Intrazonal or interzonal? Improving intrazonal travel forecast in a four-step travel demand model

Keunhyun Park[1] · Sadegh Sabouri[2] · Torrey Lyons[2] · Guang Tian[3] · Reid Ewing[2]

Abstract

Conventional four-step travel demand models, used by most metropolitan planning organizations (MPOs), state departments of transportation, and local planning agencies, are the basis for long-range transportation planning in the United States. Trip distribution—whether the trip is intrazonal (internal) or interzonal (external)—is one of the essential steps in travel demand forecasting. However, the current intrazonal forecasts based on a gravity model involve flawed assumptions, primarily due to a lack of considerations on differences in zone size, land use, and street network patterns. In this study, we first survey 25 MPOs about how they model intrazonal travel and find the state of the practice to be dominated by the gravity model. Using travel data from 31 diverse regions in the U.S., we develop an approach to enhance the conventional model by including more built environment D variables and by using multilevel logistic regression. The models' predictive capability is confirmed using k-fold cross-validation. The study results provide practical implications for state and local planning and transportation agencies with better accuracy and generalizability.

Keywords Trip distribution · Gravity model · Intrazonal trips · Built environment · Multilevel modeling

Introduction

Metropolitan planning organizations (MPOs) coordinate transportation investments from federal, state, and local sources to ensure that regional transportation plans meet performance criteria such as air quality and congestion management. One of the essential ways

✉ Keunhyun Park
  keunhyun.park@usu.edu

1 Department of Landscape Architecture and Environmental Planning, Utah State University, 4005 Old Main Hill, Logan, UT 84322, USA

2 Department of City and Metropolitan Planning, College of Architecture + Planning, University of Utah, 375 S 1530 E, Salt Lake City, UT 84112, USA

3 Department of Planning and Urban Studies, University of New Orleans, 2000 Lakeshore Drive, New Orleans, LA 70148, USA

MPOs determine how to allocate funds is the forecasting of future travel demands. Forecasts are ordinarily made using what is known as the four-step travel demand model.

Some MPOs are beginning to abandon the traditional four-step travel model in favor of activity/tour-based travel modeling (ABT). As of 2015, in the US, ABT modeling was still in its formative stages and not standard practice (Travel Forecasting Resource 2015). Atlanta Regional Commission, San Diego Association of Government, and New York Metropolitan Transportation Council are some of the pioneering MPOs using this approach. Notwithstanding nearly 30 years of promotion of activity-based modeling (ABM) in the travel modeling literature, we believe that enhancements of the conventional four-step model are still relevant and desirable. As presented in this paper, our survey of MPOs, both big and small, shows that all still use the conventional four-step model, and 20 of 25 MPOs surveyed still use the gravity model for trip distribution. The conventional model and gravity model are still near-universal among small and medium-sized MPOs. As a representative of our local MPO said, when it comes to modeling, MPOs need to be "met where they are at." Meeting MPOs at the current state of the practice and providing an incremental advancement to that practice are the goal of our suggested approach. Our method is meant to be simple and used in connection with the gravity model.

In the simplest terms, the four-step model proceeds from trip generation, to trip distribution, to mode choice, and finally to route assignment. Trip generation tells us the number of trips generated (produced or attracted) in each traffic analysis zone (TAZ). Trip distribution tells us where the trips go, matching trip productions to trip attractions by considering the spatial distribution of productions and attractions as well as the impedance (time or cost) of connections. Mode choice tells us which mode of travel is used for these trips, factoring trip tables to reflect the relative shares of different modes. Route assignment tells us what routes are taken, assigning trips to networks that are specific to each mode. The model's behaviors are estimated based on travel patterns distilled from surveyed household trips. The model is calibrated and validated by comparing the predicted trips in the base year to actual travel survey data. The four-step modeling process is visualized below in Fig. 1.

A major weakness of conventional travel-demand models is that they tend to predict intrazonal trips with poor accuracy. Trips are classified as intrazonal if their origin and destination are contained within the same TAZ. Intrazonal trips are a minor consideration in the four-step travel demand modeling process, despite the fact that they typically amount to 10% or more of all trips in household travel surveys. They are treated like any other zonal
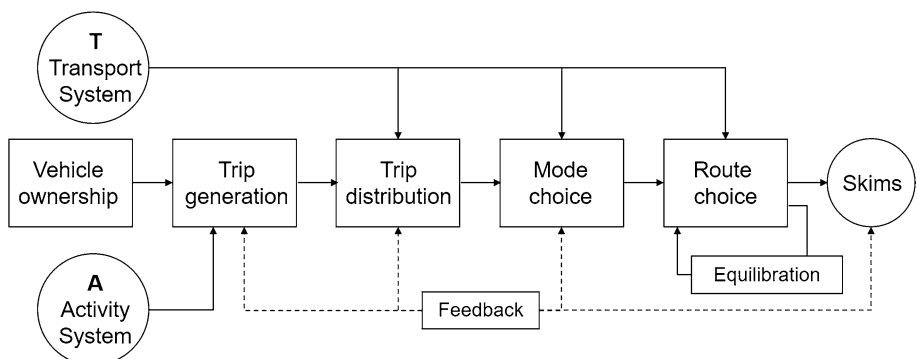

Fig. 1 Four step travel demand model (Adapted from McNally 2007)

interchange in the trip distribution step. Trip productions and attractions are modeled as occurring at a single point in the four-step model, the zone centroid, and the entire local street network on which intrazonal trips occur is reduced to one or more centroid connectors to the external street network. This means that intrazonal trips must be modeled differently than interzonal trips. To quote an unnamed reviewer of this paper, "The limitations of the gravity model are well known and it cannot be expected to deal with trips that travel what is really an unknown average distance. Practitioners have tried to overcome this limitation mostly with heuristic approaches to select an average travel distance for intrazonal trips. The main reason for this is that intrazonal trips are not particularly interesting in themselves but their number affects all the other interzonal trips estimated by destination choice models."

This paper presents a new method for modeling intrazonal trips that addresses the major identified shortcomings of traditional approaches to intrazonal trip modeling in two ways. First, we employ a novel dataset with disaggregated travel survey data coupled with TAZ-specific built environmental measurements. This rich dataset allows us to account for differences in important built environment measures like activity density, street connectivity, and mixed land uses and how they impact intrazonal trip making. The second significant improvement over standard intrazonal modeling efforts is the use of discrete choice modeling. Where traditional methods employ the gravity model which merely measures the attraction potential of a destination less its impedance from an origin on a uniform, aggregated network, discrete choice modeling actually integrates elements of behavior and utility maximization. We use binomial logistic regression, which models the decision of whether to stay within the zone or to leave, as a discrete choice dependent on built environment characteristics within the traffic analysis zone. This method more accurately represents the behavioral aspects inherent in individual travel decision making.

Our paper proceeds as follows. First, we discuss the most common method in use for trip distribution within and across TAZs, namely the gravity model, and known limitations of the method. Then we present results from a survey of 25 MPOs of different sizes across the US, determining their method-in-use for distributing trips. Then we describe our new method, developed as a substitute and improvement upon the commonly used approach. Finally, we present results using our method, validate the models, and conclude its implementation.

