attend.

Unlike the Western countries, there are many small neighborhood schools with very few classes in Tehran. This has made the school density of Tehran to be about 7.5 schools per squared KM, which is relatively high, compared to an average school density of 3.4, 2.9, 2.1, and 1.6, in Chicago, Minneapolis, New York, and Los Angeles, respectively.

To develop the models, we used four categories of data: (1) students travel information, (2) socio-economic data of the households, (3) built-environment specifications, and (4) transportation characteristics. The last two sets of data were obtained from Tehran Comprehensive Traffic and Transportation Study Co. (2013). While, an extensive survey was conducted in May 2011 in Tehran to collect students' travel information and socio-economic information of households. More than 4900 questionnaires were distributed among nearly 100 middle and high schools after a pilot survey. Forms were completed by the parents, and a response rate of 72% was obtained. Since schools in Tehran are gender-segregated, cluster sampling with respect to gender, level, and area of residence was conducted. Five common modes of transportation, including walking, cycling, school bus, private car, and public transport were considered. It should be noted that school busing in Tehran is different than in Western countries, in that it is neither free nor provided by the State. Rather, school busing is a chartered service that using minibuses, minivans, or private cars to provide door-to-door service for families that choose to pay. The cost of the service is not fixed and depends on the child's distance to the school. Since less than 1% of students reported biking to school, this data was removed from the mode choice analysis. The low share of biking may result from the lack of on- and off-street bicycle facilities and the "prohibition" of biking for girls.

To measure the distance from home to school variable, we extracted the walking travel time from Google maps using the origin and destination of each trip. We then transferred the walking travel time to commute distance, considering an average walking speed of 4.8 km/h for children. Figure 1 shows the percentage of each mode classified by travel distance from home to school. Moreover, Table 3 shows the data used in this study as well as some demographics.

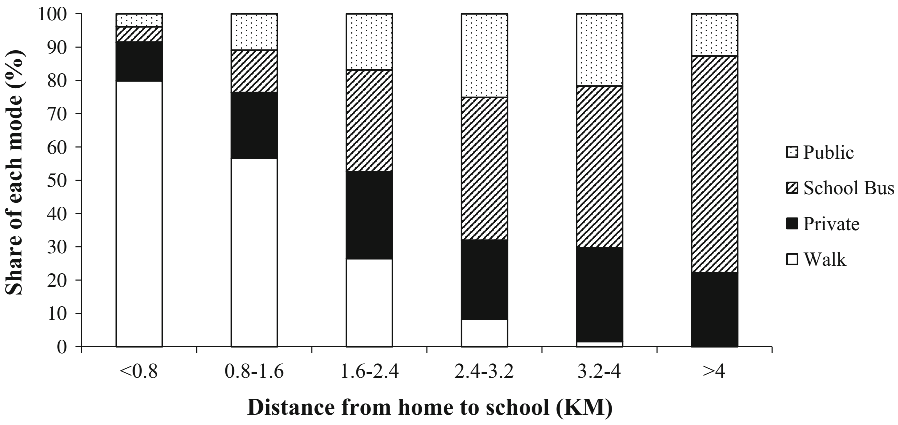


Fig. 1 Share of each mode classified by travel distance

Model

We developed a copula–based joint modeling framework for school travel mode and distance. Single models are also reported to make out the impact of modeling method on the estimated coefficients.

Joint model

Four common types of copula formulations, namely Frank, FJM, Clayton, and Joe were studied to select the appropriate model. Due to the use of non-nested copulas in this study, evaluation of different types cannot be done by the likelihood ratio test (Bhat and Eluru 2009). Thus, Bayesian Information Criterion (BIC) statistic is recommended to compare the models (Trivedi and Zimmer 2007). The model with lower BIC better fits the data. Results of the Frank copula, which has the best fit, are reported in Table 4.

One of the main goals of this study is to show that the error terms of the school trip mode and distance models are interdependent. This is well illustrated by the value of copula dependence parameters. As shown, these parameters are significant and different form zero. This demonstrates a high level of dependency between the unobserved factors of the travel mode and distance models.

The results indicate that the choice of walking is more dependent on travel distance than other modes. In other words, preference of walking regularly to school affects parents' choice of school location and vise-versa. Kendall correlation parameter is used to interpret the dependency parameter. It should be noted that interpretation of this factor varies across different copulas. Bhat and Eluru (2009) described the sign of this parameter in different copula models. In this study, the relationship between the error terms are opposite in sign, as shown in Eqs. 2 and 4. In Frank model, this negative sign indicates that unobserved factors, which increase the probability of choosing a mode, are likely to increase travel distance. Conversely, a positive sign indicates that unobserved factors cause a decrease in the probability of choosing a special mode in a longer distance. The estimated copula dependency parameter for walking and distance is negative, contrary to other transportation modes. Possible reasons for this difference could be hypothesized. For example, the

