

**Technical Report Documentation Page**

1. Report No. SWUTC/07/167262-1	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle <b>Microsimulation of Household and Firm Behaviors: Coupled Models of Land Use and Travel Demand in Austin, Texas</b>		5. Report Date <b>December 2007</b>	
7. Author(s) <b>Saurabh Kumar and Kara M. Kockelman</b>		6. Performing Organization Code	
9. Performing Organization Name and Address <b>Center for Transportation Research University of Texas at Austin 3208 Red River, Suite 200 Austin, Texas 78705-2650</b>		8. Performing Organization Report No. <b>167262-1</b>	
10. Work Unit No. (TRAIS)		11. Contract or Grant No. <b>10727</b>	
12. Sponsoring Agency Name and Address <b>Southwest Region University Transportation Center Texas Transportation Institute Texas A&amp;M University System College Station, Texas 77843-3135</b>		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes <b>Supported by general revenues from the State of Texas.</b>			
16. Abstract Households and firms are key drivers of urban growth, yet models for forecasting travel demand often ignore their dynamic evolution and several key decision processes. An understanding of household and firm behavior over time is critical in anticipating urban futures and addressing transportation, land use and other concerns. Birth and death, migration and location choice are defining events in a household's and firm's life cycle, and a study of household and firm evolution requires the estimation and application of models for each of these. Such an exercise is hindered primarily by a lack of quality micro-data. This study develops a basic framework for modeling household and firm demographics using microsimulation. Year 2005 zonal household population and employment point data for the Austin, Texas region, coupled with various, more aggregate data sets, are used to simulate household and firm evolution over time and space.			
To ensure a jobs-worker balance, the model may well merit greater synchronization of the population and firm synthesis models. The simulations also suggest a clear shift of firms and households towards more central zones, in part because of the cross-sectional nature of the data sets used to calibrate the location choice models and the lack of density restrictions or other reflections of land-availability constraints on new development. Essentially, households and firms exhibit a strong centralizing tendency, that Austin's land market simply cannot allow, due to space and other constraints on new building. Explicit expressions of such constraints should prove helpful in future implementations of this work.			
17. Key Words <b>Microsimulation, Firm Behaviors, Land Use, Travel Demand, Austin, Texas</b>		18. Distribution Statement 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) <b>Unclassified</b>	20. Security Classif.(of this page) <b>Unclassified</b>	21. No. of Pages <b>147</b>	22. Price



**Microsimulation of Household and Firm Behaviors:  
Coupled Models of Land Use and Travel Demand in Austin, Texas**

**By**

Saurabh Kumar  
Kara M. Kockelman

**Research Report SWUTC/07/167262-1**

Southwest Region University Transportation Center  
Center for Transportation Research  
University of Texas at Austin  
Austin, Texas 78712

December 2007



## **Abstract**

Households and firms are key drivers of urban growth, yet models for forecasting travel demand often ignore their dynamic evolution and several key decision processes. An understanding of household and firm behavior over time is critical in anticipating urban futures and addressing transportation, land use and other concerns. Birth and death, migration and location choice are defining events in a household's and firm's life cycle, and a study of household and firm evolution requires the estimation and application of models for each of these. Such an exercise is hindered primarily by a lack of quality micro-data. This study develops a basic framework for modeling household and firm demographics using microsimulation. Year 2005 zonal household population and employment point data for the Austin, Texas region, coupled with various, more aggregate data sets, are used to simulate household and firm evolution over time and space.

To ensure a jobs-worker balance, the model may well merit greater synchronization of the population and firm synthesis models. The simulations also suggest a clear shift of firms and households towards more central zones, in part because of the cross-sectional nature of the data sets used to calibrate the location choice models and the lack of density restrictions or other reflections of land-availability constraints on new development. Essentially, households and firms exhibit a strong centralizing tendency, that Austin's land market simply cannot allow, due to space and other constraints on new building. Explicit expressions of such constraints should prove helpful in future implementations of this work.

## Executive Summary

While microsimulation of urban systems is data and computing intensive, it provides a flexible tool for analyzing the impacts of various policy decisions as well as other, demographic, environmental and system changes. It allows transportation planners to explore the potential responses of individuals to changes in their environments and predict the long-term implications of policy decisions. This work seeks to be a bridge by providing a framework for further integrated travel demand and land use models of this type.

This study estimated a microscopic simulation model using various available data sets for the Austin, Texas region. The model specified here is used to anticipate the evolution of households and firms in Austin, Texas over a 30-year period (from 2005 to 2035). A microsimulation approach is used to track 10% of households (scaled up to 100%) and 100% of firms over time and space. The simulation approach seeks to tie evolutionary models of households and firms with models of travel behavior to provide robust forecasts. Models of life-cycle transitions for households and firms have been estimated using various available data sets. Household and firm simulation is carried out every year while travel demand models are run every 5 years, in order to update network travel times.

The model consists of household evolution, firm evolution, location choice and travel demand models. Household and firm simulation models are run at one-year time steps, in order to forecast Austin's future. The household simulation component is made up of models for birth (of children and of households), death of individuals (and other forms of household dissolution), migration, marriage and divorce decisions, children leaving home, vehicle ownership, and location choice. These models are estimated using multinomial logit and Poisson specifications. The firm simulation component consists of firm birth, death, growth and location choice models. A Markovian process is assumed in order to anticipate firm growth and contraction (across firm-size categories), along with logit and Poisson model specifications for firm location choice. Firms are categorized based on number of jobs (6 categories) and industry sector (4 sectors) they belong to.

Austin's household and commercial vehicle travel survey data were used to estimate trip generation and distribution models. Simulation results for multiple growth-rate scenarios suggest a roughly 180% increase in the Austin population over a 30-year period, 210% increase in vehicle ownership, a 230% increase in jobs, and more than a 300% increase in vehicle-miles traveled. When a 10-cent/mile flat-rate toll is applied over all links, the year 2035 VMT is predicted to be just 3% less than under the no-toll scenario. A fixed toll of 10-cents-per-mile shows a very low impact on VMT over a 30 year period than expected.

The system's travel demand model is limited by several simplifying assumptions, including the neglect of trip scheduling (i.e., time-of-day) and mode choice. Traffic assignment is run only once every 5 years, so that travel times remain unchanged in between, though location patterns, households and firms continue to evolve in those periods.

This study offers a basic simulation framework for integrated microscopic simulation models of household and firm evolution, land use and travel demand. It also provides a valuable demonstration of how a wide variety of behavioral models can be used in concert, and how individual firms and households can be tracked over several decades, using standard software and computing resources. Enhancements are coming, and opportunities for many interesting and relevant policy and growth scenarios are feasible.

## TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION .....	1
CHAPTER 2: LITERATURE REVIEW .....	3
2.1 Microsimulation .....	3
2.1.1 UrbanSim .....	4
2.1.2 ILUTE .....	5
2.2 Household and Firm Evolution .....	5
2.2.1 Firmography .....	6
2.2.2 Firm Location Choice .....	6
2.2.3 Commercial Trip Models .....	7
2.2.4 Household Demography .....	8
2.2.5 Household Location Choice .....	9
2.3 Models of Vehicle Ownership .....	9
2.4 Models of Long-Distance Travel .....	10
2.5 Summary .....	10
CHAPTER 3: DATA DESCRIPTION AND MODEL FORMULATION .....	11
3.1 Household Data Sets .....	11
3.2 Firm Data Sets .....	13
3.3 Simulation Framework .....	16
3.3.1 Population Synthesis .....	17
3.3.2 Household Transition Models .....	17
3.3.3 Model of Household Location Choice .....	23
3.3.4 Vehicle Ownership Model .....	25
3.3.5 Firm Synthesis and Evolution .....	28
3.3.6 Firm Location Choice Model .....	33
3.3.7 Microsimulation Model of Firms and Households .....	36
3.3.8 Long Distance Travel Model .....	37
3.3.9 Household Trip Generation and Distribution Model .....	38
3.3.10 Commercial Trip Generation and Distribution Models .....	39
3.4 Summary .....	41
CHAPTER 4: MODEL APPLICATION IN AUSTIN, TEXAS .....	43
4.1 Firm Evolution .....	44
4.2 Household Evolution .....	55
4.3 Travel Demand Model .....	58
4.4 Summary .....	60
CHAPTER 5: SUMMARY, CONCLUSIONS AND FUTURE WORK .....	61
5.1 Limitations and Enhancements .....	62
5.1.1 Computation-time .....	62
5.1.2 Household and Firm Evolution Model .....	63
5.1.3 Vehicle Ownership Model Limitations .....	64

5.1.4 Travel Demand Model Limitations .....	64
REFERENCES .....	65
APPENDIX A: MATLAB Microsimulation Code for Household and Firm Evolution ...	85
APPENDIX B: FIRM LOCATION CHOICE MODELS .....	135

## LIST OF FIGURES

Figure 3.1: Age Profile of Austin Population in Year 2005 .....	12
Figure 3.2: Firm Distribution by Sector .....	14
Figure 3.3: Employment Distribution by Sector .....	14
Figure 3.4: Distribution of TAZs (number and enclosed area) and Firms by Location ....	15
Figure 3.5: Frequency of Trips per Vehicle (Source: Commercial Vehicle Survey 2006 (Texas DOT)) .....	16
Figure 3.6: Household Evolution Model .....	18
Figure 3.7: Asymmetric Triangular Distribution for Income Updating .....	20
Figure 3.8: Asymmetric Triangular Distribution for Vehicle Retirement.....	28
Figure 3.9: Firm Evolution Model.....	30
Figure 3.10: Microsimulation of Households and Firms in Austin, Texas .....	37
Figure 4.1: 3-County Austin Region (TAZs by Location Type).....	43
Figure 4.2: Distribution of Austin-area Firms in 2005 (Source: CAMPO).....	44
Figure 4.3: Firm Size Profile (Low-growth) .....	45
Figure 4.4: Firm Sector Profile (Low-growth) .....	46
Figure 4.5: Firm Size Profile (High-growth).....	46
Figure 4.6: Firm Sector Profile (High-growth) .....	47
Figure 4.7: Firm Size Profile (Average growth).....	48
Figure 4.8: Firm Sector Profile (Average growth) .....	48
Figure 4.9: Job Distribution Profile by Firm Size (Average growth).....	49
Figure 4.10: Job Distribution Profile across Industry Sectors (Average growth) .....	49
Figure 4.11: Firm Density (Firms per Square Mile) by Location in Year 2005 and Forecasts for Year 2035.....	51
Figure 4.12: Employment Density (Jobs per Square Mile) by Location in Year 2005 and Forecasts for Year 2035.....	52
Figure 4.13: Firm Density (Firms per Square Mile) by Location in Year 2005 and Forecasts for Years 2020 and 2035 for Average Growth Scenario.....	53
Figure 4.14: Employment Density (Jobs per Square Mile) by Location in Year 2005 and Forecasts for Years 2020 and 2035 for Average Growth Scenario.....	54
Figure 4.15: Household Density (Households per Square Mile) by Location in Year 2005 and Forecasts for Year 2035 .....	56
Figure 4.16: VMT Associated with Each Functional Class by Time for No-toll Scenario .....	59

Figure 4.17: VMT Associated with Each Functional Class by Time for Fixed-toll Scenario ..... 59

## **LIST OF TABLES**

Table 3.1: Person Type Splits in Households in Year 2005 .....	11
Table 3.2: New Birth and Death Percentages by Firm Size (averaged over all sectors) for Low-growth and High-growth Scenarios .....	15
Table 3.3: Summary Statistics for Attributes of Birth Model .....	19
Table 3.4: Binary Logistic Model of Child Birth ( $y = 1$ if a child is born in that year) ....	19
Table 3.5: Summary Statistics for Attributes of Marriage Decision Model.....	22
Table 3.6: Binary Logistic Model of Marriage Decision .....	22
Table 3.7: Summary Statistics for Attributes of Model of Divorce Decision .....	23
Table 3.8: Binary Logistic Model of Divorce Decision .....	23
Table 3.9: Residential Location Choice Model Results (using MNL) .....	24
Table 3.10: Summary Statistics for Variables in NHTS 2001 Data Set.....	25
Table 3.11: Poisson Regression Model of Vehicle Ownership for all Households .....	26
Table 3.12: Binary Logistic Model of Vehicle Purchase in Last 12 Months .....	27
Table 3.13: Percentage Change in Number of Firms During 2001-2002 and 2002-2003 for Each Sector .....	31
Table 3.14: Percentage Change in Number of Firms During 2001-2002 and 2002-2003 For Each Size Category .....	31
Table 3.15: Markov Transition Probability Estimates for Basic Firms.....	32
Table 3.16: Markov Transition Probability Estimates for Retail Firms .....	32
Table 3.17: Markov Transition Probability Estimates for Education Firms .....	33
Table 3.18: Markov Transition Probability Estimates for Service Firms.....	33
Table 3.19: Logit and Poisson Regression Models of Firm Location Choice.....	35
Table 3.20: Binary Logistic Model of Long Distance Travel Choice .....	38
Table 3.21: Negative Binomial Model of Household Trip Generation.....	38
Table 3.22: Multinomial Logit Model of Household Trip Destination Choice.....	39
Table 3.23: Negative Binomial Model of Commercial Trip Generation by TAZ.....	40
Table 3.24: Logit Model of Commercial Trip Distribution.....	41
Table 4.1: Forecasting Evolution of Population Attributes over Time .....	57
Table B-1: Results of Firm Location Choice Model for Each Sector .....	135

## **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, in part, with general revenue funds from the State of Texas. The authors also express thanks to Dr. S. Travis Waller and fellow graduate students, Mr. Jason Lemp and Ms. Brenda Zhou for their guidance and help throughout this research. Additional thanks are expressed to Ms. Laura McWethy and Mr. Vana Prasad for their helpful suggestions and ideas about various aspects of this research, and to Mr. Mohit Chamanian for his help with coding issues.

## CHAPTER 1: INTRODUCTION

The last few decades have seen rapid urbanization, growing urban boundaries and cities getting increasingly congested across the world. The congestion problem can not be solved by adding the infrastructure alone; hence the attraction of planners has shifted towards more effective management of travel demand (Goulias and Kitamura, 1992). Answering various questions raised by urbanization and infrastructure's ability to cope with added demand poses a challenge to policymakers, planners and researchers. A proper understanding of travel behavior and traveler demographics is a critical component in devising policies to tackle this problem. Several researchers have focused on behavior of the agents involved, especially households and, to a lesser extent, firms, and their land use and transport interactions (see, e.g., Miller et al. 1998, Timmermans 2003, Waddell et al. 2003, and Salvini et al. 2005). As agents of development, households have been studied extensively, using sound behavioral specifications and data. Owing to embedded heterogeneity in firm behavior (Elgar et. al, 2006) and the lack of quality data, existing literature lacks similar studies tracking firm behavior. Since commercial trips constitute roughly 10 to 15% (Hunt et al., 2004) of inter- and intra-regional travel, an understanding of firm behavior should enhance transportation planning. Moreover, firm and household location choices are greatly affected by one another and studying their interaction may be very useful for planning purposes. Zones of high employment are major attractors (and producers) of personal and commercial trips. Anticipating the location choices of every household and firm in the study area over the next several decades should result in some useful insights that spatially and demographically aggregate models cannot provide. The effects of firmographic events (such as firm moves, growth and closure) are also generating interest among planning agencies. Thanks to changes in household and firm sizes, income and industry, the demography of firms and households may be crucial in understanding the process of urbanization and forecasting urban futures. Hence, this study aims to track firms in conjunction with households via a simulation model that forecasts future travel demand, location choices and changing demographic patterns (of households and firms).

Advances in technology and computing abilities over the last few decades have led to a shift in modeling practices, which has resulted in individual behavior becoming the subject of study rather than the behavioral patterns of an aggregate group of agents. The last decade has seen many microsimulation models (see, e.g., Goulias and Kitamura's (1992) Microanalytic Integrated Demographic Accounting System (MIDAS), and Miller et al.'s (2004) Integrated Land Use, Transport

and Environment (ILUTE) model) being formulated to study travel demand, land use and their interaction with households and firms.

This model specified here is used to anticipate the evolution of households and firms in Austin, Texas over a 30-year period (from 2005 to 2035). A microsimulation approach is used to track 10% of households (scaled up to 100%) and 100% of firms over time and space. The simulation approach seeks to tie evolutionary models of households and firms with models of travel behavior to provide robust forecasts. Model of life-cycle transitions for households and firms have been estimated using various available data sets. Household and firm simulation is carried out every 5 years, in order to update network travel times.

The following section (Chapter 2) provides a review of household and firm demography studies, microsimulation models, location choice studies and related literature. Chapter 3 provides an overview of various data sets used for this study and their characteristics. It also discusses the modeling framework used for this study and the results of various life-cycle transition models, and location choice models. Chapter 4 discusses the results of the simulation model and provides a comparison across different scenarios. Conclusions, limitation and possible future extensions of the research are summarized in Chapter 5.

## CHAPTER 2: LITERATURE REVIEW

Forecasting future travel demand and land use patterns is key to planning large cities and metropolitan areas. Planning agencies are looking for robust forecasting tools to develop policies to meet that end. Such tools require a proper understanding of life-cycle transitions in the demographic status of decision making agents - households and firms. A traditional four-step aggregate modeling technique has been used by most of the planning agencies to forecast travel demand. A better model is long overdue, but transitioning to a disaggregate model requires a much more data intensive process and faster computing abilities. Technological advances have provided faster computation but a lack of quality micro data remains a problem. For example, simulating the evolution of households and firms requires panel data to estimate models of various life-cycle transitions. Since panel studies of firms and households are not available for nearly all U.S. regions, various aggregate and disaggregate datasets have been used in this study.

### 2.1 Microsimulation

Orcutt (1957) was among the first to suggest using microsimulation in socio-economic context. Though his idea did not find many takers initially owing to its data intensive nature and complexity, the last decade has seen such an emergence. Microsimulation typically implies the tracking of individuals over time and space and enjoys increasing interest among transportation engineers. One key advantage of microsimulation stems from our desire to analyze the impacts of policies at the individual level. However, such advantages should be viewed in the context of added complexity and increased data and computational requirements (Goulias and Kitamura, 1992).

More and more aspects of travel behavior, involving both temporal and spatial dimensions, are being applied in a microsimulation framework (Miller et al., 2004). Examples include residential location choice (Otter et al 2001), household and firm demographics (Khan et al. 2002, and Maoh et al. 2005) and auto ownership (Berkovec, 1985).

With the change in modeling paradigm, activity-based and microsimulation models of travel demand are starting to replace traditional aggregate models. Disaggregate models in Bhat et al's (2004) *A Comprehensive Econometric Micro-simulator for Daily Activity-travel Patterns (CEMPDAP)*, Arentze and Timmermans' (2004) *A Learning-Based Transportation Oriented*

Simulations System (ALBATROSS), and the FHWA's (2006) Transportation Analysis Simulation System (TRANSIMS) provide footprints of behavior, while microsimulation tracks household's transitions through space and time. CEMDAP and ALBATROSS provide two distinct approaches to activity-based travel demand modeling. While CEMDAP is based on utility maximization principle, ALBATROSS proposed learning based approach to travel activity-based travel demand model. Arentze and Timmermans (2004) argue that individuals are not utility maximizers but continuously adapt to the changing environment and their choice behavior is the outcome of learning and experience gained in the process. Though these are comprehensive travel demand models, they do not capture the interaction between land use and travel demand. Land use and socio-demographic attributes are exogenous to the modeling framework.

Conventional, static equilibrium models are unable to capture the dynamics of urban system evolution, which are thought to be better explained using a dynamic disequilibrium approach (Miller et al., 2004). An agent-based modeling approach provides an effective way of modeling complex dynamic systems at a disaggregate level. Waddell et al.'s (2003) UrbanSim is another model that seeks to microsimulate land use at a highly disaggregate level.

### 2.1.1 UrbanSim

UrbanSim adopts a dynamic disequilibrium approach in modeling various economic and demographic processes, unlike the cross-sectional equilibrium approach taken in MEPLAN and TRANUS (Waddell et al., 2003). Assumption of general equilibrium is clearly an over-simplification of the complex interactions between various agents and markets. A highly disaggregate spatial resolution (typically  $150 \times 150$  meters) and modular approach for data management across model components are key features of UrbanSim (Waddell et al., 2003).

All UrbanSim models are run at 1-year intervals, but the travel demand model's outputs which are exogenous to UrbanSim, are expected to be updated less frequently. Hence, travel utilities typically remain constant between several successive runs of the travel demand model. Aggregate forecasts of economic activity and demographics also are exogenous to UrbanSim. It uses random utility maximization (RUM) theory to estimate employment and household relocation models wherein individual jobs and households are used as units of analysis. Of course, the various agents modeled by this method may not have the information required to behave as utility maximizers, which is a

potential weakness of any RUM-based paradigm. While UrbanSim lacks a travel demand model, Miller et al.'s ILUTE model seeks to integrate land use with models of travel demand.

### 2.1.2 ILUTE

Miller et al.'s (2004) ILUTE model is an agent-based integrated microsimulation model of urban land use and transportation. ILUTE seeks to forecast a region's evolution over time by simulating individuals, households and firms. It employs a range of modeling methods including logit models, rule-based (computational process) models, and their combination (Miller et al., 2004). Miller et al. (2004) identifies location choices of households and firms, which evolve within the model, as critical components of an integrated urban model. If forecasted spatial pattern is incorrect, subsequent travel demand forecasts would be misleading too. ILUTE includes microsimulation of demand-supply interactions in the residential and commercial real estate markets.

ILUTE uses a discrete time approach to evolve urban systems which are constantly adapting to changing environment (Miller and Salvini, 2001). Modeling decision processes often involves large number of alternatives as choice set. Random sampling and/or rule-based search processes are used to reduce the choice set before implementing random-utility models to predict choice behavior.

To be able to effectively forecast future land use and travel demand, an accurate model of household and firm evolution is required. Modeling the evolution pattern of these agents and their spatial distribution is critical. The following section provides a discussion on models of household and firm evolution.

## 2.2 Household and Firm Evolution

The demographic changes in the life-cycle of a household or a firm are very difficult to estimate given the highly probabilistic nature of changes, complex causality of various events, and wide variety of external factors that can affect such transitions. Another issue with modeling households is the increased complexity due to interactions among individuals within a household. Hence, a study of changes in individual households' and firms' demographics requires varying degrees of assumptions.

### **2.2.1 Firmography**

The study of firmography is multidisciplinary in nature. It can involve economists (Evans, 1987), sociologists (Hannan et al. 1988 and Carroll and Hannan 2000), geographers (Van Dijk and Pellenbarg, 2000a), and transportation researchers (Khan et al. 2002, Waddell et al. 2003, Maoh et al. 2005, and Abraham and Hunt 2000), among others. Key processes include firm birth, death, migration and location choices, as well as expansion and contraction. While migration and location choice decisions are of special interest to regional scientists, when studying spatial patterns (e.g., Pellenbarg and Steen 2003, and Dijk and Pellenbarg 2000a), firm birth, exit and growth patterns greatly interest economists, anthropologists, and others.

### **2.2.2 Firm Location Choice**

A common objective in firmographic studies is modeling the location choices of new firms and relocation decisions of existing firms. As Hayter (1997) notes, neo-classical theory assumes firms act as profit maximizers with perfect market information (see, e.g., Pellenbarg et al. 2002), while behavioral theory presumes firms are simply “satisficers” rather than maximizers (see, e.g., Brouwer et al. 2002). Past studies have modeled firms in both senses and have done a market-segmentation analysis based on industrial sector and size (De Bok and Sanders 2005, Waddell et al. 2003, Khan et al. 2002, and Elgar et al. 2006). For example, Bartik (1985), Coughlin et al. (1991), Friedman et al. (1992) and Head et al. (1995) used aggregate regions like U.S. states as choice alternatives for firm location choice models. To account for agglomeration economies and labor market effects (which are key factors and vary at much finer levels of spatial resolution and are important factors), Hansen (1987) used cities in and around São Paulo, Brazil, and Woodward (1992) used U.S. counties as choice alternatives. Guimaraés et al. (1998 and 2000) used municipalities in Puerto Rico and concelhos<sup>1</sup> in Portugal as choice alternatives. Analysis of data at a finer spatial scales results in larger choice sets, which can become impractical due to computational requirements and problems posed by multinomial logit models and their independence of irrelevant alternatives (IIA) property (Ben-Akiva and Lerman, 1985). To address such issues McFadden (1978) proposed the use of a random sample of choice alternatives, for estimation of the logit model, and showed it to result in consistent estimators. However, efficiency and replications are compromised (Train, 1986). When counts per chosen alternative are relatively high and attributes do not vary by decision maker, Coughlin et al. (2000), Guimaraés et al. (2003) and many others have used Poisson regression techniques to counter the

---

<sup>1</sup> A reasonably small administrative area, Portugal's concelhos have an average area of 322.5 km<sup>2</sup>.

challenges posed by a large set of choice alternatives. In such contexts, Poisson regression is a tractable econometric alternative to the conditional logit model, consistent with McFadden's (1974) maximization framework.

Relocation of existing firms is another aspect of location choice studied extensively in the 1970s in Europe and the U.K. (Pellenbarg et al., 2002). Data improvements and technological advances have enhanced the spatial resolution of these models, from regional to more local levels (of cities), but a study of firm mobility at the zonal or parcel level largely remains neglected. Local firm movements affect local travel patterns and land use, as well as the local economy (Van Dijk and Pellenbarg, 2000b). While most researchers have focused on factors affecting location choices of new firms, Lee (2006) used a binary probit model to examine the effect of firm and location specific factors on the relocation decision, and Van Dijk and Pellenbarg (2000b) used an ordered<sup>2</sup> logit model to study firm relocation patterns in the Netherlands.

### 2.2.3 Commercial Trip Models

Commercial trips are important to understanding traffic patterns. They account for 10 to 15% of intra-urban vehicle trips (Hunt et al., 2004), an even higher share of inter-urban trips., and roughly 12 to 15% of VMT (Hunt et al. (2004) and Spear (2006)). The recent surge in microsimulation models of household travel deserves similar models for commercial trips. Hunt et al. (2004, 2005) took a hybrid approach to model commercial movements that included a tour-based microsimulation model (accounting for two-thirds of commercial trips) along with a model of fleet allocation and movements (roughly one-thirds of commercial trips) and external-internal movements (accounting for a small fraction of remaining trips). A special survey was designed and implemented in the cities of Calgary and Edmonton, Canada to obtain required data. The survey instrument consisted of commodity flow study of business establishments, fleet allocator survey, and external truck survey carried out during years 2000 and 2002 (Hunt et al., 2004). Commercial trip patterns are strongly tied to firm location choices and this study seeks to highlight the same.

---

<sup>2</sup> Propensity to move was expressed in percentages (0%, 0–10%, 10–25%, 25–50%, 50–75%, 75–90%, 90–100% and 100%) and was used as the dependent variable in an ordered logit model.

Similar to firmography, household demographics requires attention for effective forecasts of future travel demand and land use. The following section reviews state-of-the-art and previous works on household evolution.

## **2.2.4 Household Demography**

Models of life-cycle transitions require special panel surveys to track changes in the demographics of a household. Since such surveys are rare, there have been very few models which track household evolution in great detail. Goulias and Kitamura's (1992) MIDAS is one such model, which combines models of travel behavior with a microsimulation model of household demographics. MIDAS was calibrated using the Dutch National Mobility Panel data set. Another study of interest is Calipers' STEP2 model for Nevada's Clark County (Caliper Corporation, 2003), which is closely mimicked by this study's rules of household evolution.

### **2.2.4.1 MIDAS**

MIDAS's (Goulias and Kitamura 1992) key feature is the dynamic microsimulation of household demographics based on identification of a causal structure that underlies the evolution. It estimates econometric models for life-cycle transition events, such as birth, death, and children leaving home. It uses panel data to generate Markov transition matrices for changes in employment status, and driver license holding. This study has a similar model structure, but lacks panel data for study region. However, MIDAS does not have a location choice model for households.

### **2.2.4.2 STEP2 Model**

STEP2 (Caliper Corporation, 2003) is a household-level microsimulation model of location choices and travel demand, reflecting several land use and transportation interactions. It consists of a population synthesizer, household evolution module, and a travel demand model. It lacks a vehicle ownership model but otherwise is closely mimicked by this study's model, in terms of rules for household evolution. Though simplistic in nature, these rules provide a basis for evolving households over time in the absence of any data to more directly model such transitions.

Of course, the ability to correctly forecast the future spatial distribution of population is critical for appreciating the interaction of land use and transport systems. Household location choice has been

studied extensively by transportation researchers and others, over the years. The following section briefly discusses this literature.

### **2.2.5 Household Location Choice**

Household location choice is one of the most important decisions affecting travel demand, since home-based trips are the largest component of daily trip making in urban areas. Location choices by new households and relocating households have been studied extensively using discrete choice modeling techniques using random utility theory (Bhat and Guo, 2004). Various models of household location choice have been applied by McFadden (1978), Weisbrod et al. (1980), Onaka and Clark (1983), and Gabriel and Rosenthal (1989), Waddell (1993), Borgers et al. (1993), Ben-Akiva and Bowman (1998), Sermons and Koppelman (2001), and Bhat and Guo (2004). These studies differ in their assumptions regarding the definition of choice alternatives, modeling framework and explanatory attributes. Bina et al. (2006a, 2006b) studied location choices on apartment dwellers and recent movers in Austin area. This work uses Bina et al.'s location choice model on recent home buyers in Austin.

## **2.3 Models of Vehicle Ownership**

There has been significant effort over the past couple decades to understand vehicle ownership and vehicle choices. Vehicle purchase models are generally estimated using household and vehicle attributes (see, e.g., Lave and Train 1979, Mannering et al. 2002, and Mohammadian and Miller 2003). Mohammadian and Miller (2003) studied the effects of current vehicle holding status of the household, in a nested logit structure. Kockelman and Zhao (2000) used a logit model to model the most recent vehicle purchase by households in the NHTS data.

While many researchers have examined disaggregate models of vehicle ownership, such models are not used to forecast future travel patterns. Berkovec (1985) estimated a simulation model for the U.S. automobile market, and Mannering and Winston (1985) estimated dynamic models of vehicle ownership and utilization. This work estimates a negative binomial regression model of vehicle ownership levels and a simple binomial logit model for whether a vehicle was purchased by a household in the past 12 months.

## **2.4 Models of Long-Distance Travel**

Long-distance or inter-city travel is an important part of any urban travel model, since it constitutes a considerable portion of highway traffic. However, many urban models do not consider such travel or use aggregate base year trip tables to and from external zones to develop future year forecasts which lack policy sensitivity (Erhardt et al., 2007)). Researchers have studied disaggregate models of long distance travel via different model specifications and explanatory variable sets (see, e.g., Koppelman and Hirsh 1986, Koppelman 1989, Bhat 1995, and Lee et al. 2004). While such research often has been rigorous in nature, it has not been integrated with intra-urban models of travel demand. Erhardt et al.'s (2007) model for long-distance travel integrated such estimates with a short-distance travel model for use in different applications. This work uses a binary logit model of long distance travel for individuals. Trip destinations are simply distributed to Austin's external zones based on CAMPO's matrix of external trip attractions.

## **2.5 Summary**

In the absence of panel data for the Austin area, this study seeks to develop a microsimulation model to forecast demographic characteristics of household and firm attributes using various national and local, aggregate and disaggregate data sets and under various assumptions about life-cycle events. A basic microsimulation framework has been developed which can be used in conjunction with quality micro-data to produce better forecasts of future travel demand and land use.

The various data sets used in this work and the model's simulation framework are discussed in the next chapter.

## CHAPTER 3: DATA DESCRIPTION AND MODEL FORMULATION

### 3.1 Household Data Sets

The household data for year 2005 (base year) was synthesized by McWethy (2006) using the Census 2000 5-percent Public Use Microdata Sample (PUMS) for three counties (Travis, Williamsons and Hays), as well as block-group level average data for the same region<sup>3</sup>. Public Use Microdata Sample Areas (PUMAs) are geographical areas containing at least 100,000 individuals. PUMS data contain approximately 5,000 individuals from each PUMA. The block-group level data was required to be spatially reformatted into TAZs in order to be consistent with other data sets used in this study.

In the synthesized population, seven person types were defined for analysis purposes. Table 3.1 specifies person type splits and household averages for all person types.

- Preschool children (ages 0 through 4),
- Pre-driving school-age children (ages 5 through 15),
- Driving school-age children (ages 16 and 17),
- Non-working adults,
- Student adults (full-time students only),
- Part-time working adults (1-39 work hours per week), and
- Full-time working adults (40+ hours of work per week).

Table 3.1: Person Type Splits in Households in Year 2005

Person Type	Count (%)	Mean number per household	Standard Deviation
Preschool Children	89,173 (7.77)	0.20	0.52
Pre-driving school-age children	161,293 (14.05)	0.36	0.76
Driving school-age children	29,173 (2.54)	0.06	0.26
Non-working adults	119,336 (10.39)	0.26	0.58
Student adults	87,394 (7.61)	0.19	0.53
Part-time working adults	188,963 (16.46)	0.42	0.65
Full-time working adults	472,845 (41.18)	1.05	0.94
Total Households	451,003	2.55	1.47

<sup>3</sup> A subset of this report is the paper by Kumar and Kockelman titled “Tracking the Size, Location and Interactions of Businesses: Microsimulation of Firm Behavior in Austin, Texas” and is listed in References.

Other attributes of interest that were tracked for each household in the simulation are annual household income and vehicle ownership. In year 2005, the average Austin household enjoyed an annual income of \$59,496 ( $\sigma = \$51,542$ ) and owned 1.94 vehicles ( $\sigma = 0.95$ ). Figure 3.1 shows the age distribution profile of the Austin population in year 2005.

Other data sets of interest for population evolution are the National Vital Statistics Reports (NVSR) provided by the Department of Health and Human Services and the U.S. Panel Study of Income Dynamics (PSID). The PSID is primarily sponsored by the National Science Foundation, the National Institute of Aging, and the National Institute of Child Health and Human Development and is conducted by the University of Michigan.

