Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
SWUTC/02/167523-1		

4. Title and Subtitle Uncertainty in Integrated Land Use-Transport Models: Simulation & Propagation	5. Report Date September 2002 6. Performing Organization Code
7. Author(s) Sriram Krishnamurthy and Kara Kockelman	8. Performing Organization Report No. Research Report 167523-1
9. Performing Organization Name and Address Center for Transportation Research University of Texas at Austin 3208 Red River, Suite 200 Austin, Texas 78705-2650	10. Work Unit No. (TRAIS) 11. Contract or Grant No. 10727
12. Sponsoring Agency Name and Address Southwest Region University Transportation Center Texas Transportation Institute Texas A&M University System College Station, Texas 77843-3135	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplementary Notes

Supported by general revenues from the State of Texas.

16. Abstract

This work examines the propagation of uncertainty in outputs of a standard integrated model of transportation and land use. Austin-calibrated DRAM-EMPAL predictions of residence and work locations are used as inputs to a four-step travel demand model (TDM), and the resulting travel times are fed forward into the future period's land use models. Covariance in inputs (including model parameters and demographic variables) was accommodated through multivariate Monte Carlo sampling of 200 scenarios. Variances in land use and travel predictions were then analyzed, over time, and as a function of input values. Results indicate that output variations were most sensitive to the exponent of the link performance function, the split of trips between peak and off-peak and several trip generation & attraction rates. 20 years in the future, final uncertainty levels (as measured by coefficients of variation) due solely to input and parameter estimation errors are on the order of 38% for total regional peak-period VMT, 45% for peak-period flows, and 50% and 37% for residential and employment densities, respectively.

17. Key Words	18. Distribution Statement
Uncertainty Outputs, Travel Demand Model, Land-Use, Covariance	No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	31	

Form DOT F 1700.7 **(8-72)** **Reproduction of completed page authorized**

Uncertainty in Integrated Land Use-Transport Models: Simulation & Propagation

By
Sriram Krishnamurthy
Kara Kockelman

Research Report SWUTC/02/167523-1

Southwest Region University Transportation Center

Center for Transportation Research

University of Texas at Austin

Austin, Texas 78712

September 2002

ABSTRACT

This work examines the propagation of uncertainty in outputs of a standard integrated model of transportation and land use. Austin-calibrated DRAM-EMPAL predictions of residence and work locations are used as inputs to a four-step travel demand model (TDM), and the resulting travel times are fed forward into the future period's land use models. Covariance in inputs (including model parameters and demographic variables) was accommodated through multivariate Monte Carlo sampling of 200 scenarios. Variances in land use and travel predictions were then analyzed, over time, and as a function of input values. Results indicate that output variations were most sensitive to the exponent of the link performance function, the split of trips between peak and off-peak and several trip generation & attraction rates. 20 years in the future, final uncertainty levels (as measured by coefficients of variation) due solely to input and parameter estimation errors are on the order of 38% for total regional peak-period VMT, 45% for peak-period flows, and 50% and 37% for residential and employment densities, respectively.

EXECUTIVE SUMMARY

Uncertainty in transportation systems has recently received some attention. And over the past 30 years many studies have highlighted uncertainty inherent in predictions of travel demand; however, very few have attempted to quantify this uncertainty. The difficulty is that not all sources of uncertainty are suitable for empirical analysis.

This work examines the propagation of uncertainty in outputs of a standard integrated model of transportation and land use. Austin-calibrated DRAM-EMPAL predictions of residence and work locations are used as inputs to a four-step travel demand model (TDM), and the resulting travel times are fed forward into the future period's land use models. Covariance in inputs (including model parameters and demographic variables) was accommodated through multivariate Monte Carlo sampling of 200 scenarios. Variances in land use and travel predictions were then analyzed, over time, and as a function of input values. Results indicate that output variations were most sensitive to the exponent of the link performance function, the split of trips between peak and off-peak and several trip generation & attraction rates. 20 years in the future, final uncertainty levels (as measured by coefficients of variation) due solely to input and parameter estimation errors are on the order of 38% for total regional peak-period VMT, 45% for peak-period flows, and 50% and 37% for residential and employment densities, respectively.

Simulation results indicate that the link performance parameter β_{link} is a key source of uncertainty in outputs; this is probably due to significant travel-time feedbacks to location decisions, which are fundamental to travel patterns. Population and employment growth rates only seem to have an effect in the long run. However, it should be emphasized that these results may be specific to ITLUP and such work should be done with other land use transportation models, to draw general conclusions on the impact of certain variables and parameters on uncertainty in other models' outputs.

Further work would be helpful for more fully understanding the growth in prediction uncertainties over time and across different model frameworks. Instead of random simulations, experiments could be performed by varying only one variable at a time (e.g., the population growth rate), and gauging its marginal impact on outputs. This can be time-consuming, but it certainly can assist in drawing crisper conclusions about the impacts of individual parameters and inputs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ACKNOWLEDGEMENTS

The authors recognize that support for this research was provided by a grant from the U.S. Department of Transportation, University Transportation Centers Program to the Southwest Region University Transportation Center which is funded 50% with general revenue funds from the State of Texas.

