

1
2
3 **Tracking the Size, Location and Interactions of Businesses: Microsimulation of Firm**
4 **Behavior in Austin, Texas**
5

6
7 Saurabh Kumar
8 Graduate Student Researcher
9 The University of Texas at Austin
10 B.124 E. Cockrell Jr. Hall
11 Austin, TX 78712-1076
12 saurabh@mail.utexas.edu

13
14 Kara M. Kockelman
15 (Corresponding author)
16 Associate Professor and William J. Murray Jr. Fellow
17 Department of Civil, Architectural and Environmental Engineering
18 The University of Texas at Austin
19 6.9 E. Cockrell Jr. Hall
20 Austin, TX 78712-1076
21 kcockelm@mail.utexas.edu
22 Phone: 512-471-0210
23 FAX: 512-475-8744

24 The following is a pre-print, the final publication can be found in *Transportation Research*
25 *Record* No. 2077: 113-121, 2008.
26

27 **ABSTRACT**
28

29 Firms are key drivers of urban growth along with households, yet data on businesses and
30 commercial vehicle movements can be quite difficult to obtain. An understanding of firm
31 behavior over time is critical in anticipating urban futures and addressing transportation, land use
32 and other concerns. Firm birth and death, migration and location choice are defining events in a
33 firm's life cycle, and a study of firm evolution requires estimating and applying models for each
34 of these. Such an exercise is hindered primarily by a lack of quality micro-data for businesses.
35 This paper proposes a basic framework for modeling firm demographics using a microsimulation
36 approach. Year 2005 employment point data for the Austin, Texas region and 7 years of
37 aggregate data from the Statistics of U.S. Businesses have been used to simulate firm entry, exit
38 and evolution over time and space. A Markov process is used to anticipate firm growth and
39 contraction, along with logit and Poisson models for firm location choice. Austin's commercial
40 vehicle survey data are used to estimate commercial trip generation and distribution models, and
41 applications of these are tied to forecasts of firm counts by location over time. Simulation results
42 for both low- and high-growth scenarios suggest an increasing movement of firms towards
43 central zones.
44

45 **Keywords:** Firmography, microsimulation, location choice, commercial vehicle movement,
46 Markov transition process
47
48
49
50
51
52

INTRODUCTION

Policymakers, planners and researchers are interested in answering various questions raised by urbanization and infrastructure's ability to cope with added demands. Most research has focused on studying the behavior of 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 key agents of urban development, households have been studied extensively, using sound behavioral specifications and data. The literature lacks similar studies tracking firm behavior (the focus of this study), owing to a lack of quality data (as compared to households) and greater variety (and thus uncertainty) in firm behavior (Elgar et al., 2006).

The demography of firms (firmography) may be key in understanding the process of urbanization and forecasting urban growth. Zones of high employment are major attractors and producers of personal and commercial trips. Since, commercial trips constitute a major portion of urban trips, an understanding of firm behavior should help improve transportation planning. Effect of demographic events in firms' life cycle on land use and urban travel patterns is also generating interest among urban planning agencies. Hence, this study aims to study the interaction of firm dynamics and transport infrastructure over time and spatial dimensions.

LITERATURE REVIEW

The study of firmography is multidisciplinary in nature involving 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, Wadell et al. 2003, Maoh et al. 2005, and Abraham and Hunt 2000), among others. Past studies have researched firmography from the viewpoint of a single discipline as well as from a multidisciplinary angle (Van Wissen, 2000). Key steps in the firmography process 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.

A common objective in firmographic studies is modeling the location choices of new firms and relocation decisions of existing firms. Neo-classical and behavioral schools of thought have proposed different theories to explain such location decisions (Hayter, 1997). 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 market-segmentation analysis based on industrial sector and size (De Bok and Sanders 2005, Wadell et al. 2003, Khan et al. 2002, 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 in U.S. as choice alternatives for firm location choice models. To account for agglomeration economies and labor market effects (which vary at much finer levels of spatial resolution), Hansen (1987) used cities in and around Sao Paulo, Brazil, and Woodward

1
2
3 (1992) used U.S. counties as choice alternatives. Guimaraés et al. (1998 and 2000) used
4 municipalities in Puerto Rico and concelhos¹ in Portugal as choice alternatives.
5

6
7 Analysis of data at finer spatial scales results in larger choice sets, which can become impractical
8 due to computational requirements (in popular analysis packages, like LimDep and STATA) and
9 problems posed by the independence of irrelevant alternatives property (Ben-Akiva and Lerman
10 1984). To address such issues McFadden (1978) proposed the use of a random sample of choice
11 alternatives, for estimation of the logit model, and showed it to result in consistent estimators.
12 However, efficiency and replications are compromised (Train 1986). The stochastic nature of
13 McFadden's approach also poses the problem of non-replication. Coughlin et al. (2000),
14 Guimaraés et al. (2003) and many others have used Poisson regression to counter the
15 complexities posed by a large set of choice alternatives and computational requirements. When
16 locators are homogeneous, Poisson regression is a tractable econometric alternative to the
17 conditional logit model, consistent with McFadden's (1974) maximization framework.

18
19 Relocation of existing firms is another aspect of location choice studied extensively in 1970s in
20 Europe and U.K. (Pellenbarg et al. 2002). Data improvements and technological advances have
21 improved the spatial resolution of these models, from a regional to more local levels of cities, but
22 a study of firm mobility at the truly local scale remains neglected, to a large extent. Local firm
23 movements affect local travel patterns and land use, as well as the local economy (Van Dijk and
24 Pellenbarg 2000b). While most researchers have focused on factors affecting location choices of
25 new firms, Lee (2006) used a binary probit model to examine the effect of firm and location
26 specific factors on the relocation decision and Van Dijk and Pellenbarg (1999) used an ordered²
27 logit model to study firm relocation patterns in Netherlands.

28
29 The modeling paradigm has changed considerably over the years, with more emphasis on
30 individual agents. The basis for such a change in perspective is based on the assumption that
31 tracking an individual's choices may lead to more behaviorally reasonable models and enhanced
32 forecasting. Activity-based and microsimulation models of travel demand are starting to replace
33 traditional aggregate models (e.g., Bhat et al's (2004) CEMDAP and the FHWA's (2006)
34 TRANSIMS). Microsimulation provides the flexibility of modeling individual behaviors and
35 aggregating results as needed. Disaggregate models provide footprints of behavior, while
36 microsimulation tracks individual's transitions through space and time. Microsimulation models
37 of land use include Miller et al's (2001) Integrated Land Use, Transport and Environment
38 (ILUTE) model, Waddell et al's (2003) UrbanSim, Arentze and Timmermans' (2004) A
39 Learning-Based Transportation Oriented Simulations System (ALBATROSS), and Moeckel et
40 al's (2002) Integrated Land-Use Modeling and Transportation System Simulation (ILUMASS).
41 ILUTE operates on a 100% sample of households and firms and simulates location choice as
42 well as firmography and household demographics.
43
44
45
46

47
48 ¹ As reasonably small administrative areas, Portugal's concelhos have an average area of 322.5 km².

