CHAPTER 3

Demand Model Specifications and Estimation Results

This chapter provides a detailed technical discussion of the main focus of the C04 research project, which was the specification and estimation of new advanced forms of travel demand models that aim to substantially improve how road pricing and congestion can be more fully and realistically modeled for transportation policy and planning.

This chapter describes the results of model estimation research in terms of the somewhat more general findings presented in two key subsections: Overview of Section, Approach, and Main Findings; and Summary Comparison and Synthesis.

The first subsection provides in-depth discussion of the conceptual and behavioral framework adopted for the C04 research and the wide range of possible responses considered to congestion and pricing. The rest of the chapter, which focuses on model estimation results developed with data from New York and Seattle, Washington, is organized in a two-dimensional fashion, with the major subsections organized by *model types* (route choice, time-of-day [TOD] choice, mode choice, and other choice dimensions) and the minor subsections organized by *model features* (main utility specification, segmentation, the incorporation of reliability, and other important model properties). The corresponding components developed and tested with the different models estimated in the course of the C04 research, using the data described in Chapter 2, are then presented back-to-back, focusing on each of the principal model choices and features proposed for improvement.

Given the complexity of the modeling issues and the need to adequately document the methods developed and applied for the C04 research for advanced modeling, Chapter 3 is necessarily technically detailed in nature. For a more thorough understanding of the topic, the technical reader is referred to an unabridged version of chapter with full detail in the discussion of all conceptual and technical details associated with the specification and estimation of the demand models addressed in this research project. The reader can find a full discussion of each model topic and data analysis in the unabridged, unedited Chapter 3 online: www.trb.org/Main/Blurbs/168141.aspx. Appendix A, which provides the full statistical results for the key models that were estimated for this study and are discussed in this chapter, is organized according to the specific types of models estimated (route choice, mode choice, and so forth) and the specific data sets used.

Structural Dimensions for Analysis of Congestion and Pricing Impacts on Demand

Possible Choice Dimensions

The behavioral framework adopted for the C04 research has been constructed to include a wide range of possible responses to congestion and pricing, organized as shown in Table 3.1 in an approximately hierarchical order from the short term to the long term.

Most of the existing models for pricing (both in research and practice) have been largely focused on the subset of trip-level short-term responses, including route, preroute, car occupancy, mode choice, and departure time choice (Brownstone et al. 2003; Brownstone and Small 2005; Lam and Small 2001; Mahmassani et al. 2005; Mastako 2003; Verhoef and Small 2004). Within this limited framework, there have been only a few examples of a full integration across all these choices: in the existing activity-based models (ABMs) developed for Columbus, Ohio (PB Consult, Inc. 2005) and Montreal, Quebec (PB Consult, Inc. 2003).

There are, however, many other important travel dimensions that have been less explored. Long-term impacts of congestion and pricing may include fundamental changes in travel behavior patterns that cannot be captured and understood at the single trip level. For example, in urban over congested areas like New York, Chicago, and San Francisco, many employers offer workers a compressed work schedule of four 10-hour days. This new choice dimension can have a significant impact on the amount of travel produced and its temporal distribution. This choice, however, is clearly not a trip-level decision

Table 3.1. Possible Responses to Congestion and Pricing

Choice Dimension	Time Scale for Modeling	Expected Impact
Network route choice	Short term—trip episode	Stratified response by user group
Preroute choice (toll versus nontoll)	Short term—trip episode	Stratified response by user group
Car occupancy	Short term—tour or trip episode	Planned and casual carpool
Mode choice	Short term—tour or trip episode	Shift to transit, especially to rail and for low- to medium-income groups
TOD or schedule	Short term—tour or trip episode	Peak spreading
Destination or stop location	Short term—tour or trip episode	Improved accessibility effect combined with negative pricing effect on trip distribution for nonwork trips
Joint travel arrangements	Short term—within day	Planned carpool or escorting
Tour frequency, sequence, and formation of trip chains	Short term—within day	Lower tour frequency and higher chaining propensity
Daily pattern type	Short term—weekly (day to day)	More compressed workdays and work from home
Usual locations and schedule for nonmandatory activities	Medium term—1 month	Compressed or chain patterns; weekly planned shopping in major outlets
Household or person mobility attributes (transponder, transit path, parking arrangements at work)	Medium term—1 to 6 months	Higher percentage of transponder users and parking arrangements for high incomes, higher percentage of transit path holders for low incomes
Household car ownership choice	Long term—1 year	Stratified response by income group: Higher car ownership for high incomes, lower car ownership for low incomes
School or university location and schedule	Long term—1 to 5 years	Choice by transit accessibility; flexible schedules
Job or usual workplace location and schedule	Long term—1 to 5 years	Local jobs for low incomes; compressed or flexible schedules
Residential location	Long term—5 years +	Income stratification: High-income suburbs around toll roads, low-income clusters around transit
Land use development	Long term—5 years +	Urban sprawl if no transit; otherwise shift to transit

comparable to a choice between managed and free lanes (or between toll and nontoll roads) for a particular trip. Choices such as this should be modeled within a proper behavioral framework, including an extended time scale, with a robust set of explanatory variables, and linkages to the other short-term and long-term choices (Pendyala 2005; Spear 2005).

In general, the important multiple possible behavioral responses that are beyond a traditional trip-level modeling of choices can be grouped into the following broad classes:

- Trip or tour destination choice that is equally important for both ABMs and four-step models. It is normally assumed that impacts of congestion and pricing should be captured through the generalized cost or mode choice logsum (Erhardt et al. 2003; Dehghani and Olsen 1999); however, there can be more direct and specific impacts that are worth exploring.
- Short-term choices that relate to daily activity patterns that cannot be fully captured at the elemental trip level. These choices include explicit joint travel arrangements (Vovsha et al. 2003; Vovsha and Petersen, 2005), tour formation (Parsons Brinckerhoff Quade and Douglas, Inc. et al. 2005), and daily pattern type (PB Consult, Inc. 2005) (e.g., the decision to stay at home on a given day). These choices can be applied only in an ABM framework (though there might be an additional use of this for four-step models in order to investigate congestion and pricing impacts on trip generation). It is important to address these dimensions along with the conventional trip dimensions because many of the new pricing forms are not trip based (e.g., daily area pricing schemes applied in London [Litman 2005] and currently envisioned or modeled in New York and San Francisco).
- Medium-term choices that relate to choice of usual location and schedule for nonmandatory activities (like shopping or entertainment). It might be beneficial for a deeper understanding and ability to forecast such choices to put certain choices into a medium-term framework in order to explore the impacts of congestion and pricing beyond a short-term single-trip consideration. This type of choice can be incorporated in an advanced ABM only.
- Medium- or long-term choices that relate to person or household mobility attributes like car ownership, transponder, transit path, and parking arrangements. There is a growing recognition of the importance of these choices in understanding and modeling impacts of congestion and pricing.

There have been some initial attempts to formulate and estimate choice models related to the acquisition of transponders (Yan and Small 2002) simultaneously with preroute, departure time, or car occupancy (or some combination of these factors), although the estimation was implemented at the single-trip level.

- Long-term location choices of residential place, workplace, and school, as well as land use development impacts. A special methodology for analysis of congestion and pricing impacts on these choices has not yet been developed. The existing long-term models of this type operate with standard trip-level measures of accessibility (Vovsha, Davidson et al. 2005); thus, the effect of a different and extended time scale is lost. The team plans to explore data sets that include information on long-term choices (along with trip records, of course) in order to ascertain the differential impacts of congestion and pricing over various time scales.

This classification of possible choice dimensions is incorporated in the formulation of a comprehensive conceptual model of travel behavior that served as the starting point in C04 for the specification of model systems that could be estimated with the selected data sets. Several of these choice dimensions represent relatively new choice models that have not yet been widely accepted and explored (only first attempts to formulate and estimate these models have been made and reported). These include integration of the binary preroute choice (toll versus nontoll) in the mode choice nested structure, payment type (cash, E-ZPass, transponder) and associated vehicle equipment, as well as models of carpooling mechanisms (explicit modeling of joint travel).

Functional Forms for Highway Utility (Generalized Cost)

As described in Chapter 1 in Highway Utility Components, the highway travel utility function is a basic expression that combines various level-of-service (LOS) and cost attributes as perceived by the highway user. It is directly used in the highway trip route choice (e.g., between the managed lanes and general-purpose lanes on the same facility). It also constitutes an essential component in mode and TOD choice utilities. The form of the highway utility function is also important for modeling other (upper-level) travel choices, as it serves as the basis for accessibility measures. Thus, it is essential to explore the highway travel utility function and its components before considering a simplified framework of route choice in the highway network, because the complexity builds when additional choice dimensions are considered.

In most travel demand models, including those developed for practical and research purposes, the highway utility function (U) takes the following simple form:

$$U = a \times T + b \times C \tag{3.1}$$

where

T = travel time;
C = travel cost;
$a < 0$ = coefficient for travel time;
$b < 0$ = coefficient for travel cost; and
a/b = value of time (VOT).

Coefficients for travel time and cost normally take negative values, reflecting the fact that travel in itself is an onerous function necessary only for visiting the activity location. Thus, the travel utility is frequently referred to as the "disutility" of travel. However, in some research, the negative character of travel utility has been questioned in some contexts. In particular, a positive travel utility was seen to be associated with long recreational trips on weekends (Stefan et al. 2007). Also, an interesting effect was observed for commuting trips, on which commuters seem to prefer a certain minimum time and are not interested in reducing it below a certain threshold (Redmond and Mokhtarian 2001).

More importantly, it is clear that the standard representation of highway travel utility as a linear function of two variables with constant coefficients is an extremely simplified one. A great deal of the C04 research effort has been devoted to the elaboration of this basic form in the following ways:

- Investigation of the highway user perception of travel time by congestion levels. This means that a simple generic coefficient for travel time could be replaced with the coefficients differentiated by congestion levels;
- Inclusion and estimation of additional components, of which travel time reliability has been currently identified as the most important. Reliability is seen as an additional and nonduplicating term along with the average travel time and cost; and
- Testing more complicated functional forms that are nonlinear in time and cost, as well as account for randomly distributed coefficients or VOT (in addition to any explicit segmentation accounting for the observed user heterogeneity). With these enhancements, VOT is no more assumed as a constant, but becomes a varying parameter depending on the absolute values of time and cost as well as reliability.

As a working model the team has adopted the following general expression for the highway travel utility that will be explored component-by-component in the current research:

$$U = \sum_{k=1}^{5} [a_k \times \varphi_k(T_k)] + \sum_{m=1}^{3} [b_m \times \phi_m(C_m)] + \sum_{n=1}^{3} c_n R_n \tag{3.2}$$

where

$k = 1$ represents the uncongested highway travel time component;
$k = 2$ represents the congested highway travel time component (extra delay);

$k = 3$ represents parking search time;
$k = 4$ represents walk access or egress time (e.g., from the parking lot to the trip destination);
$k = 5$ represents extra time associated with carpooling (picking up and dropping off passengers);
$T_k =$ (average) travel time by component;
$m = 1$ represents highway toll value;
$m = 2$ represents parking cost;
$m = 3$ represents vehicle maintenance and operating cost;
$C_m =$ travel cost value by component;
$n = 1$ represents disutility of time variation (first measure of reliability);
$n = 2$ represents schedule delay cost (second measure of reliability);
$n = 3$ represents utility of (lost) activity participation (third measure of reliability);
$R_n =$ reliability measures by component;
$a_k, b_m, c_n =$ coefficients to be estimated; and
$\varphi_k(\ldots), \phi_m(\ldots) =$ functions for nonlinear transformation of time and cost variables.

This formulation makes it more difficult to calculate VOT, although it is still possible. In the same way, value of reliability (VOR) can be calculated for the first type of reliability measure (assuming that this reliability measure is in minutes). VOR essentially represents travelers' willingness to pay for reduction in travel time variability in the same way as VOT represents their willingness to pay for (average) travel time savings. More specifically, in the context of willingness to pay tolls for saving time in congestion conditions, VOT can be calculated by the following general formula:

$$VOT(T_2, C_1) = \frac{\partial U/\partial T_2}{\partial U/\partial C_1} = \frac{a_2 \varphi_2'(T_2)}{b_1 \phi_1'(C_1)} \tag{3.3}$$

A similar calculation can be implemented for VOR. With nonlinear transformation functions, VOT and VOR are no longer simply constant values. They now vary and depend on the absolute values of time and cost variables at which the derivatives of the transformation functions are taken.

The innovative components of the C04 research that relate to perceived highway time, travel time reliability, and nonlinear transformations are discussed in the subsequent sections. It should be noted that some components, specifically perceived travel time and some reliability measures, might be correlated statistically (and are also conceptually duplicative to some extent). Thus, it is highly improbable that the entire formula (Equation 3.2) would ever be applied. Instead it serves instead as a conceptual framework in which proposed model structures can be derived and statistically tested.

Dimensions for Model Segmentation

Another long-term gap in the understanding and the modeling of congestion and pricing is associated with inadequate segmentation of population and travel. It has been generally recognized by the both research and practitioner communities that the profession needs to advance beyond the crude average VOT estimates (and other related behavioral parameters) obtained from aggregate analyses (Hensher and Goodwin 2003).

There is a significant amount of research providing insights into behavioral mechanisms and statistical evidence on the heterogeneity of highway users across different dimensions. Although income and trip purpose have been traditionally used in many models as the main factors that determine VOT, in reality VOT is a function of many other variables. In fact, in many cases, income and trip purpose might not even be the most important factors, especially when situational factors and time pressure come into play (Spear 2005; Vovsha, Davidson et al. 2005).

A variety of traveler and trip-type dimensions are understood to be important. The research team distinguishes between the following main groups:

- Socioeconomic Segments of Population. These characteristics are exogenous to all activity and travel choices that are modeled in the system. Thus, the corresponding dimensions can always be applied for any model, either for a full segmentation or as a variable in the utility function;
- Segmentation of Activities. These characteristics are exogenous to travel choices, but endogenous to activity-related choices. Thus, in the applied model system, it is necessary that the corresponding activity choices are modeled prior to the given model; otherwise they cannot be used for the model segmentation; and
- Travel Segmentation. These characteristics are endogenous to the system of travel choices. In model estimation, they have to be carefully related to the model structure to ensure that all dimensions and variables used in each particular model have been already modeled in the model chain.

The socioeconomic segmentation of population may best be addressed by the following:

- Income, Age, and Gender. A higher income is normally associated with higher VOT (Brownstone and Small 2005; Dehghani et al. 2003). Middle-age female status has also been associated with higher VOT (Mastako 2003; PB Consult, Inc. 2003);
- Worker Status. Employed persons (even when traveling for nonwork purposes), because of their tighter time constraints, are expected to exhibit a higher VOT than nonworkers; and
- Household Size and Composition. Larger households, with children, are more likely to carpool and take advantage of managed lanes (Stockton, Benz et al. 2000; Vovsha et al. 2003).

The segmentation of activities may best be addressed by the following:

- Travel Purpose. Work trips, and, in particular, business-related trips, normally are associated with higher VOT than trips for nonwork purposes (Dehghani et al. 2003; Parsons Brinckerhoff Quade and Douglas, Inc. et al. 2005; PB Consult, Inc. 2003). A frequently cited high-VOT trip purpose is a trip to the airport to catch an outbound flight (Spear 2005). A list of special trip purposes with high VOT might also include escorting passengers, visiting a place of worship, a medical appointment, and other fixed-schedule events (e.g., theater or sport events). A deeper understanding of the underlying mechanisms for such behavior would be valuable, including a combination of factors such as schedule inflexibility, low trip frequency, and situational time pressure.
- Day of Week. Weekday versus Weekend. There is statistical evidence that VOT for the same travel purpose, income group, and travel party size on weekends is systematically lower than on weekdays, including some examples of positive travel utility associated with long discretionary trips (Stefan et al. 2007). It is yet to be determined if these differences can be explained by situational variables, or if there is an inherent weekend type of behavior that is different from regular weekday behavior. In any case, whether directly or for a proxy for situational time pressure, it would be useful to test the differences statistically. A positive utility of travel has been found, most notably in the choice of distant destinations for discretionary activities on weekends (perhaps with a sightseeing or excursion component). This utility should be explored, however, to see if it is actually correlated with tolerance of congestion delays or unwillingness to pay tolls.
- Activity and Schedule Flexibility. Fixed-schedule activities are normally associated with higher VOT for trips to activity because of the associated penalty of being late; this has manifested itself in previous research that documents that VOT for the morning commute is higher than for the evening commute. Probably a similar mechanism for trips to airports (high penalty of being late) creates higher VOT estimates. The team also expects that schedule flexibility will be an important factor for nonwork activities; for example, a trip to a theater might exhibit a high VOT, but shopping might be more flexible.

The segmentation of travel can be best addressed with the following:

- Trip Frequency. Regular trips and their associated costs may receive more (or less) formal consideration than those that occur infrequently. For example, $1.50 for an auto trip to work may be perceived as $3.00/day (assuming a symmetric toll) and $60/month, thus receiving special consideration. This perceptional mechanism is likely to be very different for infrequent and irregular trips when the toll is perceived as a one-time payment. For intercity trips, travelers' recognition of the return trip is not obvious, since it may occur on a different day.
- TOD. Research confirms that a.m. and p.m. peak periods are associated with higher VOT than off-peak periods and that a.m. travelers (mostly commuters) are more sensitive to both travel time and reliability than p.m. commuters (who mostly are returning home) (Brownstone et al. 2003). However, few have explored how these phenomena relate to schedule flexibility, or how TOD factors affect VOT for nonwork trips;
- Vehicle Occupancy and Travel Party Composition. Although a higher occupancy normally is associated with higher VOT (though not necessarily in proportion to party size), it is less clear how travel party composition (e.g., a mother traveling with children, rather than household heads traveling together) affects a party's VOT.
- Trip Length or Distance. An interesting convex-shape function has been estimated for commuters' VOT (Steimetz and Brownstone 2005). For short distances, VOT is comparatively low since the travel time is insignificant, and delays are tolerable. For trip distances around 30 miles, VOT reaches a maximum; however, for longer commuters VOT goes down again, because they presumably have self-chosen residential and work places based on the long-distance travel. Additionally, in the context of mode choice, strong distance-related positive biases have been found for rail modes in the presence of congestion (as a manifestation of reliability [Parsons Brinckerhoff Quade and Douglas, Inc. et al. 2005]) and carpools (since carpools are associated with extra formation time).
- Toll Payment Method. This is an important additional dimension that has not been explored in detail. An analysis done by the Port Authority of New York and New Jersey has shown that the introduction of E-ZPass at its tolled crossings attracted a significant new wave of users despite a relatively small discount (Holguín-Veras et al. 2005). In the same way in which transportation analysts speak about perceived time, we should also probably speak about perceived value of money in the context of pricing. Bulk discounts and other nondirect pricing forms should be modeled at the daily pattern level rather than trip level. We also have to understand the impact of congestion on the whole daily pattern rather than by single trips, including analysis of daily time budgets and trade-offs made to overcome congestion (including work from home, compressed workweeks, compressed shopping, and moving activities to weekends).
- Situational Context. Time Pressure versus Flexible Time. This is recognized as probably the single most important factor determining VOT that has proven difficult to measure

and estimate explicitly, as well as to include in applied models (Spear 2005; Vovsha, Davidson et al. 2005). There is evidence that even a low-income person would be willing to pay a lot for travel time savings if he or she were in a danger of being late to a job interview or were escorting a sick child. This factor is correlated with the degree of flexibility in the activity schedule (inflexible activities, trips to airport, fixed schedules, and appointments will be the activities most associated with time pressure), but does not duplicate it. For example, for a high-income person traveling to the airport, the VOT might not be relatively low if this person has a 4-hour buffer before the departure time. With ABMs, the analyst could use the number of trips or activities implemented by the person in the course of a day, as well as the associated time window available for each trip or activity, as an instrumental proxy for time pressure.

In model formulation, estimation, and application, it is crucial to follow a conceptual system design and obey the rules of application of those variables that are exogenous to the current model. For example, if the TOD model is placed after mode and occupancy choice, then mode and occupancy can be used as the TOD model segmentation. However, TOD in this case cannot be used for segmentation of the mode and occupancy choice models. If the order of models is reversed (TOD choice before mode and occupancy choice), then the segmentation restrictions would also be reversed. When different models are estimated it is essential to keep a conceptual model system (or at least a holistic framework as described below) in mind in order to make these models compatible and avoid endogeneity–exogeneity conflicts.

It should be understood that all these dimensions cannot be simultaneously included in operational models as explicit segments in Cartesian combination. With a four-step model framework, this would immediately result in an unfeasibly large number of trip tables. The disaggregate basis of the ABM framework is more flexible, and theoretically can accommodate any number of segments. They are, however, limited in practical terms by the sample size of the travel survey (normally several thousands of individuals), which quickly wears thin for multidimensional segments. However, there are other ways to constructively address segmentation in operational models. They include flexible choice structures with parameterized probabilistic distribution for parameters of interests (e.g., VOT), as well as aggregation of segments by VOT for assignment and other model components that are especially sensitive to dimensionality.

It should also be understood that VOT represents only one possible behavioral parameter, and that it is essentially a derived one. In most model specifications and corresponding estimation schemes, VOT is not directly estimated, but rather derived either as the ratio of the time coefficient to cost coefficient (in simple linear models as specified in Equation 3.1) or as the marginal rate of substitution between time and cost (in a general case as specified in Equation 3.3). Thus, very different behaviors can be associated with the same VOT. For example, both time and cost coefficients can be doubled, which leaves VOT unchanged; however, there would be very different estimated responses to congestion and pricing in these two models. Large coefficients will make the model more sensitive to any network improvement or change in costs, whereas smaller coefficients will make it less sensitive.

One of the most detailed VOT segmentation analyses of the type described in the previous subsection was carried out for the Netherlands National Value of Time study (Bradley and Gunn 1990), which used 10 simultaneous segmentation variables. A similar approach was used for national studies in the United Kingdom and Sweden.

All else being equal, a more detailed segmentation typically tends to dampen the overall price sensitivity across the population, since a typical sigmoid response curve, like the logit model, has the steepest (most elastic) part in the middle, but the ends are quite flat, and market segmentation tends to move distinct groups away from the middle.

Measures of Travel Time Reliability

In general, four possible methodological approaches to quantifying travel-time reliability are either suggested in the research literature or already applied in operational models:

- Indirect Measure: Perceived Highway Time by Congestion Levels. This concept is based on statistical evidence that in congestion conditions, travelers perceive each minute with a certain weight (Small et al. 1999; Axhausen et al. 2007; Levinson et al. 2004; McCormick Rankin Corp. and Parsons Brinckerhoff 2008). Perceived highway time is not a direct measure of reliability because only the average travel time is considered, although it is segmented by congestion levels. It can serve, however, as a good instrumental proxy for reliability because the perceived weight of each minute spent in congestion is in part a consequence of associated unreliability.
- First Direct Measure: Time Variability (Distribution). This is considered as the most practical direct approach and has received considerable attention in recent years. This approach assumes that several independent measurements of travel time are known that allow for forming the travel time distribution and calculation of derived measures, such as buffer time (Small et al. 2005; Brownstone and Small 2005; Bogers et al. 2008). One of the important technical details with respect to the generation of travel time distributions is that even if the link-level time variations are known, it is a nontrivial task to synthesize the O-D–level time

distribution (reliability skims) because of the dependence of travel times across adjacent links due to a mutual traffic flow. The implementation challenge posed by this issue was specifically addressed in the course of this project.

- Second Direct Measure: Schedule Delay Cost. This approach has been adopted in many academic research works on individual behavior (Small 1982; Small et al. 1999). According to this concept, the direct impact of travel time unreliability is measured through cost functions (penalties in expressed in monetary terms) of being late (or early) compared with the planned schedule of the activity. This approach assumes that the desired schedule is known for each person and activity in the course of the modeled period. This assumption, however, is difficult to meet in a practical model setting.
- Third Direct Measure: Loss of Activity Participation Utility. This method can be thought of as a generalization of the schedule delay concept. It is assumed that each activity has a certain temporal utility profile and that individuals plan their schedules to achieve maximum total utility over the modeled period (e.g., the entire day) taking into account expected (average) travel times. Any deviation from the expected travel time due to unreliability can be associated with a loss of a participation in the corresponding activity (or gain if travel time proved to be shorter) (Supernak 1992; Kitamura and Supernak 1997; Tseng and Verhoef 2008). Recently this approach was adopted in several research works on dynamic traffic assignment (DTA) formulation integrated with activity scheduling analysis (Kim et al. 2006; Lam and Yin 2001). Similar to the schedule delay concept, however, this approach suffers from data requirements that are difficult to meet in practice. The added complexity of estimation or calibration of all temporal utility profiles for all possible activities and all person types is significant. This makes it unrealistic to adopt this approach as the main concept for the current project. This approach, however, can be considered in future research efforts.

The details of each approach are considered in the subsequent sections.

Perceived Travel Time Weights by Congestion Levels

Variations in the perceived utility of components of transit travel time have been long recognized and used in travel models. For example, in most mode choice models and transit assignment algorithms, out-of-vehicle transit time components like *wait* and *walk* are weighted compared with in-vehicle travel time. It is not unusual to apply weights in the range of 1.5 to 4.0, reflecting that travelers perceive out-vehicle time as more onerous than in-vehicle time.

