Socioeconomics of the Individual and the Costs of Driving: New Evidence for Travel Demand Modeler

A. P. Talvitie and Maria Koskenoja

Automobile travel costs have long been neglected in travel behavior analyses and travel demand models. The socioeconomic determinants of automobile travel cost choices are explored as part of a study into the effects of quality of work on work-related travel conducted in the Road and Traffic Laboratory, Technical Research Center of Finland. The study was motivated by the need to assess driving cost estimates for the mode choice model. It was approached through separate functions for fixed, variable, and total driving costs of the respondent. Within-household effects were examined through driving cost regressions of the household members of the respondent. The within-household regressions indicate that there are within-household work life effects on driving costs. The effects take place on a detailed level, which is not captured in a variable describing the household member's employment status. Rather, it appears that the quality of occupations of the household members has more effect on driving costs than the number or share of employed persons in the household. Automobile operating costs are not equal for everybody. They are an outcome of a choice and are influenced by the characteristics of the individual and his activities and the other household members and their activities. The same applies for automobile capital costs. The operating costs of automobiles had a low statistical significance on mode choice, and high capital costs were associated with automobile choice. These findings are not surprising but ones not believed nor a part of present travel demand model systems. But it is believable that automobile operating costs do not influence mode choice, or, that once automobile is chosen, it is an expensive one—comfort costs—by choice. The driving cost models indicate that travel choices are a part of complex behavioral interplay within a household in which the costs of transport have a major role.

Analyses of automobile travel costs have long been neglected in travel behavior studies and in developing travel demand models, with some exceptions (*1–4*). Models for automobile travel costs are developed and the socioeconomic determinants of automobile travel cost choices are explored as part of a study into the effects of type of work on work-related travel conducted in the Road and Traffic Laboratory, Technical Research Center of Finland.

In the present study rigid ex ante hypotheses are avoided. Rather, it is generally hypothesized that work arrangements in their entirety influence work-related travel choices and that these decisions interact with non-work-related transport decisions (*5*). In a similar vein, the models developed may not satisfy all the technical, statistical assumptions. The emphasis has been on exploring behavior, not mathematics.

A. P. Talvitie, Viatek Consulting Engineers and Architects, Pohjantie 3, Espoo, 02100 Finland. M. Koskenoja, University of California, Irvine, Calif.

The data for the study come from Finland, where both men and women work. In 1984, 49 percent of all employed people in Finland were women, of whom 88 percent work more than 30 hr per week; for men the figure is 98 percent. The percentage of the work force at the working ages is 79 for men and 72 for women. Work histories of employed men are only 2 years longer than those of employed women for persons aged 35 to 45. The difference gradually increases, but only to 7 years for persons aged 55 to 64. Thus, work and the working environment are influential factors for both sexes in the formation of preferences for everyday choices (*6–8*).

The contents, tools, and procedures of work are also changing because of changes in market structure and technology. Besides, work conditions are an object of planning by management and institutions. If there is a link between work arrangements and travel behavior, there is a need to understand the effects of such arrangements in order to respond to the changing transportation needs of different groups of working individuals.

DATA

Sampling

The sample was drawn by systematic sampling using the Population Register of Finland. The population was divided into two classes by age: adults (over 15 years) and children (7 to 14 years). Of the adults, persons were selected who are in the active work force and reported the travel diary day to be a normal workday. There were 1,518 such adult, working persons in the survey, which the diary cleaning reduced to the final sample size of 1,333 persons.

Data Collection

The transportation data were collected by the Central Statistical Office of Finland between September 1985 and February 1986. The collection procedure was as follows:

1. Persons in the sample received a letter from the Statistical Center explaining the purpose of the survey and the expectations concerning the respondent.
2. Three questionnaires to be filled out by the respondent were attached to the letter, including a note pad for the daily travel pattern, called the travel diary.

3. The interviewer contacted the respondent before the intended day of the study and made an interview appointment.
4. The data were gathered in the interview. The travel diary was translated from the note pad to the questionnaire form with help from the interviewer to avoid misperceptions.

The collected data are representative of the population, but the process of selecting diaries may bias accepted travel diaries toward young, male, transportationally active people. This was just a feeling that came when choosing the usable data diaries from the original sample.

Occupational Characteristics

Occupational characteristics come from a study conducted by the Central Statistical Office of Finland (*7*), in which 4,502 wage earners were interviewed about their conditions of work. The interviews took place in 1984, just 1 year before the transportation data collection, ensuring the compatibility of the two data sets.

In the occupational data survey the respondents gave subjective ratings about different dimensions of occupational requirements. The occupations were classified into 34 groups on the basis of similarities in activities; this, however, is not the official occupational grouping used in government statistics. The grouping of occupations causes some aggregation error, but the grouped data still include more of the variance than the normally used, single-dimension, class description of organizational status classes (blue collar, white collar, etc.).

Occupational groups were described by detailed questions. For example, the physical demands of an occupation were scaled through deviations from normal temperature, draft, dustiness of air, heavy lifts, extreme stretching, and so forth; from these an index of physical demands of work was formed. The chosen occupational dimensions are mental demands, physical demands, monotonousness, and the power to influence work rhythm.

The subjective nature of occupational ratings, together with technological and market structure changes, could soon cause the description of occupations to be outmoded if the characteristics of occupations are not updated. Fortunately, it is in the interest of labor unions and employer associations as well as university and statistical bureaucracies to produce these kinds of data. Therefore, the usage of occupational characteristics offers a possibility to use more disaggregated data, which are updated by existing institutions.

The occupational data were connected with the transportation questionnaire data by using the occupation reported in the questionnaire.

In the transportation questionnaire the respondents were also asked about weekly working hours, whether they had a shift schedule or irregular working hours, whether flextime was used, and whether they worked at home or at several locations. These variables are truly disaggregate in the combined data base.