Limitations of the gravity model

Various methods have been developed for forecasting intrazonal trips as a component of conventional four-step modeling. However, limitations of the methods raise concerns about the ability of conventional travel demand modeling to adequately account for intrazonal trips. This section considers some methods in common use and their limitations.

One of the most glaring issues with travel demand modeling and the gravity model is that it is done at a relatively aggregate level. Hamilton (1989) was one of the first to point out this issue, stating that as data become more aggregated, the model's assumptions become more and more compromised. Varying sizes of TAZs could lead to differing likelihoods that trips will be intrazonal (Hamilton 1989; Moeckel and Donnelly 2015; Okrah 2016).

Cervero (2006) provides a critique of the conventional approach to four-step modeling that makes a similar point, while also emphasizing the importance of considering localized

information on built environment characteristics. He asserts that in the conventional four-step process, "fine-grained land use mixes, local street connectivity, and pedestrian amenities, do not influence intrazonal trip estimates." This is a general criticism of four-step models, but is particularly apropos to the modeling of intrazonal trips. The failure to consider local land use and street network patterns potentially leads to an underprediction of intrazonal trip rates in densely developed areas.

Research investigating intrazonal travel empirically in relation to characteristics of the local built environment is scant, but some findings are pertinent to this discussion. Modeling intrazonal travel in Gainesville, Florida, Ewing and Tilbury (2002) found that built environment variables (the D variables of development density, land use diversity, street network design, destination accessibility, and distance to transit) rival or sometimes exceed the explanatory power of the gravity formula used to estimate intrazonal trips in a conventional four-step model. This finding has two implications: first, that conventional models are ill-suited to predict intrazonal trips, and second, that sketch planning models that account for these other variables can correct the problem to a degree. One land-use variable, an entropy measure, appeared consistently significant in their models of intrazonal travel for different trip purposes. This variable, derived from Property Appraisers' parcel-level data using GIS, captured the following mix of land uses: pedestrian-oriented retail uses; finance, insurance, and real estate offices; general office buildings; and commercial lodging. Also, highly significant in the authors' models was the presence of a grocery store (for home-based shopping and non-home-based trips) and a public school (for home-based social-recreational and other trips).

Examining intrazonal trip characteristics, Greenwald (2006) found that mode choice for these trips is affected by urban form. The choice of mode, in turn, then affects trip distribution, as non-motorized trips are more likely to stay close to their origin. However, as Greenwald cautions, there is a threshold effect in the ability of the built environment to affect travel behavior; at some point, changes to the economic diversity of a TAZ start showing decreasing impacts on mode choice.

Although research is limited on intrazonal travel measured empirically in relation to D variables, there has been more work on methods for forecasting intrazonal travel as a component of the four-step model. The trip distribution step in the conventional four-step model relies on measuring trip impedance, essentially a measure of the time it will take to travel from a trip origin to a destination. The most common method for capturing impedance is to employ a gravity model, but the standard gravity model disregards local land use and street network patterns. Facile approaches to intrazonal trip distribution are common, including the use of uniform intrazonal trip rates derived from travel surveys as well as simple runs of a gravity model. In the latter case, impedances must be estimated based on intrazonal travel times. Impedances for intrazonal trips are technically zero in the four-step model, since both origins and destinations are located at the same point in space, the zone centroid (Horner and Murray 2002; Bhatta and Larsen 2011). Therefore, intrazonal travel times must be crudely approximated, usually by factoring the size of a TAZ or travel time to adjacent zones.

The traditional four-step model treats intrazonal trips exactly like all trips within the trip distribution step. The basic approach is to use a gravity model to determine the number and proportion of trips being made from a specific origin zone to a specific destination zone. The gravity model works under the assumption that the trips produced at an origin and attracted to a destination are directly proportional to the number of trip productions at the origin and the trip attractions at the destination, and inversely proportional to the

travel time impedance between the original and destination. The standard form of the gravity model is depicted below:

$$T_{ij} = \frac{A_j F_{ij} K_{ij}}{\sum_{allzones} A_x F_{ij} K_{ix}} \times P_i$$

where T_{ij} is trips produced at i and attracted at j; P_i is total trip production at i; A_j is a total trip attraction at j; F_{ij} is the travel impedance between i and j; K_{ij} is the socioeconomic adjustment factor for interchange ij (Anas 1985).

A relatively large body of literature has been published on techniques for estimating intrazonal impedances in the gravity model, in other words for estimating the F_{ij} values in the above formula. Early methods were based on assumptions that vastly simplified the problem, such as one advanced by Batty (1976). In this method, Batty assumed a constant population density over an evenly spread circular zone. His equation for estimating intrazonal travel cost was as follows:

$$C_{ii} = \frac{r_i}{\sqrt{2}}$$

where c_{ii} is travel cost and r_i is the radius of the zone.

Venigalla et al. (1999) suggest a relatively simple method in which intrazonal trip impedance is calculated by merely dividing the trip length and time to the nearest zone centroid in half, sometimes referred to as the nearest neighbor approximation. Others have assumed that intrazonal travel time is two-thirds the time to the nearest neighboring zone, or equal to a set fraction of the average travel time to two or more adjacent zones.

These methods have obvious shortcomings, such as the necessity to make assumptions that zones are circular in shape and demonstrate homogeneous population densities. A marginal improvement to this method was made by Dowling (2005), who divided each zone into 13 concentric squares. The authors then determined mean distance by averaging the distances from the zone centroid to the perimeter of each of the squares. Finally, they used a table of speeds by area type and time of day to compute travel time from the intrazonal distances.

In some regions, the method of calculating intrazonal impedance is based on the zone's total area as well as the average travel speed of the zone. This approach is one of the earliest to be developed (Lamb et al. 1970). The average intrazonal trip distance is approximated by one half of the square root of the zone's area, and the conversion to time in minutes is made with the intrazonal speed in miles per hour and the constant 60 to convert hours into minutes (Martin and Mcguckin 1998).

$$Intrazonal\ time = \frac{0.5 \times \sqrt{(\text{zonal area})} \times 60}{\text{Intrazonal speed (area type)}}$$

Whatever approximation is used, the result flies in the face of findings from our empirical research. Using the gravity model, the larger the zone area is, the greater the impedance is and the smaller the proportion of intrazonal trips becomes. In fact, however, we determined empirically that all else being equal, larger zones capture a higher proportion of total trips generated within the zone. We discuss our research findings on this topic in more detail below.

State-of-the-practice in intrazonal travel modeling

To understand the gap between academic research and practical implementation, we conducted a survey of current intrazonal travel-modeling practices at 25 MPOs in the U.S. in May 2018. We selected MPOs with various population size: three MPOs with a service area population of less than 300,000, nine MPOs between 300,000 and 1 million, and 13 MPOs with more than 1 million population. We focused mostly on large regions because we assume that their MPOs are leaders in using new travel modeling techniques.