Propagation of Uncertainty in Transportation-Land Use Models: An Investigation of DRAM-EMPAL and UTPP Predictions in Austin, Texas

CHAPTER 1. INTRODUCTION

Uncertainty in transportation systems has recently received some attention (Pradhan and Kockelman, 2002; Niles and Nelson, 2001; Rodier and Johnston, 2001; Zhao and Kockelman, 2001). And over the past 30 years, many studies have highlighted uncertainty inherent in predictions of travel demand (e.g., Alonso, 1964; Aitken and White, 1972; Horowitz and Emsile, 1978; Mackinder and Evans, 1981; Mahmassani, 1984; Barton-Aschman et al, 1997), but very few have attempted to quantify this uncertainty. The difficulty, as Mahmassani (1984) explains, is that not all sources of uncertainty are suitable for empirical analysis.

A study by Mehndiratta et al. (2000) concluded that, most Metropolitan Planning Organizations (MPO's) do not quantify uncertainty even though they believe it to be pervasive in their planning process. The lack of legislative mandate dictating analysis of uncertainty and the increase in the complexity of the planning process, in case uncertainty analysis is performed were noted as the two common reasons for the reluctance on part of MPO's. Yet for optimal decision-making, an appreciation of uncertainty is critical. (See, e.g., Keeny and Raifa [1993].) Rodier and Johnston (2001) recently stressed that acknowledgement of uncertainty is necessary in order to maintain the credibility of transportation modeling process.

Pradhan and Kockelman (2002) argued that since many MPOs are now moving towards integrated land use-transportation models (largely in response to federal legislation), it is important to understand uncertainty propagation in these very complex models. They also suggested that reluctance of planners may be overcome by developing general results and a standard method of uncertainty analysis that can be applied to a variety of integrated models and settings.

This study furthers the work of Pradhan and Kockelman (2002) by calibrating the entire suite of models based on a single region's data set; this approach provides all parameter covariance matrices, thus allowing more realistic simulation. We investigate uncertainty propagation in an integrated model known as ITLUP (Putman, 1984), and we highlight the potential for errors in both land use and transportation model outputs.

1.1 Background

Uncertainty in travel demand predictions derives from a host of sources, including model misspecification, imperfect input information, and innate randomness in events and behaviors. Mahmassani (1984) described five sources of uncertainty that affect the evaluation of alternative transportation options. The first is uncertainty arising from political disturbances or major technological changes. The second class arises from uncertainty regarding the economic, political or social events and variables that affect the environment in which transportation system operates. Misspecification of the model and the uncertainty associated with the measurement of model parameters and inputs forms the third category. This category is most suitable for empirical analysis. The fourth class arises because of unclear evaluation criteria such as 'political desirability', which cannot be modeled using probability theory. The fifth category arises from the uncertainty regarding the basis of evaluation. This is very important because in most cases the planning process is critically dependent on the evaluation criteria.

Pradhan and Kockelman's (2002) investigations addressed only those sources of uncertainty that could be represented using probabilistic theory. The scope of their work, therefore, was limited to an examination of randomness in the predictions of land use-transportation models arising from uncertainty in model inputs and parameters. Essentially all studies have similarly limited the scope of their investigations (Zhao and Kockelman, 2001; Rodier and Johnston, 2001; Thompson et al, 1997; Harvey and Deakin, 1995). Because of the complexities that arise should any of the other categories of uncertainty be considered, and the difficulties associated with the quantification of their impacts, the current study also focuses on uncertain predictions arising from variability in model inputs and model parameters.

1.2 Computing Uncertainty

The method of moments and Monte Carlo simulation are two key methods for assessing the distribution of outputs, which are functions of random inputs. By relying on Taylor-series expansions of the output function, the method of moments approximates output moments using means, standard deviations, and other, higher-order moments of inputs.

Use of this method requires that outputs be specified as a clear, single function of inputs; this is extremely difficult, if not impossible, for almost all integrated model outputs (for reasons that will become clear in the model specification). Additionally, accuracy in approximation requires use of high-order derivatives, further complicating the analyses.

Monte Carlo techniques essentially draw input values from their multivariate distributions. These are used as model inputs, and their corresponding outputs are calculated. If inputs are drawn randomly, the resulting output values constitute a random sample from their respective probability distributions.

Due to the computational complexities inherent in integrated land use-transportation models (e.g., network equilibration), Monte Carlo methods were employed in this study. Such methods require substantial computer run time, and human input, but they produce more accurate and realistic results than method of moments. With 200 simulations of the full model, our outputs demonstrate most of their range; however, extreme events are rarely generated.

1.3 Uncertainty Analysis

There are several relatively simple but valuable techniques for examining the effects of input uncertainty on outputs (Morgan and Henrion, 1990). The most useful here is Multivariate sensitivity analysis, which helps identify the degree to which model outputs are affected by changes in inputs, while controlling for variations in other inputs. By eliminating dimensions or units, much like elasticities, standardized regression coefficients are very helpful here. After linearly regressing an output on inputs, a standardized regression coefficient SRC_i is calculated as follows:

$$SRC_i = \frac{\beta_i \times \sigma_i}{\sigma_y} \qquad (1)$$

This measures the number of standard deviations in the output one may expect from a single standard deviation's increase in input i. These coefficients will be used to report results from this work's sensitivity analyses, in a later section.