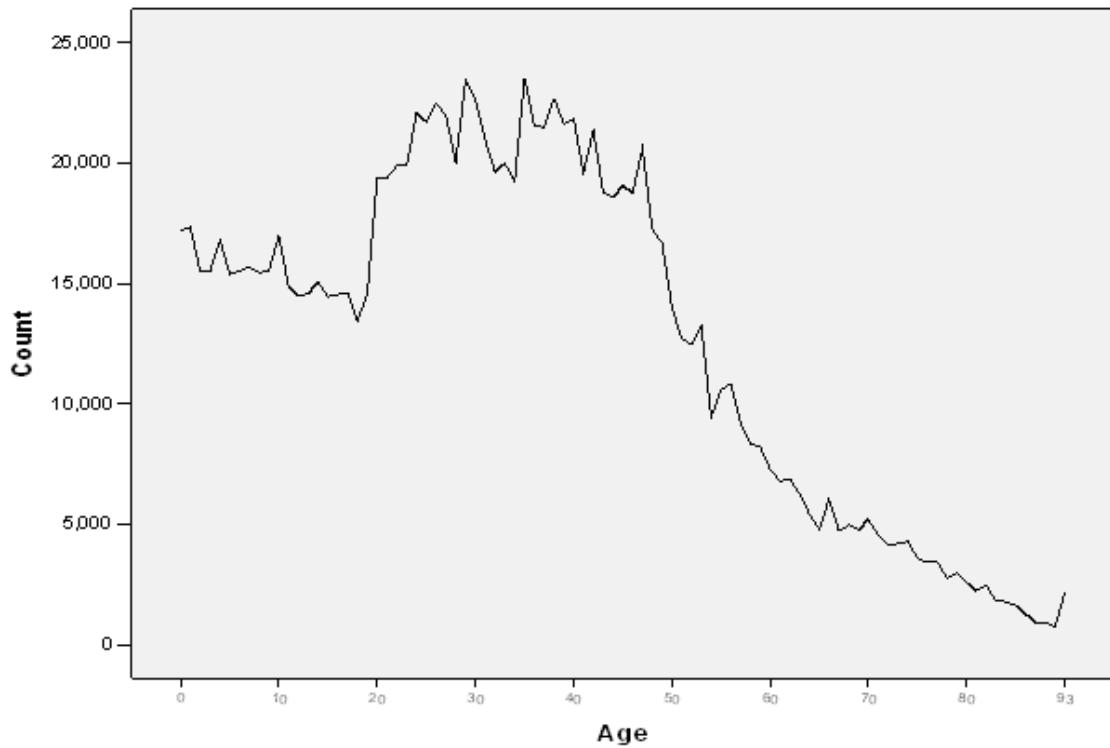


Figure 3.1: Age Profile of Austin Population in Year 2005

Information on TAZ-level land use type and (aggregate) demographic data was obtained for the year 1997 from the Capital Area Metropolitan Planning Organization (CAMPO). That data set also provides year 2007 estimates of all demographic attributes. 1997 and 2007 data were used to impute estimates for the year 2005. These data sets rely on population-equivalent density (PED) measure to classify

TAZs as rural, suburban, urban, and CBD. The definition of PED and classification criteria for TAZs is as follows:

$$PED = [Zone Pop + (Regional Pop/Regional Emp) \times Zone Emp]/Zone Acres$$

CBD if  $PED \geq 15$ , Urban if  $8 \leq PED < 15$ , Suburban if  $1 \leq PED < 8$ , and Rural if  $PED < 1$

where Pop and Emp stand for population and employment, respectively.

Transit availability is also reported at the TAZ-level data, as the fraction of area (within each TAZ) that lies within 0.25 miles of a transit stop, or within 0.50 miles of a transit stop.

Data from the Austin Travel Survey of 1997 (ATS 1997) was used to estimate various travel demand models. ATS 1997 consisted of 16,383 trips by 5,400 persons residing in 1,662 households. 1,843 (11.25%) of these trip entries were excluded due to incomplete or inconsistent information. 236 trips were reported to be external trips (either origin or destination lie outside the study area). Households make an average of 10 trips ( $\sigma = 7$ ) daily, as suggested by the ATS data.

### 3.2 Firm Data Sets

Point location data for all firms in the three-county Austin region was provided by the Texas Workforce Commission and geocoded by the CAMPO. In 2005 Hays, Travis and Williamson counties contained 32,063<sup>4</sup> firms employing 655,722 full- and part-time workers. The study region consists of 1,074 TAZs, and zone-level statistics for the year 2005 were imputed from the CAMPO's 1997 and 2007 data sets. Figures 3.2 and 3.3 show the distribution of firms and jobs by industry sector<sup>5</sup>, respectively. 57% of the firms belong to the service sector and account for 44% of the jobs, while educational establishments account for only 2% of the firms but 11% of the jobs. The average firm size is 20 workers, with a standard deviation of 155 owing largely to a few, very large educational establishments<sup>6</sup>. Figure 3.4 shows the distribution of TAZs and firms by area type. Around 45% of the

---

<sup>4</sup> Addresses of 5 firms could not be mapped to the three-county Austin region and hence were removed from the population for this study.

<sup>5</sup> The four sectors are Basic (manufacturing, construction, mining, transport services and utilities, forestry, fishing, recreation, and hunting and agricultural support), Retail (department stores, restaurants and hotels, retail and wholesale shops), Education (schools, colleges, and training institutes), and Service (information services, real estate, finance and insurance, professional and technical services, health, government services, and banks).

<sup>6</sup> The University of Texas at Austin, Texas State University and Austin Community College have 20,250, 4,884 and 3,390 workers, respectively.

Austin's TAZs are coded as rural, and these account for 85% of the three-county area but only 16% of firms and 12% of workers. In contrast, only 22% of the TAZs are coded as urban or CBD (accounting for a mere 3% of area, yet 45% of the firms and 59% of workers).

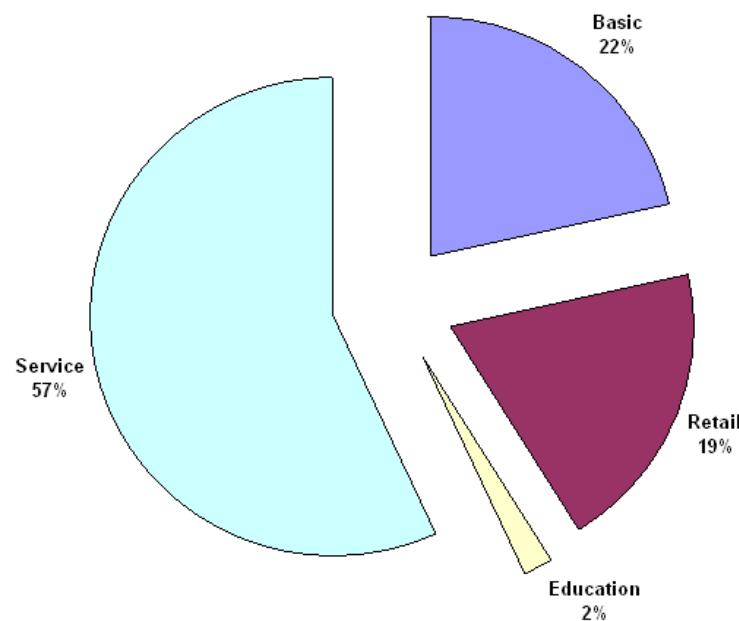


Figure 3.2: Firm Distribution by Sector

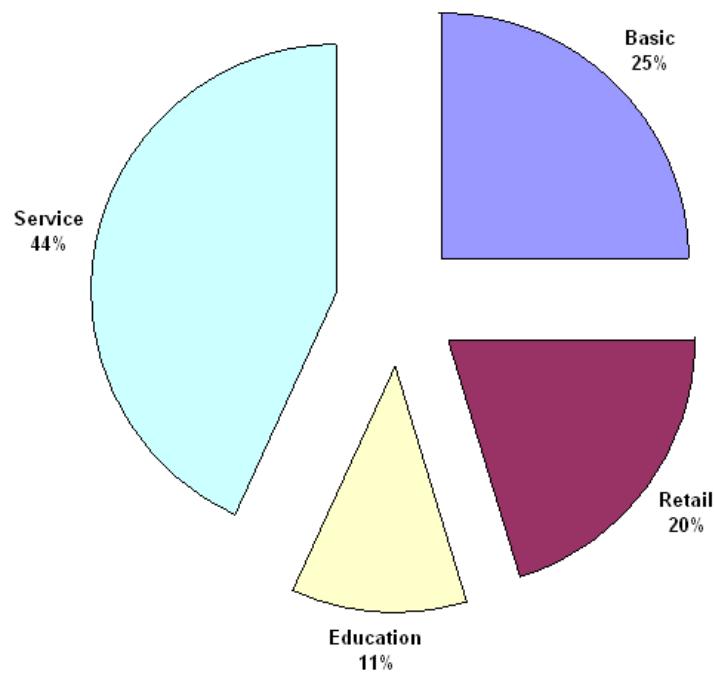


Figure 3.3: Employment Distribution by Sector

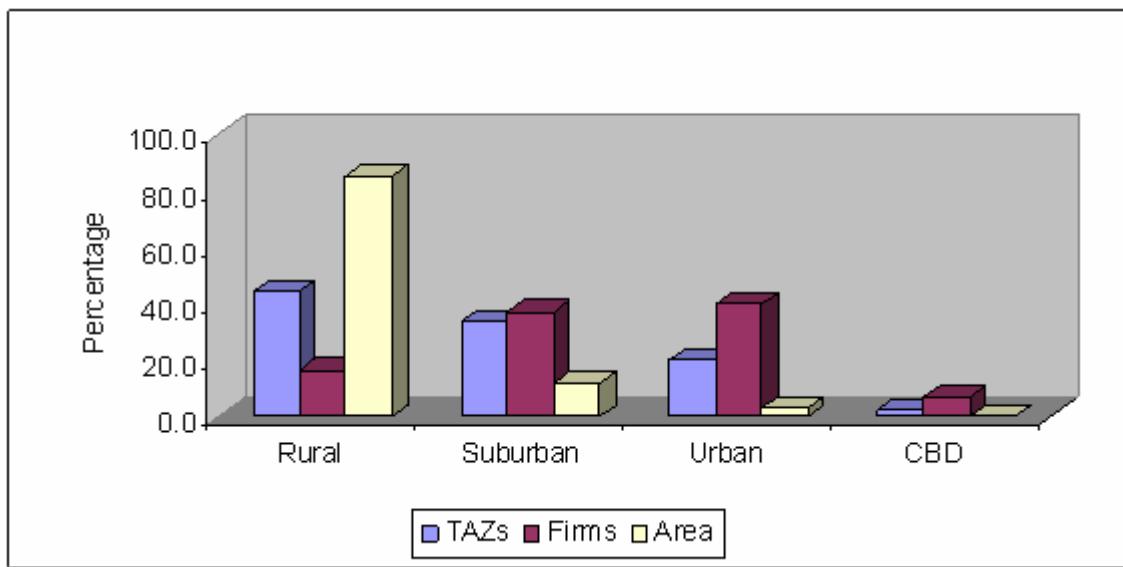


Figure 3.4: Distribution of TAZs (number and enclosed area) and Firms by Location

The Statistics of U.S. Businesses (SUSB) provide annual data on the number of firms and new birth and deaths for employment size of firm categories and industry. The data are a "snapshot" of firms at a point in time. SUSB's firm data from year 1998 to 2004 were used for generating Markov transition matrices of firm growth. 1998 to 2004 was a slow growth period for industries with a couple of years of recession. New firm birth and death data from year 2001-2002 and 2002-2003 were used for modeling firm births and deaths. While most of the industries witnessed negligible growth in 2001-2002, 2002-2003 showed higher growth rates. Table 3.2 shows new birth and death percentages by firm size and averaged over all sectors for low- and high-growth scenarios. An average growth scenario was estimated using mean birth and death rates for these two years.

Table 3.2: New Birth and Death Percentages by Firm Size (averaged over all sectors) for Low-growth and High-growth Scenarios

Firm Size	New Birth (%)	Death (%)	New Birth (%)	Death (%)
	Low-growth (2001-2002)		High-growth (2002-2003)	
1 – 4	16.77	16.81	18.23	16.02
5 – 9	8.29	8.02	9.03	7.08
10 – 19	5.83	5.91	6.06	5.36
20 – 99	3.68	3.9	3.5	3.43
100 – 499	1.05	1.2	1.07	0.87
>= 500	0.11	0.05	0.07	0.06

Data Source: Statistics of U.S. Businesses (2001-2003)

The Commercial Vehicle Survey 2006 data for the region, provided by The Texas DOT, was used to estimate commercial trip generation and distribution models. The Survey was conducted between March and November 2006. 342 vehicles were randomly selected from the study area's list of registered commercial vehicles. These 342 vehicles were based in 113 TAZs and made 2,551 trips over the survey day. Surveyed vehicles made an average of 7.5 trips ( $\sigma = 6$ ) on the day of survey. Figure 3.5 shows frequency of trips made by surveyed vehicles on the day of survey.

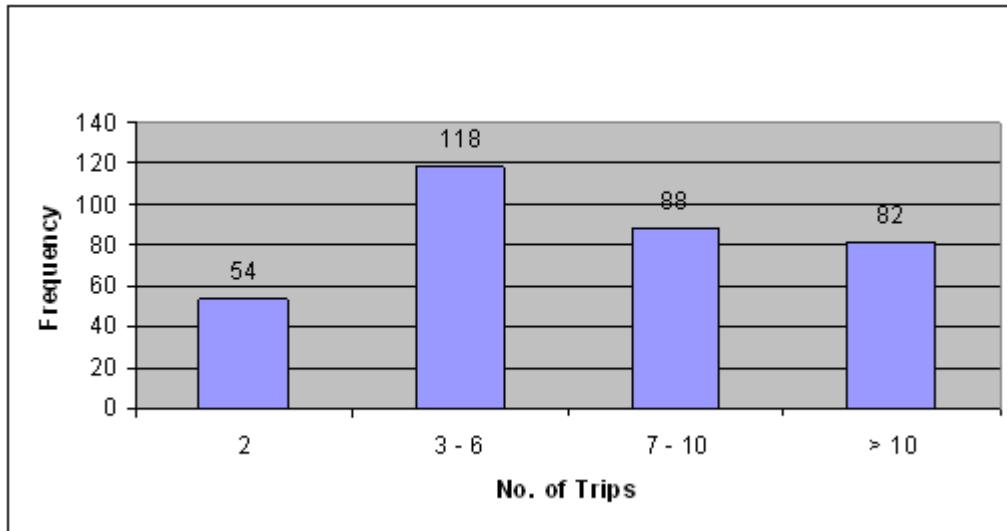


Figure 3.5: Frequency of Trips per Vehicle (Source: Commercial Vehicle Survey 2006 (Texas DOT))

A comprehensive survey, covering all aspects of the various models included in the simulation framework, is not available for any region. Hence, we make use of national and state level statistics for various attributes which are unavailable for the region. The simulation framework developed here, and discussed below, uses a diverse set of datasets mentioned above.

### 3.3 Simulation Framework

This work uses a microsimulation approach to simulate individual household and firm life-cycle events in an urban environment. Microsimulation can be essential in tracking decision makers over their respective life cycles and providing insights into individual responses to various policy changes. Central to such simulation are the statistical models of life-cycle transitions. Lack of quality data is the biggest challenge in developing such models. Of course, any model of human behavior is an abstraction of reality and the uncertainty in external environments and behavioral dependence on

various unpredictable and uncontrolled factors raises questions regarding interpretation of simulation results. Nonetheless, such modeling exercises can provide insights into how households and firms react to various changes in their urban environment, land use and transport system features.

Figure 3.6 illustrates the basic structure of household synthesis and evolution. This model takes a synthetic Austin population (of households) and uses a series of transition models to evolve the population over a 30-year period. Household transition models are applied annually. The entire simulation process is coded in MATLAB programming language, and the code used to run the model is provided here, in Appendix A.

### 3.3.1 Population Synthesis

The first step in the modeling process is generation of the synthetic population. This model uses McWethy's (2006) synthetic population for Austin households. Their process takes zonal population, household population, labor force population and average annual household income as inputs. Since only average values of household size are available at the census tract level, McWethy (2006) used a shifted Poisson distribution to generate marginal distributions for household size in each zone. Using PUMS data as seeds in the iterative proportional fitting (IPF) process, proportions of households in each category in each zone were generated. Following the IPF process, draws of actual values from the closest-matching PUMA household then provided the required demographic information for individual household members (McWethy, 2006)<sup>7</sup>. The synthetic household population is then taken through various life cycle transition models to predict future.

### 3.3.2 Household Transition Models

Figure 3.6 highlights the causal structure underlying the evolution of individuals and households. The modeling structure requires a unique identifier to match individual and household data, and cross-referencing of individual and household data files at multiple stages during the simulation process. This cross-referencing is integral to the model process but results in increased computational requirements and simulation run time. Individuals and households experience births, deaths, migration, marriages, divorces, and so forth. The set of rules for household evolution closely

---

<sup>7</sup> McWethy (2006) provides a complete description of this population synthesis process, along with R code for the IPF calculations.

mimics Caliper's STEP2 model (Caliper Corporation, 2003) for Clark County. All attributes of every simulated agent (members of households and entire firms) are updated annually.

The household simulation model starts with a child-birth model and a death model for individuals. The following section discusses the assumptions in and results of birth and death models.

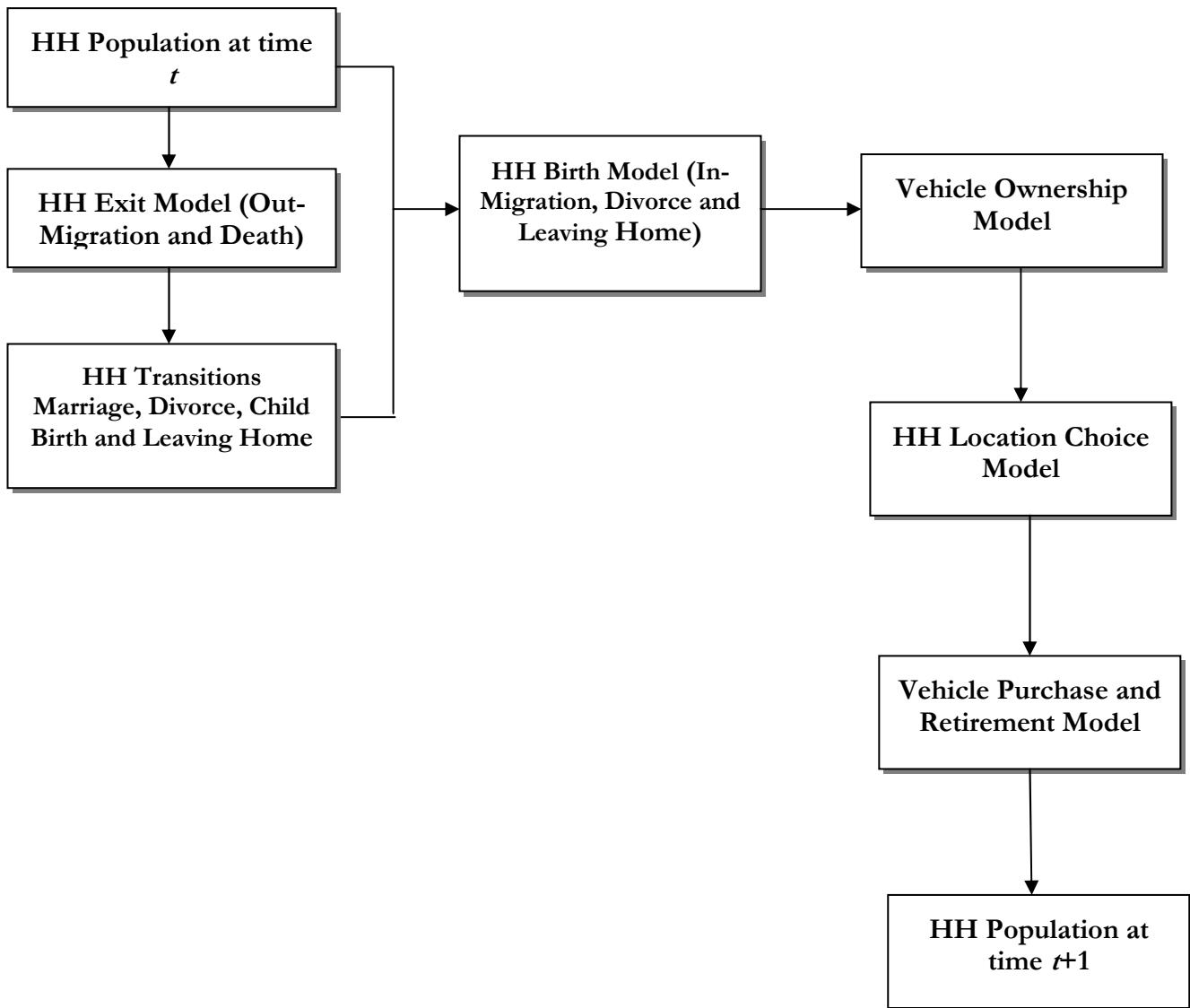


Figure 3.6: Household Evolution Model

### 3.3.2.1 Birth and Death Model

The U.S. National Vital Statistics Report (2002) provides life tables (probability of dying at a particular age) for persons stratified by gender. Using these probabilities and Monte Carlo simulation,

dying individuals were identified and removed from the population. A single-person household was removed from the population in the event of death. This process is followed by a birth-of-child model for all households with a female adult.

A binary logistic model of child birth was estimated using data from the 2005 wave of the U.S. PSID. Sample for birth model consisted of all households with an adult woman. The explanatory variables used in this model are age of mother, number of children, age of youngest child, and employment status of the family head. Any adult women can give birth to a child. Head of household is male adult or the only adult (in case of single-adult households) in the household. Table 3.3 provides summary statistics for all of these, and Table 3.4 provides the model's estimation results.

Table 3.3: Summary Statistics for Attributes of Birth Model

	Min	Max	Mean	Standard Deviation
Household size	1	13	2.70	1.47
Age of mother (in years)	18	45	33.65	6.96
Number of children in the household	0	7	1.41	1.19
Age of youngest child (in years)	0	17	4.45	4.80
Employment status of household head	1957 working and 170 nonworking			

Table 3.4: Binary Logistic Model of Child Birth ( $y = 1$  if a child is born in that year)

	Coefficients	t-statistics	Elasticities
Constant	0.591	1.13	
Age of mother	-0.1098	-7.35	-.00010
Number of children	1.893	14.67	.00167
Age of youngest child	-1.618	-10.80	-.00143
Household head's employment status	0.666	1.89	.00045*
<b>Statistics:</b>			
Pseudo r-square	0.4522		
Number of observations	2,127		

(\*) elasticity is for discrete change of dummy variable from 0 to 1

Data Source: 2005 US PSID households with at least one adult female

The Probability of a child birth in a household decreases with mothers' and youngest child's age. Households with working head are more likely to have a child birth in any particular year than households with non-working head. Child birth probability also increases with number of children already present in the household.

Person death and birth-of-child module is followed by an income updating module in the simulation model, which is discussed in the next section.

### 3.3.2.2 Income Updating

Gomme and Rupert (2004) suggest a 3% per capita annual income growth in U.S. during 1992 through 2003. Assuming a constant income increase for the entire population would be an oversimplification. For purposes of practicality and demonstration, this work assumes an asymmetric triangular distribution of annual percent income increase. Figure 3.7 shows the distribution's parameters and shape of the triangular distribution. This distribution has the following probability density function:

$$f(\theta | a, \mu, b) = \begin{cases} \frac{2}{b-a} \frac{\theta-a}{\mu-a} & \forall \theta \in [a, \mu] \\ \frac{2}{b-a} \frac{b-\theta}{b-\mu} & \forall \theta \in [\mu, b] \\ 0 & otherwise \end{cases} \quad (1)$$

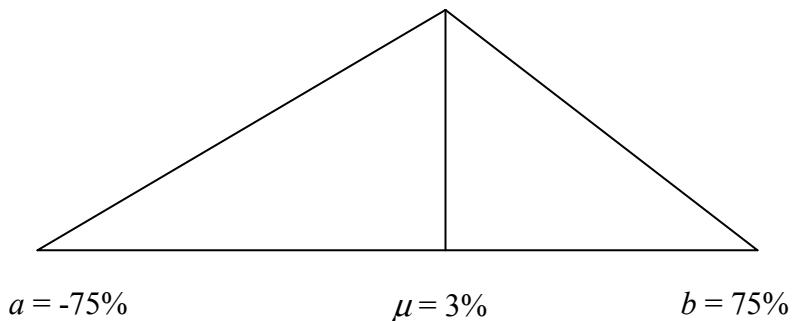


Figure 3.7: Asymmetric Triangular Distribution for Income Updating

Income updating module is followed by household migration and children-leaving-home modules. The following sections discusses these modules and assumptions that underlie them.

### 3.3.2.3 Migration Module

The migration module relies on the U.S. census data, using a positive rate of net population migration at the county-level for three-county region in year 2005. The region's year-2005 net rate of increase in population (per year) is used through year 2035. Use of a net migration rate focused on generating only new, entering households, rather than departing households implies that households never leave the study area, which is an obvious limitation of the present model. In reality, probably

more than 3% of Austin's residents depart each year, including many UT Austin students in search of new jobs. In the absence of data to estimate such a migration model, net numbers of new households are randomly drawn from the population and duplicated. Obviously, demographics of in-migrating households could be quite different from the existing households but this assumption is necessitated by data constraints. Household population in the region also changes due to children leaving home, which is discussed in the following section.

#### **3.3.2.4 Households Formed by Children Leaving Home**

Based on the observed frequency of 22 year old children living alone, 22 year olds are randomly selected to form new households. In the event of a move-out decision, the following household adjustment rules are applied:

1. The new formed household will be a single-member household.
2. Income of such new households is assumed to be the average household income of single-member households for persons aged 22 to 25.
3. The probability that a new household will have a vehicle is equal to the auto availability ratio (number of automobiles per person) in the original household.
4. Original household income and vehicles are adjusted accordingly, except that a single vehicle is never removed in a single year and income is always positive.

The following sections discuss the use of binary logistic models for marriage and divorce, as well as rules for household splits and mergers.

#### **3.3.2.5 Marriage Module**

A binary logistic model of marriage decision (Table 3.6) in a household was estimated using data from the 2005 wave of the U.S. PSID. The explanatory variables used in this model are explained in Table 3.5 and include household size, and household head's age and employment status. The probability of a marriage decreases with an increase in the household head's age. Larger households and households with working heads are more likely to have a marriage. Marriage model is applied to all households with an unmarried or divorced adult and Monte Carlo simulation is used to identify

households having marriage in a particular year. The following set of rules is applied to identify the marrying couple and formation of new households:

1. A male individual is chosen from the pool of marrying individuals first followed by a female within 3 years of his age. If there is no match available, age differences are allowed to be 5 years and then 10 years, as needed.
2. Households of the marrying couple are merged if they were the only adults in those households.
3. Probability of the marrying couple to form a new household is equal to the ratio of single-couple household to all households with at least a couple.

Table 3.5: Summary Statistics for Attributes of Marriage Decision Model

	Min	Max	Mean	Standard Deviation
Household size	1	13	1.98	1.34
Age of household head	16	99	44.55	18.21
Employment status of household head	2,304 working and 1,198 nonworking			

Table 3.6: Binary Logistic Model of Marriage Decision

	Coefficients	t-statistics	Elasticities
Constant	-1.869	-10.63	
Household size	0.129	4.27	0.0063
Age of household head	-0.032	-8.80	-0.0016
Household head's employment status	0.027	2.60	0.0013*
<b>Statistics:</b>			
Pseudo r-square		0.1409	
Number of observations		3,504	

(\*) elasticity is for discrete change of dummy variable from 0 to 1

Data Source: 2005 US PSID households with at least one unmarried adult

### 3.3.2.6 Divorce Module

A binary logistic model of divorce (Table 3.8) occurring in a household in any given year was estimated using the same 2005 U.S. PSID dataset. The explanatory variables used are described in Table 3.7. The likelihood of a divorce in a household decreases with household size and household

head's age. Households determined by the model to divorce in a given year are adjusted based on the following rules:

1. The existing household is removed from the population and two new households are created.
2. Children are randomly allotted to the two households.
3. Vehicles are split equally between the two households if there are even numbers of vehicles in the original household. In other cases, the remaining vehicle is assigned randomly to one of the households.
4. Income is split based on average income for male and female adults in the Austin area.
5. The marital status of the two adults involved in the divorce is changed, and these individuals are allowed to re-marry in the next simulation period.

Table 3.7: Summary Statistics for Attributes of Model of Divorce Decision

	Min	Max	Mean	Standard Deviation
Household size	1	19	3.26	1.30
Age of household head	16	96	46.55	14.42

Table 3.8: Binary Logistic Model of Divorce Decision

	Coefficients	t-statistics	Elasticities
Constant	-1.437	-6.85	
Household size	-0.281	-5.75	-.0083
Age of household head	-0.028	-6.81	-.0008
<b>Statistics:</b>			
Pseudo r-square	0.1326		
Number of observations	4,028		

Data Source: 2005 US PSID households with at least one married couple

New households formed as a result of in-migration, marriage, divorce, and children moving out of households as well as a fraction of existing households make location choice decisions during each simulation period. The following section discusses a multinomial logit model of household location choice.

### 3.3.3 Model of Household Location Choice

Household location choices were estimated by Bina and Kockelman (2006). Location choice was modeled at the TAZ level for recent movers, using Travis County's 544 TAZs. The choice set for move was comprised of nine randomly chosen TAZs, plus the chosen option. Bina (2005) quantified

zone size using a natural-log-of-number-of-housing-units control variable. Table 3.9 shows the results of Bina and Kockelman's (2006) location choice model<sup>8</sup> for recent home buyers. While 62.7% of movers during 2002 and 2003 were renters, they only represent 29% of U.S. households (Schachter, 2004). Location choice model is estimated for recent home buyers alone and does not consider apartment dwellers and presents scope for improvement.

Table 3.9: Residential Location Choice Model Results (using MNL)

	Coefficients	Elasticities
Suburban location	0.476	0.192
Urban location	0.351	0.096
Distance to CBD (miles)	-0.071	-0.431
Median household income (dollars)	1.10E-05	0.579
Ratio of median home value in TAZ to household's income	-0.311	-0.581
Median number of rooms in TAZ per dwelling unit	0.403	1.806
Population density (persons per sq. mile)	1.14E-04	0.297
Employment density (jobs per sq. mile)	1.86E-05	0.030
Logsum for home-based work trips	-2.73E-04	-1.189
Natural logarithm of the number of housing units in TAZ	0.814	4.333
<b>Statistics:</b>		
Log likelihood at convergence	-1541.511	
Pseudo adjusted R <sup>2</sup>	0.173	
Number of observations	811	

\* All variables are statistically significant at a 95% level of confidence

Source: Bina and Kockelman (2006)

The location choice model results suggest that central locations (closer to the CBD) are preferred, *ceteris paribus*. However, the negative coefficient on logsum of home-based work-trips suggests otherwise. As expected, more expensive locations are less likely as suggested by a negative coefficient on ratio of median home value in TAZ to household's income, which is an indicator of neighborhood affordability. However, zones with greater household incomes are more likely, which may be a proxy for other attributes such as attractive neighborhood appearance, quite and safe neighborhood. Also as expected, neighborhoods offering larger homes (a higher median number of rooms per home) are preferred. Population and employment density increases the likelihood of a TAZ. The coefficient on natural logarithm of the number of housing units is positive, as expected. Finally, high elasticities indicate that neighborhood home sizes and regional accessibility are highly significant.

---

<sup>8</sup> Bina et. al's (2006a) provides summary statistics for the variables used in the location choice model.

Vehicle ownership is one of the key attributes of household demographics for studying transport systems. This work tracks vehicle ownership levels of households during the entire simulation period. Vehicle ownership model for new households and models of recent vehicle purchase and retirement are discussed in the following sections.

### 3.3.4 Vehicle Ownership Model

The vehicle ownership model used for this study was estimated using the National Household Travel Survey (NHTS) 2001 data set. Summary statistics of variables in NHTS data set is shown in Table 3.10. A negative binomial regression model of vehicle ownership was investigated, but the overdispersion<sup>9</sup> parameter was found to be statistically insignificant (*t*-statistics = 1.03). Thus, a Poisson regression model was estimated, and its results are shown in Table 3.11.

Table 3.10: Summary Statistics for Variables in NHTS 2001 Data Set

	Min	Max	Mean	Standard Deviation
Vehicles	1	4	2.09	0.92
Household income (\$/year)	2,500	87,500	49,988	25,400
Household size	1	14	2.60	1.34
Workers	0	10	1.42	0.95
Drivers	1	10	1.90	0.73
	Count			
1 adult + youngest child (0-5 yrs)			564	
1 adult + youngest child (6-15 yrs)			1,370	
1 adult + youngest child (16-21 yrs)			568	
2+ adults			14,465	
2+ adults + youngest child (0-5 yrs)			8,348	
2+ adults + youngest child (6-15 yrs)			9,066	
2+ adults + youngest child (16-21 yrs)			3,018	
1 retired adult (65+ years)			5,497	
2+ retired adults			9,537	

---

<sup>9</sup> The negative binomial distribution can be parametrized as follows:

$$\Pr(Y = k \mid \lambda, \alpha) = \frac{\Gamma(k + \alpha^{-1})}{k! \Gamma(\alpha^{-1})} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \lambda} \right)^{\alpha^{-1}} \left( \frac{\lambda}{\alpha^{-1} + \lambda} \right)^k$$

where  $E[Y] = \lambda$  and  $\text{Var}(Y) = \lambda(1 + \lambda * \alpha)$ .

