Activity Duration and Travel time: Alternative Formulations

Sittipan Sittikariya, PhD

Transportation Engineer

DKS Associates, 719 Second Avenue, Suite 1250, Seattle, WA 98104, USA

Tel: (1-206) 382-9800, Fax: (1-206) 587-0692, Email: sxs@dksassociates.com

Venkataraman N. Shankar*, PhD, PE

Associate Professor

Department of Civil and Environmental Engineering, Pennsylvania State University,

226C Sackett Building, University Park, PA 16802, USA

Tel: (1-814) 865-9434, Fax: (1-814) 863-7304, Email: shankarv@engr.psu.edu

*Corresponding Author

Abstract: Activity analysis has been centerstage in the field of travel demand in the past decade. Attention to this field has been significant that a large body of work now places activity modeling at the hub of travel demand forecasting in the coming years. The increase in complexity in demand modeling as a result of activity model architecture brings with it some fundamental questions relating to model construction. What are plausible relationships among key demand metrics, and how does one formulate those relationships in an analytically tractable manner? In this paper we present an illustrative relationship between two basic metrics of activity models -- travel time to activity and activity duration. Using shopping activity as an illustrative example, we present alternative formulations to modeling the relationship between activity duration and travel time to activity. Primary issues relate to the behavior of key time trend effects under the assumptions of individual specific heterogeneity embodied in fixed effects formulations.

INTRODUCTION

Discretionary travel forms a large part of observed daily travel, and is subject to greater spatial and temporal variability than non-discretionary travel such as commutes. As such, a better understanding of discretionary travel is important to a greater understanding of travel in general. The discretionary travel type considered in this study is the evening shopping trip. Shopping is an important activity for daily household maintenance, and evening shopping trips occur during a time of the day when most people have the opportunity to choose to participate in them.

The methodological modeling process used in this analysis attempts to answer the following questions primarily? What is the impact of age on the activity duration engagement of a person, and how does that effect vary across persons? Relatively, what is the impact of travel time viewed in the same perspective, since travel time has been strongly tied to activity duration (Goulias and Ma, 1996).

LITERATURE REVIEW

Activity-based travel demand modeling methods attempt to develop a holistic understanding of travel by including numerous factors influencing activity decisions (Bhat and Koppelman, 1999). Among the recommendations in Bhat and Koppelman, 1999 for future research are: 1) Integrate disparate theories of time-use, such as psychology, anthropology, sociology, geography, urban planning, and economics, to develop a comprehensive theory. 2) Develop joint-activity episode generation and scheduling. 3) Develop understanding of the generation and scheduling process, the internal mechanisms leading to the revealed episode patterns we collect and study. 4) Use GIS to model the spatial contexts of activity episode patterns.

The schedules of non-work trips are more flexible than commuting trips and likely that factors such as socio-demographics and travel demand measures will be more important predictors (Steed and Bhat, 2000). Non-work trips area becoming more important parts of the activity-based travel modeling as their proportion is increasing (Steed and Bhat, 2000). It has been determined that older people avoid shopping in the evening, while higher income people avoid shopping in the morning and afternoon. Students prefer to shop during the evening peak and late evening periods, but employed people prefer to shop later in the evening, avoiding the evening peak period (Steed and Bhat, 2000). The choice to engage in a shopping activity is dependent upon the day of the week (Hirsch et al., 1986).

The time budget for total weekly discretionary activities, considering differences between weekdays and weekends along with allocation between in-home and out-of-home locations has been studied. Out-of-home activity durations tend to increase with the number of vehicles in the household, and to decrease with higher numbers of adults in the household, greater traveler age, longer work duration, and greater distance traveled to work (Bhat, 1996). Correlations between activity duration and the travel time have long been found and documented, in particular, for personal and household maintenance activities. Also, work duration is a significant predictor of activity duration and travel time (Pant and Bullen, 1982).

