

ABSTRACT

Title of Document: AN AGENT-BASED MODEL TO EXAMINE
HOUSING PRICE, HOUSEHOLD LOCATION
CHOICE, AND COMMUTING TIMES IN
KNOX COUNTY, TENNESSEE

Matthew C. Mulbrandon, Master of Science,
2007

Directed By: Professor, Kelly Clifton, Department of Civil
and Environmental Engineering and Department
of Urban Planning

The research conducted for this thesis uses an agent-based model (ABM) to simulate housing price, location, and *journey to work* (JTW) times for households in Knox County, Tennessee. The model is a unique hybrid, combining analytic functions and agents that typically have been used separately for theoretical urban research in very simplified urban landscapes. At the same time it uses data from a real urban area to run and calibrate the model, which is common for statistically-based or gravity models. There are two goals for this simulation; first to examine the feasibility of this approach in urban modeling, second to test the effect of altering transportation times and preferences on agent behavior. Results show this approach can fit real data and represent urban processes reasonably well. In addition several interesting and surprising results are reported from model runs.

**An Agent-Based Model to Examine Housing Price, Household Location Choice,
and Commuting Times in Knox County, Tennessee**

By

Matthew C. Mulbrandon

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Masters of Science
2007

Advisory Committee:
Professor Kelly Clifton, Chair
Professor Paul Schonfeld
Professor Gang-Len Chang

© Copyright by
Matthew C. Mulbrandon
2007

Acknowledgements

I would like to thank the many people who supported and assisted me in conducting my research for this thesis, my family including my sister and mother, my advisor Kelly Clifton and committee members G.L. Chang and Paul Schonfeld. Also I would like to thank my fellow graduate students.

Table of Contents

CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	4
2.1 MODELING IN TRANSPORTATION AND RESIDENTIAL LOCATION.....	4
2.2 PREVIOUS WORK USING AGENT-BASED URBAN/TRANSPORTATION MODELS	6
2.3 URBAN STRUCTURE AND PRICE GRADIENT.....	8
2.4 RESIDENTIAL SEGREGATION	9
2.5 PREFERENCE FOR THE BUILT ENVIRONMENT	10
2.6 COMMUTING (JOURNEY TO WORK).....	12
CHAPTER 3: METHODOLOGY	15
3.1 CONCEPTUAL MODEL	16
3.2 IMPLEMENTATION MODEL	23
<i>Subsection 3.2.1 Introduction of the Implementation Model</i>	<i>23</i>
<i>Subsection 3.2.2 Initial Conditions of the Model.....</i>	<i>25</i>
<i>Subsection 3.2.3 Location Choice Model.....</i>	<i>26</i>
<i>Subsection 3.2.4 Implementation Tools</i>	<i>32</i>
CHAPTER 4: DATA.....	34
4.1 KNOX COUNTY	34
4.2 DYNASMART	38
4.3 CREATING INDIVIDUAL HOUSEHOLDS FROM PUMS AND CTPP DATA.....	39
CHAPTER 5: CALIBRATION RESULTS	42
5.1 POPULATION AND HOUSEHOLD FIT OF SIMULATION RESULTS	43
5.2 INCOME FIT OF THE SIMULATION	45
5.3 PRICE PER SQUARE FOOT COMPARISONS.....	46
5.4 RACIAL SEGREGATION COMPARISONS.....	47
5.5 JOURNEY TO WORK COMPARISONS.....	51
5.6 HOUSING MODEL VARIABILITY DUE TO RANDOM VARIATION	54
CHAPTER 6: SCENARIO RUNS	57
6.1 SCENARIO RUNS FOR JTW VARIABLES	57
6.2 REMOVAL OF TRAVEL PREFERENCE.....	57
<i>Subsection 6.2.1 Preference change affect on price per square foot.....</i>	<i>58</i>
<i>Subsection 6.2.2 Average Income Results for no travel preference Run</i>	<i>59</i>
<i>Subsection 6.2.3 JTW Times No Travel Preference.....</i>	<i>61</i>
6.3 ONLY HOUSING AND SPACE PREFERENCE SCENARIO.	63
<i>Subsection 6.3.1 Price per Square Foot - Travel and Housing Space Preference</i>	<i>63</i>
<i>Subsection 6.3.2 Average Income - Only Travel and Space Preference</i>	<i>66</i>
<i>Subsection 6.3.3 JTW Time Results - Only JTW Time and Housing Space Preferences.</i>	<i>68</i>
6.4 INCREASED TRAFFIC 50 PERCENT AND DOUBLE TRAVEL TIME RUNS	70
<i>Subsection 6.4.1 Price per Square foot 50% More Traffic and Double Travel Times.....</i>	<i>70</i>
<i>Subsection 6.4.2 Average Income and Population Change for Increased Travel Time Scenarios</i>	<i>71</i>
<i>Subsection 6.4.3 JTW Times for Increase Traffic Scenarios.....</i>	<i>75</i>
6.5 INCREASED TRAVEL TIMES COMBINED WITH ONLY TRAVEL AND SPACE PREFERENCE.....	78
<i>Subsection 6.5.1 Price per Square Foot Change.....</i>	<i>78</i>
<i>Subsection 6.5.2 Average Income Increased Traffic and Only Travel and Space Preference.....</i>	<i>79</i>
<i>Subsection 6.5.3 JTW times increased traffic scenarios with only travel and space preference .</i>	<i>79</i>
6.6 MEASURING THE EFFECT OF CHANGING TRAVEL PREFERENCE	81
CHAPTER 7: SUMMARY	83