Table 3.11: Poisson Regression Model of Vehicle Ownership for all Households

	Coefficient	t-statistics	Elasticities
Constant	-0.068	-5.90	
Household income (\$/year)	2.59e-06	20.60	5.19e-06
Household size	-0.0157	-3.61	-0.031
Drivers	0.2345	34.99	0.470
Workers	0.0356	7.14	0.071
2+ adults *	0.2398	20.24	0.512**
2+ adults + youngest child (0-5 yrs) *	0.2250	14.51	0.489**
2+ adults + youngest child (6-15 yrs) *	0.2551	17.17	0.560**
1 adult + youngest child (16-21 yrs) *	0.1738	5.59	0.379**
2+ adults + youngest child (16-21 yrs) *	0.2294	13.20	0.510**
1 retired adult (65+ years) *	0.0290	1.91	0.059**
2+ retired adults *	0.2049	15.21	0.441**
<b>Statistics:</b>			
# of observations		59,553	
Pseudo r-square		0.1573	

\* Indicator variables

\*\* Elasticity is for discrete change of dummy variable from 0 to 1

Data Source: NHTS 2001

The results of the Poisson regression model suggest that household vehicle ownership decreases with household size. This is possibly because larger households may have less disposable income. As expected, number of vehicles increases with household income, drivers and workers. Elasticities suggest that number of drivers and adults in household are highly significant. Age of the youngest child in the household also was found to be very significant.

The vehicle ownership model as stated above is only for base year and new households added to the population. Changes in the vehicle ownership level of existing households are represented by vehicle purchase and retirement models.

### 3.3.4.1 Vehicle Purchase and Retirement Model

Vehicle purchase decisions of last 12 months (simulation run frequency) are modeled using a binary logistic model. This model was estimated using NHTS 2001 data set. Summary statistics of the explanatory variables are shown in Table 3.10 (shown earlier) and results of binary logit model of vehicle purchase decision are shown in Table 3.12 below.

Table 3.12: Binary Logistic Model of Vehicle Purchase in Last 12 Months

	Coefficient	t-statistics	Elasticities
Constant	-1.330	-32.62	
Vehicles	-0.067	-4.99	-.011
Household income (\$/year)	-7.17E-06	-13.45	-8.36e-07
Household size	-0.034	-2.87	-0.004
Workers	0.162	8.33	0.020
2+ adults*	-0.277	-8.23	-0.031**
1 adult + youngest child (0-5 yrs)*	0.728	7.71	0.109**
1 adult + youngest child (6-15 yrs)*	0.316	4.62	0.041**
2 adult + youngest child (6-15 yrs)*	-0.114	-2.98	-0.013**
1 adult + youngest child (16-21 yrs)*	0.265	2.58	0.035**
1 retired adult*	-0.489	-9.44	-0.050**
2+ retired adults*	-0.605	-13.66	-0.061**
# of observations		59,549	
Pseudo r-square		0.1573	

\* 1 adult is the base case

\*\* Elasticity is for discrete change of dummy variable from 0 to 1

Data Source: NHTS 2001

As expected, households with more vehicles in the household are less likely to buy a vehicle in a given year. 2+ adult households as well as retired-adult households are less likely to buy a vehicle in any given year as compared to 1 adult household. Surprisingly, larger households and high income households are less likely to buy a vehicle in any given year. Workers have relatively high elasticity value, as they are expected to be significant and increase the likelihood of a vehicle purchase in any given year.

The lack of vehicle retirement data is an obstacle in estimating any vehicle retirement model. Kim et al. (2004) noted that the median lifetime estimate of a 1990-model passenger vehicle in the U.S. is approximately 16 years. Here, vehicle retirement age was modeled as the following asymmetric triangular distribution, which yields a median lifetime of 16 years and average lifetime of 14.67 years.

$$f(\theta) = \begin{cases} \frac{2}{12} \frac{\theta - 8}{8} & \forall \theta \in [8, 16] \\ \frac{2}{12} \frac{20 - \theta}{4} & \forall \theta \in [16, 20] \\ 0 & otherwise \end{cases} \quad (2)$$

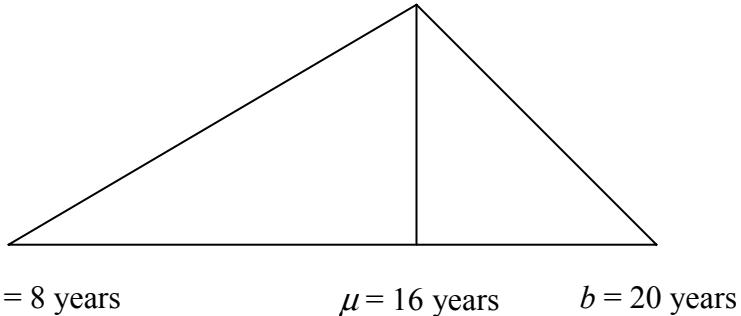


Figure 3.8: Asymmetric Triangular Distribution for Vehicle Retirement

Vehicles are retired and removed from the population based on Monte Carlo draws from this distribution. Note that very short lifetimes are usually due to serious crashes. Other distributions can be easily assumed and included in the model code, as better vehicle datasets become available. Make and model of vehicle are ignored here, but will be important to investigations of fuel use, emissions, resale, crash frequencies, and so forth.

This concludes the discussion on models of household life-cycle transitions. The following section discusses the simulation model for firm evolution in the region. It is important to track firms during the simulation period as firms are an important agent of growth and affect land use and transport system.

### 3.3.5 Firm Synthesis and Evolution

Figure 3.9 illustrates basic structure for firm synthesis and evolution. The firm population is updated every year by running a sequence of sub-models. A model of firm death/exit is applied first, and exiting firms are removed from the population. This is followed by an expansion/contraction model for firms remaining in the population. All existing firms are allowed to relocate based on utility differentials at their current location and ten most popular<sup>10</sup> TAZs. New firms are added (based on the birth model) and their locations are chosen based on a Poisson regression model of location choice. Commercial trip generation and distribution models are applied in each time period. Firm populations and travel patterns are updated every year.

---

<sup>10</sup> Popular TAZs are identified based on levels of TAZ-specific systematic utilities. The TAZ offering the highest systematic utility is defined as the “most popular”.

A lack of panel data necessitates the use of Markov decision process<sup>11</sup> for many submodules in the microsimulation process. Firms are modeled on a three-dimensional state space ( $S$ ) of size, industry affiliation and location. Hence, the Markovian probability of any event in a firm's life cycle is dependent only on its current size ( $s$ ), and industry affiliation ( $i$ ). The Markovian series of states (over time) can be expressed as:

$$\Pr(S(t) = s_t, i_t) = \sum_{s_{t-1}} \sum_{i_{t-1}} \Pr(S(t) | S(t-1)) \times \Pr(S(t-1) = s_{t-1}, i_{t-1}) \quad (3)$$

where  $\Pr(S(t) | S(t-1))$  is one step (one year) transition probability from state at time  $t-1$  to state at time  $t$ . Hence, the transition of a firm from state  $S_{t-1}$  to  $S_t$  depends on firm's existing state ( $S_{t-1}$ ) and is independent of earlier states ( $t' < t-1$ ).

The following discussion provides details of these transition and location choice models as well as models of commercial trip generation and distribution.

---

<sup>11</sup> Assumption of a Markov transition process imply that firm history (e.g., how long a firm has existed or has been in its size category) does not affect its future growth. While such assumptions limit data requirements, ideally they should be tested and improved upon, as needed.

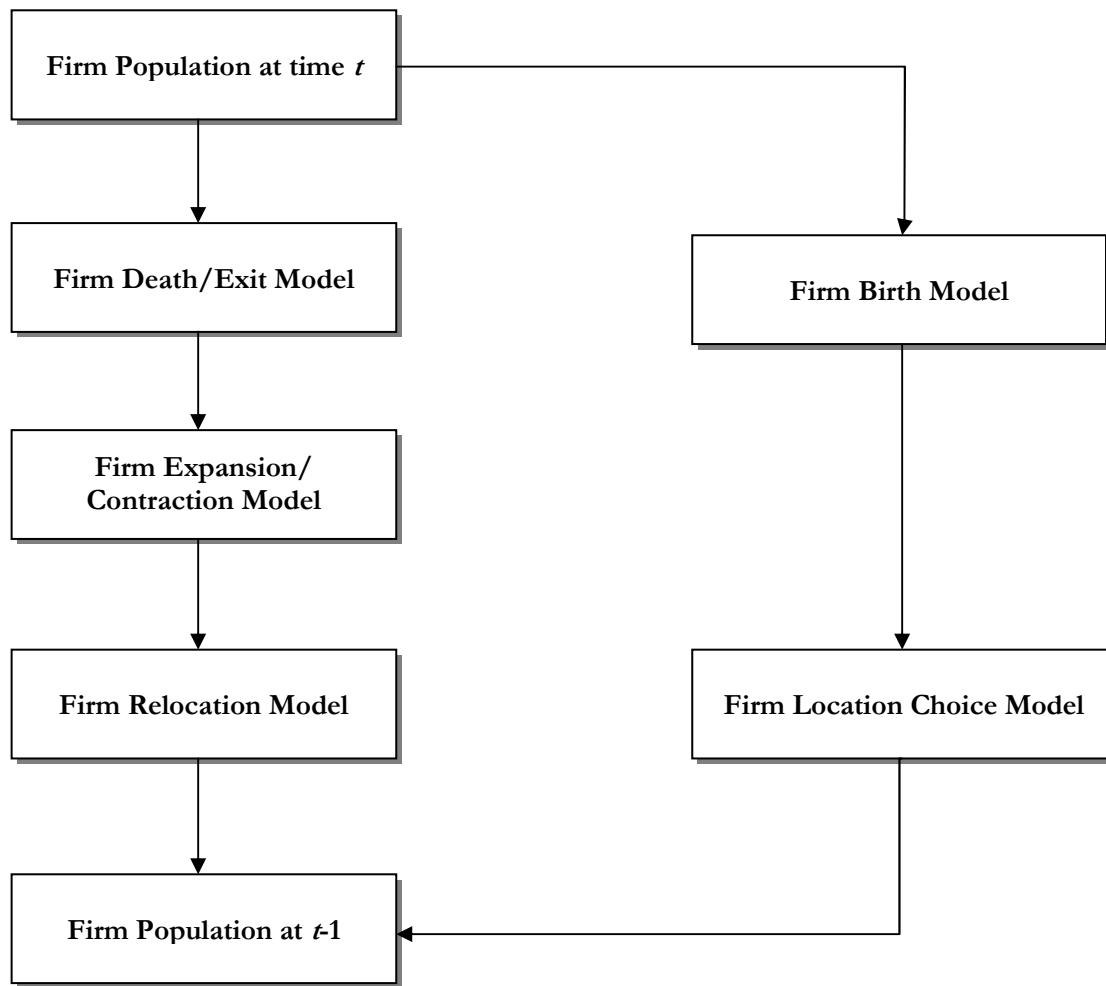


Figure 3.9: Firm Evolution Model

### 3.3.5.1 Firm Death/Exit and Birth Model

SUSB estimates of firm death/exit in the U.S. during 2001-2002 and 2002-2003 periods are provided as a function of firm size category  $s$  and industry sector  $i$ . And these were used to randomly sample firms from the population for removal. The firm-birth model is similar to the death/exit model in its use of SUSB estimates and a random sampling approach. New firm attributes are (randomly) selected from the sample space (as defined by the existing firm population). The location attribute (i.e., zone of residence) is allocated to the each firm using the firm location-choice submodule. SUSB data provides firm counts by size and sector from 1998 onwards and these were used to generate Markov transition matrices for firm expansion/contraction. It also provides data on new firm births and deaths for these years but they are not stratified based on industry sector. Since number of firm deaths and births (stratified by sector) can not be inferred from these data, 2001-2002 and 2002-2003 SUSB data were used for birth and death submodules. As shown in Table 3.13 and 3.14, net percentage increase in

number of firms in year 2001-2002 is significantly lower as compared to 2002-2003; these rates were used to simulate low- and high-growth scenarios, respectively.

Table 3.13: Percentage Change in Number of Firms During 2001-2002 and 2002-2003 for Each Sector

	Low-growth (%) (2001-2002)	High-growth (%) (2002-2003)
Basic	-1.18	2.66
Retail	0.02	0.83
Education	2.82	4.32
Service	0.94	1.19

Table 3.14: Percentage Change in Number of Firms During 2001-2002 and 2002-2003 For Each Size Category

	Low-growth (%) (2001-2002)	High-growth (%) (2002-2003)
1 - 4	1.49	1.95
5 - 9	1.89	3.07
10 - 19	0.86	2.15
20 - 99	-0.31	1.08
100 - 499	0.81	2.14
>= 500	1.48	3.10

### 3.3.5.2 Firm Expansion/Contraction Model

Using SUSB data on annual changes in firm counts (by industry sector and size) from year 1998 through 2004, a Markovian transition matrix was estimated via constrained least squares. SUSB data provides the aggregate proportions  $y_j(t)$  and  $y_i(t-1)$  of firms in size categories  $j$  and  $i$ . Using the conditional probability argument and admitting an error term to account for differences between actual and estimated proportions of  $y_j(t)$ , the following stochastic equation expresses the relation between observed and estimated proportions (Lee et al. 1970, and Jones 2005):

$$y_j(t) = \sum_i y_i(t-1) p_{ij} + \varepsilon_j(t) \quad (4)$$

where  $y_j(t)$  is the proportion of firms at time  $t$  in size category  $j$ ,  $p_{ij}$  is the transition probability from size category  $i$  to  $j$ , and  $\varepsilon_j$  is the stochastic error term. This can be written in matrix notation as:

$$\mathbf{y} = \mathbf{X}\mathbf{p} + \boldsymbol{\varepsilon} \quad (5)$$

This can be expressed as a quadratic programming problem with the following objective function and linear constraints:

$$\underset{\{p_{ij}\}}{\text{Minimize}} f(\mathbf{p}) = \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\mathbf{p})'(\mathbf{y} - \mathbf{X}\mathbf{p}) \quad (6)$$

Subject to  $\sum_i p_{ij} \leq 1 \quad \forall i \quad \& \quad 0 \leq p_{ij} \leq 1 \quad \forall i, j$

Estimates of the transition probabilities  $\{p_{ij}\}$  resulting from this optimization are biased high for large firms. One possible reason for this is the small number of large ( $\geq 500$ ) firms. The variation in  $p_{ij}$  for large firms was quite high initially (upto 8%) thus requiring more steps to reach a stationary distribution. Hence, Markov transition matrices were constrained such that  $p_{ij} = 0$  if  $|i - j| > 1$ . Tables 3.15 through 3.18 show estimates of the Markov transition matrices for basic, retail, education and service industry sectors, respectively. Using such Markovian transition probabilities and Monte Carlo sampling methods, firm and job counts are updated annually.

Table 3.15: Markov Transition Probability Estimates for Basic Firms

	1-4 employees	5-9	10-19	20-99	100-499	$\geq 500$
1-4 employees	<b>0.9257</b>	0.0743	0	0	0	0
5-9	0.0958	<b>0.8340</b>	0.0702	0	0	0
10-19	0	0.0546	<b>0.8454</b>	0.1000	0	0
20-99	0	0	0.0411	<b>0.8805</b>	0.0784	0
100-499	0	0	0	0.0734	<b>0.8810</b>	0.0456
$\geq 500$	0	0	0	0	0.0823	<b>0.9177</b>

Table 3.16: Markov Transition Probability Estimates for Retail Firms

	1-4 employees	5-9	10-19	20-99	100-499	$\geq 500$
1-4 employees	<b>0.9106</b>	0.0894	0	0	0	0
5-9	0.1034	<b>0.8105</b>	0.0861	0	0	0
10-19	0	0.0491	<b>0.8546</b>	0.0962	0	0
20-99	0	0	0.0474	<b>0.8916</b>	0.0610	0
100-499	0	0	0	0.0749	<b>0.8812</b>	0.0439
$\geq 500$	0	0	0	0	0.0704	<b>0.9296</b>

Table 3.17: Markov Transition Probability Estimates for Education Firms

	1-4 employees	5-9	10-19	20-99	100-499	$\geq 500$
1-4 employees	<b>0.9064</b>	0.0936	0	0	0	0
5-9	0.0819	<b>0.8174</b>	0.1007	0	0	0
10-19	0	0.0715	<b>0.8652</b>	0.0632	0	0
20-99	0	0	0.0465	<b>0.8965</b>	0.0570	0
100-499	0	0	0	0.0671	<b>0.8339</b>	0.0990
$\geq 500$	0	0	0	0	0.0971	<b>0.9029</b>

Table 3.18: Markov Transition Probability Estimates for Service Firms

	1-4 employees	5-9	10-19	20-99	100-499	$\geq 500$
1-4 employees	<b>0.9160</b>	0.0840	0	0	0	0
5-9	0.1095	<b>0.8347</b>	0.0558	0	0	0
10-19	0	0.0375	<b>0.8735</b>	0.0889	0	0
20-99	0	0	0.0352	<b>0.8967</b>	0.0681	0
100-499	0	0	0	0.0639	<b>0.8692</b>	0.0669
$\geq 500$	0	0	0	0	0.0473	<b>0.9527</b>

### 3.3.6 Firm Location Choice Model

While actual firm location choices are made at the level of parcels or suites within buildings, it typically is quite difficult to model choice at such resolution due to computational and data availability issues. At finer resolution, the number of choice alternatives increases considerably, and most of the statistical software packages are unable to handle such amount of data. Also, it is difficult to get location data at such high resolution. Thus, this study emphasizes firm location decisions at the TAZ level across the three-county Austin region (1074 TAZs).

In general, let the utility derived by firm  $i$  by locating at TAZ  $j$  be given by:

$$U_{ij} = \beta X_{ij} + \varepsilon_{ij} \quad (7)$$

where  $\beta$  is a vector of unknown parameters to be estimated,  $X_j$  is a vector of explanatory variables associated with the choice alternative, and  $\varepsilon_{ij}$  is the error (stochastic) term. Under a random utility maximization framework and assuming a generalized extreme value type-I (Gumbel) distribution for  $\varepsilon_{ij}$ , this results in the following probabilities:

$$P_{ij} = \frac{e^{\beta X_{ij}}}{\sum_k e^{\beta X_{ik}}} \quad (8)$$

The log-likelihood of Austin's 32,063 firms' location choices in the year 2005 across the region's 1074 TAZs then takes the following form:

$$L = \sum_i^{32,063} \sum_j^{1,074} \delta_{ij} \log P_{ij} \quad (9)$$

where  $\delta_{ij}$  equals 1 if firm  $i$  chooses to locate in TAZ  $j$  and 0 otherwise. After segmenting firms by sector (and not by size or other attributes), this study assumes that the firm location decision is based entirely on attributes specific to the zone alternatives; thus,  $X_{ij} = X_j$ . This assumption allows us to rely on a count model for firms across zones. Essentially, the log-likelihood reduces to:

$$L = \sum_j \eta_j \ln P_j \quad (10)$$

where  $\eta_j$  is the number of firms located in zone  $j$  and  $P_j$  is the probability of zone  $j$  being chosen by any single firm. If all firms choose their locations independently, equations (9) and (8) imply a Poisson model of firm counts, by zone (Guimaraés et al., 2003). In the case of a Poisson distribution for firm counts, the probabilities are as follows:

$$P(\eta_j, \lambda_j) = \frac{e^{-\lambda_j} \lambda^{\eta_j}}{\eta_j!} \quad (11)$$

Hence, the log-likelihood of firm counts across the 1074 zones takes the following form:

$$L = \sum_j^{1074} (-\lambda_j + \eta_j \log \lambda_j - \log \eta_j!) \quad (12)$$

where  $\lambda_j = \exp(\alpha + \beta X_j)$ <sup>12</sup> (for consistency with equation (8)), so that equation (12) reduces to:

$$L = \sum_j \eta_j \log P_j + constant \quad (13)$$

---

<sup>12</sup>  $\alpha$  is the scale factor recognizing total number of firms to be distributed.

The above expression differs from a logit's log-likelihood expression only in terms of its constant and hence, its maximization results in the same estimates of  $\beta$  (and similar standard errors for these  $\hat{\beta}$ 's).

This work's model of firm location choice model was estimated using data for firms across the three-county region. Logsums from a commercial trip distribution model serve as zonal accessibility indices. These accessibilities were computed as follows:

$$Access_i = \sum_j \frac{A_j}{e^{(\gamma_1 \cdot time_{ij} + \gamma_2 \cdot dist_{ij})}} \quad (14)$$

where  $A_j$  is the attraction (population and employment),  $time_{ij}$  is the peak period travel time from zone  $i$  to zone  $j$ , and  $\gamma_1$  and  $\gamma_2$  are parameters from logit models of commercial trip destination choice. Population ( $A_j$  = population in zone  $j$ ) and employment ( $A_j$  = employment in zone  $j$ ) accessibilities are estimated separately and used in the location choice model. All new-born and relocating firms are assigned a zone according to probabilities from this model of firm location choice.<sup>13</sup> The results of the single location choice model for all firms are shown in Table 3.19.

Table 3.19: Logit and Poisson Regression Models of Firm Location Choice

	Logit Model		Poisson Model	
	Coefficients	t-statistics	Coefficients	t-statistics
ln(area)	0.507	64.11	0.482	61.86
Rural*	-3.142	-74.04	-3.025	-73.65
Suburb*	-1.495	-46.67	-1.464	-47.82
Urban*	-0.816	-30.16	-0.778	-30.31
Distance to CBD (miles)	-0.004	-3.19	-	-
Median HH Income (\$/year)	1.31E-06	5.79	1.09E-06	4.93
Employment Access	0.011	39.18	0.011	40.35
Population Access	-0.010	-35.55	-0.010	-36.77
Constant	-	-	4.948	105.18
<b>Statistics:</b>				
log-likelihood (null)			-109,052	-23,817
log-likelihood (model)			-96,558	-12,319
Pseudo r <sup>2</sup>			0.115	0.483
No. of observations			32,063	1,074

\* CBD is base area type.

<sup>13</sup> While separate location choice models initially were estimated for each of the four industries, the estimates were quite similar and hence a single model of location choice is used here, without loss in generality. The results of location choice model for each sector is shown in Appendix B.

The difference in the coefficients of the logit and Poisson models is due to randomly selected choice alternatives. In case of complete choice set,  $\hat{\beta}$ 's would be identical in both the models. Poisson model of location choice is used in the simulation model; since, it is easy to estimate and provides the same estimates as logit model with full choice set. But Poisson regression restricts the use of any firm-specific attribute. Finding an alternative to Poisson model that would allow for individual-specific attributes presents scope for future research.

Firms are more likely to locate near the CBD, everything else constant. Firms also are more likely to locate around other firms or in TAZs with better job accessibility, implying positive agglomeration effects. However, they appear to be less likely to locate in TAZs with high population accessibility, possibly due to congestion, high land prices and zoning restrictions.

Firms also make relocation decisions to cut costs, benefit from agglomeration effects, and to locate near labor market. A lack of data does not allow us to estimate a relocation model for firms. The following sections discusses the assumptions this work makes for firm relocation module.

### **3.3.6.1 Firm Relocation Model**

Kroll et al. (1990) found that roughly 5% of firms relocate locally (within a county) every year. Here, relocating firms were selected randomly with a sampling rate of 15% from among the one-third of firms ( $5\% = 1/3 \times 15\%$ ) exhibiting the lowest systematic utilities, as estimated using the location choice model. New locations were assigned to these 15% of the low-utility firms based on TAZ-specific probabilities determined in the firm location choice model (using appropriately weighted Monte Carlo draws from a multinomial across all possible zones).

After evolving households and firms, models of long-distance travel, trip generation and distribution (household and commercial) are estimated. The following sections discuss these models and traffic assignment module which is run every 5 years.

### **3.3.7 Microsimulation Model of Firms and Households**

Figure 3.10 illustrates the complete microsimulation model of firms and households. The process of household and firm evolution is followed by a long-distance travel model and household and commercial trip generation and distribution models. These models are required to load network and later be able to estimate household fuel consumption, for example. An O-D matrix for all trips was

generated and loaded on Austin's 21,122-coded-links road network. The traffic assignment module was run every 5 years during the simulation process. Hence, trip times between all OD pairs are assumed to remain constant for each 5-year period (between traffic assignment module runs). OD pairs were generated for entire-day and not by time-of-day, which is one of the weaknesses of this model. However, this was done for simplicity and can be easily relaxed.

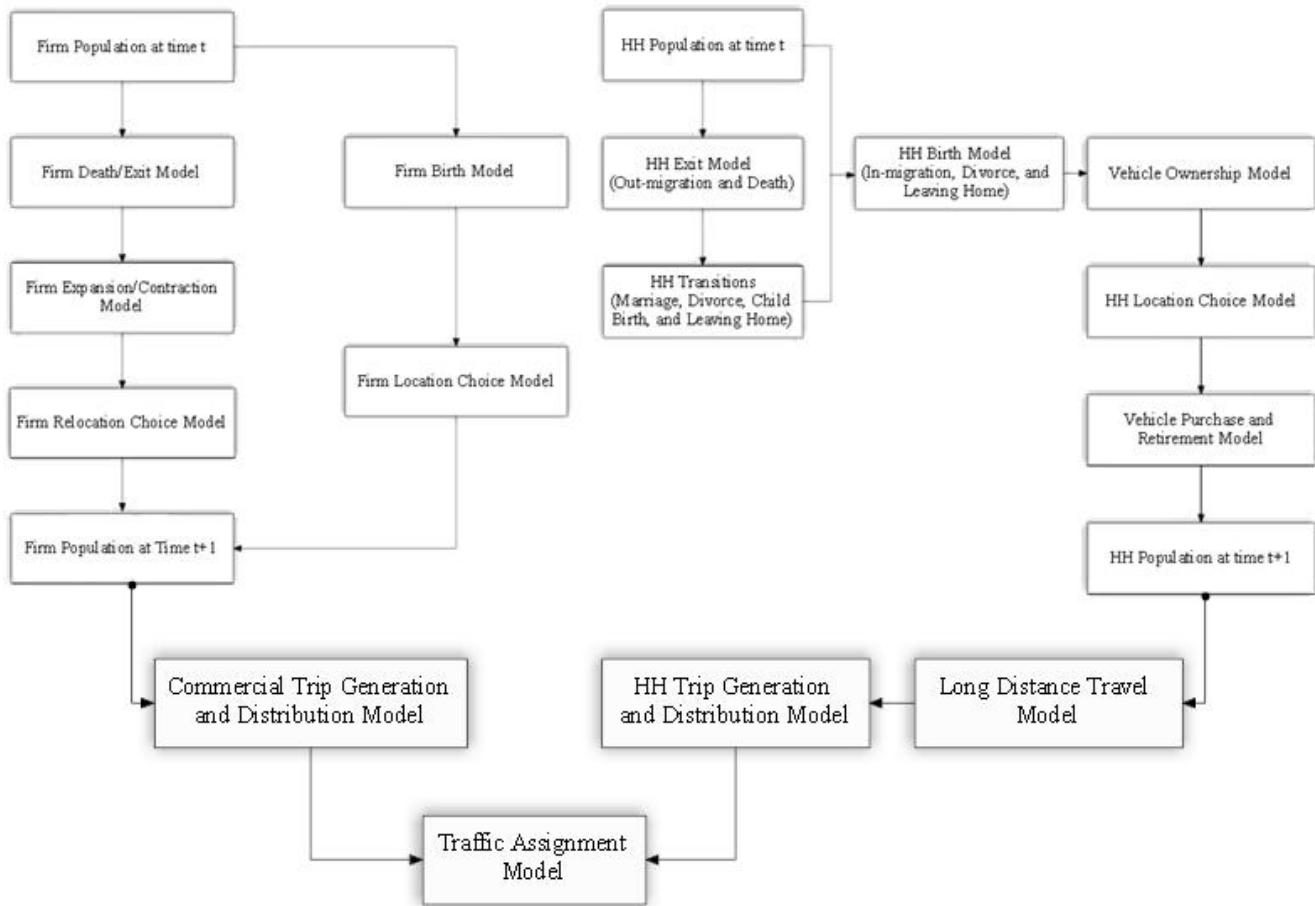


Figure 3.10: Microsimulation of Households and Firms in Austin, Texas

### 3.3.8 Long Distance Travel Model

Trips of length greater than 50 miles as well as trips that originate in the study area and are destined to an external zone are classified as long-distance trips. This study used a binary logistic model for long distance trip generation for travelers residing in the study area. Long distance trips originating in the study area were assigned an external zone as their “local” destination, based on observed external-trip frequencies in CAMPO’s trip table for year 1997. Trips originating outside the study area (external-external trips and external-internal trips) are exogenous to the simulation model

and were loaded on the road network at the traffic assignment stage. External-internal trips and external-external trip tables are provided by the CAMPO. The 1997 ATS data was used to estimate the binary logit model of long-distance trips, the results of which are shown below in table 3.20. NHTS 2001 data set has 0.66% of total person trips as long-distance trips while 1997 ATS data set reports 1.44% trips as long-distance trips. Higher percentage of long-distance trips in ATS data set is partly due to inclusion of external trips with trip length less than 50 miles as long distance trips.

Table 3.20: Binary Logistic Model of Long Distance Travel Choice

	Coefficients	t-statistics	Elasticities
Household size	-0.0819	-19.09	-0.001
Vehicles	0.1881	34.24	0.003
Household income	0.0053	3.34	0.00007
Number of preschool children	0.5073	57.19	0.007
Constant	-4.5528	-292.76	
<b>Statistics:</b>			
Pseudo r <sup>2</sup>		0.0178	
No. Of Observations		1,662	

The likelihood of making an external trip decreases with household size but increases with number of vehicles owned. Rich households and households with more kids are more likely to make an external trip. Next, models of household trip generation and distribution are discussed.

### 3.3.9 Household Trip Generation and Distribution Model

Household trip generation and distribution models were estimated using the 1997 ATS data set. Household trip generation is modeled as a negative binomial random variable in order to account for overdispersion in trip counts. The results of the household trip generation model are shown below in Table 3.21.

Table 3.21: Negative Binomial Model of Household Trip Generation

	Coefficients	t-statistics	Elasticities
Household size	0.3246	369.02	2.80
Vehicles	0.0311	26.61	0.27
Household income (\$/year)	5.91e-06	77.17	5.1e-05
(Household income) <sup>2</sup>	-2.38e-11	-52.62	-2.05e-10
Number of preschool children	-0.2688	-123.45	-2.32
Alpha (Overdispersion parameter)	1.1607	442.93	
<b>Statistics:</b>			
Pseudo r <sup>2</sup>		0.0975	
No. of Observations		1,662	

Number of trips made by a household increases with household size, vehicles, and income, as expected. Households with more preschool-age children tend to make fewer trips. Elasticities suggest that household size and number of preschool children are practically significant variables.

The household trip distribution model is a multinomial logit for destination choice. The choice set for each trip consists of 30 randomly selected destinations<sup>14</sup> plus the chosen destination. The results of the MNL model are shown in Table 3.22.

Table 3.22: Multinomial Logit Model of Household Trip Destination Choice

	Coefficients	t-statistics
Mode logsum	-0.252	-5.42
Zonal area	0.028	3.91
Basic employment	0.00011	15.78
CBD destination indicator	3.006	32.48
Rural origin and suburb destination indicator	0.877	5.57
Rural origin and urban destination indicator	0.645	4.49
Suburb origin and destination indicator	1.987	19.39
Suburb origin and urban destination indicator	2.322	19.97
Urban origin and suburb destination indicator	1.146	10.28
Urban origin and destination indicator	2.378	25.67
<b>Statistics:</b>		
Pseudo r <sup>2</sup>	0.1724	
No. Of Observations	1,662	

Household trip generation and distribution models are followed by a discussion of commercial trip generation and distribution models in the following sections.

### 3.3.10 Commercial Trip Generation and Distribution Models

Commercial trip movements were modeled using trip generation and distribution models. The model of commercial trip generation was estimated using the 2006 Austin Commercial Vehicle Survey data provided by the Texas DOT. The number of trips originating from each TAZ was regressed on various attributes of the TAZ. Trip counts were modeled using a negative binomial regression

<sup>14</sup> Statistical software packages like LIMDEP and STATA constraint the size of data sets that can be estimated. Thus, using all 1074 TAZs as choice alternatives for trip distribution was not feasible with these software programs. Since a subset of randomly selected alternatives results in consistent estimates for MNL models (McFadden, 1978), 30 TAZs were randomly selected as choice alternatives for each trip.

(recognizing their integer nature, while allowing for overdispersion due to reasons like unobserved heterogeneity). The results of the commercial trip generation model are shown below, in Table 3.23.

Table 3.23: Negative Binomial Model of Commercial Trip Generation by TAZ

	Coefficients	t-statistics	Elasticity
ln(area)	0.316	5.50	1.026
Suburban*	0.691	4.41	2.566
Urban*	0.984	4.66	4.476
Distance to CBD	-0.028	-4.00	-0.092
Number of firms	0.006	3.42	0.018
Constant	0.921	6.02	-
Alpha (Overdisp. param.)	3.161	16.32	-
<b>Statistics:</b>			
Pseudo $r^2$		0.567	
No. of Observations		1,074	

\* CBD is the base area type.

Trip counts originating from a TAZ follow a negative binomial distribution, and produce a high overdispersion parameter, of 3.161. Signs and magnitudes of coefficients are as expected: Commercial trip production falls with distance from the CBD. Urban and suburban zones are the biggest generators of commercial trips, *ceteris paribus*. However, trip counts show a very low elasticity with respect to number of firms in the zone.

Trip distribution was modeled as a logit model using proportion of trips destined to a TAZ as the dependent variable. Demographic attributes, firm and employment count of destination TAZ, trip times, and distances serve as explanatory variables in this trip distribution model. Table 3.24 shows the results of the logit model of commercial trip distribution, including a pseudo<sup>15</sup>  $r^2$  value of 0.180 (which is a very reasonable fit for disaggregate trip data). Commercial trips tend to be destined to outer zones (away from the CBD) as suggested by positive coefficient on the distance-from-CBD variable. As expected, peak-period travel time has a negative coefficient.