CHAPTER 2. DATA AND LAND USE MODEL DESCRIPTION

2.1 Data Description

The current application was undertaken for the Austin, Texas, region, which is mid-size metropolitan area of 700,000 people, spread across three counties (Hays, Travis and Williamson). The region is divided into 1074 traffic serial zones (TSZ's), on the basis of aggregations of census block groups; and the network is comprised of 16,966 links. The great majority of data used for calibrating the models was obtained from the Capital Area Metropolitan Planning Organization (CAMPO). These include population and employment by type and by zone, as well as 24-hour and two-hour peak travel times for 1997 (which were used to calibrate various land use and travel demand sub-models)

The 1990 Census of Population data also were used, in order to estimate the mean and standard deviation of household incomes by zone, as a function of Census-provided median incomes. The results of these regressions permitted predictions of household distributions across four income categories for each zone, these were then used in the land-use and trip-generation models.

Austin's 1996 Travel Survey (the ATS) was used to calibrate the four-step travel demand model (TDM). This survey consisted of all members above 5 years of age across 1939 households reporting all trips made on a single weekday. 20.1% of the trips were Home-Based Work, 49.3% were Home-Based Non Work, and 30.6% were Non Home-Based. 88.6% were by automobile, 6.1% were by transit, 3.8% were by walking, and 1% by bike.

2.2 Land Use Model Description

The current application of Putman's (1984) Integrated Transportation and Land-Use Package (ITLUP) consists of: a Disaggregate Residential Allocation Model (DRAM) and an Employment Allocation Model (EMPAL). Both are modified forms of Lowry's model (1964). DRAM uses the attractiveness of a zone and the accessibility of a zone's workers to jobs in other zones as the principal factors in allocating households to zones. EMPAL allocates employment based on the employment in the previous time period and the attractiveness of the zone for households.[1]

2.3 Disaggregate Residential Allocation Model

The original DRAM and EMPAL equations (Putman, 1984) have undergone several changes over the years, primarily due to data differences. In the current application households were allocated to zones based on the following DRAM formula:

$$\hat{N}_{i,t} = \sum_j E_{j,t} r_t \left[\frac{W_{i,t} e^{\beta_p c^p_{ji,t} + \beta_{op} c^{op}_{ji,t}}}{\sum_k W_{k,t} e^{\beta_p c^p_{ji,t} + \beta_{op} c^{op}_{ji,t}}} \right] \qquad (2)$$

where $\hat{N}_{i,t}$ is the estimated number of households in zone i at time t, $W_{i,t}$ is an attractiveness measure for zone i at time t, $c^p_{ji,t}$ is the peak travel time between zones j and i at time t, $c^{op}_{ji,t}$ and is the off-peak travel time between zones j and i at time t, r_t is the region-wide ratio of households per employee at time t, and both β_{op} and β_p are empirically derived parameters.

The attractiveness of a zone was calculated as follows:

$$W_{i,t} = (L_{i,t})^{\theta} \prod_{k=1}^{4} \left(1 + \frac{N_{i,t}^{k}}{\sum_{k=1}^{4} N_{i,t}^{k}} \right)^{\gamma_k} \quad (3)$$

where $L_{i,t}$ is the total land area of zone i at time t, $N_{i,t}^{k}$ is the number of residents in zone i who are in the k^{th} income quartile at time t, and θ and $\gamma_1 \ldots \gamma_4$ are empirically derived parameters.

There are 4 household quartiles defined by annual income and thus four versions of equations 2 and 3. Thus, there are 28 parameters used in the DRAM portion of the integrated model.

2.4 Employment Allocation Model

In this application the EMPAL equation used for allocating employment to zones is as follows:

$$E_{j,t} = \left(\frac{e^{\delta}}{1 + \mathrm{e}^{\delta}} \right) r_t^h \sum_i H_{i,t-1} \left[\frac{W_{j,t-1} e^{\beta_p c_{ki,t} + \beta_{op} t_{ki,t}}}{\sum_k W_{k,t-1} e^{\beta_p c_{ki,t} + \beta_{op} t_{ki,t}}} \right] + \left(\frac{1}{1 + e^{\delta}} \right) r_t^e E_{j,t-1} \quad (4)$$

where $E_{j,t}$ is the employment in the relevant sector (basic, retail, and service) in zone j at time t, $H_{i,t-1}$ is the number of households in zone i at time t, $W_{j,t-1}$ is the attractiveness function for zone j at time $t-1$, r_t^h is the region-wide ratio of employment at time t to households at time $t-1$, r_t^e is the region-wide ratio of employment at time t to employment at time $t-1$, and δ is an empirically derived parameter.

The attractiveness of a zone was calculated as follows:

$$W_{j,t-1} = (E_{j,t-1})^{\delta_1} \times (L_j)^{\delta_2} \quad (5)$$

where L_j is the total land area of zone j, and δ_1 and δ_2 are empirically derived parameters.

There are 3 sectors defined by employment type and thus three versions of equations 4 and 5. Thus, there are 15 parameters used in the EMPAL portion of the integrated model.

The DRAM model parameters were calibrated using the 1997 household and employment allocations across Austin's zones. Households were categorized into four groups based on estimates of average zonal annual income and its standard deviation (both as a function of available median income data); these groups were defined as having an annual household income less than or equal to $30,000, between $30,000 and $42,500, between $42,500 and $55,000, and greater than $55,000.