Table 3 Description of explanatory variables used in the modeling

Variable	Description	Average	SD
GENDER	1:Male/0:Female	0.39	0.48
AGE	Age of children between 12–17 years old	14.10	1.61
INCOME	1: less than 5/2: 5–10/3: 10–15/4: 15–20/5: 20–25/6: more than 25 million Iranian Rials* household income	2.09	–
NON_AUTO	1: Households with no car/0: Otherwise	0.20	0.40
AUTO_MOR2	1: Households with more than two car/0: Otherwise	0.18	0.38
PARTTIME	1: If work of one of parents is part time/0: Otherwise	0.333	0.47
EDUCATION	Educational level of parents 1: less than a high school diploma/2: high school diploma/3: bachelor of science/4: master of science or equivalent/5: higher degrees	2.03	–
CHILD_7	Number of school children in household (ages7–18)	1.58	0.67
NON_WRK	1: If non worker parents are in household/0: Otherwise	0.05	0.21
LOW_EDU	1: Parents have less than a high school diploma/0: Otherwise	0.33	0.47
SB_N_COST	Out-of-pocket school bus travel cost (10 Rials) divided by INCOME	2074.84	1068.43
WALKTRNT	Distance between home and the nearest bus station (meter)	571.21	449.72
POPDENS	Population density in each zone (person per m^2)	0.02	0.01
WALKSCH	1: less than 10/2: 10–20/3: 20–30/4: 30–40/5: 40–50/6:more than 50 min walk time to school	2.63	–
DISTANCE	Distance from home to school (Meter)	971.29	576.80
DURATION	1: If parents are primarily concerned about their children travel time/0: Otherwise	0.23	0.42
SAFETY	1: If parents are primarily concerned about their children travel safety/0: Otherwise	0.31	0.46
RELIABLE	1: If parents are primarily concerned about their children travel reliability/0: Otherwise	0.18	0.38
TRF_LIMIT	1: Students that live or study in a limited traffic zone/0: Otherwise	0.11	0.31
ACCESS	1: If parents have acceptable access to modes of transportation/0: Otherwise	0.18	0.39
CONVENIENCE	1: If parents are primarily concerned about their children travel convenience/0: Otherwise	0.30	0.46
COST	1: If parents are primarily concerned about their children travel cost/0: Otherwise	0.30	0.45
D_GENWSCH	1: Influence of WALKSCH variable for boys/0: Otherwise	1.07	1.65

* 11,800 Iranian Rails was equivalent to 1 USD in May 2011

preference of some parents that choose active transportation modes due to its health benefits is not considered in the systematic part of the utility. The negative sign of the dependency parameter implies that such factors that encourage walking are more likely to increase travel distance from home to school. Similarly, students who prefer to walk in order to socialize are inclined to travel longer distances than other students.

Student's t-test was used to evaluate the effect of variables. Most variables are significant at 90% of confidence interval. The smallest t-value is for *GENDER* in the regression model of walking and *DURATION* for regression model of public transportation.

Table 4 Summary of the frank copula model

Attributes	Variables	Multinomial logit model (mode)				Regression model (distance)			
		Private	Sch. Bus	Public	Walk	Private	Sch. Bus	Public	Walk
Students	CONSTANT	−2.48 (−3.72)	−1.49 (−1.92)	−5.60 (−6.49)	–	566.59 (1.02)	1616.49 (2.89)	−1613.44 (−2.61)	240.67 (1.39)
	GENDER	–	−0.39 (−2.69)	0.84 (4.65)	1.07 (3.29)	243.36 (1.93)	–	452.66 (5.01)	53.68 (1.41)
	AGE	–	00.12 (−2.78)	0.17 (3.11)	0.13 (3.14)	117.08 (3.17)	93.62 (2.44)	246.38 (6.89)	42.28 (3.84)
	D_GENWSCH	–	–	–	−0.43 (−2.81)	–	–	–	–
Household demographics	INCOME	0.19 (2.89)	0.27 (2.74)	–	–	–	105.86 (2.07)	–	−39.35 (−2.13)
	NON_AUTO	–	–	1.58 (7.07)	1.30 (6.64)	–	–	–	–
	AUTO_MOR2	0.48 (3.42)	–	–	–	–	–	–	–
	PARTTIME	0.23 (2.00)	–	–	–	–	–	–	–
	EDUCATION	–	–	−0.51 (−4.70)	−0.30 (−4.23)	233.53 (3.57)	208.7 (3.12)	–	–
	CHILD_7	–	–	0.23 (2.21)	–	–	–	–	–
	NON_WRK	–	–	–	–	415.51 (1.47)	–	–	–
	LOW_EDU	–	–	–	–	–	–	667.46 (3.11)	–
	SB_N_COST	–	−0.0002 (−1.64)	–	–	–	–	–	–