Of non-work activities, post-work activities are the majority. Post-work activity prediction is therefore important for traffic forecasting (Hamed and Mannering, 1993). Methods to model travelers' post-work activity patterns have been developed. Hamed and Mannering, 1993 analyzed five different post-work activity patterns, they considered selectivity bias correction, and modeled activity type choice, activity duration, and travel time to the activity and back home. They postulate that travel time to a post-work activity and the activity duration are, in general, dependent on each other, with a positive correlation between travel time and activity duration. They suggest that a system of simultaneous equations of travel time and activity duration may be more appropriate than models that consider activity duration and travel time independent.

Endogeneity and selectivity have been viewed to be potential causes of parameter bias in the analysis of activity duration and travel time (see for example, Rosen et al., 2004). The authors found that accounting for endogeneity using a simultaneous equations approach was warranted on the basis of a year of crossectional data from the Puget Sound Regional Council travel survey. The authors also report that selectivity in the observation of shopping trip durations and travel times can cause a shift in the estimated average durations and travel times. The use of selection bias correction methods was explored by Mannering, 1986. He found that methods employing a combination of a discrete choice model, typically a random utility model such as probit or logit, and a continuous regression model can have a significant impact on parameter estimates. Prominent work that has focused in the past on the structural relationship between activity and travel patterns includes for example Goulias and Ma, 1996 and Kuppam and Pendyala (2004). Perhaps the most comprehensive work on longitudinal analysis of travel activity in the U.S. using the Puget Sound panel is attributable to Goulias (1999, 2005). In a longitudinal sense, the questions relating to endogeneity and selectivity change: is there a time trend that is significantly affected by age and the trend varies across persons as a function of other characteristics? Certainly, other questions remain, such as the spatial coverage of activity and associated constraints. But to answer the basic trend question, a better understanding of the behavior of the age effect is necessary. To this end, some have attempted to employ between and within group model structures (see for example, Srinivasan and Athuru, 2005). The authors examine at the household level between and within group effects relating to activity allocation. Our methodological inquiry in this paper begins with a similar question addressed at the person level, hoping to leverage useful information based on all available panels from the PSTP.

DATA

The Puget Sound Transportation Panel was used to generate diaries for persons who remained in the panel over the ten waves of travel surveys in the years 1989, 1990, 1992, 1993, 1994, 1996, 1997, 1999, 2000 and 2002. Two-day travel diaries, household, and personal data were obtained from the Puget Sound Regional Council, yielding information for households and persons who remained in more than two consequent waves. It is noted here that 15% of wave one households from 1989 remained in 2002. The level of retention and participation by wave varies over the decade (see NRGI, 2003). Since observations are made across multiple waves with varying wave participation levels, the panel size varies across groups. For example, if a group is defined as a person participating in more than two waves, then group sizes can be as high as ten years of observation or as low as two. In our case, group sizes with ten years of observation. However, from a modeling standpoint, it must be noted that for any given year, multiple shopping activity measurements can be reported along with associated travel times. Activity duration here is measured as the duration per activity for any of the two days reported by an individual in their survey. Roughly 264 households from the first wave remained in the tenth wave with roughly over 900 distinct individual groups representing two days of observations for any given wave.

ANALYSIS

This section presents the stages of development of an activity-based travel demand model for joint determination of travel time to evening shopping activities and the activity duration. First a straightforward linear regression is used. It is then enhanced with corrections for self selectivity and later endogeneity. The results lead the way to an estimation of travel time and activity duration as a simultaneous system.

