Title
The Four Step Model

Author
McNally, Michael G.

Publication Date
2008-11-17

The Four Step Model

UCI-ITS-AS-WP-07-2

Michael G. McNally

Department of Civil and Environmental Engineering and
Institute of Transportation Studies
University of California, Irvine; Irvine, CA 92697-3600, U.S.A.
mmcnally@uci.edu

2007

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

To appear as Chapter 3 in Hensher and Button (eds).“Handbook of Transport Modeling”, Pergamon [2nd Ed 2007]

Chapter 3

THE FOUR STEP MODEL

MICHAEL G. McNALLY
University of California, Irvine

1. INTRODUCTION

The history of demand modeling for person travel has been dominated by the modeling approach that has come to be referred to as the four step model (FSM) (see Chapter 2). Travel, always viewed in theory as derived from the demand for activity participation, in practice has been modeled with trip-based rather than activity-based methods (as presented in Chapter 4). Trip origin-destination (O-D) rather than activity surveys form the principle database. The influence of activity characteristics decreases, and that of trip characteristics increases, as the conventional forecasting sequence proceeds. The application of this modeling approach is near universal, as in large measure are its criticisms (these inadequacies are well documented, e.g., by McNally and Recker (1986)). The current FSM might best be viewed in two stages. In the first stage, various characteristics of the traveler and the land use - activity system (and to a varying degree, the transportation system) are "evaluated, calibrated, and validated" to produce a non-equilibrated measure of travel demand (or trip tables). In the second stage, this demand is loaded onto the transportation network in a process than amounts to formal equilibration of route choice only, not of other choice dimensions such as destination, mode, time-of-day, or whether to travel at all (feedback to prior stages has often been introduced, but not in a consistent and convergent manner). Although this approach has been moderately successful in the aggregate, it has failed to perform in most relevant policy tests, whether on the demand or supply side.

This chapter extends the material in Chapter 2 by providing a concise overview of the mechanics of the FSM, illustrated with a hypothetical case study. The discussion in this chapter, however, will focus on U.S. modeling practice. Transportation modeling developed as a component of the process of transportation analysis that came to be established in the U.S.A. during the era of post-war development and economic growth. Initial application of analytical methods began in the 1950s. The landmark study of Mitchell and Rapkin (1954) not only established the link of travel and activities (or land use) but called for a comprehensive framework and inquiries into travel behavior. The initial development of models of trip generation, distribution, and diversion in the early 1950s lead to the first comprehensive application of the four-step model system in the Chicago Area Transportation Study (see Weiner, 1997) with the model sandwiched by land use projection and economic evaluation. The focus was decidedly highway-oriented with new facilities being evaluated versus traffic engineering improvements. The 1960s brought federal legislation requiring "continuous, comprehensive, and cooperative" urban transportation planning, fully institutionalizing the FSM. Further legislation in the 1970s brought environmental concerns to planning and modeling, as well as the need for multimodal planning. It was recognized that the existing model system may not be appropriate for application to emerging policy concerns and, in what might be referred to as the "first travel model improvement program", a call for improved models led to research and the development of disaggregate travel demand forecasting and equilibrium assignment methods that integrated well with the FSM and have greatly directed modeling approaches for most of the last 30 years. The late 1970s brought "quick response" approaches to travel forecasting (Sosslau *et al.*, 1978; Martin and McGuckin, 1998) and independently the start of what has grown to become the activity-based approach (see Chapter 4). The growing recognition of the misfit of the FSM and relevant policy questions in the 1980s led to the (second, but formal) Travel Model Improvement Program in 1991; much of the subsequent period has been directed at improving the state-of-the-practice relative to the conventional model while fostering research and development in new methodologies to further the state-of-the-art (see Chapter 4).

The FSM is best seen as a particular application of transportation systems analysis (TSA), a framework due to Manheim (1979) and Florian *et al.* (1988), which positions the model well to view its strengths and weaknesses. A brief presentation of this TSA framework introduces the FSM context and leads to a discussion of problem and study area definition, model application, and data requirements. The models that are perhaps most commonly utilized in the FSM are then presented in the form of a sample application.

2. TRANSPORTATION SYSTEMS ANALYSIS

The basic structure introduced by Manheim (1979) and expanded by Florian *et al.* (1988) provides a comprehensive paradigm in which to examine the four-step model (FSM). In this representation (**Figure 1**), the transportation system **T**, defined as all elements of transportation infrastructure and services, and the activity system **A**, defined as essentially everything else (the spatial distributions of land use and the demographic and/or economic activity that occurs in those land uses), serve as exogenous inputs to performance procedures **P** and demand procedures **D**, respectively. It is such demand and performance procedures that comprise the basic FSM. While some form of location procedure **L** is required, it has typically been executed independent of the FSM and rarely integrated in any formal manner within the basic equilibration procedure. Similarly, formal supply procedures **S** are virtually non-existent. Florian *et al.* characterizes formal analysis as involving the choice of analysis perspective (effectively, time frame and spatial definition) and the definition of procedures, and thus variables, which are to be specified endogenously or exogenously.