49 ² Firm propensity to move was expressed in percentages (0%, 0–10%, 10–25%, 25–50%, 50–75%, 75–90%, 90–
50 100% and 100%) and used as the dependent variable in an ordered logit model.
51
52

1
2
3 Commercial trips³ account for 10 to 15% of intra-urban vehicle trips (Hunt et al. 2004) and even
4 higher share of inter-urban trips. Commercial trips also account for roughly 12 to 15% of VMT
5 (Hunt et al. (2004) and Spear (2006)). The recent surge in microsimulation models of household
6 travel would be well supplemented by similar models for commercial trips. Hunt et al. (2004,
7 2005) estimated a tour-based microsimulation model of commercial movements in the city of
8 Calgary, Canada. Commercial trip patterns are strongly tied to firm location choices and this
9 study highlights the same.
10

11 DATA DESCRIPTION

12
13 Point location data for all firms in the three-county Austin region was provided by the Texas
14 Workforce Commission and geocoded by the Capital Area Metropolitan Planning Organization
15 (CAMPO). In 2005 Hays, Travis and Williamson counties contained 32,063⁴ firms employing
16 655,722 full- and part-time workers. The study region consists of 1,074 Traffic Analysis Zones
17 (TAZs), and zone-level statistics for the year 2005 were imputed from CAMPO's 1997 and 2007
18 data sets. 57% of the firms belong to the service sector⁵, and 26% of firms are single-worker
19 firms. The average firm size is 20 workers, with a standard deviation is 155 owing mostly to the
20 presence of a few, very large educational establishments⁶. Figure 1 shows the distribution of
21 TAZs and firms by area type. Around 45% of Austin's TAZs are coded as rural⁷; these account
22 for 85% of the three-county area but only 16% of firms and 12% of workers. In contrast, only
23 22% of the TAZs are coded as urban or CBD (accounting for a mere 3% of area yet 45% of the
24 firms and 59% of workers).
25

26 The Statistics of U.S. Businesses (SUSB) provide annual data⁸ on the number of firms and new
27 birth and deaths for employment size of firm categories by location and industry. The data are a
28 "snapshot" of firms at a point in time. SUSB's firm data from year 1998 to 2004 were used for
29 generating Markov transition matrices of firm growth. 1998 to 2004 was a slow growth period
30 for industries with a couple of years of recession. New firm birth and death data from year 2001-
31 2002 and 2002-2003 were used for modeling firm births and deaths. While most of the industries
32 witnessed slow rate of growth in 2001-2002, 2002-2003 showed higher growth rates.
33

34 The Texas Department of Transportation (DOT) provided commercial vehicle survey data for the
35 region. In 2006, 342 vehicles were randomly selected from the study area's list of registered
36

37
38 ³ Commercial vehicles were classified as trucks used for commercial or public agency-purposes with six or more
39 tires and having a curb weight of 6,000 pounds or more.

40 ⁴ Addresses of 5 firms could not be mapped to the 3-county Austin region and hence were removed from the study
41 population.

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

46 ⁶ The University of Texas at Austin, Texas State University and Austin Community College have 20,250, 4,884 and
47 3,390 workers in 2005, respectively.

48 ⁷ Austin's zone designations are based on a population-equivalent density (PED), equaling $[ZonePop +$
49 $(RegionalPop/RegionalEmp)*ZoneEmp)]/(ZoneAcres)$. A zone with $PED \geq 15$ is coded as CBD,
50 $8 \leq PED < 15$ is coded as Urban, $1 \leq PED < 8$ is designated Suburban, and $PED < 1$ is defined as Rural.

51 ⁸ The new birth and death data is available only for years 2001-2002 and 2002-2003 while aggregate number of
52 firms in each size category and sector is available for 1998 through 2004.

1
2
3 commercial vehicles. These 342 vehicles were based in 113 TAZs and made 2,551 trips over the
4 survey day.
5

6 **MODEL SPECIFICATION**

7
8 A microsimulation approach is used to simulate individual firm life cycles in an urban
9 environment. Microsimulation is critical in tracking each decision maker over its life cycle and
10 provides a better insight into firm behavior than forecasts made using aggregate cross-sectional
11 data (Goulias and Kitamura, 1992). One key component of microsimulating firmography lies in
12 estimating the statistical models of life cycle transitions. Lack of quality data is the biggest
13 challenge in developing such a model. Of course any model of human behavior is an abstraction
14 of reality and the uncertainty in business environments and behavioral dependence on various
15 unpredictable and uncontrolled factors raises questions regarding interpretation of simulation
16 results. Nonetheless, such modeling exercises can provide insights into how firms react to
17 various changes in their urban environment and land use and transport components.
18

19 The basic structure for firm synthesis and evolution used here is shown in Figure 2. The firm
20 population is updated every year by running through a sequence of sub-models. A model of firm
21 death/exit is applied first, and exiting firms are removed from the population. This is followed by
22 an expansion/contraction model for remaining firms in the population. All existing firms are
23 allowed to relocate based on utility differential at their current location and ten most popular
24 TAZs (based on levels of TAZ-specific systematic utility). New firms are added (based on the
25 birth model) and their locations are chosen based on the location choice model. Commercial trip
26 generation and distribution models are applied at time t based on the current firm population.
27 Firm population and travel patterns are updated every year.
28

29 A lack of panel data necessitates the use of Markov decision process⁹ for many submodules in
30 the microsimulation process. Firms are modeled on a three-dimensional state space (S) of size,
31 industry affiliation and location. Hence, the Markovian probability of any event in a firm's life
32 cycle is dependent only on its current size (s), and industry affiliation (i). The Markovian series
33 of states (over time) can be expressed as:
34

$$35 \Pr(S(t) = s_t, i_t) = \sum_{s_{t-1}} \sum_{i_{t-1}} \Pr(S(t) | S(t-1)) \times \Pr(S(t-1) = s_{t-1}, i_{t-1}) \quad (1)$$

36
37
38 where $\Pr(S(t) | S(t-1))$ is one step (one year) transition probability from state at time $t-1$ to state at
39 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})
40 and is independent of earlier states ($t' < t-1$).
41

42 *Firm Death/Exit and Birth Model*

43 SUSB estimates for U.S. firm death/exit (2001-2002 for low-growth and 2002-2003 for high-
44 growth estimates), as a function of firm size category s and industry sector i , were used to
45 randomly sample firms from the population for removal. Assumptions for the birth model are
46

47
48 ⁹ Assumption of a Markov transition process implies that firm history (e.g., how long a firm has existed or been in
49 its size category) does not affect its future growth. While such assumptions limit data requirements, ideally they
50 should be tested and improved upon, as needed.
51
52

similar to those for the death/exit model, in terms of the use of SUSB estimates and sampling. New firm attributes were (randomly) selected from the sample space (as defined by the existing firm population). Location information was provided each new-born firm simply by applying the firm location choice submodule (with a multinomial assignment based on random utility values).

Firm Expansion/Contraction Model

Using SUSB data on annual changes in firm counts (by industry sector and size) from years 1998 through 2004, a Markovian transition matrix was estimated via constrained least squares. SUSB data provide the aggregate proportions $y_j(t)$ and $y_i(t-1)$ of firms in size categories j and i over time. Using tenets of conditional probability 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):

$$y_j(t) = \sum_i y_i(t-1)p_{ij} + \varepsilon_j(t) \quad (2)$$

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 ε_j is the stochastic error term. This can be written in matrix notation as:

$$\mathbf{y} = \mathbf{Xp} + \boldsymbol{\varepsilon} \quad (3)$$

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

$$\begin{aligned} & \underset{\{p_{ij}\}}{\text{Minimize}} f(\mathbf{p}) = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{Xp})'(\mathbf{y} - \mathbf{Xp}) \\ & \text{Subject to } \sum_i p_{ij} \leq 1 \quad \forall i \quad \& \quad 0 \leq p_{ij} \leq 1 \quad \forall i, j \end{aligned} \quad (4)$$

As an example, estimates of transition probabilities $\{p_{ij}\}$ resulting from this optimization are shown in Table 1, for basic firms. Using such Markovian transition probabilities and Monte Carlo sampling methods, firm and job counts are updated annually.

Firm Location Choice Model

While actual location choices are made at the parcel or building suite levels, it is very difficult to model location choice at such resolution due to computational and data availability issues. This study thus emphasizes firm location decisions at the TAZ level across the 3-county Austin region.