Table 4 continued

Attributes	Variables	Multinomial logit model (mode)				Regression model (distance)			
		Private	Sch. Bus	Public	Walk	Private	Sch. Bus	Public	Walk
Built environment	WALKTRNT	–	–	−0.0006 (−3.17)	–	–	–	–	–
	POPDENS	–	–	–	6.81 (1.50)	−16174.36 (−3.28)	−26699.42 (−4.62)	–	–
	WALKSCH	−0.38 (−6.92)	–	−0.22 (−4.40)	−1.72 (−15.59)	–	–	–	–
Parental concerns	COST	–	–	1.96 (10.98)	1.09 (7.35)	–	−540.84 (−2.64)	–	166.46 (4.48)
	DURATION	0.90 (7.14)	–	–	–	–	–	−230.32 (−1.34)	−178.24 (−4.44)
	SAFETY	–	–	−0.17 (−2.08)	−1.43 (−9.05)	–	–	–	−183.44 (−3.46)
	RELIABLE	–	0.79 (5.11)	−0.45 (−1.81)	–	439.24 (2.78)	233.67 (1.72)	–	–
	TRF_LIMIT	−0.85 (−3.21)	–	1.09 (4.13)	0.40 (1.62)	–	275.81 (1.47)	–	–
	ACCESS	–	−0.76 (−4.68)	–	0.26 (2.04)	–	–	–	–
	CONVENIENCE	0.68 (4.04)	0.97 (5.47)	–	–	–	–	–	–
Copula dependency parameter						2.37 (6.36)	2.37 (6.36)	1.57 (1.96)	−6.06 (−7.20)
Kendall's measure of dependency						0.53	0.53	0.21	−0.57
Scale parameter						1328.68 (25.53)	1402.05 (30.13)	1235.48 (31.50)	573.85 (42.64)

Table 4 continued

Attributes	Variables	Multinomial logit model (mode)				Regression model (distance)			
		Private	Sch. Bus	Public	Walk	Private	Sch. Bus	Public	Walk
Likelihood at zero				−20,263.27					
Likelihood at convergence				−18,901.06					
Sample size				3272					

* t-statistics are reported in the parentheses below each value

Single models

A summary of the single models are outlined in Table 5. A multinomial logit and a linear regression model are provided, respectively, for school trip mode and distance. The explanatory variables are similar to the joint model, but the estimated coefficients and standard errors are somehow different.

In the logit model, all the explanatory variables are significant at a 95% confidence interval, except for *SAFETY*, *TRF_LIMIT*, and *POPDENS*. The latter two variables appear in the utility function of the *Walk* mode with a 90% confidence interval, and *SAFETY* appears in the *Public* mode's utility function with a 75% level of significance. Further, the model has an adjusted R-squared of 0.35.

In the regression model, on the other hand, *NON_WRK*, and *TRF_LIMIT*, respectively in the *Private*, and *Sch. Bus* models are significant at a 90% confidence interval. While, *GENDER* and *DURATION* are significant at an 85% confidence interval, respectively in the *Public*, and *Walk* models.

Joint model versus single models

Differences between the joint and single models' coefficients are reported in Table 6. Wald test is deployed to measure the statistical significance between zero and the estimated differences. In the travel mode coefficients, walk time to school, and travel safety perception in the *Walk* alternative has a statistically different coefficient, respectively, in a 95% and 90% confidence. Further, *SB_N_COST* in *Sch. Bus* mode, *WALKTRNT* in *Public* mode, and *WALKSCH* in *Private* mode have different coefficients with a 40% confidence. The remaining coefficients in the joint and single mode choice models are not statistically different. In the distance models, on the other hand, *AGE* has slightly different coefficients with a 40% confidence interval.

Discussion

This part of the study is devoted to assess the impacts of different variables on school trip modes and distance according to the joint model. Variables are examined in four categories: student characteristics, household demographics, built environment, and parental reservations.

Student characteristics

Gender and age were found significant among student characteristics. The results indicate that boys are more probable to either use public transit or walk to school. It is more probable that girls use the school bus rather than other modes of travel. This may be rooted in the safety and security concerns of parents toward girls in Tehran.

Younger students are less probable to use public transit or walk, and more probable to take school bus. This probability may stem from the sense of independence in older students and a reduction in parents' travel safety concerns (McDonald and Aalborg 2009). One may also hypothesize that older students tend to use independent modes such as public transportation and walking to socialize and be more independent in Tehran (Samimi and Ermagun 2012). Students' age and travel distance are found to have a positive relation,