We examined the MPOs' travel modeling documents if available and contacted MPO travel modelers to confirm their methods (Capital District Transportation Committee 2010; CMAP & CATS 2014; Community Planning Association of Southwest Idaho 2017; Fresno Council of Governments 2014; H-GAC 2014; Kimley-Horn and Associates, Inc. 2013, 2014; LSA Associates, Inc. 2007, 2011; METROPLAN Orlando 2016; National Capital Region Transportation Planning Board, Metropolitan Washington Council of Governments 2012; PBS&J 2010a, b; RVAMPO 2011; StanCOG 2010; The Association of Monterey Bay Area Governments 2015; The Des Moines Area MPO 2006). The survey findings are presented in Table 1 with their population size, trip distribution model, and intrazonal trip forecast method. The results of our survey show that the four-step travel demand modeling process is still being widely used for regional travel modeling. All surveyed MPOs use the conventional four-step model.

The model that is used most commonly for estimating trip distribution is the gravity model. Out of 25, 20 MPOs use the gravity model for trip distribution—both intrazonal and interzonal. The next most widely used method is the destination choice model, a type of trip distribution or spatial interaction model, typically employing logit. The destination choice model can be thought of as a generalization of the gravity model. In the gravity model, most MPOs use nearest neighbor approximations for calculating the intrazonal travel time, while the number of adjacent zones included in the equation varies from one (the nearest zone; e.g., COMPASS, StanCOG) to four (e.g., ARTS, CHCNGTPO, Memphis, Brunswick).

Basically, the MPOs treat intrazonal trips just like interzonal trips, and the only zone-specific attributes accounted for are trip productions at the zone centroid, trip attractions at the zone centroid, and a crude estimate of intrazonal travel time to create separation between the two—except for CMAP which is not based on the travel time (see Table 1). It is worth mentioning that six of them (FresnoCOG, NCTCOG, SEMCOG, OKI, NJTPA and CMAP) are working on activity-based modeling which is the state-of-the-art in travel modeling. While some of them are almost done with this process, they have not completely switched to ABM yet, as of May 2018.

Our methodology

Data

For 31 regions in the U.S. (Table 2), household travel surveys were collected from MPOs. The surveys were conducted between 2006 and 2012. While conducted by individual regional organizations such as metropolitan planning organizations (MPOs) or State Departments of Transportation, the regional household travel surveys have quite similar structure and questions, akin to U.S. DOT's National Household Travel Survey (NHTS).

Table 1 The summary of MPOs methods for calculating trip distribution and intrazonal trips (as of March 2018; sorted by population size)

MPO name	Major city	Population (2010)	Trip distribution model	Method for calculating intrazonal trips
CMAP	Chicago, IL	8,444,660	Gravity with intervening opportunities	Both inter-zonal and intra-zonal trips are modeled together based on zone size, trip cost, and available destinations, and then separated based on impedance (time, cost, etc.)
NJTPA	Newark, NJ	6,579,801	Gravity	The intrazonal time was calculated using half of the sum of time from two closest "nonzero" zones, and then multiplied it by 0.60
NCTCOG	Arlington, TX	6,417,630	Gravity	Nearest neighbor rule (0.5 of three zones)
H-GAC	Houston, TX	5,892,002	Atomistic model (a gravity-analogy-based model)	…by dividing existing zones into atoms a more realistic interchange of intrazonal trips and short (less than 5 min) trips among adjacent zones is defined
NCRTPB	Washington, DC	5,068,540	Gravity	The intra-zonal times have been set to 85% of the minimum inter-zonal time
SEMCOG	Detroit, MI	4,703,593	Destination choice model	Intra-zonal travel time is calculated based on 4 nearest neighbor zones
Boston Region MPO	Boston, MA	3,159,512	Gravity	Nearest neighbor rule (0.5 of 3 zones)
EWGCOG	St. Louis, MO	2,571,253	Gravity/destination choice model	For home-based work: gravity model/for other types: destination choice model
OKI	Cincinnati, OH	1,981,230	Gravity	Half of the average travel time to the nearest three zones
MARC	Kansas City, MO	1,895,535	Destination choice model	The nearest neighbor rule was used to estimate the intrazonal travel times
METROPLAN Orlando	Orlando, FL	1,837,385	Gravity	The nearest neighbor rule with terminal time as the constraining variable.
WFRC	Salt Lake City, UT	1,561,348	Gravity	Intrazonal travel time as a function of the area of the zone and the average travel speed
Memphis Urban Area MPO	Memphis, TN	1,077,697	Destination choice model	The intrazonal travel times are computed by taking half the average travel time to the four closest neighboring zones

Table 1 (continued)

MPO name	Major city	Population (2010)	Trip distribution model	Method for calculating intrazonal trips
FresnoCOG	Fresno, CA	930,885	Gravity	100% and 33.3% the average time to the nearest adjacent TAZ for urban and rural areas, respectively
CDTC	Albany, NY	823,239	Gravity	A travel time of 6 min is assumed for intrazonal trips (trips within the same zone)
AMBAG	Marina, CA	732,667	Gravity	Intra-zonal travel times were computed based on the average time to the nearest 3 zones
COMPASS	Meridian, ID	550,359	Gravity	Travel times: 50% time to the nearest zone
Stanislaus COG	Modesto, CA	514,453	Gravity	Intrazonal travel times are estimated based on 50% of the travel time to the nearest adjacent zone
Des Moines Area MPO	Urbandale, IA	475,855	Gravity	Three neighbor zones for the calculation of average travel time were chosen and a final factor, 0.5, was applied to the end result
ARTS	Augusta, GA	440,134	Gravity	Intrazonal times were created by the travel purpose + matrix function using half of the average travel time to the nearest four TAZ's
CHCNGTPO	Chattanooga, TN	436,669	Destination choice model	The intrazonal travel time is calculated as half the average travel time to the four closest neighboring zones
North Front Range MPO	Fort Collins, CO	433,178	Gravity	Intrazonal travel time is calculated as a function of the travel time required to reach the closest adjoining zone
Lincoln MPO	Lincoln, NE	285,407	Gravity	Intrazonal travel time has been calculated by multiplying the distance to the single nearest neighbor by 75%
RVAMPO	Roanoke, VA	227,507	Gravity	Two adjacent zones are used to compute the intra-zonal travel time during the trip distributions

Table 1 (continued)

MPO name	Major city	Population (2010)	Trip distribution model	Method for calculating intrazonal trips
Brunswick MPO	Brunswick, GA	79,626	Gravity	Intrazonal times were created by the travel purpose + matrix function using half of the average travel time to the nearest four TAZ's

To gather comprehensive data on travel and transportation patterns, the survey data consistently includes, but is not limited to, household demographic information, vehicle information, and data about one-way trips taken during a designated 24-hour period on a weekday, including travel time, mode of transportation, and purpose of trip information. The survey data have exact XY coordinates so we could geocode the precise locations of households and the precise origins and destinations of trips. The regional survey data were acquired from individual MPOs or state DOTs with confidentiality agreements. The pooled data set consists of 843,287 trips produced by 89,768 households within 25,469 traffic analysis zones (TAZs) in 31 regions.