---

<sup>15</sup> The pseudo  $r^2$  is McFadden's Likelihood Ratio Index,

$r^2 = 1 - \frac{\ln \hat{L}(\text{full})}{\ln \hat{L}(\text{intercept})}$  where  $\hat{L}(\text{full})$  is the likelihood for the full model, with predictors, and  $\hat{L}(\text{intercept})$  is for the constrained, constants-only model, without any predictors.

Table 3.24: Logit Model of Commercial Trip Distribution

	Coefficients	t-statistics
Area x Rural indicator	0.024	4.97
Area x Suburban	-0.426	-13.2
Area x Urban	-1.193	-10.69
Sqrt(HH)	0.006	3.58
Median Household income	-6.60E-06	-5.91
Number of firms	0.0005	2.05
Distance to CBD	0.007	2.7
Peak drive time	-0.075	-51.44
<b>Statistics:</b>		
Pseudo r <sup>2</sup>	0.180	
No. of Observations	2,551	

### 3.4 Summary

This chapter highlights the microsimulation framework proposed to track each household and firm in Austin region and its members' travel behavior. The assumptions underlying the life-cycle transitions have been highlighted along with a discussion of various models' strengths and weaknesses. The discussion provides insights into the nature of various datasets used in the larger model, and these data sets' limitations. The following chapter describes the results of the simulation models' application in the Austin region.



## CHAPTER 4: MODEL APPLICATION IN AUSTIN, TEXAS

The proposed microsimulation model was applied to the Austin region (Figure 4.1) in order to forecast its growth and its sensitivity to system changes and external factors. Three sets of growth assumptions were used: a relatively high-growth rate (of added firms) scenario, a relatively low-growth rate scenario, and an average growth rate scenario (2% firms are added up every year). Two tolling scenarios were investigated, both in the average-growth-rate context: a base-case/no-toll scenario and a 10 cents/mile system-wide tolling scenario. Comparisons with congestion-based tolling and/or freeway-only tolling would be useful scenarios to examine as well, but these were not considered here owing largely to time constraints.

Microsimulation of household and firm evolution was carried out using MATLAB. The results of the household and firm evolution are discussed here, in terms of changing location patterns and demographic attributes over a 30-year period.

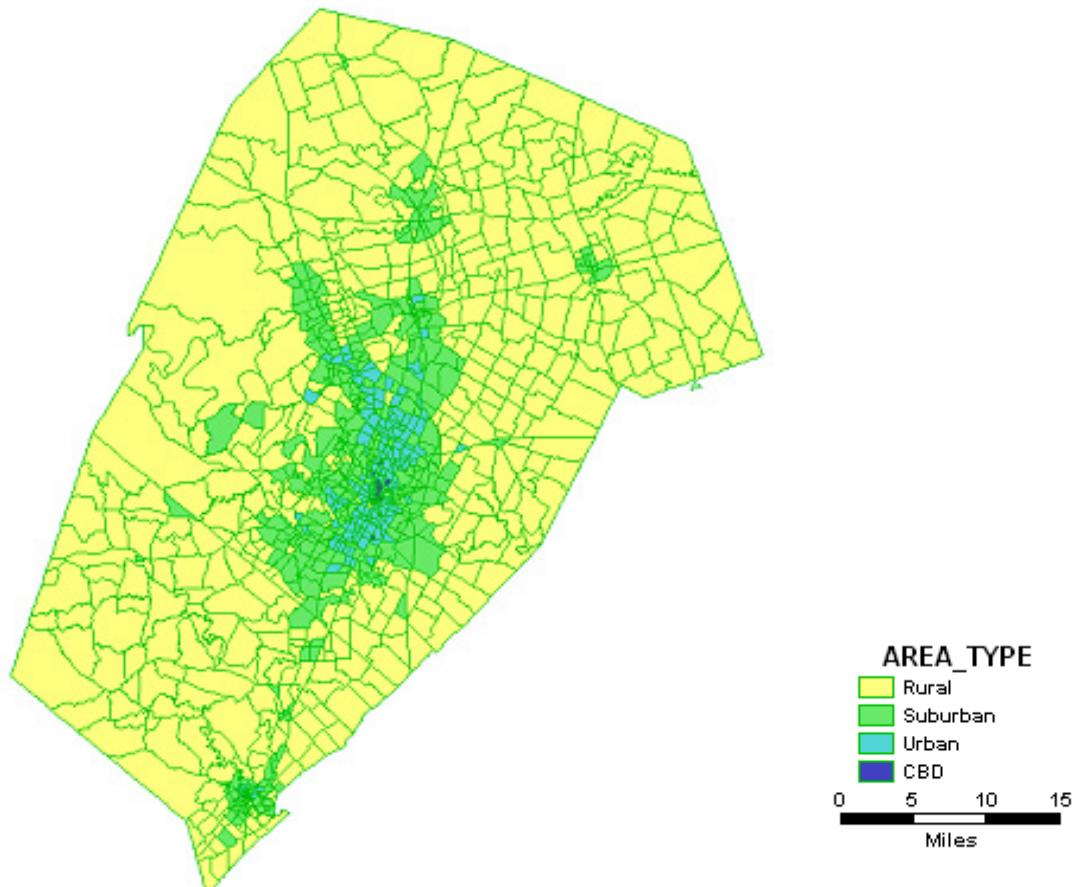


Figure 4.1: 3-County Austin Region (TAZs by Location Type)

## 4.1 Firm Evolution

An initial population of 32,063 firms and 45,180 households (10% sample) distributed over 1074 TAZs (Figure 4.2) was taken through the series of submodules at 1-year steps to forecast forward firm and household population, location and related trips 30 years into the future. Firm and household population profiles and trip generation and attraction profiles were updated at the end of each 1-year run of the simulation model for all TAZs. The following discussion focuses on firm dynamics.

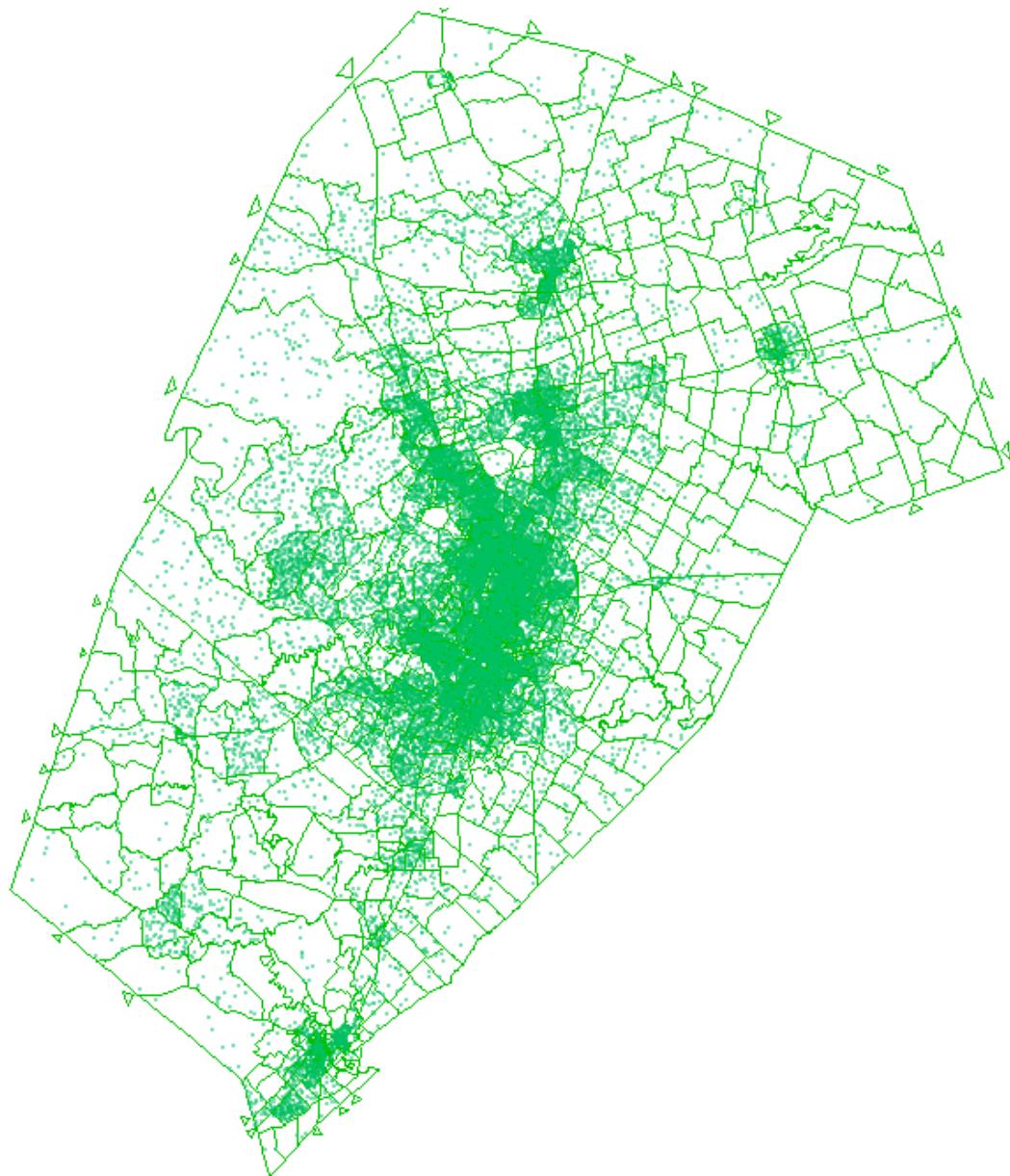


Figure 4.2: Distribution of Austin-area Firms in 2005 (Source: CAMPO)

Assuming year 2001/2002 rates of firm exit, birth and growth/contraction for different industry sectors and firm sizes, Figures 4.3 and 4.4 illustrate simulated changes in firm type profiles over time, for the low-growth scenario. These results suggest steady increase in all firm sectors and size groups, with this trend stabilizing over time. Markov transition probabilities developed from all years of the SUSB data set certainly appear skewed towards larger firms (possibly due to smaller sample sizes in this size group), producing higher growth in large firm sizes. The model ideally should be calibrated and/or parameter estimates constrained in order to produce results more consistent with expectations. Figures 4.5 and 4.6 show the anticipated changes in firm size and sector profiles over time for the high-growth scenario. As with the low-growth scenario, the simulation implies a steady increase in the number of firms in all sectors and all firm-size groups. However, the high-growth scenario's rate of change in firm counts exceeds that of the low-growth scenario.

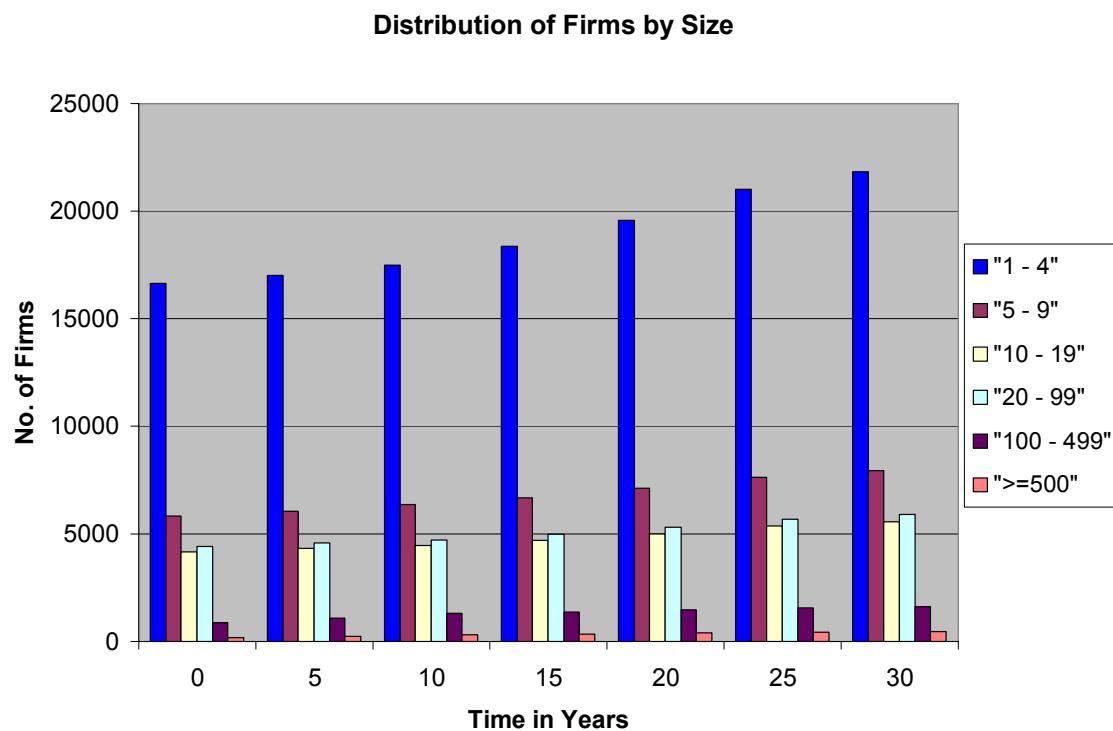


Figure 4.3: Firm Size Profile (Low-growth)

### Firm Distribution across Industry Sectors

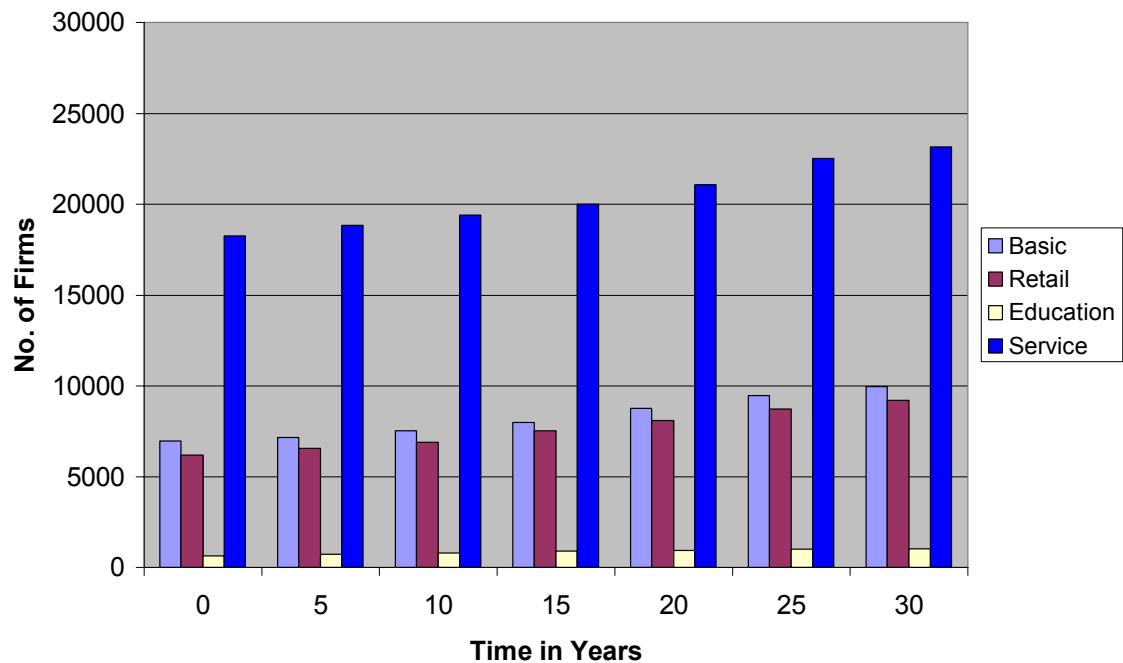


Figure 4.4: Firm Sector Profile (Low-growth)

### Distribution of Firms by Size

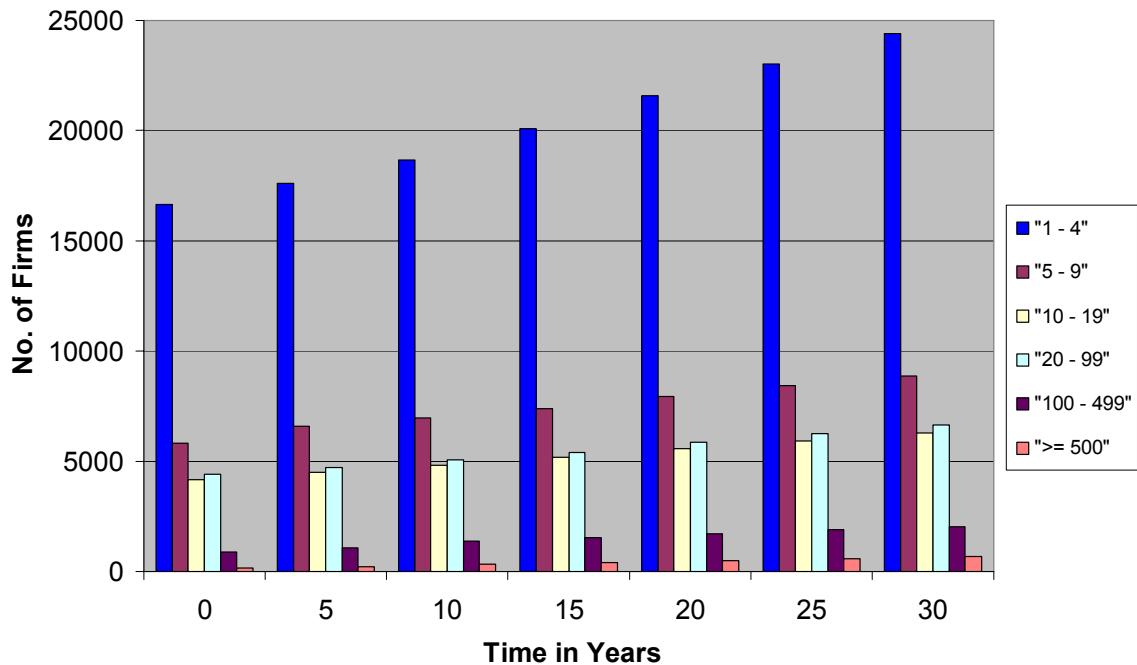


Figure 4.5: Firm Size Profile (High-growth)

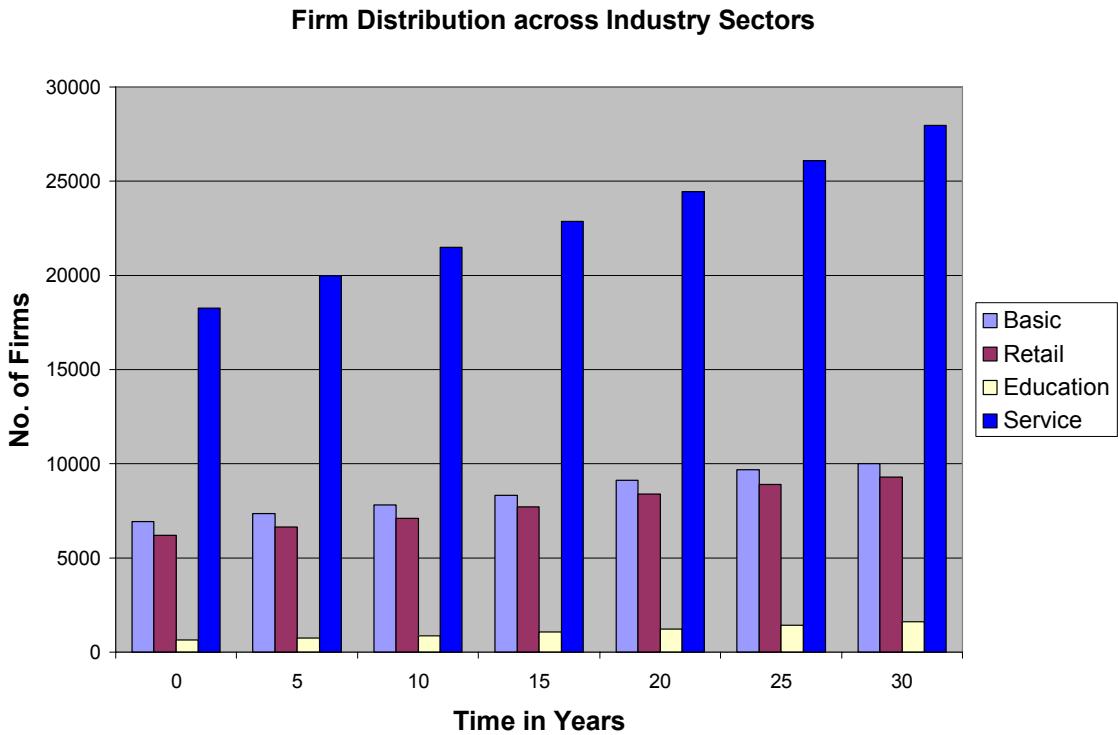


Figure 4.6: Firm Sector Profile (High-growth)

Expected changes in firm size and sector profiles over time for an average growth scenario are shown in Figures 4.7 and 4.8. The evolutionary pattern is similar to that of the low-growth and high-growth scenarios. Figures 4.9 and 4.10 show the distributions of jobs by firm size category and across industry sectors over the 30-year period. Job or employee counts per firm were randomly allotted to new firms in each size category using a right-truncated shifted exponential distribution, in order to ensure that the jobs met the category size requirements. In order to avoid job growth out-stripping population growth, maximum job growth rates in each size category were set at 3% per year. Due to Markov transition matrix estimates that favor shifts toward large-size (100+ employee) firms, job totals for the Austin region were generally simulated to grow at more than 3% a year for each of the first 20 years. Tracking jobs instead of firms and/or simultaneously tracking both jobs and firms is a possible enhancement for the model. Specifying ties between jobs and population growth models is also likely to be critical, to ensure reasonably synchronization in both sets of counts. Ultimately, in the average-growth-rate scenario, the total number of jobs was simulated to grow by 229% over the 30-year period.

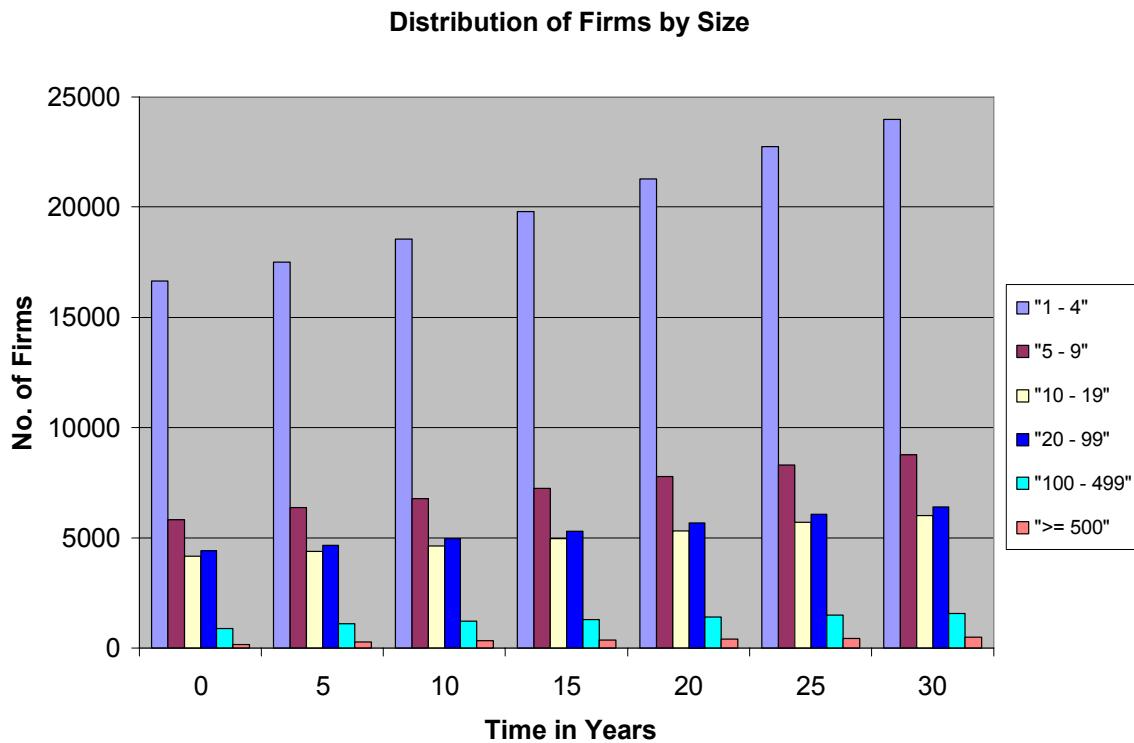


Figure 4.7: Firm Size Profile (Average growth)

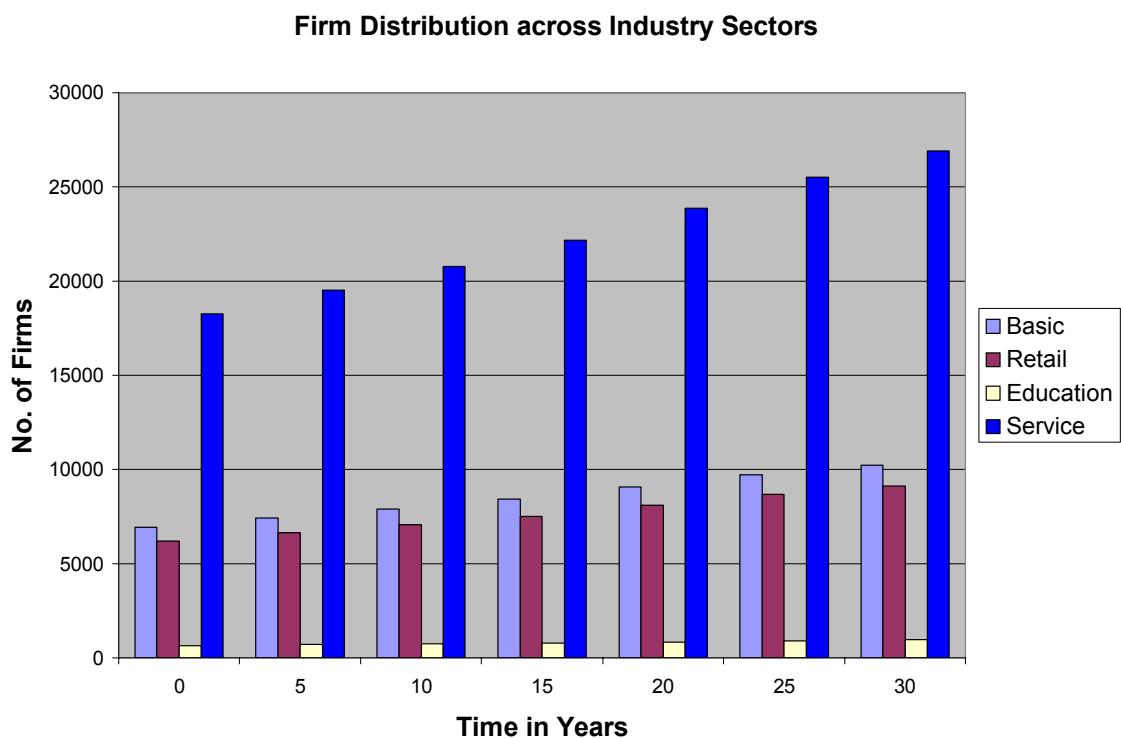


Figure 4.8: Firm Sector Profile (Average growth)

### Distribution of Jobs by Firm Size

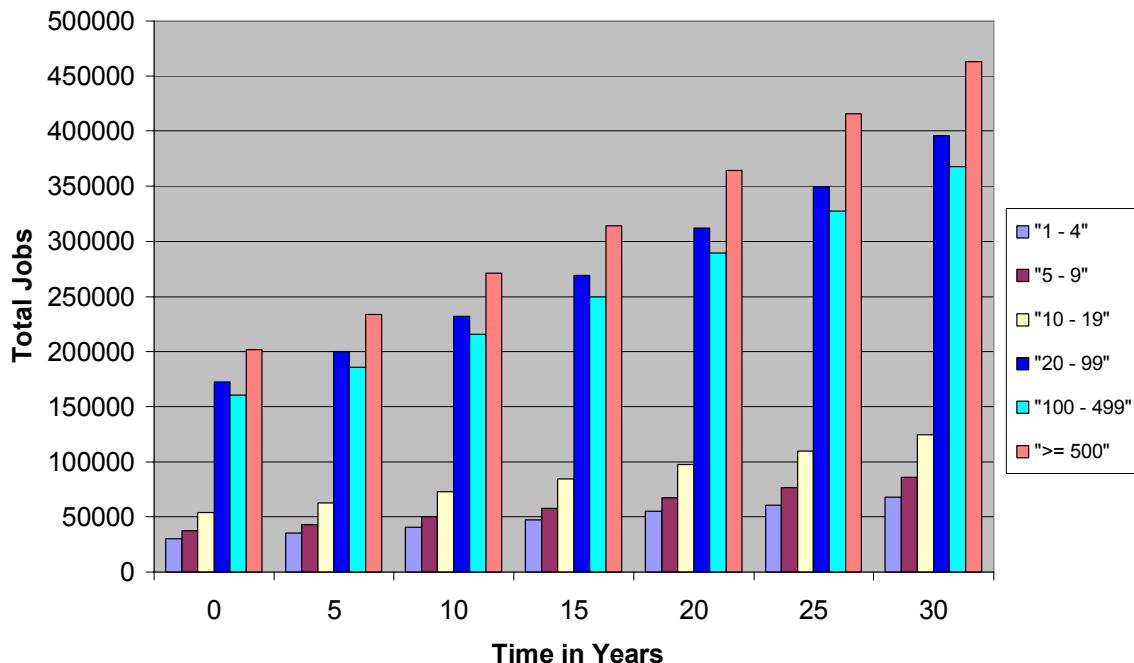


Figure 4.9: Job Distribution Profile by Firm Size (Average growth)

### Job Distribution across Industry Sectors

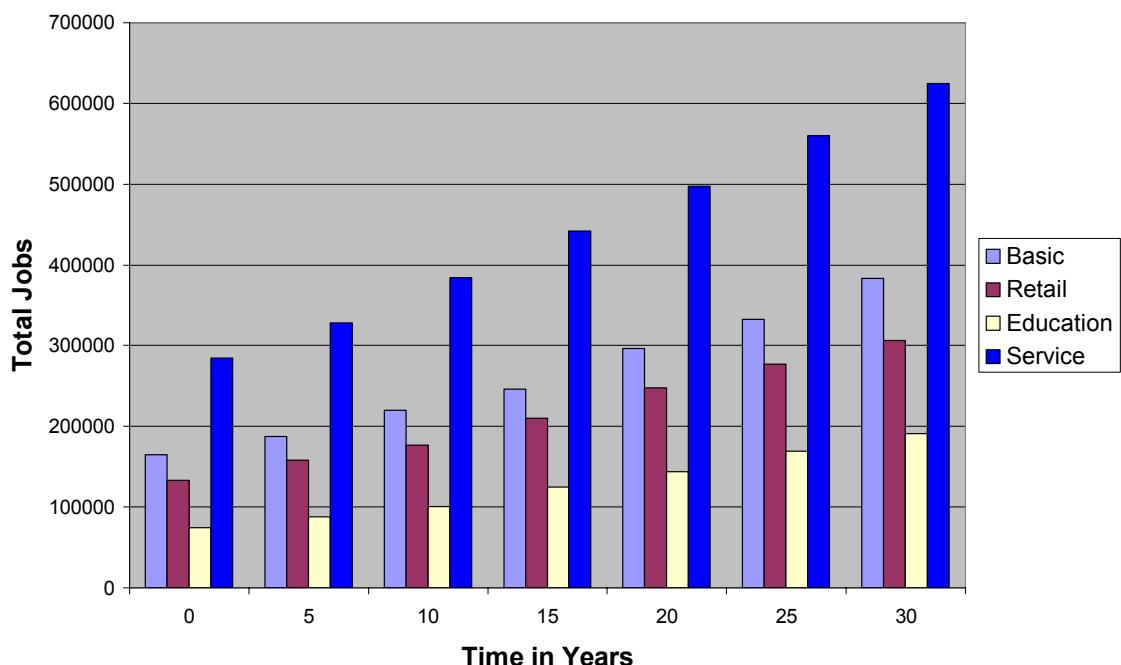


Figure 4.10: Job Distribution Profile across Industry Sectors (Average growth)

Figures 4.11 through 4.14 illustrate the model-predicted changes in the spatial distribution of firms (and thus employment) over time. The models predict greater spatial concentration than presently experienced in the Austin region, since, according to the location choice model, firms are expected to favor urban and CBD zones, in both the high-growth and low-growth scenarios. Such inward movement seems unlikely and is probably due to the use of cross-sectional data sets for calibration of the location choice model (rather than use of panel sets and a dynamic model focused on location changes) and a lack of information on land availability and development tendencies. Use of growth caps or controls and a model of land development would be very useful in this context, in order to sync demand with supply. These are useful model features to incorporate in the future.