Table 5 Summary of the single models

Attributes	Variables	Multinomial logit model (mode)				Regression model (distance)			
		Private	Sch. Bus	Public	Walk	Private	Sch. Bus	Public	Walk
Students	*CONSTANT*	−2.29 (−3.79)	−1.37 (−2.00)	−5.51 (−7.27)	–	524.34 (1.04)	1389.15 (2.75)	−1697.51 (−2.97)	284.16 (1.97)
	GENDER	–	−0.38 (−2.98)	0.86 (5.31)	1.08 (3.60)	237.54 (2.24)	–	460.02 (3.65)	53.57 (1.58)
	AGE	–	−0.11 (−3.06)	0.17 (3.62)	0.11 (3.00)	88.59 (2.79)	84.81 (2.53)	247.35 (6.42)	52.04 (5.18)
	D_GENWSCH	–	–	–	−0.43 (−3.28)	–	–	–	–
Household demographics	INCOME	0.18 (3.05)	0.27 (3.10)	–	–	–	105.55 (2.40)	-	−38.04 (−2.20)
	NON_AUTO	–	–	1.58 (8.03)	1.39 (7.99)	–	–	–	–
	AUTO_MOR2	0.49 (3.81)	–	–	–	–	–	–	–
	PARTTIME	0.23 (2.19)	–	–	–	–	–	–	–
	EDUCATION	–	–	−0.50 (−5.16)	−0.33 (−4.75)	212.30 (3.82)	207.80 (3.56)	–	–
	CHILD_7	–	–	0.22 (2.39)	–	–	–	–	–
	NON_WRK	–	–	–	–	416.98 (1.81)	–	–	–
	LOW_EDU	–	–	–	–	–	–	666.66 (3.83)	–
	SB_N_COST	–	−0.0001 (−1.96)	–	–	–	–	–	–

Table 5 continued

Attributes	Variables	Multinomial logit model (mode)				Regression model (distance)			
		Private	Sch. Bus	Public	Walk	Private	Sch. Bus	Public	Walk
Built environment	WALKTRNT	–	–	−0.0008 (−4.62)	–	–	–	–	–
	POPDENS	–	–	–	7.94 (1.77)	−18405.61 (−4.10)	−29674.34 (−5.74)	–	–
	WALKSCH	−0.42 (−8.41)	-	−0.22 (−3.38)	−1.41 (−15.36)	–	–	–	–
Parental concerns	COST	–	–	1.95 (12.38)	1.10 (8.18)	–	−552.81 (-2.92)	–	167.07 (4.95)
	DURATION	0.90 (7.88)	–	–	–	–	–	−229.80 (−1.63)	−172.10 (−4.33)
	SAFETY	–	–	−0.17 (-1.14)	−1.76 (-13.19)	–	–	–	−169.52 (−3.49)
	RELIABLE	–	0.82 (6.05)	−0.47 (−2.09)	–	426.92 (3.16)	233.99 (2.05)	–	–
	TRF_LIMIT	−0.87 (−3.59)	–	1.08 (4.71)	0.40 (1.81)	–	274.12 (1.65)	–	–
	ACCESS	–	−0.74 (−5.25)	–	0.27 (2.20)	–	–	–	–
	CONVENIENCE	0.68 (4.35)	1.00 (6.18)	–	–	–	–	–	–
Adjusted R-squared				0.354		0.101	0.172	0.152	0.064
Sample size				3,272		601	702	374	1,302

* t-statistics are reported in the parentheses below each value

Table 6 Differences between the coefficients of the joint and single models†

Attributes	Variables	Multinomial logit model (mode)				Regression model (distance)			
		Private	Sch. Bus	Public	Walk	Private	Sch. Bus	Public	Walk
Students	*CONSTANT*	0.19 (0.21)	0.12 (0.12)	0.09 (0.08)	–	42.25 (0.06)	227.34 (0.30)	84.07 (0.10)	43.49 (0.19)
	GENDER	–	0.01 (0.05)	0.02 (0.08)	0.01 (0.02)	5.82 (0.04)	–	7.36 (0.05)	0.11 (0.00)
	AGE	–	0.01 (0.18)	0.00 (0.00)	0.02 (0.36)	28.49* (0.58)	8.81 (0.17)	0.97 (0.02)	9.76* (0.65)
	D_GENWSCH	–	–	–	0.00 (0.00)	–	–	–	–
Household demographics	INCOME	0.01 (0.11)	0.00 (0.00)	–	–	–	0.31 (0.00)	-	1.31 (0.05)
	NON_AUTO	–	–	0.00 (0.00)	0.09 (0.34)	–	–	–	–
	AUTO_MOR2	0.01 (0.05)	–	–	–	–	–	–	–
	PARTTIME	0.00 (0.00)	–	–	–	–	–	–	–
	EDUCATION	–	–	0.01 (0.07)	0.03 (0.30)	21.23 (0.25)	0.90 (0.01)	–	–
	CHILD_7	–	–	0.01 (0.07)	–	–	–	–	–
	NON_WRK	–	–	–	–	1.47 (0.00)	–	–	–
	LOW_EDU	–	–	–	–	–	–	0.80 (0.00)	–
	SB_N_COST	–	0.00* (0.76)	–	–	–	–	–	–