Table 2 Percentage of Intrazonal travel by trip purpose from travel surveys

	HBW	HBShp	HBOth	NHBW	NHBNW
Albany, NY	3.2	8.5	21.9	9.5	15.0
Atlanta, GA	3.4	9.8	17.4	10.6	15.9
Boston, MA	2.9	7.3	15.3	10.6	12.6
Burlington, NC	4.5	4.4	13.1	10.3	11.0
Dallas, TX	2.3	6.4	15.9	7.7	11.6
Denver, CO	2.8	4.6	11.5	8.0	11.6
Detroit, MI	2.0	8.9	9.6	6.2	9.9
Eugene, OR	1.3	3.2	7.4	7.1	8.2
Greensboro, NC	1.9	5.0	15.1	8.7	12.0
Hampton Roads–Norfolk, VA	2.8	7.8	19.4	11.4	14.6
Houston, TX	3.1	8.4	14.7	6.5	11.8
Indianapolis, IN	2.5	3.8	11.0	7.4	12.7
Kansas City, MO	4.8	11.0	16.8	9.8	15.1
Madison, WI	5.9	4.8	13.8	12.6	13.0
Miami, FL	1.7	5.0	13.4	6.7	10.9
Minneapolis–St. Paul, MN-WI	3.0	5.2	9.0	7.8	12.3
Orlando, FL	2.1	6.2	21.8	9.5	12.5
Palm Beach, FL	2.6	8.0	26.0	9.3	11.9
Phoenix, AZ	2.8	10.5	20.2	9.3	13.5
Portland, OR	3.3	7.8	14.9	16.7	17.1
Provo-Orem, UT	3.3	4.6	19.1	6.6	10.5
Richmond, VA	2.2	5.6	17.9	9.9	11.1
Rochester, NY	2.8	5.7	9.3	5.8	12.2
Salem, OR	2.4	0.9	8.7	6.7	9.6
Salt Lake City, UT	2.7	4.2	15.0	6.2	10.6
San Antonio, TX	2.8	5.5	10.9	6.2	10.7
Seattle, WA	1.5	7.0	11.1	10.5	10.0
Springfield, MA	4.0	8.2	15.2	16.3	17.5
Syracuse, NY	1.4	5.9	15.7	7.6	10.7
Tampa, FL	4.2	8.3	21.5	8.3	12.7
Winston-Salem, NC	3.2	4.5	14.0	5.7	11.1
Total	2.9	6.9	14.4	9.2	12.8

The 843,287 trips were classified as either intrazonal (produced and attracted within the same TAZ) or interzonal trips (produced in one TAZ and attracted to another). On average, intrazonal trips account for 10.7% of total trips. This is a significant share of total trips. We computed intrazonal trip shares by trip purpose from the regional household travel surveys. The result is presented in Table 2. The shares vary from region to region. For example, intrazonal home-based work trips make up 2.9% of all home-based work trips on average, ranging from 1.3% in Eugene to 5.9% in Madison. Intrazonal home-based other trips (excluding work and and shopping-related ones) make up 14.4% of all home-based other trips on average, ranging from 7.4% in Eugene to 26.0% in Palm Beach. This large variance may reflect differences in zone size, land use and street network patterns, or even socio-demographics. The need to model intrazonal travel, in terms of these variables, is evident. In this paper, we show results from modeling intrazonal travel in relation to the D variables for the 31 regions, based on the regional household travel surveys.

Also, we collected land use data at the parcel level with detailed land use classifications, so we could study land use intensity and mix down to the parcel level for the same year as the household travel survey. We also gathered GIS data layers for streets, population and employment for TAZs, and travel times between zones by different modes, again for the same years as the household travel survey. Built environmental variables were computed for each TAZ and assigned to households within the TAZ.

Variables

In this study, the D variables of the built environment were measured and used to predict the intrazonal travel. The measurement of the D variables and their expected effect on travel behavior are summarized in Table 3. Some dimensions capture closely related qualities (e.g., diversity and destination accessibility). Still, it is a useful framework used to organize the empirical literature and provide order-of-magnitude insights (Ewing and Cervero 2010). The dependent and independent variables used in this study are defined in Table 4. Sample sizes and descriptive statistics are also provided.

For home-based trip (home-based-work, home-based-shopping, and home-based-other) models, the D variables of the TAZ where the home is located were used to characterize the built environment of the TAZ. For the non-home-based-work trip model, the D variables of the TAZ where the workplace is located were used to characterize the built environment of the TAZ. For the non-home-based-non-work trip model, the D variables of the TAZ where the trip origin is located were used to characterize the built environment of the TAZ.

Analysis methods

We treated intrazonal/interzonal travel as a binary choice, and hence modeled it with multilevel binomial logistic regression. We modeled intrazonal travel for the 31 regions. A binomial logistic regression predicts the probability that an observation falls into one of two categories of a dichotomous dependent variable (intrazonal or interzonal travel, in this case) based on multiple independent variables (in our case, the TAZ-level D variables).

A three-level model was required to represent the nested nature of the dataset, with multiple trips nested within TAZs and TAZs nested within regions. Multilevel modeling accounts for dependence among observations. All trips within a given TAZ share TAZ characteristics and all TAZs within a given region share regional characteristics. This

Table 3 The D variables (Ewing et al. 2015)

D variable	Measurement
Density	Density is always measured as the variable of interest per unit of area. The area can be gross or net, and the variable of interest can be population, dwelling units, employment, or building floor area. Population and employment are sometimes summed to compute an overall activity density per areal unit
Diversity	Diversity measures pertain to the number of different land uses in a given area and the degree to which they are balanced in land area, floor area, or employment. Entropy measures of diversity, wherein low values indicate single-use environments and higher values more varied land uses, are widely used in travel studies. Jobs-to-housing or jobs-to-population ratios are also used. The latter measures can be more useful for a prediction purpose because specific land use compositions are difficult to predict
Design	Design measures include average block size, proportion of four-way intersections, and number of intersections per square mile. Design is also occasionally measured as sidewalk coverage (share of block faces with sidewalks); average building setbacks; average street widths; or numbers of pedestrian crossings, street trees, or other physical variables that differentiate pedestrian-oriented environments from auto-oriented ones
Destination accessibility	Destination accessibility measures ease of access to trip attractions. It may be regional or local (Handy 1993). In some studies, regional accessibility is simply distance to the central business district. In others, it is the number of jobs or other attractions reachable within a given travel time, which tends to be highest at central locations and lowest at peripheral ones. The gravity model of trip attraction measures destination accessibility. Local accessibility is a different animal. Handy (1993) defines local accessibility as distance from home to the closest store
Distance to transit	Distance to transit is usually measured as an average of the shortest street routes from the residences or workplaces to the nearest rail station or bus stop. Alternatively, it may be measured as transit route density, distance between transit stops, or the number of stations per unit area. In this literature, frequency and quality of transit service are overlooked

dependence violates the independence assumption of standard regression. Standard errors of regression coefficients will consequently be underestimated. Moreover, coefficient estimates will be inefficient. Multilevel models overcome these limitations, producing more accurate coefficient and standard error estimates (Raudenbush and Bryk 2002). The three-level model used in this study partitions variance among the trip level (Level 1), the TAZ level (Level 2), and the regional level (Level 3) and uses level-specific variables to explain the variance at each level.