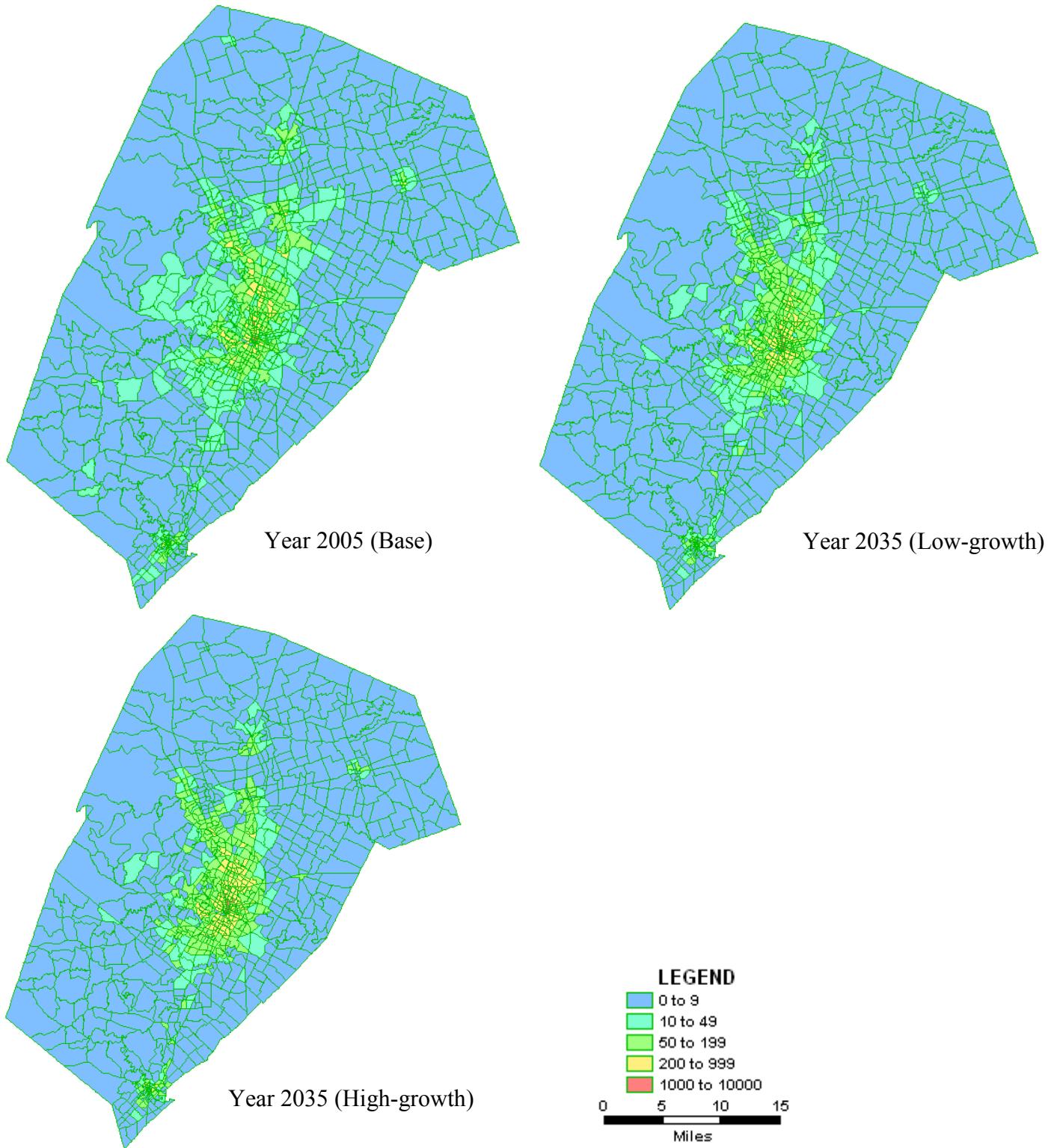


Figure 4.11: Firm Density (Firms per Square Mile) by Location in Year 2005 and Forecasts for Year 2035

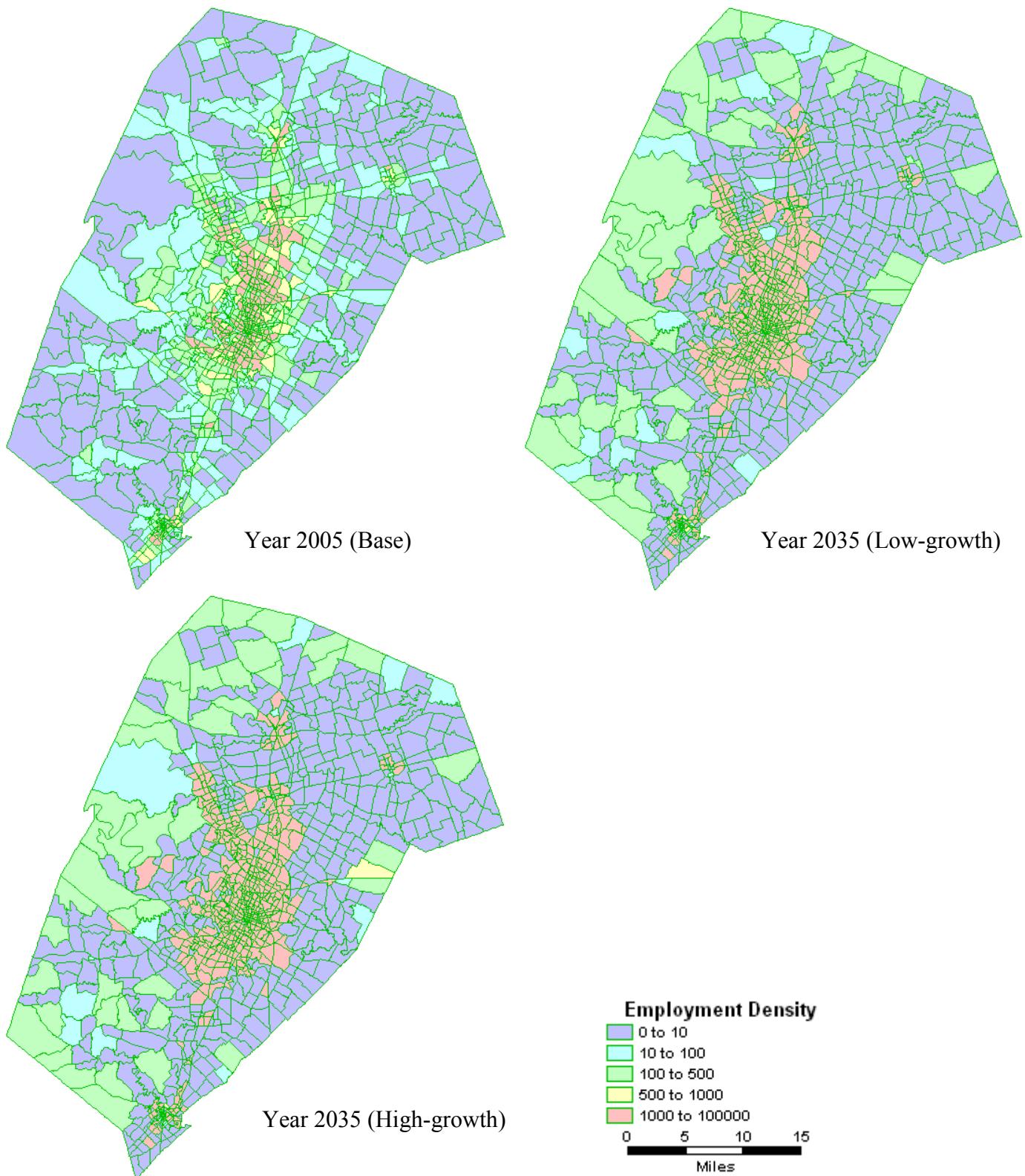


Figure 4.12: Employment Density (Jobs per Square Mile) by Location in Year 2005 and Forecasts for Year 2035

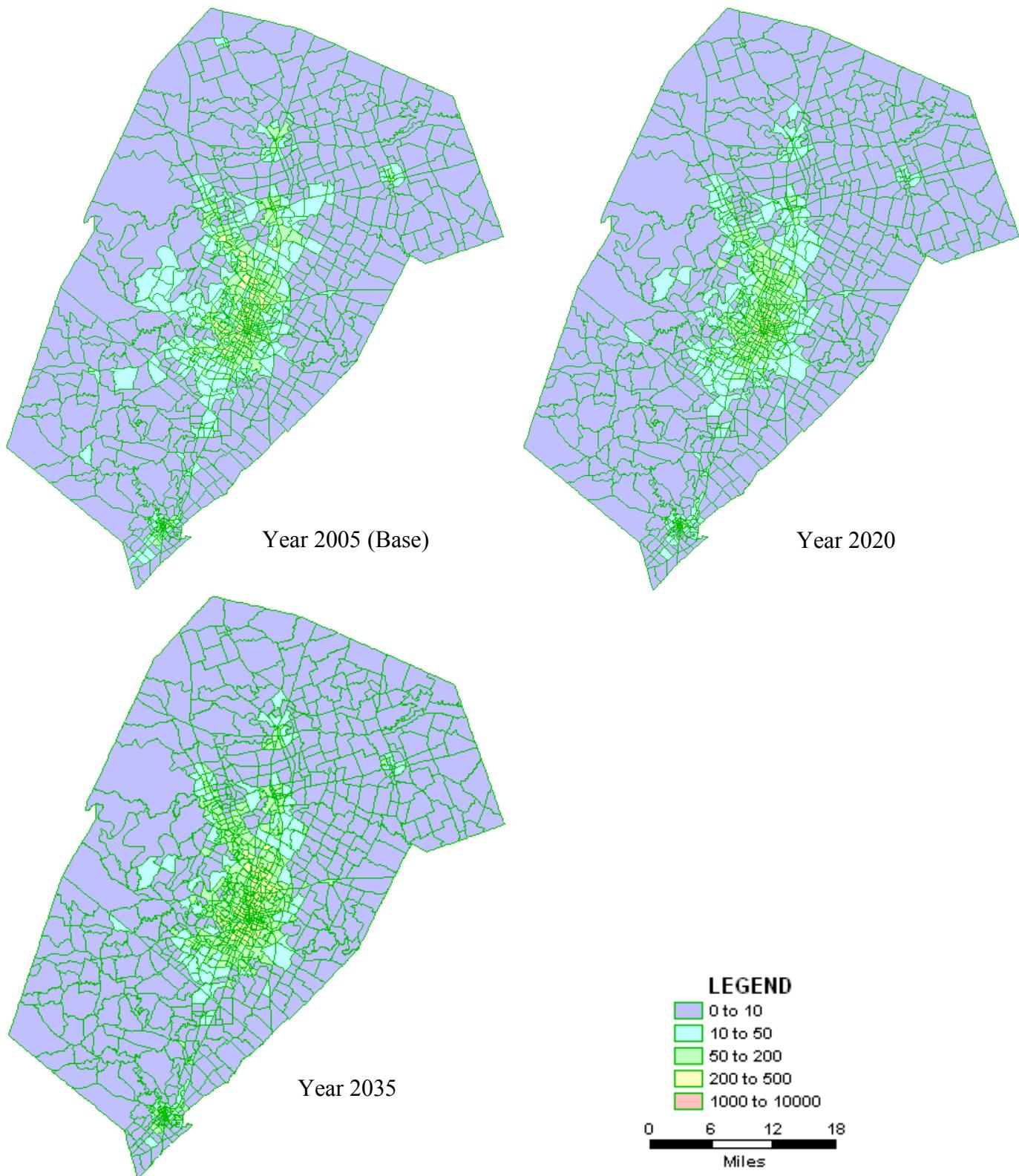


Figure 4.13: Firm Density (Firms per Square Mile) by Location in Year 2005 and Forecasts for Years 2020 and 2035 for Average Growth Scenario



Figure 4.14: Employment Density (Jobs per Square Mile) by Location in Year 2005 and Forecasts for Years 2020 and 2035 for Average Growth Scenario

Of course, households were simulated together with firms at 1-year steps to forecast Austin's future population patterns<sup>16</sup>. The following section discusses the results of household evolution demographically and spatially.

## 4.2 Household Evolution

Households are evolved using a 10-percent sample due to computational-time constraints<sup>17</sup>. An initial population of 115,268 individuals belonging to 45,180 households was taken through the series of submodules at 1-year steps to forecast household population, location and related trips 30 years into the future. Household structure and attributes of each individual were updated at the end of each run of the simulation. Household and individual population was scaled up by a factor of 10 while presenting results of the simulation.

Figure 4.15 shows the changing location patterns predicted for Austin households from 2005 through 2035. The models predict greater spatial concentration for households than currently visible in Austin, similar to the location results for firms. However, the concentration is much less pronounced than that for firms. While firms tend to segregate toward the urban and CBD zones, household patterns exhibit a much more widespread growth.

---

<sup>16</sup> A subset of this report is the paper by Kumar and Kockelman titled "Tracking the Size, Location and Interactions of Businesses: Microsimulation of Firm Behavior in Austin, Texas" and is listed in References.

<sup>17</sup> Run time for Austin's 100% household simulation model over 30 years required more than 3 weeks on a 1-GB RAM machine with 1.66 GHz processor.



Figure 4.15: Household Density (Households per Square Mile) by Location in Year 2005 and Forecasts for Year 2035

Table 4.1 shows statistics characterizing trends in population attributes for the average growth and no-toll scenario. Households and population are forecast to grow by 176% and 185% respectively, over the 30-year period. Total (personal) vehicles exhibit an increase of 209%, while average vehicle ownership (per household) is simulated to rise just 19%. Average household income is expected to increase a striking 93% during the same period, which implies an annual growth rate of 2.2%, consistent with assumptions.

Table 4.1: Forecasting Evolution of Population Attributes over Time

	Year 2005		Year 2015		Year 2035	
	451,003		601,214		793,239	
	1,148,177		1,575,103		2,129,252	
	Std.		Std.		Mean	Std. Dev
	Mean	Dev	Mean	Dev	Mean	Std. Dev
Households	2.55	1.47	2.62	1.58	2.68	1.66
Population	1.94	0.95	2.09	1.01	2.31	1.09
HH size	\$59,496	\$51,542	\$81,408	\$82,345	\$114,902	\$117,064
Vehicles						
Income						
Fraction of households with...						
Pre-school children	0.20	0.52	0.21	0.54	0.22	0.58
Pre-driving school-age children	0.36	0.76	0.37	0.77	0.39	0.78
Driving school-age children	0.06	0.26	0.07	0.26	0.07	0.28
Non-working adults	0.26	0.58	0.27	0.60	0.27	0.60
Student adults	0.19	0.53	0.21	0.54	0.23	0.55
Part-time working adults	0.42	0.65	0.43	0.66	0.43	0.66
Full-time working adults	1.05	0.94	1.07	0.99	1.08	1.07

Similar to commercial trip model applications, household trip generation and distribution models were updated at the end of each simulation period. However due to run-time considerations, the traffic assignment module was run only every 5 years; hence, travel times remained constant between successive runs of the assignment module. Since population and employment distributions changed every year, of course, accessibility indices and densities were updated yearly, impacting firm

and household location choices. The following section discusses the results of the travel demand model and various assumptions that underlie it.

### 4.3 Travel Demand Model

Household and commercial trip generation and distribution models were run during each simulation period; while, external-internal and external-external trips were exogenous to the simulation model (which is a limitation of the current model and provides scope for improvement) and added to the OD matrix for traffic assignment. External-internal trips and external-external trips were assumed to increase by a constant factor (5% every year). This work used a static traffic assignment routine and neglected time-of-day effects. For simplicity, no mode choice model was applied and traffic is based on a 24-hour count and a vehicle occupancy factor of 1.2. Link capacities are assumed to be 35% of actual capacities. This neglect of time-of-day and mode-choice are simplifying assumptions that provide opportunities for future model improvements.

The average growth-rate scenario (for firm additions to the region) was used for variations in the travel demand scenarios. Recall that firms do not have a major influence on travel demand, since only 8.43% of Austin's estimated vehicle miles traveled (VMT) in 2005 come from commercial trips<sup>18</sup>. Figures 4.16 and 4.17 show the comparison of link flow data for years 1997, 2005 (base year), 2020 and 2035 for no-toll and fixed-toll scenarios, respectively. The 1997 link flow data came from 1997 traffic counts performed by CAMPO in the Austin region. Link flow volumes show a steady increase for all functional class of roadways in both scenarios. However, imposition of a fixed 10 cents-per-mile toll on all links causes VMT to rise at a lower rate. VMT shows a 327% and 319% rise over the 30-year period for no-toll and fixed-toll scenarios, respectively.

---

<sup>18</sup> According to CAMPO's truck/commercial trip tables for year 1997, approximately 9.5% of Austin's total VMT comes from trucks/commercial trips.

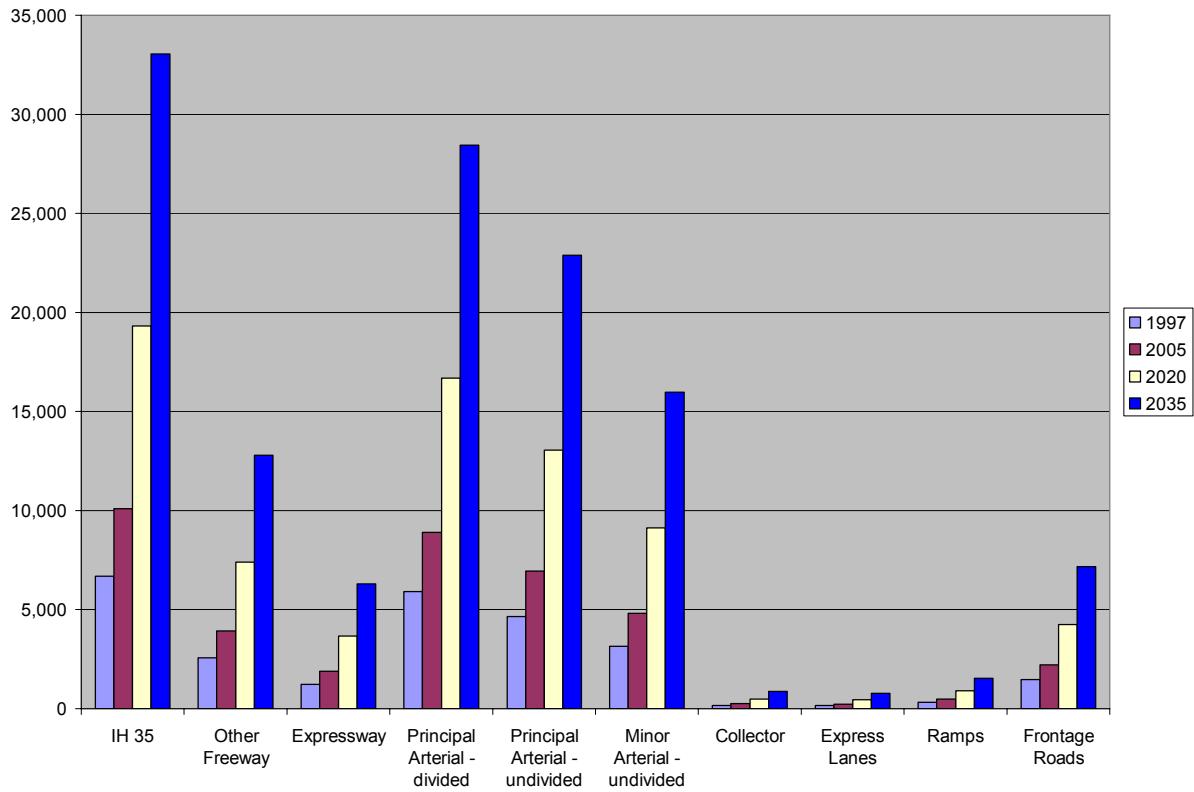


Figure 4.16: VMT Associated with Each Functional Class by Time for No-toll Scenario

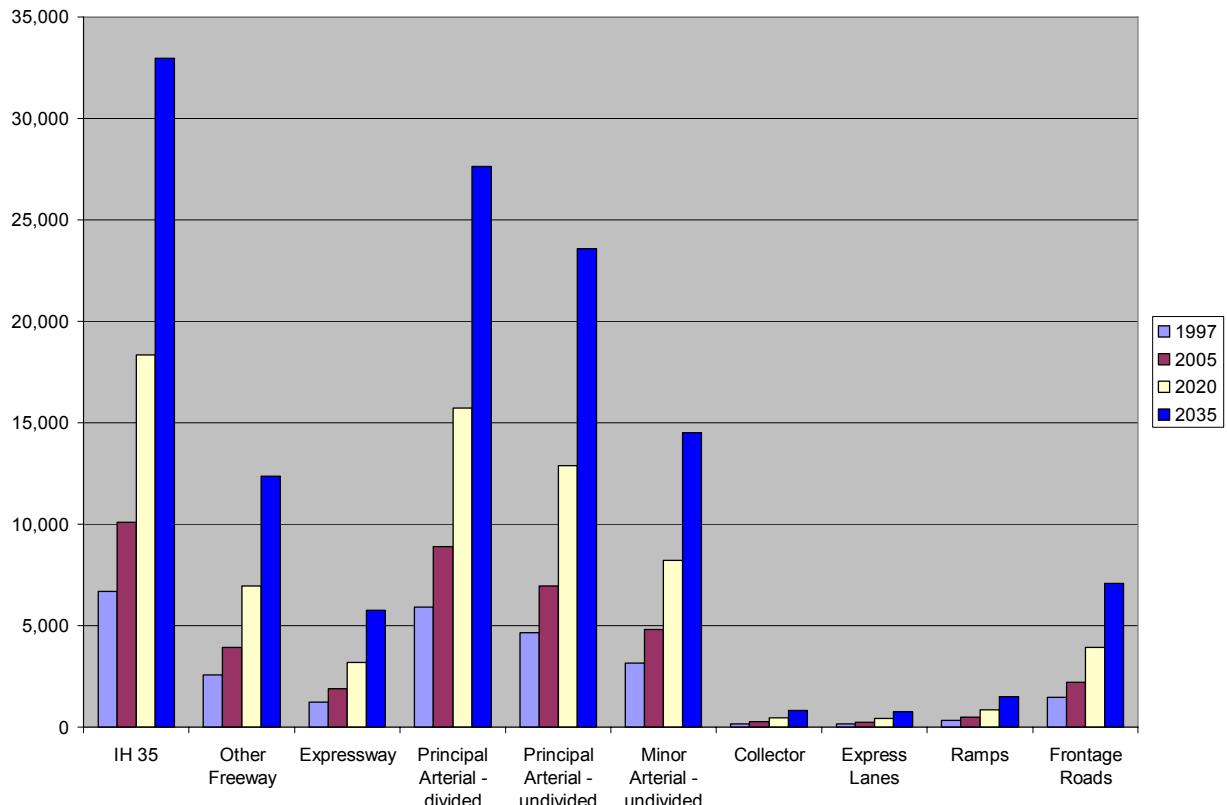


Figure 4.17: VMT Associated with Each Functional Class by Time for Fixed-toll Scenario

## 4.4 Summary

The application of the microsimulation model in Austin region provides unexpected results with respect to household and firm population forecasts. While many of the life-cycle transition models are based on simplifying assumptions, they easily can be enhanced, as data become available. Application of the microsimulation model in Austin region based on cross-sectional local data shows firms to have an unlikely high preference for central zones when selecting new locations. This is no doubt due to the lack of a built-space or land-development model and static nature of the location choice models' calibration. Estimation of Markov transition matrices using aggregate firm data also was necessitated by a lack of appropriate data, and transition probabilities were found to be skewed towards large firm size. The household evolution model is built on a set of rules based on many assumptions about birth, marriage and household formation behavior. Various simplifying assumptions made in the household evolution model, partly due to data constraints and partly due to computation time constraints, can be easily updated, as appropriate data become available. Both the no-toll and fixed (10 cents-per-mile toll) scenarios exhibit a steady increase in VMT on all functional classes of roadways over the 30-year period, with the VMT increases being greater in the no-toll scenario, as expected. The following chapter provides conclusions as well as ideas for future research directions.

## **CHAPTER 5: SUMMARY, CONCLUSIONS AND FUTURE WORK**

This work developed a basic microsimulation framework for household and firm evolution, coupled with models of household and commercial travel demand. It provides details for all underlying firm and household evolution models as well as discussion of their current limitations. It provides a detailed comparison of logit and Poisson models of location choice, and uses a constrained optimization technique to estimate Markovian transition probabilities of firms by size from aggregate data. Most importantly, it demonstrates that a microsimulation model of firms and households is feasible using largely existing datasets and standard desktop computing.

In terms of actual results, the number of firms in all sectors and firm-size groups show a steady increase with varying rates for all growth scenarios. Job and firm counts exhibit similar evolutionary patterns over time. In the average-growth-rate scenario, the total number of jobs is estimated to rise 229% over the 30-year period. The models predict greater spatial concentration than presently experienced in the Austin region for firms and households. Households and population are forecast to grow by 176% and 185% respectively, over the 30-year period. Total (personal) vehicles exhibit an increase of 209%, while average vehicle ownership (per household) is simulated to rise just 19%. Average household income is expected to increase a striking 93% during the same period, which implies an annual growth rate of 2.2%, consistent with assumptions. Both the no-toll and fixed (10 cents-per-mile toll) scenarios exhibit a steady increase in VMT on all functional classes of roadways over the 30-year period, with the VMT increases being greater in the no-toll scenario, as expected.

This study estimated a microscopic simulation model using various available data sets for the Austin, Texas region. While such microsimulation is data and computing intensive, it provides a flexible tool for analyzing the impacts of various policy decisions as well as other, demographic, environmental and system changes. It allows transportation planners explore the potential responses of individuals to changes in their environments and predict the long-term implications of policy decisions. This work seeks to be a bridge for further integrated travel demand and land use models of this type. Key items for near-term model improvement are discussed here now.

## **5.1 Limitations and Enhancements**

This work pulls together a variety of detailed models and assumptions for anticipating the population, firm, land use and traffic patterns of a region's future. Existing integrated travel demand and land use models vary tremendously in their levels of spatial and temporal detail. There is always room for improvement, and the present modeling endeavor is no exception. The biggest challenge in developing a microsimulation model of household and firm evolution is the lack of data. Obviously, a model can only be as good as the data that underlies it. In theory, microsimulation models require detailed panel data on households and firms, but few such data sets exist. In modeling firm evolution, panel data would be useful for estimating models of birth, death, expansion/contraction and dynamic models of location choice. Similarly, models of various transition events in the life of households require panel data sets. Modeling freight traffic requires data on commercial vehicle movements, which also is scarce. However, as such data becomes available, various assumptions made in Chapter 3 can be relaxed.

As alluded to, almost every model estimated as part of the simulation model presents scope for improvement, in terms of incorporating more attributes. Many simplifying assumptions such as random sampling, Markov processes, constant birth and death rates, and constant migration rates can be relaxed. Assumption of asymmetric distributions for income updating and vehicle retirement is too simplistic and requires more research. Currently the long-distance travel model does not track distance traveled during each trip or anticipate (in a behavioral fashion) the external zones used, and external-internal and external-external trips are exogenous to the modeling framework. On the firm simulation part, more research is required in estimating Markov transition matrices. Also, firm relocation model and birth and death models need to be estimated using some panel data source. Even the travel demand model provides scope for future improvement provided the simplifying assumptions that underlie them. These and other opportunities for model enhancements are discussed below.

### **5.1.1 Computation-time**

The computational requirements of microsimulation model are regularly of concern, and provided a major challenge in this work. Tracking 450,003 households and more than 1 million individuals over a span of 30 years is tedious. A full run of the household evolution model required more than 3 weeks on a standard desktop machine (1-GB RAM with 1.66 GHz processor), while the model of firm evolution required 5 hours for a complete simulation run. This necessitated the use of a

reduced (10%) sample of households here. Similarly, most statistical software has limitations on the amount of data it can handle. The use of Poisson regression models, rather than logit models, for estimation of firm and household location choice parameters was partly a response to this limitation. While a Poisson model is a highly tractable alternative to a logit model, it does not allow the use of decision-maker-specific attributes. Estimating a microsimulation model at this scale requires estimating more than 100 parameters and tracking hundreds of attributes associated with each individual, household, firm and TAZ.

Due to time constraints, only one run of the complete simulation was run. Ideally more simulation should be run, to exploit and appreciate the variability inherent in the Markov processes underlying these models' applications, though these are time-consuming.

### **5.1.2 Household and Firm Evolution Model**

One important limitation of this model is its reliance on cross-sectional data to estimate models of location choice. This model also lacks a land-use model for space development, to ensure location demands can be met and/or send price signals via rents. It does not track property prices or homes and offices. Absence of a microsimulation model to track these attributes and a lack of market feedback mechanism is a handicap in effective modeling of land use. Assumption of a Markov process for firm sizes is also a potentially significant issue: firms that have endured longer will experience different 1-year shifts than those that have not been around so long. Skewed Markov transition matrix probabilities were observed initially (leading to heavy growth in the largest firm sizes), and the simplifying assumption of  $p_{ij} = 0$  when  $|i - j| > 1$  helped achieve reasonable transition probabilities. The model's random selection of households and firms for death and birth decisions also could be enhanced, as data sets become available.

Relaxing the various simplifying assumptions described in Chapter 3 – such as simple triangular distributions for income updating and vehicle retirements, Markov processes for firms – is an obvious next step. Assuming a constant rate of firm birth and growth is a weakness, and models for these may well enhance the behavioral foundation of the larger model. Location choice model for firms can be enhanced with and without the use of panel data by introducing firm-specific behaviors. Assuming fixed proportion of firms and households relocating every year is a significant simplification, and better models of relocation will be useful in understanding the changing land-use patterns of a region. Migration module for households assumes a net in-migration rate implying that

any household in the study area never leaves it. This is a major simplifying assumption as the household profile of in-migrants can be quite different from that of out-migrants.

### **5.1.3 Vehicle Ownership Model Limitations**

Vehicle ownership decisions also would be useful for anticipating a region's greenhouse gas emissions and travel patterns. A dynamic model of vehicle ownership also would be useful. In order to ascertain carbon futures, vehicle ownership models need to be estimated at the level of vehicle makes and models. Estimating a vehicle retirement model and tracking vehicle exchanges over the simulation period could enhance forecasting abilities of the model. Long-distance travel also contributes a fair share of carbon and other emissions, and forms an important part of inter-state highway traffic loads. Distance and mode choice models of long-distance travel as a function of gas prices are needed to provide estimates of transport's share in greenhouse gas emissions. Tying a model of household energy demands to the microsimulation model also would enhance estimation of households' contribution to greenhouse gas emissions. The microsimulation model discussed in this work provides a basic framework and demonstrates the feasibility of such models. There are ample opportunities to enhancement and extension.

### **5.1.4 Travel Demand Model Limitations**

The system's travel demand model is limited by several simplifying assumptions, including the neglect of trip scheduling (i.e., time-of-day) and mode choice. Traffic assignment is run only once every 5 years, so that travel times remain unchanged in between, though location patterns, households and firms continue to evolve in those periods.

In summary, this work offers a basic simulation framework for integrated microscopic simulation models of household and firm evolution, land use and travel demand. It also provides a valuable demonstration of how a wide variety of behavioral models can be used in concert, and how individual firms and households can be tracked over several decades, using standard software and computing resources. Enhancements are coming, and opportunities for many interesting and relevant policy and growth scenarios are feasible.

## REFERENCES

- Abraham, J. E., and J. D. Hunt. 2000. Firm Location in the MEPLAN Model of Sacramento. *Transportation Research Record* 1685: 187-198.
- Arentze, T.A., and H. J. P. Timmermans. 2004. A Learning-Based Transportation Oriented Simulation System. *Transportation Research B* 38: 613-633.
- Bartik, T. 1985. Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States. *Journal of Business and Economic Statistics* 3: 14–22.
- Ben-Akiva, M., and S. R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, Cambridge.
- Ben-Akiva, M., and J. L. Bowman. 1998. Integration of an activity-based model system and a residential location model. *Urban Studies* 35(7): 1131-1153.
- Berkovec, J. 1985. Forecasting Automobile Demand Using Disaggregate Choice Models. *Transportation Research Part B* 19(4): 315-329.
- Bhat, C.R. 1995. A Heteroskedastic Extreme Value Model of Intercity Travel Mode Choice. *Transportation Research Part B: Methodological* 29(6): 471-483.
- Bhat, C.R., and J. Y. Guo. 2004. A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling. *Transportation Research Part B* 38(2): 147-168.
- Bhat, C.R., J. Y. Guo, S. Srinivasan, and A. Sivakumar. 2004. A Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns. *Transportation Research Record* 1894: 57-66.
- Bina, M., D. Suescun, and K. Kockelman. 2006a. Location Choice vis-à-vis Transportation: The Case of Recent Home Buyers. Proceedings of the 85th Annual Meeting of the Transportation Research Board, Washington, D.C.
- Bina, M., V. Warburg, and K. Kockelman 2006b. Location Choice vis-à-vis Transportation: The Case of Apartment Dwellers. *Transportation Research Record* 1977: 93-102.
- Borgers, A., and H. J. P. Timmermans. 1993. Transport facilities and residential choice behavior: a model of multi-person choice processes. *Papers in Regional Science* 72 (1): 45–61.
- Brouwer, A. E., I. Mariotti, and J. N. van Ommeren. 2002. The Firm Relocation Decision: A Logit Model. Paper presented at the 42nd ERSA conference, Dortmund, Germany.
- Carroll G. R., and M. T. Hannan. 2000. *The Demography of Corporations and Industries*. Princeton University Press, Princeton.
- Coughlin, C., J. Terza, and V. Arromdee. 1991. State Characteristics and the Location of Foreign Direct Investment within the United States. *this REVIEW* 73: 675–683.