Let the utility derived by firm i by locating at TAZ j be given by:

$$U_{ij} = \beta X_{ij} + \varepsilon_{ij} \quad (5)$$

Where β is a vector of unknown parameters to be estimated, X_j is a vector of explanatory variables associated with the choice alternative, and ε_{ij} is the error (stochastic) term. Under a

random utility maximization framework and assuming a generalized extreme value type-I (Gumbel) distribution for ε_{ij} , this results in the following probabilities:

$$P_{ij} = \frac{e^{\beta X_{ij}}}{\sum_k e^{\beta X_{ik}}} \quad (6)$$

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:

$$L = \sum_i^{32,063} \sum_j^{1,074} \delta_{ij} \log P_{ij} \quad (7)$$

where δ_{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. Essentially, the log-likelihood reduces to:

$$L = \sum_j \eta_j \log P_j \quad (8)$$

Where η_j is the number of firms choosing to locate 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 (7) and (8) imply a Poisson model of firm counts, by zone. In the case of a Poisson probability for firm counts, the probabilities are as follows:

$$P(\eta_j, \lambda_j) = \frac{e^{-\lambda_j} \lambda_j^{\eta_j}}{\eta_j!} \quad (9)$$

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

$$L = \sum_j^{1074} (-\lambda_j + \eta_j \log \lambda_j - \log \eta_j!) \quad (10)$$

Where $\lambda_j = \exp(\alpha + \beta X_j)$ (for consistency with equation (8)), so that equation (10) reduces to:

$$L = \sum_j \eta_j \log P_j + \text{constant} \quad (11)$$

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 for β (and similar standard errors for the β 's). Guimaraes et al. (2003) compared logit model of location choice with Poisson regression model and concluded that logit model coefficients are equivalent (as long as one includes the full set of alternatives in the probabilities of the likelihood function).

1
2
3
4 This work's model of firm location choice model is estimated using the employment point data
5 provided by CAMPO for the 3-county region. Logsums from a commercial trip distribution
6 model serve as zonal accessibility indices. The accessibilities were computed as follows:
7

$$8 \quad Access_i = \sum_j \frac{A_j}{e^{(\gamma_1 \cdot time_{ij} + \gamma_2 \cdot dist_{ij})}} \quad (12)$$

9
10
11 where A_j is the attribute under consideration (such as population or employment), $time_{ij}$ is the
12 peak period travel time from zone i to zone j , $dist_{ij}$ is the distance between centroids of zone i and
13 zone j , and γ_1 and γ_2 are parameters from logit models of commercial trip destination choice.
14 All new-born and relocating firms are assigned a zone according to probabilities from this model
15 of firm location choice.¹⁰
16

17 *Firm Relocation Model*

18 Kroll et al. (1990) found that roughly 5% of firms relocate locally (within a county) every year.
19 Relocating firms were selected randomly (with a sampling rate of 15%) from among the one-
20 third of firms exhibiting the lowest systematic utilities, as estimated using the location choice
21 model. New locations were assigned to these 15% of the low-utility firms based on firm-specific
22 probabilities determined in the firm location choice model (using appropriately weighted Monte
23 Carlo draws from a multinomial across all possible zones).
24
25

26 *Commercial Trip Generation and Distribution Models*

27 Commercial trip movements were modeled using trip generation and distribution models. A
28 model of commercial trip generation was estimated using 2006 Austin Commercial Vehicle
29 Survey data provided by the Texas DOT. The number of trips originating from each TAZ was
30 regressed on firmographic and location attributes of the TAZ. Trip counts were modeled using a
31 negative binomial regression (recognizing their integer nature, while allowing for overdispersion
32 due to things like unobserved heterogeneity).
33

34 Trip distribution was modeled as a logit model, using the proportion of trips destined to a TAZ as
35 the dependent variable. Demographic attributes and firm and employment counts at the
36 destination TAZs serve as explanatory variables, along with trip times and distances.
37

38 **MODEL RESULTS**

39
40 Microsimulation of firm evolution was carried out using MATLAB. An initial population of
41 32,063 firms was taken through the series of models at one-year steps to forecast firm population,
42 location and related trips 30 years into the future. Firm population profiles and commercial trip
43 generation and attraction profiles were updated at the end of each one-year run of the simulation
44 model for all TAZs.
45
46
47

48 ¹⁰ While separate location choice models were initially estimated for each of the four industries, the estimates were
49 quite similar, so a single model of location choice is used here, without loss in generality.
50
51
52

Table 2 lists the results of Poisson regression model of firm location choices. 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, everything else constant, they appear to be less likely to locate in TAZs with high population accessibility, possibly due to congestion, higher land prices and zoning restrictions.

As discussed earlier commercial trip generation and destination choice models were estimated for the region using 2006 commercial vehicle survey data. Trip counts originating from a TAZ follow a negative binomial distribution, and produce a high overdispersion parameter, of $\alpha = 3.161$ ¹¹. Table 3 shows the results of this model. 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. Table 4 shows the results of the logit model of commercial trip distribution, including a pseudo¹² r^2 value of 0.180 (which is a very reasonable fit for a logit model of 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.

Assuming year 2001/2002 rates of firm exit, birth and growth/contraction for different sectors and sizes, Figure 3 illustrates simulated changes in firm size by sector over time, for the low-growth scenario. These results suggest a steady increase in service firms while other sectors exhibit are predicted to endure losses in firm numbers, with this trend stabilizing over time. Large firms (> 100 workers) show a sharp increase at the cost of smaller firms initially, but this trend also stabilizes.

Of course, every simulation is different, and it is important to get a sense of the simulation's own variability. 25 runs of each scenario were performed, and these exhibit less than 10% variation (from lowest simulated value to highest value) in the total number of firms in different categories for both scenarios by the year 2035. Much greater variation in firm counts generally is evident at the level of individual zones, due to their smaller size and the discrete nature of firm existence and location choice. However, in congested parts of the region, a 10% variation can remain. For example, in the high-growth scenario, CBD zones are estimated to end up with a total of 3,612 to 3,927 firms, while in the low-growth scenario, these numbers vary between 2,884 and 3,118 (versus 2,163 firms in the year 2005).

¹¹ Negative binomial distributions can be written as follows:

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

where $E[Y] = \lambda$, $\text{Var}(Y) = \lambda(1 + \lambda\alpha)$, and α is called the overdispersion parameter.

¹² McFadden's Likelihood Ratio Index also referred to as a pseudo- R^2 : $r^2 = 1 - \frac{\ln \hat{L}(\text{full})}{\ln \hat{L}(\text{intercept})}$ where $\hat{L}(\text{full})$ is

estimated likelihood for model with predictors and $\hat{L}(\text{intercept})$ is for the no-information model (with only constant terms).

1
2
3 Employment counts in each category, larger firms, and educational firms showed relatively
4 higher variance than others. Markov transition probabilities developed from all years of the
5 SUSB data set appear skewed towards larger firms (possibly due to smaller sample sizes in this
6 size group), producing strong growth in large firm sizes at the expense of smaller firms.
7

8
9 Figure 4 illustrates changes in firm size and sector profiles over time for the high-growth
10 scenario. The simulation implies a steady increase in number of firms in all sectors. However,
11 the increase in basic and retail firms is much less pronounced than in that seen in education and
12 service firms.

13
14 Figure 5 shows the changes in the spatial distribution of firms and employment over time. The
15 models predict greater spatial concentration than presently experienced in the Austin region,
16 since firms are expected to favor urban and CBD zones, in both the high- and low-growth
17 scenarios. Such inward movement seems unlikely and is probably due to the use of cross-
18 sectional data sets for calibration of the location choice model (rather than use of panel sets and a
19 dynamic model focused on location changes). A model of land development also would be very
20 useful in this context, in order to synchronize demand with supply. This will be a useful model
21 feature to incorporate in the future.