Household Structure Variables

The age, gender, and employment status of the household members are known. Also, the personal income class and the income class of the household were asked. These data were used to form the life cycle variable and an indicator of whether the person was self-sustaining, the spouse of the primary earner, or a dependent. Intrafamily effects are not well covered, because there are virtually no transportation data from other household members. The relationships of daily tour combinations of household members remains to be investigated.

There was no clear indicator of the main wage earner of the household. This would have been desirable for determining the economic power positions of the respondents inside the household and for deducing the life cycle of that household. The decision rule used for selecting the main wage earner was the following: the respondent was classified as the main earner if she or he earned at least half of the total income of the household. If that was not the case, the presence of other working household members between ages 19 and 64 was determined. If they existed, the main earner was deemed to be the other in-law, mother, father, or a spouse, in that order. For example, if there were both a father and a spouse present, the spouse was classified as the main earner. This information was condensed to identify the respondent's position in the household: self-sustaining, spouse, or dependent.

The household life cycle has five classes. The first life cycle class contains childless households in which the main earner is younger than 35. Households are classified in the second class if the youngest child is less than 7 years old, irrespective of the age of the main earner. The third life cycle class has a youngest child between 7 and 16 (the school age), again regardless of the age of the main earner. A household without children under 17 is classified in the fourth class if the main earner is between 35 and 54. The fifth class contains households in which the main earner is 55 or older and there are no children under 17.

The classification of life cycles based on the age of the youngest child was selected because it was thought that the youngest person restricts mobility the most. This assumption is commonly made when forming life cycle variables. Zimmerman (*9*) discusses alternative ways to select life cycle variables.

Diary Data Cleaning

The following information concerning starting place and destination was collected in the diary: addresses, place codes, and times of day. The purpose of the trip, travel mode, and waiting time for public transportation were also asked.

The checking of the data had to be done manually, case by case, to ensure the right corrections. It took almost 1 year to clean the data. Distances were checked for illogical or not reported entries and, when necessary, measured from maps. In addition, starting places were corrected to be the same as previous destinations. Diaries that contained intractable obscurities in the paths were deleted. Diaries were also deleted if any work trip was done on an unusual mode (plane, boat, or "other"). The modes included in the analyses were automobile driver, automobile passenger, bus, tram, train, bicycle, and walk.

The data contained diaries that were otherwise tractable but in which the distance of a home-based walk round-trip (with purpose sport or unspecified) was unreported. These diaries were accepted with a change of distance of 1.0 km for the walking trip. Most of these walks were the last trips in the diary, so they are "evening walks" (taking the dog out, window shopping). Finally, a diary was deleted from the data if the respondent's travel day started from an implausible place, such as school, shop, other place of errands, or a day-care center.

Even though numerous judgments had to be made in preparing the data for model estimation, they are believed to be no worse or less accurate than data normally used in transportation models. That references to data-cleaning work are rarely made in the literature or, worse, that the analyst is not aware of how data were prepared, does not improve data quality. The view taken is that a thorough knowledge of data improves the appraisal of model results.

DRIVING COSTS

In developing a mode choice model, the time and money costs of driving must be approximated for the sampled individuals. The driving costs of cars are normally obtained first, then the driving costs of trips are estimated using the information about the cars available to the person in question. This figure is an estimate of the realized driving costs for persons who had a car available that day.

The method just described can also be used to approximate the driving costs for a person who at the moment did not have a car. The question to be answered in that case is, If the person had a car, what kind of driving costs would he or she have incurred per driven kilometer? This approximation to driving cost is different from the one customarily used in travel model studies, where driving costs are assumed to be constant regardless of the car driven and, indeed, regardless of whether or not a car is available.

Driving Costs per Car—Assumptions and Initial Values

The cost approximations that appear in this section are based on research conducted in the Technical Research Center of Finland about car prices and factors affecting driving costs. To calculate driving costs, cars are divided into five size groups and three age groups:

Size Class	*Size (cc)*	*Purchase Price, New (FIM)*	*Age Class (years)*
I	$x < 1300$	40,000	0–3
II	$1300 < x < 1600$	55,000	4–9
III	$1600 < x < 2000$	75,000	>10
IV	$2000 < x < 2500$	120,000	
V	$2500 < x$	150,000	

For simplicity it is assumed that all cars are driven 17 000 km/year. This was the average annual mileage in Finland in 1985. It is casually believed that most kilometers are driven with new cars. If this is true, their driving costs will be estimated high and the driving cost for old cars low. Unfortunately, the yearly mileage figures were not available. This lack of better data is regrettable but not unusual in transportation studies as the analogous example by Train (*10*), reported later, illustrates.

Fixed costs include depreciation, interest, mandatory traffic and car insurance, anticorrosion treatment, storage, and parking at home. Variable costs include tires, repair and maintenance, and fuel.

Depreciation of capital during the first 3 years is set at 15 percent of the remaining capital. The rest is depreciated over the following 6 years. When the car reaches the age of 10 years, all the original capital is depreciated. Whatever value the car may have thereafter is considered to be because of spare parts and repair and maintenance put into it. The interest rate used is 6 percent per the average capital of the year in question. The rate is lower than the market rate because interest costs are partly tax deductible; 6 percent is estimated to be the part of interest the owner really has to pay.

Anticorrosion treatment is calculated to be 450 FIM/year regardless of car size or age. The costs of tires and insurance rise with car size class as follows:

Size Class	*Tires (FIM)*	*Insurance (FIM)*
I	650	2,014
II	700	2,362
III	800	2,709
IV	900	3,057
V	900	3,751

Storage and parking costs depend on the respondent's living environment. In the central city they are 150 FIM/month, in the suburbs they are 75 FIM/month, and in towns and rural areas storage and parking costs are assumed to be nil.