- Coughlin, C., and E. Segev. 2000. Location Determinants of new Foreign-owned Manufacturing Plants. *Journal of Regional Science* 40(2): 323-351.
- De Bok, M. and F. Sanders. 2005. Firm Location and the Accessibility of Locations: Empirical Results from the Netherlands. Proceedings of the 84th Annual Meeting of Transportation Research Board, Washington DC.
- Elgar, I., and E. J. Miller. 2006. A Conceptual Model of Small Office Firm Location. Proceedings of the 85<sup>th</sup> Annual Meeting of the Transportation Research Board, Washington DC, 22-26 January.
- Erhardt, G. D., J. Freedman, A. Stryker, H. Fujioka, and R. Anderson. 2006. The Ohio Long Distance Travel Model. Proceedings of the 86th Annual Meeting of the Transportation Research Board, Washington, D.C.
- Evans, D. 1987. The Relationship between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries. *Journal of Industrial Economics* 35: 567–581.
- FHWA. 2006. TRANSIMS Home Page. URL: <http://tmip.fhwa.dot.gov/transims/>. Accessed 10<sup>th</sup> October 2006.
- Firm Size Data. Office of Advocacy, U. S. Small Business Administration. URL: <http://www.sba.gov/advo/research/data.html#us>. Accessed 10<sup>th</sup> July 2007.
- Friedman, J., D. Gerlowski, and J. Silberman. 1992. What Attracts Foreign Multinational Corporations? Evidence from Branch Plant Location in the United States. *Journal of Regional Science* 32: 403–418.
- Gabriel, S. A., and S. S. Rosenthal. 1989. Household Location and Race: Estimates of a Multinomial Logit Model. *The Review of Economics and Statistics* 71(2): 240-249.
- Goulias, K. G., and R. Kitamura. 1992. Microsimulation for Travel Demand Forecasting: A Dynamic Model System of Household Demographics and Mobility. Institute of Transportation Studies Research Report, UCD-ITS-RR-92-4, University of California, Davis.
- Guimaraés, P., O. Figueirido, and D. Woodward. 2000. Agglomeration and the Location of Foreign Direct Investment in Portugal. *Journal of Urban Economics* 47: 115–135.
- Guimaraés, P., O. Figueirido, and D. Woodward. 2003. A Tractable Approach to the Firm Location Design Problem. *The Review of Economics and Statistics* 85(1): 201–204.
- Guimaraés, P., R. Rolfe, and D. Woodward. 1998. Regional Incentives and Industrial Location in Puerto Rico. *International Regional Science Review* 21(2): 119–138.
- Hannan, M. T., G. R. Carroll, S. D. Dobrev, and J. Han. 1988. Organizational Mortality in European and American Automobile Industries Part I: Revisiting the Effects of Age and Size. *European Sociological Review* 14: 279-302.

- Hansen, E. 1987. Industrial Location Choice in Sao Paulo, Brazil: A Nested Logit Model. *Regional Science and Urban Economics* 17: 89–108.
- Head, K., J. Ries, and D. Swenson. 1995. Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States. *Journal of International Economics* 38: 223–247.
- Hayter, R. 1997. *The Dynamics of Industrial Location. The Factory, The Firm and The Production System*. New York: Wiley.
- Hunt, J. D., D. S. Kriger, and E. J. Miller. 2005. Current Operational Urban Land Use—Transport Modelling Frameworks: A Review. *Transport Reviews* 25(3): 329–376.
- Hunt, J. D., K. J. Stefan, A. T. Brownlee, J. D. P. McMillan, A. Farhan, K. Tsang, D. Atkins, and M. Ishani. 2004. A Commercial Movement Modelling Strategy for Alberta's Major Cities. Paper prepared for presentation at the 2004 Annual Conference of the Transportation Association of Canada, Québec City, Québec. <http://www.tac-atc.ca/english/pdf/conf2004/hunt.pdf>. Accessed 8<sup>th</sup> March 2007.
- Jones, M. T. 2005. Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk. *IMF working paper*, Monetary and Financial Systems Department. <http://www.imf.org/external/pubs/ft/wp/2005/wp05219.pdf>. Accessed 10<sup>th</sup> April 2007.
- Kim, H. C., M. H. Ross, and G. A. Keoleian. 2004. Optimal Fleet Conversion policy from a Life Cycle Perspective. *Transportation Research Part D* 9: 229-249.
- Khan, A.S., J.E. Abraham, and J.D. Hunt. 2002. Agent-based Microsimulation of Business establishments. Paper presented at the 42nd ERSA Congress, Dortmund. <http://rwp.raumplanung.uni-dortmund.de/en/index.htm>. Accessed 10<sup>th</sup> November 2006.
- Kockelman, K.M., and Y. Zhao. 2000. Behavioral Distinctions: The Use of Light-Duty Trucks and Passenger Cars. *Journal of Transportation Statistics* 3(3): 47-60.
- Koppelman, F.S., and M. Hirsh. 1986. Intercity Passenger Decision Making: Conceptual Structure and Data. *Transportation Research Record* 1085: 70-75.
- Koppelman, F.S. 1989. Multidimensional Model System for Intercity Travel Choice Behavior. *Transportation Research Record* 1241: 1-8.
- Kumar, Saurabh and Kara Kockelman. "Tracking the Size, Location and Interactions of Businesses: Microsimulation of Firm Behavior in Austin, Texas." Tentatively accepted for publication in *Transportation Research Record*, Journal of the Transportation Research Board, February 2008.
- Kroll, C., and J. Landis. 1990. Housing Prices, Other Real Estate Factors and the Location Choice of Firms. *Fisher Center for Real Estate & Urban Economics*. <http://repositories.cdlib.org/iber/fcreue/reports/0990>. Accessed 26<sup>th</sup> April 2007.
- Lave, C.A., and K. Train. 1979. A Disaggregate Model of Auto-Type Choice. *Transportation Research Part A* 13 (1): 1-9.

- Lee, J.H., K. S. Chon, and C. Park. 2004. Accommodating Heterogeneity and Heteroscedasticity in Intercity Travel Mode Choice Model: Formulation and Application to Honam, South Korea, High-Speed Rail Demand Analysis. *Transportation Research Record* 1898: 69-78.
- Lee, T. C., G. G. Judge, and A. Zellner. 1970. *Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data*. Amsterdam: North Holland.
- Lee, Y. 2006. Relocation Patterns in U.S. Manufacturing. *Working Paper 06-24*, Federal Reserve Bank of Cleveland. [www.clevelandfed.org/research](http://www.clevelandfed.org/research). Accessed 9<sup>th</sup> January 2007.
- Mannering, F., and C. Winston. 1985. A Dynamic Empirical Analysis Of Household Vehicle Ownership And Utilization. *Rand Journal of Economics* 16(2): 215-236.
- Mannering, F., C. Winston, and W. Starkey. 2002. An Exploratory Analysis of Automobile Leasing by US Households. *Journal of Urban Economics* 52: 154-176.
- Maoh, H. F., and P. S. Kanaroglou. 2005. Agent-Based Firmographic Models: A Simulation Framework for the City of Hamilton. Proceedings of PROCESSUS Second International Colloquium on the Behavioural Foundations of Integrated Land-use and Transportation Models: Frameworks, Models and Applications, Toronto. [http://www.civ.utoronto.ca/sect/traeng/ilute/processus2005/PaperSession/Paper08\\_Maoh-Kanaroglou\\_Agent-basedFirmographicModels\\_CD.pdf](http://www.civ.utoronto.ca/sect/traeng/ilute/processus2005/PaperSession/Paper08_Maoh-Kanaroglou_Agent-basedFirmographicModels_CD.pdf). Accessed 8<sup>th</sup> November 2006.
- McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. (pp. 105–142), in P. Zarembka (Ed.). *Frontiers in Econometrics*. New York: Academic Press.
- McFadden, D. 1978. Modeling the Choice of Residential Location. (pp. 75–96), in A. Karquist, L. Lundqvist, F. Snickars, and J. Weibull (Eds.). *Spatial Interaction Theory and Planning Models*. Amsterdam: North-Holland.
- McWethy, L. M. 2006. Comparing Microscopic Activity-Based and Traditional Models of Travel Demand: An Austin Area Case Study. Master's Thesis. Department of Civil Engineering, University of Texas at Austin.
- Miller, E. J., and P. A. Salvini. 1998. The Integrated Land Use, Transportation, Environment (ILUTE) Modeling System: A Framework. Proceedings of the 77<sup>th</sup> Annual Meeting of the Transportation Research Board, Washington DC, 12-15 January.
- Miller, E. J., and P.A. Salvini. 2001. The Integrated Land Use, Transportation, Environment (ILUTE) Microsimulation Modelling system: Description & Current Status. Chapter 41, in D. Hensher (ed.) *The Leading Edge in Travel Behaviour Research*. Amsterdam: Pergamon, 711-724.
- Miller, E.J., J.D. Hunt, J.E. Abraham and P.A. Salvini. 2004. Microsimulating Urban Systems. *Computers, Environment and Urban Systems* special issue, “Geosimulation: Object-Based Modeling of Urban Phenomena” 28: 9-44.
- Moeckel, R. 2005. Simulating Firmography. Fourth Oregon Symposium on Integrating Land Use and Transportation Models, 15-17 November.

[http://www.oregon.gov/ODOT/TD/TP/docs/Modeling/4symp/1117\\_1045.pdf](http://www.oregon.gov/ODOT/TD/TP/docs/Modeling/4symp/1117_1045.pdf). Accessed 5<sup>th</sup> November 2006.

Mohammadian, A., and E. J. Miller. 2003. Empirical Investigation of Household Vehicle Type Choice Decisions. *Transportation Research Record* 1854: 99-106.

Onaka, J. L., and W. A. V. Clark. 1983. A Disaggregate Model of Residential Mobility and Housing Choice. *Geographical Analysis* 15: 287-304.

Orcutt, G. A new type of Socio-economic System. *Review of Economics and Statistics* 58: 773-797.

Otter, H. S., A. van der Veen, and H. J. de Vriend. 2001. ABLOoM: Location Behaviour, Spatial Patterns, and Agent-based Modelling. *Journal of Artificial Societies and Social Simulation* 4(4). <http://www.soc.surrey.ac.uk/JASSS/4/4/2.html>. Accessed 5<sup>th</sup> May 2007.

Pellenbarg, P. H., and P. J. M. Van Steen. 2003. Spatial Perspectives on Firm Dynamics in the Netherlands. *Tijdschrift voor Economische en Sociale Geografie* 94(5): 620-630.

Pellenbarg, P. H., L. J. G. Van Wissen, and J. van Dijk. 2002. Firm Relocation: State of the Art and Research Prospects. *SOM Research Report 02D31*, University of Groningen. <http://som.eldoc.ub.rug.nl/FILES/reports/themeD/2002/02D31/02D31.pdf>. Accessed 15<sup>th</sup> December 2006

Salvini, P.A., and E. J. Miller. 2005. ILUTE: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems. *Networks and Spatial Economics* 5: 217-234.

Schachter, J. 2004. Geographic Mobility: 2002 to 2003. Washington, D.C.: US Census Bureau. <http://www.census.gov/prod/2004pubs/p20-549.pdf>. Accessed 10<sup>th</sup> May 2006.

Sermons, M. W., and F. S. Koppelman. 2001. Representing the Differences between Female and Male Commute Behavior in Residential Location Models. *Journal of Transport Geography* 9(2): 101-110.

Spear, B. D. 2006. Freight Modeling in Urban Areas – State of the Practice. [www.trb.org/Conferences/2006/FDM/Spear.pdf](http://www.trb.org/Conferences/2006/FDM/Spear.pdf). Accessed 10<sup>th</sup> November 2007.

Timmermans, H. 2003. The saga of integrated land use-transport modeling: How Many More Dreams Before We Wake Up? Conference Keynote Paper at the 10th International Conference on Travel Behaviour Research, Lucerne, 10-15 August. <http://www.ivt.baug.ethz.ch/allgemein/pdf/timmermans.pdf>. Accessed 10<sup>th</sup> October 2006.

Train, K. 1986. *Qualitative Choice Analysis: Theory, Econometrics , and an Application to Automobile Demand*. Cambridge, Massachusetts: The MIT Press.

Van Dijk, J., and P. H. Pellenbarg. 2000a. Spatial Perspectives on Firm Demography. *Papers in Regional Science* 79: 107-110.

Van Dijk, J., and P. H. Pellenbarg. 2000b. Firm Relocation Decisions in The Netherlands: An Ordered Logit Approach. *Papers in Regional Science* 79: 191-219.

Van Wissen, L. J. G. 2000. A Micro-Simulation Model of Firms: Applications of Concepts of the Demography of the Firm. *Papers in Regional Science* 79: 111-134.

Van Wissen, L. J. G. 2002. Demography of the Firm: A Useful Metaphor? *European Journal of Population* 18: 263–279.

Waddell, P. 1993. Exogenous workplace choice in residential location models: is the assumption valid? *Geographical Analysis* 25(1): 65-82.

Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke, and G. Ulfarsson. 2003. Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim. *Networks and Spatial Economics* 3(1): 43-67.

Weisbrod, G., M. Ben-Akiva, and S. Lerman. 1980. Tradeoffs in Residential Location Decisions: Transportation versus Other Factors. *Transportation Policy and Decision-Making* 1(1): 13-26.

Woodward, D. 1992. Locational Determinants of Japanese Manufacturing Start-Ups in the United States. *Southern Economic Journal* 58: 690–708.

## APPENDIX A: MATLAB Microsimulation Code for Household and Firm Evolution

The MATLAB microsimulation code here does not include logsum calculations which were carried out every 5 years and the traffic assignment model which was estimated using TransCAD.

```
%READING DATA FROM FILES
%Input files for household simulation

    %Reading person file
    person=dlmread('individual.dat');

    %Reading hh file
    hh=xlsread('hh.xls');

    %Reading TAZ Characteristics file file
    houseval=xlsread('TAZhouserule.xls');
    tazcharhh=xlsread('tazcharHH.xls');

    %Reading life table
    deathprob=xlsread('Death Probabilities.xls');

    %Reading logsums file
    logsumhh=xlsread('logsum.xls');

    %End of reading hh input files

    %Input files for firmography simulation

    %Employment point data
    empdata = xlsread('employment data.xls');

    %Input files for birth and death rates by sector and size
    Basic = xlsread('Birth and Death exp.xls', 'Basic');
    Retail = xlsread('Birth and Death exp.xls', 'Retail');
    Education = xlsread('Birth and Death exp.xls', 'Education');
    Service = xlsread('Birth and Death exp.xls', 'Service');

    %Input files for location choice model
    tazlocprob=xlsread('TAZLocProb.xls', 'TazLocProb');
    locchoice=xlsread('location choice 24 june.xls', 'location choice 24 june');
    logsum=dlmread('logsum 21 july.dat');

    %Commercial trip movement model files
```

```

tripgen=xlsread('1996 TG and TD Models.xls', 'TG');
tdshare=xlsread('TDTAZshare26june_2.xls');

%TAZ characteristics file
tazstats=xlsread('TAZ stats.xls');

% Markov transition matrices
markov_basic=xlsread('Markov Transition Matrix.xls', 'Markov_Basic');
markov_retail=xlsread('Markov Transition Matrix.xls', 'Markov_Retail');
markov_education=xlsread('Markov Transition Matrix.xls', 'Markov_Education');
markov_service=xlsread('Markov Transition Matrix.xls', 'Markov_Service');
empgrowth=xlsread('firm data setup employment.xls', 'employment growth');

%End of reading input files for firms
%END OF READING DATA FIRM FILES

%Household simulation code

year=30;

for time=1:year %iteration loop

    [a,b]=size(person);
    [c,d]=size(hh);

    %Updating income data without addition or subtraction of new workers
    %Specifying triangular distribution

    step=1.5/c;
    x=-0.75:step:0.75;
    y=trimf(x,[-0.75 0.03 0.75]);

    %End of specifying triangular distribution

    for i=1:c
        hh(i,4)=hh(i,4)*(1+y(i));
    end

    for i=1:c
        j=i;
        count=0;
        while j<a
            if hh(i,1)==person(j,1)
                person(j,9)=hh(i,4);
                count=count+1;
            elseif count>0

```

```

        j=a;
    end
    j=j+1;
    end
end

%End of updating income data without addition or subtraction of new workers

%Specifying income of the newadults-npart and nfull

incomepart=0;
incomefull=0;
countpart=0;
countfull=0;
for i=1:a
    if person(i,4)==1
        if person(i,35)>=18 && person(i,35)<25
            if person(i,41)==1
                incomepart=incomepart+person(i,9);
                countpart=countpart+1;
            elseif person(i,42)==1
                incomefull=incomefull+person(i,9);
                countfull=countfull+1;
            end
        end
    end
meanincpart=incomepart/countpart;
meanincfull=incomefull/countfull;

%End of specifying income of new adults-npart and nfull

%Updating person file
person(:,35)=person(:,35)+1;
if person(:,35)==5
    person(:,10)=person(:,10)-1;
    person(:,11)=person(:,11)+1;
    person(:,36)=0;
    person(:,37)=1;
elseif person(:,35)==16
    person(:,11)=person(:,11)-1;
    person(:,12)=person(:,12)+1;
    person(:,37)=0;
    person(:,38)=1;
end
npschtot=sum(hh(:,15));

```

```

npdritot=sum(hh(:,16));
ndritot=sum(hh(:,17));
nnwrktot=sum(hh(:,18));
nstudtot=sum(hh(:,19));
nparttot=sum(hh(:,20));
nfulltot=sum(hh(:,21));
nadulttot=nnwrktot+nstudtot+nparttot+nfulltot;
[a,b]=size(person);
for i=1:a
    newadult=rand;
    if person(i,35)==18
        person(i,12)=person(i,12)-1;
        person(i,38)=0;
        if newadult<=(nnwrktot/nadulttot)
            person(i,13)=person(i,13)+1;
            person(i,39)=1;
        elseif newadult>(nnwrktot/nadulttot) && newadult<=((nnwrktot+nstudtot)/nadulttot)
            person(i,14)=person(i,14)+1;
            person(i,40)=1;
        elseif newadult>((nnwrktot+nstudtot)/nadulttot) &&
newadult<=((nnwrktot+nstudtot+nparttot)/nadulttot)
            person(i,15)=person(i,15)+1;
            person(i,41)=1;
        elseif newadult>((nnwrktot+nstudtot+nparttot)/nadulttot)
            person(i,16)=person(i,16)+1;
            person(i,42)=1;
        end
    end
end

for i=1:7
    if person(:,35+i)>0
        person(:,43)=i;
    end
end

for i=10:16
    if person(:,i)>0
        person(:,i+7)=1;
    end
end

%End of updating person file

%HH file updating
[c,d]=size(hh);

```

```

for i=1:c
    j=i;
    count=0;
    while j<a
        if hh(i,1)==person(j,1)
            if sum(person(j,10:12))>0
                hh(i,5)=1;
            end
            hh(i,8:14)=person(j,17:23);
            hh(i,15:21)=person(j,10:16);
            j=j+1;
            count=count+1;
        elseif count>0
            j=a;
        end
        j=j+1;
    end
end
%End of updating hh file

%DDeath module
[a,b]=size(person);
random=rand(a,1);
for i=1:a
    for j=1:100
        if person(i,35)==deathprob(j,1)
            %male=person(34)
            if person(i,34)==1
                if person(i,4)==1
                    if random(i)<=deathprob(j,2)
                        random(i)=-999;
                    end
                else
                    if random(i)<=deathprob(j,2)
                        random(i)=999;
                    end
                end
            else
                if person(i,4)==1
                    if random(i)<=deathprob(j,3)
                        random(i)=-999;
                    end
                else
                    if random(i)<=deathprob(j,3)
                        random(i)=999;
                    end
                end
            end
        end
    end
end

```

```

        end
    end
end
end
end
end

%person(36,37,38,39,40,41,42)=presch,predri,driving,nonwk,stud,part,full
for i=1:a
    x=person(i,4);
    if random(i)==999
        if i<=24 %24 is max hhsiz
            for k=1:i+x
                if person(i,1)==person(k,1)
                    %updating hhsiz
                    person(k,4)=person(k,4)-1;
                    %end of updating hhsiz

                    %updating income
                    if person(i,42)==1
                        person(k,9)=person(k,9)-(person(k,9)/(person(k,16)+(person(k,15)*0.5))*0.75);
                    end
                    if person(i,41)==1
                        person(k,9)=person(k,9)-(person(k,9)/(person(k,15)+(2*person(k,16)))*0.75);
                    end
                    %end of income update

                    %updating n/ppresch,n/predri,n/pdriv,n/ponwk,n/pstud,n/ppart,n/pfull
                    for j=36:42
                        if person(i,j)==1
                            person(k,j-26)=person(k,j-26)-1;%npresch etc.
                            if person(k,j-26)>0%ppresch etc.
                                person(k,j-19)=1;
                            else
                                person(k,j-19)=0;
                            end
                        end
                    end
                    %end of n/ppresch,n/predri,n/pdriv,n/ponwk,n/pstud,n/ppart,n/pfull update
                end
            end
        elseif i<=a-24
            for k=i-x:i+x
                if person(i,1)==person(k,1)
                    %updating hhsiz
                    person(k,4)=person(k,4)-1;
                    %end of updating hhsiz

```

```

%updating income
if person(i,42)==1
    if (person(k,16)+(person(k,15)*0.5))*0.75>0
        person(k,9)=person(k,9)-
(person(k,9)/(person(k,16)+(person(k,15)*0.5))*0.75);
    else
        person(k,9)=person(k,9);
    end
end
if person(i,41)==1
    if (person(k,15)+(2*person(k,16)))*0.75>0
        person(k,9)=person(k,9)-
(person(k,9)/(person(k,15)+(2*person(k,16)))*0.75);
    else
        person(k,9)=person(k,9);
    end
end
%end of income update

%updating n/ppresch,n/ppredri,n/pdriv,n/ponwk,n/pstud,n/ppart,n/pfull
for j=36:42
    if person(i,j)==1
        person(k,j-26)=person(k,j-26)-1;%npresch etc.
        if person(k,j-26)>0%ppresch etc.
            person(k,j-19)=1;
        else
            person(k,j-19)=0;
        end
    end
    %end of n/ppresch,n/ppredri,n/pdriv,n/ponwk,n/pstud,n/ppart,n/pfull update
end
elseif i>a-24
    for k=i-x:a
        if person(i,1)==person(k,1)
            %updating hhsiz
            person(k,4)=person(k,4)-1;
            %end of updating hhsiz

            %updating income
            if person(i,42)==1
                if (person(k,16)+(person(k,15)*0.5))*0.75>0
                    person(k,9)=person(k,9)-
(person(k,9)/(person(k,16)+(person(k,15)*0.5))*0.75);

```

```

        else
            person(k,9)=person(k,9);
        end
    end
    if person(i,41)==1
        if (person(k,15)+(2*person(k,16)))*0.75>0
            person(k,9)=person(k,9)-
(person(k,9)/(person(k,15)+(2*person(k,16)))*0.75);
        else
            person(k,9)=person(k,9);
        end
    end
%end of income update

%updating n/ppresch,n/ppredri,n/pdriv,n/ponwk,n/pstud,n/ppart,n/pfull
for j=36:42
    if person(i,j)==1
        person(k,j-26)=person(k,j-26)-1;%npresch etc.
        if person(k,j-26)>0%ppresch etc.
            person(k,j-19)=1;
        else
            person(k,j-19)=0;
        end
    end
    %end of n/ppresch,n/ppredri,n/pdriv,n/ponwk,n/pstud,n/ppart,n/pfull update
end
end
end
end

%Deleting dying people
i=1;
count=0;
while i<=a
    if random(i)==-999 || random(i)==999
        person(i-count,:)=[];
        count=count+1;
    else
        person(i-count,:)=person(i-count,:);
    end
    i=i+1;
end
%End of deleting dying people
%End of death module

```

```

%Household file updating at the end of death module
[a,b]=size(person);
[c,d]=size(hh);
hhdeathflag=zeros(c,1);
for i=1:c
j=1;
count=1;
while j<a
if hh(i,1)==person(j,1)
hh(i,3)=person(j,4);%hhsiz
hh(i,4)=person(j,9);%income
if person(j,17)==1 || person(j,18)==1 || person(j,19)==1 %preskids
hh(i,5)=1;
else
hh(i,5)=0;
end
for k=17:23 %ppresch etc.
if person(j,k)==1
hh(i,k-9)=1;
else
hh(i,k-9)=0;
end
end
for k=10:16
hh(i,k+5)=person(j,k); %onpresch etc.
end
j=a;
else
j=j+1;
count=count+1;
end
end
if count==a;
hhdeathflag(i)=999;
end
end

%Deleting dying households
i=1;
count=0;
while i<=c
if hhdeathflag(i)==999
hh(i-count,:)=[];
count=count+1;
else

```

```

hh(i-count,:)=hh(i-count,:);
end
i=i+1;
end
%End of deleting dying households
%End of household file updating after death module

%Birth module
%Individual file updating
person_birth=person;
[a,b]=size(person_birth);
birth=zeros(a,8);
i=1;
while i<a
if person_birth(i,34)==1
i=i+1;
else
if person_birth(i,35)>=18
birth(i,1)=person_birth(i,1);%hhid
birth(i,2)=person_birth(i,35);%age of mother
birth(i,3)=sum(person_birth(i,10:12));%no. of children
if i<25%age of youngest child
x=100;
hhsiz=person(:,4);
for j=1:i+hhsiz(i)
if birth(i,1)==person_birth(j,1)
if birth(i,3)>0
birth(i,4)=min(x,person_birth(j,35));
else
birth(i,4)=0;
end
x=birth(i,4);
end
end
elseif i>=25 && i<a-24
x=100;
for j=i-hhsiz(i):i+hhsiz(i)
if birth(i,1)==person_birth(j,1)
if birth(i,3)>0
birth(i,4)=min(x,person_birth(j,35));
else
birth(i,4)=0;
end
x=birth(i,4);
end
end
end

```

```

elseif i>=a-24
x=100;
for j=i-hhszie(i):a
    if birth(i,1)==person_birth(j,1)
        if birth(i,3)>0
            birth(i,4)=min(x,person_birth(j,35));
        else
            birth(i,4)=0;
        end
        x=birth(i,4);
    end
end
if person_birth(i,22)==1 || person_birth(i,23)==1%employment status
    birth(i,5)=1;
else
    birth(i,5)=0;
end

%Birth model
birth(i,6)=exp(0.591-0.1098*birth(i,2)+1.893*birth(i,3)-
1.618*birth(i,4)+0.666*birth(i,5));
birth(i,7)=birth(i,6)/(1+birth(i,6));
%End of birth model

end
i=i+1;
end
end
%End of individual file updating

%Births and characteristics attributing
for i=1:a
    if birth(i,1)>0
        if birth(i,7)>rand
            birth(i,8)=999;
        end
    end
    i=i+1;
end
count=1;
while i<a
    if birth(i,8)==999
        person_birth(a+count,1)=birth(i,1);
        person_birth(a+count,2:3)=person_birth(i,2:3);
        person_birth(a+count,4)=person_birth(i,4)+1;
    end
end

```

```

person_birth(a+count,5:9)=person_birth(i,5:9);
person_birth(a+count,10)=person_birth(i,10)+1;
person_birth(a+count,11:16)=person_birth(i,11:16);
person_birth(a+count,17)=1;
person_birth(a+count,18:29)=person_birth(i,18:29);
person_birth(a+count,30)=person_birth(a+count,4);
person_birth(a+count,31)=person_birth(a+count,1)*100+person_birth(a+count,30);
person_birth(a+count,32)=person_birth(i,32);
person_birth(a+count,33)=person_birth(a+count,32)*100+person_birth(a+count,30);
male=rand;
if male<=0.5
    person_birth(a+count,34)=1;
else
    person_birth(a+count,34)=0;
end
person_birth(a+count,35)=1;
person_birth(a+count,36:42)=person_birth(i,36:42);
person_birth(a+count,43)=1;
count=count+1;
end
i=i+1;
end
%End of births and characteristics attributing

%Adding new borns to the original person file
[y,z]=size(person);
[u,v]=size(person_birth);
born=u-a;
for i=1:born
    person(y+i,:)=person_birth(a+i,:);
end
%End of adding new borns to the original person file

%HH file updating
[a,b]=size(person);
[c,d]=size(hh);
for i=(a-born):a
    for j=1:c
        if person(i,1)==hh(j,1)
            hh(j,3)=hh(j,3)+1;
            hh(j,5)=1;
            hh(j,8)=1;
            hh(j,15)=hh(j,15)+1;
        end
    end
end

```

```

%End of HH file updating

%Migration module
approach=0;
%approach=1;
if approach==0
%Identifying persons approach
[a,b]=size(person);
migrate=(2.127/100);
migr=rand(a,1);
for i=1:1074
    houseval(i,3)=rand;
end
count=1;
for i=1:a
    if migr(i)<=migrate
        person(a+count,:)=person(i,:);
        count=count+1;
    end
end
%End of identifying persons approach
else
%Identifying Households approach
[c,d]=size(hh);
migrate=(2.127/100);
migr=rand(c,1);
for i=1:1074
    houseval(i,3)=rand;
end
count=1;
for i=1:c
    if migr(i)<=migrate
        hh(c+count,:)=hh(i,:);
        count=count+1;
    end
end
%End of identifying households approach
end

if approach==0
%If persons approach, THEN
%Household size and income distribution calculation
hhsizel=0;
inc1=0;
hhsizel2=0;
inc2=0;

```

```

hhsiz3=0;
inc3=0;
hhsiz4=0;
inc4=0;
hhsiz5=0;
inc5=0;
hhsiz6=0;
inc6=0;
hhsiz7=0;
inc7=0;
hhsiz8=0;
inc8=0;
for i=1:c
    if hh(i,3)==1
        hhsiz1=hhsiz1+1;
        inc1=inc1+hh(i,4);
        %avinc1=inc1/hhsiz1;
    elseif hh(i,3)==2
        hhsiz2=hhsiz2+1;
        inc2=inc2+hh(i,4);
        %avinc2=inc2/hhsiz2;
    elseif hh(i,3)==3
        hhsiz3=hhsiz3+1;
        inc3=inc3+hh(i,4);
        %avinc3=inc3/hhsiz3;
    elseif hh(i,3)==4
        hhsiz4=hhsiz4+1;
        inc4=inc4+hh(i,4);
        %avinc4=inc4/hhsiz4;
    elseif hh(i,3)==5
        hhsiz5=hhsiz5+1;
        inc5=inc5+hh(i,4);
        %avinc5=inc5/hhsiz5;
    elseif hh(i,3)==6
        hhsiz6=hhsiz6+1;
        inc6=inc6+hh(i,4);
        %avinc6=inc6/hhsiz6;
    elseif hh(i,3)==7
        hhsiz7=hhsiz7+1;
        inc7=inc7+hh(i,4);
        %avinc7=inc7/hhsiz7;
    elseif hh(i,3)>=8
        hhisze8=hhsiz8+1;
        inc8=inc8+hh(i,4);
        %avinc8=inc8/hhsiz8;
    end

```

```

end
avinc1=inc1/hhsiz1;
avinc2=inc2/hhsiz2;
avinc3=inc3/hhsiz3;
avinc4=inc4/hhsiz4;
avinc5=inc5/hhsiz5;
avinc6=inc6/hhsiz6;
avinc7=inc7/hhsiz7;
if hhsiz8==0
    avinc8=avinc7;
else
    avinc8=inc8/hhsiz8;
end
%End of household size and income distribution calculation

%Determining hhsiz and making households from migrating persons
c1=0;
for i=a-count:a
    x=rand;
    if person(i,35)>=18
        %Single membered migrating families updated except for location choice and auto
ownership
        if x<=hhsiz1/c
            c1=c1+1;
            person(i,4)=1;%household size of single person household
            person(i,9)=avinc1;%Income of single person household
            for j=1:7
                if person(i,35+j)==1
                    person(i,9+j)=1;
                    person(i,16+j)=1;
                else
                    person(i,9+j)=0;
                    person(i,16+j)=0;
                end
            end
        end
        %Single membered migrating families updated except for location choice and auto
ownership
        elseif x>hhsiz1/c && x<=(hhsiz1+hhsiz2)/c
            person(i,4)=2;
        elseif x>(hhsiz1+hhsiz2)/c && x<=(hhsiz1+hhsiz2+hhsiz3)/c
            person(i,4)=3;
        elseif x>(hhsiz1+hhsiz2+hhsiz3)/c && x<=(hhsiz1+hhsiz2+hhsiz3+hhsiz4)/c
            person(i,4)=4;
        elseif x>(hhsiz1+hhsiz2+hhsiz3+hhsiz4)/c &&
x<=(hhsiz1+hhsiz2+hhsiz3+hhsiz4+hhsiz5)/c
            person(i,4)=5;
    end
end

```

```

        elseif           x>(hysize1+hysize2+hysize3+hysize4+hysize5)/c      &&
x<=(hysize1+hysize2+hysize3+hysize4+hysize5+hysize6)/c
            person(i,4)=6;
        elseif           x>(hysize1+hysize2+hysize3+hysize4+hysize5+hysize6)/c      &&
x<=(hysize1+hysize2+hysize3+hysize4+hysize5+hysize6+hysize7)/c
            person(i,4)=7;
        else
            person(i,4)=8;
        end
    end
end

cnt2=1;
cnt3=1;
cnt4=1;
cnt5=1;
cnt6=1;
cnt7=1;
cnt8=1;
for i=a-count:a
    if person(i,4)==2
        per2(cnt2,1)=person(i,1);
        per2(cnt2,2)=person(i,35);
        cnt2=cnt2+1;
    elseif person(i,4)==3
        per3(cnt3,1)=person(i,1);
        per3(cnt3,2)=person(i,35);
        cnt3=cnt3+1;
    elseif person(i,4)==4
        per4(cnt4,1)=person(i,1);
        per4(cnt4,2)=person(i,35);
        cnt4=cnt4+1;
    elseif person(i,4)==5
        per5(cnt5,1)=person(i,1);
        per5(cnt5,2)=person(i,35);
        cnt5=cnt5+1;
    elseif person(i,4)==6
        per6(cnt6,1)=person(i,1);
        per6(cnt6,2)=person(i,35);
        cnt6=cnt6+1;
    elseif person(i,4)==7
        per7(cnt7,1)=person(i,1);
        per7(cnt7,2)=person(i,35);
        cnt7=cnt7+1;
    elseif person(i,4)==8
        per8(cnt8,1)=person(i,1);

```

```

per8(cnt8,2)=person(i,35);
cnt8=cnt8+1;
end
end

```

```

sortrows(per2,-2);
sortrows(per3,-2);
sortrows(per4,-2);
sortrows(per5,-2);
sortrows(per6,-2);
sortrows(per7,-2);
sortrows(per8,-2);

```

```

[a2,b2]=size(per2);
[a3,b3]=size(per3);
[a4,b4]=size(per4);
[a5,b5]=size(per5);
[a6,b6]=size(per6);
[a7,b7]=size(per7);
[a8,b8]=size(per8);