Of critical importance to this approach is an understanding of the units of analysis for these procedures, defined spatially and temporally. Demand procedure **D** typically produces person trips, defined as the travel required from an origin location to access a destination for the purpose of performing some activity. These trips reflect units of time and space (such as daily person trips per household or peak-hour person trips per zone). Performance procedure **P** nearly always reflects mode-specific trips (person or vehicle) defined as a link volume (e.g., freeway vehicle trips per hour or boardings per hour for a particular transit route segment). The equilibration process must resolve demand and performance procedures defined at different spatial levels. Demand procedures defined at the zonal level and performance procedures defined at the link level are interconnected by the path level: paths comprise sequences of links that connect O-D pairs.

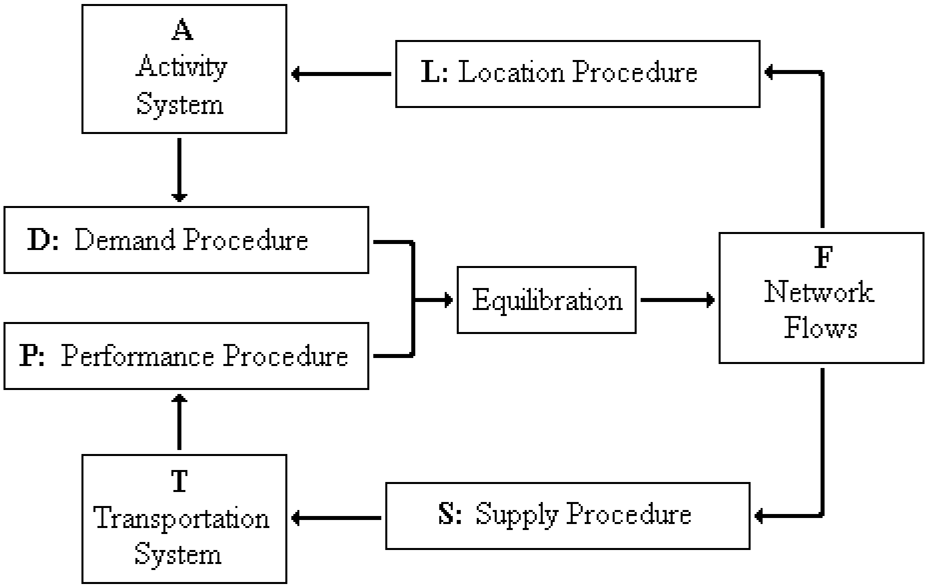


Figure 1. The Manheim/Florian Transportation Systems Analysis Framework

3. PROBLEMS, STUDY AREAS, MODELS, AND DATA

The four step model is the primary tool for forecasting future demand and performance of a transportation system, typically defined at a regional or sub-regional scale (smaller scales often apply simplified models). The FSM must be suitably policy-sensitive to allow for the comparison of alternative interventions to influence future demand and performance. The models system was developed for evaluating large scale infrastructure projects and not for more subtle and complex policies involving management and control of existing infrastructure or introduction of policies that directly influence travel behavior. Application of travel forecasting models is a continuous process. The period required for data collection, model estimation, and subsequent forecasting exercises may take years, during which time the activity and transportation systems change as do policies of interest, often requiring new data collection efforts and a new modeling effort. Little time is apparently available for systematic study of the validity of these models after the fact.

3.1 Study Area Definition

Once the nature of the problem at hand is identified, the study area can be defined to encompass the area of expected policy impact; a cordon line defines this area. The area within the cordon is composed of Traffic Analysis Zones (TAZs) and is subject to explicit modeling and analysis. Interaction with areas outside the cordon is defined via external stations (ESs) that effectively serve as doorways for trips into, out of, and through the study area. The Activity System for these external stations is defined directly in terms of trips that pass through them, and the models that represent this interaction are separate from and less complex than those that represent interactions within the study area (typically, growth factor models are used to forecast future external traffic).

The internal *Activity System* **A** is typically represented by socio-economic, demographic, and land use data defined for TAZs or other convenient spatial units. The number of TAZs, based on model purpose, data availability, and model vintage, can vary significantly (from several hundred to several thousand). The unit of analysis, however, varies over stages of the FSM and might be at the level of individual persons, households, TAZs, or some larger aggregation.

The *Transportation System* **T** is typically represented via network graphs defined by links (one-way homogeneous sections of transportation infrastructure or service) and nodes (link endpoints, typically intersections or points representing changes in link attributes). Both links and nodes have associated attributes (for example, length, speed, and capacity for links and turn prohibitions and penalties for nodes). The activity system **A** is interfaced with the transportation system **T** via centroid connectors that are abstract links connecting TAZ centroids to realistic access points on the physical network (typically mid-block and not at nodes).