Bibliography	86
---------------------------	-----------

List of Tables

TABLE 4.1-1 POPULATION BY RACE (1990 AND 2000).....	36
TABLE 4.1-2 AVERAGE INCOME (1999) AND (1989) IN 1999 DOLLARS	37
TABLE 4.3-1 VARIABLE TYPES AND RESOLUTION.....	41
TABLE 6.2-1 COMPARISON BASELINE VS. NO-TRAVEL PREFERENCE RUN (CHANGE)	58
TABLE 6.3-1 COMPARISON TRAVEL SPACE PREFERENCE ONLY RUNS WITH THE BASE-LINE.....	64
TABLE 6.4-1 PRICE PER SQ FOOT FOR INCREASED TRAFFIC RUNS	70
TABLE 6.4-2 AVERAGE INCOME RESULTS FOR INCREASED TRAFFIC RUN.....	71
TABLE 6.4-3 JTW TRAVEL TIME COMPARISON WITH DIFFERENT TRAFFIC LEVELS.....	76
TABLE 6.5-1 SQUARE FOOT COMPARISONS: TRAVEL AND SPACE PREFERENCES WITH INCREASED TRAVEL TIMES.....	78
TABLE 6.5-2 AVERAGE INCOME COMPARISONS: TRAVEL AND SPACE PREFERENCE WITH INCREASED TRAVEL TIMES	79
TABLE 6.5-3 JTW TIME COMPARISONS: TRAVEL AND SPACE PREFERENCES WITH INCREASED TRAVEL TIMES.....	80
TABLE 6.6-1 JTW TIMES AND TRAVEL PREFERENCE CHANGE	82

List of Figures

FIGURE 3.1-1 DIAGRAMS OF INTERACTION OF HOUSEHOLD LOCATION CHOICE.....	18
FIGURE 3.1-2 ILLUSTRATION OF CONCEPTUAL MODEL.....	21
FIGURE 4.1-1 KNOXVILLE MSA 1990.....	36
FIGURE 4.2-1 DYNASMART NETWORK TRAFFIC FLOWS	39
FIGURE 5.1-1 RATIO HOUSEHOLDS: SIMULATED VS. REAL DATA BY NUMBERS OF HOUSEHOLDS FOR INDIVIDUAL TAZS	43
FIGURE 5.1-2 RATIO POPULATION: SIMULATED VS. REAL DATA BY NUMBERS OF HOUSEHOLDS FOR INDIVIDUAL TAZS	44
FIGURE 5.1-3 RATIO OF POPULATION: BASELINE RUN VS. REAL DATA.....	45
FIGURE 5.2-1 RATIO OF INCOME: BASELINE RUN VS. REAL DATA	46
FIGURE 5.3-1 PRICE PER SQUARE FOOT BASELINE RUN.....	47
FIGURE 5.4-1 A COMPARISON OF PERCENT BLACK IN TAZS: REAL VS. BASELINE RUN	48
FIGURE 5.4-2 PERCENT BLACK KNOX COUNTY 1990.....	49
FIGURE 5.4-3 PERCENT BLACK BASELINE RUN	49
FIGURE 5.4-4 PERCENT BLACK WITH NO BLACK PREFERENCE INCLUDED FOR HOUSEHOLD AGENTS.....	50
FIGURE 5.5-1 RATIO OF BASELINE VS. REAL JTW TIMES (ALL MODES) BY NUMBER OF HOUSEHOLDS PER TAZ	51
FIGURE 5.5-2 RATIO JTW TIME: BASELINE VS. REAL (ALL MODES)	52
FIGURE 5.5-3 RATIO JTW TIME: BASELINE VS. REAL (DRIVE ALONE).....	53
FIGURE 5.6-1 ROOT MEAN SQUARE ERROR FOR AVERAGE INCOME BETWEEN BASELINE SIMULATION RUNS	55
FIGURE 5.6-2 ROOT MEAN SQUARE ERROR FOR PRICE PER SQUARE FOOT BETWEEN BASELINE SIMULATION RUNS	56
FIGURE 6.2-1 RATIO PRICE PER SQ FOOT: NO-TRAVEL PREFERENCE VS. BASELINE	59
FIGURE 6.2-2 RATIO AVERAGE INCOME: NO-TRAVEL PREFERENCE VS. BASELINE	60
FIGURE 6.2-3 RATIO POPULATION: NO-TRAVEL PREFERENCE VS. BASELINE.....	61
FIGURE 6.2-4 RATIO JTW TIME: NO-TRAVEL PREFERENCE VS. BASELINE.....	62
FIGURE 6.3-1 PRICE PER SQ FOOT ONLY TRAVEL AND SPACE PREFERENCE.....	65
FIGURE 6.3-2 RATIO PRICE PER SQ FOOT: TRAVEL AND SPACE PREFERENCE VS. BASELINE RUN	66
FIGURE 6.3-3 AVERAGE INCOME ONLY TRAVEL AND SPACE PREFERENCE.....	67
FIGURE 6.3-4 RATIO PEOPLE PER TAZ: ONLY TRAVEL AND SPACE PREFERENCE VS. BASELINE	68
FIGURE 6.3-5 RATIO JTW TIME: TRAVEL AND SPACE PREFERENCE VS. BASELINE.....	69
FIGURE 6.4-1 RATIO AVERAGE INCOME: 150%*TRAFFIC VS. BASELINE	72
FIGURE 6.4-2 RATIO AVERAGE INCOME: 2*TRAVEL TIME VS. BASELINE	73
FIGURE 6.4-3 RATIO POPULATION: 150%*TRAFFIC VS. BASELINE.....	74
FIGURE 6.4-4 RATIO POPULATION: 2*TRAVEL TIME VS. BASELINE RUN	75
FIGURE 6.4-5 RATIO JTW: 2*TRAVEL TIME VS. BASELINE	77
FIGURE 6.5-1 RATIO JTW TIME: TIME AND SPACE PREFERENCE (2*TRAVEL TIME VS. BASELINE)	80