In contrast to transit modeling practice, virtually all travel models used for highway analysis include a single generic term for highway time; that is, the same coefficient is applied for each minute of highway time regardless of travel conditions. However, there is some compelling statistical evidence that highway users perceive travel time differently by congestion levels. For example, driving in free-flow conditions is likely to be perceived less negatively than driving in heavily congested (stop-and-go) conditions. It is an intuitive and behaviorally appealing notion that highway users driving in congested conditions might perceive the longer travel time as an additional delay or penalty on top of free-flow (or some expected reasonable) time. With a segmentation of travel time coefficients by congestion levels, the time spent on links with congested conditions is expected to have a larger disutility. A larger disutility associated with congestion would have at least two behavioral interpretations:

- A negative psychological perception that is similar to the weight for walking to or waiting for transit service; and
- A simplified operational proxy for reliability that should be explored in combination with the explicit reliability measures.

Several research studies report statistical evidence of quite high perceptional weights that highway users put on travel time in congested conditions (Small et al. 1999; Axhausen et al. 2007; Levinson et al. 2004; McCormick Rankin Corp. and Parsons Brinckerhoff 2008; Wardman et al. 2009). Also, there have been multiple indications in recent analyses of travel surveys that the perception of the time saved by respondents in the revealed preference (RP) survey is about twice the actual measured time saved (Small et al. 2005; Sullivan 2000). In the RP framework, this might well be a manifestation that travelers operate with perceived travel times in which time spent traveling through congested segments is psychologically doubled.

In order to illustrate the magnitude of the possible weights, as well as possible approaches to differentiate travel time by congestion levels, three examples of estimated perceptions of travel time are discussed below. It should be noted that in both cases, the approaches are very simple to implement on the supply side. The network simulation can be performed, and the required LOS skims can be generated by static assignment methods, although DTA could offer additional benefits. This technique can be easily applied with both ABMs and four-step models.

In the first example (Small et al. 1999), travel time was broken into two parts:

- Time in uncongested conditions (LOS A to D); and
- Time in congested conditions (LOS E to F, i.e., close to the stop-and-go condition).

Table 3.2. Cost of Shifting 1 Minute from Uncongested to Congested Time

Total Travel Time (min)	Cost of Shifting 1 min from Uncongested to Congested Time ($)	Equivalent in VOT ($/h)
10	0.77	46.2
15	0.51	30.6
20	0.30	18.0
30	0.26	15.6
45	0.17	10.2
60	0.13	7.8

The choice framework presented in the stated preference (SP) survey context included only route choice. Travel time and cost variables were not estimated, but were stated in the SP questionnaires. The highway utility expression included total time, cost, and percentage of congested time. Using the previously introduced notation, the adopted utility specification can be written in the following way:

$$U = a \times (T_1 + T_2) + b \times C + c \times \frac{T_2}{T_1 + T_2} \tag{3.4}$$

This expression is different from the suggested formula (Equation 3.2), but it could be transformed into an equivalent formula with certain assumptions (fixed total travel time). The estimation results confirmed a very high significance for the additional term of percentage of congested time. The authors translated it into a recommended mark-up value of 2.5 to VOT savings under congested conditions compared with uncongested conditions. More detailed estimation results are summarized in Table 3.2. By virtue of the specified utility function, the cost of shifting 1 minute from uncongested to congested time is dependent on the total travel time. For an average time of 30 minutes, the VOT equivalent of the additional perceived burden associated with only congestion itself is about $15/hour, which is roughly equal to the average commuting VOT applied in most models.

The second example is taken from the recently completed travel demand model for the Ottawa–Gatineau, Canada, region (McCormick Rankin Corp. and Parsons Brinckerhoff 2008). The model framework, choice context, and utility formulation were different from those used in the Small et al. (1999) study. However, the bottom-line results look similar in many respects. In this study, a mode choice model was estimated for five travel purposes and two TOD periods (a.m. and p.m.) based on the RP data from the large household travel survey (23,870 households representing a 5% sample). Travel time and cost variables were provided from static assignment equilibrium skims from the modeled network.

The highway utility included travel cost with one generic coefficient and travel time broken into the following two components—note that this breakdown of travel time is different from the one adopted for Small et al.:

- Free-flow (minimal) time; and
- Extra delay, calculated as congested time minus free-flow time for the entire origin–destination (O-D) path.

The highway utility function had the following form:

$$U = a_1 \times T_1 + a_2 \times T_2 + b \times C + \sum_s (d_s \times h_s) \tag{3.5}$$

where

s = additional mode-specific constants and household or zonal variables;
h_s = values of additional variables; and
d_s = estimated coefficients.

The estimation results are shown in Table 3.3 as translated into VOT terms. They indicate that for several segments, specifically a.m. and p.m. work trips, as well as p.m. discretionary trips, each minute of congestion delay is perceived as about twice as onerous as the free-flow (minimal) time component. For other segments, however, statistical tests did not show a significant difference between free-flow and congestion time components; hence, two coefficients were pooled together.

Table 3.3. VOT Estimates for Free-Flow Time and Congestion Delay

	VOT ($/h)			
	a.m.		p.m.	
Trip Purpose	Free-Flow Time	Congestion Delay	Free-Flow Time	Congestion Delay
Work	22.2	42.7	19.4	40.0
University	10.0	10.0	11.0	11.0
School	5.1	5.1	5.1	5.1
Maintenance	10.7	10.7	12.1	12.1
Discretionary	9.0	9.0	11.4	29.3

Table 3.4. Highway Time Weight by Congestion Levels

Travel Time Conditions	United Kingdom	United States
Free flow	1.00	1.00
Busy	1.05	1.03
Light congestion	1.11	1.06
Heavy congestion	1.31	1.20
Stop and start	1.20	1.38
Gridlock	1.89	1.79

The third example is taken from the research work of Wardman et al. (2009), who provided new evidence on the variation in the valuation of motorists' travel time savings across a finer gradation of traffic-condition types (six levels of congestion) than had been previously attempted by means of analyzing SP data collected from different tolled roads in the United Kingdom and the United States. The summary of the time relativities is presented in Table 3.4. The study further supports a finding that a reasonable value for the perceived time weight in congested conditions lies in the range 1.3 to 2.0.

Mean Variance, Buffer Time, and Other Time Variability Measures

Time variability can be measured by any compact measure associated with a travel time distribution (e.g., any combination of the mean, dispersion, or higher moments). Taking into account such considerations as behavioral realism and simplicity of the model estimation (specifically, the formulation of SP alternatives) and application, three main forms have been proposed and tested to date (Batley et al. 2008):

- *Standard deviation* is a symmetric measure that assumes that being early or late is equally undesirable (probably not a realistic assumption for many trips and underlying activities);
- The difference between the 80th, 90th, or 95th and the 50th percentile (median) of travel times is frequently referred to as *buffer time.* This is an asymmetric and more behaviorally appealing measure because it specifically targets late arrivals and is less sensitive to early arrivals; and
- Simplified asymmetric measures in terms of *probability of certain delays* with delay thresholds such as 15 or 30 minutes are frequently used in the SP framework.

An illustrative example of the standard deviation approach is provided in Small et al. (1999) in the context of a binary route choice. The following form of utility function was adopted:

$$U = a \times T + b \times C + c \times SD(T) \tag{3.6}$$

where $SD(T)$ is the standard deviation of travel time.

Table 3.5. Value of Reliability Measured as Standard Deviation of Time

Trip Purpose and Income Group	Value of Reliability As SD(T)	
	($/min)	($/h)
Work trips, higher income	0.258	15.5
Work trips, lower income	0.215	12.9
Nonwork trips, higher income	0.210	12.6
Nonwork trips, lower income	0.167	10.0

Note: SD(T) = standard deviation of time.

Standard deviation of travel time was calculated based on the set of five travel times presented in the SP questionnaire for each highway route alternative. The estimation results showed that highway users assign a very high value to each minute of standard deviation, comparable with or even higher than the VOT associated with average travel time itself (i.e., $c \geq a$). In addition, a certain logical variation across trip purposes and income groups was captured as summarized in Table 3.5 (for one of several reported model specifications).

A good example of the second type of variability measure was presented by Small et al. (2005). The adopted quantitative measure of variability was the upper tail of the distribution of travel times, such as the difference between the 80th and 50th percentile travel times (see Figure 3.1). The authors argue that this measure is better than a symmetric standard deviation, because in most situations, being late is more crucial than being early, and many regular travelers will tend to build a safety margin into their departure times that will leave them an acceptably small chance of arriving late (i.e., planning for the 80th percentile travel time would mean arriving late for only 20% of the trips).

The choice context included binary route choice between the managed (tolled) lanes and general-purpose (free) lanes on a section of SR-91 in Orange County, California. The survey included actual users of the facility, and the model was estimated on the mix of RP and SP data. The variation of

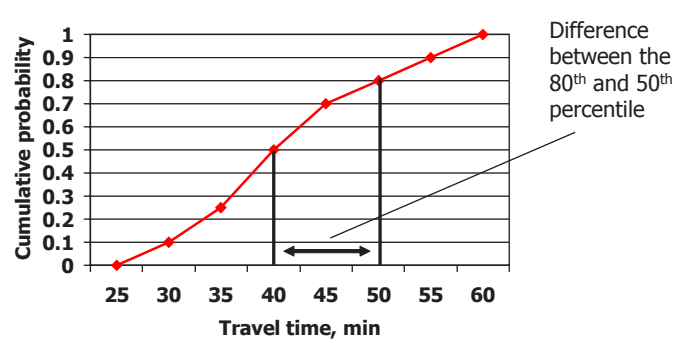

Figure 3.1. Travel time variability measure.

travel times and tolls was significantly enriched by combining RP data from actual choices with SP data from hypothetical situations that were aligned with the pricing experiment. Distribution of travel times was calculated based on the independently observed data. The measures were obtained from field measurements on SR-91 taken at many times of day on 11 days. It was assumed that this distribution was known to the travelers based on their past experience. The utility function was specified by the following formula:

$$U = a \times T + b \times C + c \times R(T) \tag{3.7}$$

where $R(T)$ is the difference between the 80th and 50th percentile.

Reliability, as defined above, proved to be valued by travelers as highly as the median travel time (VOT was roughly equal to VOR; i.e., $a \approx c$). This particular model form, with the condition of equal VOT and VOR, has a very interesting and intuitive interpretation; it could be used for a model formulation in a slightly simplified form if it were assumed from the outset that $a = c$. Indeed, if it is assumed that the willingness to pay for saving 1 minute of average travel time (the 50th percentile) is equal to the willingness to pay for 1 minute of reduction of the difference in time between the 80th and 50th percentiles, then both terms can be combined in the highway utility function because they have the same coefficient. This means that the underlying decision-making variable is the travel time value at the 80th percentile. This variable essentially combines both average travel time and time variation measure.

An example in Table 3.6 illustrates this possible approach. In the example, it is assumed that the highway user has to choose between two roads for commuting that are characterized by different time distributions. Road 1 is longer but more reliable; the travel time varies from 41 to 50 minutes. Road 2 is shorter, but travel time is less predictable and varies from 29 to 52 minutes. It is assumed that the highway user is familiar with both roads and makes his or her choice based on a rational consideration of the known distributions. In practical terms, this can be interpreted as a recollection of at least 10 trips on each road in the past, sorted by travel times from the best to worst.

Table 3.6. Illustration of Reliability Impact on Route Choice

Percentile	Travel time (min)		Preference
	Road 1	Road 2	
10	41	29	
20	42	30	
30	43	35	
40	44	39	
50	45	40	Road 2 by conventional approach
60	46	41	
70	47	45	
80	48	50	Road 1 by suggested approach
90	49	51	
100	50	52	

Although Road 2 has a better (lower) average travel time and would be preferred in most conventional modeling procedures, Road 1 has a better 80th percentile measure. In reality, the user would probably prefer Road 1 as the more reliable service. This choice framework with a single measure can be used as a simplified version of the approach. Rather than estimating two terms (average travel time and additional time associated with 80th–50th percentile), a single measure of the 80th percentile (or any other percentile larger than 50th if it yields a better statistical fit) could be used. For example, in a similar context, a 90th percentile measure was used by Brownstone and Small (2005). This framework is based on a plausible assumption that travelers under congestion conditions, characterized by travel time uncertainty, behave as rational risk minimizers. They do not base their decisions on the average values. However, they do not adopt the extreme mini–max approach (minimize risk and choose according to the worst possible case), either. The decision point probably lies somewhere between the 80th and 90th percentiles.

It is important to note that making this approach operational within the framework of regional travel models requires explicitly deriving these measures from simulation of travel time distributions, as well as adopting assumptions regarding the ways in which travelers acquire information about the uncertain situation they are about to experience. DTA and traffic microsimulation tools are crucial for the application of models that include explicit travel time variability, because static assignment can only predict average travel times.

Other approaches for measuring variability of travel time can also be considered. They are similar to the approach described above in conceptual terms, but they use a different technique in both the model estimation and the application stages. For example, in the travel model developed by PB Consult, Inc. (2003), the probability of delays longer than 15 and 30 minutes was introduced in the SP questionnaires for trucks. The subsequent estimation of the choice model revealed a very high significance of this variable that was comparable with the total trip time (in line with the VOR estimation of Small et al. [2005]). Application of this model required special probability-of-delay skims that were calculated based on the observed statistics of delays as a function of the modeled volume-to-capacity (V/C) ratio. Although this technique requires a multiday survey of travel times and speeds, it can be applied in combination with the static assignment method. Many regions with continuous traffic monitoring equipment

now have such data available for important highway segments. A problem yet to be resolved, however, is that when calculating the travel time reliability measure over the entire O-D path, the highway links cannot be considered independent.

Reliability is closely intertwined with VOT. In RP models, if variability is not measured explicitly and included as a variable, this omission will tend to inflate the estimated value of average time savings. In reality, variability in travel time tends to be correlated with the mean travel time, and people are paying for changes in both variables, so omitting one will tend to attribute the total effect to the other. Consequently, an important use of SP data sets that include reliability is to use them in combination with RP data sets for which good objective estimates of travel time variability can be derived.

It should be mentioned that the direct use of travel time variability in the framework of behavioral modeling is not the most appealing approach when compared with the other two approaches (discussed below). The principal conceptual drawback of this approach is that it does not explicitly consider the nature of underlying activities and mechanisms that create the disutility. Needless to say, the largest part of the disutility associated with unreliable travel time is being late (or too early) at the activity location, and consequently losing some part (or in some cases all) of one's participation in the planned activity. The clear practical advantage of the time variability approach, however, is in its relative simplicity and exclusive reliance on the data supplied by the transportation networks.

Schedule Delay Cost Approach

This approach has been widely accepted by the research community since its inception (Small 1982). According to this approach, the impact of travel time (un)reliability is measured by the explicit cost associated with the delayed or early arrival at the activity location. This approach considers a single trip at a time and assumes that the preferred arrival time that corresponds to zero schedule cost is known. The essence of the approach is that the trip cost (i.e., disutility) can be calculated as a combination of the following three components:

α = value of travel time and cost;
β = cost of arriving earlier than the preferred schedule; and
γ = cost of arriving later than the preferred schedule.

By definition, only one of the schedule costs can have a nonzero value in each particular case, depending on the actual arrival time versus the preferred one. There can be many analytical forms for the schedule cost as a function of the actual time difference (delay or early arrival). It is logical to assume that both functions should monotonically increase with respect to the time difference. It is also expected, in most cases, that the schedule delay function should be steeper than the early arrival function for most activities (being late is more onerous than being early). The details, however, depend on the activity type, person characteristics, and situational context.

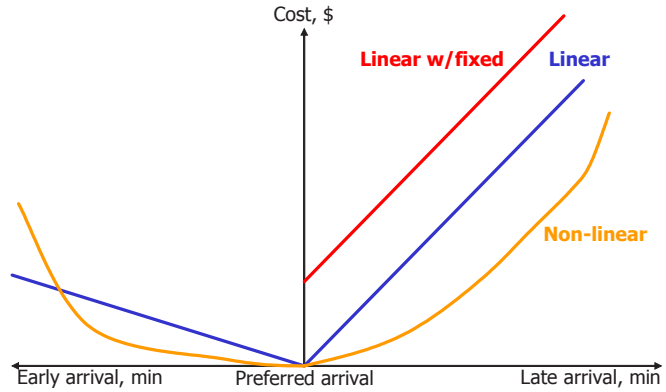

Figure 3.2. Schedule delay cost functions.

The most frequently used forms include simple linear function (i.e., constant schedule delay cost per minute), nonlinear convex function (assuming that large delays are associated with a growing cost per minute), and various piecewise functions accounting for fixed cost associated with any delay along with a variable cost per minute, as shown in Figure 3.2.

An example of a schedule delay model estimated in a highway route choice context with a specially designed SP survey is given in Small et al. (1999). The utility function was specified in the following way:

$$U = a \times T + b \times C + c \times SD(T) + \beta(\Delta t) + \gamma(\Delta t) \tag{3.8}$$

where

Δt = difference between actual and preferred arrival time;
$\beta(\Delta t)$ = early arrival cost specified as a nonlinear convex function; and
$\gamma(\Delta t)$ = late arrival cost specified as a linear function with a fixed penalty.

The estimation results with respect to the schedule delay cost are summarized in Table 3.7 for one of the tested model

Table 3.7. Estimation Example for Schedule Delay Cost

Component	Marginal Value ($)
Early Arrival (nonlinear)	
By 5 min	0.028/min
By 10 min	0.078/min
By 15 min	0.128/min
Late Arrival Dummy	
Work trips	2.87
Nonwork trips	1.80
Late Arrival (linear)	0.310/min
Extra Late Arrival Dummy	0.98

specifications. Interestingly, as reported by the authors, in the presence of explicit schedule delay cost, the travel time variability measure (standard deviation) lost its significance. The authors concluded that in models with a fully specified set of schedule costs, it is unnecessary to include the additional cost of unreliability of travel time (standard deviation).

Schedule delay cost should be distinguished from TOD choice and the associated disutility of shifting the planned (preferred) trip departure or arrival time, although in practical estimation analysis the data might mix these two factors. To clearly distinguish between the planned schedule and schedule delay, the person should explicitly report actual and preferred arrival times for each trip. Schedule delay cost assumes that the person has planned a certain schedule, but in the implementation process on the given day the delay occurs to disturb this plan. TOD choice relates to the stage of schedule planning. The outcome of this process is the preferred arrival time.

In comparing schedule delay to time variability as two different measures of time reliability, it should be noted that the schedule delay approach provides a better behavioral insight than travel time variability. It explicitly states the reasons and attempts to quantify the factors of the disutility associated with unreliable travel time, specifically perceived penalties associated with not being at the activity location on time. The schedule delay approach, however, has its own theoretical limitations as identified in the following:

- The approach is applied separately for each trip made by a person during the day, and it is assumed that the schedule delay cost for each subsequent trip is independent of the previous trip. Technically this approach is based on a fixed departure time and a preferred arrival time for each trip. In general, this is not a realistic assumption, since the activity duration requirements would create a dependence of the departure time for the next trip on the arrival time for the previous trip.
- This approach does not consider activity participation explicitly, though it makes a step toward such a consideration that the travel time variability approach ignores.
- If applied for the evaluation of user benefits from travel time savings, this approach must incorporate TOD choice (i.e., travelers' reconsideration of departure time in response to the changed congestion). Otherwise, travel time savings can result in early arrival penalties overweighting the value of saved travel time.

On the practical side, in order to be implementable, the schedule delay approach imposes several requirements that are not easy to meet, especially with conventional RP surveys:

- For each trip, in addition to the actual arrival time, the preferred arrival time should be identified. Although the preferred arrival time is generally known to the traveler (or perceived subconsciously), it is generally not observed by the modeler using RP-type data. To explore this phenomenon and estimate models that address it, the SP framework proved to be very effective, since the preferred arrival time and schedule delays can be stated in the design of alternatives. In some research, simplified assumptions about the preferred arrival time were adopted. For example, in Tseng and Verhoef (2008), the preferred arrival time was calculated as a weighted average between the actual departure time and would-be arrival time under free-flow traffic conditions.
- Application of this model for forecasting would again require input in the form of preferred arrival times. This could be accomplished either by means of external specification of the usual schedules on the activity supply side (which would probably be possible for work and fixed nonwork activities) or by means of a planned schedule model on the demand side. The latter would generate individual schedule plans (departure times) based on the optimal activity durations conditional on the average travel times. The subsequent simulation (plan implementation) model would incorporate schedule delay cost based on the simulated travel times.

Loss of Activity Participation Utility: Temporal Utility Profiles for Activity Participation

The third approach is based on a concept of time-dependent utility profile by activity type (Supernak 1992; Kitamura and Supernak 1997). Recently this approach was adopted in several research works on DTA formulation integrated with activity scheduling analysis (Kim et al. 2006; Lam and Yin 2001). The essence of this approach is that each individual has a certain temporal utility profile for each activity that is characterized by function $U(t)$. The utility profile can be estimated as a parametric or a nonparametric function of time, and time can be modeled in either continuous or discrete form. The utility profile represents an instant utility of participation in the activity at the given point of time (or during the discrete time unit that starts at the given point of time). The total utility of participation in the activity can be calculated by integrating the utility profile from the arrival time (τ) to departure time (π):

$$U(\tau, \pi) = \int_{\tau}^{\pi} u(t)\,dt \tag{3.9}$$

Simple utility profiles are independent of the activity duration. In this case, it is assumed that the marginal utility of each activity at each point of time is independent of the time already spent on this activity. This might be too simplifying an assumption, at least for certain activity types like household maintenance needs, in which the activity loses its value after the errands have been completed. More complicated utility

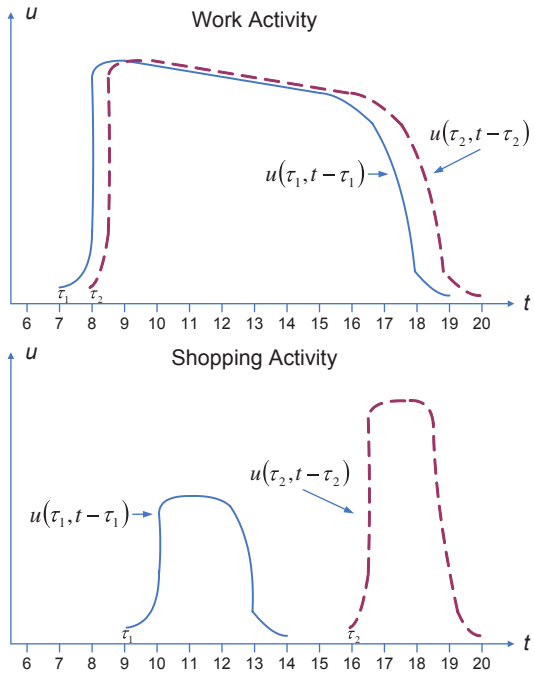

Figure 3.3. Examples of temporal profiles of activity participation utility.

profiles can be specified as two-dimensional functions $U(t,d)$, where d denotes the activity duration until moment t. In this case, the total utility of activity participation can be written as

$$U(\tau, \pi) = \int_{\tau}^{\pi} u(\tau, t-\tau)dt \tag{3.10}$$

Hypothetical but typical temporal utility profiles specified in a discrete space with an hourly resolution are shown in Figure 3.3. The work activity profile is adjusted to reflect the fixed schedule requirements (higher utility to be present at 8:00 a.m. and 5:00 p.m.). The shopping activity profile is much more uniform, with an additionally assumed convenience to undertake this activity after usual work hours. In both cases the utility is measured versus staying at home (i.e., not participating in any out-of-home activity that would require travel) as the reference (zero) utility. Thus, the utility profile can take both positive and negative values.

The concept of utility profiles is instrumental in understanding how individuals construct their daily activity schedules. According to this concept, each individual maximizes a total daily utility of activity participation. If a predetermined sequence of activity episodes is considered, it can be said that individuals switch from activity to activity when the time profile of the second activity exceeds the time profile of the previous activity. Travel episodes are placed between activity episodes in such a way that the whole individual daily schedule represents a continuous sequence of time intervals, as shown in Figure 3.4.

The effect of unreliability of travel times can be directly measured by comparing the planned and actual total daily utility of the schedule, which includes all activity and travel episodes. For simplicity, but without essential loss of generality, it is assumed that the sequence of activity episodes and trip departure times are fixed. It is also assumed that travel time delay never exceeds the planned duration of the subsequent activity; thus, activities cannot be cancelled as a result of unreliable travel time. Thus, unreliability affects only travel times and arrival times. In this context, the reliability measure can be expressed as the loss of activity participation in the following way:

$$L = \sum_i (U_i^P - U_i^A) \tag{3.11}$$

where

L = total user loss (disutility) over the whole schedule;
U_i^P = utility of the trip and subsequent activity with preferred arrival time; and
U_i^A = utility of the trip and subsequent activity with actual arrival time.