3.2 Models

The FSM provides a mechanism to determine equilibrium flows as illustrated in **Figure 1**. For elementary networks, direct demand functions can be estimated and, together with standard link performance functions and path enumeration, can provide the desired flows. For any realistic regional application, an alternative model is required due to the complexity of the network. The FSM was developed to deal with this complexity by formulating the process as a sequential four step model (**Figure 2**). First, in *trip generation*, measures of trip frequency are developed providing the propensity to travel. Trips are represented as trip ends, productions and attractions, which are estimated separately. Next, in *trip distribution*, trip productions are distributed to match the trip attraction distribution and to reflect underlying travel impedance (time and/or cost), yielding trip tables of person-trip demands. Next, in *mode choice*, trip tables are essentially factored to reflect relative proportions of trips by alternative modes. Finally, in *route choice*, modal trip tables are assigned to mode-specific networks. The time dimension (time of day) is typically introduced after trip distribution or mode choice where the production-attraction tables are factored to reflect observed distributions of trips in defined periods (such as the a.m. or p.m. peaks). In route choice, performance characteristics are first introduced, thus, the FSM in its basic form only equilibrates route choices. In other words, total "demand", as specified through generation, distribution, mode choice, and time-of-day models, is fixed, with only the route decision to be determined. Most applications of the FSM feedback equilibrated link travel times to the mode choice and/or trip distribution models for a second pass (and occasionally more) through the last three steps, but no formal convergence is guaranteed in most applications. Integrated location procedures (land use and transportation models) are absent in most U.S. applications. The future activity system is forecasted independently with no feedback from the FSM (see Chapter 9).

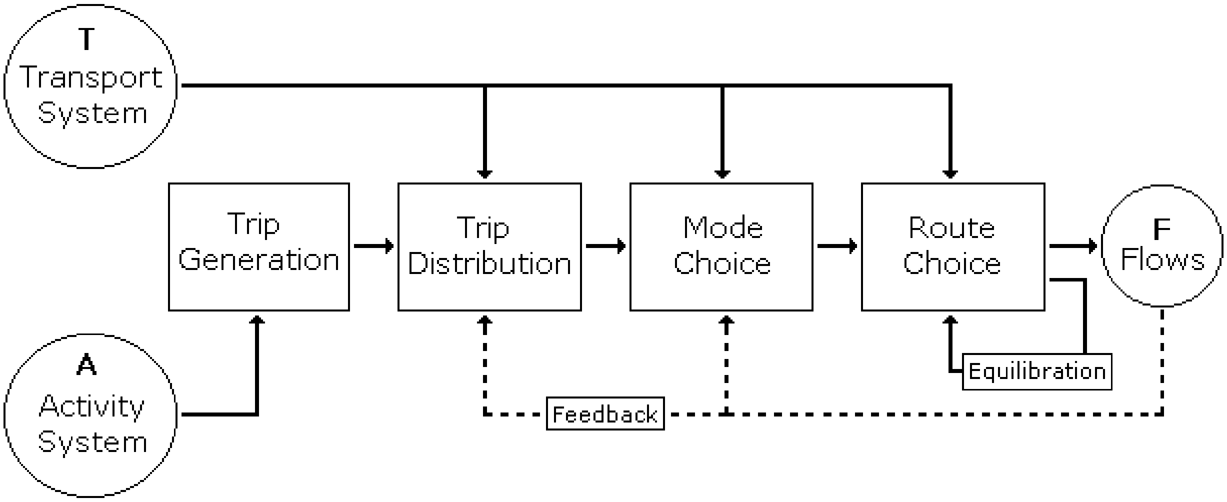


Figure 2. The Four Step Model

3.3 Data

The FSM has significant data demands in addition to that required to define the activity and transportation systems. The primary need is data that defines travel behavior and this is gathered via a variety of survey efforts. Household travel surveys with travel-activity diaries provide much of the data that is required to calibrate the FSM. These data, and observed traffic studies (counts and speeds), also provide much of the data needed for validation.

Household travel surveys provide: (i) household and person-level socio-economic data (typically including income and the number of household members, workers, and cars); (ii) activity-travel data (typically including for each activity performed over a 24-hr period activity type, location, start time, duration, and, if travel was involved, mode, departure time, and arrival time; and (iii) household vehicle data. This survey data is utilized to validate the representativeness of the sample, to develop and estimate trip generation, trip distribution, and mode choice models, and to conduct time-in-motion studies.