A multi-level model can be interpreted the same way as a single-level model; values of the independent variables are substituted for the variables in equations, multiplied by coefficients, and summed to get the log odds. Then, by exponentiating the log-odds, we can compute the odds of an intrazonal trip and the probability of an intrazonal trip, which is equal to [odds of intrazonal trips/(1 + odds of intrazonal trips)].

The final models were chosen based on three considerations—(1) whether the sign of a coefficient is expected or not (for example, total population of a TAZ is expected to have a positive relationship with the share of intrazonal trips. If not, we drop that variable), (2) statistical significance of the explanatory variable, and (3) the overall model fit based on the pseudo-R^2.

Table 4 Descriptive statistics for our variables

Variable	Description	N	Mean	Median	SD
Outcome variable (level 1: trip)					
Intrazonal	Trip remaining internal to TAZ (1 = intrazonal, 0 = interzonal)	843,287	0.11	0.00	0.31
Trip purpose	Trip purpose: home-based-work (HBW), home-based-shopping (HBShp), home-based-other (HBOth), non-home-based-work (NHBW), non-home-based-non-work (NHBNW)	–	–	–	–
Explanatory variables (level 2: TAZ)					
totpop	Total population within TAZ	25,396	1,832.76	1406.00	1,664.44
totemp	Total employment within TAZ	25,396	611.60	283.00	1,065.82
area	Gross land area of TAZ in square miles	25,396	1.82	0.50	10.57
actden	Activity density within TAZ (pop + emp per square mile in 1000 s)	25,396	7.05	4.29	21.14
jobpop[a]	Job-population balance within TAZ	25,396	0.55	0.57	0.28
intden	Intersection density within TAZ	25,396	98.39	85.47	80.52
pct4wy	Percentage of 4-way intersections within TAZ	25,396	25.80	20.22	20.10
pctemp10a	Percentage of regional employment within 10 min by car	25,396	6.93	3.06	11.01
pctemp20a	Percentage of regional employment within 20 min by car	25,396	27.4	18.9	25.2
pctemp30a	Percentage of regional employment within 30 min by car	25,396	49.3	50.5	30.2
pctemp30t	Percentage of regional employment within 30 min by transit	25,396	16.81	7.84	21.26

[a]JOBPOP = 1 – [ABS(employment – 0.2 × population)/(employment + 0.2 × population)], where ABS is absolute value of expression in parentheses (Ewing et al. 2015). The value 0.2, representing a balance of employment and population, was found through trial and error to maximize the explanatory power of the variable

Table 5 Home-based-work models

	Coef.	SE	z value	p value	Odds ratio
intercept	−4.683	0.112	−41.706	< 0.001	0.007
totemp	0.0003	0.00003	10.430	< 0.001	1.0003
area	0.009	0.003	3.111	0.002	1.010
pctemp20a	−0.007	0.002	−3.290	0.001	0.993

Sample size: level 1—121,200; level 2—19,656; level 3—31

Log likelihood: −13,033; AIC: 26,078; pseudo-R^2: 0.01

Table 6 Home-based-shopping models

	Coef.	SE	z value	p value	Odds ratio
intercept	−4.426	0.121	−36.532	< 0.001	0.012
totemp	0.0003	0.00002	14.841	< 0.001	1.0003
totpop	0.0001	0.00001	3.605	< 0.001	1.0001
area	0.004	0.002	1.994	0.046	1.004
jobpop	0.754	0.104	7.276	< 0.001	2.125
intden	0.001	0.000	2.961	0.003	1.001
pct4way	0.007	0.002	4.103	< 0.001	1.007
pctemp20a	−0.005	0.002	−2.920	0.004	0.995

Sample size: level 1—134,454; level 2—20,301; level 3—31

Log likelihood: −27,701; AIC: 55,422; pseudo-R^2: 0.02

Model validation

To test how well the intrazonal models are able to predict intrazonal travel, we evaluated the predictive performance of our five models—one for each trip purpose—by running k-fold cross-validation on our datasets (Fielding and Bell 1997; Hair et al. 1998). Using the same data to estimate parameters and to test predictive accuracy may overestimate model validity. In k-fold cross-validation, the data are divided into k equal partitions. One partition is withheld, and the model is fitted with the remaining data. As Borra and Di Ciaccio (2010) suggest, data were randomly divided into tenfolds: 90% of the data (training data) used for model fitting and 10% of the data withheld for model validation in each iteration.

The receiver operating characteristic (ROC) curves and the areas under ROC curves (AUC) are appropriate measures to evaluate prediction capability of logistic regression models (Greiner et al. 2000; Hanley and McNeil 1982; Meng 2014; Zweig and Campbell 1993). For the ROC curves, the rate of true-positives is plotted on the vertical axis and the rate of false-positives is plotted on the horizontal axis. Then the ROC statistics, AUC, provide the predictive accuracy of the logistic models, with values from 0.5 (no predictive power) to 1.0 (perfect prediction). In this study, the ROC curves were first used to visualize prediction capability of our models using only the left-out partition that was not used in model fitting. Predictive accuracy is then assessed by calculating the areas under ROC curves (AUC). This procedure is repeated for each of the k partitions, and the AUC values are averaged to obtain the mean AUC value.

In addition to the k-fold validation, we also validated our models against a conventional practice—the gravity model. How much more accurate is our model than the gravity model? Instead of modeling it, there are a few regions using a constant value, a region-wide proportion

Table 7 Home-based-other models

	Coef.	SE	z value	p value	Odds ratio
intercept	−2.744	0.088	−31.297	< 0.001	0.064
totemp	0.0001	0.00001	7.397	< 0.001	1.0001
totpop	0.0001	0.00001	10.689	< 0.001	1.0001
area	0.005	0.001	3.285	0.001	1.005
jobpop	0.333	0.059	5.673	< 0.001	1.395
intden	0.0004	0.0002	2.015	0.044	1.0004
pctemp10a	−0.006	0.002	−2.716	0.007	0.994

Sample size: level 1—256,004; level 2—22,273; level 3—31

Log likelihood: −92,914; AIC: 185,845; pseudo-R^2: 0.01

of intrazonal trips by trip purpose, to estimate intrazonal trip distribution. Is our model better than that simplest approach?

To prove the validity of our model, we compare our model with two other models—a gravity model and a constant model (using a region-wide average proportion of intrazonal trips by trip purpose) using data from two regional MPOs—Wasatch Front Regional Council (WFRC) and Mountainland Association of Governments (MAG). Two regions were selected because we could obtain intrazonal proportions by TAZ from their gravity models. Thus, our unit of analysis is the TAZ. The modeled values were compared against the actual proportion of intrazonal trips by trip purpose by TAZ from the 2012 Utah Household Travel Survey.

