3.0 REVIEW OF RELEVANT LITERATURE

3.1 Introduction

For accomplishing the research objectives as stated in chapter 2.0, review of relevant literature was performed. The focus of this effort was on the synthetic O-D trip table estimation models, and the application of socio-economic/census data to transportation planning. This chapter provides a brief overview of some of the approaches that estimate O-D trip tables from ground counts and old/prior trip table information, with particular emphasis on The Highway Emulator (THE) and the Linear Programming (LP) models. Both THE and LP models were tested extensively on the Pulaski network for the earlier research effort and were found to show encouraging results. Therefore, they have been selected for further evaluation. One of the research objectives was to study the models’ performances with improved target/seed table. To establish the improved target/seed table, it was important to incorporate more realistic travel characteristic in the target/seed table. To accomplish this task, the socio-economic/census data were investigated. Census Transportation Planning Package (CTPP), which contains the 1990 census data for different areas was also considered as a potential source for obtaining the relevant data.

3.2 Review of O-D Trip Tables Estimation Models Using Link Volumes

1970s saw the beginning of the theoretical approaches for estimating O-D trip tables using link traffic counts among transportation planners and engineers. The desire of planning agencies and transportation professionals for alternatives to traditional, long term approaches for O-D trip tables estimation, initiated the research in this field.
Limitations in budget, time and man power resources further enhanced the research towards development of theoretical models for trip table estimation. Since then, different approaches have evolved, incorporating various desirable features and refinements. In general the synthetic O-D trip table estimation models use the available link volume and old/prior target/seed table information for estimating the current trip table.

THE and LP models have been reviewed in detail below, since they were the focus of this research work.

3.2.1 The Highway Emulator (THE)

THE model (Bromage, 1991) is a microcomputer highway traffic simulation model for modeling of individual communities, corridors, sections of counties and analysis of small sections of major cities. THE incorporates two distinct modeling approaches, the first utilizes the traditional four-step urban transportation planning methodology, and second utilizes the maximum entropy algorithm for estimating trip tables from link traffic volumes. The second approach was of primary interest for this research, as it extracts a most likely trip table, from the observed link volume and target/seed table information. The trip table estimation program is based on the maximum entropy formulation and algorithm detailed by Van Zuylen and Willumsen (1980). A maximum entropy algorithm attempts to define a trip table with the maximum degree of disorder or random exchange between zones. Besides regional and sub regional modeling efforts, this model can be applied to estimate origin-destination tables for traffic circles, one way streets in a CBD, freeways, toll roads and other limited access facilities.
3.2.1.1 Trip Table Estimation Procedure

The program first assigns a seed trip table to the specified network. While the assignment is in progress the program creates a link use probability file. These probabilities are used to make adjustments to the trip exchange matrix in order to duplicate the observed volumes. The next step is the actual application of the trip table estimator algorithm. The algorithm is iterative and may take up to one hour for each iteration, for large networks. The output of the trip table estimation program is a new trip table which closely duplicates the observed traffic counts for the given probabilities. The above steps are repeated to satisfy the number of calibration iterations. For microcomputer applications, with respect to handling large networks, the trip table estimation program which can handle a maximum of 300 traffic zones and up to 500 links (Bromage, 1991).

3.2.1.2 Data Requirements

THE model requires the following data as input for running the trip table estimation program: node numbers at each end, link length in miles, free flow speed on the link, link delays (if any), whether link is one way or two way, any additional impedance (such as a zone connector terminal time), hourly capacity for each direction, link volumes (preferably for all links, but partial counts also accepted), and prior (target/seed) trip table, if desired to guide the solution (a structural table could also be used).

The coding of network is based on traditional method, and is also outlined in NCHRP Report # 187 (NCHRP Report 187, 1978). It is also consistent with FHWA’s UTPS coding scheme. It must be noted that it cannot accept a node with more than four connections. In such cases, dummy links are needed. However, turn links can be coded. In terms of user parameters to
be supplied, THE requires the specification of number of iterations for maximum entropy process, assignment and calibration. However, default values for these can also be used. For the Pulaski network 15 iterations were specified for carrying out the trip table estimation program.

### 3.2.1.3 Model Restrictions

The Highway Emulator has number of restrictions while running the trip table estimation program. While few of the restrictions are minor and/or internal to the model, certain others have implication in data collection and manipulation. These are briefly described below (Bromage, 1991).