EMPAL's parameters were calibrated using the 1997 household data, as well as 1997 and CAMPO-predicted 2007 employment data. The implicit assumption of this strategy is that EMPAL responds first to prior employment and household allocations and then households respond to current employment allocations. Employment was categorized into three types: basic

(Standard Industrial Classification (SIC) code 1-5199), retail (SIC code 5200-5999) and service (SIC code 6000-9799). Maximum likelihood estimation of all the parameters was performed using Gauss software (version 3.2.34 [Aptech, 1998]). The parameter estimates and their associated t-statistics are shown in Table 1. All the parameter values are consistent with expectations. The positive peak travel time parameters suggest that households and jobs are attracted to regions, which are congested during the peak period but this effect is compensated to a slight extent off-peak effects.

The standard ITLUP formulation has several limitations. A major one is that it does not account for land use intensity constraints. ITLUP models assign jobs and households to zones even if they do not have the capacity to accommodate more jobs or households. The current application deals with this limitation by reallocating excess allocation to zones which have the area to absorb more jobs and households. Maximum allowable residential and commercial densities used in this work are 25 households per residential acre and 100 jobs per commercial acre. These two maximum densities were calculated based on an envelope analysis of CAMPO's 1995 land use data set. Unfortunately, the land use data were only available for 549 out of the 1074 TSZ's, so a multinomial logit model for fraction of land by category (residential, commercial, and vacant), as a function of zonal area and network distance to the CBD, was developed to estimate the land use distribution in the remaining zones in the year 1995. The household and job allocations obtained from EMPAL and DRAM are first used to fill up the commercial and residential area in a zone. If, the numbers require more area, any vacant area is allocated to jobs and households in proportion of their demand for land. If a zone cannot accommodate more development, that zone is removed from consideration and the remaining jobs and/or households are allocated to the other zones by re-running the DRAM and EMPAL models.

Another limitation of ITLUP is that DRAM and EMPAL models are applied sequentially, neglecting simultaneous interactions between jobs and households. In addition, ITLUP does not consider land prices and commodity flows in allocating jobs and households; so it lacks important relationships and variables of great interest to planners, policymakers, and the public. However, its relative simplicity permits a fairly transparent analysis of uncertainty. In contrast, for example, Waddell et al.'s UrbanSim (2000) requires on the order of 1,000 parameters and tens of thousands of input values that extremely few regions possess (such as development history and land prices by hectare). Analysis of uncertainty under that model was performed by Pradhan and Kockelman (2002), but only with limited input combinations and reliance on given parameter sets (without covariance matrices).

---

[1] DRAM and EMPAL specifications may appear somewhat counter to one's general expectations. For example, households are assumed to locate where employers would be most attracted to them (rather than in locations where they are most attracted to employers). And jobs locate in positions where households would find them most accessible (rather than in positions where they find households to be most accessible). Both perspectives make sense. Our work aims to follow the traditional paradigm as closely as possible.

CHAPTER 3. TRAVEL DEMAND MODEL

A traditional four-step travel demand model (TDM) was adopted to link the land use allocations of jobs and households to the Austin transportation network. The trip purposes considered in this application are Home-Based Work (HBW), Home-Based Non Work (HBNW) and Non Home-Based (NHB) trips. Details of the four basic model steps follow here.

3.1 Trip Generation

Linear regression models for trip production and attraction were developed based on the 1996 ATS data. Trip production models for HBW and HBNW trips were developed at the household level, whereas the trip production model for NHB trips was developed at the zonal level.

The model specifications are given below:

$$P_{HBW} = f(Inc_1, Inc_2, Inc_3, Inc_4)$$

$$P_{HBNW} = f(Inc_1, Inc_2, Inc_3, Inc_4)$$

$$P_{NHB} = f(Basic, Retail, Service)$$

where P_{HBW} and P_{HBNW} are the number of person trips (per day) produced by a household, and P_{NHB} are the numbers of NHB trips produced/generated by a zone. Inc_i is the indicator variable for the i^{th} income category, and *Basic*, *Retail* and *Service* are the numbers of these jobs in a zone.

Trip-attraction models also were developed at the zonal level, and their general functional specifications are the same as for the production of NHB trips (i.e., as a function of numbers by basic, retail, and service jobs per attractive zone). The parameter estimates for both production and attraction models are shown in Table 2.

3.2 Trip Distribution

Multinomial logit models of destination choice were calibrated for each trip purpose and for peak and off-peak times (Ben Akiva and Lerman, 1985). The natural log of the total number of attracted trips (as estimated using Table 2's values) and travel times were used as explanatory variables. The coefficient on the total number of trips term attracted was constrained to equal one, so that the model form is the same as a gravity model; and size is accommodated proportionally, which is consistent with probabilistic theory. The model specification is shown below:

$$p_{ij} = \frac{A_j e^{\beta_1 t_{ij}}}{\sum_k A_k e^{\beta_1 t_{ik}}} \qquad (6)$$

where p_{ij} is the percentage of trips produced in zone i that are attracted to zone j, A_j is the number of person-trips attracted to zone j, and t_{ij} is the travel time from zone i to zone j (for both peak and off-peak periods, depending on the time of day). Calibration results are given in Table 3.