22 **CONCLUSIONS AND EXTENSIONS**

23
24 This paper presents a modeling framework for tracking events in firm life cycles using data from
25 Austin, Texas. Microsimulation is a powerful tool, but models are only as good as the data that
26 underlie them. A lack of quality business data is a severe handicap in modeling firm behavior
27 effectively, throughout the world. The results of this study provides a good starting point for
28 developing such models and offer insights into the effects of various attributes on firm behavior.
29

30
31 This paper uses Poisson regression as an alternative to logit models, facilitating the estimation of
32 location choice models at finer spatial resolution. Microsimulation of firm size and location
33 choices was carried out by industry sector. Certain Markovian process assumptions to model
34 such changes were necessitated by data constraints but can be easily relaxed within the modeling
35 framework once longitudinal data on individual firms becomes available. A major limitation of
36 this paper is the use of SUSB data for Texas to model firm evolution, birth and death patterns,
37 when Austin may not be a representative of entire Texas.

38
39 With availability of high-quality data, microsimulation can help forecast the future effectively
40 while formulating policies for directing and managing urban growth. These models of firm
41 behavior may serve as a valuable supplement to household simulation models; together such
42 agents dominate urban travel and land use patterns. Moreover, an appreciation of commercial
43 vehicle movements can be very useful in mitigating urban congestion, especially in industrially
44 and commercially developed areas. The approaches used here permit close examination of
45 various policy and infrastructure scenarios.

46
47 Key extensions of this study include integration of models of household behavior, and iteration
48 with models of travel demand, in order to appreciate changes in location access and
49 attractiveness over time. Recognition of the land development process and supply interactions
50
51
52

1
2
3 also is critical for system forecasting. Use of dynamic data to avoid pitfalls in interpretation of
4 cross-sectional models of location choice may also be key.
5

6 Firms and households are the two most important drivers of urban growth, and an integrated
7 model explaining both household and commercial travel patterns will provide policy makers a
8 handy tool to design policies to sustain economic and social systems while combating problems
9 of congestion and sprawl. The work seeks to facilitate adoption of freight modeling methods by
10 MPOs and others.
11

12 REFERENCES

13 Abraham, J. E., and J. D. Hunt. 2000. Firm Location in the MEPLAN Model of Sacramento.
14 *Transportation Research Record*, 1685: 187-198.
15

16 Arentze, T.A., and H. J. P. Timmermans. 2004. A Learning-Based Transportation Oriented
17 Simulation System. *Transportation Research B*, 38: 613-633.
18

19 Bartik, T. 1985. Business Location Decisions in the United States: Estimates of the Effects of
20 Unionization, Taxes, and Other Characteristics of States. *Journal of Business and Economic*
21 *Statistics* 3: 14–22.
22

23 Ben-Akiva, M., and S. R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to*
24 *Travel Demand*. The MIT Press, Cambridge.
25

26 Bhat, C.R., J. Y. Guo, S. Srinivasan, and A. Sivakumar. 2004. A Comprehensive Econometric
27 Microsimulator for Daily Activity-Travel Patterns. *Transportation Research Record*, 1894: 57-
28 66.
29

30 Brouwer, A. E., I. Mariotti, and J. N. van Ommeren. 2002. The Firm Relocation Decision: A
31 Logit Model. Paper presented at the 42nd *ERSA conference*, Dortmund, Germany.
32

33 Carroll G. R., and M. T. Hannan. 2000. *The Demography of Corporations and Industries*.
34 Princeton University Press, Princeton.
35

36 Coughlin, C., J. Terza, and V. Arromdee. 1991. State Characteristics and the Location of Foreign
37 Direct Investment within the United States. *this REVIEW* 73: 675–683.
38

39 Coughlin, C., and E. Segev. 2000. Location Determinants of new Foreign-owned Manufacturing
40 Plants. *Journal of Regional Science*, 40(2): 323-351.
41

42 De Bok, M. and F. Sanders. 2005. Firm Location and the Accessibility of Locations: Empirical
43 Results from the Netherlands. Proceedings of the 84th Annual Meeting of Transportation
44 Research Board, Washington DC.
45

46 De Bok, M., and M. Bliemer. 2005. Using Firm Demographic Microsimulation for Land Use and
47 Transport Scenario Evaluation: Model Calibration. Presented at the 45th European Regional
48 Science Association Congress, Amsterdam, 23-27 August.
49
50
51
52

1
2
3
4 Elgar, I., and E. J. Miller. 2006. A Conceptual Model of Small Office Firm Location.
5 Proceedings of the 85th Annual Meeting of the Transportation Research Board, Washington DC,
6 22-26 January.
7

8 Evans, D. 1987. The Relationship between Firm Growth, Size and Age: Estimates for 100
9 Manufacturing Industries. *Journal of Industrial Economics* 35: 567–581.

10 FHWA. 2006. TRANSIMS Home Page. URL: <http://tmip.fhwa.dot.gov/transims/>. Accessed 10th
11 October 2006.
12

13 Firm Size Data. Office of Advocacy, U. S. Small Business Administration. URL:
14 <http://www.sba.gov/advo/research/data.html#us>. Accessed 10th July 2007.
15

16 Friedman, J., D. Gerlowski, and J. Silberman. 1992. What Attracts Foreign Multinational
17 Corporations? Evidence from Branch Plant Location in the United States. *Journal of Regional*
18 *Science* 32: 403–418.
19

20 Goulias, K. G., and R. Kitamura. 1992. Microsimulation for Travel Demand Forecasting: A
21 Dynamic Model System of Household Demographics and Mobility. Institute of Transportation
22 Studies Research Report, UCD-ITS-RR-92-4, University of California, Davis.
23

24 Guimaraés, P., O. Figueirido, and D. Woodward. 2000. Agglomeration and the Location of
25 Foreign Direct Investment in Portugal. *Journal of Urban Economics* 47: 115–135.
26

27 Guimaraés, P., O. Figueirido, and D. Woodward. 2003. A Tractable Approach to the Firm
28 Location Design Problem. *The Review of Economics and Statistics* 85(1): 201–204.
29

30 Guimaraés, P., and R. Lindrooth. 2005. Dirichlet-Multinomial Regression. *Economics Working*
31 *Paper* Archive at WUSTL, Econometrics, 0509001. [http://ideas.repec.org/p/wpa](http://ideas.repec.org/p/wpa/wuwpem/0509001.html)
32 [/wuwpem/0509001.html](http://ideas.repec.org/p/wpa/wuwpem/0509001.html). Accessed 18th April 2007.
33

34 Guimaraés, P., R. Rolfe, and D. Woodward. 1998. Regional Incentives and Industrial Location in
35 Puerto Rico. *International Regional Science Review* 21(2): 119–138.
36

37 Hannan M. T., G. R. Carroll, S. D. Dobrev, and J. Han. 1988. Organizational Mortality in
38 European and American Automobile Industries Part I: Revisiting the Effects of Age and Size.
39 *European Sociological Review* 14: 279-302.
40

41 Hansen, E. 1987. Industrial Location Choice in Sao Paulo, Brazil: A Nested Logit Model.
42 *Regional Science and Urban Economics* 17: 89–108.
43

44 Head, K., J. Ries, and D. Swenson. 1995. Agglomeration Benefits and Location Choice:
45 Evidence from Japanese Manufacturing Investments in the United States. *Journal of*
46 *International Economics* 38: 223–247.
47
48
49
50
51
52

1
2
3 Hayter, R. 1997. *The Dynamics of Industrial Location. The Factory, The Firm and The*
4 *Production System*. New York: Wiley.
5

6
7 Hunt, J. D., D. S. Kriger, and E. J. Miller. 2005. Current Operational Urban Land Use–
8 Transport Modelling Frameworks: A Review. *Transport Reviews* 25(3): 329–376.