The problem with this approach is that many TAZs have no or only a few trips. This raises sampling error issues, meaning that the small number of trips in the survey cannot represent all trips occurring in that TAZ. For example, if a TAZ has only one trip (which is internal) from the survey, it gets 100% intrazonal trip probability. Thus, we tried different values in the minimum number of trips in a TAZ to minimize the sampling error and determined 20 as a threshold for model validation purposes.

Root mean square error (RMSE) is an appropriate measure of model prediction quality between two continuous variables (in this case, the proportion of intrazonal trips in the survey vs. a model). RMSE is a frequently used measure of the differences between values predicted by a model and the values actually observed. RMSE is a measure of accuracy, to compare forecasting errors of different models for a particular dataset. The smaller the RMSE, the more accurate the model (and the better the predictive power).

Results

Intrazonal trip share models

Tables 5, 6, 7, 8 and 9 show the results of multilevel binomial logistic regressions for intrazonal trips by trip purpose. The intercept in the tables is the constant of the models, which is the expected mean value of log-odds of Y (intrazonal trip share) when all independent variables are zero. The coefficients are log-odds of a trip being intrazonal not interzonal for a one-unit change in the specific independent variable. By exponentiating the log-odds, we can compute the odds of intrazonal trip and the probability of intrazonal trip, which is equal to [odds of intrazonal trips/(1 + odds of intrazonal trips)].

Table 8 Non-home-based-work models

	Coef.	SE.	z value	p value	Odds ratio
intercept	− 2.603	0.084	− 31.053	< 0.001	0.074
totemp	0.00005	0.00002	2.672	0.008	1.00005
actden	0.003	0.001	2.564	0.010	1.003
pct4way	0.003	0.001	3.003	0.003	1.003
pctemp30a	− 0.003	0.001	− 2.717	0.007	0.997

Sample size: level 1—86,763; level 2—16,200; level 3—31

Log likelihood: − 25,060; AIC: 50,136; pseudo-R^2: 0.002

Table 9 Non-home-based-non-work models

	Coef.	SE	z value	p value	Odds ratio
intercept	− 2.096	0.040	− 52.431	< 0.001	0.123
totemp	0.00004	0.00001	3.848	< 0.001	1.00004
totpop	0.00001	0.00001	2.299	0.021	1.00001
area	0.004	0.001	4.137	< 0.001	1.004
pctemp10a	− 0.004	0.001	− 2.457	0.014	0.996
pctemp30t	− 0.002	0.001	− 3.196	0.001	0.998

Sample size: level 1—183,066; level 2—20,156; level 3—31

Log likelihood: − 67,680; AIC: 135,375; pseudo-R^2: 0.002

Different D variables are shown to be significant predictors of intrazonal trips for different trip purposes. All relationships are as expected. To summarize, total employment (demographic variable) is positively associated with the share of intrazonal trips for all five trip purposes. Total population (demographic variable) is positively associated with the share of intrazonal trips for home-based-shopping, home-based-other, and non-home-based-non-work purposes. Area size has a positive association with the intrazonal trip likelihood for home-based-work, home-based-shopping, home-based-other, and non-home-based-non-work trips. Activity density is only included in non-home-based-work model. A land use diversity variable, job-population balance, is positively related to the share of intrazonal trips for all home-related trip purposes but home-based-work trips. Destination accessibility—the percentage of jobs available within 10-min, 20-min, or 30-min by car or 30-min by transit—is negatively associated with the share of intrazonal trips for all five trip purposes. This implies that the more jobs immediately outside of the given TAZ, the more likely a trip crosses the zone boundary for specific trip types. A measure of street network design—the percentage of four-way intersections—is positively associated with intrazonal trip likelihood only for home-based-shopping and non-home-based-work trips. Lastly, regional variables—total population, total employment, and population density—are not statistically significant in any models, and so were dropped.

Model validation results

After fitting the models with the full data, we assessed the predictive power of the five intrazonal models using tenfold cross-validation. Travel data were randomly split into ten
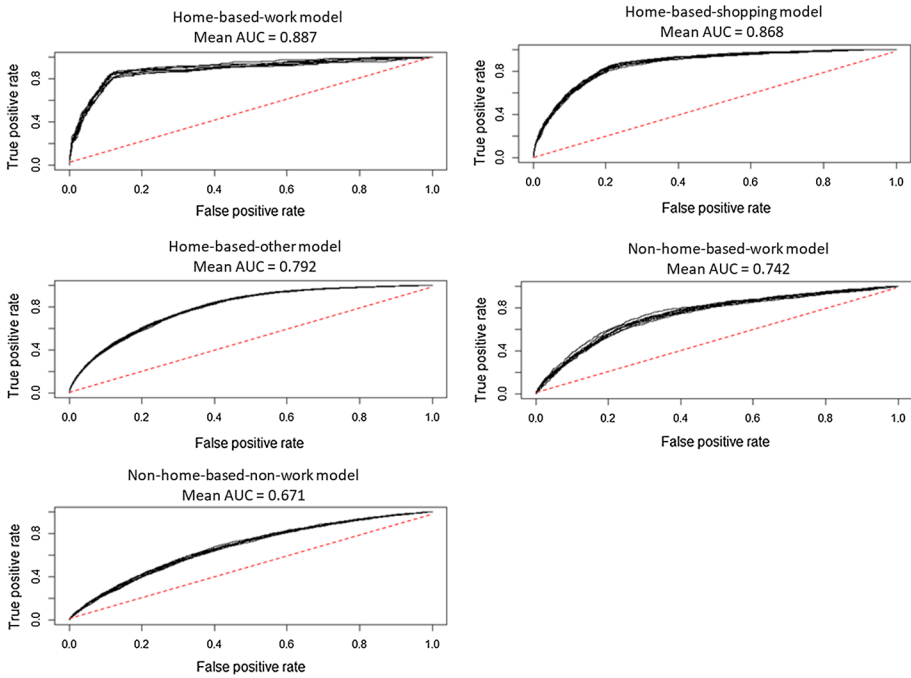
Fig. 2 Model validation (1): receiver operating characteristic (ROC) curves and the area under the ROC (AUC) statistics for measuring predictive power of the models

equal-sized groups. The validation data set, 10% of the data, was used to validate the model which was fitted using the other 90% of the data through multilevel logistic regression.

As a result of the tenfold cross-validation, we obtained average AUCs by trip purpose. The average AUCs range from 0.671 for the non-home-based-non-work model to 0.887 for the home-based-work model (Fig. 2). The AUC provides the predictive accuracy of the logistic models, with values from 0.5 (no predictive power) to 1.0 (perfect prediction). Following Swets (1988) and Manel et al. (2001), models with an AUC value ranging between 0.7 and 0.9 are treated as 'useful applications' and those with values greater than 0.9 as being of 'high accuracy.' Thus, most models can be considered useful applications. The non-home-based-non-work is lower than the threshold of 0.7, implying a need for a different, more advanced modeling approach such as generalized additive model (Hastie and Tibshirani 1990).