**Balanced flow at nodes:** THE manual states that for the trip table estimation program, it is necessary to provide balanced traffic counts, except for roadway segments intersected by traffic zone connectors. This means that the total traffic volume entering an intersection node must be equal to the total volume exiting it, i.e., conservation of flow should be maintained. While this is a restriction, it was learnt from Mr. Bromage that the model could still be run without the balanced flows. However, this may affect the quality of the results.

**Assignment to Produce Observed Counts:** Another constraint for THE model trip table estimation program, as stated in the manual, is that the assignment of the estimated trip table should produce the observed volumes. The algorithm works by adjusting a seed trip table slightly in order to match the observed volumes. If a trip table cell has some or all of its volumes assigned to a network link for which there exist observed volumes, then that cell is adjusted so that the above mentioned constraint is met.

**Counts on links:** THE manual states that “the trip table program works very well if traffic counts are available for every network link, including those that connect traffic zones.
Unfortunately, traffic volume data is usually known on only a fraction of the network links which comprise the highway network. Without traffic count data on zone connectors and on virtually every network link, the trip table estimation program will take great liberty in defining what volume should be on the zone connector”. The manual recommends that the trip table estimation program should be used in conjunction with a gravity model distribution in situations where traffic volume data is not known for traffic zone connectors.

For this research approach THE was used for estimating the trip table from the observed link volume and prior target/seed trip table. The Pulaski network details i.e., link length, link speed, link distance, free flow time, link capacity and observed link volumes were suitably coded, before running the trip table estimation program. The modeled trip tables obtained were then evaluated with the VDOT surveyed trip tables.

### 3.3 Linear Programming (LP) Model

LP is one of the latest theoretical model that estimates trip tables from ground counts and old/prior trip table information. With the objective of overcoming certain weaknesses of earlier models, this new linear programming approach was developed at Virginia Tech. A detailed description of the development this model and preliminary test results are presented in the Ph. D. dissertation of Sivanandan (1991) and in the paper by Sherali, Sivanandan, and Hobeika (1994). The model is based on a non-proportional-assignment, user-equilibrium principle and incorporates the linear programming approach for estimating O-D trip tables from available data on link traffic volumes. The model is designed to determine a traffic equilibrium network flow solution that reproduces the link volume data, if such a solution exists. The model’s solution also has a tendency to match a specified prior target trip table as closely as possible. The model recognizes that due to incomplete information, although the
individual user is driven by the choice of a least impedance path, the actual flow may not exactly conform to a user equilibrium solution. Moreover, due to inherent inconsistencies in the link traffic data, there might not exist a trip table that can exactly duplicate the link flows. Accordingly, these features are accommodated into the model through suitable artificial variables and objective penalties. However, if there does exist a user equilibrium solution that reproduces the link flows, the model, with suitable penalty parameters, will determine such a solution along with the corresponding O-D trip table. Additionally, due to the potentially large number of alternative paths to be considered between the different O-D pairs, an efficient column generation technique that utilizes shortest-path sub-problems in order to determine an optimal solution to the linear programming model has been developed. The model is also designed to handle the situation in which a prior target trip table is specified, and it is required to find a solution that, in addition to the foregoing considerations, has a tendency toward reproducing this table as closely as possible.

The linear programming model is formulated as below (Sherali et al., 1994):

Minimize
\[
\sum_{(i,j) \in OD} \sum_{k=1}^{n_{ij}} \tilde{t}_{ij}^k x_{ij}^k + M e \cdot (y^+ + y^-) + M \sigma \sum_{(i,j) \in OD} (Y_{ij}^+ + Y_{ij}^-) \tag{3.1}
\]

subject to
\[
\sum_{(i,j) \in OD} \sum_{k=1}^{n_{ij}} p_{ij}^k x_{ij}^k + y^+ - y^- = \tilde{f} \tag{3.2}
\]
\[
\sum_{k=1}^{n_{ij}} x_{ij}^k + Y_{ij}^+ - Y_{ij}^- = Q_{ij} \quad \forall (i,j) \in OD \tag{3.3}
\]
\[
x \geq 0, \quad y^+ \geq 0, \quad y^- \geq 0, \quad Y^+ \geq 0, \quad Y^- \geq 0. \tag{3.4}
\]

where,

\(OD\) is set of origin-destination (O-D) pairs comprising the trip table

\(\overline{OD}\) is some key O-D pairs for which target trip values are specified