9
10 Hunt, J. D., K. J. Stefan, and J. E. Abraham. 2003. Modeling Retail and Service Delivery
11 Commercial Movement Choice Behaviour in Calgary. Paper presented at the 10th International
12 Conference on Travel Behaviour Research, Lucerne, 10-15 August.
13 <http://www.ivt.baug.ethz.ch/allgemein/slides/hunt.pdf>. Accessed 25th April 2007.

14
15 Hunt, J. D., K. J. Stefan, A. T. Brownlee, J. D. P. McMillan, A. Farhan, K. Tsang, D. Atkins, and
16 M. Ishani. 2004. A Commercial Movement Modelling Strategy for Alberta’s Major Cities. Paper
17 prepared for presentation at the 2004 Annual Conference of the Transportation Association of
18 Canada, Québec City, Québec.

19
20 Jones, M. T. 2005. Estimating Markov Transition Matrices Using Proportions Data: An
21 Application to Credit Risk. *IMF working paper*, Monetary and Financial Systems Department.
22 <http://www.imf.org/external/pubs/ft/wp/2005/wp05219.pdf>. Accessed 10th April 2007.

23
24 Khan, A.S., J.E. Abraham, and J.D. Hunt. 2002. A System for Microsimulating Business
25 Establishments: Analysis, Design and Results. *Access to Activities and Services in Urban*
26 *Canada: Behavioral processes that condition equity and sustainability*, Laval University,
27 Quebec City.

28
29 Kroll, C., and J. Landis. 1990. Housing Prices, Other Real Estate Factors and the Location
30 Choice of Firms. *Fisher Center for Real Estate & Urban Economics*.
31 <http://repositories.cdlib.org/iber/fcreue/reports/0990>. Accessed 26th April 2007.

32
33 Lee, T. C., G. G. Judge, and A. Zellner. 1970. *Estimating the Parameters of the Markov*
34 *Probability Model From Aggregate Time Series Data*. Amsterdam: North Holland.

35
36 Lee, Y. 2006. Relocation Patterns in U.S. Manufacturing. *Working Paper 06-24*, Federal Reserve
37 Bank of Cleveland. www.clevelandfed.org/research. Accessed 9th January 2007.

38
39 Maoh, H. F., and P. S. Kanaroglou. 2005. Agent-Based Firmographic Models: A Simulation
40 Framework for the City of Hamilton. Proceedings of *PROCESSUS Second*
41 *International Colloquium on the Behavioural Foundations of Integrated Land-use and*
42 *Transportation Models: Frameworks, Models and Applications*, Toronto.
43 [http://www.civ.utoronto.ca/sect/traeng/ilute/processus2005/PaperSession/Paper08_Maoh-](http://www.civ.utoronto.ca/sect/traeng/ilute/processus2005/PaperSession/Paper08_Maoh-Kanaroglou_Agent-basedFirmographicModels_CD.pdf)
44 [Kanaroglou_Agent-basedFirmographicModels_CD.pdf](http://www.civ.utoronto.ca/sect/traeng/ilute/processus2005/PaperSession/Paper08_Maoh-Kanaroglou_Agent-basedFirmographicModels_CD.pdf). Accessed 8th November 2006.

45
46 McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. (pp. 105–142),
47 in P. Zarembka (Ed.). *Frontiers in Econometrics*, New York: Academic Press.

48
49 McFadden, D. 1978. Modeling the Choice of Residential Location. (pp. 75–96), in A.
50
51
52

1
2
3 Karquist, L. Lundqvist, F. Snickars, and J. Weibull (Eds.). *Spatial Interaction Theory and*
4 *Planning Models*. Amsterdam: North-Holland.
5

6
7 Miller, E. J., and P. A. Salvini. 1998. The Integrated Land Use, Transportation, Environment
8 (ILUTE) Modeling System: A Framework. Proceedings of the 77th Annual Meeting of the
9 Transportation Research Board, Washington DC, 12-15 January.

10
11 Miller, E. J., and P.A. Salvini. 2001. The Integrated Land Use, Transportation, Environment
12 (ILUTE) Microsimulation Modelling system: Description & Current Status. Chapter 41, in D.
13 Hensher (ed.) *The Leading Edge in Travel Behaviour Research*, Amsterdam: Pergamon, 711-724.

14
15 Moeckel, R. 2005. Simulating Firmography. Fourth Oregon Symposium on Integrating Land Use
16 and Transportation Models, 15-17 November.
17 http://www.oregon.gov/ODOT/TD/TP/docs/Modeling/4symp/1117_1045.pdf. Accessed 5th
18 November 2006.

19
20 Moeckel, R., C. Schurmann, and M. Wegener. 2002. Microsimulation of Urban Land Use.
21 Proceedings of the 42nd European Congress of the Regional Science Association, Dortmund, 27-
22 31 August. <http://www.ersa.org/ersaconfs/ersa02/cd-rom/papers/261.pdf>. Accessed 5th
23 November 2006.

24
25 Nerella, S., and C.R. Bhat. 2004. Numerical Analysis of Effect of Sampling of Alternatives in
26 Discrete Choice Models. *Transportation Research Record* 1894: 11-19.

27
28 Pellenbarg, P. H., and P. J. M. Van Steen. 2003. Spatial Perspectives on Firm Dynamics in the
29 Netherlands. *Tijdschrift voor Economische en Sociale Geografie*, 94(5): 620-630.

30
31 Pellenbarg, P. H., L. J. G. Van Wissen, and J. van Dijk. 2002. Firm Relocation: State of the Art
32 and Research Prospects. *SOM Research Report 02D31*, University of Groningen.
33 <http://som.eldoc.ub.rug.nl/FILES/reports/themeD/2002/02D31/02D31.pdf>. Accessed 15th
34 December 2006

35
36 Salvini, P.A., and E. J. Miller. 2005. ILUTE: An Operational Prototype of a Comprehensive
37 Microsimulation Model of Urban Systems. *Networks and Spatial Economics*, 5: 217-234.

38
39 Spear, B. D. 2006. Freight Modeling in Urban Areas – State of the Practice.
40 www.trb.org/Conferences/2006/FDM/Spear.pdf. Accessed 10th November 2007.

41
42 Stefan, K. J., J. D. P. McMillan, and J. D. Hunt. 2004. An Urban Commercial Vehicle Movement
43 Model for Calgary. Proceedings of the 84th Annual Meeting of the Transportation Research
44 Board, Washington DC.

45
46 Timmermans, H. 2003. The saga of integrated land use-transport modeling: How Many More
47 Dreams Before We Wake Up? Conference Keynote Paper at the 10th International Conference
48 on Travel Behaviour Research, Lucerne, 10-15 August.
49 <http://www.ivt.baug.ethz.ch/allgemein/pdf/timmermans.pdf>. Accessed 10th October 2006.
50
51
52

1
2
3
4 Train, K. 1986. *Qualitative Choice Analysis: Theory, Econometrics , and an Application to*
5 *Automobile Demand*. Cambridge, Massachusetts: The MIT Press.
6

7 Van Dijk, J., and P. H. Pellenbarg. 2000a. Spatial Perspectives on Firm Demography. *Papers in*
8 *Regional Science* 79: 107-110.
9

10 Van Dijk, J., and P. H. Pellenbarg. 2000b. Firm Relocation Decisions in The Netherlands: An
11 Ordered Logit Approach. *Papers in Regional Science* 79: 191-219.
12

13 Van Wissen, L. J. G. 2000. A Micro-Simulation Model of Firms: Applications of Concepts of the
14 Demography of the Firm. *Papers in Regional Science* 79: 111-134.
15

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

19 Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke, and G. Ulfarsson. 2003.
20 Microsimulation of Urban Development and Location Choices: Design and Implementation of
21 UrbanSim. *Networks and Spatial Economics* 3(1): 43-67.
22