Chapter 1: Introduction

One of the most studied and debated topics in urban transportation is the link between transportation and urban land use. This research will examine a piece of the land use transportation link in a mid size urban area, Knox County, Tennessee. Specifically the research for this thesis will model the housing price, location choice, and the *Journey to Work* (JTW) at the individual and household level. Two *agent-based models* (ABMs) are used in the research. One, for modeling housing price and location choice, was developed for this thesis. The other, pre-existing agent-based software called *Dynasmart*, is used to estimate transportation times utilizing driver agents on a road network.

The research uses *agent-based models* in order to replicate processes that are occurring at the individual or household scale not possible with other methods. Agents are a form of artificial life and are defined as entities that are separate from their surrounding environment, can move through space and time, have goals, and can evaluate alternatives and take action based on those goals (Torrens 2001). Agent actions can be governed by anything a modeler can program into a computer. The simulations will model these urban processes at the micro-scale treating each person, household, and driver as an agent. Ideally agents will react to change in a given model run in a similar manner as in real life. This will enable a bottom up approach to modeling, creating aggregated results from many individual decisions. Many simplifying assumptions are made although like any model the goal is to retain important processes so the model will be a useful experimental platform.

Essentially accessibility will be determined from road network speeds given utilization rates using *Dynasmart*. These utilization rates are exogenously given from empirical data and not determined within the run although they can be altered prior to model runs. Only working households are modeled and will use this travel time information along with other variables to bid for square feet in urban neighborhoods. These neighborhoods are defined by Traffic Analysis Zones (TAZs) that have approximately 2000 people each. Preferences are not a known commodity therefore

they are assigned to agents based on the modelers best interpretation of present knowledge but are adjusted in order to fit real data during model runs. Importantly, the preference values will not be identical for every agent.

The households will determine, using individual preference and budget constraints, how much housing space they will buy in each TAZ.. Based on their individual preferences the relative utility can be calculated and a move decision made. As the household agents bid the price changes in each TAZ resulting in feedback and complexity so a solution cannot be calculated by pure analytic methods. The global solution will be attained numerically by having households move and bid iteratively in a simulation run until sufficient occupancy rates are reached in each TAZ. The sum total of agent actions in a model run creates a sort of equilibrium or steady state. As every household adjusts to model parameters the final results represents a long-term change. While people do not make decisions in as a precise manner as an analytic equation in housing they are hypothesized to use some sort of heuristic to make a reasonably beneficial selection.

The model has similar properties as the analytical equation models of Alonso (1964) or Greenshields (1956) but also allows for much more complicated and realistic interactions that exist in *complex systems* (Forrester 1969; Shelling 1971). In the model several variables like demographics are held constant, even though in real life they will change, in order to isolate transportation effects. This will be explained further in Chapter 3.

The empirical goal of this research is to demonstrate the direction and magnitude of change in housing location caused by altering transportation travel times and/or preferences in the *Journey to Work*. Previous agent housing models focused on simplified synthetic landscapes or statistical agent drivers that could not account for feedback effects adequately. The second goal is to demonstrate the ability of present agent methods and computational abilities to simulate complex urban processes using real life data. This model is the first to model housing location and price at the

household level using nonstatistical agents. This should help lay the ground-work for future research on similar topics using agent-based micro simulation.

The rest of the thesis will be organized as follows. Chapter two is a review of literature on the present knowledge of land use, transportation, location choice, and agent-based methods. In Chapter three the conceptual and implementation model are explained. Chapter four presents a background to the study area and the data used in this model. In Chapter five results of model calibrations for Knox County are presented. In Chapter six transportation scenarios are run and analyzed. A summary of results and conclusions are given in Chapter seven as well as a discussion of future research.