Table 6 continued

Attributes	Variables	Multinomial logit model (mode)				Regression model (distance)			
		Private	Sch. Bus	Public	Walk	Private	Sch. Bus	Public	Walk
Built environmen	WALKTRNT	–	–	0.00* (0.78)	–	–	–	–	–
	POPDENS	–	–	–	1.13 (0.18)	2231.25 (0.33)	2974.92 (0.38)	–	–
	WALKSCH	0.04* (0.54)	–	0.00 (0.00)	0.31**** (2.16)	–	–	–	–
Parental concerns	COST	–	–	0.01 (0.04)	0.01 (0.05)	–	11.97 (0.04)	–	0.61 (0.01)
	DURATION	0.00 (0.00)	–	–	–	–	–	0.52 (0.00)	6.14 (0.11)
	SAFETY	–	–	0.00 (0.00)	0.33*** (1.60)	–	–	–	13.92 (0.19)
	RELIABLE	–	0.03 (0.15)	0.02 (0.06)	–	12.32 (0.06)	0.32 (0.00)	–	–
	TRF_LIMIT	0.02 (0.06)	–	0.01 (0.03)	0.00 (0.00)	–	1.69 (0.01)	–	–
	ACCESS	–	0.02 (0.09)	–	0.01 (0.06)	–	–	–	–
	CONVENIENCE	0.00 (0.00)	0.03 (0.12)	–	–	–	–	–	–

† t-statistics are reported in the parentheses below each value

* Significant in a 40% confidence interval

** Significant in a 60% confidence interval

*** Significant in a 85% confidence interval

**** Significant in a 95% confidence interval meaning that older students tend to travel longer distances. This change in distance is more common for students who use public transportation more than other modes.

Socio-demographic variables

A variety of socio-demographic variables are found to have a significant effect on travel mode and distance. Among these, income, education, car ownership, and travel cost are found to be more influential.

The results showed that students from high-income households are more probable to use private cars and school bus. This is justified by the high travel cost of school bus in Tehran. As alluded to previously, unlike the Western countries, school bus in Tehran is a door-to-door expensive travel mode. Thus, it is not surprising that the prevalence of this mode is high among high-income families. High income households who choose school bus, also, do not show a significant sensitivity to travel distance. High income households that prefer to walk to school, however, preferred shorter distances. We can conclude that high-income families in Tehran are more probable to select private cars and school bus modes, and if they are willing to allow their children to walk to school they select a closer school.

Many studies (Yarlagadda and Srinivasan 2008; Ermagun and Samimi 2015) have discussed the impact of car ownership on mode choice. All of them found a positive correlation between car ownership and tendency to use personal vehicles for school trips. Two dummy variables of *NON_AUTO* and *AUTO_MOR2* were used to investigate this influence in households with no car and households with more than two. The first variable indicates the probability of using public transportation and walking increases for households that have no access to private car. Also, households with more than two private cars are found to be more willing to use their car for dropping off their children at school.

Parents' occupation was found to affect both travel mode and distance. This variable is more or less overlooked in previous studies (Wen et al. 2008). However, in line with some previous studies (Yarlagadda and Srinivasan 2008; Wen et al. 2008), we found non-workers are less inclined toward the use of active modes for their children's school trips. Also, unemployed parents that prefer motorized modes for their children's travel to school are not necessarily more inclined toward shorter distances.

As noted in past studies (Martin et al. 2007), parents' education affects students' mode choice. Educated parents, for instance, are found to discourage use of public transport or walking to school. One reason for this behavior could be attributed to the relationship between education and income. An interesting finding is that as the level of education increases, the tendency to walk declines with a more moderate rate than tendency to use public transit. Educated parents' awareness of health benefits of regular walking to school could be a reason for this behavior. The results, also, showed that an increase in the education level of parents that prefer private car or school bus leads to an increase in travel distance.

Undoubtedly, travel cost of public transportation is influential on the probability of using this mode. Some studies (Ermagun and Samimi 2015) have pointed to the importance of this variable on the probability of choosing different transportation modes. It is expected that high-income households are less affected by travel cost fluctuations. Therefore, we divided cost of school bus by the household income (*SB_N_COST*), to identify systematic taste variations among households.

Built environment variables

Most of the past studies (Wilson et al. 2010; Mitra et al. 2010; Ewing et al. 2004) found that the tendency to use public transportation and walking decreases as home-to-school distance increases. We brought in home-to-school distance in the utility of different modes by *WALKSCH*. Results showed that the utility of walking, public transport, and private car reduces with different magnitudes when the distance increases.

Distance from home to the closest public transit station is another built environment factor that has been disregarded in many previous studies. Understandably, our results showed that the higher the public transit accessibility, the higher the utility to use this mode for school trips. More illustration for this explanatory variable is provided in the next section.