3.4 A Sample Problem

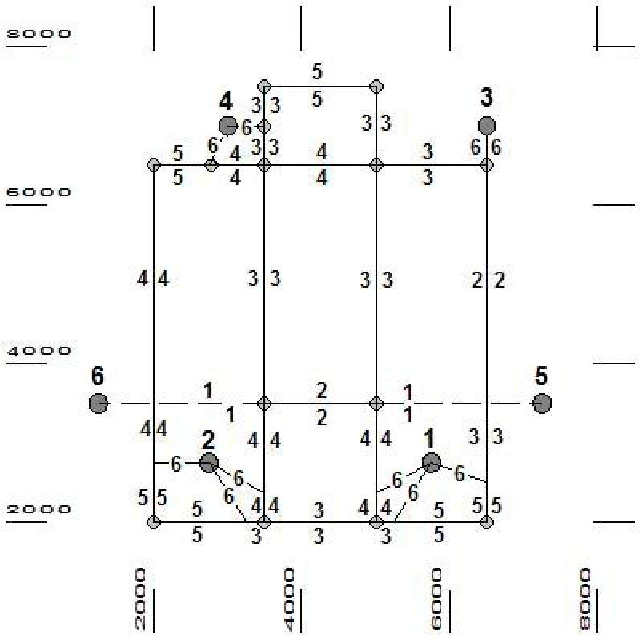
An example of the FSM is provided to illustrate a typical U.S. application. Initial tasks define the transportation and activity systems in a form compatible with the FSM software being utilized, tasks increasingly facilitated by Geographical Information Systems (GIS) (see Chapters 14-16). **Table 1** and **Figure 3** depict the transportation network for the study area and **Table 2** summarizes key population-level demographic variables for the area's four TAZs (1-4). There are also two external stations (5-6), the associated trips of which are separately modeled and appended to the study area trip tables.

In this hypothetical example, a home interview survey was "conducted" for a five percent sample of households in each TAZ, "yielding" 1852 total trips in the 200 households (see **Table 3**). The sample size in this example is too small to ensure that it is representive of the population, and the estimation of FSM models will violate the minimum category counts required for statistical significance, but this should not limit the utility of the sample problem. The stages of the FSM are presented below in sequence (Sections 4-7).

Table 1. Network Characteristics

	Link Type (all links 1-way)	Speed (kph)	Number of Lanes	capacity per lane	capacity (veh/hour)
1	freeway	90	2	200	400
2	primary arterial	90	2	100	200
3	major arterial	60	2	100	200
4	minor arterial	45	2	100	200
5	collector street	45	1	100	100
6	centroid connector	30	9	100	900

Figure 3. The Transportation Network for the Study Area



[link annotation is facility type (see Table 1)]

Table 2. Zonal Socio-Economic Data
(total number of households per zone and total number of employees per zone)

Internal Zone	Total Zonal Households	Total Zonal Employment			
		Retail	Service	Other	Total
1	1400	800	400	800	2000
2	1200	800	400	400	1600
3	800	200	400	200	800
4	600	200	200	0	400
Total	4000	2000	1400	1400	4800

Table 3. Household Demographic Data
(number of households per zone by household car ownership and household income)

	Zone 1			Zone 2			Zone 3			Zone 4		
	L	M	H	L	M	H	L	M	H	L	M	H
0 cars	40	80	80	20	40	40	0	0	0	0	0	0
1 car	120	320	360	80	260	160	20	80	100	0	40	60
2 cars	40	200	160	100	300	200	80	220	330	0	160	340

Note: L=low income; M=middle income; H=high income

4. TRIP GENERATION

The objective of this first stage of the FSM process is to define the magnitude of total daily travel in the model system, at the household and zonal level, for various trip purposes (activities). This first stage also explicitly translates the FSM from activity-based to trip-based, and simultaneously severs each trip into a production and an attraction, effectively preventing network performance measures from influencing the frequency of travel. Generation essentially defines total travel in the region and the remaining steps are effectively share models.

4.1 Process

Separate generation models are estimated for productions $f_P^p(\mathbf{A})$ and attractions $f_A^p(\mathbf{A})$ for each trip type (purpose) p:

$$P_i^p = f_P^p(\mathbf{A} \text{ activity system characteristics}) \tag{1}$$

and

$$A_j^p = f_A^p(\mathbf{A} \text{ activity system characteristics}) \tag{2}$$

where: P_i^p are the total trip productions generated for trip type p for analysis unit i and A_j^p are the total trip attractions for trip type p for analysis unit j.

Virtually all model applications are for discrete spatial systems typically defined by on the order of 100-2000 traffic analysis zones. Typically, at least three different trip purposes are defined, often home-based work trips (HBW), home-based other (or non-work) trips (HBO), and non-home-based trips (NHB). The majority of trips are typically home-based, having their origin or destination at home. NHB trips have neither trip end at home (these trips are thus linked trips and part of a home-based trip chain, although this distinction is usually ignored in the FSM). Trip ends are modeled as productions or attractions. The home-end of a trip is always the production -- it is the household and its activity demands that gives rise to, or produce, all trips; the non-home end is the attraction (for NHB trips, the origin is the production and the destination is the attraction).