Regression Models of Evening Shopping Activity Duration

The analysis begins with a fixed effects model of shopping activity duration that captures within-individual and across-individual variation due to the influence of several factors such as travel time to activity, age, gender, trip activity, car mode use and household income. The fixed effects model is a useful tool for the assessment of variation of heterogeneity across and within individuals. By definition then, this approach creates a somewhat computationally cumbersome problem, which is related to the size of the individual specific coefficient vector. While this is a more practical consideration in the feasibility of fixed effects models, a more important constraint relating to the use of dynamic models with linear specifications stems from the notion of endogeneity. Random effects models are alternatives to fixed effects approaches – for our consideration in this paper, we will limit the discussion to the treatment of the effect as an individual-specific intercept that varies across individuals versus an error term that varies across individuals. One is not interested in computing a vector of individual specific coefficients in random effects models, so while this advantage is appealing, the behavioral intuition behind the use of one versus the other is not that distinguishing. It has been empirically suggested that panels with greater than five years of observation typically will follow a fixed effects structure. Wooldridge, 2002 suggests that attention should be paid to the nature of the regressors in the way they relate to the random error term. The random effects structure implicitly supposes complete exogeneity, which in the context of measurement errors is virtually unavoidable in the modeling of activity duration as a dependent variable. Measurement errors are not only likely to occur in the reporting of travel time if travel time were to be used as a regressor; but furthermore, there is a likelihood of travel time being correlated with error term. In this sense, which is fairly fundamental from a modeling construct, we propose a fixed effects structure. The availability of a decade long panel empirically supports the previously suggested notion that a longer panel might be better suited for fixed effects analysis. Prior to the formal presentation of the fixed effects structure, two points deserve mention. An ordinary least squares approach to the same problem will most certainly overestimate the magnitude of trip related factors while underestimating the magnitude of individual factors, mainly due to the fact that it is not set up to model growth or change.

A Fixed Effects Model of Individual Shopping Activity Duration

The linear regression model with fixed effects is

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + \alpha_i + \delta_t + \varepsilon_{it};\ t = 1,...,T(i),\ i = 1,...,N,$$

$$E[\varepsilon_{it}|\mathbf{x}_{i1},\mathbf{x}_{i2},...,\mathbf{x}_{iT(i)}] = 0,$$

$$\mathrm{Var}[\varepsilon_{it}|\mathbf{x}_{i1},\mathbf{x}_{i2},...,\mathbf{x}_{iT(i)}] = \sigma^2.$$

The parameters of the linear model with fixed individual effects can be estimated by the 'least squares dummy variable' (LSDV). This is computed by least squares regression of *the deviation of the activity duration measurement for individual “i” in year “t” from the individual mean for the decade on the deviation of the exogenous regressors in year “t” from the decadal mean*. The individual specific dummy variable coefficients can be estimated using group specific averages of residuals (Greene, 2003). Essentially since it is a difference estimator, it is often referred to as a within group estimator; however, we note here that we can observe the impact of across individual factor variations on activity duration, once the within individual variation is accounted for. Of some interest here is the impact of parameter uncertainty on the magnitudes and direction of exogenous factor effects on activity duration. In passing, we also note that the fixed effects structure allows for the individual effects to be correlated with the exogenous regressors. Isolating the fixed effects may still be problematic. Since, we are not entirely sure if the estimated fixed effect is capturing just individual heterogeneity or time heterogeneity subsumed as part of an overall heterogeneity, a time effect is included as well. The effect of the time variable is to dampen the magnitude of the age effect which otherwise would be overestimated. Furthermore, we would like to be able to estimate the mean of the dependent variable in the absence of all independent regressors. We do not have here a truly dynamic model, which may aid in the capture of lagged effects, one that would be a natural extension of the structure presented. One may argue that the dependent variable is a positive valued measurement, in which case a nonlinear fixed effects structure involving an exponential function may be appropriate. A slew of methodological issues still remain if one were to pursue a nonlinear structure. The object of this paper is to do exploratory analysis on decadal data with a parsimonious structure that provides some insight into the marginal impact of travel time and other key socio economic variables on individual activity duration.