23 Woodward, D. 1992. Locational Determinants of Japanese Manufacturing Start-Ups in the
24 United States. *Southern Economic Journal* 58: 690–708.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

LIST OF TABLES AND FIGURES

Figure 1: 2005 Distribution of TAZs (Number and Enclosed Area) and Firms by Area Type

Figure 2: Flowchart for Model of Firm Behaviors (Firm Evolution, Location Choice and Trip Patterns)

Figure 3: Firm Size Profile and Distribution of Firms across Industries (Low-growth Scenario)

Figure 4: Firm Size Profile and Distribution of Firms across Industries (High-growth Scenario)

Figure 5: Firm Distributions in Year 2005 and Forecast for Year 2035

Table 1: Markov Transition Probability Estimates for Basic Firms

Table 2: Poisson Regression Model of Firm Location Choice

Table 3: Negative Binomial Model of Commercial Trip Generation by TAZ

Table 4: Logit model of Commercial Trip Distribution

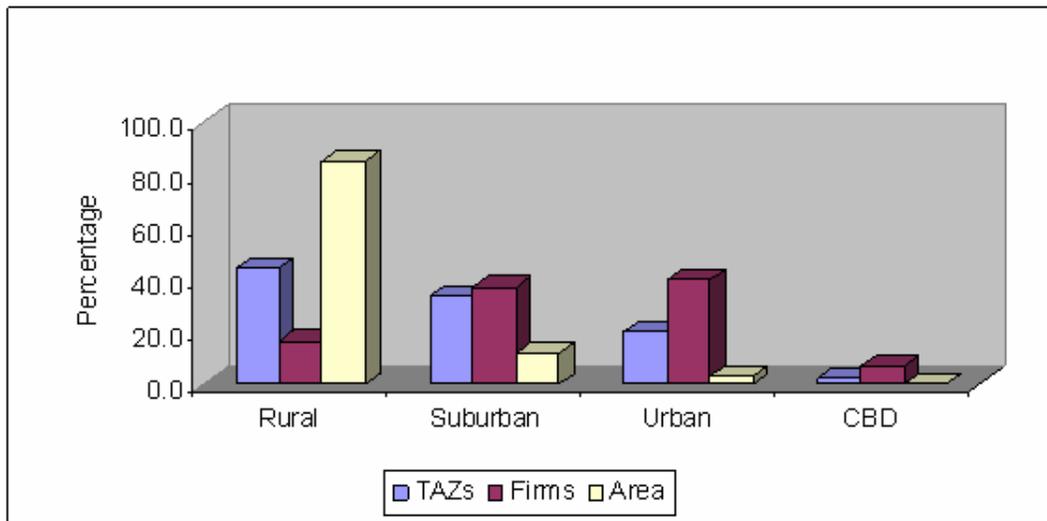


Figure 1: 2005 Distribution of TAZs (Number and Enclosed Area) and Firms by Area Type

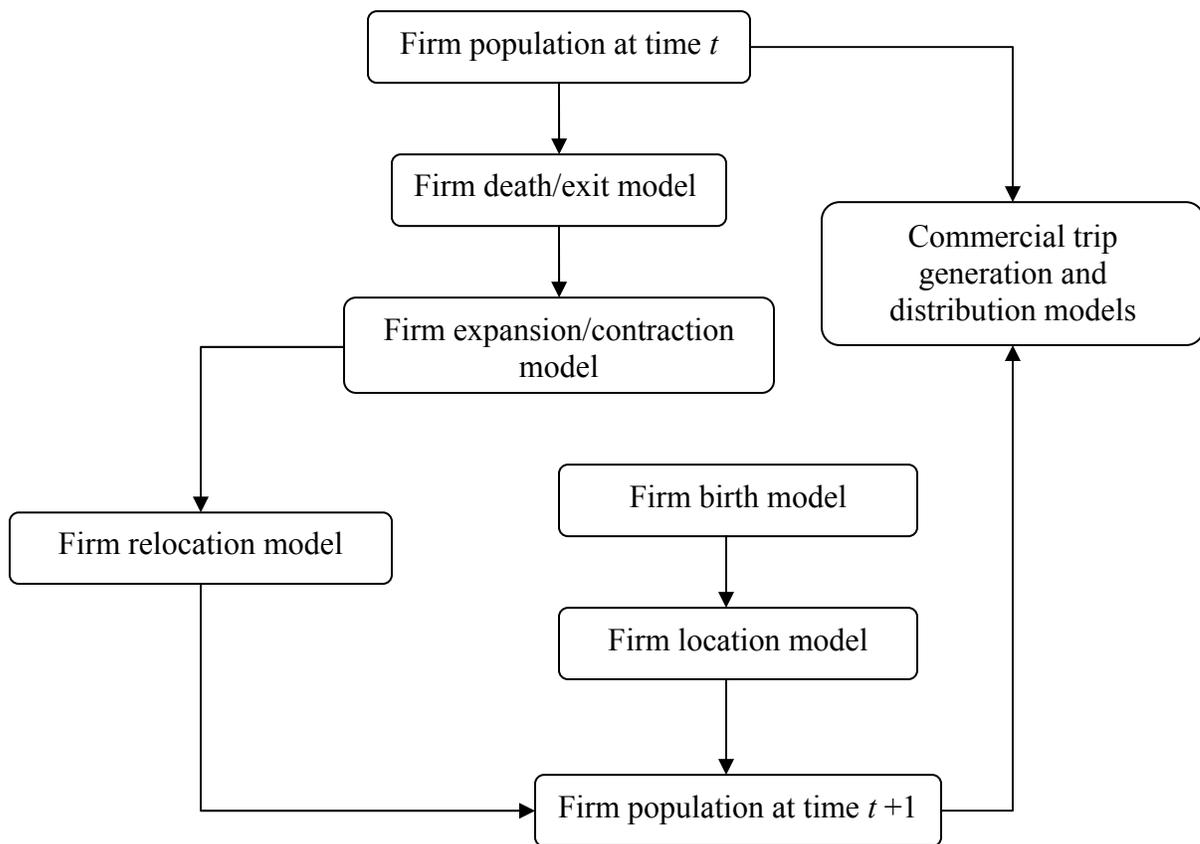


Figure 2: Flowchart for Model of Firm Behaviors (Firm Evolution, Location Choice and Trip Patterns)