Repair and maintenance costs are calculated as a percentage of the purchase price of the car. The percentage is lower for new cars, reflecting their better condition, and also lower for bigger cars due to their better quality. The following percentages apply in each age and size group; fuel consumption is also given:

Size Class	*New Cars (Age 0–3 years) (percent)*	*4–9 year-old-cars (percent)*	*Fuel consumption (L/100 km)*
I	2	4	8
II	2	4	9
III	1.75	3.5	10
IV	1.5	3	12
V	1.5	3	14

For cars more than 10 years old, the repair percentage is a flat 10 percent in every size class. The fuel price is 3.54 FIM/L, which was the price of regular gasoline in October 1985.

Fixed and variable costs are summed and divided by 17,000. The results are shown in Table 1. The three different costs for each car size class are due to the different storage and parking costs in city centers (c), suburbs (s), and villages or rural areas (r).

After the driving costs per car are approximated, the cost per driven kilometer for different persons can be estimated.

Driving Costs of Respondents with Cars at Their Disposal

The data indicate the number and makes of vehicles in the household and their primary users. One of the many alter-

TABLE 1 VARIABLE AND FIXED COSTS OF DRIVING IN FINLAND (FIN/km) ($1 U.S. = 4.0 FIM)

Size Class		I	II	III	IV	V
* CAR AGE 0-3 years						
fixed	c:	0.64	0.80	1.02	0.47	1.80
	s:	0.58	0.75	0.96	1.42	1.75
	r:	0.53	0.70	0.91	1.36	1.69
variable	:	0.36	0.42	0.48	0.58	0.68
tot.cost	c:	1.00	1.22	1.50	2.05	2.48
	s:	0.94	1.17	1.44	2.00	2.44
	r:	0.89	1.12	1.37	1.94	2.37
* CAR AGE 4-9 years						
fixed	c:	0.52	0.65	0.80	1.12	1.38
	s:	0.47	0.59	0.75	1.07	1.33
	r:	0.42	0.54	0.70	1.01	1.27
variable	:	0.41	0.49	0.56	0.69	0.81
tot.cost	c:	0.93	1.14	1.36	1.81	2.19
	s:	0.88	1.08	1.31	1.76	2.14
	r:	0.83	1.03	1.26	1.70	2.08
* Car AGE 10+ years						
fixed	c:	0.25	0.27	0.29	0.31	0.35
	s:	0.20	0.22	0.24	0.26	0.30
	r:	0.14	0.17	0.19	0.21	0.25
variable	:	0.58	0.68	0.84	1.18	1.43
tot.cost	c:	0.83	0.95	1.13	1.49	1.79
	s:	0.78	0.90	1.08	1.44	1.73
	r:	0.72	0.85	1.03	1.39	1.68

natives was that use of the car was shared. A small proportion of the respondents indicated this to be the case. The most common case was that the car belonged to a certain person in the household. Often one person was the main user of several cars.

The cost allocation problem arises because the diary data do not indicate which car the person was driving or got a lift on. The following rules are used to approximate the personal driving cost of the respondent:

1. If the person reports one or more cars to be at his personal disposal and there are no cars in common use, the fixed costs are the sum of the cars that the person has.
2. If no one else can use the cars, the fixed costs are all summed, because the owner can drive only one car at a time, but the time runs on all of the cars. The variable costs are estimated to be a mean of the variable costs of those cars.
3. If the household has one or more cars in shared use, the fixed cost is the sum of all those cars divided by the number of driving licenses in the household. The variable costs in shared use are again a mean of the variable costs of the cars.
4. If the person has some cars at his or her personal disposal and some in shared use, only the cars at personal disposal are counted in estimating costs. This is a simplifying assumption, but in most cases the vehicle in shared use was a van and the vehicles in personal use were cars. Thus the vans drop off when the mode of travel is driving a car, because in the questionnaire there is another mode for driving a truck or van.
5. Total driving costs are the sum of fixed and variable costs.

Driving Costs for Respondent's Household Members with Access to a Car

Because cars are allocated to certain household members, it is possible to formulate models describing the effect of the respondent's characteristics on the chosen driving cost level of some other household member, provided he or she has access to a car. To gain this understanding of the effects of a person on another person's driving cost choices, the data were used in reverse order: the cost choices of the respondent's household members were regressed on the characteristics of the respondent.

Of interest here are the respondent's work type or socioeconomic characteristics that may cause changes in the household members' driving costs. For example, the young age of a respondent could induce someone else in the household to change driving costs. For example, a British study (*11,12*) reports higher values of time when there are small children in the car. It is plausible that this affects the driver's choice of driving costs, that is, the choice of car type.

In the data there is much information about the personal characteristics of the respondent. However, there is not much information about the other household members. In the linear regression model all variables are introduced to account for both the effect of those household characteristics that are common to all of the household members (e.g., location of the household) and those that vary among individuals (an example of this type of variable is "main earner of the household"). As in the previous section, three cost variables were created: others' fixed costs (OFC), others' variable costs (OVC), and others' total costs (OTC). They refer not to the respondent, but to another household member with access to cars. The following rules were adopted for calculating OFC, OVC, and OTC: OFC is the mean of the fixed costs of the cars belonging to the other household members. OVC is the mean of the variable costs of these cars. OTC is simply OFC plus OVC.

Correction for Self-Selectivity Bias in Modeling Driving Costs for Persons Who Do Not Have a Car

Marketing science suggests that different persons drive different kinds of cars. There may be combinations of driving costs a particular person would not even consider as an alternative.

In estimating potential driving costs for persons who at present are nondrivers, it is important to know what determines the chosen costs of driving for those who are drivers. This information can then be applied to the nondrivers to produce a realistic driving cost alternative for them.

If this calculation is done simply on the basis of the observed characteristics of the drivers, it will yield biased estimates. This is because the variables that affect the chosen driving cost level also affect the choice to be or not to be a driver. This bias is called the self-selectivity bias. The correction for this bias is presented next.