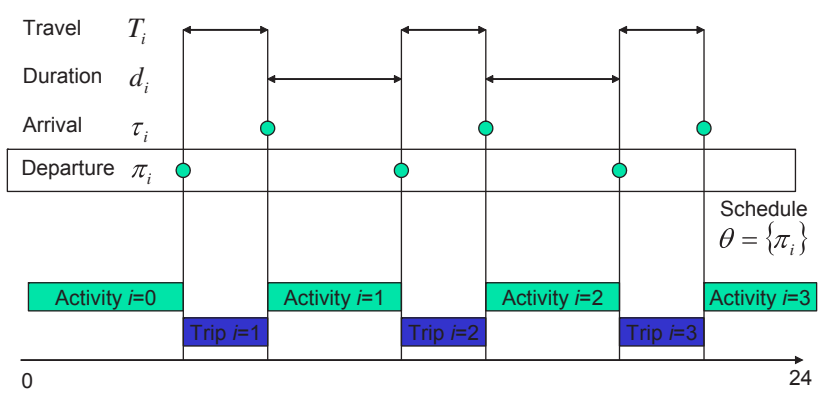

Figure 3.4. Consistent individual daily schedule.

The planned and actual utilities can be written as shown in Equations 3.12 and 3.13, respectively:

$$U_i^P(\tau_i^P) = a \times T_i^P + b \times C_i^P + \int_{\tau_i^P}^{\pi_{i+1}} U_i(t)\,dt \tag{3.12}$$

$$U_i^A(\tau_i^A) = a \times T_i^A + b \times C_i^A + \int_{\tau_i^A}^{\pi_{i+1}} U_i(t)\,dt \tag{3.13}$$

where

$$T_i^P = \tau_i^P - \pi_i;\ T_i^A = \tau_i^A - \pi_i \tag{3.14}$$

By substituting Equation 3.14 into Equations 3.12 and 3.13, and then substituting Equations 3.12 and 3.13 into the basic expression (Equation 3.11), Equation 3.15 is obtained:

$$L = \sum_i \left[a \times \left(\tau_i^P - \tau_i^A\right) + b \times \left(C_i^P - C_i^A\right) + \int_{\tau_i^P}^{\tau_i^A} U_i(t)\,dt \right] \tag{3.15}$$

where the last term (integral) represents the loss of activity participation, and the first two terms represent extra travel time and cost.

A logical relationship between temporal activity profiles of utilities and schedule delay cost was explored by Tseng and Verhoef (2008) that led to an insightful general framework. It can be shown that these two approaches are not independent. The schedule delay cost functions can always be consistently derived from the temporal utility profiles; thus, the schedule delay approach can be thought of as a particular transformation of the temporal utility profile approach. Interestingly, the opposite is true; that is, temporal utility profiles could be fully restored from the schedule delay cost functions only under some specific assumptions.

Accounting for Unobserved Heterogeneity and Situational Variability

Increasingly, travel demand analysts are looking beyond average user responses to travel costs, travel times, and other attributes toward accounting for heterogeneity or differences in user response across the population. Capturing heterogeneity in user valuation of attributes, such as travel costs and travel times, is important in order to correctly predict overall (market) responses to measures such as pricing, as well as to provide policy makers with information about the impacts of policies on different segments of the population. For example, the money VOT for users may vary considerably across a population, and policies based on the assumption of a mean money VOT may not produce the anticipated impacts (Sillano and Dios de Ortúzar 2005). As shown in the previous sections, a well-specified demand model will attempt to include as many as possible of the observable factors that can be shown to affect travel time valuation in a systematic way. However, these factors may not always be known to the analyst, and in many cases various other sources can account for the varying valuations across the population; this is referred to as *unobserved heterogeneity.*

Major advances in choice model formulation and estimation over the past decade have produced relatively robust methods to incorporate unobserved heterogeneity, particularly in the form of random coefficients (i.e., model parameters that are assumed to follow a distribution across the user population). This section provides a general framework for accounting for unobserved heterogeneity in travel demand models, specifically discrete choice models. The following subsections discuss heterogeneity, both unobserved and observed, and how to account for them within a discrete choice modeling framework. Model specification and estimation issues are briefly discussed, followed by an example to illustrate the range of questions that can be addressed with a model that accounts for unobserved heterogeneity.

Accounting for Observed and Unobserved Heterogeneity

The response of users toward attributes, such as travel time savings and cost, of different alternatives varies in general over the population of users. For example, low-income individuals are probably more concerned about and sensitive to toll prices than high-income individuals. From a practical standpoint, a common method for capturing such heterogeneity, controlling for other factors such as trip purpose, is to segment the sample of users based on exogenous criteria, such as income level, trip length, and TOD (peak versus nonpeak). Separate models are then estimated for each segment. Another practical approach is to interact attributes of the alternatives with exogenous criteria. Consider a user's choice between taking a toll or a nontoll route to work. Assume the only two attributes observed are travel cost TC_j for route type j and travel time TT_j. The importance that users place on these two attributes, reflected in the coefficients α_n and, may vary over the population, with the utility for each alternative written as

$$U_{nj} = \alpha_n TT_j + \beta_n TC_j + \varepsilon_{nj} \tag{3.16}$$

where j is either toll or free, and α_n and β_n are parameters specific to individual n. One common method for accounting for heterogeneity in response is interacting the travel time or cost terms with exogenous criteria such as income. Assuming that the importance of travel cost is inversely related to the observed income of the users (I_n), with low-income individuals placing more importance on travel costs, the coefficient for travel cost can be expressed as

$$\beta_n = \theta / I_n \tag{3.17}$$

where θ can be regarded as the mean value or importance placed on cost across all users.

An alternative approach is to allow further differences by expressing the parameters that represent preference weights (α_n and β_n) as random parameters, as opposed to point estimates, such that the distribution for these preference weights can obtained and used in the derivation of value of travel time, which in turn will be distributed across the population. These distributions can also be a function of exogenous variables. Randomness in preference weights results from a variety of reasons, possibly just because people are inherently different. Assume that users' response to travel costs, reflected by the parameter α_n, varies across all users, but is linked to unobserved factors or is intrinsically random. Examples of these unobserved or latent factors may include differences in familiarity with the network or differences in general stress levels. The resulting α_n can be expressed as

$$\alpha_n = \rho + \mu_n \tag{3.18}$$

where ρ reflects the mean response of users to travel time, and μ_n is a randomly distributed term that captures deviations from this mean value. Substituting Equation 3.17 and Equation 3.18 back into Equation 3.16 provides a utility expression that reflects both unobserved and observed heterogeneity in user responses to travel costs and travel times, as shown below:

$$\begin{aligned} U_{nj} &= \underbrace{(\rho + \mu_n)_j + (\theta / I_n) TC_j + \varepsilon\varepsilon_{nj}}_{\text{Observed or Systematic}} \\ &= \underbrace{\rho TT_j + (\theta / I_n) TC_j + \mu_n TT_j + \varepsilon_{nj}}_{\text{Unobserved or Random}} \end{aligned} \tag{3.19}$$

The above utility expression accounts for both observed and unobserved heterogeneity. The response of users toward travel costs varies systematically according to observed income, as expressed in Equation 3.17, but users' response to travel times varies randomly according to unobserved factors, as expressed in Equation 3.18. The analyst needs to specify a distribution for the random coefficient α_n. For example, the analyst may assume this distribution is normal, with mean and variance to be estimated in order to make inferences and gain insight on the distribution of users' response to travel times, including related measures such as money VOT.

Given an estimated distribution of VOT, the proportion of a population P that decides to pay a toll C_{toll} is given by the proportion with VOTs saved greater than C_{toll}:

$$P_{C_{toll}} = \int_{C_{toll}}^{\infty} (VOT) \tag{3.20}$$

The analyst selects the distribution $f(.)$ of VOT in order to find a satisfactory representation of the "true" empirical distribution. This is illustrated below in Figure 3.5, in which the proportion of payers is the blue area, given that the toll is set to $20.

The shaded area to the right is the measure of the number of people who have VOT savings exceeding the toll charged and would therefore pay it. In the case of the substantially skewed lognormal distribution, the mean is not the center of the distribution, and for the case shown in Figure 3.5, there will be fewer people in the population actually ready to pay for the toll.

The next subsection discusses the estimation in relation to discrete choice models that can capture unobserved heterogeneity and the forms these models can take.

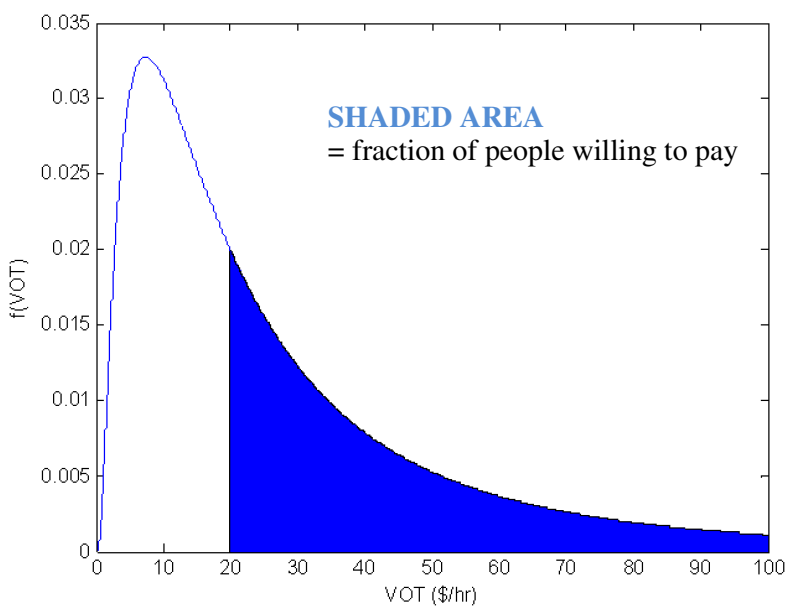

Figure 3.5. Proportion of payers with lognormal distribution for VOT for a toll of $20.

Discrete Choice Model Form and Estimation Issues

The model form these models take is dictated partly by assumptions on the error terms, which in turn are dictated by the need to account for unobserved heterogeneity. In Equation 3.19, since μ_n is not observed, the term $\mu_n TT_j$ becomes part of the unobserved component of the utility $\widetilde{\varepsilon_{nj}}$, so it can be expressed as

$$U_{nj} = \rho TT_j + (\theta / I_n) C_j + \widetilde{\varepsilon_{nj}} \tag{3.21}$$

where

$$\widetilde{\varepsilon_{nj}} = \mu_n TT_j + \varepsilon_{nj} \tag{3.22}$$

The example above illustrates the concept of heterogeneity in terms of the value individuals place on the attributes of alternatives. Heterogeneity can be captured within a discrete choice framework by linking this variation to observed or unobserved characteristics. As an example of observed heterogeneity, individual response to cost β_n was linked to an individual's income level, such that low-income individuals were more sensitive relative to high-income individuals. If individual response is linked to unobserved variables or is purely random, then the analyst would need to account for unobserved heterogeneity. In the example above, variation in response to travel time α_n was assumed to be random, as expressed in Equation 3.19, leading to total error for the utility expressed in Equation 3.22.

The type of heterogeneity present and accounted for dictates the type of choice model that is appropriate. If only observed heterogeneity is captured and accounted for, and the error term ε_{nj} is still distributed independently and identically Gumbel, a logit formulation can be used. If unobserved heterogeneity is accounted for, a logit formulation cannot be used because the total error term $\widetilde{\varepsilon_{nj}}$ is no longer distributed independently and identically. If the heterogeneity in tastes is linked to unobserved variables and is random, a logit model form would be a misspecification. As an approximation, the logit model may capture average tastes fairly well, but it cannot provide information on the distribution or heterogeneity of tastes around the average. This distribution is very important in many situations, such as forecasting the market share for tolled routes that appeal to a minority of people rather than to those of average tastes. To incorporate random taste variation appropriately and fully, probit or mixed logit model forms, or both, may be used instead.

The mixed logit and probit models are particularly well suited for incorporating unobserved heterogeneity. Continuing the previous example, assuming the coefficient for travel time varies randomly over individuals, the utility is expressed in Equation 3.18, where α_n is assumed to be distributed with a density $f(\alpha)$ with parameters θ, which can consist of a mean b and a covariance W.

The goal of estimation is to determine values for b and W. Several distributions can be assumed, both continuous and discrete. The analyst observes the travel times TT_j, but not the individual specific parameters α_n or the errors $\widetilde{\varepsilon_{nj}}$. If α_n were known, then the choice probability for an alternative, conditional on knowing α_n, would be

$$Pr_{ni}(\alpha_n) = \frac{\exp\left(\alpha_n TT_i + \left(\frac{\theta}{I_n}\right) TC_i\right)}{\sum_j \exp\left(\alpha_n TT_j + \left(\frac{\theta}{I_n}\right) TC_j\right)} \tag{3.23}$$

However, since the analyst often does not know α_n, he cannot condition on α. The unconditional choice probability is therefore the integral of $Pr_{ni}(\alpha_n)$ over all possible values of α_n:

$$Pr_{ni} = \int \left(\frac{\exp\left(\alpha_n TT_i + \left(\frac{\theta}{I_n}\right) TC_i\right)}{\sum_j \exp\left(\alpha_n TT_j + \left(\frac{\theta}{I_n}\right) TC_j\right)} \right) \cdot f(\alpha) \cdot d\alpha \tag{3.24}$$

The parameters in the choice probability in Equation 3.24 are estimated using simulated maximum likelihood estimation. This is accomplished by taking several draws from the distribution of α and averaging the choice probability Pr_{ni} across all these draws. The probabilities expressed in Equation 3.23 are approximated through simulation for given parameter values. This average simulated probability is expressed as

$$\check{Pr}_{ni} = \frac{1}{R} \sum_{r=1}^{R} Pr_{ni}(\alpha_n^r) \tag{3.25}$$

$$Pr_{ni}(\alpha_n^r) = \frac{\exp\left(\alpha_n^r TT_i + \left(\frac{\theta}{I_n}\right) TC_i\right)}{\sum_j \exp\left(\alpha_n^r TT_j + \left(\frac{\theta}{I_n}\right) TC_j\right)} \tag{3.26}$$

where R is the number of draws. The simulated probabilities (Equations 3.11 and 3.12) are inserted into the log likelihood function to give a simulated log likelihood function (SLL):

$$SLL = \sum_{n=1}^{N} \sum_{j=1}^{J} d_{nj} \ln(Pr_{ni})$$

where $d_{nj} = 1$ if person n chose j, and $d_{nj} = 0$ otherwise. The maximum simulated likelihood estimator is the value of the parameters that maximizes SLL.

Distributions for Travel Time Coefficient

Several distributions may be assumed for the travel time coefficient β^{time}, although commonly for VOT studies, this is assumed to be a truncated normal or truncated lognormal

distribution. The normal distribution has been shown to cause some problems when applied to coefficients of undesirable attributes, such as travel time and cost, due to the possibility of positive coefficient values for these attributes (Hensher and Greene 2000; Cirillo and Axhausen 2006). To circumvent this, the normal is usually truncated to ensure that coefficients are negative for undesirable attributes and positive for desirable attributes. The lognormal distribution has the useful property of being bounded below by zero. It is useful for coefficients of attributes that are liked (or disliked) by all users. The sign is reversed for undesirable attributes, such as a travel time variable, such that the coefficient is necessarily negative. In studies of willingness to pay, the lognormal distribution has been shown to produce large and unreasonable variances and means (Hensher and Greene 2000; Hess et al. 2005). Evidence of this can be seen in the estimation results below for an "unbounded" lognormal distribution. To circumvent this, the lognormal distribution may need to be truncated to ensure reasonable means and variances.

The researcher specifies a distribution for the coefficients and estimates the parameters of that distribution. In most applications, $f(\alpha)$ is specified to be normal or lognormal:

$\alpha \sim N(b, W)$ or

$[\ln(\alpha)] \sim N(b, W)$ with parameters b and W that are estimated.

The lognormal distribution is useful when the coefficient is known to have the same sign for every decision maker, such as a travel time coefficient that is known to be negative for everyone. Triangular and uniform distributions have also been used (Hensher and Greene 2003). With the uniform density, β is distributed uniformly between $b - s$ and $b + s$, where the mean b and spread s are estimated. The triangular distribution has a positive density that starts at $b - s$, rises linearly to b, and then drops linearly to $b + s$, taking the form of a triangle. The mean b and spread s are estimated, as with the uniform distribution, but the density is peaked instead of flat. These densities have the advantage of being bounded on both sides, thereby avoiding the problem that can arise with normals and lognormals having unreasonably large coefficients for some share of decision makers.

One way around the unbounded nature of the normal and lognormal distributions is to truncate these distributions, specifying either a lower or upper bound, or both. In studies of willingness to pay, the lognormal distribution has been shown to produce large and unreasonable variances and means (Hensher and Greene 2000; Hess et al. 2005). Evidence of this can be seen in the estimated models for an "unbounded" lognormal distribution presented in Chapter 4. To circumvent this, the lognormal distribution may need to be truncated to ensure reasonable means and variances. Another alternative is to use an *Sb*-Johnson distribution, which requires specifying an upper and lower bound. The *Sb* distribution is useful for a variety of purposes. *Sb*-Johnson densities can be shaped like lognormals, but with an upper bound and with thinner tails below the bound. *Sb* densities are more flexible than lognormals: they can be shaped like a plateau with a fairly flat area between drop-offs on each side and can even be bimodal (Train and Sonnier 2004). When a lower bound other than zero is specified, the distribution is useful for an attribute that some people like and others dislike but for which there is a limit for how much the person values having or avoiding the attribute. In general, the analyst should specify a distribution that results in plausible behavior and provides good fit to the data.

Route-Type Choice: Revealed Preference Framework (New York Model)

Overview of Section, Approach, and Main Findings

Auto route choice in the highway network represents the simplest and most basic platform for understanding and modeling the behavior of highway users and their underlying generalized cost functions. Route choice is essentially a trip-level decision with no significant tour-level effects or constraints. In this choice context, it is assumed that trip origin, destination, departure time, and auto occupancy are fixed and are taken from the corresponding decisions that were modeled earlier in the model system hierarchy. Thus, the choice set consists of different highway network routes that may differ by time, cost, distance, reliability, or other measures, while the effects of person and household variables are included via interactions with route variables or as segmentation variables. This specification allows the analyst to focus on the basic form of the highway utility (generalized cost) function that is incorporated as part of the (more complicated) mode and TOD utility expressions.

Despite the attractiveness of the route choice framework as a platform for analyzing the highway utility function, there is limited supporting evidence in the literature on estimated route choice models. This is primarily due to the lack of available data sets with actual auto route itineraries reported or recorded. The common practice in most travel surveys is to collect only trip origins and destinations. Even if the actual route can be restored from some indicators on major facilities used, the identification of reasonable alternative routes to form a choice set is not a trivial issue. These problems were resolved to some extent with the two RP data sets available and extended for the current research.

The first route choice data set was based on the Household Travel Survey in the 28-county New York region, collected in

1997 and 1998. In the survey, each auto trip has an attribute of toll value paid. In the New York region most tolls are clearly associated with major facilities like bridges and tunnels around Manhattan or the New Jersey Turnpike, and specific facilities are clearly the best tolled options for certain subsets of origin and destination zone pairs. A significant number of auto trip records in the survey have origins and destinations for which both a tolled route and a free route are feasible and reasonably competitive in terms of generalized cost. Thus, it proved to be possible to from a binary route-type choice model (tolled versus free) and to support it by the corresponding set of skims for the reported TOD period, including time, cost, distance, and reliability measures (generated by the method described in Chapter 2), along with the possibility of segmentation by congestion levels and facility type. The synthetic congested travel time estimates used for this model are, of course, subject to the limitations of static traffic assignment procedures, although the simulations were implemented for each hour of the day separately.

The second data set was created from the Seattle Traffic Choices Study from 2006 (see Chapter 2), for which global positioning system (GPS) time and location data streams from travel on actual routes were available. The chosen route types were identified, and the best alternative routes were constructed to support the same binary route-type choice model (a tolled freeway route versus a nonfreeway route with a lower toll cost). Travel time distributions were calculated based on the actual average time variability across the GPS traces for each network link pair (further aggregated to O-D pair) over all weekdays in the 12-month survey period. The route-type choice framework with the Seattle data was extended to incorporate TOD (departure time dimension).

As is the case with practically all RP data sets, only the first two types of reliability measures described earlier (perceived highway time and travel time distribution) were available. Analysis of the other two reliability measures (schedule delays and temporal utility profiles) could not be supported by the available data in the survey, and no reasonable way of generating these measures synthetically was found within the research project framework.

The estimation results for these models are analyzed in the following subsections and compared with other relevant results reported in the literature. The analysis begins with the most basic linear specification. In each subsequent subsection, one specific aspect is analyzed individually in relation to the base specification. In the penultimate subsection, the best features are combined in one recommended specification of the highway generalized cost function that is the main constructive outcome of the current stage of research. This form of the generalized cost function linearly combines mean travel time, cost, and travel time reliability with a consideration of the nonlinear effects of distance, income, and car occupancy on these three main terms. This form is used as a seed construction that is further analyzed in this chapter as part of extended choice models that include mode and TOD dimensions.

In the final subsection, this specification is additionally analyzed with respect to unobserved heterogeneity when some of the coefficients were estimated as random rather than as deterministic values.

Time-of-Day Choice and Joint Time-of-Day and Route-Type Choice: Revealed Preference Framework (Seattle)

Overview of Section, Approach, and Main Findings

This section explores another primary dimension—TOD choice—in the most basic trip framework. Two data sets with different model specifications were used. The first is based on the Household Travel Survey in Seattle in 2000. This data set was used to estimate a trip TOD (departure time) model.

The second data set was created from the Seattle Traffic Choices Study from 2006, for which GPS time and location data streams from travel on actual routes were available. The chosen route types were identified, and the best alternative routes were constructed to support the same binary route-type choice model (a tolled freeway route versus a nonfreeway route with a lower toll cost), as was estimated for New York and discussed earlier in Chapter 3. Travel time distributions were calculated based on the actual average time variability across the GPS traces for each network link pair (further aggregated to O-D pair) over all weekdays in the 12-month survey period. The route-type choice framework with the Seattle data was extended to incorporate TOD (departure time dimension).

The combined route-type and TOD choice model estimated for Seattle is not equivalent to the pure route-type choice model estimated for New York. However, some comparisons across the coefficients that describe the route dimension are possible. As is the case with practically all RP data sets, only the first two types of reliability measures described earlier (perceived highway time and travel time distribution) were available. Again, the analysis begins with the most basic linear specification, and in each subsequent subsection one particular aspect of the base specification is analyzed.

The following main findings regarding TOD choice are summarized as follows:

- The coefficients for the main model variables of average time and cost proved to be in a reasonable range relative to previous studies. Extra delay variables (for time longer than 1.2 of free-flow time) proved to have an additional impact on TOD as a result of avoiding driving in congestion conditions.

- A direct measure of travel time reliability like standard deviation of travel time or standard deviation of travel time per unit distance proved to be statistically significant and performed better than more elaborate measures such as buffer time (the difference between the 90th and 50th percentiles).
- TOD choice is subject to many person and household variables. In particular, variables such as full-time versus part-time work status and income proved to have a significant effect on work schedules. Part-time workers and low-income workers have shorter activity durations compared with full-time, high-income workers. The longer work activity (tour) duration corresponds to earlier departures from home and later arrivals back home for higher incomes. Interestingly, after controlling for worker status and income, age and gender proved to have only minor impacts on work schedules. More than 80% of part-time workers are female, which can explain why the gender variable might be significant if work status is not included as a variable.
- Carpools for nonwork purposes tend to have a later schedule than drive-alone nonwork trips. The majority (about 75%) of the nonwork carpools correspond to joint travel by household members for which the schedule consolidation (especially if workers are involved) required this trip to be pushed to a later (after work) hour.

These effects are explored later in Chapter 3 in a more general framework of joint mode and TOD choice with cross comparison between the New York and Seattle regions.

The following main findings regarding route-type choice can be summarized with a special emphasis on generic impacts that proved to be common for both New York and Seattle:

- When compared with the basic specification of the New York route-type choice model, in general, the travel time coefficients across travel purpose and regions proved to be in a reasonable range (from –0.02 to –0.07), with a tendency for work purpose to have a greater coefficient than nonwork purpose. However, the VOTs obtained for New York ($19–$30/hour) are significantly higher than VOTs for Seattle ($7–$12/hour).
- In the previously discussed results for New York, travel time segmentation between arterial and local roads versus highways and freeways resulted in a statistically significant difference in coefficients. Arterial and local roads were characterized by a significantly higher (negative) coefficient than for highways and freeways. The Seattle model formulation adds an additional important facet to this analysis. The advantage of driving highways and freeways manifest itself only if a substantial portion of the overall trip can be driven on highway and freeways. If the freeway component is very small it loses its advantage, since the access to and egress from the freeway become as onerous as driving through intersections and stopping at traffic lights.
- A direct measure of travel time reliability such as standard deviation of travel time or standard deviation of travel time per unit distance proved to be statistically significant and performed better than more elaborate measures like buffer time (the difference between the 90th and 50th percentiles). However, the coefficients for standard deviation and standard deviation per unit distance obtained with the New York data were significantly larger than those obtained with the Seattle data. The corresponding reliability ratio for New York exceeded 1.0 in many cases, but it stands significantly below 1.0 for Seattle.
- These results contribute to the general observation from the multitude of previous studies that simple models are in general not easily transferable. Depending on the regional conditions, model specification, and the manner in which reliability measures were generated, the reliability ratio can range between 0.5 and 2.0. For this reason, in the final synthesis and recommendations the team does not follow either the New York model or Seattle model directly but rather considers them as somewhat extreme examples.