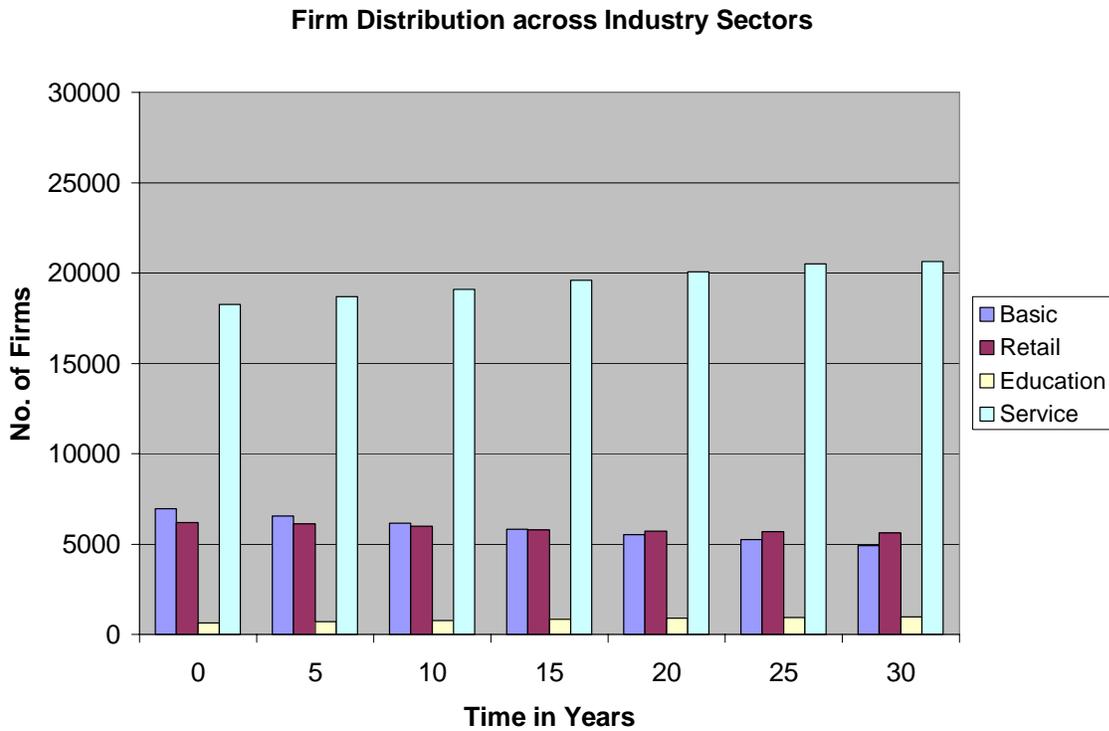
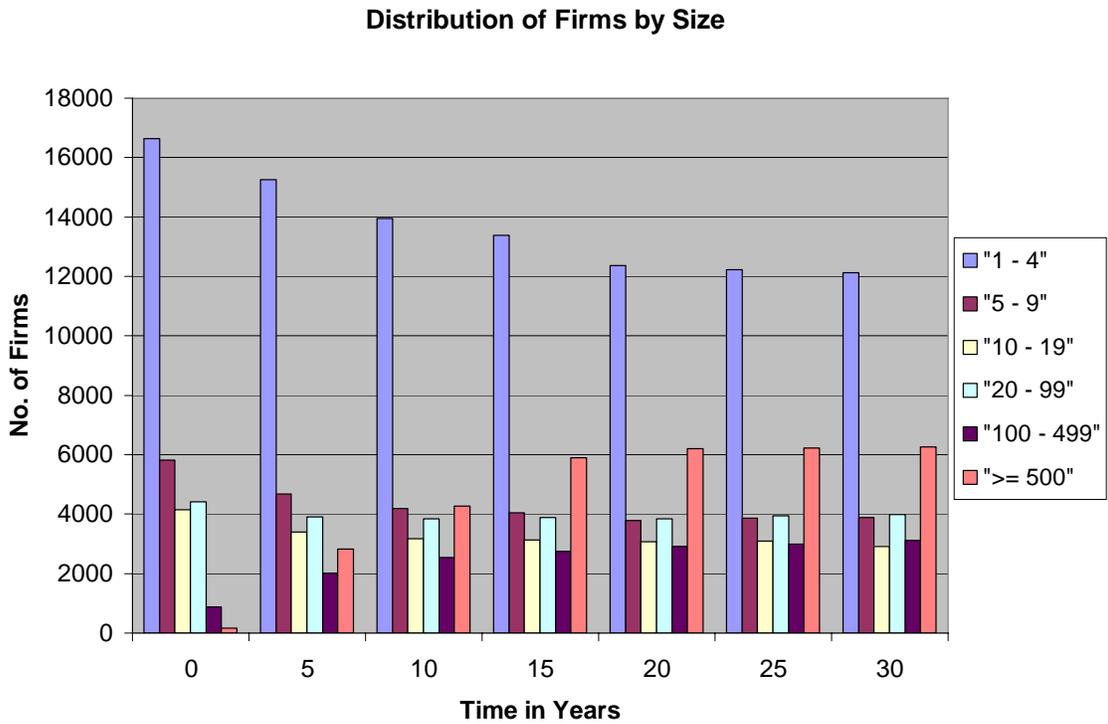


Figure 3: Firm Size Profile and Distribution of Firms across Industries (Low-growth Scenario)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

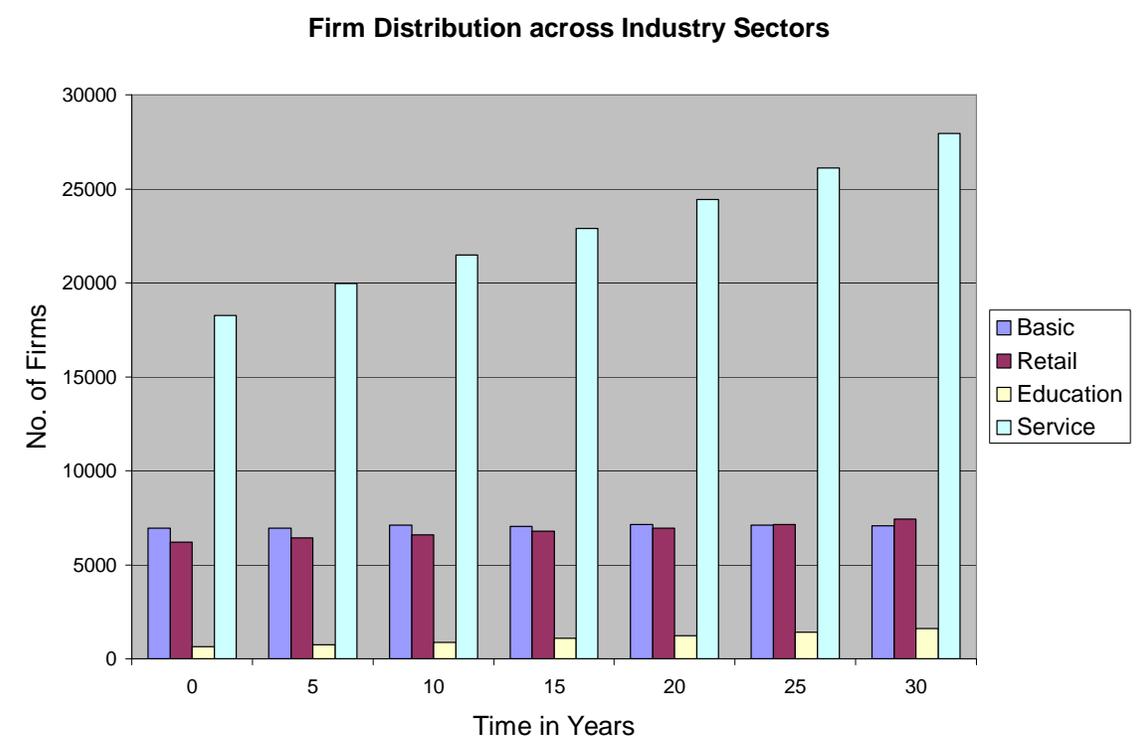
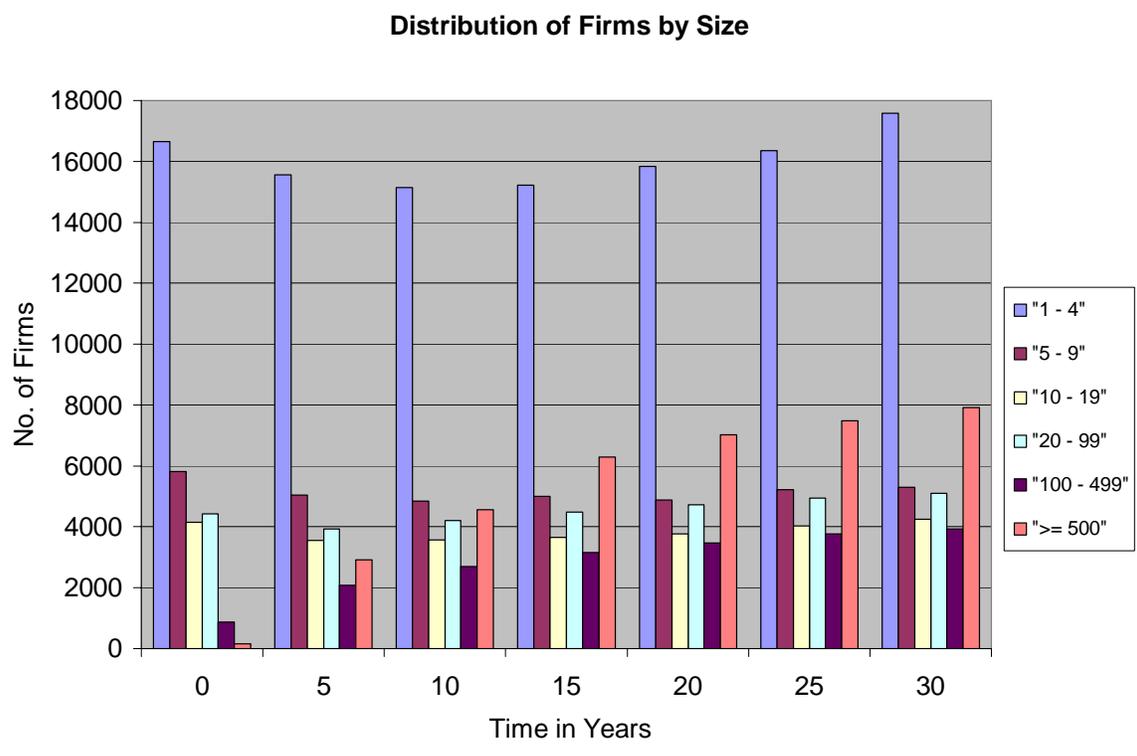