Table 10 Model validation (2): root mean square error (RMSE)

	HBW	HBShp	HBOth	NHBW	NHBNW
WFRC gravity model	0.076	0.101	0.199	0.055	0.112
Constant model	0.047	0.082	0.170	0.064	0.090
Our model	0.007	0.010	0.017	0.020	0.029

The smaller the RMSE, the more accurate the model and the better the predictive power

In addition to the k-fold validation, we validated our models against a conventional practice—the gravity model. We compare our model with two other models—a gravity model and a constant model (using a region-wide average proportion of intrazonal trips by trip purpose) using travel survey data from the 2012 Utah Household Travel Survey.

Table 10 shows that our model outperforms other models for all five trip purposes. The error rate of gravity model is significantly higher than that of our model (more than tenfold in most models), and even higher than the constant model using an identical region-wide value of intrazonal proportion for each trip purpose.

Conclusions

Conventional four-step models, used by virtually all metropolitan planning organizations (MPOs), state departments of transportation, and local transportation planning agencies to forecast future travel patterns, are the basis for long-range transportation planning in the United States. Trip distribution is one of the critical steps in travel demand forecasting. In the model structure, it consists of two categories—intrazonal trips and interzonal trips. As Bhatta and Larsen (2011) explained, intrazonal trips cannot be ignored, due to the impact they have on important aspects of transportation, such as congestion and pollution. For modeling intrazonal trips, there are two important components: (1) predicting whether a trip will be intrazonal and (2) determining the impedance of intrazonal trips. Little attention has been given to the former component, and in this study, we developed an approach to enhance the conventional gravity model for predicting intrazonal trips by including more built environment D variables and using a more robust modeling method.

In the first step, we surveyed 25 MPOs about how they model intrazonal travel. The finding shows the dominance of the gravity model with nearest neighbor assumptions, while a few regions are currently in the process of shifting to activity-based modeling. However, the current model involves validation errors, probably due to differences in zone size, land use, and street network patterns, none of which should be overlooked. The need to model intrazonal travel in terms of the built environment variables is evident. Thus, by using multilevel binomial logistic regression models and regional household travel survey data from 31 U.S. regions, we proved that different D variables are significant predictors of intrazonal trips for different trip purposes. Model validation results confirm that our models are useful for prediction purposes.

There is broad interest in the planning and policy communities in developing accurate tools to predict the consequences of land use and transportation strategies on travel demands. State, regional, and local organizations such as state departments of transportation and MPOs, public health organizations, transit agencies, and city and county planning commissions are also eager to have a reliable means of evaluating growth scenarios and planning alternatives. To this end, the results of this study could be used in travel demand modeling practice, especially in the hundreds of medium- and small-sized MPOs. It is worthwhile to note that two regional MPOs, Wasatch Front Regional Council (WFRC) and Mountainland Association of Governments (MAG), are incorporating our models into their four-step models in the transportation modeling software, Cube, to improve the accuracy of travel forecasts. Because we estimated models based on 31 region database, the models have external validity, and are generalizable for future changes on land use and transport toward more compact, mixed-use, and transit-supportive developments.

A limitation to this study is the fact that we are proposing a novel approach to the less than novel practice of four-step travel demand modeling. A travel modeler must understand the limitations of our modeling approach—advanced intrazonal trip models with built environment variables. For one thing, there can be a trade-off between a unified destination choice model and a gravity model separating intrazonal and interzonal trips. This separation of trips clearly affects the accuracy of a calibrated distribution model. However, we contend that intrazonal travel is qualitatively different than interzonal travel, even to a nearby zone, because intrazonal travel time is only crudely approximated by the gravity model. Furthermore, the decision to stay within a zone is highly affected by the built environment characteristics of the zone. Our use of D variables to model intrazonal trip shares is essentially a refinement of the intrazonal travel time estimate for vehicle trips and an add on for intrazonal travel by non-motorized modes. In a subsequent study, one could add D variables as zonal attributes in a more sophisticated utility function for a destination choice model and compare the results.

In addition, as we described in the introduction, the state-of-the-art is activity-based modeling (ABM). Many of the shortcomings of the trip-based approach to travel modeling such as the inability to consider the potential sequencing of trips, are rectified by the application of ABM. However, while ABM is the state-of-the-art in travel demand modeling, trip-based modeling is still the state-of-the-practice for small to medium-sized MPOs and many large MPOs surveyed. While we acknowledge that our small sample of large MPOs seems to have some bias toward those which are still using the four-step model, our survey indicates that some of the largest MPOs with the highest capacities continue to use the four-step model. An incremental improvement to the tool that is currently the most ubiquitous among travel modelers is a valuable contribution to practice.

Authors' Contributions KP: data analysis and validation, manuscript writing and editing. SS: MPO survey, literature review, manuscript writing. TL: literature search and review, manuscript writing and editing. GT: data collection and analysis, manuscript editing. RE: research design, manuscript editing.

Compliance with Ethical Standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Anas, A.: The combined equilibrium of travel networks and residential location markets. Reg Sci Urban Econ **15**(1), 1–21 (1985)

Batty, M.: Urban modeling: algorithms, calibrations, predictions. Cambridge University Press, London (1976)

Bhatta, B.P., Larsen, O.I.: Are intrazonal trips ignorable? Transp. Policy **18**(1), 13–22 (2011)

Borra, S., Di Ciaccio, A.: Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. Comput. Stat. Data Anal. **54**(12), 2976–2989 (2010)

Capital District Transportation Committee: The CDTC STEP MODEL, appendix E: documentation of CDTC step model and planning assumptions (2010)

Cervero, R.: Alternative approaches to modeling the travel-demand impacts of smart growth. J Am Plan Assoc **72**(3), 285–295 (2006)

CMAP & CATS: Go to 2040: travel model documentation for Chicago Metropolitan Agency for Planning (2014)

Community Planning Association of Southwest Idaho: Regional travel demand forecast model calibration and validation report for Ada and Canyon County in Idaho. Report number 06-2017

Dowling, R.G.: Predicting air quality effects of traffic-flow improvements: final report and user's guide. Transportation Research Board, Washington (2005)

Ewing, R., Tilbury, K.: Sketch models for integrated transportation and land use planning. Metropolitan Transportation Planning Organization, Gainesville (2002)

Ewing, R., Cervero, R.: Travel and the built environment: a meta-analysis. J Am Plan Assoc **76**(3), 265–294 (2010)

Ewing, R., Tian, G., Goates, J.P., Zhang, M., Greenwald, M.J., Joyce, A., et al.: Varying influences of the built environment on household travel in 15 diverse regions of the United States. Urban Stud **52**(13), 2330–2348 (2015)

Fielding, A.H., Bell, J.F.: A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. **24**(1), 38–49 (1997)