The decision to have a car and the decision of driving cost level are made jointly. According to Train (*10*), models of car ownership and driving quantity should be linked to avoid the self-selectivity bias. He proposes an instrumental variable estimation, where the probability of car ownership is estimated first and this estimate is used to form the correction term in the driving quantity regression. In Train's model the driving cost was fixed and the amount of driving was estimated. In the model here, the amount of driving is fixed and the cost of driving varies. The logic of the correction term is not violated by this change.

Following Train, if the choice probabilities are logit and the error term of the driving cost function is normal, the correction term is

$$E(e_c) = -(\sqrt{6}\sigma/\pi) * p_c[P_q \ln P_q/(1 - P_q) + \ln P_c]$$

where

- e_c = error term in the driving cost equation;
- $E(e_c)$ = $e_c - \tau$, selectivity correction term, which equals the error term in the driving cost function—normally distributed;
- P_q = probability of owning one or more cars;
- P_c = probability of owning no cars;
- σ = standard deviation of e in the entire population (not conditional on the choice of the number of cars); and
- p_c = correlation of e with the unobserved utility associated with owning no cars.

When this correction term is added to the driving cost regression, the rest of the parameters are unbiased:

$$\text{vc} = \boldsymbol{\beta}\mathbf{s} + \Gamma C_c + \tau$$

$$\text{fc} = \boldsymbol{\alpha}\mathbf{s} + \Gamma C_c + \tau$$

where

- vc = variable cost conditional upon being a driver,
- $\boldsymbol{\beta}$ = vector of parameters to be estimated,
- $\mathbf{s}$ = vector of characteristics of the person and other explanatory variables,
- τ = a normally distributed error term,
- C_c = $[P_q \ln P_q/(1 - P_q) + \ln P_c]$, and
- Γ = $-(\sqrt{6}\sigma/\pi) * p_c$.

Because the value of C_c can be calculated, the value of p_c comes from the estimation.

Train also proposes that the utility of choosing the car (make or model and vintage) from the varying number of alternatives inside each class should be accounted for by a correction term. This reflects the assumption that if a car is selected from a class containing several alternatives, the choice will be closer to the optimum than when there are only a few makes or models to choose from. Because this kind of correction is not done here, the estimates may be biased. Because each class is likely to contain a large number of alternatives, the bias is small.

Train estimated the models for the probability of having no car, one car, and two cars and used these estimates in the correction term of the equation for operating costs. In his model the exogenous variables for predicting the operating costs of households' cars were gas price in the area of residence, household income, household size, type of housing unit, number of adults and adolescents, number of workers, age of household head, education level of household head, sex of household head, distance to work, population of household's area of residence, and number of transit trips in the area of residence.

The variables in this study are expected to be different because the equations are estimated not for a household but for an individual in the household. An important variable, distance to work, does not exist in the present data.

Logit Model for Respondent's Car Availability

To obtain an estimate for the correction term to avoid the self-selectivity bias, a binomial logit model was estimated. The alternatives were no cars available (nondriver) and one or more cars available (driver). If the person shared the use of at least one car, he or she was classified as a driver. The set of variables that were considered important is given in Table 2, and a "best" model estimated using these variables is given in Table 3.

The model predicts the probability that the respondent is a driver, that is to say, is the primary user or shares the use of at least one car. The automobile availability model is needed

TABLE 2 EXPLANATORY VARIABLES IN CAR AVAILABILITY MODELS

Biological Variables:	
family life cycle	age of the youngest hh member
year of birth of the respondent	sex (female=1, male=0)
Time connected Variables:	
works outside home	works flexible hours
several places of employment	hours worked in a week
works regular hours	works in shift
works irregular hours (=deadline)	
Organizational Variables:	
farmer	self-employed
upper white collar	lower white collar
blue collar worker	
Industry Variables:	
primary production	manufacturing
energy and utilities	construction
commerce, restaurant and hotel	finance and insurance
transportation, and communications	services
Occupational Variables:	
mentally demanding	physically demanding
much/little impact on the work rhythm	monotonous
Relative Economic Power of the Household Members:	
monthly personal income	monthly household income
main wage earner	spouse of the main earner
dependent (=Neither of the above)	
Household Structure:	
number of persons in the household	number of working persons
number of persons not active in work life	
Distances:	
distance to the nearest bus stop	distance to the food store
distance to a bank	distance to a post office
Locational variables:	
traffic volume of home street	home in the central city
home in a suburb	home in rural area
home in a small town or village	
House type:	
single family or a detached house	townhouse apartment
owns a summer cottage	can use somebody's summer cottage
Driving variables:	
r(espondent) has a car available	r's household member has a car available r has a
drivers licence	no. of driver licenses in the household
others' driving costs	

TABLE 3 LOGIT MODEL FOR RESPONDENT'S CAR AVAILABILITY (DEPENDENT VARIABLE: RESPONDENT HAS A CAR AVAILABLE = 1)

Independent Variable	Estimated Coefficient	t-Statistic
constant	-13.997	-6.907
female	-1.895	-11.076
self-employed	1.659	3.749
works outside home	0.619	1.870
several employments	0.499	1.882
hours worked	0.000	2.300
log of pers income	0.883	3.020
log of hh income	0.621	2.289
dependent	-0.785	-2.321
others drive	-2.743	-11.194
# of other licenses	0.939	5.831
rural area	0.963	3.722
one family house	0.930	4.177
row house	0.817	2.986
auxiliary statistics	at convergence	initial
log likelihood	-461.674	-779.097
number of observations	1124	
percent correctly predicted	82.2	

in estimating car driving costs. A few comments on the model follow.

Females have a smaller probability of being drivers than do males. Self-employment, workplace outside home, several employers, and the hours worked during a week all increase the probability of being a driver. The work location variables are statistically significant at one-sided 0.05 level, and they were included in the model because of their expected sign and because the effect of work type on travel behavior is the subject of this study.