\((i,j)\) is O-D pair from origin \(i\) to destination \(j\)

\(k\) is path identifier between different O-D pairs
$n_{ij}$ is number of paths between O-D pair $(i,j)$

$\hat{t}_{ij}^k$ is weighted impedance on route $k$ between O-D pair $(i,j)$

$$\hat{t}_{ij}^k = \begin{cases} 
    t_{ij}^k & \text{if } k \in K_{ij} \\
    M_{ij}t_{ij} & \text{if } k \notin K_{ij}
\end{cases}$$

$t_{ij}^k$ is impedance on route $k$ between O-D pair $(i,j)$

$K_{ij}$ is set of paths between O-D pair $(i,j)$ whose path costs are equal to the shortest path for the O-D pair $(i,j)$

$x_{ij}^k$ is flow on path $p_{ij}^k$, for each $k = 1, \ldots, n_{ij}$, $(i,j) \in OD$

e is vector of ones

$y^+$ ($y^-$) is vector of positive (negative) deviations in link flows

$Y^*$ ($Y^-$) is vector of positive (negative) deviations from targeted trip table values

$p_{ij}^k$ is $k^{th}$ path between O-D pair $(i,j) \in OD$

$\bar{f}$ is vector of observed link volumes

$Q_{ij}$ is Prior (target) trip table value for O-D pair $(i,j) \in OD$

$M_i$ and $M_o$ are scalar penalty parameters.

A modified column generation solution technique has been presented to optimally solve the above problem. The preliminary test results indicated the superiority of this model over the maximum entropy model (Sherali et al., 1994a,b).

The above model had a limitation in that it required the specification of volume information for all the links of the network. Since this may not always be available, the model was enhanced during the earlier research at Center for Transportation Research by adding the capability to estimate O-D trip tables even when only a “partial set” of link volume information was available. The proposed approach
formulates a sequence of linear programs to approximate a fundamentally nonlinear optimization problem that is employed to estimate O-D flows, given incomplete network flow information (Narayanan, 1995).

### 3.3.1 Data Requirements

LP requires the following data input for running the trip table estimation program: zones, nodes and link numbers (traditional network coding method), hourly capacities, free-flow speeds and lengths of links (alternatively, if current travel times on links are known, that will suffice), whether link is one way or two way, link volumes (preferably for all links, but partial counts also accepted), prior (target/seed) trip table data (optional) (a structural target could also be used), and any link delays and additional impedance, if appropriate, may be included in the current link travel times.

For the LP model, user also needs to input the value for the parameter sigma (σ) to reflect the relative degree of importance in minimizing the trip table deviations (modeled vs. targeted) versus the link flow deviations (modeled vs. observed). Hence, suitable value of sigma (0 \leq \sigma \leq 1), ranging from unimportant (σ=0) to equally important (σ=1) can be specified, to reflect the relative degree of importance placed in the target/seed trip table for guiding the LP models solution. For this research, since the target/seed table was developed using the socio-economic/census data, greater confidence was placed in the target/seed table. To reflect this in the model runs, a relative high σ value (0.9) was used for the LP runs, to minimize the modeled trip table deviations from the input target/seed trip table.
3.3.2 Model Restrictions

The LP model accommodates inconsistencies in link volumes and does not require balanced flows at nodes. However, the model will yield better results in the absence of such discrepancies. It can also accept more than four links joining at a node.

Counts on Links: The original version of the model required that traffic count data on all the network links be given as input. However, during the course of earlier research, the model was enhanced to overcome this restriction, and can now estimate O-D tables with partial set of volume data. But superior results can be expected with complete (or close to it) volume information.

Network Size: The limitation of LP model with regard to network size arises mainly due to limitations in allocating array sizes in the computer program. The array sizes in the model are influenced by the number of constraints of the optimization program, which are based on the number of links and the number of O-D interchanges to be targeted. However, since this restriction is compiler and computer based, it can be overcome through the use of an enhanced compiler and a powerful computer. Also, the computation time for model runs tend to grow rapidly with an increase in network size.

3.4 Use of Census Data in Transportation Planning

From its beginning with the 1960 census to the present, the planning process has become increasingly dependent on the census data, including all of the supporting socio-economic data from the census that provided flexible small area population characteristics for input to trend analysis and forecasts (Pisarski, 1994). There are several elements of data contained in
the census that have relevance to urban transportation planning. A brief overview on the use of census data in transportation planning is shown below.