Basic Specification, Segmentation, and Associated Value of Time

When compared with the basic specification of the New York route-type choice model, in general, the travel time coefficients across travel purpose and regions proved to be in a reasonable range (from –0.02 to –0.07), with the tendency for work purpose to have a greater coefficient than nonwork purpose. For New York, this also resulted in an expected higher VOT for work trips compared with nonwork trips, which is also the most common result with many other models. However, for Seattle a different result was obtained, with the work VOT being lower than nonwork VOT. Also, in general, the VOT obtained for New York ($19–$30/hour) is significantly higher than VOT for Seattle ($7–$12/hour). The overall difference between the regions can be easily explained by the difference in average income (and income segmentation is not applied yet). The reversed ratio between the work and nonwork VOT in Seattle is difficult to substantiate and it may be a consequence of a relatively small subset of nonwork trips with tolled routes in the Seattle Traffic Choices Study.

Impact of Congestion Levels and Facility Type

The New York route-type choice model used a different specification from the Seattle model for the facility-type analysis. Thus, a direct comparison of the facility-type impacts between the two models is difficult. However, the analysis in both regions supported somewhat complementary results. In the previously discussed results for New York, travel time

segmentation between arterial and local roads versus highways and freeways resulted in a statistically significant difference in coefficients (at least for the nonwork travel purpose). Arterial and local roads were characterized by a significantly higher (negative) coefficient compared with highways and freeways. This implies a general user preference for highways and freeways compared with arterials and local roads, which is in line with the common consideration that intersections and traffic lights, in addition to travel time itself, are perceived negatively by drivers. It might also be tempting to interpret the higher (negative) travel time coefficient for arterials and local roads as a proxy for travel time reliability (which is the case with travel time segmentation by congestion levels). However, this is questionable because the freeway congestion levels are as significant as for arterials and local roads.

The Seattle model formulation adds an additional facet to this analysis. The advantage of driving highways and freeways manifests itself only if a substantial portion of the trip can be driven on highway and freeways. If the freeway component is very small it loses its advantage since the access to and egress from the freeway become as onerous as driving through intersections and stopping at traffic lights. This finding is behaviorally appealing. In general, the team believes that further research should be encouraged with respect to segmentation by facility type and the construction of a route utility function that includes variables like facility type, intersection type, and presence of traffic lights in addition to travel time and cost.

With the New York route-type choice model discussed above, a significant differentiation of time by congestion levels was found. It was technically implemented by dividing the total auto time into free-flow time and congestion delay. It is only slightly different from the Seattle RP formulation, in which the time breakdown point was 1.2 rather than 1.0 of the free-flow time. In the Seattle RP formulation this segmentation did not work directly in the trip departure-time choice context, but the delay variable proved to be statistically significant as a shift variable. This means that highway users not only tend to avoid routes with higher congestion levels, but also tend to adjust their schedule to avoid driving in the congestion periods. However, these effects may only be proxies for direct impacts of reliability measures.

Incorporation of Travel Time Reliability Measures and Value of Reliability Estimation

Overall, the Seattle results with the Traffic Choices Study confirm the main findings described for the New York model in that standard deviation of time and standard deviation of time per unit distance performed better than other (more elaborate) measures of travel time reliability, such as a buffer time (difference between the 90th percentile and median). Standard deviation of travel time per unit distance has a significant practical advantage over a simple unscaled standard deviation because the latter is frequently correlated with the mean travel time. This is not a conceptual advantage per se, but it is a significant practical constraint that is difficult to resolve in the RP setting. (This constraint can be overcome in the SP setting, however, by controlling the input LOS data.)

The coefficients for standard deviation and standard deviation per unit distance obtained with the New York data were significantly greater in magnitude than those obtained with the Seattle data. The corresponding reliability ratio for New York exceeded 1.0 in many cases, but the reliability ratio stands significantly below 1.0 for Seattle. These results contribute to the general observation from the multitude of previous studies that simple models are in general not easily transferable. Depending on the regional conditions, model specification, and the way the reliability measures were generated, the reliability ratio can be between 0.5 and 2.0, or even exceed these limits for some particular cases (Li et al. 2010; Concas and Kolpakov 2009). For this reason, in the final synthesis and recommendations the team does not follow either the New York model or Seattle model directly, but rather considers them as somewhat extreme examples. New York is characterized by extremely high congestion levels and notoriously unpredictable travel times. Coupling this with a relatively short average travel distance for auto trips (the majority of long-distance commuters in New York use transit), a reliability ratio greater than 1.0 is behaviorally justified. Seattle has generally lower congestion levels across the region; hence, the entire unreliability scale is set differently versus the average travel time.

Impact of Gender, Age, and Other Person Characteristics

Due to the data limitations of the Seattle Traffic Choices Study, it was impossible to directly compare the results with New York in the route-type choice context. However, it should be noted that even with the New York data, for which a rich set of person and household variables was available, only some gender effects, in the form of additional toll-averse bias, proved to be statistically significant. Gender, age, worker status, and other person characteristics manifested strongly in TOD choice. The cross comparisons between New York and Seattle with respect to TOD choice are discussed with a full specification of joint mode and TOD choice model.

Impact of Income

Again, the limitations of the Seattle Traffic Choices data set prevented direct comparisons with the New York analysis. With the New York model, as discussed above, the team substantiated a general functional form of highway generalized cost for which the cost variable was scaled down by income

powered by 0.6. This formulation will be further tested in the extended choice frameworks of mode and TOD choice. Income also has a strong direct impact on TOD choice. The cross comparisons between New York and Seattle with respect to TOD choice are discussed below, with a full specification of joint mode and TOD choice model.

Impact of Car Occupancy

Again, the limitations of the Seattle Traffic Choices data set prevented direct comparisons with the New York analysis. With the New York model, as previously discussed, the research team substantiated a general functional form of highway generalized cost for which the cost variable was scaled down by car occupancy powered by 0.6. This formulation was further tested in extended choice frameworks of mode and TOD choice. Car occupancy (and joint household travel) also has a direct impact on TOD choice. The cross comparisons between New York and Seattle with respect to TOD choice are discussed below, with a full specification of joint mode and TOD choice model.

Nonlinear Level of Service and Trip-Length Effects

The previously discussed analysis with the New York data substantiated a seed functional form for an interactive term between auto time and distance for work trips. This form results in a parabolic function for VOT in which the maximum VOT is associated with a commuting distance of about 30 miles; for shorter and longer trips, VOT is reduced. An attempt to replicate this effect with the Seattle data resulted in somewhat inconclusive functional forms, with the key coefficients being statistically insignificant. Part of the problem was that the Seattle data, unlike the New York data, did not provide a sufficiently large set of long travel (commuting) distances. Although the average commuting distance in the New York metropolitan region is relatively short (7.5 miles), the household survey of 11,000 households provided a significant number of observations of commuting distances beyond 30 miles. Thus, this particular model component needed further exploration and cross-regional comparisons in the mode and TOD choice frameworks, as discussed below.

Mode and Car Occupancy Choice: Revealed Preference Framework

Overview of Section, Approach, and Main Findings

The models of mode and car occupancy choice represent the next tier of statistical analysis in which the highway travel utility (generalized cost) is considered in the multimodal context. All aspects described above for route choice are also relevant for mode choice, as well, because the highway modes and route types represent alternatives in mode choice. However, because the choice framework is substantially extended to include transit modes, there are many more potential impacts, factors, and variables that come into play. Also, the mode choice framework naturally includes a much wider set of travelers, including transit users who may have very different perceptions of travel time, cost, and reliability. Additionally, the mode choice models estimated for New York used in this section are tour based, which means that two-directional LOS variables are considered (for the corresponding outbound and inbound TOD periods). A tour framework is essential for analyzing mode preferences, since many mode constraints and relative advantages of different modes cannot be seen at the level of a single trip.

A central research question at this stage is whether the main findings regarding the functional form of highway travel utility from the route choice analysis described above would hold in the more general framework of tour mode choice. In the subsections that follow, the team applies the same approach as for the previously discussed route choice. Each major factor and its associated impacts are analyzed one at a time and are progressively incorporated into the final model structure. Each factor is statistically tested with the New York data and Seattle data, while trying to keep the model structures as close and compatible as possible. Each subsection concludes with a synthesis of main findings that proved to be common for both regions.

The following main findings regarding mode choice can be summarized with a special emphasis on the generic impacts that proved to be common for both New York and Seattle and were also similar to the route choice and mode choice frameworks:

- Both mode choice models have a rich set of explanatory variables, including LOS variables, as well as various person and household variables. The overall scale of time and cost coefficients (specifically for auto time that is in the focus of the current study) is reasonable. It should be taken into account that the LOS variables in a tour model should be approximately doubled when compared with a trip mode choice model. Thus, the corresponding coefficients for time and cost need to be halved for a trip mode choice model when compared directly with a tour mode choice model. This is the case for auto in-vehicle time; for example, for work-related travel, it is –0.014 for the New York tour mode model and –0.029 for the Seattle trip mode model. For work tours in New York and work trips in Seattle, the base model specifications showed a relatively low VOT for auto users of \$6–\$7/hour. This value is not recommended for use in other models. However, the team decided not to enforce a more

reasonable VOT at this stage, but rather to continue testing of more elaborate forms for generalized cost. For nonwork travel, the VOT values are more reasonable, although there was a significant difference between New York ($6/hour) and Seattle ($11/hour). This can be explained by the model specification differences.

- Segmentation of travel time by congestion levels brought very different results. With the New York data a statistically significant effect was confirmed and actually manifested itself in the mode choice framework much more strongly than in the route-type choice framework. The congestion delay component of travel time proved to be weighted 1.8 to 3.5 versus the free-flow time. A similar test with the Seattle data did not bring reasonable results. It may be concluded that travel time segmentation by congestion levels works better in extremely congested areas, but is questionable for less-congested regions.
- With respect to direct reliability measures, the most promising model estimated with the New York data is the model for nonwork tours in which a standard deviation of travel time per unit distance was used. The corresponding reliability ratio is about 1.5 at a 10-mile distance. The most promising models estimated with the Seattle data included a formulation with buffer time per unit distance for work and nonwork trips, although a formulation with standard deviation of travel time per unit distance for nonwork trips had the right sign for all LOS variables.
- The main common effects that relate to the impact of car ownership on mode choice can be summarized as follows. There is a common tendency for carpooling to be negatively correlated with car sufficiency. Bigger households (in terms of number of workers and in terms of overall size) with fewer cars are the most frequent carpoolers. For a subchoice between transit modes, zero-car households are logically characterized by a strong propensity to walk to transit rather than drive to transit access. The probability of walk to transit is highly affected by absence of cars, or low car sufficiency, in the household. Households from these categories constitute the majority of transit users; many of them are transit captives since they either do not have cars at all, or have fewer cars than workers; hence, at least some of them become transit captives.
- Several different approaches to account for income were explored with both models, including scaling the cost variable by income (powered by a scaling parameter that should be between zero and one) and segmentation of the cost variable coefficient by income group. Although in many cases segmentation by income group resulted in better likelihood values, the team believes that the income-scaling version is more behaviorally appealing. With the New York model, a scaling parameter value of 0.8 was established for work tours and 0.6 for nonwork tours, which is in line with the previously discussed findings for route-type choice (0.6 and 0.5, respectively). The fact that the VOT elasticity with respect to income proved to be somewhat higher in the mode choice framework compared with route-type choice framework can be explained. The mode choice framework includes transit users who in general have a lower VOT and income. The corresponding version of the Seattle model, with the coefficient values corresponding to the New York route-type choice model, justified the specification with all coefficients having the right sign and being statistically significant.
- Several alternative specifications were tried with both the New York and Seattle data to capture the best cost-sharing mechanisms for carpools statistically. They included cost scaling by the powered occupancy, as well as occupancy-specific cost coefficient. The scaling strategy prevailed in New York; segmentation of the cost coefficient by occupancy was less successful. The scaling values of 0.8 for work tours and 0.7 for nonwork tours were eventually adopted for New York because those values are in line with the route choice findings. The results for Seattle indicate that the cost sharing reflected in the Seattle RP data is perhaps less strong than in the New York data.
- With the New York data set, a dummy variable that represents person status categorized by three major types (worker, adult nonworker, and child) proved to be statistically significant and was included in the base model specification described above for nonwork travel. A richer set of behavioral impacts with respect to person characteristics was found with the Seattle model specification, including some related effects on VOT of gender, age, and part-time worker status. In this regard, the New York model and Seattle model provide complementary examples of specifications that can be combined and hybridized in many ways.
- With the New York model, the shape of the distance-effect curves, which were similar to the shape obtained for work trips in the route-type choice framework, was statistically confirmed in the more general mode choice framework. Depending on the highest order of polynomial function used in the model specification (squared or cubed), the inverted U effect can be less or more prominent, with a very small impact on the overall model fit. The explanation given above can be reiterated for the same effect in the route choice framework; that is, the lower VOT for long-distance commuters is a manifestation of restructuring the daily activity–travel pattern. The team obtained roughly the same shape for both home-based work trips (HBW) and home-based other trips (HBO) with the Seattle model, with VOT rising to a maximum at a distance of about 25 miles and then decreasing, but the effect is much more pronounced for HBO. For HBO, the maximum VOT is about twice as high as the VOT for very short trips, but for

HBW, the maximum VOT is only about 20% higher than for very short trips.

- It is important to account for the main land use and density effects in the mode choice framework to ensure a reasonable background for analysis of LOS impacts and to separate these effects from the pure effects of travel time, cost, and reliability. In the New York regional conditions, the primary effects were found by segmenting trips to and from Manhattan (strongly dominated by transit) and internal trips within Manhattan (dominated by transit, walk, and taxi). These effects were captured by stratified mode-specific constants without an impact on VOT. The Seattle data indicate a somewhat similar effect for trips to the central business district (CBD).

Basic Specification, Segmentation, and Associated Value of Time

Both mode choice models have a rich set of explanatory variables including LOS variables, as well as various person and household variables. This provides a reasonable background for further tests with different functional forms for the generalized cost. The overall scale of time and cost coefficients (specifically for auto time, which is the focus of the current study) is reasonable. It must be taken into account that the LOS variables in a tour model should be approximately doubled when compared with a trip mode choice model. Thus, the corresponding coefficients for time and cost should be halved for a trip mode choice model when it is directly compared with a tour mode choice model. This is the case for auto in-vehicle time; for example, for work-relate travel, it is –0.014 for the New York tour mode model and –0.029 for the Seattle trip mode model.

VOT is directly comparable between tour and trip models. For work tours in New York and work trips in Seattle, the base model specifications showed a relatively low VOT for auto users of \$6–\$7/hour. This value is not, however, recommended for use in other models. The team decided not to enforce a more reasonable VOT at this stage but rather to continue testing of more elaborate forms for generalized cost. For nonwork travel, the VOT values are more reasonable, although there is a significant difference between New York (\$6/hour) and Seattle (\$11/hour). This can be explained by the model specification differences. Although the New York model has generic time coefficients, cost coefficients, and VOTs, the Seattle model explicitly distinguishes between auto users and transit users by employing mode-specific time and cost coefficients. This distinction must be taken with caution, because in the choice framework, utilities are not directly associated with mode users. In fact, every traveler is exposed to all modes. However, in reality, many auto users and transit users are repetitive in their choices. Thus, the chosen modes create a latent segmentation of the users themselves, which is partially captured by the estimated mode-specific coefficients.

Travel Time Segmentation by Congestion Levels and Facility Type

Segmentation of travel time by congestion levels brought very different results in the two regions. With the New York data, a statistically significant effect was confirmed and actually manifested itself in the mode choice framework, much more strongly than in the route-type choice framework. The congestion delay component of travel time proved to be weighted 1.8 to 3.5 versus the free-flow time. It is logical that mode choice framework provides a better statistical support for this phenomenon compared with route-type choice framework. In the New York region, transit share for trips to and from Manhattan constitutes 80%, but it is less than 10% for the rest of the region, and the corresponding auto trips have the biggest congestion delay. Thus, in the mode choice framework a congestion-averse attitude of transit users in addition to auto users is captured.

A similar test with the Seattle data did not bring such reasonable results. It should be noted that the Seattle model operates with a different segmentation of time than the New York model. In the Seattle model, links are broken into two categories (overcongested and other), but in the New York model, entire trip travel time is broken into a free-flow component and congestion delay. However, in both specifications the same phenomenon is captured, and in general, the trips with a greater number of links with $V/C > 1.2$ should have the biggest congestion delay. Thus, although the two segmentation schemes are not equal, the results should be strongly correlated. The team believes that the failure of this particular component with the Seattle data is the consequence of very different regional conditions compared with New York. It may be concluded that travel time segmentation by congestion levels works well in extremely congested areas, but is questionable for less congested regions where the differences between different trips in terms of congestion are somewhat blurred by the crudeness of synthetic skims.

As mentioned above in the route-type choice context, the team does not propose this method as the main vehicle for the current research, despite the strong statistical evidence from the New York data. In general, highway travel time segmentation is only a proxy for direct measures of travel time reliability.

Incorporation of Travel Time Reliability and Value of Reliability

Introduction of direct reliability measures in both models proved to be difficult, and many attempted specifications in the models failed to produce reasonable and statistically significant results. In general, it was difficult to simultaneously obtain the

right (negative) sign on average travel time, cost, and travel time reliability measures. This model specification is inherently fragile with RP data because of the correlation between all three variables (although using a standard deviation per unit distance significantly alleviates this problem). As mentioned above, part of the problem is the synthetic nature of the reliability measures and the quality of the other LOS skims. However, with some particular specifications, it proved possible to generate a logical model structure with all three variables in place.

The most promising model estimated with the New York data is the model for nonwork tours in which a standard deviation of travel time per unit distance was used. The corresponding reliability ratio is about 1.5 at a 10-mile distance. The most promising models estimated with the Seattle data included a formulation with buffer time per unit distance for work and nonwork trips, although a formulation with standard deviation of travel time per unit distance for nonwork trips had the right sign for all LOS variables. At this stage the decision was made to continue with the most promising specifications and to explore additional effects and impacts that could interact with the impacts of LOS variables.

Impact of Household Car Availability

The rich set of explanatory variables in the New York and Seattle models results in many logical impacts of congestion and pricing on mode choice. The main common effects that relate to the impact of car ownership on mode choice with respect to auto and transit modes can be summarized as follows:

- For both work and nonwork travel, there is a common tendency for carpooling to be negatively correlated with car sufficiency. Bigger households (both in terms of number of workers and overall size) with fewer cars are the most frequent carpoolers. It is important to note that about 80% of the observed carpools are intrahousehold in both regions;
- For both work and nonwork travel, drive to transit requires a car. Thus, for a subchoice between transit modes, zero-car households are logically characterized by a strong propensity to use walk to transit rather than drive to transit access; and
- For both work and nonwork travel, walk to transit is highly related to the absence of cars or low car sufficiency. Households from these categories provide the majority of transit users. Many of these transit users are transit captives because they either do not have cars at all or have fewer cars than workers; hence, at least some of them become transit captives.

The New York model provides some interesting behavioral insights about using taxis, which is a very frequently used mode in Manhattan. However, taxi is not a frequent mode in Seattle, and it was not included in the Seattle model formulation. The Seattle model includes nonmotorized modes and provides some additional insights with regard to them. The New York mode choice model includes only motorized modes, because the split between motorized and nonmotorized travel is modeled in the New York Metropolitan Transportation Council's best practice model by a separate model that was added to the New York model system due to a large proportion of walk trips in Manhattan. It was found that a mode choice framework that includes motorized and nonmotorized modes is less effective in the extreme conditions of New York, where many trips are generated as nonmotorized and are not subject to mode choice per se. However, these special modes are not in the focus of the current research.

Impact of Household or Person Income

Several approaches were explored with models in both regions, including scaling the cost variable by income and segmentation of the cost variable coefficient by income group. Although segmentation by income group resulted in many cases in better likelihood values (as illustrated with the Seattle model), the team believes that the income-scaling version is more behaviorally appealing. Segmentation by income group requires an arbitrary setting of income categories that can be quite broad. Also, it does not guarantee a smooth monotonic effect across all categories.

With the New York model, scaling parameter values of 0.8 and 0.6 were established for work tours and nonwork tours, respectively. These values are in line with the previously discussed findings for route-type choice (0.6 and 0.5, respectively), although they are not identical. The fact that VOT elasticity with respect to income proved to be somewhat higher in the mode choice framework than the route-type choice framework can be explained. The mode choice framework includes transit users, who in general have a lower VOT and income. Thus, with the generic specification of the LOS variables and cost-scaling parameters, this might result in the higher sensitivity to cost. This means that the constant elasticity to cost across a wide range of income groups and modes is still an analytically convenient simplification, and some more elaborate cost-scaling forms should be explored.

The corresponding version of the Seattle model, with the coefficient values corresponding to the New York route-type choice model, justified the specification, with all coefficients having the right sign and being statistically significant. Thus, this scaling strategy for income was adopted as the main approach in the further statistical tests. As mentioned above, this functional form is also consistent with the prevailing view on VOT elasticity with respect to income (Abrantes and Wardman 2011; Borjesson et al. 2012). This formula corresponds to the constant elasticity with the coefficient less than one (i.e., weaker than a linear function).

Impact of Joint Travel and Car Occupancy

Several alternative specifications were tried with both the New York and Seattle data in order to capture the best cost-sharing coefficient statistically. They included cost scaling by the powered occupancy as well as an occupancy-specific cost coefficient. The scaling strategy prevailed in New York; segmentation of the cost coefficient by occupancy was less successful. The values of 0.8 for work tours and 0.7 for nonwork tours were eventually adopted for New York because they are in line with the route choice findings.

The results for Seattle indicate that the cost sharing reflected in the Seattle RP data is perhaps less strong than in the New York data. Similar to the results for income, the exponents appear somewhat too high to fit the Seattle data, leaving one to question whether lower values would be more appropriate or whether other changes to the model specification could be found to better match the sensitivity found in the New York data.

As discussed above in the corresponding section on the impact of car occupancy on route-type choice, some additional dimensions within this effect could be further explored. In particular, intrahousehold and interhousehold carpools can have different cost-sharing mechanisms. It is expected that cost sharing should be higher for interhousehold carpools (that means the power coefficient close to 1.0) and lower for intrahousehold carpools (power coefficient close to zero). Additionally, cost sharing between adults might be stronger than between adults and children.

Impact of Gender, Age, and Other Person Characteristics

With the New York data set, a dummy variable representing person status, categorized by three major types (worker, adult nonworker, and child), proved to be statistically significant and was included in the base model specification described above for nonwork travel. Workers are characterized by a higher propensity to use household cars for solo driving, except for commuter rail, which workers use more frequently (as it is also the commuting mode for many of them). Nonworkers carpool and use transit more frequently. Children are the most frequent carpool and taxi passengers compared with both workers and nonworkers.

A richer set of behavioral impacts with respect to person characteristics was found with the Seattle model specification, including some related effects on VOT of gender, age, and part-time worker status. For both HBW and HBO, females have a marginally significant negative coefficient, which when added to the main travel time coefficient results in VOT for females about 20% to 25% higher than for males. As age increases from a base of 18 years old, a slight increase is seen in VOT for HBW trips, and a slight decrease is seen in VOT for nonwork trips. For part-time workers, a somewhat lower VOT for HBW trips is seen, with a magnitude similar to the gender effect, but in the opposite direction and less significant.

In this regard, the New York model and Seattle model provide complementary examples of specifications that can be combined and hybridized in many ways. The best existing specifications were preserved for the further testing. It is difficult to recommend one particular model structure that would fit all possible regional conditions. Person variables can be effectively used to stratify mode choice constants (as the New York model has shown) or to stratify VOT (as the Seattle model has shown), or both.

Effect of Tour or Trip Length

With the New York model, the shape of the distance-effect curves was estimated to be similar to the shape obtained for work trips in the route-type choice framework and was statistically confirmed in the more general mode choice framework. Depending on the highest order of polynomial function used in the model specification (squared or cubed), the inverted U effect can be less or more prominent, with a very small impact on the overall model fit. The explanation given above for the same effect in the route choice framework can be reiterated, that the lower VOT for long-distance commuters is a manifestation of restructuring the daily activity–travel pattern. Long-distance commuters tend to simplify their patterns and not to have many additional out-of-home activities on the day of regular commute because the work activity and commuting take the lion's share of the daily schedule. To compensate for this, commuters tend to have compressed work weeks or telecommute more frequently, which provides an opportunity to combine nonwork activities in one particular day of the week (most frequently, Friday) when they do not commute to work.

The team obtained roughly the same shape for both HBW and HBO trips with the Seattle model, with VOT rising to a maximum at a distance of about 25 miles and then decreasing, but the effect is much more pronounced for HBO. For HBO, the maximum VOT is about twice as high as the VOT for very short trips, but for HBW, the maximum VOT is only about 20% higher than for very short trips.

Given the consistent statistical evidence from both regions with respect to work travel, the polynomial (quadratic) form was adopted as the main structure for the subsequent tests. Adoption of the same form for nonwork travel was problematic given its failure with the New York data.