Trips can be modeled at the zonal, household, or person level, with household level models most common for trip productions and zonal level models most common for trip attractions. For household production models, all trips are initially generated at the home location, and NHB trips must be re-allocated to be "produced" in the actual origin zone of the trip. Such production models can reflect a variety of explanatory and policy-sensitive variables (e.g., car ownership, household income, household size, number of workers). Category models are more common that regression-based models and provide a reasonably accurate measure of trip frequency at the household level and, once aggregated, at the zonal level (person-level models are similar in structure). The independent modeling of trip ends has limited the ability to integrate measures of accessibility into generation models (few if any models have achieved significant inclusion of accessibility variables despite the intuitive appeal that such variables should affect trip frequency, a result that eliminates potential feedback from route choice models). Trip attraction models serve primarily to scale the subsequent destination choice (trip distribution) problem. Essentially, these models provide a measure of relative attractiveness for various trip purposes as a function of socio-economic and demographic (and sometimes land use) variables. The estimation is more problematic, first because regional travel surveys sample at the household level (thus providing for more accurate production models) and not for non-residential land uses and second because the explanatory power of attraction variables is usually not as good. For these reasons, factoring of survey data is required prior to relating sample trips to population-level attraction variables, typically via regression analysis. Subsequent attraction levels, while typically normalized to production levels for each trip purpose, should nonetheless be carefully examined if the totals vary significantly from that for productions. Special generators are introduced to independently model trips at locations (such as universities) that are not well-represented in the standard models.

The above discussion refers to internal trips (resident trips with both ends in the study area). Non-resident trips within the study area and external trips (including both through trips and trips with one end outside of the study area) are modeled separately (but must not double count resident trips already reflected in the regional travel survey). External-internal trips typically are modeled with the production at the external station. Internal attractions are scaled to total internal productions plus the difference between external productions and external attractions. Growth factors, often reflecting traffic counts at the external stations, are used to factor current external totals for forecasting purposes. External and external-internal trips, typically vehicle trips, are integrated in the vehicle trip tables prior to route assignment.

4.2 A Sample Household Trip Production Model

A category model was estimated for household trips by purpose from the trip data and demographic characteristics of the 200 sample households. Category variables are selected based on ability to significantly discriminate trip rates between categories, general policy sensitivity, and the availability of future data. Category models are less restrictive than regression models but require that the joint distribution of population variables be forecast. A variety of methods have been used with iterative proportional fitting perhaps the most direct. The role of activity system forecasts is clear, as is the need for quality forecasts of automobile ownership since this variable is typically most highly correlated with total trips per household. The resulting estimated trip rates are displayed in **Table 4** (to simplify presentation, rates from Martin and McGuckin (1998) are utilized). Aggregation proceeds directly since the model is linear. Once the joint distribution of households is known for the base or forecast year, the cell counts are multiplied by the estimated trip rates to obtain the total number of trips per zone.

Table 4. Sample Estimated Household Trip Production Model (daily person trips per HH)

Cars per HH	Household Income	HBW	HBO	NHB	Total
Cars = 0	Low HH Income	0.5	2.0	0.9	3.4
	Mid HH Income	1.1	3.0	1.2	5.3
	High HH Income	1.4	3.9	1.8	7.1
Cars = 1	Low HH Income	0.8	3.2	1.3	5.3
	Mid HH Income	1.5	3.9	1.6	7.0
	High HH Income	1.8	4.9	2.2	8.9
Cars = 2	Low HH Income	1.4	5.2	2.1	8.7
	Mid HH Income	2.1	5.7	2.3	10.1
	High HH Income	2.5	6.6	3.1	12.4
Source: based on Martin and McGuckin (1998), Table 7, pg. 27.					

Note: HBW=home based work; HBO=home based other; NHB=non-home based

4.3 A Sample Zonal Attraction Model

The sample model estimate relative attractiveness by regressing factored values of sample trips (aggregated to the zone level) on relevant zonal characteristics. The choice of explanatory variables is constrained in a manner similar to trip productions models - model significance, policy sensitivity, and forecastability. These models are summarized in **Box 1**.

Box 1. Sample Estimated Zonal Trip Attraction Models

Example of Estimated Trip Attraction Models	
Zonal HBW Attractions =	1.45 * Total Employment
Zonal HBO Attractions =	9.00 * Retail Employment + 1.70 * Service Employment + 0.50 * Other Employment + 0.90 * Number of Households
Zonal NHB Attractions =	4.10 * Retail Employment + 1.20 * Service Employment + 0.50 * Other Employment + 0.50 * Number of Households
Source: Martin and McGuckin (1998), Table 8, pg. 28.	

4.4 Application to the Base Population

There is no internal consistency between the production and attraction models. With productions models in general being more reliable, attractions by purpose are typically normalized to productions (this process may need to reflect internal-external trips if they are included in the rate model). The estimated models are applied to population-level data for the study area (the zonal data in **Table 2** and the estimated population-level equivalent of **Table 3**); these values are displayed in **Table 5**. Estimates for NHB trips are listed for the zone in which the household is located and these trips must be re-allocated to the actual zone of origin.

Table 5. Sample Trip Generation Results

	HBW		HBO		NHB[a]		Total	
Zone	P	A[b]	P	A[b]	P	A[b]	P	A[b]
1	2320	2900	6464	9540	2776	4859	11560	17299
2	2122	2320	5960	9160	2530	4559	10612	16039
3	1640	1160	4576	3300	1978	1800	8194	6260
4	1354	580	3674	2680	1618	1359	6646	4619
Total	7436	6960	20674	24680	8902	12577	37012	44217

Note: a. tabulated NHB trips are not yet re-allocated; b. attractions not normalized

4.5 Time of Day

Trip generation can reflect time of day with productions and attractions being generated for specific time periods; this is often the case when compiled land use trip rates are utilized since these rates are typically defined by time of day. Adjustments for time of day, however, are more common after subsequent FSM steps.