The occupational variables (mentally demanding, physically demanding, monotonous, and possibility of influencing work rhythm) all lost their significance when other variables were introduced. Both physically demanding and monotonous occupations are correlated with not being a driver. Contrary to expectations, flexible working hours, shift, or variable (deadline) working hours did not have a statistically significant impact on car availability.

Both the logarithm of personal income and the logarithm of household income are positive and statistically significant. Personal income would have been statistically significant also without the logarithmic transformation, but the logarithmic transformation was preferred, to reflect the declining marginal utility of income.

The variables describing the person's relative economic strength in the household—main wage earner, spouse, and dependent—are mutually exclusive dummy variables. "Dependent" receives a value of 1 when the person in question is not the head of the household (is not the main wage earner) and is not the spouse of the head. "Spouse" marks all persons whose spouse is the main wage earner.

"Dependent" receives a statistically significant negative value indicating that, in spite of income, this social standing works toward not having a car at one's disposal. If "dependent" and "spouse" are both included in the model, they both get a negative coefficient, but for "spouse" it is smaller and not statistically significant ($t = 1.3$).

Car availability in the household by persons other than the respondent has a strong negative effect on the respondent's car availability and suggests transferability of driving tasks. Conversely, when the number of driving licenses in the household increases, so does the respondent's probability of acquiring a car. The explanation for this could be a possibility of more efficient car usage, learning the way of life from other household members, or both.

Finally, living in a rural area and in a one-family detached house or a town house is positively correlated with the respondent's car availability.

Logit Model for Car Availability of the Household Members of the Respondent

The estimated logit model for the respondent's household members to have a car available is given in Table 4. The variables, which all describe the respondent, were chosen with three primary concerns in mind: they should be statistically significant, interpretable, and preferably the same variables as in the respondent's car availability model, to see their effect inside the household and in relation to the outside of the household.

The other (nonrespondent) members of the household tended to have a car available if the respondent was female. The nonavailability of the respondent for household tasks (much employment and large number of hours worked weekly) is one inducement for the other members of the household to acquire a car. Self-employment of the respondent works in the same direction. The explanation could be lack of time for household tasks or, just as well, a learned way of life. Again, flexible working hours, shift, or variable (deadline) working hours did not have any impact.

The probability that a household member has a car available increases with household income, as expected. However, the probability diminishes with increasing income of the respondent. This is not totally explained by the income differentials, because those are captured in the dummy variables "main wage earner" and "dependent." This characteristic is statis-

TABLE 4 LOGIT MODEL FOR CAR AVAILABILITY OF HOUSEHOLD MEMBERS OF RESPONDENT (DEPENDENT VARIABLE: RESPONDENT'S HOUSEHOLD MEMBER HAS A CAR AVAILABLE = 1)

Independent Variable	Estimated Coefficient	t - Statistic
constant	-11.485	-5.540
female	0.841	4.116
sevrl employments	0.365	1.446
hours worked	0.000	2.521
self-employed	0.657	2.015
log of hh income	3.110	9.163
log of pers inc	-2.216	-6.241
main earner	0.603	1.939
dependent	0.793	2.046
traffic volume	-0.173	-2.683
rural area	0.173	0.739
summer cottage	0.827	4.262
one family house	1.493	6.753
townhouse	1.057	4.009
r has a car avail	-2.611	-10.334
has a licence	0.874	3.544
auxiliary statistics	at convergence	initial
log likelihood	-483.621	-765.234
number of observations	1104	
percent correctly predicted	79.9	

tically significant and does not change when other variables are changed. It needs further investigation.

Car availability is negatively correlated with the traffic volume on the street of residence, perhaps reflecting the level of service of public transportation. Curiously enough, the distance to the nearest bus stop, which was thought to indicate the level of service, did not have any significance. The same locational variables that were significant in driver model estimation are significant here: living in a rural area and in a one-family house or a townhouse are all associated with an increased probability that the household member has a car available.

Another similarity to the driver model is that the driver status of the respondent reduces the probability that other household members are drivers, but the driving license of the respondent increases it. This again could be interpreted as indicating the transferring of travel-related tasks inside the household.

Driving Cost Models for Respondents

The driving cost models were estimated by standard least squares estimation with the previously discussed correction term included in the model. The correction term's logic is briefly presented here; see Figure 1. The term corrects the regression for "unobserved observations," that is, for the would-be-chosen cost levels that are not possible because of technical or market reasons such as one shown in Figure 1.

If income were the only factor influencing the chosen driving cost level, and the threshold amount of income were the one marked with the horizontal line, then the observed driving cost level and income pairs would be the ones above the line. If a regression were run on these pairs only, the estimated line would not be steep enough and the intercept would be higher than in the true regression. The reason for this is the double influence of income: it affects the decision to have or not have a car and the level-of-cost decision. If totally different variables affected the automobile and the cost-level choice, the correction term would be unnecessary.

The cost regressions were estimated separately for fixed costs, variable costs, and total costs. The division between these costs is interesting because decisions about these costs are supposedly made on different time horizons—fixed costs for longer horizons and variable costs for shorter. This assumption is supported in the empirical findings from longitudinal studies that the transfer cost of selling the old car and buying a new one is considered high (*10*). This means that people do not buy and sell their cars to adjust to the optimal combination of costs, but rather decide the fixed-cost level and let the variable costs have their impact on the amount of driving.

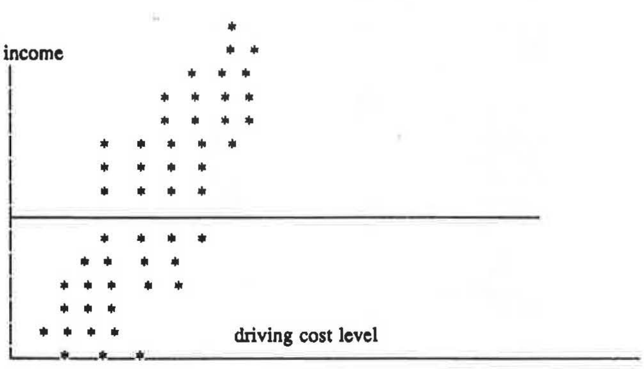

FIGURE 1 Rationale for correction term due to "unobserved" observations in driving cost regression.