Impact of Urban Density and Land Use

This component is somewhat peripheral to the main purpose of the current research. However, it was important to account for the main land use and density effects in the mode choice

framework to ensure a reasonable background for the analysis of LOS impacts and to separate these effects from the pure effects of travel time, cost, and reliability. In the New York region, the primary effects were found by segmenting trips to and from Manhattan (strongly dominated by transit) and internal trips within Manhattan (dominated by transit, walk, and taxi). These effects were captured by stratified mode-specific constants without an impact on VOT. The Seattle data indicate a somewhat similar effect for trips to the CBD, but this region does not have a metropolitan core comparable to Manhattan; thus further analysis for internal trips within the CBD made little sense.

Joint Mode and Time-of-Day Choice: Revealed Preference Framework

Overview of Section, Approach, and Main Findings

Linkage Between Mode and TOD Choice

In practice, mode choice models are often estimated separately from TOD choice models. Typically, mode choice models are estimated using the auto and transit LOS variables for the actual chosen TOD. In model application, the mode choice model can be applied conditional on the choice from the TOD choice model and, ideally, mode choice logsums for each TOD alternative are passed up to the TOD model, as well. Several of the ABM systems in use have applied this approach. Alternatively, a TOD outcome can be drawn stochastically from the aggregate shares before the application of the mode choice model, and then the TOD model can be applied conditional on the prediction from the mode choice model, predicting the TOD using the LOS for the predicted mode. The ABMs used in Denver and Sacramento use this latter approach. As it is not obvious whether TOD should be predicted conditional on mode choice or vice versa, the best approach is to estimate joint TOD and mode choice models and empirically investigate nesting structures between mode and TOD. That is the approach used for this research, with the results described in this section.

Both this model structure and the way in which utility components related to TOD choice were formed are discussed step by step below.

Seed Hybrid Time-Of-Day Choice–Duration Structure

The seed structure used in this research with the New York data is a model for scheduling travel tours that can predict departure-from-home and arrival-back-home time for each tour with enhanced temporal resolution. The model formulation is fully consistent with the tour-based modeling paradigm and is designed for application within an individual microsimulation framework. TOD choice models of this type have been estimated and applied as a part of the activity-based travel demand model system developed in the regions in and around Columbus, Ohio; Atlanta, Georgia; and Sacramento, San Diego, and the San Francisco Bay Area, California.

The model is essentially a discrete choice construct that operates with tour departure-from-home and arrival-back-home time combinations as alternatives (Vovsha and Bradley 2004). The utility structure, which is based on "continuous shift" variables, represents an analytical hybrid that combines the advantages of a discrete choice structure (flexible in specification and easy to estimate and apply) with the advantages of a duration model (parsimonious structure with a few parameters that support any level of temporal resolution, including continuous time). The model is applied with a temporal resolution of 1 hour. It is expressed in 20 alternatives for departure and arrival times from 5:00 a.m. through 11:00 p.m. as follows:

1. Earlier than 5:00 a.m.
2. 5:00–5:59 a.m.
3. 6:00–6:59 a.m.
4. 7:00–7:59 a.m.
5. 8:00–8:59 a.m.
6. 9:00–9:59 a.m.
7. 10:00–10:59 a.m.
8. 11:00–11:59 a.m.
9. 12:00–12:59 p.m.
10. 1:00–1:59 p.m.
11. 2:00–2:59 p.m.
12. 3:00–3:59 p.m.
13. 4:00–4:59 p.m.
14. 5:00–5:59 p.m.
15. 6:00–6:59 p.m.
16. 7:00–7:59 p.m.
17. 8:00–8:59 p.m.
18. 9:00–9:59 p.m.
19. 10:00–10:59 p.m.
20. 11:00 p.m. or later.

This is expressed in $(20 \times 21)/2 = 210$ hour-by-hour departure–arrival time alternatives. Only feasible combinations in which the arrival hour is equal to or later than the departure hour are considered.

Analogue Between Discrete Choice and Duration Models Through Shift Variables

Consider a discrete set of time-related alternatives, such as alternative duration for some activity in hours $t = 1, 2, 3, \ldots, n$.

A general form for the probabilistic model that returns the probability of activity duration is

$$P(t) = f(t) \tag{3.27}$$

where $f(t)$ represents a probability density function for duration. This general form is not really operational because it incorporates any possible parametric or nonparametric density function and does not suggest any constructive method for model estimation.

Duration models operate with a special function $0 < \lambda(t) < 1$ that represents a termination rate (frequently called *hazard* in the literature) at time t assuming that the activity has not been terminated before (i.e., at one of the time points $1, 2, \ldots, t-1$). The probability density function for a duration model in discrete space takes the following form:

$$P(t) = \lambda(t)\prod_{s=1}^{t-1}[1-\lambda(s)] \tag{3.28}$$

There is a direct correspondence between the general-form density function and the continuous duration model. Any duration model has the correspondent density function calculated by Equation 3.28, and any density function has the underlying termination rate calculated by the following formula:

$$\lambda(t) = \frac{f(t)}{1-\sum_{s=1}^{t-1} f(s)} \tag{3.29}$$

The duration-type formulation (Equation 3.28) has both operational and meaningful advantages over the general model formulation because the termination rate function $\lambda(t)$ is frequently easier to parameterize, estimate, and interpret than the density function itself. These advantages are especially clear when modeling processes with duration-related conditionality. In addition, having the termination rate $\lambda(t)$ as an analytical function of t makes the duration model equally practical for any units of t.

Formulation of the duration model as a discrete choice model employs the following analytical form, assuming a multinomial logit model in this case:

$$P(t) = \frac{\exp(V_t)}{\sum_s \exp(V_s)} \tag{3.30}$$

where V_t denotes the utility function that is a linear-in-parameters function of independent variables:

$$V_t = \sum_k \beta_{kt} x_{kt} \tag{3.31}$$

where

$k \in K$ = household-, person-, zonal-, and duration-related variables;

x_{kt} = values of the variables for each alternative; and

β_{kt} = coefficients for the variables.

There is again a direct correspondence between the choice model (Equation 3.30) and the general-form density function (Equation 3.27). Any choice model has the corresponding density function calculated by Equation 3.30, and any density function (Equation 3.27) has an underlying set of utilities that are calculated by the following formula:

$$V_t = \ln f(t) \tag{3.32}$$

As in the case of duration models, discrete choice models (Equation 3.30) have advantages over the general formulation (Equation 3.27) because utility expressions (Equation 3.31) are easier to parameterize, estimate, and interpret than the density function itself. However, when the utility expression (Equation 3.31) is formulated in a general way with all alternative-specific coefficients and variables, the choice model (Equation 3.30) gets more complex with the addition of temporal resolution, which is not the case with the duration model (Equation 3.28). Also, the multinomial-logit formulation with independent alternative-specific variables suffers from the IIA property (independence from the irrelevant alternatives) with respect to those variables, ignoring the fact that the duration alternatives are naturally ordered.

Both of these deficiencies of the discrete choice formulation can be overcome using a certain specification of the utility function (Equation 3.31). This specification stems from an analogy that can easily be established between the duration model (Equation 3.28) and discrete choice model (Equation 3.30). Consider a ratio of densities for two subsequent points in time stemming from the two models, and restrict it to be equal in both cases:

$$\frac{P(t+1)}{P(t)} = \frac{\lambda(t+1)\times[1-\lambda(t)]}{\lambda(t)} = \exp(V_{t+1} - V_t) \tag{3.33}$$

Equation 3.33 contains several interesting and analytically convenient particular cases that lead to operational models that can be equally written and estimated in either duration form (Equation 3.28) or discrete choice form (Equation 3.30). Consider only one (actually, the simplest) case that corresponds to a duration model with a constant termination rate λ. With this assumption, Equation 3.33 is simplified to the following formula:

$$\exp(V_{t+1} - V_t) = 1 - \lambda \tag{3.34}$$

This means that there is a constant decrement in the utility function for each subsequent time point compared with the previous one, and it is the equivalent of the constant termination rate parameter of the duration model. The negative utility increment corresponds to the value of $1 - \lambda$ that is less than one. To ensure that the utility increment is independent of the time point, the variables x_{kt} and coefficients β_{kt} should be set in the utility expression (Equation 3.31) in a specific way. One of the possible ways to do this is to define all coefficients as

generic across duration alternatives ($\beta_{kt} = \beta_k$), while the variables are assumed to have the following form:

$$x_{kt} = t \times x_k \tag{3.35}$$

This formulation for the variables is not very restrictive since most of the household, person, and zonal characteristics in the TOD choice model are naturally generic across time alternatives. However, it is not true for network LOS variables that vary by TOD and should be specified as alternative specific. These variables, which are essentially time specific, violate the constant termination rate assumption. However, the discrete choice framework allows for easy hybridization of both types of variables (generic and time specific).

Using generic coefficients and variables of this type (Equation 3.35) creates a compact structure of the choice model in which the number of alternatives can be arbitrarily large (depending on the chosen time unit scale), but the number of coefficients to estimate is limited to the predetermined set *K*. These variables can be interpreted as continuous shift factors that parameterize the termination rate in such a way that a positive coefficient means the termination rate is getting lower, and the whole distribution is shifted to the longer durations. Negative values work in the opposite direction, collapsing the distribution toward shorter durations.

In the current research, the team also considered a nonlinear generalization of shift variables in the following forms:

$$x_{kt}^1 = t \times x_k; \quad x_{kt}^2 = t^2 \times x_k \tag{3.36}$$

where x_{kt}^1 and x_{kt}^2 are used in the utility function as independent variables with estimated coefficients β_k^1 and β_k^2 consequently. This extension of model structure allows for capturing some nonlinear effects, in particular saturation effects, in which the impact of a certain variable x_k is expressed in differential shifts along the duration time line. Essentially, the resultant multiplier for original variable x_k in the utility function ($t \times \beta_k^1 + t^2 \times \beta_k^2$) represents the timing profile for the impact of this variable.

In addition to nonlinear shifts in the current research, the team also applied various referencing and constraining schemes for shift variables. Referencing means that the shift is calculated relative to a certain point in time, and differential shifts can be applied for being earlier or later. Referencing can be formalized in the following way:

$$x_{kt}^3 = \min(t - t_k, 0) \times x_k; \quad x_{kt}^4 = \max(t - t_k, 0) \times x_k \tag{3.37}$$

where

t_k = reference time point (alternative) for the variable;
x_{kt}^3 = variable corresponding to shifts to an earlier time than the reference alternative; and
x_{kt}^4 = variable corresponding to shifts to a later time than the reference alternative.

Constrained shifts are only applied for a certain subset of adjacent alternatives, rather than for all 20 alternatives. For example, some peak-spreading effects can be localized within a specific peak period, such as 6:00–10:00 a.m., and are not relevant to the later hours.

In the process of model estimation, all types of shift variable transformations are applied in combination, including nonlinear effects, referencing, and constraining. The best combined form is defined by statistical fit and also by meaningful behavioral interpretations. The resulting impact of each variable x_k is referred to as its *timing profile.* It essentially singles out the impact of this variable on TOD choice. If the variable itself is a dummy (like female gender or income group indicator), the timing profile is expressed in utility units. This is the most common case with a straightforward interpretation. If the variable is continuous (like travel time or distance), then the interpretation of timing profile is more complicated and is expressed as a relative impact of each minute or mile of travel on TOD choice.

Time-of-Day Model Formulation for a Tour

Scheduling of an entire travel tour requires that the choice alternatives are formulated as tour departure-from-home (g) and arrival-at-home (h) hour combinations (g, h). Tour duration is derived as the difference between the arrival and departure hours ($h - g$). In the current research, tour duration incorporates both the activity duration and travel time to and from the main tour activity, including intermediate stops.

The tour TOD choice utility for a single tour can be operationalized in the following general form (Vovsha and Bradley 2004; Abou-Zeid et al. 2006; Popuri et al. 2008):

$$V_{gh} = V_g + V_h + D_{h-g} \tag{3.38}$$

where

g, h = departure-from-home and arrival-back-home times;
V_g = departure time choice–specific component;
V_h = arrival time choice–specific component; and
D_{h-g} = duration-specific component.

Departure hour– and arrival hour–specific components are estimated using generic shift-type variables (household, person, and zonal characteristics) according to Equations 3.35 and 3.36 with a limited set of TOD period-specific constants. Just as duration shift variables are multiplied by the duration of the alternative, departure shift variables are multiplied by the departure alternative, and arrival shift variables are multiplied by the arrival alternative.

Note that the index of the duration component is ($h - g$) rather than ($g \times h$), making the estimation procedure much simpler, since the number of duration alternatives is much less than the number of departure–arrival combinations. It should also be noted that none of the estimated components of the utility function (Equation 3.36) has an index with dimensionality ($g \times h$). Thus, the number of coefficients that

have to be estimated is in general fewer than the number of alternatives. This parsimonious structure, however, outperformed a model with a full set of ($g \times$ h) alternative specific constants (Vovsha and Bradley 2004).

Joint Time-of-Day and Mode Model Formulation

Model generalization to incorporate the mode choice dimension is straightforward and results in adding one more component to the utility function:

$$V_{ghm} = V_g + V_h + D_{h-g} + W_m(gh) \tag{3.39}$$

where

V_{ghm} = combined utility function for TOD and mode choice;

m = travel modes, car occupancy categories, and route types; and

$W_m(gh)$ = mode utility component with TOD-specific LOS variables.

Although the combined structure has a very large number of alternatives (210 × 13 = 2,730), the complexity of model estimation is approximately equal to the sum of efforts corresponding to TOD choice and mode choice due to the additive utility function. The mode-related component structures with all pertinent variables were adopted from the final version of estimated mode choice model discussed above. Thus, this mode choice construct is used as the starting point. All coefficients however, were reestimated in the more general choice framework, including travel time, cost, and reliability coefficients. It should be noted that all mode choice coefficients in the previously discussed mode choice model and this combined model were specified as generic across TOD periods to keep the model structure manageable. The mode-related utility components $W_m(gh)$ differ across TOD periods because the LOS variables were generated for each TOD (gh) specifically.

The mode choice coefficients were estimated simultaneously with the coefficients related to TOD choice. Despite the complexity of joint estimation, this method offers the significant advantage of the possibility of exploring different nested structures between the TOD and mode dimensions, as well as within each of them. Previously estimated tour TOD models of this type were fed by precalculated mode choice logsums for each TOD period. In the process of joint mode estimation, lower-level logsums were calculated automatically and adjusted according to the mode choice coefficient estimates.

Main Findings

This subsection summarizes the main model components and corresponding behavioral impacts that proved to be common for the New York and Seattle models. The main research question at this stage was whether adopting the more general framework of joint mode and TOD choice would change the main findings of the previous sections with respect to the seed form of the generalized cost of highway modes. The previously substantiated functional forms were subject to a series of additional statistical tests in which the utility function included both mode and TOD components.

The main conclusion that could be made at this stage was that for both the New York and Seattle models, the extension of the model to include TOD choice dimension in addition to mode dimension did not violate the main impacts of LOS and other variables. In particular, all main LOS components previously substantiated for more limited frameworks of route-type choice and mode choice proved to be statistically significant, with the right sign, and mostly with a similar magnitude in the more general choice context that included the TOD dimension. This confirms the main hypothesis of the C04 project, that there is a generic form of highway generalized cost that can combine mean travel time, cost, and travel time reliability (standard deviation of travel time per unit distance) measures, and that this form can be used as a seed component in the utility function through the entire hierarchy of main travel choices. This finding is encouraging, because using the same seed formulation for generalized cost from bottom up in the travel model system ensures consistency of the model system elasticities and responses to congestion and pricing.

In both the New York and Seattle models, the mode choice part of the utility for highway modes included trip-length effects on VOT through the interaction terms between travel time and distance substantiated previously for the route-type choice and mode choice frameworks. In this regard, all effects associated with trip distance captured in the mode choice framework discussed above (including nonlinear impacts on VOT) were preserved in the more holistic framework of integrated TOD and mode choices.

Several interesting direct effects of tour distance on TOD choice were captured with the New York data. The composite effect on departure time from home shows that with each additional mile of commuting, the probability of earlier departure will grow across all hours, with the strongest shift between the hours of 11 p.m. and 6 a.m. In a similar vein, each additional commuting mile proved to stretch the departure time for the return trip to home toward later hours, with the highest elasticity between 2 and 6 p.m.

The impact of car availability on the results of the New York and Seattle models proved to be very similar. The impact of household car availability on combined choice of TOD and mode was captured through the mode utility component, which proved to be very similar to the impact on pure mode choice. With the New York data, adding the TOD dimension did not change the main effects that were expressed through mode preferences by four car-sufficiency groups. In

the same vein, the mode preference effects included in the mode choice models based on the Seattle RP data were once again included in the mode and TOD choice models, and the results were much the same.

Income has several important impacts on joint choice of TOD and mode. The first effect relates to mode preferences. With the New York model, these impacts in the joint choice framework proved to be very similar to the mode preferences discussed above for the pure mode choice model. In the same vein, for the Seattle models, the team repeated the same tests that were done for the mode choice models reported above: segmenting the cost coefficient by income and vehicle occupancy versus assuming the same power function that was adopted for the analyses on the New York data. Similar to the tests with the New York model, the results were virtually unchanged from what was found for the mode choice–only models. The second income impact relates to schedule preferences. For example, the New York data showed that low-income commuters tend to have later schedules (departure from home after 9 a.m.) more frequently than medium- and high-income workers. The most prominent feature of high-income commuters is that they avoid very early starts (before 7 a.m.) compared with medium- and low-income workers.

Similar to the income variable, the joint travel variable has two intertwined effects on combined mode and TOD choice. The first effect is captured by the mode-specific utility components that enter each TOD choice alternative with the corresponding LOS variables. These effects are explained by the mode choice component, and they remain very similar in the joint mode and TOD choice formulation to the mode choice effects (specifically the subsplit between single-occupant vehicle [SOV] and high-occupancy vehicle [HOV]) discussed above. However, in the joint mode and TOD formulation, a second (direct) effect of car occupancy on TOD choice was captured with the New York model that is more related to carpool organization factors and the associated schedule constraints of the participants. Carpooling commuters are characterized by a later departure from home when very early hours, from 5 to 7 a.m., are avoided. In a similar vein, carpoolers avoid very late arrival times back home (after 7 p.m.), as late arrivals might not be convenient for at least some members of the travel party.

With respect to additional person characteristics, the team first summarized the impacts of person and household characteristics on the mode choice component, comparing it with the previously estimated pure mode choice formulations above. For the Seattle model, with cost divided by functions of income and vehicle occupancy, the team also included additive travel time variables specific to females and part-time workers, as well as an additive travel time variable multiplied by age minus 18 (set at a minimum of zero to apply only to those over 18 years of age). The results were somewhat less significant than were found in the mode choice–only models, with no significant differences in VOT related to gender, age, or part-time employment status. Second, there are significant mode preferences and TOD shift variables related to these characteristics, as reported in previous sections. Similarly, with the New York model, several TOD-related effects of person characteristics were captured in addition to the person variables included in the mode choice portion of the combined utilities. The most important distinction that strongly affects TOD choice for commuters is worker status. Part-time workers are characterized by later departure-from-home time than full-time workers. With respect to work tour duration, part-time workers are characterized by significantly shorter schedules than full-time workers. For arrival time back home, most of the effects will be derived as a composition of departure time and duration effects. For example, longer durations for full-time workers will naturally create later arrivals, all else being equal.

Travel time reliability was explored with respect to impacts on two travel dimensions. As discussed above, travel time unreliability, measured as a standard deviation of travel time per unit distance, affects choice of modes as it was found statistically significant in mode choice utilities for highway modes, in addition to the mean travel time. With the New York data, this effect becomes even more statistically significant when mode choice is considered jointly with TOD choice. In general, the greater the level of unreliability of highway travel time, the lower is the share of highway modes versus transit and other modes, all else being equal. It was also important to explore a possible direct impact of travel time reliability on TOD choice, in addition to the effect incorporated in mode choice logsums. For this purpose, with the New York model, travel time reliability in the measure was explored statistically as a shift variable in the TOD portion of the utility. The results confirmed two logical and statistically significant effects. The first effect relates to the shift of commuting departure time to hours earlier than 8 a.m., which is progressively stronger for each earlier hour. This statistical evidence fully confirms the fact that commuters have to take into account a certain extra (buffer) time in the presence of travel time unreliability. A similar symmetric effect was found for arrival time back home after work. Travel time unreliability resulted in later arrivals, with a progressive effect between 5 and 9 p.m.

Several effects associated with urban density and land use type were explored with the New York data. Some effects were already incorporated in the mode choice utilities as discussed above. The effects remained stable after extension of the choice dimensions to include TOD in addition to mode. The additional direct effect on TOD choice captured by the Manhattan dummy is associated with a significantly longer duration of work tours. Manhattan jobs are characterized primarily by office and managerial occupations that are associated with longer durations and more flexible arrangements like a compressed work week.

Basic Specification, Segmentation, and Associated Value of Time

The main conclusion that could be made at this stage was that for both the New York and Seattle models, the extension of the model to include a TOD choice dimension in addition to mode dimension did not violate the main impacts of LOS and other variables. In particular, all main LOS components previously substantiated for more limited frameworks of route-type choice and mode choice proved to be statistically significant, with the right sign, and mostly with a similar magnitude, in a more general choice context that included the TOD dimension. This finding confirms the main hypothesis of the C04 project that there is a generic form of a highway generalized cost function that can combine mean travel time, cost, and travel time reliability (standard deviation of travel time per unit distance) measures and that this form can be used as a seed component in the utility function through the entire hierarchy of main travel choices. This finding is encouraging because using the same seed formulation for generalized cost from bottom up in the travel model system ensures consistency of the model system elasticities and responses to congestion and pricing.

There are some particularly subtle effects associated with a joint consideration of mode and TOD choice compared with a pure mode choice model with fixed TOD. The primary difference, which manifests itself more strongly in a congested area like New York, is that when the TOD choice dimension becomes endogenous, it makes congestion-averse behavior more explicit. In the New York model it resulted in a stronger impact of travel time reliability. In general, due to a strong interdependence between mode and TOD choice, it is desirable to estimate these models jointly rather than sequentially. The current research has proven that this is both possible and practical, even though this practice results in a complicated choice structure, with a large number of alternatives, a large number of coefficients in the combined utility functions, and several nesting levels to explore. In the model application, the model can still be broken into a sequence of submodels by nesting levels that is equivalent to a fully joint model if the lower-level logsums are properly carried up.

Nonlinear Level of Service, Trip Length, and Location Effects

In both the New York and Seattle models, the mode choice part of the utility for highway modes included trip-length effects on VOT through the interaction terms between travel time and distance that were substantiated previously for the route-type choice and mode choice frameworks. In this regard, all effects associated with trip distance captured in the mode choice framework discussed above (including nonlinear impacts on VOT) were preserved in the more holistic framework of integrated TOD and mode choices. With the Seattle RP data, compared with the mode choice models results, the curve is less pronounced for HBO and reaches a maximum at a higher distance (around 40 miles) for HBW.

In addition, several interesting direct effects of tour distance on TOD choice were captured with the New York data. Commuting distance proved to have a direct impact on departure time from home and arrival time back home that was captured by linear and squared shift variables. The composite effect on departure time from home shows that with each additional mile of commuting, probability of earlier departure will grow across all hours, with the strongest shift between the hours of 11 p.m. and 6 a.m. In a similar vein, each additional commuting mile proved to stretch the departure time from home toward later hours, with the highest elasticity between 2 and 6 p.m.

Impact of Congestion Levels

In both the New York and Seattle regions the extension of the choice dimensions to TOD and mode did not significantly change the previous results with respect to auto time segmentation by congestion levels. Overall, the results were statistically unstable or insignificant, or both. Thus, these results only reinforced the decision to apply direct measures of travel time (un)reliability, like standard deviation of travel time per unit distance, rather than use indirect measures, like auto time weights differentiated by congestion levels.

Impact of Household Car Availability

The impact of car availability yielded very similar results with both the New York and Seattle models. The impact of household car availability on combined choice of TOD and mode was captured through the mode utility component in a way that proved to be very similar to the impact on pure mode choice. With the New York data, adding the TOD dimension did not change the main effects that were expressed through mode preferences by four car-sufficiency groups. In the same vein, the mode preference effects included in the mode choice models based on the Seattle RP data were once again included in the mode and TOD choice models, and the results were much the same, with the exception that many of the effects were estimated even more significantly.

Impact of Household or Person Income

Income has several important impacts on joint choice of TOD and mode. The first set of impacts relates to mode preferences. With the New York model, these impacts in the joint choice framework proved to be very similar to the mode preferences discussed above for the pure mode choice model. In the same

vein, for the Seattle models, the same tests were repeated that were done for the mode choice models that are reported above: segmenting the cost coefficient by income and vehicle occupancy as opposed to assuming the same power function that was adopted for the analyses on the New York data. Similar to the tests with the New York model, the results were virtually unchanged from what was found for the mode choice–only models.

The second set of income impacts tested relates to schedule preferences. For example, the New York data showed that low-income commuters tend to have later schedules (departure from home after 9 a.m.) more frequently than medium- and high-income workers. The most prominent feature of high-income commuters is that they avoid very early starts (before 7 a.m.) compared with medium- and low-income workers. This finding correlates with the nature of corresponding occupations and schedule flexibility.