5. TRIP DISTRIBUTION

The objective of the second stage of the process is to recombine trip ends from trip generation into trips, although typically defined as production-attraction pairs and not origin-destination pairs. The trip distribution model is essentially a destination choice model and generates a trip matrix (or trip table) T_{ij} for each trip purpose utilized in the trip generation model as a function of activity system attributes (indirectly through the generated productions P_i and attractions A_j) and network attributes (typically, interzonal travel times).

5.1 Process

The general form of the trip distribution model as the second step of the FSM is:

$$\begin{aligned} T_{ij} &= f_{TD}(\mathbf{A}, t_{ij}) \\ &= f_{TD}(P_i, A_j, t_{ij}) \qquad (3) \end{aligned}$$

where t_{ij} represents a measure of travel impedance (travel time or generalized cost) between the two zones (note that the index p describing trip purpose is dropped for simplicity). For internal trips, perhaps the most common model is the so-called gravity model:

$$T_{ij} = a_i\, b_j\, P_i\, A_j\, f(t_{ij}) \qquad (4)$$

where:

$$a_i = [\Sigma_j\, b_j\, A_j\, f(t_{ij})]^{-1}$$

$$b_j = [\Sigma_i\, a_i\, P_i\, f(t_{ij})]^{-1}$$

and $f(t_{ij})$ is some function of network level of service (LOS)

The production-constrained gravity model sets all b_j equal to one and defines W_j in place of A_j as a measure of relative attractiveness. The impedance term, $f(t_{ij})$, essentially provides a structure for the model with the balancing terms scaling the resulting matrix to reflect the input productions and attractions. The estimation of gravity models involves the estimation of this function. While various intuitively and empirically-supported functional forms have been used, for many years the most common estimation technique involved the iterative fitting of "friction factors" reflecting the observed travel frequency distributions from the household travel survey. The friction factors were subsequently smoothed to exponential, power, or gamma distributions. Most software now allows for direct estimation of these functions, although the implication is that one or two parameters are responsible for overall distribution of each trip purpose. The estimated impedance function is assumed to capture underlying travel behavior and to be stable in the future to allow its use in forecasting. Discrete choice models also have occasionally been utilized for destination choice (see Chapter 5). Growth factor models are utilized primarily to update existing matrices for external trips but are not used for internal trips since measures of level-of-service are not incorporated. The most common of these (Furness or Fratar) is identical to the doubly-constrained gravity model with an existing trip matrix replacing the impedance function and essentially providing the structure by which the model functions.

5.2 Travel Impedance and Skim Trees

Most trip generation models unfortunately to not reflect travel impedance or any general measure of accessibility due to the splitting of trips into productions and attractions. Travel impedance is explicitly utilized in subsequent stages, thus, skim trees (interzonal impedances) must be generated prior to subsequent steps. Free flow automobile travel times are most often used for the initial (and sometimes only) pass through the FSM. Ideally, these skim trees would reflect generalized costs appropriately weighted over all modes in subsequent steps. Only interzonal impedances are directly computed. Intrazonal impedance is estimated via a weighted average of interzonal impedance to one or more neighboring zones. The skim matrix is usually updated to reflect terminal time for access and egress at either end of the trip. **Table 6** depicts the resulting skim trees. When there is feedback from the route choice stage, travel times estimated from link performance functions using assigned link volumes are utilized instead of initial link travel times and new skim trees are developed. Since the results of assignment are period specific care must be exercised when returning to the distribution stage in determining what value of link impedance should be used.

5.3 A Sample Gravity Model

The 1852 trips from the household survey were used to construct an observed trip length frequency distribution and, together with minimum path skim trees, were used to estimate gravity models for each of the three trip types (HBW, HBO, and NHB). A gamma impedance function was estimated (see **Table 7**).

Table 6. Minimum Travel Time Paths (Skim Trees)

Skim t_{ij}	TAZ 1	TAZ 2	TAZ 3	TAZ 4	ES 5	ES 6
TAZ 1	**1**	**4**	**5**	**8**	**4**	**5**
TAZ 2	**4**	**2**	**9**	**7**	**5**	**4**
TAZ 3	**5**	**9**	**2**	**6**	**7**	**8**
TAZ 4	**8**	**7**	**6**	**3**	**7**	**6**
ES 5	**4**	**5**	**7**	**7**	**0**	**4**
ES 6	**5**	**4**	**8**	**6**	**4**	**0**

Table 7. Sample Estimated Impedance Function for the Gravity Model: $f(t_{ij}) = a\, t_{ij}^{b} \exp(c\, t_{ij})$

Trip Purpose	Parameter a	Parameter b	Parameter c
Home-based Work (HBW)	28,507	-0.020	-0.123
Home-based Other (HBO)	139,173	-1.285	-0.094
Non-home-based (NHB)	219,113	-1.332	-0.100
Source: Martin and McGuckin (1998). Table 14, pg.41.			