ESTIMATION RESULTS

The data sample involved over 64 percent female shoppers. Mean activity duration for shopping was approximately 33 minutes, while mean travel time to shopping was approximately 13 minutes, while total trips reported on the day of shopping activity were 6.71 on the average. Approximately 54 percent used car as the mode of travel to shopping, while the average age was 53 years. Roughly 4.3 percent of individuals reported household incomes in excess of 75,000 per year. Table 1 shows the estimation results for a fixed effects model of shopping activity duration with and without time dummies. As the empirical results show, the absence of the time dummy significantly overestimates the mean duration, while the introduction of time dummies removes the time heterogeneity from the intercept estimate. It should be noted here the intercept estimates being discussed (93.592 and 40.160 respectively) refer to the mean in the absence of all regressors, common to all individuals. When one considers the variation of individual effects, they are significant. In the interest of brevity, we do not report the over 900 coefficients estimated. But the range of the effects varies considerably, from a low of -25.5 to 209.70. The impact of travel time is to increase activity duration. The magnitude of this effect varies in the absence of the time dummy. Without the time dummy, the effect is underestimated (0.575 compared to 0.621). The positive effect of travel time confirms the findings of Hamed and Mannering, 1993.

The effect of other travel measures such as total trip making on the day of observed shopping is to decrease activity duration. Similar to the travel time effect, this variable is also underestimated in the absence of the time dummy.

Individual socio-economic characteristics tend to be overestimated (-1.920 compared to -0.201) in the absence of the time dummy. Age in particular is significantly overestimated in its dampening effect. This is consistent with expectation that age would pick up spurious time effects.

Evening shopping activity duration generally tends to increase with income, as noted in prior studies such as the work of Niemeier and Morita, 1996. However it is tempered by a negative effect for people in households with high incomes, in our case, in excess of 75,000 dollars per year. Evening shopping activity duration is shorter for men, which confirms the findings of Niemeier and Morita, 1996. Both effects are overestimated in the absence of the time dummy, but not as significantly as the age effect. Use of the car mode does not appear to shift estimates significantly.

CONCLUSIONS AND RECOMMENDATIONS

We present a fixed effects approach to the analysis of activity duration in a decadal panel of individuals found to be engaged in shopping activity. Albeit parsimonious in our structure, our exploratory analysis suggests that parsing out true time effects and controlling for within-individual variation over time, allows for a better understanding of the marginal influences of key socio-economic and travel factors. It is evident from our analysis that travel factors such as travel time and trip making behavior appear to be under-estimated when time and individual specific effects are not accounted for. In the presence of under-reporting growth in the measurement of trip making, it is difficult to speculate if the shifts in parameter magnitudes would be significant. Socio-economic factors, including age are overestimated without time effects, as one would expect in a growth model. The key is to determine the nature of this growth trend; whether a time dummy in and of itself suffices, and if so in what functional form. We have enforced a linear form in the analysis of this effect. We believe this is limiting in the understanding of true change over time of individuals' shopping duration, or for that matter any activity duration measurement. Other issues remain – such as the assessment of parameter uncertainties and the ability of such models to predict and be feasible forecasting tools. The conditional fixed effects structure was not particularly amenable to forecasting; with unconditional fixed effects structures, this limitation can be avoided.

Model parsimony is an issue – however, when considers the impact of heterogeneity, the real issue of interest is the separation of heterogeneity from state dependence. Heterogeneity can be introduced by the absence of variables; as long as those variables are not significant correlates of the individual's state (example, prior activity duration or activity choice), the problem is more tractable. This is not assured in the context of activity analysis. The simple recourse would be to try a lagged effects model with a "dynamic"component that could vary along two dimensions – within individual across multiple activities on a given day, and across days. A final and important issue for further consideration, is the selectivity bias in the reported sample. We observe shopping duration only for those who reported shopping activity. As prior research (Rosen et al., 2004) suggests, selection bias might be a significant issue in the analysis of shopping duration. The findings from this analysis are by no means definitive; yet, they offer some direction in terms of the analysis of long-term endogeneity and selection bias in longitudinal data. It was important to set up a structure initially to identify the nature of individual heterogeneity and parse out time effects if necessary prior to exploring endogeneity and selection bias. Second, it was important to provide some exploratory basis for a more detailed assessment of the nature of the growth in shopping activity duration within individuals and how that varied across individuals.