Impact of Joint Travel

Similar to the income variable, the joint travel variable has two intertwined effects on combined mode and TOD choice. The first effect is captured by the mode-specific utility components that enter each TOD choice alternative with the corresponding LOS variables. These effects are explained by the mode choice component and, for the New York model, they remain very similar in the joint mode and TOD choice formulation to the mode choice effects (specifically, the subsplit between SOV and HOV), as discussed above. In a similar way, with the Seattle model, the carpooling impacts were virtually unchanged from what was found for the mode choice–only models discussed above, with less sensitivity of the cost coefficient to vehicle occupancy than what is represented in the assumed power functions, particularly in the HBO models.

However, in the joint mode and TOD formulation, a second (direct) effect of car occupancy on TOD choice was captured with the New York model. Although the first (logsum-related) effect is sensitive to LOS variables and corresponding policies like HOV and high-occupancy toll (HOT) lanes, the second effect is more related to carpool organization factors and the associated schedule constraints of the participants. Carpooling commuters are characterized by a later departure from home when very early hours, from 5 to 7 a.m., are avoided. In a similar vein, carpoolers avoid very late arrivals back home (after 7 p.m.), because late arrivals might not be convenient for at least some members of the travel party.

Impact of Gender, Age, and Other Person Characteristics

This section first summarizes the impacts of person and household characteristics on the mode choice component, comparing it with the previously estimated pure mode choice formulations discussed above. For the Seattle model, with cost divided by functions of income and vehicle occupancy, the team also included additive travel time variables specific to females and part-time workers, as well as an additive travel time variable multiplied by age minus 18 (set at a minimum of zero to apply only to those over 18 years of age). The results were somewhat less significant than those found in the mode choice–only models, with no significant differences in VOT related to gender, age, or part-time employment status.

Second, there are significant mode preference and TOD shift variables related to these characteristics, as reported in previous sections. Similarly, with the New York model, several TOD-related effects of person characteristics were captured in addition to the person variables included in the mode choice portion of the combined utilities. The most important distinction that strongly affects TOD choice for commuters is worker status, for which three main person types (full-time worker, part-time worker, and nonworker) are considered. The last category includes some commuters for job interviews, occasional work, and volunteers. Part-time workers and nonworkers are characterized by later departure-from-home time than full-time workers. With respect to work tour duration, both part-time workers and nonworkers are characterized by significantly shorter schedules than full-time workers. For arrival time back home, most of the effects will be derived as a composition of departure time and duration effects. For example, longer durations for full-time workers will naturally create later arrivals, all else being equal. However, in addition to the derived effects for full-time workers, one direct arrival time–related effect proved to be significant. Full-time workers rarely arrive back home before 3 p.m., in contrast to part-time workers and nonworkers.

Incorporation of Travel Time Reliability and Value of Reliability Estimation

Travel time reliability was explored with respect to impacts on two travel dimensions. As discussed above, travel time unreliability, measured as a standard deviation of travel time per unit distance, affects choice of modes as it was found statistically significant in mode choice utilities for highway modes, along with the mean travel time. With the New York data, this effect becomes even more statistically significant when mode choice is considered jointly with TOD choice. In general, the greater the level of unreliability of highway travel time, the lower is the share of highway modes versus transit and other modes, all else being equal. Through TOD-specific mode choice logsums, this impact also has an effect on TOD choice. However, the results for the Seattle model were less successful. For the HBW models, the reliability variables all have the incorrect sign, except when included as the buffer travel time (90th percentile minus

median) divided by distance, which has a significant negative coefficient. For the HBO models, the buffer travel time variable is again the only reliability variable with a significant negative coefficient, but this time when *not* divided by trip distance. It may be concluded that the distribution of level of congestion and associated variation in travel time reliability measures in the Seattle data was not rich enough.

It was also important to explore a possible direct impact of travel time reliability on TOD choice in addition to the effect incorporated in mode choice logsums. For this purpose, with the New York model, the travel time reliability measure was explored statistically as a shift variable in the TOD choice utility (departure from home and arrival back home components), in addition to inclusion of travel time reliability in the mode choice logsum. The results confirmed two logical and statistically significant effects. The first effect relates to the shift of commuting departure time to hours earlier than 8 a.m., which is progressively stronger for each earlier hour. This statistical evidence fully confirms the fact that commuters take into account a certain extra (buffer) time in the presence of travel time unreliability. A similar symmetric effect was found for arrival time back home after work. Travel time unreliability resulted in later arrivals, with a progressive effect between 5 and 9 p.m.

Impact of Urban Density and Land Use

Several effects associated with urban density and land use type were explored with the New York data. Some of the effects were already incorporated in the mode choice utilities as discussed above. The effects remained stable after extension of the choice dimensions to include TOD in addition to mode.

The most prominent new effect on TOD choice was associated with a simple Manhattan job dummy that captured the principal difference between commuting to Manhattan and the rest of the metropolitan area. The additional direct effect on TOD choice captured by the Manhattan dummy is associated with a significantly longer duration of work tours. This spans durations from very short to 14 hours. Two main behavioral mechanisms can explain this phenomenon. First, Manhattan jobs are characterized primarily by office and managerial occupations that are associated with longer durations and more flexible arrangements like a compressed work week. Second, in this analysis the team operated with the entire tour duration (from departure from home until arrival back home), rather than with duration of the work activity itself. Thus, additional activities (stops) on the way to and from work come into play. Logically, commuting tours to Manhattan are characterized by a higher frequency of stops, primarily in Manhattan, because of the great variety of opportunities for shopping and discretionary activities there.

Route Type, Time-of-Day, and Mode Choice: Stated Preference Framework

This section summarizes the findings from models estimated on data from the Seattle Puget Sound Regional Council (PSRC) mode choice SP experiment, the San Francisco County Transportation Authority (SFCTA) cordon pricing study SP, and the Los Angeles HOT lane SP study. These three SP experiments are described in Chapter 2, and Table 3.8 summarizes their main design characteristics. Some key differences between the experiments are as follows:

- The Seattle and Los Angeles experiments recruited people who had actually made recent trips in relevant highway corridors in the region, and then presented experiments (sent via a survey form customized to their actual trip) that offered hypothetical tolled options in that corridor. For the Los Angeles experiment, the tolled option was offered as a HOT lane or express lane alongside free general-purpose

Table 3.8. Summary of Design Characteristics of Three SP Experiments

Characteristic	Seattle SP	Los Angeles SP	San Francisco SP
SP choice context	Introduction of tolls on route (general)	Introduction of HOT or express lanes	Introduction of toll to enter downtown area
Recruitment method	Recent trips on relevant highways, from HH travel survey sample	Recent trips on relevant highways, from telephone recruit survey	Recent trips to downtown SF, recruited at parking locations
Offered nontolled auto alternative?	Yes, could be on a different highway	Yes, free general lanes on same highway	No, all auto trips to downtown pay toll
Offered different prices in the off-peak period?	Yes, same peak periods for all respondents	Yes, varied peak toll periods across sample	Yes, varied peak toll periods across sample
Offered transit mode alternative?	No	Yes	Yes
Included a travel time reliability variable?	Yes, varied frequency of extra delay, fixed at 15+ minutes duration	No	Yes, varied duration of delay, presented as “1 in 5” or “1 in 10” trips

Note: HH = household.

lanes. In the Seattle experiment, the free option could be on a different route, requiring a different travel distance;

- The San Francisco experiment recruited people who made recent auto trips and parked downtown, and then presented hypothetical options with a cordon toll charged to enter the downtown area. The experiment was customized to the actual trips and presented via computer-based interview screens, either in-person on laptops or via the internet. No free auto alternative was offered;
- All three experiments offered different prices in the peak and off-peak periods, but the San Francisco and Los Angeles experiments also customized and randomly varied the definition of the peak period across the sample in order to better estimate TOD switching preferences;
- The San Francisco and Los Angeles experiments offered a transit alternative, but no transit option was included in the Seattle experiment; and
- The Seattle and San Francisco experiments included a travel time variability and reliability attribute, but the Los Angeles experiment (by design) did not. For the Seattle SP, the duration of extra delay was fixed at 15 minutes and the frequency was varied. For the San Francisco SP, the frequency of extra delay was randomly set at either one in five trips or one in 10 trips for each respondent, and the duration of the extra delay was varied within respondents.

Basic Specification, Segmentation, and Associated Value of Time

The team attempted to make the framework for estimation of the SP data sets somewhat consistent with the framework used in the preceding chapters for the RP-based analyses. The team started with basic models and then incrementally added detail about specific SP attributes and segmentation variables. In contrast to RP data models, however, the models estimated for SP data are largely determined by the design of the SP experiment; that is, one can include only those choice alternatives and LOS attributes that were portrayed in the choice scenarios. For example, a variable for travel time reliability and variability can be included only if that attribute was explicitly included in the SP design. Furthermore, all of the variables that were included in the SP experimental design must be included; otherwise, their omission may bias the estimates of the included attributes.

Table 3.9 summarizes results for the Seattle, San Francisco, and Los Angeles SP experiments for a basic model specification including only the basic SP attributes, along with appropriate alternative-specific constants and nesting logsum parameters. For each experiment, two models are shown: one for work trips and another for nonwork trips. In this model specification, there is no segmentation by income or auto occupancy, and all relationships are assumed to be linear. For purposes of comparison, Table 3.9 generally shows only results in the form of ratios of the coefficients. Unless otherwise stated, the coefficient estimates were significantly different from zero (although the team did not account for repeated measurements within respondents in estimation, so the standard errors will be somewhat underestimated). The values of all estimated coefficients and *t*-statistics can be found in the appendix.

All SP models were estimated as nested models. For the experiments that offered both tolled and nontolled auto alternatives (Seattle and Los Angeles), the tolled and nontolled route options were nested under each TOD period, and the logsum parameters were all significantly less than 1.0, generally around 0.4. Although one would not necessarily expect to estimate the same nesting coefficients for RP and SP data (due to differences in the way the choice sets are specified), it is interesting that this same nesting of route type under time period was also found for the New York RP and Seattle Traffic Choices data sets.

For the experiments that offered a transit alternative (San Francisco and Los Angeles), it was also found best to nest the auto alternatives across time periods and put the auto versus transit choice at the highest level. The logsum parameters were estimated at around 0.6 for the HBW models, but for HBO, there was some instability in estimating logsum coefficients; coefficients constrained to 0.5 gave a better fit than a nonnested model. Note that this nesting of TOD under mode is different from the results obtained for the Seattle and New York RP data. The SP experiments, however, were offered only to auto users in the context of actual trips they had made by auto, so there will naturally be less tendency to choose transit than one would find in a representative RP sample.

A key finding of the SP experiments is the overall willingness to pay for auto travel time savings in the form of the ratio of the auto in-vehicle and cost coefficients. Table 3.9 shows very similar estimates for the Seattle and Los Angeles data sets; both are in the range of \$11–\$12/hour for HBW trips and \$9–\$10/hour for HBO trips. These values are in a range typically estimated for highway users in the context of a toll project. For the San Francisco SP, however, the team estimated values that are about 50% higher than for the other experiments (in the range of \$15–\$18/hour). This result could be due to higher incomes, on average, for those who travel to downtown San Francisco; this possibility is explicitly tested in the next section. It could also be due, however, to the different context of cordon pricing. There is no nontolled option except for switching mode (or destination), so auto users may be more willing to pay a toll, particularly in the case of downtown San Francisco, where the cost of parking is already high by comparison.

It is worth noting that all three SP experiments show very similar overall VOT for those making work trips versus nonwork trips, with VOT for work trips 10%–20% higher in each case. This is in contrast to the team's findings for the two RP studies: Although the New York RP study found higher

Table 3.9. Basic Specifications for Three SP Experiments

	SP Experiment					
	Seattle	LA	SF	Seattle	LA	SF
	HBW	HBW	HBW	HBO	HBO	HBO
Summary of Results	**psspw1**	**laspw1**	**sfspw1**	**psspn1**	**laspn1**	**sfspn1**
VOT, auto in-vehicle time ($/hour)	12.0	11.2	17.7	9.9	9.4	15.8
Values in Equivalent Minutes Auto In-vehicle Time						
Toll route constant (min)	7.0	13.5	NA	8.8	14.5	NA
Distance to avoid toll (min/mi)	0.47	NA	NA	1.02	NA	NA
Average extra delay (min/min)	2.42	NA	0.42	2.65	NA	2.91
Shift earlier in a.m. (min/min)	0.17	0.40	0.21	0.05	0.62	0.29
Shift later in a.m. (min/min)	0.28	0.91	0.66	0.25	0.97	0.09
Shift earlier in p.m. (min/min)	0.20	0.36	0.09	0.03	0.39	0.27
Shift later in p.m. (min/min)	0.10	0.79	1.71	0.00	0.72	0.24
Transit total travel time (min/min)	NA	1.36	NA	NA	0.99	NA
Transit in-vehicle time (min/min)	NA	NA	0.72	NA	NA	1.10
Transit out-of-vehicle time (min/min)	NA	NA	1.09	NA	NA	0.97
Transit service frequency (min/min)	NA	0.43	NA	NA	0.73	NA
Transit transfers (min/transfer)	NA	15.7	9.7	NA	15.9	15.9
Transit mode constant (min)	NA	57.4	28.0	NA	193.1	35.6
Nesting Logsum Parameters						
Toll or nontoll nested under TOD	0.402 (–4.4)	0.387 (–15.2)	na	0.463 (–6.2)	0.264 (–17.8)	na
OTOD (*t*-statistic versus 1.0)						
TOD periods nested under modes (*t*-statistic versus 1.0)	na	0.581 (–3.5)	0.669 (–3.8)	na	0.50 (constr)	0.50 (constr)
Summary Statistics						
Observations	1,355	2,976	2,357	1,507	2,932	2,722
Rho-squared with respect to 0	0.247	0.297	0.166	0.247	0.276	0.109
Final log likelihood	–1,414.4	–3,907.5	–2,723.7	–1,574.0	–3,961.5	–3,360.4

Note: LA = Los Angeles; SF = San Francisco; HBW = home-based work trips; HBO = home-based other trips; na = not applicable; and NA = not available.

VOT for work trips, the Seattle RP study found higher VOT for nonwork trips. Standard practice is to use much higher VOT for work trips than for nonwork trips, but such a result is rarely found in SP-based studies. According to household welfare economics, one would expect VOT for any personal trip, commuting or otherwise, to be proportional to the value of spending time in the leisure activity that the saved time would be devoted to (i.e., the value of leisure time at the margin) relative to the value of spending time driving a vehicle. This, however, assumes that travelers can schedule their travel and activities predictably, and that time is fully substitutable between activities. In reality, there are a number of reasons why those conditions may not always hold:

- There may be unexpected delays or conditions that cause time to be taken from more valuable activities, such as work or more highly valued leisure activities;
- In response to unreliability, travelers may leave more of a buffer time between activities, particularly those activities with a high penalty or disutility for arriving late. Adding buffer time to travel times results in a suboptimal scheduling of activities and, possibly, a lower-valued use of time savings; and
- All alternative uses of time are not available at all times of day. For example, many leisure activities may not be possible during the early morning hours, and it may not be possible to shift work schedules to allow travel time saved in the morning to be used later in the day. Leisure activities often need to be scheduled in coordination with others, which further limits the possibility to schedule them at any given TOD.

These various conditions may differ for work travel relative to nonwork travel. In particular, the effects of reliability and

unexpected delays will, on average, tend to be stronger for work trips than for leisure trips (although they may also be quite strong for specific types of nonwork trips). This means that the more that reliability can be explicitly accounted for in models, the more one would expect to see that the net value of travel time is similar between work and nonwork travel.

Toll Route Constant

In Table 3.9, the remaining estimates are reported as ratios relative to the auto in-vehicle time coefficient, normalized to equivalent minutes of travel time. The first row is for the toll route constant, which was significant and negative for both travel purposes in the Seattle and Los Angeles experiments. (All auto alternatives were priced in the San Francisco case, so no constant was estimated.) Negative toll route constants are typically found in SP studies, and often in RP studies, as well (including the RP studies on the New York and Seattle data described above). All else equal, the constant on the tolled route is equivalent to about 8 minutes of extra in-vehicle time for the Seattle SP, and about 14 minutes in the Los Angeles SP.

Detour Distance to Avoid Tolls

In the Seattle SP, the nontolled path was not always on the same facility and could involve driving additional distance. Each mile of extra distance was valued negatively, above and beyond the time required to drive it. The value is twice as high for nonwork as for work trips, suggesting that work trips are more willing to search for alternative paths to avoid tolls (perhaps due to more familiarity with alternative routes). The value for work trips is equivalent to about 0.5 minutes per mile. If the average speed on the extra distance were 30 mph, then it would take 2 minutes to drive each extra mile, so this extra term would increase the disutility by 25% in that case.

Extra Delay and Reliability

As mentioned above, the Seattle and San Francisco SP experiments both included attributes related to the frequency and duration of extra delays above the usual travel time. Ideally, respondents are presented with a distribution of day-to-day travel times to obtain estimates comparable to the RP-based results given in the previous sections. Some recent SP studies have done this by presenting respondents with a series of five or 10 possible travel times for each alternative instead of a single time. Both the Seattle and San Francisco experiments opted for simpler approaches. The Seattle SP obtained more significant estimates, suggesting that people find it easier to understand the approach with fixed duration of extra delay and comparing different frequencies (e.g., one in 10 trips versus one in 20 trips), rather than vice versa. In either case, it is possible to multiply the frequency and duration to obtain the expected minutes of extra delay, which is somewhat comparable to a standard deviation measure (if travel times were never shorter than the typical time). Apart from the work trip result for the San Francisco SP, which was not statistically significant, the other three estimates indicate that each expected minute of extra delay is equivalent to about 2 to 3 minutes of expected travel time. If the expected extra delay is comparable to the standard deviation, this result suggests a reliability ratio between two and three, which does not seem out of the question. In general, it is expected that an extra delay minute would be valued more than a minute of the mean travel time (Li et al. 2010; Concas and Kolpakov 2009). Standard deviation is a symmetric measure that reflects both cases of being early and late. The relationship between standard deviation and expected lateness depends on the distribution of travel time. For symmetric distributions around the mean, the expected lateness is equal to $1/\sqrt{2}$ of the standard deviation.

Shifting Out of the Peak Pricing Period

In addition to the various period-specific constants that are reported in Appendix A, all of the SP models include shift variables for people who actually traveled in the peak pricing period and who could shift out of it to pay a lower toll. In contrast to RP-based TOD shift variables that are cross-sectional in nature, the SP-based variables are pseudolongitudinal, measuring before and after responses to hypothetical system changes. In general, the Seattle measures are lower in magnitude and less statistically significant than measures from the other studies, primarily because only one definition of the peak pricing period was used for the entire sample in the Seattle experiment. This restriction does not provide much variation for identifying shifting preferences; in contrast, the other experiments used peak period definitions that were semirandomly customized to respondents' trips.

In general, the largest resistance to shifting trip times is to shift later in the a.m. peak, particularly for work trips. As many individuals have to be at work by a specific time, this result makes sense. For the Los Angeles and San Francisco experiments, the disutility of each minute of shifting later in the a.m. is almost as large as the disutility of a minute of travel time, which is quite a high result. In the p.m., there seems to be somewhat more resistance to shifting later than shifting earlier. These results will also tend to vary depending on the current departure time; for example, someone already traveling at 7 a.m. may be less willing to shift earlier than someone traveling at 8 a.m. They may also be nonlinear, with some travelers having thresholds at which the shift becomes more difficult to schedule. These aspects of behavior are investigated in Chapter 4.

Shifting to Transit

A number of different variables were used to represent the transit alternative, which is necessary to give respondents some clear idea of how attractive the transit alternatives would be for their particular trips. This includes travel time broken down into components, as well as transfers, frequency, and fare. Because the SP samples only included actual auto users and transit was offered as an alternative to paying tolls, it would not be expected that these experiments would provide the most accurate or representative measures of the value of various transit service levels; RP and SP data from representative samples are better for that. Nevertheless, it can be seen that each minute of travel by transit has estimated values fairly similar to the value of auto time. Each transit transfer has a disutility equal to about 15 minutes of extra travel time, which is in a range typically estimated. The most interesting result is perhaps for the residual mode constant for transit. As one might expect, the resistance to switching to transit, all else equal, is higher for the Los Angeles experiment than for the downtown San Francisco experiment, equivalent to about 60 minutes and 30 minutes of auto travel time, respectively, for work trips. There is a wide selection of transit options into downtown San Francisco, whereas many of the highway corridors studied in the Los Angeles experiment currently have very little transit service, and many of the Los Angeles respondents may have never used transit for those trips. The resistance to shifting to transit is particularly high for the Los Angeles nonwork trips.

Impact of Household or Person Income

The discussion in this section and the following sections reflects the estimation of a second set of models that included the following additional variables:

- Segmentation of the travel cost coefficient by income quartile;
- Segmentation of the auto in-vehicle time coefficient by occupancy (SOV versus HOV); and
- Segmentation of the TOD shift variables by actual TOD, plus estimation of nonlinear functions rather than simple linear effects.

The income effects on the cost coefficient and VOT are shown in detail in the appendix and are summarized in Table 3.10 and Figure 3.6. The differences between income groups are generally significant and in the expected direction, with cost having a more negative coefficient (lower VOT ratio) for lower-income groups. When plotted in the graph, all of the experiments and purposes show a fairly similar trend of increasing VOT across quartiles. (Roughly, the lowest quartile is below $30,000, the second is $30,000–$60,000, the third is $60,000–$100,000, and the highest is over $100,000; these division vary somewhat by sample.) Curiously, after the change of model specification, the San Francisco sample shows somewhat lower VOT than the other regions, except in the highest income quartile. The graph also uses dotted lines to show approximately what the curve would look like if the VOT trend conformed to a power function of 0.4, 0.5, 0.6, or 0.7. In general, the curves of 0.4 and 0.5 seem closest in slope to the estimated trends, slightly lower than the exponents of 0.5 and 0.6 that were assumed in the RP analyses for this project.

Impact of Joint Travel

After various specification tests, in the SP experiments the effect of vehicle occupancy (SOV versus HOV) on willingness to pay was captured better by segmenting the travel time coefficient rather than the cost coefficient. As summarized in Table 3.11 (the detailed estimation results are given in Appendix A), willingness to pay is higher for HOV than for SOV in all models, particularly for nonwork trips. In general, however, the effect is less than linear with vehicle occupancy, and less than the power function exponents assumed in the

Table 3.10. Cost Coefficients and VOT by Income in Three SP Experiments

	SP Experiment					
	Seattle	LA	SF	Seattle	LA	SF
	HBW	HBW	HBW	HBO	HBO	HBO
Income Quartile	psspw2	laspw2	sfspw2	psspn2	laspn2	sfspn2
Lowest	8.6	7.6	6.0	6.7	6.9	4.4
Second	10.2	8.3	7.5	7.9	7.2	6.7
Third	13.2	10.5	9.2	9.1	8.5	7.1
Highest	17.7	16.2	11.6	12.7	10.3	16.6

Note: VOT is for auto in-vehicle time.

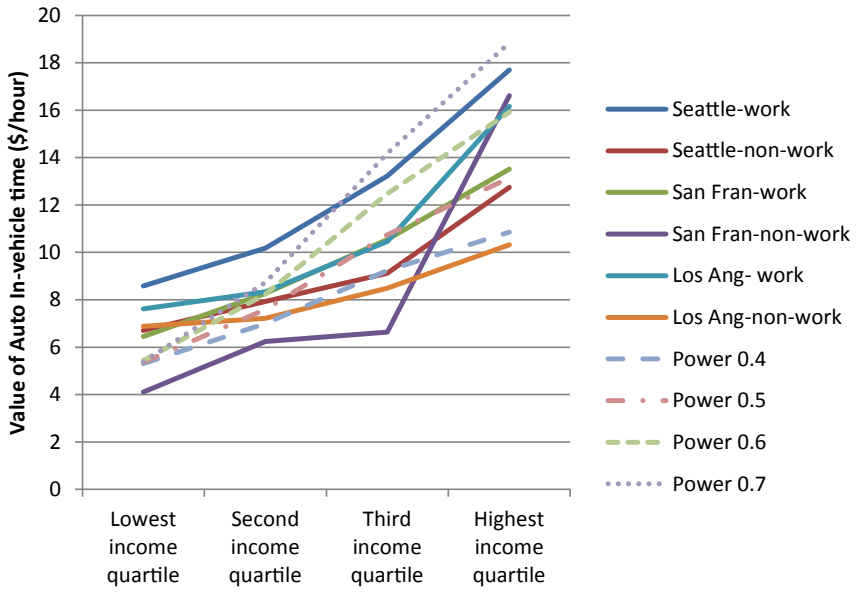

Figure 3.6. Summary of estimated VOT by income group.

RP analysis. Typically, SP samples only include the vehicle driver and not the other vehicle occupants, and it is not always clear to what extent respondents are answering only on their own behalf and to what extent they are answering for the entire traveling party, particularly with regard to sharing payment of tolls or other travel costs. As a result, SP results may be more accurate and representative for SOV trips than for HOV trips.