Parental reservation

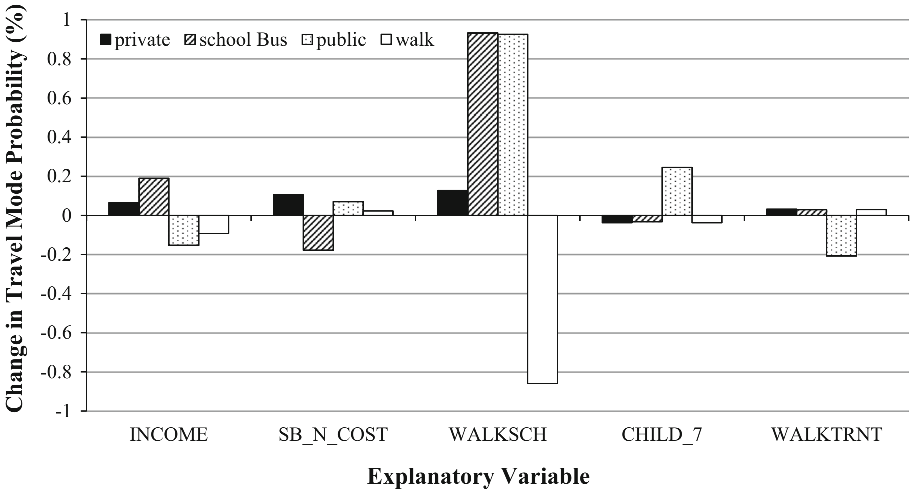
Safety, convenience, reliability, travel time, and traffic limits are the parental concerns whose effects on school travel mode and distance are examined. Apart from the first two variables (safety and convenience), other parental reservations have been hardly considered in the previous studies (McDonald and Aalborg 2009; McMillan 2007). We found that parents, who are concerned about their children's safety, discourage public transportation or walking as a mode of school travel. Moreover, those who have safety concerns but also prefer to walk to school are more probable to live closer to schools.

As expected, parents who are concerned about their children's travel expenses, mostly choose walking and public transportation compared to school bus and private car. Also, households that choose walking due to cost-related reasons are more probable to travel longer distances. In contrast, households that have decided to use school bus and have travel expenses as one of their priorities, try to choose schools that are close to their residence. Further, parents who have travel time concerns are more probable to take their children by private car more than any other mode. Also, if they choose walking or public transit for any reasons, they select schools that are located nearby.

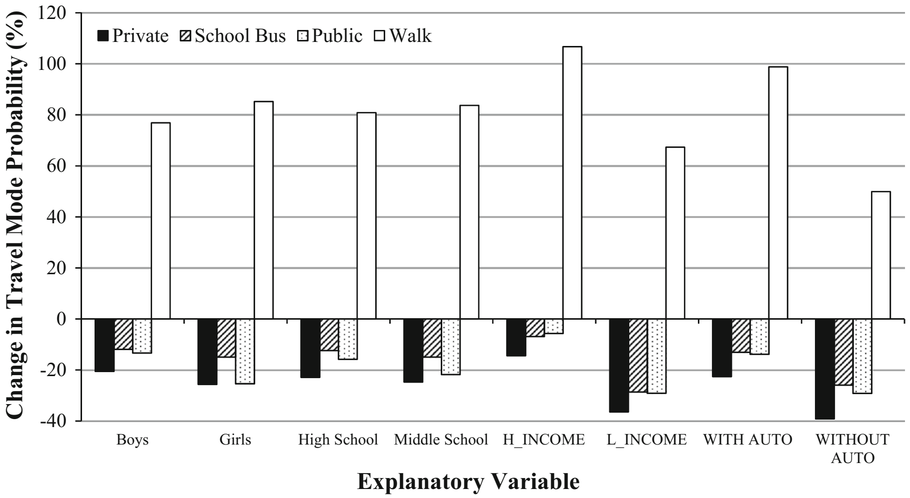
Another important concern is reliability. The results showed that attractiveness of public transit and school bus, respectively, decreases and increases when parents are concerned about school trip reliability. This is justified by the poor timing and frequency of public transit system in Tehran. Also, parents who are more inclined toward private car or school bus due to reliability concerns are less sensitive to travel distance.

Policy implication

Safety, traffic limits, distance from home to school, and cost of different transportation modes are policy sensitive variables that we examined. Figure 2 shows the elasticities of different variables on transportation mode choice, derived from the joint model results. Elasticity is measured as the percentage change in the mode choice probability when a given independent variable increases by one present. We examined elasticity of continuous variables (e.g. income and home-to-school distance) at the average points. Effects of dummy variables (e.g. safety perception) are computed as the percentage change in mode choice probability, if that concern is eliminated completely. Elasticity of different variables was examined on various population clusters such as girls, boys, and low-income