5.4 Adjustments

The calibration process is driven by the underlying trip length frequency distribution. In the basic process, either this distribution or its mean is used to judge convergence. The relative distribution of trip interchanges (matrix cells) is not directly considered. Individual cells can be adjusted via estimation of K_{ij} factors, but opinions vary as to the use of what are essentially fudge factors. On one hand, it is difficult to relate any policy variables to these factors, thus, it is difficult to assess their validity in the future. On the other hand, the resultant base trip matrix will more closely reflect observed behavior.

The trip matrices are at this stage defined as production to attraction (P-A) flows. Depending on the treatment of mode choice, these matrices may be converted from P-A format to O-D format (which is required in the route choice step). Conversions may also be made at this stage to reflect time-of-day, particularly if the subsequent mode choice models are period-dependent. In this sample application, these adjustments are made prior to mode choice. P-A to O-D conversion typically reflects the observed travel data. When surveys are analyzed to develop base distributions of observed trips by purpose, the proportion of trips from the production zone to the attraction zone are also computed. These rates are depicted in **Table 8.** While 24-hour factors are usually equivalent, period specific-factors vary significantly (for example, many more HBW trips are generally heading from the work attraction to the home production in the PM peak period than the reverse). Each cell of the O-D trip table is computed by adding the product of the corresponding cell of the P-A trip table multiplied the appropriate P-to-A factor to the corresponding cell of the transposed P-A trip table multiplied by the appropriate A-to-P factor (**Table 8**).

Table 8. Time-of-Day and P-A/O-D Conversion Factors and Average Vehicle Occupancy

	HBW Trips		HBO Trips		NHB Trips	
Period	P to A	A to P	P to A	A to P	P to A	A to P
2-hr a.m. peak	0.30	0.00	0.06	0.02	0.04	0.04
3-hr p.m. peak	0.03	0.30	0.10	0.15	0.12	0.12
Off-peak	0.17	0.20	0.34	0.33	0.34	0.34
1-hr p.m. peak	0.02	0.15	0.04	0.07	0.06	0.06
Average Occupancy	1.10 persons/veh		1.33 persons/veh		1.25 persons/veh	

6. MODE CHOICE

Mode choice effectively factors the trip tables from trip distribution to produce mode-specific trip tables. These models are now almost exclusively disaggregate models often estimated on separate choice-based samples and reflecting the choice probabilities of individual trip makers. While in U.S. applications, transit is less of a factor, many recent mode choice models reflect current policies such as carpooling choices resulting from high occupancy vehicle facilities and the presence of tolls on automobiles. The most common model estimated is the nested logit model (see Chapters 5 and 13). These mode choice models can reflect a range of performance variables and trip-maker characteristics, but produce disaggregate results that must then be aggregated to the zonal level prior to route choice (see Ortuzar and Willumsen, 2001).

Due to space limitation, in lieu of a formal mode choice model the sample application instead utilizes a simplified factoring of the person trip tables to allow for the development of vehicle trip tables. Essentially, average vehicle occupancies reflecting total person trips versus total vehicle trips are used to produce the trip table of automobile trips while ignoring trips by other modes. This of course would only be valid if the proportion of trips by other modes was very small, but it does allow for the illustration of how vehicle trip tables are then assigned to the highway network; transit trips, if computed, would be assigned to the corresponding transit network. Some software allows for the simultaneous equilibration of true multimodal networks and these methods should be utilized when significant choices exist. Here, average occupancies are "determined" from the hypothetical travel surveys and are included in **Table 8**.

7. ROUTE CHOICE

In this last of four major steps of the FSM, an equilibration of demand and performance is finally present. Modal O-D trip matrices are loaded on the modal networks usually under the assumption of user equilibrium where all paths utilized for a given O-D pair have equal impedances (for off-peak assignments, stochastic assignment has been used, which tends to assign trips across more paths better reflecting observed traffic volumes in uncongested periods).

7.1 Process

The basic UE solution is obtained by the Frank-Wolfe algorithm that involves the computation of minimum paths and all-or-nothing (AON) assignments to these paths. Subsequent AON assignments (essentially linear approximations) are weighted to determine link volumes and thus link travel times for the next iteration (see Chapters 10 and 11). The estimated trip tables are fixed, that is, they do not vary due to changing network performance.

Although combined models integrating any or all of the four stages have been developed, they have rarely been applied in practice (in part due to the non-availability of standard software and agency inertia). Rather, informal and heuristic feedback mechanisms have been introduced. With congestion effects explicitly captured at the route choice level, the most common feedback is to mode and destination choice where congested link travel times are used to determine congested paths for subsequent re-distribution of trips.