Chapter 2: Literature Review

This chapter will examine work relevant to this research question. It will describe what methods and research have been completed on residential location and urban models. It will provide an assessment of the strengths and weaknesses of the previous modeling approaches.

2.1 Modeling in transportation and residential location

Both residential location choice and transportation have been conceptualized as a system. Modeling of human behavior as a system took hold in the 1950s and 60s (Benenson and Torrens 2004). Systems modeling examines not just the object itself but also the interactions between different objects and how they relate to one another.

Some early examples of urban models from the 60s include spatial interaction such as the Lowery, and Forester models (Forrester 1969). These models depended on interactions based on gravity models or entropy to forecast future events. While they could be calibrated to match real data their theoretical and predictive qualities were questioned, most notably by Lee 1973. After this period this modeling practice continued to be exported to urban areas around the world (Batty 1989).

Ultimately urban models during this period, until recently, proved to have several difficulties: (from Torrens <http://www.geosimulation.org/geosim/>)

1. Centralized in nature: often, it is assumed that all activity in the city revolves around the downtown.
2. Relatively static: time moves in "snapshots", sometimes several years.
3. Rest on an unsteady theoretical footing: they commonly contain very limiting assumptions.

4. Highly aggregate: model developers often break a city into a few hundred units in a model.

5. (Unnecessarily) complicated: the inner workings of models are not easily conveyed to users.

More recently the advancement of desktop computing power has sparked an increased use of systems modeling in many social science disciplines including economics, sociology, and political science (Benenson 2004; Benenson and Torrens 2004). Complexity theory and the computational power to use it have given a systems approach new life or replaced it depending on one's interpretation (Benenson and Torrens 2004). Complexity is hard to define. Several factors that contribute to complexity in urban modeling include: feedback, non-linear relationships, and an explosive number of permutations (Benenson 2004). Additionally sensitivity test on all permutations is computationally impractical.

Agent-based modeling has reduced the need for simple aggregate measures by allowing the individual (the agent) to be explicitly considered within the model. Artificial life methods developed in computer science and biology have opened new avenues of research exploration. Improvements and availability of GIS has enabled far more sophisticated spatial analyses than were available 30 years ago (Benenson 2004; Sheppard 2001; Torrens 2001). Thus models no longer have to be simplistic in terms of aggregation across geography, time, and demographic categories. Individual traits and how they interact to create larger patterns can be modeled. Assumptions of equilibrium, which may not necessarily apply to complex systems, no longer have to be used by researchers (Sheppard 2001).

2.2 Previous work using Agent-Based urban/transportation models

Most of the serious attempts at using agent-based simulations for residential location have been recent, due to memory and computational requirements. The Schelling model, the first agent-based residential location model demonstrated, using simple neighborhood geometry and assumptions, if there were only a slight difference in preference for the racial makeup of a neighborhood that neighborhoods would continually change their racial makeup (Schilling 1971). Work on examining residential location that has not used statistical methods has usually been theoretical using so called 'toy models' (Torrens 2002; Fosset 2006). Many of these have also examined residential segregation using only one decision variable, while statistical models using disaggregate techniques have tended to use a Multinomial Logit (MNL) technique or some variation, and have examined real urban areas using many different variables thought to be important (Waddell 1993).

One of the most comprehensive of the agent-based urban models is Urbansim (Waddell 2003). It uses the Multinomial Logit (MNL) approach to simulate housing choice, development, and price in a comprehensive urban model designed to forecast urban growth modeling households and developers (Waddell 2000). It has been implemented in several US cities. The residential location decision in the simulation uses several commonly accepted factors such as price and distance to work. Households move into square feet as the unit of allocation. It does not take into account schools or crime since they cannot be determined within the model (Waddell 2000). The coefficients for the model are determined empirically (using LIMDEP) and are simply input into the model (Waddell and Ulfarsson 2004).

Agent-based models have been increasingly used in urban modeling (Batty and Torrance 2001). Some examples include sprawl-sim in which agents and Cellular Automata used to model urban expansion including residential location (Torrens 2002). There is a mover and developer in the model along with basic geography of a few hundred cells. ABLOoM simulates household and firm locations showing that it

is possible to get macro effects, such as sprawl and clustering, from micro-level decisions (Otter 1999). The data used is generated, not real data, but it uses a combination of economics as well as decision rules to determine agent decisions. It allows for learning or adapting agents, a concept common in complexity literature. Other work has tried to examine von Thunen's Location Theory using simulation (Sasaki and Box 2003). These kinds of approaches or methodologies are under current development for a wide range of urban issues as well as ecological and social organizational research (Batty and Torrance 2001). With increases in digitization of urban databases and continued advances in IT technology this trend will likely continue (Torrens 2001).