(a) Elasticity effects of continuous variables

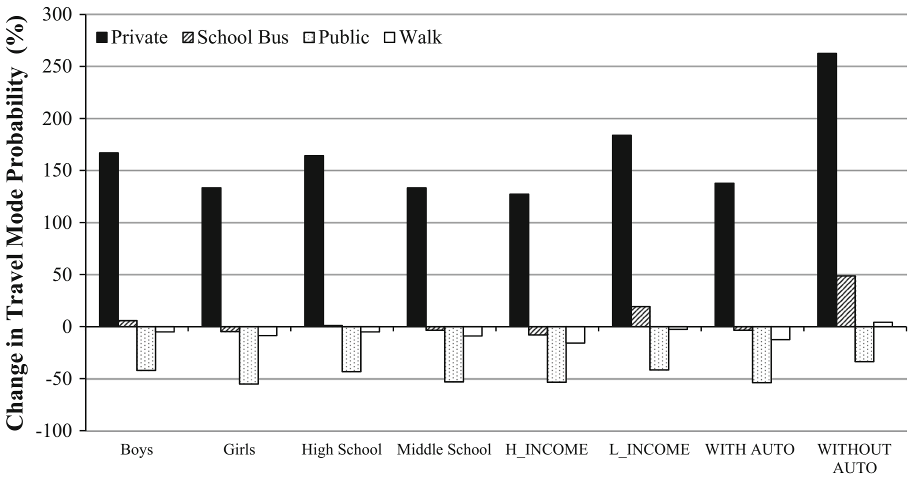


(b) Elasticity effects of *SAFETY*

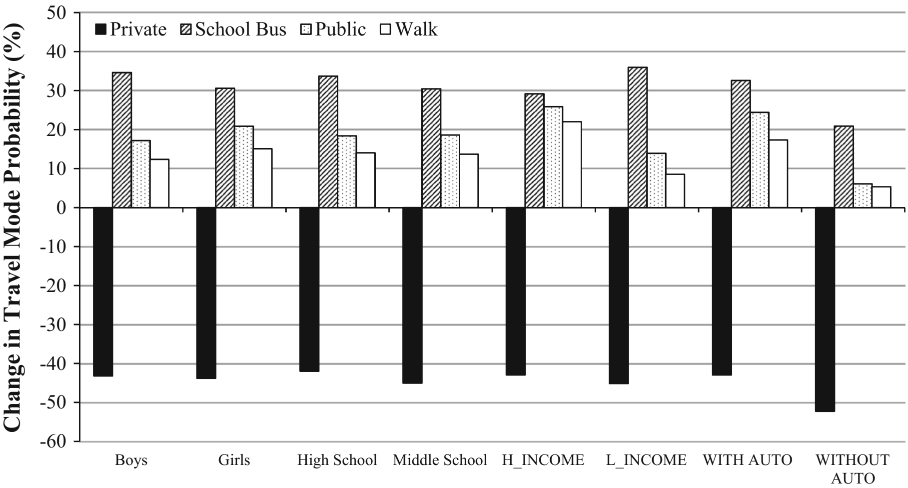
Fig. 2 Elasticities of policy sensitive variables

households. Such findings provide useful information for policy makers to better target policies for different segments of the society.

As shown in Fig. 2a, when household income increases by 1%, the probability of using school bus and private car, respectively, rises by 0.19% and 0.07%. Further, probabilities of using public transportation and walking decrease by 0.15 and 0.09%, respectively. The cost of school bus is normalized by households' income so one can observe a systematic taste variation among families. This variable indicates that a 1% increase in the cost of school bus decreases the probability of school bus choice by 0.17% and 0.08%, respectively, for low and high income families. Such results are useful for policy makers who try to decrease the usage of non-active modes by diverse pricing scenarios.



(c) Elasticity effects of *TRF_LIMIT*



(d) Elasticity effects of *DURATION*

Fig. 2 continued

As indicated earlier, a paramount variable in school travel mode is distance from home to school. If home-to-school distance increases by 1%, probability of walking will reduce by 0.85%, and propels parents toward non-active modes, especially the school bus. Therefore, policy makers that try to promote active modes of transportation should notice that a greater number of neighborhood schools can meet their desire in lieu of regional schools.

Availability of public transportation has not received due attention in previous studies. Elasticity of access to transit station (*WALKTRNT)* showed that a 1% improvement in the access increases the students' propensity to take transit by 0.20%. It should be noted that public transportation is a hybrid mode that implicitly includes some sort of physical activity. Therefore, promoting this mode is an option for policy makers that strive to ameliorate society's health condition.

Figure 2 illustrates how parental reservations on travel safety, traffic limits, and travel duration influence their mode choice decisions. We found that safety is the main concern of parents regarding their children's school trip. Should this concern be addressed entirely, the propensity to walk is expected to increase by 82%. Figure 2b shows the effect of safety on different household types. For example, the utility of walking will increase 106% for high-income and 67% for low-income families by removing travel safety concerns. Also, families are more concerned about their girls' safety. Elimination of safety concerns can elevate walking propensity by 85% among girls and 76% among boys. Policy makers with this kind of information can attain better results by applying different policies to different demographic groups and locations. For example, policies such as safe route to school (McDonald and Aalborg 2009) or walk-bus to school (Staunton et al. 2003) can help alleviate parents' concerns about their children's safety. As illustrated in Fig. 2c, setting traffic limits leads to a huge diminution in the use of private car. Results indicated that removing traffic limits can boost the use of private cars by up to 262%. The consequences of this variable are far-reaching. Therefore, considering appropriately-set traffic limits, urban designers can potentially help in reducing the use of motorized modes for school travel.