Fresno Council of Governments: Fehr & peers transportation consultant: travel demand model: model description & validation report for Fresno Council of Governments (2014)

Greenwald, M.J.: The relationship between land use and intrazonal trip making behaviors: evidence and implications. Transp Res Part D Transp Environ **11**(6), 432–446 (2006)

Greiner, M., Pfeiffer, D., Smith, R.D.: Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev Vet Med **45**(1–2), 23–41 (2000)

H-GAC: 2012 model validation and documentation report, Houston-Galveston Area Council regional travel models, appendix 4 (2014)

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: Multivariate data analysis (vol. 5, no. 3), pp. 207–219. Prentice Hall, Upper Saddle River (1998)

Hamilton, B.W.: Wasteful commuting again. J. Polit. Econ. **97**(6), 1497–1504 (1989)

Handy, S.L.: Regional versus local accessibility: implications for non-work travel. Transp. Res. Rec. **1400**, 58–66 (1993)

Hanley, J.A., McNeil, B.J.: The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology **143**(1), 29–36 (1982)

Hastie, T.J., Tibshirani, R.J.: Generalized additive models. Chapman & Hall/CRC, Boca Raton (1990)

Horner, M.W., Murray, A.T.: Excess commuting and the modifiable areal unit problem. Urban Stud **39**(1), 131–139 (2002)

Kimley-Horn and Associates, Inc.: Chattanooga travel demand model development highway and transit assignment, technical memorandum #6, final version 3.0 (2013)

Kimley-Horn and Associates, Inc., Cambridge Systematics, Inc. and HNTB: Direction 2040: long range transportation plan: travel demand model documentation for Memphis urban area, appendix G (2014)

Lamb, G. M.: Introduction to transportation planning: 4-trip distribution and model split. Traffic engineering and control, April, pp. 628–630 (1970)

LSA Associates, Inc.: North front range regional travel model: model process, parameters, and assumptions (2007)

LSA Associates, Inc.: Travel demand model: model development and validation report for Lincoln Metropolitan Planning Organization (2011)

McNally, M.G.: The four-step model. In: Hensher, D.A., Button, K.J. (eds.) Handbook of transport modelling, 2nd edn., pp. 35–53. Emerald Group Publishing Limited, Bingley (2007)

Manel, S., Ceri Williams, H., Ormerod, S.J.: Evaluating presence–absence models in ecology: the need to account for prevalence. J. Appl. Ecol. **38**, 921–931 (2001)

Martin, W.A., McGuckin, N.A.: Travel estimation techniques for urban planning, vol. 365. National Academy Press, Washington (1998)

Meng, Q.: Modeling and prediction of natural gas fracking pad landscapes in the Marcellus Shale region, USA. Landsc Urban Plan **121**, 109–116 (2014)

METROPLAN Orlando: 2040 long range transportation plan, technical report 8: model development and application guidelines, final adopted plan (2016)

Moeckel, R., Donnelly, R.: Gradual rasterization: redefining spatial resolution in transport modelling. Environ. Plan. **42**(5), 888–903 (2015)

National Capital Region Transportation Planning Board, Metropolitan Washington Council of Governments: Calibration report for the TPB travel forecasting model, version 2.3, on the 3,722-Zone Area System (2012)

Okrah, M.B.: Handling non-motorized trips in travel demand models. In: Wulfhorst, G., Klug, S. (eds.) Sustainable mobility in metropolitan regions, pp. 155–171. Springer, Wiesbaden (2016)

PBS&J: ARTS 2035 long range transportation plan, appendix B: the travel demand model for the ARTS MPO (2010a)

PBS&J: The travel demand model for the Brunswick MPO (2010b)

Raudenbush, S.W., Bryk, A.S.: Hierarchical linear models: applications and data analysis methods, 2nd edn. Sage Publications, Thousand Oaks (2002)

RVAMPO: Constrained long-range transportation plan 2035 (2011). http://rvarc.org/chapters-4-and-5-technical-process-clrtp-2035/

StanCOG: The StanCOG transportation model program: documentation and validation report, version 1 (2010). http://www.stancog.org/trans-model.shtm

Swets, J.A.: Measuring the accuracy of diagnostic systems. Science **240**, 1285–1293 (1988)

The Association of Monterey Bay Area Governments: Association of Monterey bay area governments regional travel demand model technical report (2015)

The Des Moines Area MPO: Travel demand model: documentation & user guide (2006)

Travel Forecasting Resource: Category: activity-based models. http://tfresource.org/Category:Activity-based_models. Accessed 2 Apr 2019

Venigalla, M.M., Chatterjee, A., Bronzini, M.S.: A specialized equilibrium assignment algorithm for air quality modeling. Transp Res Part D Transp Environ **4**(1), 29–44 (1999)

Zweig, M.H., Campbell, G.: Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin. Chem. **39**(4), 561–577 (1993)

Keunhyun (Keun) Park is an assistant professor in the Department of Landscape Architecture and Environmental Planning at Utah State University. With degrees in landscape architecture and urban planning and design, Keun's research focuses on digital technologies and spatial data analytics in planning and design—1) technology-driven behavioral research in public space (drone, VR/AR, sensor, etc.) and 2) spatial research on smart growth and its behavioral outcomes (through geoprocessing, statistical modeling, visualization, etc.).

Sadegh Sabouri is a Ph.D. student in the Metropolitan Planning, Policy, and Design program at the University of Utah. He received his bachelor and master of urban planning from the University of Tehran and currently, he is working as a research analyst at the Metropolitan Research Center at the University of Utah. Sabouri's research focuses on the technologies impacts on transportation, ride-sourcing services, and advanced data analytics in urban and transportation planning.

Torrey Lyons is a Ph.D. candidate in the Metropolitan Planning, Policy and Design program at the University of Utah. He received his Bachelor of Science in Environmental Studies from the University of Vermont, and a Master of Public Policy from the University of Utah. His dissertation examines social equity in transit service. He has also written on the determinants of transit ridership, travel and the built environment, and smart growth policy. Torrey has taught research methods, program and policy analysis, and land use and transportation planning at the University of Utah.

Guang Tian is an assistant professor in the Transportation Study Program at the University of New Orleans. With a master and bachelor degrees in Geographical Information System (GIS), Guang's research focuses on GIS, urban planning and studies, transportation and its social, environmental and economic impacts.

Reid Ewing is a Distinguished Professor of City and Metropolitan Planning at the University of Utah, associate editor of the Journal of the American Planning Association, and columnist for Planning magazine. His 10 books include Pedestrian and Transit Oriented Design, co-published by the Urban Land Institute and American Planning Association; Growing Cooler: Evidence of Urban Development and Climate Change, published by the Urban Land Institute; and Best Development Practices, listed by the American Planning Association (APA) as one of the 100 "essential" books in planning over the past 100 years. Ewing's research

focuses on the built environment at five different scales and its impacts on quality of life. He has studied the built environment at scales ranging from the individual block and pedestrian activity, to the MXD and internal capture, to the metropolitan region and mode shares.