3.3 Time of Day Characteristics

Peak and off-peak times of day were used in this study. According to CAMPO, the peak period lasts two hours in the morning (from 7:15 a.m. to 9:15 a.m.). The trip production-attraction (PA) matrices were converted to origin-destination (OD) matrices by time of day (peak and off-peak) using departure and arrival rates during the peak period, as shown in equation 7.

$$HBNW_{OD} = HBNW_{PA} \times \Psi_{HBNW} + HBNW_{PA}^{T} \times \Theta_{HBNW} \tag{7}$$

where Ψ_{HBNW} is the departure rate for HBNW trips in the peak period, Θ_{HBNW} is the return rate for HBNW trips in the peak period. Similar equations were used to calculate the OD matrices for HBW and NHB trips. These rates were calculated from the 1996 ATS data.

3.4 Mode Split

Binary logit models of mode choice were calibrated to assign person-trips to automobile and non-automobile modes. Six different models were estimated: one for each trip purpose and time of day. But each model form's probability for auto choice is structured the same, as follows:

$$p_{auto} = \frac{e^{\alpha_{auto} + \beta_m t_{ij}^{auto}}}{e^{\alpha_{auto} + \beta_m t_{ij}^{auto}} + e^{\beta_m t_{ij}^{non-auto}}} \tag{8}$$

The parameters and goodness-of-fit measures for these 6 mode-choice (MC) models are given in Table 3.

All parameter estimates are consistent with expectations. But the low goodness-of-fit measures suggest that the preference for automobile use is independent of the CAMPO-provided travel times (which may be measured with significant error). In other words, simply a mode-specific constant will do almost as well in explaining mode choice as the available travel time data. As a result of this, the β_m was not found to be important in sensitivity analysis, as described later in this paper.

3.5 Vehicle Occupancy

Vehicle occupancy for each trip purpose was calculated from the 1996 ATS data. Average vehicle occupancy levels of 1.20, 1.99, and 1.85 were for HBW, HBNW, and NHB trips, respectively.

3.6 Traffic Assignment

All automobile trips were assigned to the Austin network using the Stochastic User Equilibrium (SUE) (Sheffi, 1982) assignment method available in TransCAD (Caliper 2001). The following settings were chosen for the assignment: a maximum of 30 iterations, a probit route-choice model with 5% error, and a convergence criterion of 0.01.[2]

The link performance function is of the following type (based on the original Bureau of Public Roads (1964) formula, but using different parameter values for α and β_{link}):

$$t = t_f \left[1 + \alpha \left(\frac{v}{c} \right)^{\beta_{link}} \right] \tag{9}$$

where t_f is the free-flow travel time, and α and β_{link} are link performance parameters, provided by CAMPO. There are 8 sets of these parameters, depending on road type. No information on these parameters' uncertainty was available, so they were assumed to have standard deviations that are 30% of their given values and assumed to follow independent normal distributions.

3.7 Background Flow

Austin's 43 external zone's trip counts, as obtained from CAMPO's 2007 OD matrix predictions, was used to load background flows on to the network. For future model applications, these trip values were assumed to grow at the (randomly drawn) population growth rate.

---

[2] Only 30 iterations were performed due to the time required. An average of 18 minutes was needed for each traffic assignment. Two assignments (peak and off-peak) were done for each run, and four runs were performed for each of the 200 simulations (i.e., for the future years 2002, 2007, 2012, and 2017), Thus, on average 2.4 hours were required to perfom essentially just the traffic assignment step for each of the 200 simulations. A convergence criterion of 1% refers to the maximum percentage flow difference between successive iterations. An error of 5% represents the percentage error for the error term used in stochastic user equilibrium assignment.

CHAPTER 4. RESULTS

4.1 Simulations

200 full Monte Carlo simulations of input sets (including both model parameters and starting distributions of jobs and households, by type and zone) were performed. The demographic variables varied were population and employment growth rates. The means and the standard deviations of these rates were taken to be 3.3% +/- 0.5% (population growth rate), and 3.1% +/- 0.5% (employment growth rate) [3](CAMPO, 2002).

All model parameters were varied assuming multivariate normal distributions, based on their estimated variance-covariance matrices (an output of the software codes used to calibrate the various ITLUP and TDM submodels). The link-performance parameters, and the peak/off-peak splits were also varied, assuming independent normal distributions and coefficients of variation (mean divided by standard deviation) of 0.3.

In all 95 parameters and 2 demographic variables were varied, and 200 simulations were performed. To observe the evolution of uncertainty over time, the land use and travel demand model were run every 5 years, and four such runs were performed for each simulation, for a total of 20 years of forecasts of population, employment, and travel, across the region.

4.2 Results

Sensitivity analysis was performed by regressing various key outputs of the integrated model on input parameters. Standardized coefficients and p-values were used to gauge the practical and statistical impact of variables on the model outputs. The outputs considered for analysis were weighted residential density, weighted commercial density, VMT and VHT for both peak and off-peak periods and link flows during peak and off-peak periods.

The final model specifications were obtained by removing the variables, which were not significant in any model. Different combinations of the remaining variables were considered to arrive at the specifications shown in Tables 4, 5, and 6.

The results for off-peak and peak VMT and VHT are shown in Table 4. The results indicate that the employment and population growth rates do not impact VMT and VHT in a significant way. However, population and employment totals are very important in predicting VHT and VMT The exponent (β_{link}) on the volume-to-capacity term in the link performance function is the most significant variable for predicting both peak and off-peak VHT. This makes sense, since VHT is very dependent on link travel times. The coefficients in the trip production and trip attraction models are also highly significant.