Regarding travel time concerns, the results indicate that as parental concern for travel time declines so does use of private vehicle. Figure 2d shows this decline fluctuates between 42.0% and 52.3% among different socio-economic family strata. As a result, the probability of school busing and public transit use increases significantly. Another interpretation of this analysis is that the parents who have concerns for travel time prefer to use private vehicle. In the case that private vehicle is unavailable, they prefer school bus over public transit and walking.

Conclusion

While researchers and practitioners have focused on analyzing effective explanatory variables on travel mode choice decisions as the key explanation for travel behavior in school trips, many are shifting attention toward the commute distance and school location decisions. It is essential to have a robust evaluation of effective variables on both school travel mode and home-to-school distance, in order to encourage active transportation among students. We tested a copula-based joint discrete–continuous model to unravel mode choice and travel distance in a joint framework. A confirmatory evidence for our hypothesis demonstrates that the error term of the school travel mode and distance are significantly inter-related. Joint estimation of the models made a significant difference in the effect of travel distance on willingness to walk to school. In fact, the absolute value of the travel distance coefficient in the mode choice model increased by 22% when a joint formulation was adopted instead of the conventional single estimations. Further, we found a significant decrease of 19% in the coefficient of travel safety perception in the joint mode choice model compared to the single model. Thus, single estimation of the models is technically a model misspecification that could result in distorted estimates. This dependency between the errors is more pronounced for the walking mode, according to Kendall's dependency parameter.

Furthermore, this study evaluated the effect of a range of variables such as distance from home to school, travel cost, safety and traffic limits on the school travel mode choice behavior among different demographic segments. For example, we evidenced that eliminating parents' concern about safety can double the propensity of walking to school. In the final model, the demographic heterogeneity was considered in evaluating the effect of each variable. High and low income families, for instance, are expected to respectively show 106% and 67% increase in the probability of walking, if their safety concerns are eliminated. Such findings can further provide information to policy makers for devising targeted policy measures to increase physically active travel to school among the different segments of the society.

While the results of this research offer some insights into the role of mode choice and home-school distance by using a copula-based modeling framework, there are limitations to the study that could be addressed in the future.

- Walkability of the neighborhood should be added to the model to avoid potential model misspecification. Built-environment variables such as density of green space, school, sidewalk and other land-use variables that are expected to promote walking trips were not utilized, due to unavailability of disaggregated data in this study.
- The way that students are accompanied to school could be modeled jointly with the travel mode and distance in a three dimensional joint model.
- The choice of school distance is really the choice of school location or school itself. The school location choice is affected by many explanatory variables such as quality of school, registration cost, and quality of the school's neighborhood. However, these variables are not publically available in our case study to take into account in the modeling process.

Acknowledgements We would like to thank Professor Abdul Pinjari for the helpful comments. He offered detailed advices on the modeling procedure and the theme of the paper. We also thank Ali Shamshiripour for providing assistance in the code development process, and three anonymous reviewers for their constructive comments.

References

Alemu, D.D., Tsutsumi, J.I.G.: Determinants and spatial variability of after-school travel by teenagers. J. Transp. Geogr. **19**(4), 876–881 (2011)

Bhat, C.R., Eluru, N.: A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. Transp. Res. Part B Methodol. **43**(7), 749–765 (2009)

Broberg, A., Sarjala, S.: School travel mode choice and the characteristics of the urban built environment: the case of Helsinki, Finland. Transp. Policy **37**, 1–10 (2015)

City Mayors. http://www.citymayors.com. Accessed 29 July 2013

Cooper, A.R., Andersen, L.B., Wedderkopp, N., Page, A.S., Froberg, K.: Physical activity levels of children who walk, cycle, or are driven to school. Am. J. Prev. Med. **29**(3), 179–184 (2005)

Council of Educational Facility Planners, International. Creating Connections: The CEFPI Guide to Educational Facility Planning. Council of Education Facility Planners International, Scottsdale AZ (2004)

De Jong, G., Ben-Akiva, M.: A micro-simulation model of shipment size and transport chain choice. Transp. Res. Part B Methodol. **41**(9), 950–965 (2007)

D'Haese, S., De Meester, F., De Bourdeaudhuij, I., Deforche, B., Cardon, G.: Criterion distances and environmental correlates of active commuting to school in children. Int. J. Behav. Nutr. Phys. Act **8**, 88 (2011)

Easton, S., Ferrari, E.: Children's travel to school—the interaction of individual, neighbourhood and school factors. Transp. Policy **44**, 9–18 (2015)