3.4.1 The Census Transportation Planning Package (CTPP) and its Use in Establishing Target Trip Tables

There are several elements of data contained in the census that have relevance to urban transportation planning. Following are some of the uses of census data in the application of urban transportation planning models (Sosslau, 1984; ITE 1991):

- Current socio-economic data, such as population, dwelling units, and income, can be used as input to determine current trip generation with existing models.
- Census data can serve as a benchmark for checking updated long-and short-range land use and socio-economic data.
- Journey-to-Work (JTW) census questions can yield responses that can be used as a secondary source for checking the validity of trip length frequency distributions, trip ends, and work trip tables.
- The data can be used for the calibration and development of urban transportation planning models.

A good source of census data for transportation planning applications is the 1990 Census Transportation Planning Package (CTPP) (FHWA, 1995). The CTPP is a collection of summary tables that contain information about population and household characteristics, worker characteristics, and characteristics of the Journey-to-Work (JTW). These tables have been designed specifically for transportation planning analyses.

The CTPP is organized into the following two elements: Statewide and Urban. The Statewide Element of the CTPP consists of data summaries for all places of 2,500 or more
population, the balance of the county, the county as a whole, and entirety of the state. The Urban Element of CTPP consists of data summaries for urbanized areas with a population of 50,000 or more. This data is available at the level of Traffic Analysis Zones (TAZ) or the Census tract. Data is also included for census-defined urbanized areas, the MPO defined study area, and the Metropolitan Statistical Area in which the MPO is located.

The CTPP can be used in the Urban Transportation Planning Process (UTPP) by utilizing census data to estimate new or calibrate pre-existing trip generation models based on socio-economic data and travel characteristics stratified by TAZ. Once available at the TAZ level, trip generation and trip distribution models can use the data to calculate trip tables.

A brief overview of the CTPP in the UTPP has been presented below. The use of CTPP data for carrying out the traditional trip generation and trip distribution steps of the four-step planning process has also been explained. In addition the use of CTPP for Pulaski study area has also been discussed.

### 3.4.2 CTPP Use in UTPP

Due to its inherent advantages in terms of availability and implementation, census data has been used for transportation planning by researchers and planners over the years. Based on similar context the data in CTPP can readily be employed for carrying out the traditional trip generation and trip distribution steps of the travel demand modeling process. The general approach for using CTPP in the four-step planning process has been explained below.

Most trip generation procedures use independent variables i.e., automobile ownership or income or both and/or the number of households or population or both. For trip-attractions, employment by industry is most often used. The relationship of the percentage of households by income and by car availability is derived from census data. From Parts 1 and 2 of the
CTPP, the JTW trips can be examined with reference to characteristics of the residential and employment ends of the trip. Trip generation estimates can easily be derived from trip end data, though the census geography of the CTPP is not fine-grained enough to produce site-level rates, as found in the Institute of Transportation Engineer’s Manual on Trip Generation (Cervero, 1994). From CTPP-1, work trip production rates can be estimated by indexing total daily vehicle trips in a zone to the number of the household or total acreage. At place of employment, work trip attraction rates, expressed in terms of total workers or acreage can be estimated for CBDs and large suburban employment centers (using CTPP-2).

The trip distribution models can be calibrated using the adjusted observed home-based work (HBW) person trip tables and network databases. CTPP can also be utilized to check a locally calibrated gravity model (Clarke, 1984).

3.5 Examination of CTPP for Use in the Pulaski Case Study

The use of CTPP data for the town of Pulaski was examined in the context of this research. Being a small town, Pulaski fell under the Statewide Element of CTPP. The available data was shown aggregated for the whole town, and was not amenable to the research approach adopted in this study. It was not possible to stratify the data in terms of TAZs, as was needed for this research. In short, the socio-economic data required at the TAZ level, was not available for Pulaski in the CTPP database. CTPP provided the census data in terms of TAZs, only for the urbanized areas with population of more than 50,000. Hence, even though CTPP would be the preferred choice for accomplishing the objective of establishing a target/seed table in this research approach, it could not be used for the Pulaski case study. Hence, investigation of other potential data sources were undertaken to obtain the relevant socio-economic/census data at the TAZ level, for carrying the travel modeling for the town of Pulaski.