```

```

c2=floor(a2/2);
remove=a2-2*c2;
for i=1:remove
    per2(i,:)=[];
end
for i=1:c2
    per2(c2+i,1)=per2(i,1);
end

```

```

c3=floor(a3/3);
remove=a3-3*c3;
for i=1:remove
    per3(i,:)=[];
end
for i=1:c3
    per3(c3+i,1)=per3(i,1);
    per3(2*c3+i,1)=per3(i,1);
end

```

```

c4=floor(a4/4);
remove=a4-4*c4;
for i=1:remove
    per4(i,:)=[];
end
for i=1:c4

```

```

per4(c4+i,1)=per4(i,1);
per4(2*c4+i,1)=per4(i,1);
per4(3*c4+i,1)=per4(i,1);
end

c5=floor(a5/5);
remove=a5-5*c5;
for i=1:remove
    per5(i,:)=[];
end
for i=1:c5
    per5(c5+i,1)=per5(i,1);
    per5(2*c5+i,1)=per5(i,1);
    per5(3*c5+i,1)=per5(i,1);
    per5(4*c5+i,1)=per5(i,1);
end

c6=floor(a6/6);
remove=a6-6*c6;
for i=1:remove
    per6(i,:)=[];
end
for i=1:c6
    per6(c6+i,1)=per6(i,1);
    per6(2*c6+i,1)=per6(i,1);
    per6(3*c6+i,1)=per6(i,1);
    per6(4*c6+i,1)=per6(i,1);
    per6(5*c6+i,1)=per6(i,1);
end

c7=floor(a7/7);
remove=a7-7*c7;
for i=1:remove
    per7(i,:)=[];
end
for i=1:c7
    per7(c7+i,1)=per7(i,1);
    per7(2*c7+i,1)=per7(i,1);
    per7(3*c7+i,1)=per7(i,1);
    per7(4*c7+i,1)=per7(i,1);
    per7(5*c7+i,1)=per7(i,1);
    per7(6*c7+i,1)=per7(i,1);
end

c8=floor(a8/8);
remove=a8-8*c8;

```

```

for i=1:remove
    per8(i,:)=[];
end
for i=1:c8
    per8(c8+i,1)=per8(i,1);
    per8(2*c8+i,1)=per8(i,1);
    per8(3*c8+i,1)=per8(i,1);
    per8(4*c8+i,1)=per8(i,1);
    per8(5*c8+i,1)=per8(i,1);
    per8(6*c8+i,1)=per8(i,1);
    per8(7*c8+i,1)=per8(i,1);
end

[a2,b2]=size(per2);
[a3,b3]=size(per3);
[a4,b4]=size(per4);
[a5,b5]=size(per5);
[a6,b6]=size(per6);
[a7,b7]=size(per7);
[a8,b8]=size(per8);

for i=1:a2
    for j=1:a
        if per2(i,1)==person(j,1)
            person(a+i,:)=person(j,:);
            person(a+i,4)=2;
            person(a+i,9)=avinc2;
        end
    end
end

for i=1:a3
    for j=1:a
        if per3(i,1)==person(j,1)
            person(a+a2+i,:)=person(j,:);
            person(a+a2+i,4)=3;
            person(a+a2+i,9)=avinc3;
        end
    end
end

for i=1:a4
    for j=1:a
        if per4(i,1)==person(j,1)
            person(a+a2+a3+i,:)=person(j,:);
            person(a+a2+a3+i,4)=4;
        end
    end
end

```

```

        person(a+a2+a3+i,9)=avinc4;
    end
end
end

for i=1:a5
    for j=1:a
        if per5(i,1)==person(j,1)
            person(a+a2+a3+a4+i,:)=person(j,:);
            person(a+a2+a3+a4+i,4)=5;
            person(a+a2+a3+a4+i,9)=avinc5;
        end
    end
end

for i=1:a6
    for j=1:a
        if per6(i,1)==person(j,1)
            person(a+a2+a3+a4+a5+i,:)=person(j,:);
            person(a+a2+a3+a4+a5+i,4)=6;
            person(a+a2+a3+a4+a5+i,9)=avinc6;
        end
    end
end

for i=1:a7
    for j=1:a
        if per7(i,1)==person(j,1)
            person(a+a2+a3+a4+a5+a6+i,:)=person(j,:);
            person(a+a2+a3+a4+a5+a6+i,4)=7;
            person(a+a2+a3+a4+a5+a6+i,9)=avinc7;
        end
    end
end

for i=1:a8
    for j=1:a
        if per8(i,1)==person(j,1)
            person(a+a2+a3+a4+a5+a6+a7+i,:)=person(j,:);
            person(a+a2+a3+a4+a5+a6+a7+i,4)=8;
            person(a+a2+a3+a4+a5+a6+a7+i,9)=avinc8;
        end
    end
end

%newper=c1+a2+a3+a4+a5+a6+a7+a8;
[tempa,tempb]=size(person);

```

```

newper=tempa-a;
%End of determining household sizes and making families

%Updating household file
%Identifying new households by HHno.
newfam=c1+c2+c3+c4+c5+c6+c7+c8;
countfam=0;
for i=a+1:a+newper
    j=i+1;
    while j<a+newper
        if person(i,1)==person(j,1)
            countfam=countfam+1;
            hh(c+countfam,1)=person(i,1);
            hh(c+countfam,3)=person(i,4);
            hh(c+countfam,4)=person(i,9);
            j=a+newper;
        else
            j=j+1;
        end
    end
end
%End of identifying new households by HHno.

%Household file updated
%Sorting household and person file for the migrating entries
hh_temp=zeros(countfam,39);
for i=c+1:c+countfam
    hh_temp(i-c,:)=hh(i,:);
end
person_temp=zeros(newper,b);
for i=a+1:a+newper
    person_temp(i-a,:)=person(i,:);
end
sortrows(hh_temp,1);
sortrows(person_temp,1);
for i=c+1:c+countfam
    hh(i,:)=hh_temp(i-c,:);
end
for i=a+1:a+newper
    person(i,:)=person_temp(i-a,:);
end
%End of sorting migrating households and persons

for i=c+1:c+countfam
    j=a+1;
    tempcount=0;

```

```

while j<a+newper
  if hh(i,1)==person(j,1)
    tempcount=tempcount+1;
    hh(i,40)=0;
    if person(j,4)==1%single size household
      hh(i,2)=0;% did not assign taz till now
      hh(i,3)=1;
      hh(i,4)=person(j,9);
      if sum(person(j,36:38))>0
        hh(i,5)=1;
      else
        hh(i,5)=0;
      end
      hh(i,6:7)=0;% did not assign vehicles till now
      hh(i,8:14)=person(i,17:23);
      hh(i,15:21)=person(i,10:16);
      hh(i,22:39)=0;%did not assign these characteristics either
      if person(i,35)>=65
        hh(i,40)=1;
      end
    end
  else%more than one person in household
    hh(i,2)=0;
    %hh(i,3)=no. of persons in the household, already allotted above for all households
    hhsiz=0;
    k=j;
    while hh(i,1)==person(k,1)
      hhsiz=hhsiz+1;
      k=k+1;
    end
    hh(i,8:14)=0;
    npsch=0;
    npdri=0;
    ndri=0;
    nnwrk=0;
    nstud=0;
    npart=0;
    nfull=0;
    nret=0;
    for l=1:hhsiz
      if person(k+l-1,36)==1
        hh(i,8)=1;
        npsch=npsch+1;
      elseif person(k+l-1,37)==1
        hh(i,9)=1;
        npdri=npdri+1;
      elseif person(k+l-1,38)==1

```

```

hh(i,10)=1;
ndri=ndri+1;
elseif person(k+l-1,39)==1
    hh(i,11)=1;
    nnwrk=nnwrk+1;
    if person(k+l-1,35)>=65
        nret=nret+1;
    end
elseif person(k+l-1,40)==1
    hh(i,12)=1;
    nstud=nstud+1;
    if person(k+l-1,35)>=65
        nret=nret+1;
    end
elseif person(k+l-1,41)==1
    hh(i,13)=1;
    npart=npart+1;
    if person(k+l-1,35)>=65
        nret=nret+1;
    end
elseif person(k+l-1,42)==1
    hh(i,14)=1;
    nfull=nfull+1;
    if person(k+l-1,35)>=65
        nret=nret+1;
    end
end
hh(i,15)=npsch;
hh(i,16)=npdri;
hh(i,17)=ndri;
hh(i,18)=nnwrk;
hh(i,19)=nstud;
hh(i,20)=npart;
hh(i,21)=nfull;
hh(i,40)=nret;
end
elseif tempcount>0
    j=a+newper;
end
j=j+1;
end
end
%End of updating household file

```

%Changing the index of new persons and households in person and hh file

```

for i=c+1:c+countfam
    hh(i,1)=max(hh(1:i-1,1))+1;
end
for j=a+1:a+newper
    person(i,1)=max(hh(1:i-1,1))+1;
end
%End of changing the index of new persons and households in person and hh file

sortrows(person,[1,30]);
sortrows(hh,1);

%Updating sen and hhperno
person(:,30)=1;
for i=1:a+newper
    j=i+1;
    count=1;
    while j<a+newper
        if person(i,1)==person(j,1)
            count=count+1;
            person(j,30)=count;
            j=j+1;
        else
            j=a+newper;
        end
    end
    person(:,31)=person(i,1)*100+person(:,30);
%End of updating sen and hhperno
%If persons approach, STOP; ELSE: if households approach
else
%Updating individuals file
    newper=0;
    for i=1:count
        hhmatch=0;
        for j:1:a
            if hh(c+i,1)==person(j,1)
                person(a+i+newper+hhmatch,:)=person(j,:);
                match=match+1;
            end
        end
        newper=newper+hhmatch;
    end
%End of updating individuals file

%Updating unique identifiers for households and individuals
    for i=c+1:c+count

```

```

hh(i,1)=max(hh(1:i-1,1))+1;
end
for j=a+1:a+newper
    person(i,1)=max(hh(1:i-1,1))+1;
end
%End of updating unique identifiers for households and individuals

%Updating sen and hhperno
person(:,30)=1;
for i=1:a+newper
    j=i+1;
    countn=1;
    while j<a+newper
        if person(i,1)==person(j,1)
            countn=countn+1;
            person(j,30)=countn;
            j=j+1;
        else
            j=a+newper;
        end
    end
    person(:,31)=person(i,1)*100+person(:,30);
%End of updating sen and hhperno
end

%APPLY LOCATION CHOICE MODEL
for j=1:1074
    houseval(j,20)=exp(0.476*houseval(j,19)+0.351*houseval(j,18)-
0.071*houseval(j,2)+0.110/10000*houseval(j,14)-
0.311*houseval(j,15)/houseval(j,5)+0.114*houseval(j,10)/1000+0.0186*houseval(j,11)/1000)-
2.73/10000*logsumhh(j,2);
end
for j=1:1074
    houseval(j,21)=houseval(j,20)/sum(houseval(:,20));
end
for j=1:1074
    houseval(j,22)=sum(houseval(1:j,21));
end
for i=c+1:c+countfam
    x=rand;
    if x<=houseval(1,22)
        hh(i,2)=1;
    else
        for j=2:1074
            if x>houseval(j-1,22) && x<=houseval(j,22)

```

```

        hh(i,2)=houseval(j,1);
    end
end
end
for i=c+1:c+countfam
    for j=1:1074
        if hh(i,2)==tazcharhh(j,1)
            hh(i,22:39)=tazcharhh(j,2:19);
        end
    end
end
%END of LOCATION CHOICE MODEL

```

```

%Adjust all the characteristics
for i=c+1:c+countfam
    for j=a+1:a+newper
        if hh(i,1)==person(j,1)
            person(j,2)=hh(i,2);
            person(j,4)=hh(i,3);
            person(j,6)=hh(i,6);
            if person(j,6)>person(j,4)
                person(j,7)=person(j,6)-person(j,4);
            else
                person(j,7)=0;
            end
            if person(j,6)<person(j,4)
                person(j,8)=person(j,6)-person(j,4);
            else
                person(j,8)=0;
            end
        end
    end
end
%End of adjusting all the characteristics

```

```

[a,b]=size(person);
[c,d]=size(hh);
hhdummy=zeros(c,3);

for i=1:c
    if hh(i,3)>1
        if sum(hh(i,18:21))>1
            hhdummy(i,1)=1;
        end
    end

```

```

end

person(:,b+1)=0;
hh(:,d+1)=0;

for i=1:c
    if hhdummy(i,1)==1
        j=i;
        while j<a
            if hh(i,1)<person(j,1)
                j=j+1;
            elseif hh(i,1)==person(j,1)
                if person(j,35)>=18 && person(j+1,35)>=18
                    person(j:j+1,b+1)=1;
                    hh(i,d+1)=1;
                    hhdummy(i,2)=person(j,35);
                end
                j=a;
            elseif hh(i,1)>person(j,1)
                j=a;
            end
        end
    end
end

```

```

%DIVORCE MODULE
count=0;
for i=1:c
    if hh(i,d+1)==1
        util=exp(-1.437-0.281*hh(i,3)-0.028*hhdummy(i,2));
        hhdummy(i,3)=util/(1+util);
        if rand<=hhdummy(i,3)
            hhdummy(i,1)=999;
        end
        if hhdummy(i,1)==999
            count=count+1;
            hh(c+count,1)=hh(c+count-1,1)+1;
            hh(c+count,3)=1;
            hh(i,3)=hh(i,3)-1;
            hh(i,4)=hh(i,4)/2;
            hh(c+count,4)=hh(i,4);
            hh(c+count,5)=0;
            if hh(i,3)>2
                j=i;
                while j<a

```

```

if hh(i,1)==person(j,1)
    for k=j+2:j+hh(i,3)
        if rand<=0.5
            person(k,1)=hh(c+count,1);
            hh(c+count,3)=hh(c+count,3)+1;
            hh(i,3)=hh(i,3)-1;
        end
    end
    j=a;
else
    j=j+1;
end
end
end
end
end
end
%End of divorce module

```

hh(:,d+1)=[];

person(:,b+1)=[];

% APPLY LOCATION CHOICE MODEL here for divorced households

```

for i=1:c+count
    x=rand;
    if x<=houseval(1,22)
        hh(i,2)=1;
    else
        for j=2:1074
            if x>houseval(j-1,22) && x<=houseval(j,22)
                hh(i,2)=houseval(j,1);
            end
        end
    end
end
for i=1:c+count
    for j=1:1074
        if hh(i,2)==tazcharhh(j,1)
            hh(i,22:39)=tazcharhh(j,2:19);
        end
    end
end

```

% END OF LOCATION CHOICE MODEL for divorced households

%Updating person and household files after divorce

for i=1:c+count

```

if hhdummy(i,1)==999 || i>c
    hh(i,5)=0;
    hh(i,8:21)=0;
    count1=0;
    for j=1:a
        if hh(i,1)==person(j,1)
            person(j,2)=hh(i,2);
            person(j,24:27)=hh(i,25:28);
            count1=count1+1;
            if person(j,43)==1 || person(j,43)==2 || person(j,43)==3
                hh(i,5)=1;
            end
            for m=1:7
                if person(j,43)==m
                    hh(j,7+m)=1;
                    hh(j,14+m)=hh(j,14+m)+1;
                end
            end
            if hh(i,6)==1
                if rand<=0.5
                    hh(i,6)=1;
                    hh(c+count1,6)=0;
                else
                    hh(i,6)=0;
                    hh(c+count1,6)=1;
                end
            elseif hh(i,6)==2 || hh(i,6)==4 || hh(i,6)==6 || hh(i,6)==8
                hh(i,6)=hh(i,6)/2;
                hh(c+count1,6)=hh(i,6);
            elseif hh(i,6)==3 || hh(i,6)==5 || hh(i,6)==7 || hh(i,6)==9
                if rand<=0.5
                    hh(i,6)=round(hh(i,6)/2);
                    hh(c+count1,6)=floor(hh(i,6)/2);
                else
                    hh(c+count1,6)=round(hh(i,6)/2);
                    hh(i,6)=floor(hh(i,6)/2);
                end
            elseif hh(i,6)==0
                hh(i,6)=0;
                hh(c+count1,6)=0;
                hh(i,7)=1;
                hh(c+count1,7)=1;
            end
            person(j,4)=hh(i,3);
            person(j,6)=hh(i,6);
            if person(j,6)>person(j,4)

```

```

        person(j,7)=person(j,6)-person(j,4);
    else
        person(j,7)=0;
    end
    if person(j,6)<person(j,4)
        person(j,8)=person(j,6)-person(j,4);
    else
        person(j,8)=0;
    end
    person(j,9)=hh(i,4);
    person(j,10:16)=hh(i,15:21);
    person(j,17:23)=hh(i,8:14);
end
end
end
%End of updating person and household files after divorce

[a,b]=size(person);
[c,d]=size(hh);

sortrows(hh,1);
sortrows(person,[1,-35]);

%Updating sen and hhperno
i=1;
while i<a
    j=i+1;
    count=1;
    while j<a
        if person(i,1)==person(j,1)
            count=count+1;
            person(i,30)=1;
            person(j,30)=count;
            j=j+1;
        else
            j=a;
        end
    end
    i=i+count;
end
person(:,31)=person(:,1)*100+person(:,30);
person(:,33)=person(:,32)*100+person(:,30);
%End of updating sen and hhperno

%Marraige module

```

```

perdummy=zeros(a,6);
for i=1:a
    if person(i,4)==1 && person(i,43)>=4
        if person(i,34)==1
            perdummy(i,1)=1;
        else
            perdummy(i,2)=1;
        end
        perdummy(i,3)=person(i,35);
    elseif person(i,4)>1 && person(i,43)>=4
        if i<=24
            for j=1:i+person(i,4)
                if person(i,1)==person(j,1) && i~=j
                    if person(i,35)>=18 && abs(person(i,34)-person(j,34))==1
                        flag=-999;
                    end
                end
            end
        elseif i>24 && i<a-24
            for j=i-person(i,4):i+person(i,4)
                if person(i,1)==person(j,1) && i~=j
                    if person(i,35)>=18 && abs(person(i,34)-person(j,34))==1
                        flag=-999;
                    end
                end
            end
        elseif i>=a-24
            for j=i-person(i,4):a
                if person(i,1)==person(j,1) && i~=j
                    if person(i,35)>=18 && abs(person(i,34)-person(j,34))==1
                        flag=-999;
                    end
                end
            end
        end
    end
    if flag~-=-999
        if person(i,34)==1
            perdummy(i,1)=2;
        else
            perdummy(i,2)=2;
        end
    end
    perdummy(i,3)=person(i,35);
end
end
for i=1:a

```

```

if sum(perdummy(i,1:2))>0
    util=exp(-1.869+0.129*person(i,4)-0.032*perdummy(i,3)+0.025*sum(person(i,41:42)));
    perdummy(i,4)=util/(1+util);
    if rand<=perdummy(i,4)
        perdummy(i,4)=999;
    end
end
end

%Identifying couples for marraige
count=0;
for i=1:a
    if perdummy(i,5)==0
        if perdummy(i,4)==999
            if perdummy(i,1)>0
                j=1;
                while j<a
                    if perdummy(i,4)==0
                        if perdummy(j,4)==999
                            if perdummy(j,2)>0
                                if abs(perdummy(j,3)-perdummy(i,3))<=3%age difference 3 years
                                    count=count+1;
                                    perdummy(i,5)==count;
                                    perdummy(j,5)==count;
                                    perdummy(i,6)=person(i,1);
                                    perdummy(j,6)=person(j,1);
                                    j=a;
                                end
                            end
                        end
                    end
                end
                j=j+1;
            end
            if count==0
                j=1;
                while j<a
                    if perdummy(i,4)==0
                        if perdummy(j,4)==999
                            if perdummy(j,2)>0
                                if abs(perdummy(j,3)-perdummy(i,3))<=5%age difference 5 years
                                    count=count+1;
                                    perdummy(i,5)==count;
                                    perdummy(j,5)==count;
                                    perdummy(i,6)=person(i,1);
                                    perdummy(j,6)=person(j,1);
                                    j=a;
                                end
                            end
                        end
                    end
                end
            end
        end
    end
end

```

```

%Making new households out of marriage
%Updating household file
for i=1:a
    j=1;
    while j<a
        if perdummy(i,6)==perdummy(j,6)
            k=1;
            while k<c
                if person(i,1)==hh(k,1)
                    m=k;
                    k=c;
                else

```

```

        k=k+1;
    end
end
l=1;
while l<c
    if person(i,1)==hh(l,1)
        m=l;
        l=c;
    else
        l=l+1;
    end
end
hh(k,3:4)=hh(k,3:4)+hh(l,3:4);
if hh(k,5)==0 && hh(l,5)==0
    hh(k,5)=0;
else
    hh(k,5)=1;
end
hh(k,6)=hh(k,6)+hh(l,6);
if hh(k,7)==1 && hh(l,7)==1;
    hh(k,7)=1;
else
    hh(k,7)=0;
end
if sum(hh(k,8:14))==0 && sum(hh(l,8:14))==0
    hh(k,8:14)=0;
else
    hh(k,8:14)=1;
end
hh(k,15:21)=hh(k,15:21)+hh(l,15:21);
hh(l,1)=-999;
j=a;
else
    j=j+1;
end
end
end
i=1;
while i<c
    if hh(i,1)==-999
        hh(i,:)=[];
    else
        i=i+1;
    end
end

```

```
%End of updating household file
```

```
%Updating person file
```

```
for i=1:a
    tempcount=0;
    j=1;
    while j<c
        if perdummy(i,6)==hh(j,1)
            tempcount=tempcount+1;
            person(i,2)=hh(j,2);
            person(i,4)=hh(j,3);
            person(i,6)=hh(j,6);
            if person(i,6)>person(i,4)
                person(i,7)=person(i,6)-person(i,4);
            else
                person(i,7)=0;
            end
            if person(i,6)<person(i,4)
                person(i,8)=person(i,6)-person(i,4);
            else
                person(i,8)=0;
            end
            person(i,9)=hh(j,4);
            person(i,10:16)=hh(j,15:21);
            person(i,17:23)=hh(j,8:14);
            person(i,24:27)=hh(j,25:28);
        elseif tempcount>0
            j=c;
        end
        j=j+1;
    end
end
```

```
%End of updating person file
```

```
[a,b]=size(person);
[c,d]=size(hh);
```

```
sortrows(hh,1);
sortrows(person,[1,-35]);
```

```
%Updating sen and hhperno
```

```
i=1;
while i<a
    j=i+1;
    count=1;
    while j<a
```

```

if person(i,1)==person(j,1)
    count=count+1;
    person(i,30)=1;
    person(j,30)=count;
    j=j+1;
else
    j=a;
end
end
i=i+count;
end
person(:,31)=person(:,1)*100+person(:,30);
person(:,33)=person(:,32)*100+person(:,30);
%End of updating sen and hhperno
%End of making new households from marraige

%Moving out module
%Working with person file
count=0;
for i=1:a
    if i<=24
        if person(i,4)>1 && sum(person(i,13:16))>2
            if person(i,35)==18
                count=count+1;
                x=person(i,4);
                for j=1:i+x;
                    if person(i,1)==person(j,1)
                        person(j,4)=person(j,4)-1;
                        autoratio=person(j,6)/(sum(person(j,12:16)));
                        if autoratio>0.5
                            person(j,6)=person(j,6)-1;
                            person(i,6)=1;
                        else
                            person(j,6)=person(j,6);
                            person(i,6)=0;
                        end
                    end
                end
            end
            person(i,44)=person(i,1);
            person(i,1)=person(a,1)+count;
            person(i,7)=0;
            if person(i,6)==1
                person(i,8)=0;
            else
                person(i,8)=-1;
            end
        end
    end
end

```

```

person(i,9)=avinc1;
person(i,10:12)=0;
person(i,13:16)=0;
y=rand;
if y<=nnwrktot/nadulttot
    person(i,13)=1;
elseif y>nnwrktot/nadulttot && y<=(nnwrktot+nstudtot)/nadulttot
    person(i,14)=1;
elseif y>(nnwrktot+nstudtot)/nadulttot
    y<=(nnwrktot+nstudtot+nparttot)/nadulttot
    person(i,15)=1;
elseif y>(nnwrktot+nstudtot+nparttot)/nadulttot
    person(i,16)=1;
end
person(i,17:19)=0;
for k=20:23
    if person(i,k-7)==1
        person(i,k)=1;
    else
        person(i,k)=0;
    end
end
person(i,30)=1;
person(i,31)=person(i,1)*100+person(i,30);
person(i,33)=person(i,32)*100+1;
person(i,36:42)=person(i,17:23);
for k=1:7
    if person(i,35+k)==1
        person(i,43)=k;
    end
end
end
endif i>24 && i<a-24
if person(i,4)>1 && sum(person(i,13:16))>2
    if person(i,35)==18
        count=count+1;
        x=person(i,4);
        for j=i-x:i+x;
            if person(i,1)==person(j,1)
                person(j,4)=person(j,4)-1;
                if sum(person(j,12:16))==0
                    autoratio=0;
                else
                    autoratio=person(j,6)/(sum(person(j,12:16)));
                end
            end
        end
    end

```

```

if autoratio>0.5
    person(j,6)=person(j,6)-1;
    person(i,6)=1;
else
    person(j,6)=person(j,6);
    person(i,6)=0;
end
end
person(i,44)=person(i,1);
person(i,1)=person(a,1)+count;
person(i,7)=0;
if person(i,6)==1
    person(i,8)=0;
else
    person(i,8)=-1;
end
person(i,9)=avinc1;
person(i,10:12)=0;
person(i,13:16)=0;
y=rand;
if y<=nnwrktot/nadulttot
    person(i,13)=1;
elseif y>nnwrktot/nadulttot && y<=(nnwrktot+nstudtot)/nadulttot
    person(i,14)=1;
elseif y>(nnwrktot+nstudtot)/nadulttot
    person(i,15)=1;
elseif y>(nnwrktot+nstudtot+nparttot)/nadulttot
    person(i,16)=1;
end
person(i,17:19)=0;
for k=20:23
    if person(i,k-7)==1
        person(i,k)=1;
    else
        person(i,k)=0;
    end
end
person(i,30)=1;
person(i,31)=person(i,1)*100+person(i,30);
person(i,33)=person(i,32)*100+1;
person(i,36:42)=person(i,17:23);
for k=1:7
    if person(i,35+k)==1
        person(i,43)=k;
    end
end

```

```

        end
    end
end
elseif i>=a-24
if person(i,4)>1 && sum(person(i,13:16))>2
    if person(i,35)==18
        count=count+1;
        x=person(i,4);
        for j=i-x:a;
            if person(i,1)==person(j,1)
                person(j,4)=person(j,4)-1;
                autoratio=person(j,6)/(sum(person(j,12:16)));
                if autoratio>0.5
                    person(j,6)=person(j,6)-1;
                    person(i,6)=1;
                else
                    person(j,6)=person(j,6);
                    person(i,6)=0;
                end
            end
        end
    end
    person(i,44)=person(i,1);
    person(i,1)=person(a,1)+count;
    person(i,7)=0;
    if person(i,6)==1
        person(i,8)=0;
    else
        person(i,8)=-1;
    end
    person(i,9)=avinc1;
    person(i,10:12)=0;
    person(i,13:16)=0;
    y=rand;
    if y<=nnwrktot/nadulttot
        person(i,13)=1;
    elseif y>nnwrktot/nadulttot && y<=(nnwrktot+nstudtot)/nadulttot
        person(i,14)=1;
    elseif y>(nnwrktot+nstudtot)/nadulttot
        &&
y<=(nnwrktot+nstudtot+nparttot)/nadulttot
        person(i,15)=1;
    elseif y>(nnwrktot+nstudtot+nparttot)/nadulttot
        person(i,16)=1;
    end
    person(i,17:19)=0;
    for k=20:23

```

```

if person(i,k-7)==1
    person(i,k)=1;
else
    person(i,k)=0;
end
end
person(i,30)=1;
person(i,31)=person(i,1)*100+person(i,30);
person(i,33)=person(i,32)*100+1;
person(i,36:42)=person(i,17:23);
for k=1:7
    if person(i,35+k)==1
        person(i,43)=k;
    end
end
end
end
%end of working with person file

```

```

%working with hh file
dummyperson=person;
for i=a-count:a
    for j=1:c
        if person(i,44)==hh(j,1)
            dummyperson(i,22:39)=hh(j,22:39);
            hh(j,3)=hh(j,3)-1;
            if person(i,6)==1
                hh(j,6)=hh(j,6)-1;
            end
            hh(j,17)=hh(j,17)-1;
            if hh(j,17)>0
                hh(j,10)=1;
            else
                hh(j,10)=0;
            end
        end
    end
    hh(c+i-a+count,1)=person(i,1);
    hh(c+i-a+count,2)=person(i,2);
    hh(c+i-a+count,3)=1;
    hh(c+i-a+count,4)=person(i,9);
    hh(c+i-a+count,5)=0;
    hh(c+i-a+count,6)=person(i,6);
    if hh(c+i-a+count,6)==0

```

```

hh(c+i-a+count,7)=1;
else
    hh(c+i-a+count,7)=0;
end
hh(c+i-a+count,8:14)=person(i,17:23);
hh(c+i-a+count,15:21)=person(i,10:16);
hh(c+i-a+count,22:39)=dummyperson(i,22:39);
end
%end of working with hh file
%end of moving out module

%Auto ownership model
indicator=zeros(countfam,7);
for i=c+1:c+countfam
    if sum(hh(i,18:21))>=2
        indicator(i-c,1)=1;
    end
    if indicator(i-c,1)==1 && hh(i,8)==1
        indicator(i-c,2)=1;
    end
    if indicator(i-c,1)==1 && hh(i,8)==0 && hh(i,9)==1
        indicator(i-c,3)=1;
    end
    if sum(hh(i,18:21))==1 && hh(i,8)==0 && hh(i,9)==0 && hh(i,10)==1
        indicator(i-c,4)=1;
    end
    if sum(hh(i,18:21))>=2 && hh(i,8)==0 && hh(i,9)==0 && hh(i,10)==1
        indicator(i-c,5)=1;
    end
    if hh(i,40)==1
        indicator(i-c,6)=1;
    end
    if hh(i,40)>=2
        indicator(i-c,7)=1;
    end
end

%Reading coeffs. of poisson regression model of vehicle ownership
vehown=xlsread('vehownershippoissoncoeff.xls');
%End of reading coeffs. of poisson regression model of vehicle
%ownership

for i=c+1:c+countfam
    %hh(i,41)=0.0000147*hh(i,4)-
    0.062*hh(i,3)+1.767*sum(hh(i,17:21))+0.277*sum(hh(i,20:21))+0.973*indicator(i-

```

```

c,1)+0.857*indicator(i-c,2)+0.967*indicator(i-c,3)+0.580*indicator(i-c,4)+0.786*indicator(i-
c,5)+0.160*indicator(i-c,6)+0.873*indicator(i-c,7);

hh(i,41)=exp(vehown(1)+vehown(2)*hh(i,4)+vehown(3)*hh(i,3)+vehown(4)*sum(hh(i,17:21))+
vehown(5)*sum(hh(i,20:21))+vehown(6)*indicator(i-c,1)+vehown(7)*indicator(i-
c,2)+vehown(8)*indicator(i-c,3)+vehown(9)*indicator(i-c,4)+vehown(10)*indicator(i-
c,5)+vehown(11)*indicator(i-c,6)+vehown(12)*indicator(i-c,7));

if hh(i,41)<=3.648
    hh(i,6)=1;
elseif hh(i,41)>3.648 && hh(i,41)<=6.583
    hh(i,6)=2;
elseif hh(i,41)>6.583 && hh(i,41)<=8.291
    hh(i,6)=3;
else
    hh(i,6)=4;
end
hh(i,7)=0;
end
for i=1:c+countfam
    if hh(i,6)<0
        hh(i,6)=0;
    end
end
%End of vehicle ownership model

```

```

% TRIP GENERATION MODULE
[c,d]=size(hh);
trip=zeros(c,18);
if year==5
    for i=1:c
        trip(i,1)=hh(i,2);
        trip(i,2)=exp(0.3246*hh(i,3)+0.0311*hh(i,6)+5.91*10^-6*hh(i,4)-2.38*10^-
11*(hh(i,4)^2)-0.2688*hh(i,15));
        %trip(i,2)=0.622+3.786*hh(i,3)+0.208*hh(i,6)-3.491*hh(i,15);
        trip(i,3)=0;
        %if trip(i,2)<0
        %    trip(i,2)=10;
        %end
    end
elseif year==10
    for i=1:c
        trip(i,4)=hh(i,2);
        trip(i,4)=exp(0.3246*hh(i,3)+0.0311*hh(i,6)+5.91*10^-6*hh(i,4)-2.38*10^-
11*(hh(i,4)^2)-0.2688*hh(i,15));
    end