Model for Fixed Costs of Driving a Car

The fixed-cost model is discussed first (see Table 5). The correction term, which carries the effect of presently having/not having access to a car, has the expected sign; the probability of having a car available increases the fixed costs. The coefficient has a low significance, but the term was left in the regression on theoretical grounds.

The life cycle dummy (cycle2) marks households in which the youngest child is under school age. People living in households with small children are driving cars with lower fixed costs. (Fixed costs are lower for smaller and older cars.) Female drivers have lower fixed cost.

The following occupational attributes are statistically significant: irregular working hours (deadline), self-employed, working in the energy and utility industry, and mentally demanding occupation. These attributes increased the chosen fixed driving costs.

Working irregular hours increases the person's scope for planning the day's activities and the number of alternatives from which to choose. It is consistent with the theory that if the person can choose an alternative closer to the optimum, the value of time is higher. This also implies that the chosen monetary cost is higher. The MVA Consultancy (*11*) found that the value of time was higher for people working irregular hours.

The occupational variable concerning impact on work rhythm did not have significance. The effect of self-employment may explain that variation. It is usual for entrepreneurs to describe the positive sides of their work as independence and flexibility in the sense that the person can agree about the deadlines.

TABLE 5 MODEL FOR FIXED COSTS OF DRIVING A CAR (DEPENDENT VARIABLE: FIXED DRIVING COSTS OF RESPONDENT)

Independent Variable	Estimated Coefficient	t-Statistic
correction term	0.023	1.016
constant	0.358	7.607
cycle2	-0.072	-2.640
female	-0.101	-2.459
deadline	0.113	3.290
self-employed	0.166	3.803
energy industry	0.131	1.961
mentally demanding	0.002	2.685
personal income	0.000	2.731
summer cottage	0.071	2.480

Number of Observations	604
R-squared	0.16
Sum of Squared Residuals	50.570
Standard Error of the Regression	0.292
Mean of Dependent Variable	0.566

Working in the energy and utility industry may increase fixed costs simply because the locations visited during the day may be hard to reach by other modes; for example, distant farmhouses, voltage leveling stations, or power lines are rarely served by public transit. Mental and physical demands of occupations have a strong negative correlation with each other. Of the two, mental demands had a clearer effect on driving cost.

Higher personal income increased the chosen fixed-cost level, as expected. Income segmentation was also tried by dividing the sample into three income classes and estimating separate coefficients for all of them. The nonsegmented and segmented incomes were also transformed to see which would best reflect the true effect on income. None of the combinations was significantly better. The same procedure was executed on household income. No form was statistically significant. Household income was left out of this regression, and personal income was used unsegmented and untransformed.

Owning a summer cottage increased the chosen fixed-cost level. The possibility of using a summer cottage of a relative or a friend did not have any effect on the chosen fixed-cost level.

Model for Variable Costs of Driving a Car

The model is given in Table 6. Even though the R^2 is very low, the model has statistical significance. The variable costs are nearly the same, captured by the constant term of the regression; however, many other factors influence the variable costs in a statistically significant way.

The correction term is right-signed and significant; owning a car increases variable costs. Cycle1, marking childless households with a young head; working in the transportation, storage, or communications industry; and the number of persons in the household are all associated with higher variable costs. The only variable associated with lower than "constant" variable costs is living in a town house.

Because maintenance costs are included in variable costs, the accepted level of variable costs could be connected with the accepted level of risking the car trip. Thus, young and childless adults may find it acceptable to change mode and timetables; working in the transportation or communication industry may be connected with the ability to perform unexpected repairs, and so forth.

The number of persons in the household may work its way through the effect that accepting higher variable costs may be the only way to acquire a big enough car. On the other hand, households with small children had lower-than-average fixed driving costs. It may be that the increase in family size causes a change to bigger but older cars. Living in a town house may be a proxy for a certain way of life, or for traffic environment; at this point no other possible explanation was found.

TABLE 6 MODEL FOR VARIABLE COSTS OF DRIVING A CAR (DEPENDENT VARIABLE: VARIABLE DRIVING COSTS OF RESPONDENT)

Independent Variable	Estimated Coefficient	t-Statistic
correction term	0.018	2.064
constant	0.509	22.605
cycle1	0.051	2.553
transp industry	0.048	2.214
# of persons in hh	0.011	2.008
townhouse	-0.047	-2.613
Number of Observations	614	
R-squared	0.04	
Sum of Squared Residuals	14.946	
Standard Error of the Regression	0.157	
Mean of Dependent Variable	0.536	

Model for Total Costs of Driving a Car

Total cost, the sum of fixed and variable costs, is given in Tables 7 and 8. For easier interpretation, the union of variables in fixed-cost and variable-cost regressions was first introduced for regression. After dropping the statistically nonsignificant variables the model in Table 7 was obtained.

To determine whether there are some different factors that influence the choice of total driving cost level, all the available variables were introduced to the model and then dropped if proven insignificant. The resulting model is given in Table 8.