Peak-period VHT and VMT are estimated to be highly sensitive to peak-off peak splits, which is consistent with our expectation. Peak and off-peak VMTs also are sensitive to α_i, the coefficient on the volume-to-capacity term in the link performance function.

The three links chosen for analysis are: (1) IH35 Northbound, near Cameron Road; (2) Loop 1Southbound, south of 5th and 6th Streets; and (3) IH35 Southbound, south of US290 and SH 71. These are critical links for the network and each carried close to maximum flowin each simulation. Averages of flows on these three links were used to perform sensitivity analysis. The link flow results shown in Table 6 indicate that β_{link} is the most significant parameter in the initial years for off-peak period link flows. In the long term, however population growth rate has a significant impact on the link flows. As with peak period VHT and VMT, peak period link flows are also significantly dependent on peak-off peak splits.

The results from land use models considered for analysis were population and jobs weighted averages of residential and commercial densities. The regression results are shown in Table 5, results indicate that the average commercial density is significantly influenced by the parameters in DRAM and EMPAL models, and trip production and attraction rates. The weighted residential density is significantly impacted by the parameters in the EMPAL model, trip production and attraction rates and mode choice model parameters.

The evolution of uncertainty over time is shown in Figure 1. Since VHT is the final output from the model it is not surprising that maximum uncertainty is associated with VHT, since uncertainty gets compounded across sub models. Peak-period outputs show greater variability as compared to off-peak outputs. This might be due to the effects of congestion, since higher congestion implies greater travel times and hence more response in exponential fashion. These findings are consistent with Pradhan and Kockelman's investigations of ITLUP for Eugene-Springfield, (Pradhan 2001; Pradhan and Kockelman 2002).

---

[3] The mean growth rates were obtained from CAMPO's newsletter, and standard deviation values were assumed.

CHAPTER 5. CONCLUSIONS

This work investigated the dependence of future location and travel choice predictions on integrated-model inputs and parameters. Results indicate that output variations were most sensitive to the exponent of the link performance function, the split of trips between peak and off-peak periods, and several trip generation and attraction rates. 20 years in the future, final uncertainty levels (as measured by coefficients of variation) due solely to input and parameter estimation errors were found to be on the order of 38% for total regional peak-period VMT, 45% for peak period flows, and 50% and 37% for residential and employment densities, respectively.

This work builds on Zhao and Kockelman's (2002) investigations of four-step travel demand models and Pradhan and Kockelman's (2002) investigations of integrated models by adding realism to parameter distributions (through controlled calibrations) and model specification (through detailed submodel assembly and an integration of land use and travel behaviors). In contrast to Zhao and Kockelman's work, it examples the evolution of prediction uncertainties over time and across model stages, and travel conditions are permitted to impact location choices. Due to these distinctions, simulation results indicate that the link performance parameter β_{link} is a key source of uncertainty in outputs; this is probably due to significant travel-time feedbacks to location decisions, which are fundamental to travel patterns. Population and employment growth rates only seem to have an effect in the long run. However, it should be emphasized that these results may be specific to ITLUP; in UrbanSim Pradhan and Kockelman (2002) observed that demographic inputs were principal sources of uncertainty in the short and long terms.

The time-consuming nature of the simulations and a lack of data affected the investigations. For example, the lack of employment data for two past time periods required we use 2007's predicted employment data for calibrating the EMPAL models, which could have some bearing on the parameter estimates and sensitivity analysis. Also, correlation information for certain variables (like peak/off-peak split and link-performance parameters) were not available, so these were assumed. External trips were not modeled explicitly; instead we relied on CAMPO's estimates to load the network with growing background flows. And vehicle occupancy rates were not randomly varied.

Further work would be helpful for more fully understanding the growth in prediction uncertainties over time and across different model frameworks. Instead of random simulations, experiments could be performed by varying only one variable at a time (e.g., the population growth rate), and gauging its marginal impact on outputs. This can be time-consuming, but it certainly can assist in drawing crisper conclusions about the impacts of individual parameters and inputs. Also, such work should be done with other land use transportation models, to draw general conclusions on the impact of certain variables and parameters on uncertainty in other models' outputs.

REFERENCES

Aptech Systems Inc. 1996. GAUSS Maximum Likelihood Estimation Module. Aptech Systems Inc., Maple Valley, Washington.

Ben-Akiva, M. and S.Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, Massachusetts: MIT Press.

Bureau of Public Roads. 1964. *Traffic Assignment Manual*. Washington D.C.

Caliper Corporation. 2001. Travel Demand Modeling With TransCAD 4.0. Caliper Corporation, Newton, Massachusetts.

CAMPO. 2002. New Population and Employment Forecasts. CAMPO, Austin, Texas.

Keeney, R.L, and H. Raiffa. 1993. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* Cambridge, Cambridge University Press.

Mahmassani, H. S. 1984. Uncertainty in Transportation Systems Evaluation:Issues and Approaches, *Transportation Planning and Technology,* Vol. 9, pp. 1-12.

Mehndiratta, S. R., D. Brand, and T. E. Parody. 2000. How Transportation Planners and Decision Makers Address Risk and Uncertainty. *Transportation Research Record 1706,* TRB, National Research Council, Washington, D.C., pp. 46-53.