7.2 A Sample Assignment of Vehicle Trip Tables to the Highway Network

After adjusting the P-A trip tables to O-D format, converting to time-of-day, and factoring to reflect vehicle occupancy, the trip tables by purpose are aggregated for assignment. Estimates of external vehicle traffic are then appended (see **Table 9**). The user equilibrium assignment resulting from loading this trip table on the sample network is depicted in **Figure 4** (links depict volume capacity ratios). No feedback was attempted; these results represent a single pass through the FSM sequence. Significant congestion in the p.m. peak hour is apparent. Resultant link volumes and travel times must be validated versus ground counts on an intersection, link, or corridor (screenline) basis prior to accepting the model system as valid for forecasting purposes.

Table 9. PM Peak Vehicle Trip O-D Matrix

T_{ij}	TAZ 1	TAZ 2	TAZ 3	TAZ 4	ES 5	ES 6	Origins
TAZ 1	829	247	206	108	100	100	1590
TAZ 2	235	725	104	158	100	100	1422
TAZ 3	137	72	343	89	0	0	641
TAZ 4	59	98	76	225	0	0	458
ES 5	0	0	100	100	0	500	700
ES 6	0	0	100	100	500	0	700
Destinations	1260	1142	929	780	700	700	5511

8. SUMMARY

This chapter has provided an overview of the application of the conventional model of travel forecasting, commonly referred to as the four step model. The text by Ortuzar and Willumsen (2001) represents the best current overall reference on the FSM. From the perspective of the state-of-the-practice, the choice of this approach is not that it is the best available but because it is the only approach available, given current institutional requirements and financial limitations. Many of the criticisms of this approach are addressed in Chapter 4, which presents activity-based approaches that have been developed to better represent underlying travel behavior and thus hold promise to forward the science and art of transportation forecasting.

REFERENCES

Florian,M., M.Gaudry, and C.Lardinois (1988). A two-dimensional framework for the understanding of transportation planning models, *Transportation Research B*, **22B**, 411-419.

Manheim,M.L. (1979). *Fundamentals of Transportation Systems Analysis*, MIT Press, Cambridge, MA.

Martin,W.A. and N.A.McGuckin (1998). *Travel Estimation Techniques for Urban Planning*, NCHRP Report 365, Transportation Research Board, Washington, DC.

McNally,M.G. and W.W.Recker (1986). *On the Formation of Household Travel/Activity Patterns: A Simulation Approach*, Final Report to USDoT, University of California.

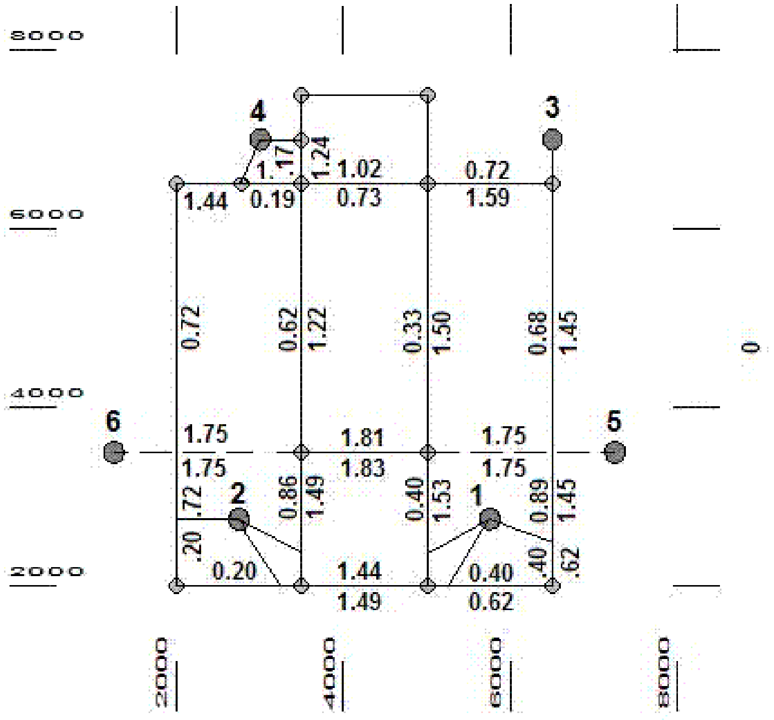
Mitchell,R.B. and C.Rapkin (1954). *Urban Traffic: A Function of Land Use*, Columbia University Press, New York, NY.

Ortuzar,J.de D. and L.G.Willumsen (2001). *Modelling Transport* (third edition), Wiley, Chichester.

Sosslau,A.B., A.B.Hassan, M.M.Carter, and G.V.Wickstrom (1978). *Quick Response Urban Travel Estimation Techniques and Transferable Parameters: User Guide*, NCHRP Report 187, Transportation Research Board, Washington, DC.

Weiner,E. (1997). *Urban Transportation Planning in the United States: An Historical Overview* (fifth edition), Report DOT-T-97-24, US Department of Transportation, Washington, DC

Figure 4. Assignment of Vehicle Trip Tables to the Highway Network



[link volume/capacity ratios are shown]