TABLE 7 MODEL FOR TOTAL COSTS OF DRIVING A CAR, RESTRICTED SET OF VARIABLES (DEPENDENT VARIABLE: TOTAL DRIVING COSTS OF RESPONDENT)

Independent Variable	Estimated Coefficient	t-Statistic
correction term	0.019	0.822
constant	0.954	21.047
female	-0.132	-3.142
deadline	0.092	2.520
self-employed	0.213	4.694
personal inc	0.000	3.747
summer cottage	0.079	2.645
Number of Observations	611	
R-squared	0.15	
Sum of Squared Residuals	57.693	
Standard Error of the Regression	0.309	
Mean of Dependent Variable	1.103	

TABLE 8 MODEL FOR TOTAL COSTS OF DRIVING A CAR, ALL VARIABLES POSSIBLE (DEPENDENT VARIABLE: TOTAL DRIVING COSTS OF RESPONDENT)

Independent Variable	Estimated Coefficient	t-Statistic
correction term	0.021	0.883
constant	1.076	18.116
female	-0.150	-3.529
self-employed	0.229	5.123
manufact industry	-0.071	-2.379
physic demanding	-0.001	-2.172
personal income	0.000	2.287
summer cottage	0.072	2.396
Number of Observations	604	
R-squared	0.15	
Sum of Squared Residuals	56.055	
Standard Error of the Regression	0.307	
Mean of Dependent Variable	1.103	

The variable "deadline" drops off and the new variables "working in manufacturing industry" and "physically demanding occupation" appear as significant. The coefficients of the two models are similar, which may indicate that the mentioned variables measure the same variance. Because the purpose of these regressions is to assign estimated driving costs to persons who are not drivers currently, the model with fewer variables is preferred. The first model is selected.

These regressions will be used for assessing the driving costs for persons who do not have cars available but who in principle could have one and, thus, choose a car alternative in a mode choice situation (model).

Driving Cost Models for Household Members

This section discusses relationships inside the household that may affect travel behavior. The internal dependency is analyzed through the variables OFC, OVC, and OTC. The regressions are run on the characteristics of the respondent to determine their connections with driving cost levels of the other household members.

OFC of Driving a Car

The model is presented in Table 9. It was created by the same procedure as the previous models. All the variables were introduced and their significance investigated.

The sample was segmented by the person's own income and the household's income. These segmentations did not prove to be worthwhile, and it was deemed best to keep the income variable as simple as possible. The logarithmic transformation of income did not improve the regression.

Among the occupational characteristics, monotonousness, physical demand level, and "little possibility to influence work rhythm" first showed strong negative impact in the regression. When the organizational and location variables were later introduced, the occupational variables lost their statistical significance. In the regression the organizational "blue/white collar" dummy variable effectively captured the variation, even though occupation and organization hardly measure the same things. It may be noted here that Cubukgil and Miller (*13*) found, in a Toronto study, that persons from different organizational status groups have different journey patterns.

As an aside, Cubukgil and Miller divided the blue collar worker group into two according to skill level and found the groups to be located and to behave differently. The location variables may, thus, "eat up" the effects of occupation. It is an interesting chicken-and-egg proposition to ask whether people live where their occupation leads them or choose the kind of occupation that is common in the neighborhood.

The other household members tended to have higher fixed driving costs if the respondent was a woman. This is true even when the correction for the probability of having or not having a car is taken into account. Household income has the expected sign, and the respondent's income is not significant in the model.

The only locational variable that kept its significance was the distance between home and the nearest grocery store. The negative sign of it is unexpected, implying that the further away the nearest store, the less people spend on the fixed cost of driving. The only interpretation at this moment is that the correction term really works here as it should work: it accounts for the probability of having a car (which is presumably higher in a place far away from shops). When this effect is accounted for, the net effect of a long distance to shop may indeed be connected with a lower fixed driving cost.

OVC of Driving a Car

Again, as with variable costs regression for the respondent, the R^2 is very low. The correction term is kept in the regression for theoretical reasons (see Table 10).

The significant variables are "respondent works in manufacturing industry," "house in area of high traffic volume," and "the respondent has high variable driving costs." The last variable correlates with the same phenomenon as "living in a one-family house." Either of these variables gets a statistically significant coefficient. Living in a one-family house enables maintenance work on and storing of cars that in other living arrangements would simply be sold. That the respondent has higher variable costs implies that mechanics may be a hobby or at least that there is some car repair know-how

TABLE 9 MODEL FOR OFC OF DRIVING A CAR (DEPENDENT VARIABLE: FIXED DRIVING COSTS OF THE HOUSEHOLD MEMBERS OF THE RESPONDENT)

Independent Variable	Estimated Coefficient	t-Statistic
correction term	0.021	1.396
constant	0.423	7.073
female	0.103	3.727
shift	-0.061	-1.927
blue-collar	-0.074	-3.074
hh income	0.000	2.675
distance to grocery	-0.000	-2.485
Number of Observations		473
R-squared		0.13
Sum of Squared Residual		26.843
Standard Error of the Regression		0.240
Mean of Dependent Variable		0.494

TABLE 10 MODEL FOR OVC OF DRIVING A CAR (DEPENDENT VARIABLE: VARIABLE DRIVING COSTS OF THE HOUSEHOLD MEMBERS OF THE RESPONDENT)

Independent Variable	Estimated Coefficient	t-Statistic
correction term	0.004	0.430
one	0.477	25.720
manufac ind	0.040	2.172
traffic volume	0.017	2.884
variabl driv costs	0.071	2.340
Number of Observations	474	
R-squared	0.04	
Sum of Squared Residuals	11.201	
Standard Error of the Regression	0.155	
Mean of Dependent Variable	0.539	

in the household, which can be relied on in case of need. These two states appear to coincide.

OTC of Driving a Car

Finally, the total driving cost regression of the respondent's household members is presented.

In the model (Table 11), household characteristics, inner city location, and household income increase the total driving costs. This is expected, because parking and storage costs are higher in a city, and household income is the classical explanation for owning expensive cars.

Physically demanding occupation decreases the respondent's household members' chosen level of driving costs. The physically demanding occupation was also significant in the "total driving costs of the respondent" regression. This characteristic "spills over" its influence to the other members of the household, too.