Morgan, M.G, and M.Henrion 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press, Cambridge.

Niles, J. S. and D. Nelson. 2001. Identifying Uncertainties in Forecasts of Travel Demand. Preprint for the Transportation Research Board's 80th Annual Meeting, Washington D.C.

Pradhan, A 2001. *Uncertainty Propagation in Land Use-Transportation Models: An Investigation of UrbanSim and ITLUP*. Masters Thesis Report, The University of Texas, Austin.

Pradhan, A., and K. Kockelman. 2002. Uncertainty Propagation in an Integrated Land Use-Transport Modeling Framework: Output Variation via UrbanSim, forthcoming in *Transportation Research Record.*

Pradhan, A., and K. Kockelman. 2002. Uncertainty Propagation in Land Use-Transportation Models, Proceedings of the 13th Euro Mini Conference on Handling Uncertainty in the Analysis of Traffic and Transportation Systems, Bari, Italy.

Putman, S. 1983. *Integrated Urban Models: Policy Analysis of Transportation and Land Use*. Pion, London.

Rodier, C. J. and R. A. Johnston. 2001. Uncertain Socioeconomic Projections Used in Travel and Emission Models: Could Plausible Errors Result in Air Quality Nonconformity? Preprint for the Transportation Research Board's 80th Annual Meeting, Washington D.C.

Sheffi, Y. 1984. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programing Methods.* Engelwood Cliffs, New Jersey.

Sheffi, Y. and W. Powell. 1982. An Algorithm for the Equilibrium Assignment Problem with Random Link Times. *Networks* 12, pp. 191-207.

Waddell. P., A. Borning, M. Noth, N. Freier, M. Becke, and G. Ulfarsson. 2001. UrbanSim: A Simulation System for Land Use and Transportation. (Accessed at the UrbanSim Website: http://www.urbansim.org/Papers/UrbanSim_NSE_Paper.pdf).

Zhao, Y., and K. M. Kockelman. 2002. The Propagation of Uncertainty Through Travel Demand Models: An Exploratory Analysis, *Annals of Regional Science* 36 (1).

Table 1. DRAM and EMPAL Model Estimation Results.

	Group 1	Group 2	Group 3	Group 4	Basic	Retail	Service
β_p	0.0490 (20.3)	0.0264 (7.55)	0.018 (2.72)	0.0170 (8.92)	0.0937 (18.2)	0.0980 (19.4)	0.0322 (14.8)
β_{op}	-0.3782 (-94.6)	-0.2859 (-56.5)	-0.236 (-28.9)	-0.1842 (-102)	-0.2781 (-29.4)	-0.3078 (-38.1)	-0.1308 (-35.9)
θ	0.6548 (157)	0.6340 (104)	0.5985 (96.6)	0.5491 (112)			
γ_1	71.49 (47.2)	56.33 (30.9)	44.81, (13.70)	12.60 (0.999)			
γ_2	58.16 (54.7)	49.19 (34.8)	35.96 (14.02)	14.13 (1.55)			
γ_3	13.21 (34.5)	17.16 (35.1)	21.02 (34.1)	14.67 (12.9)			
γ_4	67.19 (45.9)	54.35 (30.6)	44.22 (13.8)	18.16 (1.48)			
δ					-1.6350 (-82.5)	-1.5762 (-64.5)	0.633 (-33.44)
δ_1					0.2743 (36.7)	0.2018 (19.9)	0.4056 (74.9)
δ_2					0.8725 (49.7)	1.2590 (56.0)	0.5969 (56.8)
Log Lik	731.6	272.3	238.4	801.2	1618	617.6	1698
LRI	0.556	0.531	0.511	0.513	0.941	0.856	0.907
N_{obs}	1074	1074	1074	1074	1074	1074	1074

Table 2. Trip Production and Attraction Model Parameters.

Production Models			
	HBW	**HBNW**	**NHB**
Inc_1	1.325 (21.32)	4.049 (25.3)	
Inc_2	2.092 (23.6)	4.423 (19.4)	
Inc_3	2.115 (26.8)	4.881 (24.0)	
Inc_4	2.443 (30.4)	5.547(26.8)	
Basic			0.4621 (4.12)
Retail			5.186 (12.5)
Service			1.481 (10.0)
Adj. R^2	0.559	0.525	0.538
N_{obs}	1939	1939	586
Attraction Models			
Basic	0.4892 (8.79)		0.4093 (3.33)
Retail	1.668 (7.26)	5.419 (8.41)	5.659 (12.4)
Service	1.020 (12.5)	1.723 (7.69)	1.287 (7.94)
Adj. R^2	0.526	0.303	0.491
N_{obs}	613	665	568

Table 3. Destination and Mode Choice Model Parameters.

Destination Choice Models						
	HBW		HBNW		NHB	
	Off Peak	Peak	Off Peak	Peak	Off Peak	Peak
β_1	-0.0804 (-22.3)	-0.0637 (-22.4)	-0.1449 (-42.9)	-0.1201 (-40.0)	-0.1401 (-44.6)	-0.1299 (-28.4)
Adj. R^2	0.374	0.349	0.500	0.515	0.446	0.475
N_{obs}	971	1037	2571	2001	3307	1142
Mode Choice Models						
$\alpha_{\alpha\upsilon\tau o}$	1.944 (6.15)	1.935 (6.61)	1.421 (11.4)	1.442 (8.75)	2.011 (9.36)	2.046 (5.25)
β_m	-0.0225 (-2.58)	-0.0306 (-2.99)	-0.0091 (-2.62)	-0.0004 (-0.08)	-0.0216 (-3.12)	-0.0264 (-1.88)
Adj. R^2	0.0154	0.0182	0.0022	0.0015	0.0084	0.0086
N_{obs}	1085	1145	3154	1374	2507	792

Table 4. VMT and VHT Regression Results.