```

```

%trip(i,5)=0.622+3.786*hh(i,3)+0.208*hh(i,6)-3.491*hh(i,15);
trip(i,6)=0;
%if trip(i,5)<0
%  trip(i,5)=10;
%end
end
elseif year==15
for i=1:c
  trip(i,7)=hh(i,2);
  trip(i,6)=exp(0.3246*hh(i,3)+0.0311*hh(i,6)+5.91*10^-6*hh(i,4)-2.38*10^-
11*(hh(i,4)^2)-0.2688*hh(i,15));
  %trip(i,8)=0.622+3.786*hh(i,3)+0.208*hh(i,6)-3.491*hh(i,15);
  trip(i,9)=0;
  %if trip(i,8)<0
  %  trip(i,8)=10;
  %end
end
elseif year==20
for i=1:c
  trip(i,10)=hh(i,2);
  trip(i,8)=exp(0.3246*hh(i,3)+0.0311*hh(i,6)+5.91*10^-6*hh(i,4)-2.38*10^-
11*(hh(i,4)^2)-0.2688*hh(i,15));
  %trip(i,11)=0.622+3.786*hh(i,3)+0.208*hh(i,6)-3.491*hh(i,15);
  trip(i,12)=0;
  %if trip(i,11)<0
  %  trip(i,11)=10;
  %end
end
elseif year==25
for i=1:c
  trip(i,13)=hh(i,2);
  trip(i,10)=exp(0.3246*hh(i,3)+0.0311*hh(i,6)+5.91*10^-6*hh(i,4)-2.38*10^-
11*(hh(i,4)^2)-0.2688*hh(i,15));
  %trip(i,14)=0.622+3.786*hh(i,3)+0.208*hh(i,6)-3.491*hh(i,15);
  trip(i,15)=0;
  %if trip(i,14)<0
  %  trip(i,14)=10;
  %end
end
elseif year==30
for i=1:c
  trip(i,16)=hh(i,2);
  trip(i,12)=exp(0.3246*hh(i,3)+0.0311*hh(i,6)+5.91*10^-6*hh(i,4)-2.38*10^-
11*(hh(i,4)^2)-0.2688*hh(i,15));
  %trip(i,17)=0.622+3.786*hh(i,3)+0.208*hh(i,6)-3.491*hh(i,15);
  trip(i,18)=0;

```

```

%if trip(i,17)<0
%  trip(i,17)=10;
%end
end
end

%Aggregating trip data by taz
for i=1:c
  for j=1:6
    if trip(i,3j-1)<0
      trip(i,3j-1)=10;
    end
  end
end
hhtripgen=zeros(1074,2);
for i=1:1074
  hhtripgen(i,1)=i;
end
if year==5
  hhtripgen(:,2)=0;
  for i=1:1074
    for j=1:c
      if hhtripgen(i,1)==trip(j,1)
        hhtripgen(i,2)=hhtripgen(i,2)+trip(j,2);
      end
    end
  end
elseif year==10
  hhtripgen(:,2)=0;
  for i=1:1074
    for j=1:c
      if hhtripgen(i,1)==trip(j,4)
        hhtripgen(i,2)=hhtripgen(i,2)+trip(j,5);
      end
    end
  end
elseif year==15
  hhtripgen(:,2)=0;
  for i=1:1074
    for j=1:c
      if hhtripgen(i,1)==trip(j,7)
        hhtripgen(i,2)=hhtripgen(i,2)+trip(j,8);
      end
    end
  end
elseif year==20

```

```

hhtripgen(:,2)=0;
for i=1:1074
    for j=1:c
        if hhtripgen(i,1)==trip(j,10)
            hhtripgen(i,2)=hhtripgen(i,2)+trip(j,11);
        end
    end
end
elseif year==25
    hhtripgen(:,2)=0;
    for i=1:1074
        for j=1:c
            if hhtripgen(i,1)==trip(j,13)
                hhtripgen(i,2)=hhtripgen(i,2)+trip(j,14);
            end
        end
    end
elseif year==30
    hhtripgen(:,2)=0;
    for i=1:1074
        for j=1:c
            if hhtripgen(i,1)==trip(j,16)
                hhtripgen(i,2)=hhtripgen(i,2)+trip(j,2);
            end
        end
    end
end

%End of aggregating trip data by taz
%End of trip generation module

hh(:,40:41)=[];
if time==5
    person1=person;
    person2=person;
    person3=person;
    person4=person;
    hh1=hh;
    hh2=hh;
    trip1=trip;
    trip2=trip;
    if a<65000
        person_output=xlswrite('individual5.xls',person);
    elseif a>=65000 && a<130000
        person1(65000:a,:)=[];
    end
end

```

```

person_output1=xlswrite('individual5_1.xls',person1);
person2(1:64999,:)=[];
person_output2=xlswrite('individual5_2.xls',person2);
elseif a>=130000 && a<195000
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual5_1.xls',person1);
    person2(130000:a,:)=[];
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual5_2.xls',person2);
    person3(1:129999,:)=[];
    person_output3=xlswrite('individual5_3.xls',person3);
else
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual5_1.xls',person1);
    person2(130000:a,:)=[];
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual5_2.xls',person2);
    person3(1:129999,:)=[];
    person3(195000:a,:)=[];
    person_output3=xlswrite('individual5_3.xls',person3);
    person4(1:194999,:)=[];
    person_output4=xlswrite('individual5_4.xls',person4);
end
if c<65000
    hh_output=xlswrite('hh5.xls',hh);
else
    hh1(65000:c,:)=[];
    hh_output1=xlswrite('hh5_1.xls',hh1);
    hh2(1:64999,:)=[];
    hh_output2=xlswrite('hh5_2.xls',hh2);
end
if c<65000
    trip_output=xlswrite('trip5.xls',trip);
else
    trip1(65000:c,:)=[];
    trip_output1=xlswrite('trip5_1.xls',trip1);
    trip2(1:64999,:)=[];
    trip_output2=xlswrite('trip5_2.xls',trip2);
end
xlswrite('hhtripgen5.7nov.exp.xls',hhtripgen);
elseif time==10
    person1=person;
    person2=person;
    person3=person;
    person4=person;
    hh1=hh;

```

```

hh2=hh;
trip1=trip;
trip2=trip;
if a<65000
    person_output=xlswrite('individual10.xls',person);
elseif a>=65000 && a<130000
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual10_1.xls',person1);
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual10_2.xls',person2);
elseif a>=130000 && a<195000
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual10_1.xls',person1);
    person2(130000:a,:)=[];
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual10_2.xls',person2);
    person3(1:129999,:)=[];
    person_output3=xlswrite('individual10_3.xls',person3);
else
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual10_1.xls',person1);
    person2(130000:a,:)=[];
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual10_2.xls',person2);
    person3(1:129999,:)=[];
    person3(195000:a,:)=[];
    person_output3=xlswrite('individual10_3.xls',person3);
    person4(1:194999,:)=[];
    person_output4=xlswrite('individual10_4.xls',person4);
end
if c<65000
    hh_output=xlswrite('hh10.xls',hh);
else
    hh1(65000:c,:)=[];
    hh_output1=xlswrite('hh10_1.xls',hh1);
    hh2(1:64999,:)=[];
    hh_output2=xlswrite('hh10_2.xls',hh2);
end
if c<65000
    trip_output=xlswrite('trip10.xls',trip);
else
    trip1(65000:c,:)=[];
    trip_output1=xlswrite('trip10_1.xls',trip1);
    trip2(1:64999,:)=[];
    trip_output2=xlswrite('trip10_2.xls',trip2);
end

```

```

xlswrite('hhtripgen10.7nov.exp.xls',hhtripgen);
elseif time==15
    person1=person;
    person2=person;
    person3=person;
    person4=person;
    hh1=hh;
    hh2=hh;
    trip1=trip;
    trip2=trip;
    if a<65000
        person_output=xlswrite('individual15.xls',person);
    elseif a>=65000 && a<130000
        person1(65000:a,:)=[];
        person_output1=xlswrite('individual15_1.xls',person1);
        person2(1:64999,:)=[];
        person_output2=xlswrite('individual15_2.xls',person2);
    elseif a>=130000 && a<195000
        person1(65000:a,:)=[];
        person_output1=xlswrite('individual15_1.xls',person1);
        person2(130000:a,:)=[];
        person2(1:64999,:)=[];
        person_output2=xlswrite('individual15_2.xls',person2);
        person3(1:129999,:)=[];
        person_output3=xlswrite('individual15_3.xls',person3);
    else
        person1(65000:a,:)=[];
        person_output1=xlswrite('individual15_1.xls',person1);
        person2(130000:a,:)=[];
        person2(1:64999,:)=[];
        person_output2=xlswrite('individual15_2.xls',person2);
        person3(1:129999,:)=[];
        person3(195000:a,:)=[];
        person_output3=xlswrite('individual15_3.xls',person3);
        person4(1:194999,:)=[];
        person_output4=xlswrite('individual15_4.xls',person4);
    end
if c<65000
    hh_output=xlswrite('hh15.xls',hh);
else
    hh1(65000:c,:)=[];
    hh_output1=xlswrite('hh15_1.xls',hh1);
    hh2(1:64999,:)=[];
    hh_output2=xlswrite('hh15_2.xls',hh2);
end
if c<65000

```

```

trip_output=xlswrite('trip15.xls',trip);
else
    trip1(65000:c,:)=[];
    trip_output1=xlswrite('trip15_1.xls',trip1);
    trip2(1:64999,:)=[];
    trip_output2=xlswrite('trip15_2.xls',trip2);
end
xlswrite('hhtripgen15.7nov.exp.xls',hhtripgen);
elseif time==20
    person1=person;
    person2=person;
    person3=person;
    person4=person;
    hh1=hh;
    hh2=hh;
    trip1=trip;
    trip2=trip;
    if a<65000
        person_output=xlswrite('individual20.xls',person);
    elseif a>=65000 && a<130000
        person1(65000:a,:)=[];
        person_output1=xlswrite('individual20_1.xls',person1);
        person2(1:64999,:)=[];
        person_output2=xlswrite('individual20_2.xls',person2);
    elseif a>=130000 && a<195000
        person1(65000:a,:)=[];
        person_output1=xlswrite('individual20_1.xls',person1);
        person2(130000:a,:)=[];
        person2(1:64999,:)=[];
        person_output2=xlswrite('individual20_2.xls',person2);
        person3(1:129999,:)=[];
        person_output3=xlswrite('individual20_3.xls',person3);
    else
        person1(65000:a,:)=[];
        person_output1=xlswrite('individual20_1.xls',person1);
        person2(130000:a,:)=[];
        person2(1:64999,:)=[];
        person_output2=xlswrite('individual20_2.xls',person2);
        person3(1:129999,:)=[];
        person3(195000:a,:)=[];
        person_output3=xlswrite('individual20_3.xls',person3);
        person4(1:194999,:)=[];
        person_output4=xlswrite('individual20_4.xls',person4);
    end
    if c<65000
        hh_output=xlswrite('hh20.xls',hh);

```

```

else
    hh1(65000:c,:)=[];
    hh_output1=xlswrite('hh20_1.xls',hh1);
    hh2(1:64999,:)=[];
    hh_output2=xlswrite('hh20_2.xls',hh2);
end
if c<65000
    trip_output=xlswrite('trip20.xls',trip);
else
    trip1(65000:c,:)=[];
    trip_output1=xlswrite('trip20_1.xls',trip1);
    trip2(1:64999,:)=[];
    trip_output2=xlswrite('trip20_2.xls',trip2);
end
xlswrite('hhtripgen20.7nov.exp.xls',hhtripgen);
elseif time==25
person1=person;
person2=person;
person3=person;
person4=person;
hh1=hh;
hh2=hh;
trip1=trip;
trip2=trip;
if a<65000
    person_output=xlswrite('individual25.xls',person);
elseif a>=65000 && a<130000
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual25_1.xls',person1);
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual25_2.xls',person2);
elseif a>=130000 && a<195000
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual25_1.xls',person1);
    person2(130000:a,:)=[];
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual25_2.xls',person2);
    person3(1:129999,:)=[];
    person_output3=xlswrite('individual25_3.xls',person3);
else
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual25_1.xls',person1);
    person2(130000:a,:)=[];
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual25_2.xls',person2);
    person3(1:129999,:)=[];

```

```

person3(195000:a,:)=[];
person_output3=xlswrite('individual25_3.xls',person3);
person4(1:194999,:)=[];
person_output4=xlswrite('individual25_4.xls',person4);
end
if c<65000
    hh_output=xlswrite('hh25.xls',hh);
else
    hh1(65000:c,:)=[];
    hh_output1=xlswrite('hh25_1.xls',hh1);
    hh2(1:64999,:)=[];
    hh_output2=xlswrite('hh25_2.xls',hh2);
end
if c<65000
    trip_output=xlswrite('trip25.xls',trip);
else
    trip1(65000:c,:)=[];
    trip_output1=xlswrite('trip25_1.xls',trip1);
    trip2(1:64999,:)=[];
    trip_output2=xlswrite('trip25_2.xls',trip2);
end
xlswrite('hhtripgen25.7nov.exp.xls',hhtripgen);
elseif time==30
    person1=person;
    person2=person;
    person3=person;
    person4=person;
    hh1=hh;
    hh2=hh;
    trip1=trip;
    trip2=trip;
    if a<65000
        person_output=xlswrite('individual30.xls',person);
    elseif a>=65000 && a<130000
        person1(65000:a,:)=[];
        person_output1=xlswrite('individual30_1.xls',person1);
        person2(1:64999,:)=[];
        person_output2=xlswrite('individual30_2.xls',person2);
    elseif a>=130000 && a<195000
        person1(65000:a,:)=[];
        person_output1=xlswrite('individual30_1.xls',person1);
        person2(130000:a,:)=[];
        person2(1:64999,:)=[];
        person_output2=xlswrite('individual30_2.xls',person2);
        person3(1:129999,:)=[];
        person_output3=xlswrite('individual30_3.xls',person3);
    end
end

```

```

else
    person1(65000:a,:)=[];
    person_output1=xlswrite('individual30_1.xls',person1);
    person2(130000:a,:)=[];
    person2(1:64999,:)=[];
    person_output2=xlswrite('individual30_2.xls',person2);
    person3(1:129999,:)=[];
    person3(195000:a,:)=[];
    person_output3=xlswrite('individual30_3.xls',person3);
    person4(1:194999,:)=[];
    person_output4=xlswrite('individual30_4.xls',person4);
end
if c<65000
    hh_output=xlswrite('hh30.xls',hh);
else
    hh1(65000:c,:)=[];
    hh_output1=xlswrite('hh30_1.xls',hh1);
    hh2(1:64999,:)=[];
    hh_output2=xlswrite('hh30_2.xls',hh2);
end
if c<65000
    trip_output=xlswrite('trip30.xls',trip);
else
    trip1(65000:c,:)=[];
    trip_output1=xlswrite('trip30_1.xls',trip1);
    trip2(1:64999,:)=[];
    trip_output2=xlswrite('trip30_2.xls',trip2);
end
xlswrite('hhtripgen30.7nov.exp.xls',hhtripgen);
end
end
%End of household simulation

```

#### %FIRM PART

```

t=30;
yearcount=0;
for year_firm=1:t
    yearcount=yearcount+1;
    empdata_initial=empdata;

% Updating firm size
[m,n]=size(empdata);
rand1=(1:m)';

```

```

for i = 1:m
    rand1(i)=rand;
    sizecat(i)=empdata(i,4);
end
for i = 1:m
    if empdata(i,2)==1
        j=empdata(i,4);
        if rand1(i)<=markov_basic(j,1)
            empdata(i,4)=1;
        elseif rand1(i)>markov_basic(j,1) && rand1(i)<=sum(markov_basic(j,1:2))
            empdata(i,4)=2;
        elseif rand1(i)>sum(markov_basic(j,1:2)) && rand1(i)<=sum(markov_basic(j,1:3))
            empdata(i,4)=3;
        elseif rand1(i)>sum(markov_basic(j,1:3)) && rand1(i)<=sum(markov_basic(j,1:4))
            empdata(i,4)=4;
        elseif rand1(i)>sum(markov_basic(j,1:4)) && rand1(i)<=sum(markov_basic(j,1:5))
            empdata(i,4)=5;
        elseif rand1(i)>sum(markov_basic(j,1:5)) && rand1(i)<=1
            empdata(i,4)=6;
        end
    elseif empdata(i,2)==2
        j=empdata(i,4);
        if rand1(i)<=markov_retail(j,1)
            empdata(i,4)=1;
        elseif rand1(i)>markov_retail(j,1) && rand1(i)<=sum(markov_retail(j,1:2))
            empdata(i,4)=2;
        elseif rand1(i)>sum(markov_retail(j,1:2)) && rand1(i)<=sum(markov_retail(j,1:3))
            empdata(i,4)=3;
        elseif rand1(i)>sum(markov_retail(j,1:3)) && rand1(i)<=sum(markov_retail(j,1:4))
            empdata(i,4)=4;
        elseif rand1(i)>sum(markov_retail(j,1:4)) && rand1(i)<=sum(markov_retail(j,1:5))
            empdata(i,4)=5;
        elseif rand1(i)>sum(markov_retail(j,1:5)) && rand1(i)<=1
            empdata(i,4)=6;
        end
    elseif empdata(i,2)==3
        j=empdata(i,4);
        if rand1(i)<=markov_education(j,1)
            empdata(i,4)=1;
        elseif rand1(i)>markov_education(j,1) && rand1(i)<=sum(markov_education(j,1:2))
            empdata(i,4)=2;
        elseif rand1(i)>sum(markov_education(j,1:2)) && rand1(i)<=sum(markov_education(j,1:3))
            empdata(i,4)=3;
        elseif rand1(i)>sum(markov_education(j,1:3)) && rand1(i)<=sum(markov_education(j,1:4))
            empdata(i,4)=4;
        end
    end
end

```

```

    empdata(i,4)=4;
    elseif rand1(i)>sum(markov_education(j,1:4)) &&
rand1(i)<=sum(markov_education(j,1:5))
        empdata(i,4)=5;
    elseif rand1(i)>sum(markov_education(j,1:5)) && rand1(i)<=1
        empdata(i,4)=6;
    end
elseif empdata(i,2)==4
    j=empdata(i,4);
    if rand1(i)<=markov_service(j,1)
        empdata(i,4)=1;
    elseif rand1(i)>markov_service(j,1) && rand1(i)<=sum(markov_service(j,1:2))
        empdata(i,4)=2;
    elseif rand1(i)>sum(markov_service(j,1:2)) && rand1(i)<=sum(markov_service(j,1:3))
        empdata(i,4)=3;
    elseif rand1(i)>sum(markov_service(j,1:3)) && rand1(i)<=sum(markov_service(j,1:4))
        empdata(i,4)=4;
    elseif rand1(i)>sum(markov_service(j,1:4)) && rand1(i)<=sum(markov_service(j,1:5))
        empdata(i,4)=5;
    elseif rand1(i)>sum(markov_service(j,1:5)) && rand1(i)<=1
        empdata(i,4)=6;
    end
end
empdata(:,5)=empdata(:,2)*100+empdata(:,4);
for i=1:m
    j=empdata(i,4)-sizecat(i);
    if j~=0
        if empdata(i,4)==1
            empdata(i,3)=unidrnd(4);
        elseif empdata(i,4)==2
            empdata(i,3)=4+unidrnd(5);
        elseif empdata(i,4)==3
            empdata(i,3)=9+unidrnd(10);
        elseif empdata(i,4)==4
            empdata(i,3)=19+unidrnd(80);
        elseif empdata(i,4)==5
            empdata(i,3)=99+unidrnd(400);
        elseif empdata(i,4)==6
            if empdata(i,2)==1
                %empdata(i,3)=499+poissrnd(1146);
                empdata(i,3)=500;
            elseif empdata(i,2)==2
                %empdata(i,3)=499+poissrnd(547);
                empdata(i,3)=500;

```

```

elseif empdata(i,2)==3
    %empdata(i,3)=499+poissrnd(2019);
    empdata(i,3)=500;
elseif empdata(i,2)==4
    %empdata(i,3)=499+poissrnd(628);
    empdata(i,3)=500;
% for size cat = 6 and sizes=(size-500)
% basic: mean=1146 (stdev=1288)
% retail: mean=547 (stdev=653)
% education: mean=2019 (stdev=4880)
% service: mean=628 (stdev=727)
    end
end
r=empdata(i,2);
empdata(i,5+r)=empdata(i,3);
end
% Firm size updated

%Birth Module
%empdata2=empdata;
[m,n]=size(empdata);
rand2=(1:m)';
for i = 1:m
    rand2(i)=0;
end
for i = 1:m
    if empdata(i,5)==101
        rand2(i)=rand;
    elseif empdata(i,5)==102
        rand2(i)=2+rand;
    elseif empdata(i,5)==103
        rand2(i)=4+rand;
    elseif empdata(i,5)==104
        rand2(i)=6+rand;
    elseif empdata(i,5)==105
        rand2(i)=8+rand;
    elseif empdata(i,5)==106
        rand2(i)=10+rand;
    elseif empdata(i,5)==201
        rand2(i)=100+rand;
    elseif empdata(i,5)==202
        rand2(i)=102+rand;
    elseif empdata(i,5)==203
        rand2(i)=104+rand;
    elseif empdata(i,5)==204

```

```

rand2(i)=106+rand;
elseif empdata(i,5)==205
    rand2(i)=108+rand;
elseif empdata(i,5)==206
    rand2(i)=110+rand;
elseif empdata(i,5)==301
    rand2(i)=200+rand;
elseif empdata(i,5)==302
    rand2(i)=202+rand;
elseif empdata(i,5)==303
    rand2(i)=204+rand;
elseif empdata(i,5)==304
    rand2(i)=206+rand;
elseif empdata(i,5)==305
    rand2(i)=208+rand;
elseif empdata(i,5)==306
    rand2(i)=210+rand;
elseif empdata(i,5)==401
    rand2(i)=300+rand;
elseif empdata(i,5)==402
    rand2(i)=302+rand;
elseif empdata(i,5)==403
    rand2(i)=304+rand;
elseif empdata(i,5)==404
    rand2(i)=306+rand;
elseif empdata(i,5)==405
    rand2(i)=308+rand;
else
    rand2(i)=310+rand;
end
end
birth1=(1:m)';
for i = 1:m
    birth1(i)=0;
end
for i=1:m
    for j=0:5
        if rand2(i) > (2*j-1) && rand2(i) < 2*j+(Basic(j+1,4)/100)
            birth1(i)=j+1;
        end
    end
end
for i=1:m
    for j=0:5
        if rand2(i) > 100+(2*j-1) && rand2(i) < 100+2*j+(Retail(j+1,4)/100)
            birth1(i)=100+j+1;
        end
    end
end

```

```

    end
end
for i=1:m
for j=0:5
if rand2(i) > 200+(2*j-1) && rand2(i) < 200+2*j+(Education(j+1,4)/100)
    birth1(i)=200+j+1;
end
end
end
for i=1:m
for j=0:5
if rand2(i) > 300+(2*j-1) && rand2(i) < 300+2*j+(Service(j+1,4)/100)
    birth1(i)=300+j+1;
end
end
end
end
%birth2=xlswrite('birth1.0.xls',birth1);
%End of Firm Birth Module

% Updating Location Choice Module
% Travel time increases by 3% on all routes/links

% Location Choice Scenario 1 & 2 and current scenario
[m,n]=size(locchoice);
[x,y]=size(logsum);
if yearcount==1
else
logsum(:,2)=1.03*logsum(:,2);
end

for i=1:x
if logsum(i,5)==0
    logsum(i,6)=0;
else
    logsum(i,6)=log(logsum(i,5))/exp(.0896124*logsum(i,2)-.0238147*logsum(i,3));
end
if logsum(i,4)==0
    logsum(i,7)=0;
else
    logsum(i,7)=log(logsum(i,4))/exp(.1028167*logsum(i,2)-.0459523*logsum(i,3));
end
end
locchoice(:,21)=0;
locchoice(:,22)=0;
for i=1:m

```

```

for j=1:x
    if locchoice(i,1)==logsum(j,1)
        locchoice(i,21)=locchoice(i,21)+logsum(j,6);
        locchoice(i,22)=locchoice(i,22)+logsum(j,7);
    else
        locchoice(i,21)=locchoice(i,21);
        locchoice(i,22)=locchoice(i,22);
    end
end
end

locchoice(:,23)=0.507*log(locchoice(:,2))-3.142*locchoice(:,3)-1.495*locchoice(:,4)-
0.816*locchoice(:,5)-0.004*locchoice(:,16)+0.00000131*locchoice(:,11)+0.011*locchoice(:,21)-
0.010*locchoice(:,22);
%locchoice(:,23)=4.948+0.482*log(locchoice(:,2))-3.025*locchoice(:,3)-
1.464*locchoice(:,4)-0.778*locchoice(:,5)+0.00000109*locchoice(:,11)+0.011*locchoice(:,21)-
0.010*locchoice(:,22);
locchoice(:,24)=exp(locchoice(:,23));

locchoice(:,25)=locchoice(:,24)/sum(locchoice(:,24));

tazlocprob(:,22)=locchoice(:,25);

for i=1:m
    tazlocprob(i,21)=sum(tazlocprob(1:i,22));
end
% Location Choice Module updated

%Location Choice Module for new born firms
[m,n]=size(empdata);
rand3=(1:m)';
for i = 1:m
    rand3(i)=-1;
end
for i = 1:m
    if birth1(i)~=0
        rand3(i)=rand;
    end
end
[tazrow,tazcol]=size(tazlocprob);
count=0;
for i=1:m
    if rand3(i)~-=-1
        count=count+1;
        if rand3(i)<tazlocprob(1,21)
            empdata(m+count,1)=yearcount*100000+count;
        end
    end
end

```

```

empdata(m+count,2:9)=empdata(i,2:9);
empdata(m+count,10:29)=tazlocprob(1,1:20);
else
    for j=2:tazrow
        if rand3(i)<tazlocprob(j,21) && rand3(i)>tazlocprob(j-1,21)
            empdata(m+count,1)=yearcount*100000+count;
            empdata(m+count,2:9)=empdata(i,2:9);
            empdata(m+count,10:29)=tazlocprob(j,1:20);
        end
    end
end
end
end
%empdata3=xlswrite('empdata3.0.xls',empdata2);
%End of Location Choice Module for new born firms

% Death/Exit Module
[m,n]=size(empdata_initial);
rand1=(1:m)';
for i = 1:m
    rand1(i)=0;
end
for i = 1:m
    if empdata(i,5)==101
        rand1(i)=rand;
    elseif empdata(i,5)==102
        rand1(i)=2+rand;
    elseif empdata(i,5)==103
        rand1(i)=4+rand;
    elseif empdata(i,5)==104
        rand1(i)=6+rand;
    elseif empdata(i,5)==105
        rand1(i)=8+rand;
    elseif empdata(i,5)==106
        rand1(i)=10+rand;
    elseif empdata(i,5)==201
        rand1(i)=100+rand;
    elseif empdata(i,5)==202
        rand1(i)=102+rand;
    elseif empdata(i,5)==203
        rand1(i)=104+rand;
    elseif empdata(i,5)==204
        rand1(i)=106+rand;
    elseif empdata(i,5)==205
        rand1(i)=108+rand;
    elseif empdata(i,5)==206

```

```

rand1(i)=110+rand;
elseif empdata(i,5)==301
    rand1(i)=200+rand;
elseif empdata(i,5)==302
    rand1(i)=202+rand;
elseif empdata(i,5)==303
    rand1(i)=204+rand;
elseif empdata(i,5)==304
    rand1(i)=206+rand;
elseif empdata(i,5)==305
    rand1(i)=208+rand;
elseif empdata(i,5)==306
    rand1(i)=210+rand;
elseif empdata(i,5)==401
    rand1(i)=300+rand;
elseif empdata(i,5)==402
    rand1(i)=302+rand;
elseif empdata(i,5)==403
    rand1(i)=304+rand;
elseif empdata(i,5)==404
    rand1(i)=306+rand;
elseif empdata(i,5)==405
    rand1(i)=308+rand;
elseif empdata(i,5)==406
    rand1(i)=310+rand;
end
end
death1=(1:m)';
for i = 1:m
    death1(i)=0;
end
for i=1:m
    for j=0:5
        if rand1(i) > (2*j-1) && rand1(i) < 2*j+(Basic(j+1,5)/100)
            death1(i)=j+1;
        end
    end
end
for i=1:m
    for j=0:5
        if rand1(i) > 100+(2*j-1) && rand1(i) < 100+2*j+(Retail(j+1,5)/100)
            death1(i)=100+j+1;
        end
    end
end
for i=1:m

```

```

for j=0:5
    if rand1(i) > 200+(2*j-1) && rand1(i) < 200+2*j+(Education(j+1,5)/100)
        death1(i)=200+j+1;
    end
end
for i=1:m
    for j=0:5
        if rand1(i) > 300+(2*j-1) && rand1(i) < 300+2*j+(Service(j+1,5)/100)
            death1(i)=300+j+1;
        end
    end
end
%death2=xlswrite('death1.0.xls',death1);
i=1;
count=0;
while i<=m
    if death1(i)~=0
        empdata(i-count,:)=[];
        count=count+1;
    else
        empdata(i-count,:)=empdata(i-count,:);
    end
    i=i+1;
end
%empdata1=xlswrite('empdata1.0.xls',empdata);
% End of Death/Exit Module

%Location Choice Module for relocating firms
reloc=tazlocprob;
[x,y]=size(reloc);
reloc=sortrows(reloc,22);
utilindex=sum(reloc(x-10:x,22))/10;
firmcount=sum(reloc(:,20));
flag=round(firmcount*0.05);
count=0;
relocindex=ones(x,1);
relocindex(:)=0;
for i=1:x
    if count<(flag*3)
        count=count+reloc(i,20);
        relocindex(i)=reloc(i,1);
    end
end
[m,n]=size(empdata);

```

```

temp(1:m,1)=0;
for i=1:x
    for j=1:m
        if relocindex(i)==empdata(j,10)
            temp(j)=rand;
        end
    end
end
for i=1:m
    if temp(i)>0.067
        temp(i)=rand;
    else
        temp(i)=0;
    end
end
[tazrow,tazcol]=size(tazlocprob);
for i=1:m
    if temp(i)<tazlocprob(1,21) && temp(i)>0
        empdata(i,10:29)=tazlocprob(1,1:20);
    else
        for j=2:tazrow
            if temp(i)<tazlocprob(j,21) && temp(i)>tazlocprob(j-1,21)
                empdata(i,10:29)=tazlocprob(j,1:20);
            end
        end
    end
end
%empdata4=xlswrite('empdata4.0.xls', empdata3);
%End of Location Choice Module for relocating firms

%Trip Generation Module
%empdata3(:,25)=(n-5)=Trip Gen Percentage by TAZ
[x,y]=size(tazlocprob);
temptg=ones(x,1);
double(temptg);
for i=1:x

temptg(i)=tripgen(1)+tripgen(2)*tazlocprob(i,2)*tazlocprob(i,3)+tripgen(3)*tazlocprob(i,2)*tazlocprob(i,4)+tripgen(4)*tazlocprob(i,2)*tazlocprob(i,5)+tripgen(5)*tazlocprob(i,20)+tripgen(6)*tazlocprob(i,18)+tripgen(7)*tazlocprob(i,17)+tripgen(8)*tazlocprob(i,16)+tripgen(9)*tazlocprob(i,15)+tripgen(10)*tazlocprob(i,14)+tripgen(11)*tazlocprob(i,13);
end
temptgsum=sum(temptg);
temptg=(temptg/temptgsum)*100;
tazlocprob(:,25)=temptg;
if yearcount==5

```

```

commtripgen=zeros(x,2);
for i=1:x
    commtripgen(i,1)=i;
    commtripgen(i,2)=2*temptg(i);
end
end
%End of Trip Generation Module

%Updating FirmCount and Employment Stats
%empdata4=empdata3;
[m,n]=size(empdata);
empdata(:,28)=0;
empdata(:,29)=0;
for i=1:m
    for j=1:m
        if empdata(i,10)==empdata(j,10)
            empdata(i,29)=empdata(i,29)+1;
            empdata(i,28)=empdata(i,28)+empdata(j,3);
        else
            empdata(i,28)=empdata(i,28);
            empdata(i,29)=empdata(i,29);
        end
    end
end
% End of Firm count and employment status updation

if yearcount==5
    xlswrite('empdata5.7nov.exp.xls',empdata);
    xlswrite('commtripgen5.7nov.exp.xls',commtripgen);
elseif yearcount==10
    xlswrite('empdata10.7nov.exp.xls',empdata);
    xlswrite('commtripgen10.7nov.exp.xls',commtripgen);
elseif yearcount==15
    xlswrite('empdata15.7nov.exp.xls',empdata);
    xlswrite('commtripgen15.7nov.exp.xls',commtripgen);
elseif yearcount==20
    xlswrite('empdata20.7nov.exp.xls',empdata);
    xlswrite('commtripgen20.7nov.exp.xls',commtripgen);
elseif yearcount==25
    xlswrite('empdata25.7nov.exp.xls',empdata);
    xlswrite('commtripgen25.7nov.exp.xls',commtripgen);
elseif yearcount==30
    xlswrite('empdata30.7nov.exp.xls',empdata);
    xlswrite('commtripgen30.7nov.exp.xls',commtripgen);
end
%end

```

% End of printing output at every 5 years  
%Updating trip times (adding 3% every year)

end

%end iteration loop

## APPENDIX B: FIRM LOCATION CHOICE MODELS

Table B-1: Results of Firm Location Choice Model for Each Sector

	<b>Basic Firms</b>		<b>Retail Firms</b>		<b>Education Firms</b>		<b>Service Firms</b>	
	30 alternatives		30 alternatives		30 alternatives		30 alternatives	
	Coeff.	t-stats	Coeff.	t-stats	Coeff.	t-stats	Coeff.	t-stats
RURAL*	-1.354	-13.99	-3.183	-67.38	-1.942	-7.57	-2.243	-26.91
SUBURB*	-0.779	-8.87	-1.915	-53.56	-1.251	-6.09	-1.188	-18.49
URBAN*	-0.474	-5.52	-1.278	-38.86	-0.913	-4.65	-0.533	-8.87
TRAN25	<b>-0.770</b>	-15.97	<b>-0.662</b>	-22.45	<b>0.296</b>	1.88	<b>0.201</b>	3.92
SQRTHH	0.039	36.07	0.036	57.63	0.035	10.21	0.021	19.15
CBDDIST	0.008	4.19	<b>-0.019</b>	-11.94	0.016	2.78	0.016	7.08
DISTI35	-	-	<b>0.028</b>	12.36	-	-	<b>-0.011</b>	-2.97
Log-likelihood	-22,209.811		-53,970.730		-2,016.855		-19,018.948	
Adj. R <sup>2</sup>	0.060		0.131		0.087		0.098	
No. of Observations	<b>6,944</b>		<b>6,203</b>		<b>650</b>		<b>18,266</b>	