Statistical models (non-simulations) have been attempted many times some with restrictive assumptions such as only examining one-worker households (Waddell 1993). Other residential location models look at two-worker households and residential location (Ben-Akiva and Bowman 1998; Abraham and Hunt 1997). Both studies attempt to link individual and household choice in a single framework. Because of restrictive assumptions and the complex nature of cities these models could not uncover causal relationships between its variables and residential location decisions.

Increasingly detailed traffic simulation models using small time steps on realistic road networks have been developed within the last 15 years. These models are not all agent-based and include DynaMIT (Ben-Akiva, Bierlaire, Koutsopoulos, and Mishalani 1998), TRANSIMS (Barrett et al. 1995), Dynasmart (Hu and Mahmassani 1995), Paramics (Quadstone 2000), and Corsim (Owen et al. 2000). Dynasmart, Transims, and Corsim are microscopic simulations meaning they model individual cars or other vehicles. Dynasmart and Transims are designed to model traffic at the urban scale building up from small scale interactions rather than simply model specific road links as with Corsim. Transims is the most detailed model simulating all journeys and destinations in an urban area at the individual level. Dynasmart and DynaMIT are designed to simulate intelligent transportation options. Information is

given to travelers with predetermined destinations, they then react to the information and the results are analyzed.

2.3 Urban Structure and Price Gradient

Urban structure is defined as the spatial distribution of activities or processes occurring throughout a metropolitan area including: housing, workplace location, entertainment, shopping, and just about any other urban activity. Some well-known examples of simple models of 'Urban Structure' include the concentric ring model (Burgess 1925), sector model (Hoyt 1939), and the multi-nuclei model (Harris and Ullam 1945). These models place urban functions in specific areas of the cities based on a process thought to govern urban location. In the case of concentric ring model an ecological view of invasion and succession was used to explain urban location. These models were based on theoretical but nonquantitative analysis and the models implemented as schematic drawings rather than computer simulations. Later using analytical equilibrium models researchers attempted to explain land price and the location of workers and jobs (Alonso 1961; Muth 1969). These models emphasize the role of jobs in residential bidding for space with the assumption of monocentric job locations. While none of these models fit observed data precisely they do provide a general framework to study cities.

There have been numerous studies examining the housing price gradient theorized to exist starting at the CBD and declining gradually toward the periphery (Heikkila et al. 1989; Anas et al. 1998; Waddell 1993). Depending on the area being studied researchers have found different results; Los Angeles showed a weak CBD gradient effects while other cities like Chicago show a stronger effect. Not surprisingly cities with a more established larger CDB have tended to show a more pronounced land price gradient than decentralized ones like Houston and Los Angeles (Giuliano and Small 1991).

2.4 Residential Segregation

There is no question that American cities are segregated by race and income (Denton and Massey 1988). The reasons for this are varied and several explanations have been hypothesized. People of one group may not wish to live next to people of another race or different income level (racism, snobbery creating hostile environments). People may simply prefer to live among people who share common attributes but not actively dislike other socio-economic groups (Clark 1986, Clark 1989). Market forces may cause income segregation (and race segregation because blacks and Hispanics tend to have lower incomes) as people with lower resources cannot afford, or will not make sacrifices in space, to live in high priced neighborhoods. Another theory is that people of a particular race may be steered into neighborhoods by real estate agents or other intermediaries (Gladston 2005). Lastly, other factors associated with race and income like school quality and crime may be the driving force in segregation (Clark 1989). If one group is not making a choice to live in a neighborhood because of race the housing price and JTW may be affected and influence results in any model that examines commuting.

Segregation by income may seem a natural result of a market oriented housing supply. However, many times actors and institutions have made efforts to affect segregation. This may include limiting non-auto transportation options, zoning out low cost housing, limiting numbers of non-family members in housing units raising cost, and limiting rentals in single family homes. Housing and transportation are only part of the urban build environment. According to Molotch (1976) land-based elites drive urban development through political means in order to expand the local economy thereby increasing their wealth. Other academics looking at local politics have introduced notions of control of discourse and language by elites in driving the nature of what urban growth should look like and who should control it (Harvey 1973). Thus even if a location choice is caused by so called 'market forces' non-market regulations or other political forces may reinforce segregation.

In the past three and a half decades, there has been a plethora of studies that have tried to determine the effect of race on segregation (Emerson et. al. 2001). There are two main camps on this subject. One that race (black and white) plays an important independent factor in determining segregation (Emerson 2001; Charles 2000; Denton and Massey 1988; Farley 1977; Massey and Denton 1993; Taeubur 1965; Zubrinsky and Bobo 1996). The other camp believes that, although preference surveys show that racial preferences are real, other factors correlated with the racial makeup of neighborhoods are the main force behind segregation (Clark 1986, 1989, 1991 and Frey 1979).

Typically survey evidence shows that whites prefer neighborhoods that are no more than 30 percent black although many find 50 percent acceptable. Blacks tend to prefer 50 percent black neighborhoods (Clark 1991). Other work indicates this is true, but to a lesser extent, for white attitudes toward Hispanics and Asians, although one study found no clear evidence of segregation in Los Angeles based on Hispanic and Asian percentages in neighborhoods (Emerson 2001).