	Offpeak VHT		Peak VHT	
Parameters	**2017**		**2017**	
	Std Coeff.	p - value	Std Coeff.	p - value
Emprate	0.043	0.497	0.069	0.281
Poprate	0.091	0.159	-0.02	0.753
β_{link}	0.284	0	0.291	0
α_6	0.317	0		
Ψ_{HBNW}			0.323	0
Ψ_{HBW}			0.193	0.003
θ (Group 4)	0.232	0.008		
Inc_3 (HBNW)	-0.117	0.078		
Inc_3 (HBW)			0.101	0.111
Basic (NHB Prod)	-0.15	0.026		
Retail (HBNW)			-0.162	0.041
Service (HBNW)			-0.168	0.034
Retail (HBW)	0.18	0.007		
Basic (NHB Attr)	-0.12	0.066	-0.127	0.094
Adj. R^2	0.233		0.278	
	Offpeak VMT		**Peak VMT**	
Emprate	-0.147	0.198	0.011	0.87
Poprate	0.196	0.002	0.098	0.15
β_{link}	-0.072	0.001		
α_6	0.079	0		
α_8			0.592	0.008
Ψ_{HBNW}			0.298	0
Inc_3 (HBNW)	-0.147	0.025		
Retail (HBNW)	2.004	0.19		
Basic (HBW)			-0.123	0.076
Retail (HBW)	0.196	0.003		
Basic (NHB Attr)	0.142	0.016		
β_1 (HBNW Offpeak)	-0.27	0.005		
Adj. R^2	0.235		0.185	
N_{obs}	200		200	

Table 5. Weighted Residential and Commercial Density Regression Results

	Weighted Commercial Density		Weighted Residential Density	
Parameters	**2017**		**2017**	
	Std Coeff.	p - value	Std Coeff.	p - value
Emprate	0.037	0.549	0.059	0.394
Poprate	0.048	0.447	0.031	0.661
α_2	0.156	0.017		
α_3	0.099	0.113		
Ψ_{NHB}	-0.094	0.134		
θ (Group 1)	0.179	0.004	0.122	0.091
γ_1 (Group 3)	-3.293	0	-1.152	0.166
γ_3 (Group 3)	0.21	0.021		
γ_4 (Group 3)	3.081	0	0.998	0.231
θ (Group 4)	0.24	0.02		
γ_3 (Group 4)	0.746	0.001		
γ_4 (Group 4)	-0.758	0.001		
β_{op} (Group 4)	0.176	0.219		
β_p (Group 4)	0.383	0.031		
δ_2 (Basic)			0.181	0.041
δ_1 (Basic)			0.209	0.071
β_{op} (Basic)	-0.174	0.066	-0.266	0.05
δ_1 (Retail)	0.222	0.011	0.14	0.153
β_{op} (Retail)	-0.189	0.052	-0.208	0.056
β_p (Retail)	-0.282	0.001	-0.142	0.123
δ (Basic)	-0.18	0.056	-0.156	0.142
δ_2 (Service)			-0.124	0.139
β_{op} (Service)	0.073	0.231	-0.165	0.045
Inc_3 (HBNW)			0.157	0.023
Inc_1 (HBW)			0.097	0.151
Retail (HBW)	0.21	0.001	0.138	0.056
Service (HBW)			0.121	0.092
Retail (NHB)	0.09	0.158	0.188	0.009
Basic (NHB Attr)	0.17	0.024		
Service (NHB Attr)	0.122	0.096		
B_1 (HBW Offpeak)			0.11	0.111
α_{auto} (HBNW Offpeak)	0.156	0.013	0.126	0.068
α_{auto} (NHB Offpeak)			-0.219	0.02
β_m (NHB Offpeak)			0.189	0.044
α_{auto} (HBW Peak)	-0.232	0.01	-0.247	0.01
β_m (HBW Peak)	0.293	0.001	0.247	0.011
Adj. R^2	0.32		0.171	

Table 6. Average Flow Regression Results.

	Offpeak Average Link Flow		Peak Average Link Flow	
Parameters	**2017**		**2017**	
	Std Coeff.	p - value	Std Coeff.	p - value
Emprate	0.106	0.107	-0.003	0.962
Poprate	0.255	0	0.113	0.102
β_{link}	0.156	0.022		
Ψ_{HBNW}			0.252	0
Ψ_{HBW}			0.24	0
Basic (NHB Prod)	-0.113	0.09		
Retail (HBW)	0.205	0.002		
β_m (HBW Offpeak)	-0.112	0.093		
α_{auto} (NHB Offpeak)	-0.196	0.03		
α_{auto} (HBW Peak)	-0.117	0.079		
Adj. R^2	0.216		0.239	
N_{obs}	200		200	

Figure 1. Evolution of Uncertainty Over Time (TDM+LUM) Outputs.