Figure 4: Firm Size Profile and Distribution of Firms across Industries (High-growth Scenario)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

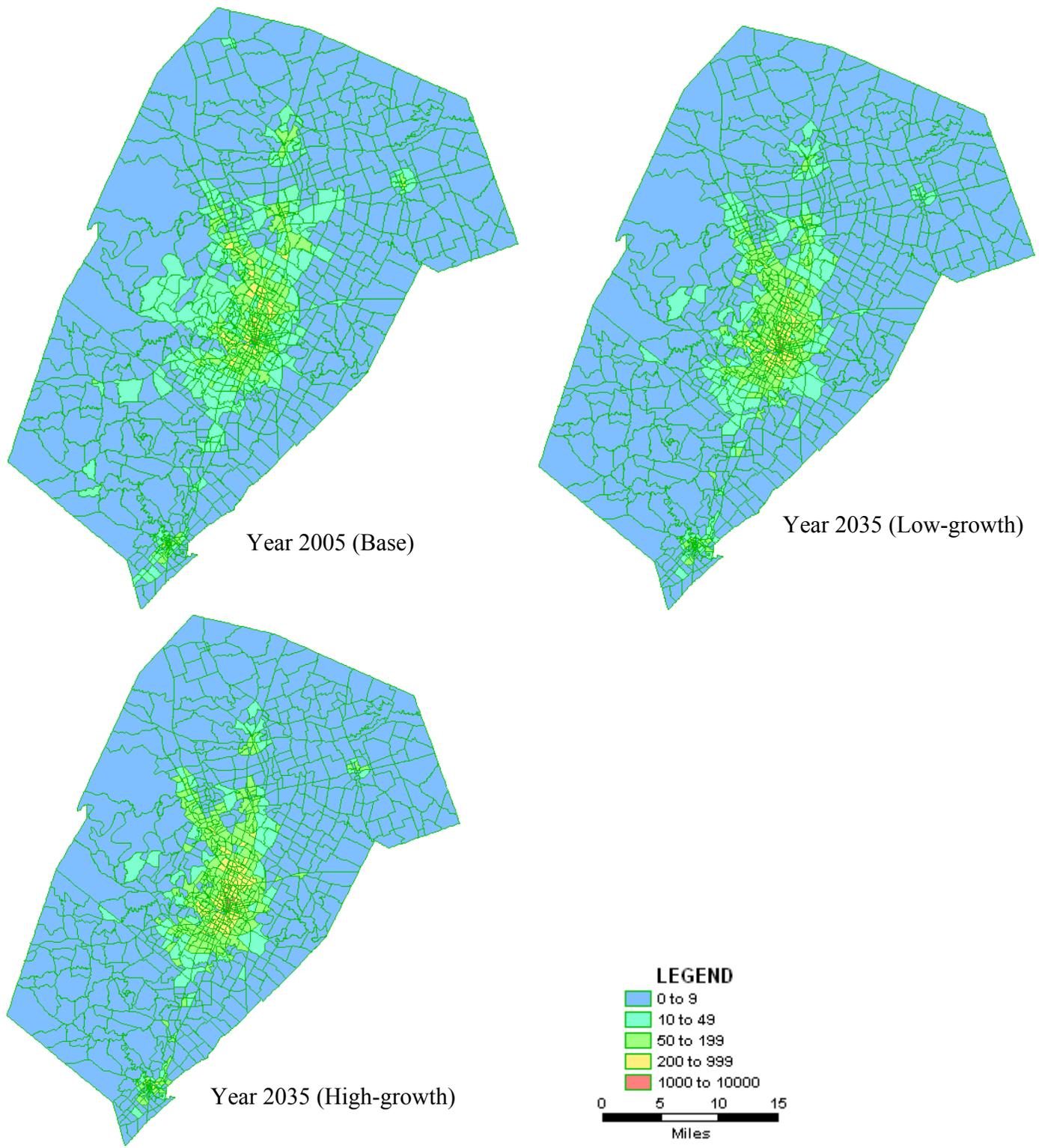


Figure 5: Firm Density by Location in Year 2005 and Forecasts for Year 2035

Table 1: Markov Transition Probability Estimates for Basic Firms

	1-4 employees	5-9	10-19	20-99	100-499	≥ 500
1-4 employees	0.880	0.073	0.021	0.015	0.004	0.008
5-9	0.090	0.779	0.068	0.052	0.011	0.001
10-19	0.039	0.052	0.769	0.096	0.018	0.025
20-99	0.007	0.011	0.069	0.818	0.047	0.048
100-499	0.008	0.008	0.012	0.059	0.760	0.153
≥ 500	0.022	0.001	0.003	0.005	0.080	0.890

Table 2: Poisson Regression Model of Firm Location Choice

	Poisson Model	
	Coefficients	t-statistics
Constant	4.948	105.18
ln(area)	0.482	61.86
Rural*	-3.025	-73.65
Suburban*	-1.464	-47.82
Urban*	-0.778	-30.31
Distance to CBD	-	-
Median HH Income	1.09E-06	4.93
Job Access	0.011	40.35
Population Access	-0.010	-36.77
log-likelihood (null)	-23,817	
log-likelihood (model)	-12,319	
Pseudo r^2 (McFadden's LRI)	0.483	
No. of observations	1,074	

* CBD is base area type.

Table 3: Negative Binomial Model of Commercial Trip Generation by TAZ

	Coefficients	t-statistics	Elasticity
Constant	0.921	6.02	-
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
No. of Firms	0.006	3.42	0.018
Alpha (Overdisp. param.)	3.161	16.32	-
log-likelihood (null)	-4,945.78		
log-likelihood (model)	-2,143.50		
Pseudo r^2 (McFadden's LRI)	0.567		
No. of Observations	1,074		

Note: CBD is the base area type. Elasticities for indicator variables are based on a discrete change in x , from 0 to 1.

Table 4: 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 HH Income	-6.60E-06	-5.91
No. of Firms	0.0005	2.05
Distance to CBD	0.007	2.7
Peak Drive Time	-0.075	-51.44
log-likelihood (null)	-11,441	
log-likelihood (model)	-9,377	
Pseudo r^2 (McFadden's LRI)	0.180	
No. of Observations	2,551	