Although the effect of race and income on segregation is disputed, the reality of racial and ethnic segregation is not (Denton and Massey 1993; Charles 2000). There are different ways to measure segregation levels. Two of the most popular methods are the dissimilarity and exposure indices. Both are susceptible to change due to the scale of measurement. This problem is known as the *modifiable aerial unit problem* (MAUP) (Openshaw 1978). There have been attempts to ameliorate this problem using kernel density or other methods (Reardon and O'Sullivan 2004).

2.5 Preference for the Built Environment

The concept of 'home' in American society has been described and analyzed by historians, architects, geographers, sociologists, and many others. The dominant ideal

of a middle class life style with the single-family home in a peaceful neighborhood has been documented by authors such as Davis (1992), Baxandall and Ewen (2000), and Andres (2001).

These works help explain housing surveys which consistently show a preference for single-family detached homes (Levine 2005). Low-density neighborhoods typically are preferred according to many studies of American adults (Myers and Gearin 2001). In several studies town homes were favored by 10 to 17 percent of respondents (Myers and Gearin 2001). When asked to make decisions involving values of new urbanism vs. traditional suburbia the majority of people will pick 'suburban values' of long commutes, more housing space, and lower density neighborhoods. For example 83 percent of consumers would prefer single-family detached homes with longer commutes compared with having smaller homes situated near transit hubs. However, half of consumers say they value walkability in their neighborhood and 33 percent support transit accessibility (Myers and Gearin 2001). Other studies showed similar patterns with Atlanta having 26 percent of residents favoring transit accessibility while 40 percent of Boston residents have this preference (Levine 2005). Older people tend to favor neighborhoods built along the precepts of New Urbanism more than people in their child-bearing years, as the baby boomers age this can be a factor in urban growth (Danielsen et. al. 1999). New Urbanism generally attempts to revive communities through public spaces, i.e. parks, sidewalks, transit accessibility, and window-shopping, themes associated with turn of the century American urbanism.

The trend in North American cities is continual decentralization of urban areas and all of its functions. This results in the decreasing importance of the traditional urban core both relatively and, in some cities, in absolute terms. Greater land use per capita and larger housing units are also national trends (Levine 2005). The focus of this thesis is on residential distribution, not housing development itself, but the two are linked..

2.6 Commuting (journey to work)

The commuting distance results from the home and work location, and the network connecting the two. This is important in terms of the effect on geographic location of home and work as well as where and what time congestion will occur. Commuting can be considered a sub-group of urban structure and commuting has been used to examine the validity of theoretical urban models (Hamilton and Roell 1982). Time of day is also a factor in commuting because people have to be at work during a particular time frame, which creates peak period traffic.

Several approaches have been used to determine the linkage between commuting and home locations. A statistical approach using hedonic price models in which commuting is one factor of many in the location choice examined the effect of being close to employment centers on apartment rents. Proximity to freeway intersections, schools, etc. has been examined statistically (Waddell and Hoch 1993). Typically it is found that being within a quarter of one mile has a large effect, either positive or negative, for these kinds of facilities.

Another tactic used to examine commuting is measuring "Excess Commuting" first undertaken by Hamilton and Roell (1982). This is the excess commuting over the amount necessary if every household were to minimize its journey to work for given housing options. Hamilton and Roell (1982) ran two experiments one assuming that commuters minimized commute, subject to a sum of land rents determined by Muth's 1969 formulation. Hamilton finds an average minimized commute in twelve American cities of 1.1 miles vs. 8.7 miles (actual). Second a completely random assignment of household location and jobs gives a commute of 12.09 miles. These findings have been disputed by further studies. Cropper and Gordon (1991) find minimum commutes in Baltimore should be about 5 miles vs. the 10.2 miles actual commute when neighborhood characteristics are added. White (1988) finds that on average 11 percent of commute time is wasteful and there is less wasteful commuting

in smaller cities on average. Other studies including Hamilton (1989) and Kim (1995) both found large amounts of wasteful commuting.

It is difficult to gauge the degree of importance of commuting from preference surveys (Johnson and Nelson 1991; Timmermans 1994). Utility functions are particularly difficult to discern. For instance people may be making joint decisions by assuming suburbia equals good schools and low crime even if the questions are parsed (Timmersman 1994). Additionally as variables like commuting time are drastically increased they can become more important in location decisions. For instance less than 8 percent of commuters took longer than 60 minutes according to census data in 2003. Few Americans travel more than 90 minutes, 2.5 percent in 2003, but this is up significantly from 1990 (AHS 2003). These so-called extreme commuters are the fastest growing segment of commuter from 1990 (Census 2000). The ten counties with the largest numbers of extreme commuters are located in large metro areas of New York, San Francisco, Washington, and Chicago in which extreme commuters comprise 4 to 10 percent of all commutes in those ten counties. Thus, at a high level, increasing congestion could have a non-linear effect on housing or work locations as people adjust to new realities.