Incorporation of Departure Time Shift Effects

The San Francisco and Los Angeles SP data sets allowed detailed analysis on the willingness to shift out of peak pricing periods as a function of the toll level, the amount of time shift necessary, and the current time of travel. The exact coefficients are given in the appendix, and the functions are plotted below. As an example, the two graphs in Figure 3.7 show the disutility of shifting departure time either earlier or later to avoid the a.m. peak pricing period for trips to work. The results indicate a stronger resistance to moving work departure time later versus moving it earlier, at least for smaller shifts. In the range of 0 to 45 minutes, the second chart has a slope of more than 1 minute versus SOV travel time; that is, each minute of moving departure time earlier has a disutility worth more than 1 minute of in-vehicle time. At higher levels, however, the slope flattens out. Presumably, once one is very late for work, additional shifts do not make as much difference.

The shifts in the curves for different actual departure times indicate that those who actually go to work very early in the morning are more resistant to changing departure time in either direction, earlier or later. It is understandable that these people would be more averse to shifting earlier, since they would need to start their day very early to do so. The fact that early risers are also more averse to moving later may be due to the fact that they have less flexible work schedules. That the same trend is found for both experiments provides evidence that this finding is not an anomaly.

For nonwork trips in the a.m. peak, the picture is not as clear, as shown in Figure 3.8. For the San Francisco experiment, more resistance is again found to moving later than moving earlier, but with a different picture by TOD. In this case, it is those who are already traveling later in the a.m. who are most resistant to moving even later. For moving earlier, there is very little difference related to the actual TOD. For Los Angeles, the pattern for nonwork trips in the a.m. looks more similar to the pattern for work trips. The resistance to

Table 3.11. Time Coefficients and VOT Segmented by Occupancy for Three SP Experiments

	SP Experiment					
	Seattle	LA	SF	Seattle	LA	SF
	HBW	HBW	HBW	HBO	HBO	HBO
Summary of Results	psspw2	laspw2	sfspw2	psspn2	laspn2	sfspn2
Ratio of shared ride (HOV) VOT	1.03	1.28	1.12	1.39	1.34	1.90

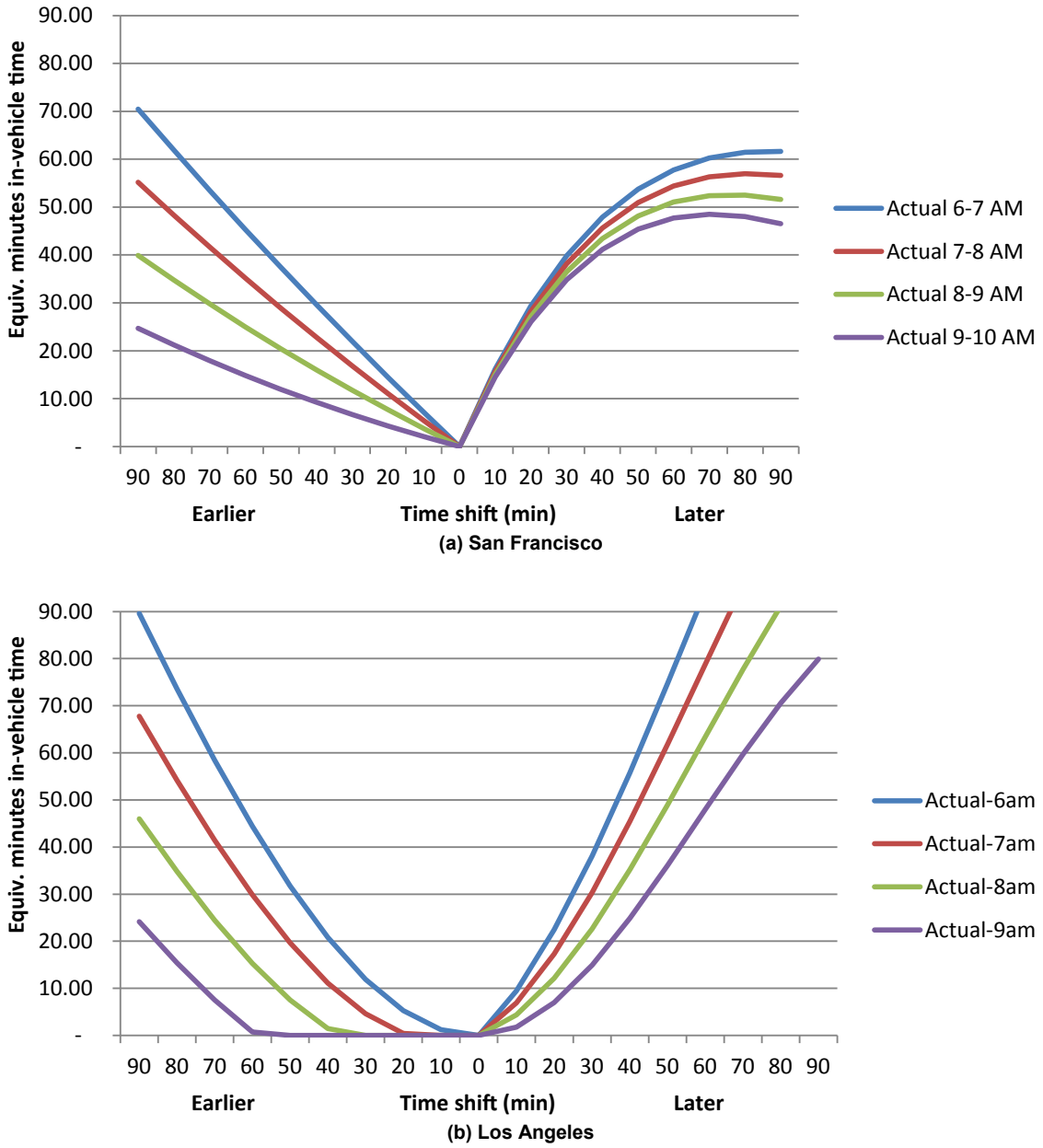

Figure 3.7. Resistance to shifting time to respond to TOD for a.m. peak work trips.

moving later levels off after about 45 minutes (the bend in the curve is an artifact of the cubic function adopted).

For nonwork trips in the p.m., both San Francisco and Los Angeles show similar sensitivities, with generally somewhat less resistance to shifting times than in the a.m. peak, as shown in Figure 3.9. For San Francisco, it is once again seen that those who travel earlier in the day are somewhat more resistant to changing times. The trips made earlier in the afternoon may be more likely to be for fixed appointments than for trips made after usual office hours. For Los Angeles, it can be seen that those traveling earlier are more averse to shifting earlier, but those traveling later in the afternoon are more averse to shifting later.

Incorporating Unobserved Heterogeneity

To investigate unobserved heterogeneity in route-type choice and TOD, data from the Seattle SP toll choice experiment were used. Four alternatives were available to individuals:

- Peak period + free;
- Peak period + toll;
- Nonpeak period + free; and
- Nonpeak period + toll.

The peak period occurred 6–9 a.m. or 3–7 p.m. Due to overlapping of the alternatives, the nesting structure shown in Figure 3.10 was assumed for modeling.

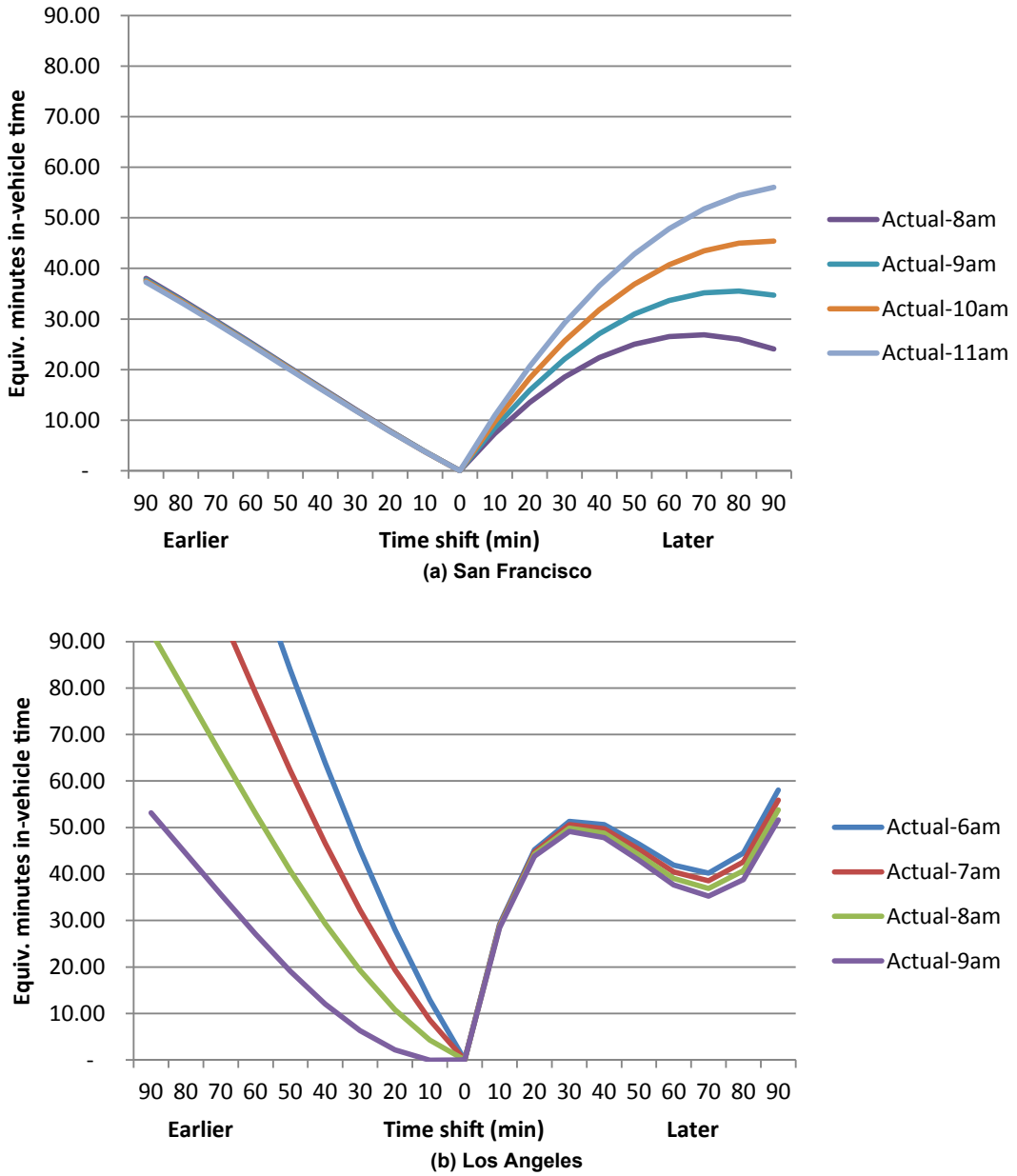

Figure 3.8. Resistance to shifting time to respond to TOD for a.m. peak nonwork trips.

The estimation results, which are shown in Table 3.12, illustrate some important insights that can be gained by accounting for unobserved heterogeneity in travel time response. By capturing the distribution of individuals' VOT, the proportion of the population with a specific VOT can be determined. In contrast with assuming all individuals have the same VOT represented by the average, Figure 3.11 shows that individuals vary greatly in their VOT.

By examining the estimation results in Table 3.12, the impact of accounting for unobserved heterogeneity in choice models is realized. First, notice that as more unobserved heterogeneity is captured, the log likelihood in general decreases. For a mixed logit model that captures serial correlation in addition to unobserved heterogeneity in the travel time coefficient, the log likelihood improves from –3095.5678 to –2789.5533. The log likelihood did not improve when just adding a random coefficient for travel time; this lack of improvement may be attributed to fixing the nesting parameter. Second, depending on the type of correlation and heterogeneity captured (e.g., randomness in the travel time coefficient or both random coefficient and serial correlation across observations), VOT varies over the population differently. Initially, by only capturing variation or heterogeneity in the travel time coefficient, the variance of VOT is larger, relative to whether both serial correlation and random travel time perception are captured. One possible explanation for

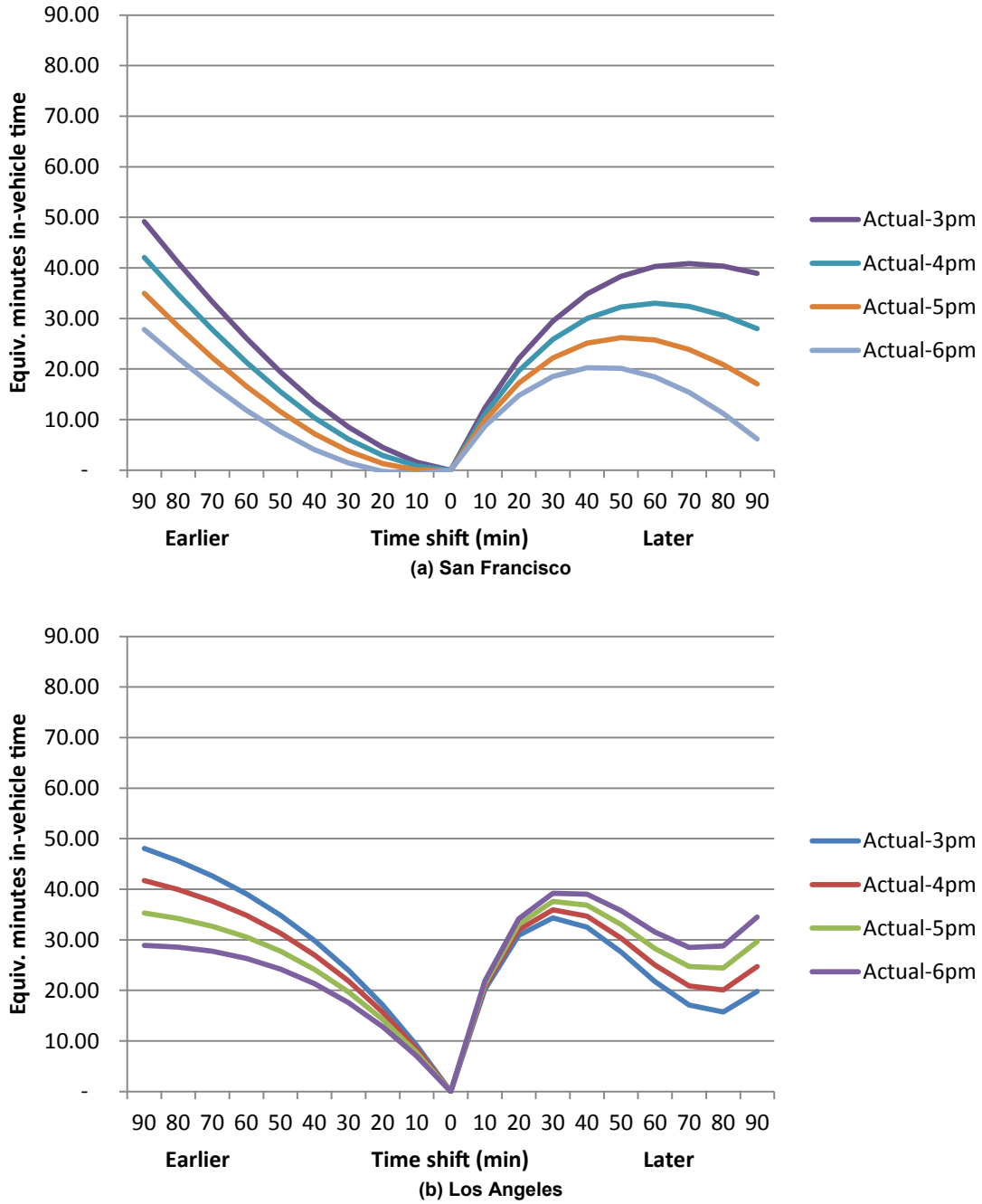

Figure 3.9. Resistance to shifting time to respond to TOD for p.m. peak nonwork trips.

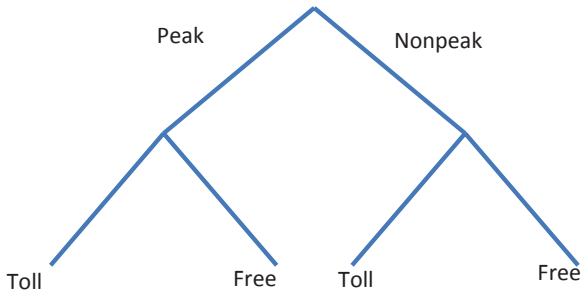

Figure 3.10. Nesting structure for Seattle SP TOD + route-type choice.

this larger variance is that more of the variance is captured by serial correlation. By not accounting for this serial correlation, VOT exhibits greater variance.

Examining the cumulative distribution functions of VOT under different assumed correlations shows that assuming only a random coefficient for travel time gives a steeper initial cumulative distribution relative to the case with both serial correlation and random coefficients.

An interesting result is that similar to the mode choice model, as more correlations are captured in the model, the variance of VOT decreases relative to the case in which no correlations are

Table 3.12. Estimation Results for RP Route-Type Choice

Model	Nested Logit		Mixed Logit		Mixed Logit	
Distribution	na		Lognormal		Lognormal	
Observations	2862		2862		2862	
Final Log Likelihood	–3041.6584		–3095.5678		–2789.55	
Rho-squared (const)	0.234		0.234		0.234	
Rho-squared (zero)	0.214		0.214		0.214	
Variable	**Coefficient**	**T-statistic**	**Coefficient**	**T-statistic**	**Coefficient**	**T-statistic**
Toll Cost ($)	–0.5580	–7.91	–0.5766	–6.54	–1.0110	–8.59
Toll Cost/Income ($/K$)	–6.9949	–2.65	–9.6423	–2.78	–9.6595	–2.10
Toll Cost∗#Passengers ($)	0.0660	2.03	0.0219	0.47	–0.2077	–3.29
Travel Time (min)	–0.1252	–13.36	–0.1821	na	–0.3558	na
Travel Distance (miles)	–0.0928	–3.53	–0.0887	–3.08	0.0263	0.55
Fraction of Times Late	–8.2644	–5.55	–8.8527	–10.16	–10.4306	–7.94
Fraction of Times Late Squared	6.564	2.52	7.0237	4.58	6.2875	2.32
Off-Peak∗actual minutes after 6 AM	–0.0152	–5.22	–0.0157	–5.21	–0.0202	–5.45
Off-Peak∗actual minutes before 9 AM	–0.0363	–12.08	–0.0372	–10.90	–0.0407	–10.35
Off-Peak∗actual minutes after 3 PM	–0.0083	–4.83	–0.0084	–4.79	–0.0106	–4.47
Off-Peak∗actual minutes before 7 PM	–0.0056	–3.34	–0.0060	–3.46	–0.0051	–2.11
Off-Peak∗actual off-peak	3.1297	43.41	3.1482	17.00	3.3892	13.99
Toll route constant	–1.0408	–11.85	–1.0755	–10.24	–1.4469	–8.76
Toll Nesting Parameter	0.3862	19.65	0.3862	na	0.3862	na
Error Term Parameters						
Variance of Beta-Travel Time	na	na	0.0835	na	0.0792	na
Variance Alternative 1	na	na	na	na	2.7751	1.98
Variance Alternative 2	na	na	na	na	3.7015	5.82
Variance Alternative 3	na	na	na	na	9.0000	na
Variance Alternative 4	na	na	na	na	2.2841	94.16
Covariance Alternative 1 Time Lag	na	na	na	na	1.1139	1.51
Covariance Alternative 2 Time Lag	na	na	na	na	1.9792	1.15
Covariance Alternative 3 Time Lag	na	na	na	na	2.7712	2.20
Covariance Alternative 4 Time Lag	na	na	na	na	2.2793	5.00
Mean Value of Time ($/hour)	11.22		15.49		18.7291	
Std. Deviation Value of Time ($/hour)	na		24.59		14.8155	

Note: na = not applicable.

captured. This is similar to the comparison between VOT for route-type choice only and the nested model, which captured both mode and route-type choice. This suggests that much of the variance associated with VOT across a population may be due to not capturing other choice dimensions, in addition to inherent taste variation across users.

Other Choice Dimensions

The choice framework described above, which includes such dimensions as TOD, mode, car occupancy, and route type, can also be effectively employed to incorporate congestion and pricing effects on all other choice dimensions, including destination choice, tour and trip frequency, daily activity patterns, and car ownership. This technique, which is based on various derived accessibility measures, has been already successfully employed in many ABMs in practice. The advantage of using accessibility measures is that all LOS variables, including travel time reliability measures included in the route, mode, or TOD utility components, will be automatically incorporated in all upper-level choice models that include these accessibility measures as explanatory variables. This technique, however, does not preclude using some relevant congestion, pricing, and reliability effects in the upper-level choice model directly. This

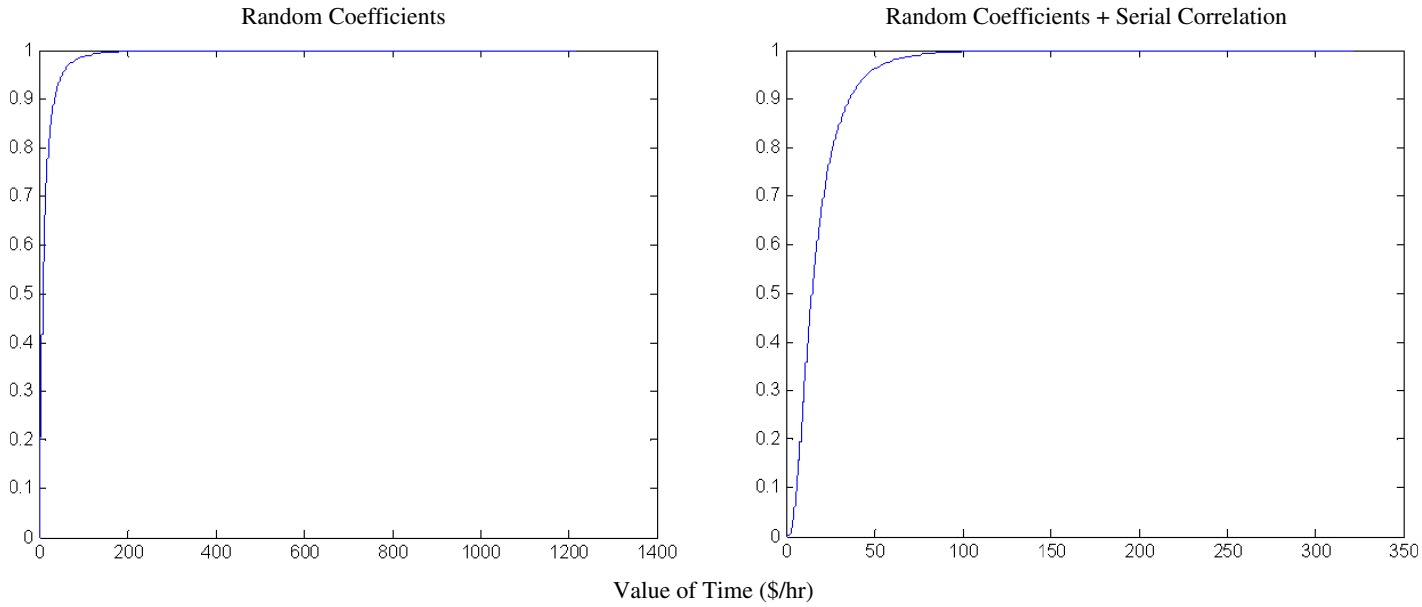

Figure 3.11. Cumulative distribution functions for VOT for* (left) *random coefficients and* (right) *random coefficients and serial correlation.

direction has been less explored and represents a sound possible topic for future research. In the current report, the team further describes the approach based on accessibility measures derived from the lower-level tour and trip models that are immediately implementable with the highway utility (generalized cost) functions described above.

General Forms of Accessibility Measures

Multiple accessibility measures have been applied in the recently developed ABMs for such metropolitan regions as Sacramento and San Diego, California; and Phoenix, Arizona. Most of the applied accessibility measures represent simplified destination choice logsums, which is the composite utility of travel across all modes to all potential destinations from an origin zone to all destination zones in different TOD periods. This way the accessibility measure is essentially a zonal characteristic that can be stored as a vector indexed by a traffic analysis zone (TAZ). Another type of accessibility measure that is calculated in the process of calculations for the zonal measure is the measure of impedance between the zones. Accessibilities of this type have to be stored as TAZ-to-TAZ matrices.

These accessibility measures are primarily needed to ensure that the upper-level models in the ABM hierarchy, such as car ownership, daily activity pattern, and (nonmandatory) tour frequency, are sensitive to improvements of transportation LOS across all modes, as well as changes in land use. Accessibility measures are similar in nature to density measures and can be thought of as continuously buffered "fuzzy" densities.

Accessibility measures are needed because it is infeasible to link all choices by full logsums due to the number of potential alternatives across all dimensions (activities, modes, time periods, tour patterns, and daily activity patterns). Accessibility measures reflect the opportunities to implement a travel tour for a certain purpose from a certain origin (residential or workplace). They are used as explanatory variables in the upper-level models (daily activity pattern type and tour frequency), and the corresponding coefficients are estimated along with the coefficients for person and household variables.

The Sacramento, Phoenix, and San Diego ABMs are among the first advanced travel models that completely avoid the "flat" area-type dummies (such as CBD, urban, suburban, and rural dummies) that are frequently used in other models to explain such choice as car ownership, tour or trip frequency, and mode choice. These qualitative labels have been completely replaced by the physical measures of accessibility sensitive to travel time, cost, (and potentially) reliability.