The other occupational characteristic, little influence on work rhythm, was also significant in the fixed driving costs of the household members' regression. There it covaried with the shift work dummy and was replaced by it. In the total driving cost regression, however, the occupational characteristic is stronger.

CONCLUSIONS

This study started from the need to assess driving cost estimates for the mode choice model. It was approached through separate functions for fixed, variable, and total driving costs of the respondent. More insight was gained by examining within-household effects, which was done by estimating the driving cost regressions of the household members of the respondent.

It is not known which of the costs should be entered in the mode choice logit model. Should one use only the fixed, the variable, the total, or both fixed and variable costs, or perhaps a combination of total and variable costs? Every alternative has a plausible explanation for having a best fit and a reason to it.

TABLE 11 MODEL FOR OTC OF DRIVING A CAR (DEPENDENT VARIABLE: TOTAL DRIVING COSTS OF THE HOUSEHOLD MEMBERS OF THE RESPONDENT)

Independent Variable	Estimated Coefficient	t-Statistic
correction term	0.013	0.808
constant	0.974	15.185
female	0.095	3.469
phys dem work	-0.001	-2.501
little influence	-0.004	-2.462
hh income	0.000	2.904
home in city	0.085	2.379
Number of Observations		470
R-squared		0.12
Sum of Squared Residuals		26.846
Standard Error of the Regression		0.240
Mean of Dependent Variable		1.033

The within-household regressions indicate that there are within-household work life effects on driving costs. These effects take place on a detailed level, which is not captured in a variable describing the household member's employment status. On the basis of these regressions, it appears that the types of occupations of the household members have more effect on the driving costs than the number or share of the employed persons in the household as such.

It remains to be seen whether this information can be used to improve the mode choice, automobile ownership and availability, and daily tour combination models. The estimated cost models provide, however, an explanation for the poor performance of cost variables in situations where the cost is a priori regarded as "the" explanation but approximated to be the same for all drivers. The operating costs are not a flat figure, equal for everybody. They are an outcome of a choice, influenced by the characteristics and activities of the individual and those of other household members. The same applies to capital costs. Capital costs are not equal for everyone, for drivers and nondrivers, or for persons in different occupations, household situations, or stages of life.

It may be mentioned as an aside that the mode choice model developed using the concepts and models presented here provided some surprises. The coefficient of the variable (operating) costs of automobiles was near zero and had a very low statistical significance on mode choice; high capital costs were associated with automobile choice. These findings are not really surprising, but ones not believed nor a part of present travel demand model systems. But it is believable that automobile variable costs do not influence mode choice, or, that once automobile is chosen it is an expensive one—comfort costs—by choice.

Other demand models of the nested system, automobile driver status and travel diary (pattern) choice, are not yet ready. It is known from the mode choice model, however, that the daily travel pattern, travel diary, affects mode choice. That is, mode choices of daily trips of travelers are interdependent.

The main determinants of the other two travel choices in the model system are still unknown. The driving cost models indicate that all travel choices, including the cost of travel, are a part of complex behavioral decisions within a household. Because the market offers a wide variety of transportation choices in terms of costs, there is no surprise involved in finding that the chosen cost level is a part of the mode choice, that is, it belongs to the left-hand side of the equation, and, therefore, is not significant in the right-hand side of a marginal mode choice model. And conversely, assigning equal—but fundamentally arbitrary—automobile costs to all users captures, with that variable and its parameter, effects that, in fact, are behavioral functions and not parameters.

ACKNOWLEDGMENT

Research reported in this paper was conducted when the second author was employed by the Technical Research Center of Finland and received support from the center, the Helsinki Regional Council, and the Finland Highway Administration. The first author was supported, in part, by Yrjö Jahnsson Foundation.

REFERENCES

1. D. McFadden, A. Talvitie, et al. *Demand Model Estimation and Validation. Final Report Series, Vol. V*. Institute of Transportation Studies, University of California, Berkeley, 1977.
2. A. M. Malecki. Perceived and Actual Costs of Operating Cars. *Transportation*, Vol. 7, 1978, pp. 403–415.
3. A. Talvitie and Y. Dehghani. Models for Transportation Level of Service. *Transportation Research B*, Vol. 14B, 1980, pp. 87–99.
4. C. Winston. Conceptual Developments in the Economics of Transportation: An Interpretive Survey. *Journal of Economic Literature*, Vol. 22, March 1985, pp. 57–94.
5. R. Kitamura. A Panel Analysis of Household Car Ownership and Mobility. *Proc., Japan Society of Civil Engineers*, JSCE No. 383/IV-7, July 1987, pp. 13–27.
6. A.-M. Lehto. Naisten ja Miesten Työolot. Helsinki, Tilastokeskus, Tutkimuksia, 1988, 138.
7. Työvoimatutkimus 1984, Työvoimaan kuuluvuus 1985. Helsinki, Tilastokeskus, Tilastotiedotuksia TY 1985:28.
8. Työolotutkimus 1984, Työn haitat 1985. Helsinki, Tilastokeskus, Tilastotiedotuksia TY 1985:18.
9. C. A. Zimmermann. The Life-Cycle Concept as a Tool for Travel Research. *Transportation*, Vol. 11, 1982, pp. 51–69.
10. K. Train. *Qualitative Choice Analysis—Theory, Econometrics, and an Application to Automobile Demand*. MIT Press, Cambridge, Mass., 1986.
11. MVA Consultancy et al. *Value of Travel Time Savings*. Department of Transport, Great Britain. Newbury, Policy Journals, 1987.
12. R. E. Schuler and J. W. Coulter. The Effect of Socio-Economic Factors on the Value of Time in Commuting to Work. *Transportation*, Vol. 7, 1978, pp. 381–401.
13. A. Cubukgil and E. J. Miller. Occupational Status and the Journey-to-Work. *Transportation*, Vol. 11, 1982, pp. 251–276.

Publication of this paper sponsored by Committee on Traveler Behavior and Values.