It is commonly assumed that commuting time is a disutility. One survey of San Francisco found 7 percent of commuters actually wanted longer commute times with about 40 percent saying their commute time was close to ideal although some of this may have been rationalization (Mokhtarian and Salomon 1999). Many commuters may in fact find commutes to be too short. However, from the data above it is clear that this preference for longer commutes will only result in a few very long commutes particularly in small urban areas. Additionally this preference for longer commutes may be a weak desire although no one knows the precise nature of commuter preferences.

Work on bridging the gap between transportation and housing has either been statistically based, such as UrbanSim, or relied on purely synthetic data. The statistical approach relies on measuring relationships between aggregated variables the underlying behavioral drivers remain unknown. While more theoretical or toy models are difficult to validate or apply to real situations and often use only one or two variables to determine location behavior. Network transportation models like Corsim and have been able to be both predictive and theoretical although certainly they are not perfect. The approach in this research will combine the progress in network models with a mechanistic agent model for housing choice.

Chapter 3: Methodology

In order to determine the effect of job accessibility on household location and housing price an *agent-based*¹ urban model has been developed for this thesis. The model is used to examine the interaction of residential location, housing price, and access time to work for households with at least one worker. Both the housing and transportation model use the individual behavior of agents assuming maximizing utility by minimizing travel time in order to model aggregate results in the Knox County transportation system. Both the move and housing bid decision are calculated using analytic utility functions for household agents in the housing model. The analysis is at the household level for location decisions and at the individual level for network times.

The commuting time is determined using a shortest path algorithm for individual drivers in Dynasmart. The road network represented in Dynasmart uses individual cars at the link level to determine initial travel times. All work trips are assumed to be done using a private automobile. Times can be change by adding or subtracting automobiles but there is not a *travel demand model* so the traffic changes must be an exogenous input. Congestion levels are fixed for a given time period but can be altered if desired. So there is not a direct feedback in the traffic model and move decision since there is not a travel demand model in the simulation, although as previously stated congestion levels can be altered through user input.

Housing supply that is rented by sq feet and demographic variables are pre-determined for Traffic Analysis Zones (TAZs). The TAZs represent neighborhoods and provide neighborhood boundaries, although neighboring TAZs are grouped in the simulation. In the case of commuting time they are grouped into what is called *super TAZs* in Dynasmart, these consist of 2 to 4 TAZs that are adjacent and have similar demographics. In the model runs traffic will be increased in order to examine

¹ Agent-Based is defined as a simulation which uses Agents that interacted with each other and their environment (in this case urban households) to make decisions based on probability, math functions, or rules.

congestion effects on location decisions. As with all models there are deviations in the model set up, parameters, and data from real life. The method is designed so that model behavior will accurately reflect the direction and scale of effects for given scenarios. It is not a predictive model.

The goal is to build a model that can simulate the effect of changing preference for neighborhood attributes and real travel times on housing price, household location, and JTW times in Knox County. The model is set up so that the city and its residents interact in a market by bidding for housing space in TAZs. The model assumes location and evaluations are rational economic behavior and do not require moving cost, intervening institutions or other factors that would alter economic adjustments. This allows the model to represent what residents ‘would’ do in the long-term in response to change in simulation values. As we observe in real life, housing and commuting is not as simple as the model proposed. Factors that may alter the move decisions such as moving cost and imperfect information, like steering from real-estate agents, are not explicitly modeled. It is assumed that despite non-market forces people are trying to better their circumstance based on utility measures. With this view, model results can be thought of as an upper bound on real life decisions, meaning results of different scenario runs are unlikely to produce a greater real life change than model results over a few years assuming the assumptions in the model are valid. Additionally, as this is the first time a mathematical agent model has been used to simulate residential location in an urban area this research will allow the examination of the relative strengths and weakness of an analytically driven agent approach to modeling complex urban processes. This will assist future developments in urban modeling.

3.1 Conceptual Model

It is assumed that households have perfect information and will choose their location in a rational, economic decision process and bid for housing by some mechanism that approximates a neoclassical market. Households have different preferences for goods

from one another. Also they have limited resources to rent housing. Household-agents will attempt to maximize utility out of all possible choices. Households will ‘rent’ space by the sq-foot, where the total number of square feet is fixed for each TAZ, not individual housing units. Also new developments will not be permitted (although there is the ability to add it). All household workers will commute by car during congested time periods the same number of days a week. Commuter-agents will choose a least-cost or shortest path to and from their work location. The sum-total of these decisions along with the existing road network create the time between TAZs that we observe.

Only residential location, housing price, and commuting time are determined within this model. There are several dynamic processes that occur within the model system, each at different geographic and time scales. Housing location assignment and price represent long-term change as stated previously. For each simulation run many different actions are made as agents bid and move to square feet within each TAZ. These can be thought of as virtual moves meaning that only the final result represents actual moves while the intermediate effects on price during the simulation are ‘negotiations’ of household agents representing what they would do under given conditions. Each agent’s bid affects the housing price which in turn affects how many and which households will move into each TAZ.