The applied zonal accessibility measures have the following general form:

$$A_i = \ln\left[\sum_{j=1}^{I} S_j \times \exp\left(TMLS_{ij}\right)\right] \tag{3.40}$$

where

$i, j \in I$ = origin and destination zones;

A_i = accessibility measure calculated for each origin zone;

S_j = attraction size variable for each potential destination zone; and

$TMLS_{ij}$ = TOD and mode choice logsum as the measure of impedance.

The composite travel impedance between zones, which can be referred to as an *O-D accessibility measure,* is calculated as a two-level logsum taken over the TOD periods and modes:

$$TMLS_{ij} = \mu \ln\left[\sum_{t=1}^{2} \exp\left(MLS_{tij} + \alpha_t\right)\right] \tag{3.41}$$

where

$t = 1, 2 =$ TOD periods (currently peak and off-peak are used);

$MLS_{tij} =$ mode choice logsum for a particular TOD period;

$\alpha_t =$ TOD-specific constant; and

$\mu =$ nesting coefficient for mode choice under TOD choice.

In this form, the destination choice accessibility measure is essentially a sum of all attractions in the region discounted by the travel impedance. Note that this measure is sensitive to travel improvements in both peak and off-peak periods. The relative impact of each period is regulated by the TOD-specific constant that is estimated for each travel segment (or activity type).

Accessibility measures are linearly included in a utility function of an upper-level model. To preserve consistency with the random-utility choice theory, the coefficient for any accessibility measure should be between zero and one, although it is not as restrictive as in a case of a proper nested logit model.

The general logic of inclusion of accessibility measures in travel models is as follows. For models that generate activity patterns, tours, and trips for which specific destinations are not known yet, zonal accessibility measures should be applied that describe the density of the supply of potential activity locations. For models in which accessibility to an already known location (modeled prior in the model chain) is evaluated, O-D measures should be used. In this case, there is no need for a size variable.

Size Variables by Activity Type

Size variables are prepared for each TAZ and segmented by activity type (trip purpose). The zonal size variables are calculated as linear combinations of the relevant land use variables. The corresponding coefficients can be preestimated by means of regressions of the expanded observed trip ends on the available land use variable, primarily employment types. In this sense, the size variables are similar to conventional trip attraction models.

A more theoretically consistent but also more complicated procedure would involve a simultaneous estimation of the size terms and impedance functions in the destination choice context by Equation 3.40. The estimation results for all activity types with the Phoenix data are presented in Table 3.13 for nonwork purposes (numbered from 4 through 9), a special reserved at-work subtours purpose (10), and all home-based nonwork (nonmandatory) purposes (by a combined nonwork attraction measure, 11). The explanatory variables in the rows are referred to by the tokens used in the model application, in which "nxx" implies employment for the North American Industry Classification System (NAICS) code "xx." The resulting size variables in the columns are designated by the purpose number and an abbreviation of the purpose.

Table 3.13. Zonal Size Variables for Accessibility Measures by Activity Type

Explanatory Variable		Size Variable by Activity Type							
Variable	**Description**	**4 = escort**	**5 = shop**	**6 = maint**	**7 = eating**	**8 = visit**	**9 = discr**	**10 = at work**	**11 = all**
Total_HH	Total number of households	1.0000	na	na	na	0.1421	0.3595	na	0.5016
retail	Retail employment (*n*44 + *n*45)	na	4.2810	1.4185	1.2908	na	0.4387	0.5403	7.4291
n51	Information	na	na	0.7091	na	na	na	na	0.7091
n52	Finance and insurance	na	na	na	na	na	na	0.1265	na
n53	Real estate rental leasing	na	na	2.4753	na	na	na	na	2.4753
n55	Management of companies and enterprises	na	na	na	na	na	na	1.3759	na
n56	Administrative and support	na	na	na	na	na	na	0.2357	na
n62	Health care, social assistance	na	na	1.0618	na	0.2349	na	na	1.2968
n71	Arts, entertainment, recreation	na	na	na	0.3224	na	0.9049	na	1.2273
n72	Accommodation, food services	na	1.1224	na	1.0458	na	0.4422	0.2809	2.6104
n92	Public administration	na	na	0.5356	na	na	na	0.2265	0.5356
total_emp	Total employment	na	na	na	na	na	na	0.1578	na

Note: All = combined nonwork attraction measure; at work = at-work subtours purpose; discr = discretionary purpose; eating = eating-out purpose; escort = escorting purpose; maint = (household) maintenance purpose; shop = shopping purpose; visit = visiting relatives and friends purpose.

For escorting purpose (purpose = 4), the size variable is set to the total population. This is a special purpose for which accessibility to a potential destination does not directly relate to the household decision to escort one of the household members (most frequently a child). Also, despite the fact that escorting is most frequently associated with the school purpose for the escorted person (child), the density of schools around the respondent's residence does not mean that escorting would occur more frequently. On the contrary, if a child can walk to the nearby school escorting will not be needed. Population density (accessibility to population can be viewed as continuously buffered population density) is somewhat the most reasonable zonal size variable that affects probability of escorting (all else being equal, meaning the household composition and necessity of escorting). Population density is the only accessibility measure for which both negative and positive signs can be accepted in the tour and activity frequency model. All other accessibility measures are accepted only if they have a logical positive sign.

For shopping purpose (purpose = 5), the main attractions are logically associated with retail employment and food services. Food services are frequently intertwined with shopping and it is difficult to completely separate these two land use types. It is equally true for both major shopping malls and small street shops or restaurants. It is recommended that in the future shopping size variables should be enriched with such explanatory variables as floor area to better distinguish between large shopping malls and small street shops. The (household) maintenance purpose (purpose = 6) includes a range of activities, such as personal business, banking, and visiting a post office, doctor, dentist, or lawyer, and is scattered over a wide range of related employment types including retail, information, real estate, rental, leasing, health care, social assistance, and public administration.

Eating out (purpose = 7) and discretionary (purpose = 9) purposes are closely intertwined and frequently combined in the same tour. They share the same attraction variables that relate to retail employment, recreation and entertainment, and food services, although the coefficients are logically different. In addition, discretionary purpose includes population as an additional attraction factor that serves as a proxy for such factors as sport facilities and playing grounds. It is recommended that in the future nonemployment variables like land or public parks and floor areas for sport facilities should be added, as these would enrich the attraction model for discretionary activities.

The purpose visiting relatives and friends (purpose = 8) is a special purpose for which the major attraction factor is population (number of households). In addition, visiting frequently occurs at a hospital, which is measured by employment in health. Attraction factors for trips that originated from the workplace (purpose = 10) include many variables, because at-work travel comprises three main purposes. First, at-work travel includes eating out during the lunch break, which is reflected in such attractions as retail employment and food services. Second, it may include business trips for meetings, which is reflected in such employment categories as management of companies and administration (the most probable places for business meetings) and some proportion of total employment. Third, workers might use the lunch break for personal business and shopping, which is reflected in such employment categories as finance, insurance, and public administration. Finally, a size variable that expresses total attractions for all nonmandatory home-based purposes (purposes 4–9) includes a mix of all corresponding employment types and population. Logically, retail employment plays a major role in this mix.

In addition to the complex size variables for nonmandatory activities, an ABM requires several size variables for zonal accessibility measures to mandatory activities. These are primarily used in the choice models for work from home and schooling from home. These size variables are simpler because they include all relevant variables with a coefficient of 1.0. For work from home, the size variable is employment for the relevant occupation (divided into the five categories used in the 2008 National Household Travel Survey used to estimate the Phoenix ABM). These five categories are related to employment by the NAICS codes used as the source of explanatory variables. For schooling from home, the size variable is enrollment in the corresponding school type broken into three categories: grades K–8 (elementary or middle school), grades 9–12 (high school), and university or college. The corresponding size variables are summarized in Table 3.14.

Impedance Functions by Person, Household, and Activity Type

Impedance functions are calculated as O-D matrices of logsums over modes and TOD periods (peak and off peak) according to Equation 3.41. The calculation is based on mode choice utilities that have to be calculated for all modes and TOD periods as the first step. These utilities are then combined in the composite logsum at the second step. Both steps are described below.

Mode Utilities

For calculation of accessibility measures, the set of modes is simplified to include five main modes: 1 = SOV, 2 = HOV, 3 = walk to transit (WT), 4 = drive to transit (DT), and 5 = nonmotorized (NM). The WT and DT utilities are based on the best transit skims implemented for the entire transit network including all modes. Mode utilities are also calculated separately for each of the four aggregate travel purposes: 1 = work, 2 = university, 3 = school, and 4 = other. Segmentation by travel purpose is essential because each travel purpose is characterized by a different set of mode preferences. For example, DT

Table 3.14. Zonal Size Variables for Mandatory Activities

Explanatory Variable		Size Variable	
Variable	**Description**	**Variable**	**Description**
n42	Wholesale trade	p12_whom1	Sales or marketing
n52	Finance and insurance		
n44	Retail trade	p13_whom2	Clerical, administrative, or retail
n45	Retail trade		
n53	Real estate and rental and leasing		
n71	Arts, entertainment, and recreation		
n72	Accommodation and food services		
n92	Public administration		
n11	Agriculture, forestry, fishing, hunting	p14_whom3	Production, construction, manufacturing, or transport
n21	Mining, quarrying, oil and gas extraction		
n22	Utilities		
n23	Construction		
n31	Manufacturing		
n32	Manufacturing		
n33	Manufacturing		
n48	Transportation and warehousing		
n49	Transportation and warehousing		
n51	Information	p15_whom4	Professional, managerial, or technical
n54	Professional, scientific, and technical services		
n55	Management of companies and enterprises		
n56	Administrative and support and waste management and remediation services		
n61	Educational services		
n62	Health care and social assistance		
n81	Other services (except public administration)	p16_whom5	Person care and services
Enroll1	Enrollment K–8	p17_shom1	Enrollment primary and middle
Enroll2	Enrollment 9–12	p18_shom2	Enrollment high school
Enroll3	Enrollment university and college	p19_shom3	Enrollment university and college

is frequently chosen for work purpose, but it is practically not observed for the purposes of school trips or other trips. All nonwork purposes are aggregated for calculation of impedances, although they are separated with respect to size variables. Additional important segmentation relates to household car sufficiency. The team distinguishes at this stage between three household groups: 1 = household without cars, 2 = household with fewer cars than workers, and 3 = households with number of cars greater than or equal to number of workers. This distinction is important because car sufficiency strongly affects mode availability and preferences.

Overall, by combining five aggregate modes with four travel purposes, three car-sufficiency groups, and two TOD periods, a set of $5 \times 4 \times 3 \times 2 = 120$ mode utilities was precalculated for all O-D pairs. The components of the mode utility functions and corresponding coefficients are summarized in Table 3.15. The coefficients shown were adopted for the San Diego and Phoenix ABMs. All coefficients are generic across TOD periods. The distinction between peak and off-peak utilities is due to different LOS variables. Mode utilities can incorporate STD, perceived highway time, or any other measure of travel time reliability if is supported by the network simulation and skimming procedures. This is not currently the case with the ABMs in practice, primarily because of the lack of effective network procedures that could generate reliability measures. This issue is the focus of such SHRP 2 projects as L04 and C10, which are currently under way. However, the current research lays down a complete methodology for incorporating reliability in travel demand models that is fully compatible with the potential network procedures.

Table 3.15. Components and Coefficients of Mode Utilities

Variable	SOV	HOV	WT	DT	NM
Work Travel Purpose					
SOV time (min)	–0.03	na	na	na	na
HOV time (min)		–0.03	na	na	na
Highway distance (mi)	–0.015	–0.01	na	na	–1.5
Highway distance >3 mi, dummy	na	na	na	na	–999
WT weighted time (min)	na	na	–0.03	na	na
WT fare (cents)	na	na	–0.002	na	na
WT in-vehicle time <1 min, dummy	na	na	–999	na	na
DT weighted time (min)	na	na	na	–0.03	na
DT fare (cents)	na	na	na	–0.002	na
DT in-vehicle time <1 min, dummy	na	na	na	–999	na
Zero-car household	–999	–3.0	na	na	na
Cars fewer than workers	–1.5	–2.0	na	na	na
Cars greater than or equal to workers		–2.5	na	na	na
University Travel Purpose					
SOV time (min)	–0.03	na	na	na	na
HOV time (min)	na	–0.03	na	na	na
Highway distance (mi)	–0.03	–0.02	na	na	–1.5
Highway distance >3 mi, dummy	na	na	na	na	–999
WT weighted time (min)[a]	na	na	–0.03	na	na
WT fare (cents)	na	na	–0.004	na	na
WT in-vehicle time <1 min, dummy	na	na	–999	na	na
DT weighted time (min)[b]	na	na	na	–0.03	na
DT fare (cents)	na	na	na	–0.004	na
DT in-vehicle time <1 min, dummy	na	na	na	–999	na
Zero-car household	–999	–2.0	na	na	na
Cars fewer than workers	–1.5	–1.0	na	na	na
Cars greater than or equal to workers	0	–1.5	na	na	na
School Travel Purpose					
SOV time (min)	–0.05	na	na	na	na
HOV time (min)	na	–0.05	na	na	na
Highway distance (mi)	–0.06	–0.04	na	na	–1.5
Highway distance >3 mi, dummy	na	na	na	na	–999
WT weighted time (min)[a]	na	na	–0.03	na	na
WT fare (cents)	na	na	–0.006	na	na
WT in-vehicle time <1 min, dummy	na	na	–999	na	na
DT weighted time (min)[b]	na	na	na	–0.03	na
DT fare (cents)	na	na	na	–0.004	na
DT in-vehicle time <1 min, dummy	na	na	na	–999	na
Zero-car household	–999	–1.0	na	–5.0	2.0
Cars fewer than workers	–1.5	0	na	–5.0	2.0
Cars greater than or equal to workers	0	–0.5	na	–5.0	2.0

(continued on next page)

Table 3.15. Components and Coefficients of Mode Utilities (continued)

Variable	SOV	HOV	WT	DT	NM
Other Travel Purpose					
SOV time (min)	–0.03	na	na	na	na
HOV time (min)	na	–0.03	na	na	na
Highway distance (mi)	–0.03	–0.02	na	na	–1.5
Highway distance >3 mi, dummy	na	na	na	na	–999
WT weighted time (min)[a]	na	na	–0.03	na	na
WT fare (cents)	na	na	–0.004	na	na
WT in-vehicle time <1 min, dummy	na	na	–999	na	na
DT weighted time (min)[b]	na	na	na	–0.03	na
DT fare (cents)	na	na	na	–0.004	na
DT in-vehicle time <1 min, dummy	na	na	na	–999	na
Zero-car household	–999	–3.0	na	–5.0	na
Cars fewer than workers	–1.5	–2.0	na	–5.0	na
Cars greater than or equal to workers	0	–2.5	na	–5.0	na

Note: SOV = Single Occupancy Vehicle; HOV = High Occupancy Vehicle; WT = Transit with Walk access; DT = Transit with Drive Access; NM = Non-motorized modes; and na = not applicable.

[a] WT weighted time includes in-vehicle time and out-of-vehicle time with weight = 2.5. Out-of-vehicle time includes initial wait, transfer wait, access walk, transfer walk, egress walk, and a 4-minute penalty for each transfer.

[b] DT weighted time additionally includes access drive in out-of-vehicle time.

Mode and TOD Choice Logsums

After mode utilities have been calculated for each mode, purpose, car-sufficiency group, and TOD period, they are combined into composite O-D accessibility measures; that is, mode and TOD choice logsums are derived by Equation 3.41. The list of logsum measures that have to be prepared to support various accessibility measures is summarized in Table 3.16.

Overall, 21 O-D accessibility measures are prepared to support the various zonal accessibility measures needed for different submodels of the MAG ABM. The structure of each logsum and associated parameters are summarized in Table 3.17. This table essentially represents a control file for the impedance (O-D) part of the program that calculates accessibility measures.

Each impedance measure is associated with a certain aggregate travel purpose (1–4) for which the mode utilities are calculated according to the coefficients in Table 3.15. Depending on the type of accessibility measure, car sufficiency is then taken into account. If a general accessibility measure is calculated that will be applied in the model system before the car-ownership model, the mode utilities are averaged across all car-sufficiency groups with the weight that reflects the observed proportion between different car-sufficiency groups in the region. If an accessibility measure is calculated for a specific car-sufficiency group (i.e., it will be applied after the car-ownership model), then the mode utilities for this specific group are used.

Not every mode is included in each logsum. The set of modes is restricted for two reasons. First, some modes are not observed for some of the trip purposes. For example, DT is relevant for work trips only. Second, certain modes are made unavailable in order to calculate a specific (mode-restricted) type of accessibility needed for a particular behavioral model. For example, mode-specific accessibilities that are used in the car-ownership model are based on a single representative mode each. Accessibilities that describe individual activities should logically exclude HOV. Accessibilities that describe joint activities naturally exclude SOV. Accessibilities that describe auto dependency include only modes that need an auto (SOV, HOV, and DT). Accessibilities that describe auto nondependency include only modes that do not need an auto (WT and NM).

Finally, to complete the logsum calculation across TOD periods, a bias constant for off-peak periods is specified (the peak period is used as the reference alternative with zero bias). This constant is set to replicate the observed proportion of trips in the peak period versus the off-peak period.

List of Zonal Accessibility Measures Adopted for Advanced Activity-Based Models

The set of zonal accessibility measures incorporated in the Sacramento, San Diego, and Phoenix ABMs (with some simplifications to create a common denominator across different models) is summarized in Table 3.18. The variety of measures stems from the combination of different size variables segmented by

Table 3.16. List of Mode and TOD Choice Logsums

Impedance	Accessibility from the Given (Residential) Zone to	Token
1	Workplace by all modes for all car-sufficiency groups	Work
2	University by all modes for all car-sufficiency groups	Univ
3	School by all modes for all car-sufficiency groups	Scho
4	Nonmandatory activity location by auto	Auto
5	Nonmandatory activity location by WT	Tran
6	Nonmandatory activity location by NM (walk)	Nonm
7	Nonmandatory activity by all modes, individual travel, zero-car household	Indi_0
8	Nonmandatory activity by all modes, individual travel, cars < workers	Indi_1
9	Nonmandatory activity by all modes, individual travel, cars ≥ workers	Indi_2
10	Nonmandatory activity by all modes, joint travel, zero-car household	Join_0
11	Nonmandatory activity by all modes, joint travel, cars < workers	Join_1
12	Nonmandatory activity by all modes, joint travel, cars ≥ workers	Join_2
13	Escort accessibility, joint travel, zero-car household	Esco_0
14	Escort accessibility, joint travel, cars < workers	Esco_1
15	Escort accessibility, joint travel, cars ≥ workers	Esco_2
16	Workplace by auto modes for all car-sufficiency groups (auto dependency)	Wrkad
17	University by auto modes for all car-sufficiency groups (auto dependency)	Unvad
18	School by auto modes for all car-sufficiency groups (auto dependency)	Schad
19	Workplace by nonauto modes (nonauto dependency)	Wrknad
20	University by nonauto modes (nonauto dependency)	Unvnad
21	School by nonauto modes (nonauto dependency)	Schnad

Table 3.17. Structure of Mode and TOD Choice Logsums

Token	Purpose	Car Sufficiency			Modes Included					Off-Peak Constant
		Zero Cars	Cars < Workers	Cars ≥ to Workers	SOV	HOV	WT	DT	NM	
Work	1 = Work	0.05	0.35	0.6	1	1	1	1	1	–0.9
Univ	2 = Univ	0.05	0.35	0.6	1	1	1	na	1	–0.5
Scho	3 = Scho	0.05	0.35	0.6	1	1	1	na	1	–1.2
Auto	4 = Other	0.05	0.35	0.6	1	na	na	na	na	0.5
Tran	4 = Other	0.05	0.35	0.6	na	na	1	na	na	0.5
Nonm	4 = Other	0.05	0.35	0.6	na	na	na	na	1	0.5
Indi_0	4 = Other	1	na	na	1	na	1	na	1	0.5
Indi_1	4 = Other	na	1	na	1	na	1	na	1	0.5
Indi_2	4 = Other	na	na	1	1	na	1	na	1	0.5
Join_0	4 = Other	1	na	na	na	1	1	na	1	0.5
Join_1	4 = Other	na	1	na	na	1	1	na	1	0.5
Join_2	4 = Other	na	na	1	na	1	1	na	1	0.5
Esco_0	4 = Other	1	na	na	na	1	na	na	1	–0.5
Esco_1	4 = Other	na	1	na	na	1	na	na	1	–0.5
Esco_2	4 = Other	na	na	1	na	1	na	na	1	–0.5
Wrkad	1 = Work	0.05	0.35	0.6	1	1	na	1	na	–0.9
Unvad	2 = Univ	0.05	0.35	0.6	1	1	na	1	na	–0.5
Schad	3 = Scho	0.05	0.35	0.6	1	1	na	1	na	–1.2
Wrknad	1 = Work	0.05	0.35	0.6	na	na	1	na	1	–0.9
Unvnad	2 = Univ	0.05	0.35	0.6	na	na	1	na	1	–0.5
Schnad	3 = Scho	0.05	0.35	0.6	na	na	1	na	1	–1.2

Table 3.18. Zonal Accessibility Measures

Measure	Size Variable		Impedance Measure		Model in Which Applied
	No.	Token	No.	Token	
1	12	Whom1	1	Work	Work from home
2	13	Whom2	1	Work	Work from home
3	14	Whom3	1	Work	Work from home
4	15	Whom4	1	Work	Work from home
5	16	Whom5	1	Work	Work from home
6	17	Shom1	3	Scho	Schooling from home
7	18	Shom2	3	Scho	Schooling from home
8	19	Shom3	2	Univ	Schooling from home
9	11	AllNM	4	Auto	Car ownership
10	11	AllNM	5	Tran	Car ownership
11	11	AllNM	6	Nonm	Car ownership
12	11	AllNM	7	Indi_0	Coordinated daily activity–travel pattern
13	11	AllNM	8	Indi_1	Coordinated daily activity–travel pattern
14	11	AllNM	9	Indi_2	Coordinated daily activity–travel pattern
15	11	AllNM	10	Join_0	Coordinated daily activity–travel pattern
16	11	AllNM	11	Join_1	Coordinated daily activity–travel pattern
17	11	AllNM	12	Join_2	Coordinated daily activity–travel pattern
18	5	Shop	10	Join_0	Joint tour frequency
19	5	Shop	11	Join_1	Joint tour frequency
20	5	Shop	12	Join_2	Joint tour frequency
21	6	Main	10	Join_0	Joint tour frequency
22	6	Main	11	Join_1	Joint tour frequency
23	6	Main	12	Join_2	Joint tour frequency
24	7	Eati	10	Join_0	Joint tour frequency
25	7	Eati	11	Join_1	Joint tour frequency
26	7	Eati	12	Join_2	Joint tour frequency
27	8	Visi	10	Join_0	Joint tour frequency
28	8	Visi	11	Join_1	Joint tour frequency
29	8	Visi	12	Join_2	Joint tour frequency
30	9	Disc	10	Join_0	Joint tour frequency
31	9	Disc	11	Join_1	Joint tour frequency
32	9	Disc	12	Join_2	Joint tour frequency
33	4	Esco	13	Esco_0	Allocated tour frequency
34	4	Esco	14	Esco_1	Allocated tour frequency
35	4	Esco	15	Esco_2	Allocated tour frequency
36	5	Shop	7	Indi_0	Allocated tour frequency
37	5	Shop	8	Indi_1	Allocated tour frequency
38	5	Shop	9	Indi_2	Allocated tour frequency
39	6	Main	7	Indi_0	Allocated tour frequency
40	6	Main	8	Indi_1	Allocated tour frequency
41	6	Main	9	Indi_2	Allocated tour frequency
42	7	Eati	7	Indi_0	Individual tour frequency

(continued on next page)

Table 3.18. Zonal Accessibility Measures (continued)

Measure	Size Variable		Impedance Measure		Model in Which Applied
	No.	Token	No.	Token	
43	7	Eati	8	Indi_1	Individual tour frequency
44	7	Eati	9	Indi_2	Individual tour frequency
45	8	Visi	7	Indi_0	Individual tour frequency
46	8	Visi	8	Indi_1	Individual tour frequency
47	8	Visi	9	Indi_2	Individual tour frequency
48	9	Disc	7	Indi_0	Individual tour frequency
49	9	Disc	8	Indi_1	Individual tour frequency
50	9	Disc	9	Indi_2	Individual tour frequency
51	10	Atwo	7	Indi_0	Individual subtour frequency
52	10	Atwo	9	Indi_2	Individual subtour frequency

the underlying activity type with different impedance measures segmented by trip purpose and person and household type. Models such as car ownership (mobility attributes), work and schooling from home, and coordinated daily activity–travel pattern are very good illustrations for zonal accessibility measures with some components that relate to O-D accessibility measures. Models such as usual workplace and school location are based on O-D accessibility measures.

The 52 zonal accessibility measures combine 19 size variables (numbered and tokenized in Tables 3.13 and 3.14) and 15 impedance measures (numbered and tokenized in Tables 3.16 and 3.17). There are six impedance measures (16–21) that are used only as O-D accessibilities. Multiple examples of impacts of the accessibility measures on different aspects of travel behavior can be found in model estimation reports for the Sacramento, San Diego, and Phoenix ABMs.