REFERENCES

Bhat, C.R. and Koppelman, F.S. (1999) 'Retrospective and prospective survey of time-use research', *Transportation*, **26(**2), pp. 119–139

Bhat, C.R. (1996) 'A hazard-based duration model of shopping activity with nonparametric baseline specification and nonparametric control for heterogeneity', *Transportation Research—Part B*, **30**(3), pp. 189–207

Goulias, K.G. and Kim, T. (2005) 'Behavioral dynamics in activity participation, travel, and information and communications technology', *Transportation and Traffic Theory: Flow, Dynamics and Human Interaction*, Elsevier, pp. 467-486

Goulias, K.G. (1999) 'Longitudinal analysis of activity and travel pattern dynamics using generalized mixed Markov latent class models', *Transportation Research—Part B*, **33**(8), pp. 535-558

Goulias, K.G. and Ma, J. (1996) 'Analysis of longitudinal data from the Puget Sound transportation panel - Task B: integration of PSTP databases and PSTP codebook', Final Report. Submitted to Federal Highway Administration and Battelle Memorial Institute Work Order BAT 94-016, PTI 9619, University Park, PA

Greene, W.H. (2003) *Econometric Analysis*, Prentice Hall, New Jersey

Hamed, M.M. and Mannering, F.L. (1993) 'Modeling travelers' post work activity involvement: toward a new methodology', *Transportation Science*, **27(**4), pp. 381–394

Hirsch, M., Prashker, J.N. and Ben-Akiva, M. (1986) 'Day-of-the-week models of shopping activity patterns', *Transportation Research Record*, **1085**, pp. 63-69

Kuppam, A.R. and Pendyala, R.M. (2004) 'A structural equations analysis of commuters' activity and travel patterns', *Transportation*, **28(**1), pp. 33-54

Mannering, F.L. (1986) 'Selectivity bias in models of discrete and continuous choice: an empirical analysis', *Transportation Research Record*, **1085**, pp. 58–62

Niemeier, D.A., and Morita, J.G. (1996) 'Duration of trip-making activities by men and women: a survival analysis', *Transportation*, **23(**4), pp. 353–371

Northwest Research Group Inc. (2003) *Puget Sound Transportation Panel Wave 10*, October 2002 – February 2003, Summary Report

Pant, P.D. and Bullen, A.G.R. (1982) ‘Urban activities, travel, and time: relationships from national time-use survey’, *Transportation Research Record*, **750**, pp.1–6

Rosen, D.M., Shankar, V.N. and Ulfarsson, G.F. (2004) ‘The relationship between shopping travel time and activity duration and spatial census tract-level and transportation network factors’, *Transportation Research Record*, **1894**, pp. 160-169

Srinivasan, K.K. and Athuru, S.R. (2005) ‘Analysis of within-household effects and between-household differences in maintenance activity allocation’, *Transportation*, **32(**5), pp. 495-521

Steed, J.L. and Bhat, C.R. (2000) ‘On modeling departure-time choice for home-based social / recreational and shopping trips’, *Transportation Research Record*, **1706**, pp. 152–159

Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press

Table 1 Fixed Effects Model of Shopping Activity Duration With and Without Time Effects

Variable	**Fixed Effects Without Time Effects+**	**Fixed Effects With Time Effects+**
	β* **σ **** **t*****	**β*** **σ **** **t*****
Constant	93.592 12.963 7.220	40.160 12.600 3.187
Travel time to shopping activity	0.575 0.049 11.840	0.621 0.046 13.397
Total number of trips on the same day of shopping activity	-1.338 0.219 -6.112	-1.512 0.209 -7.245
Age of individual	-1.204 0.228 -5.287	-0.201 0.073 -2.771
Indicator, 1 if individual is female	17.891 8.208 2.180	9.516 2.399 3.967
Indicator, 1 if individual's income > $75,000 per year	-8.856 3.495 -2.506	-6.133 3.111 -1.971
Indicator, 1 if individual travels by car	-8.423 1.249 -6.745	-8.208 1.179 -6.692
R^2	0.381	NA
Adjusted R^2	0.252	NA

+ Additional Constant for Time Effects of each Individual

* Estimated Coefficient, ** Standard Error, *** t-statistic