The traffic flow on the network is modeled in 5-minute time steps. The individual cars are modeled with a predetermined *origin and destination* (O-D). The scale covers primary and secondary roads in Knox County. To account for variability in the model the average of 10 Dynasmart runs is used to calculate the travel times between TAZs . These stay constant during a housing simulation run.

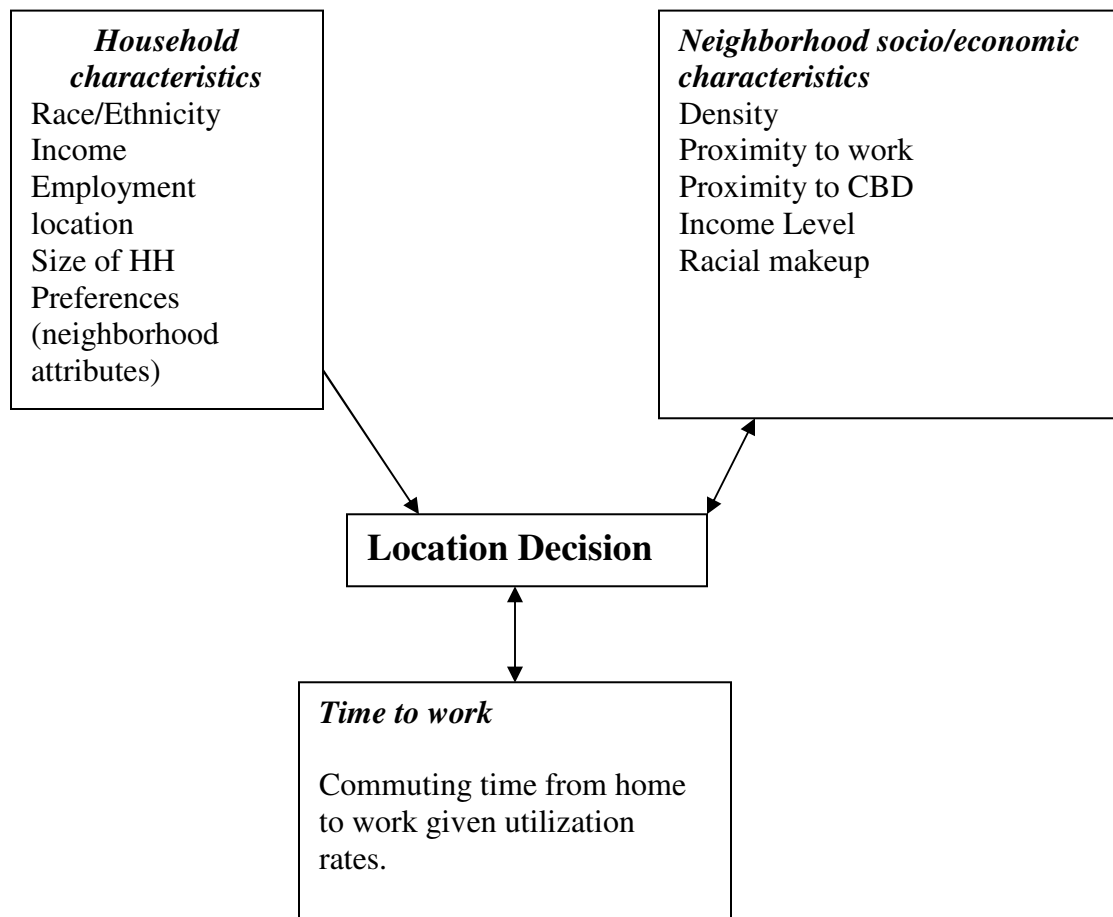
Given these assumptions there are three main interrelated elements feeding into a location decision. These include:

1. Household Characteristics

2. Neighborhood Characteristics

3. Time to Work

Figure 3.1-1 Diagrams of Interaction of Household Location Choice



Households are autonomous agents that act to maximize perceived preferences. They move until equilibrium is reached. The city is conceptualized as a complex system² with three direct interactions:

1. Residential location decisions
2. Housing Price (per square foot)
3. The network route by commuters given utilization by other drivers (trips are an exogenous input in the implementation model)

Decisions made by household agents affect one another resulting in feedback (Fig 3.1-1). Each of these boxes can be subdivided. There are three main elements that form a household characteristic; one, geographic/economic variables, which includes the employment neighborhood of people in a household and the location of the household; two, demographic variables including household size, head of household, household income, and number of children; three, the 'preference' of a household for neighborhood characteristics (Fig 3.1-1). Some preferences are influenced by the demographics, i.e. the race preference is affected by the race of the family in addition to individual variation.

The neighborhood characteristics have several components as well (Fig 3.1-2). There is a demographic component, meaning the socio-economic characteristics of people living in a given neighborhood. A neighborhood's real-estate components include total space, as well as a price per sq-foot. The price per sq-foot is determined by a bidding process that households interested in moving will use to determine their location. Also there is a geographic location that each neighborhood possesses resulting in proximity or accessibility to city amenities and employment.

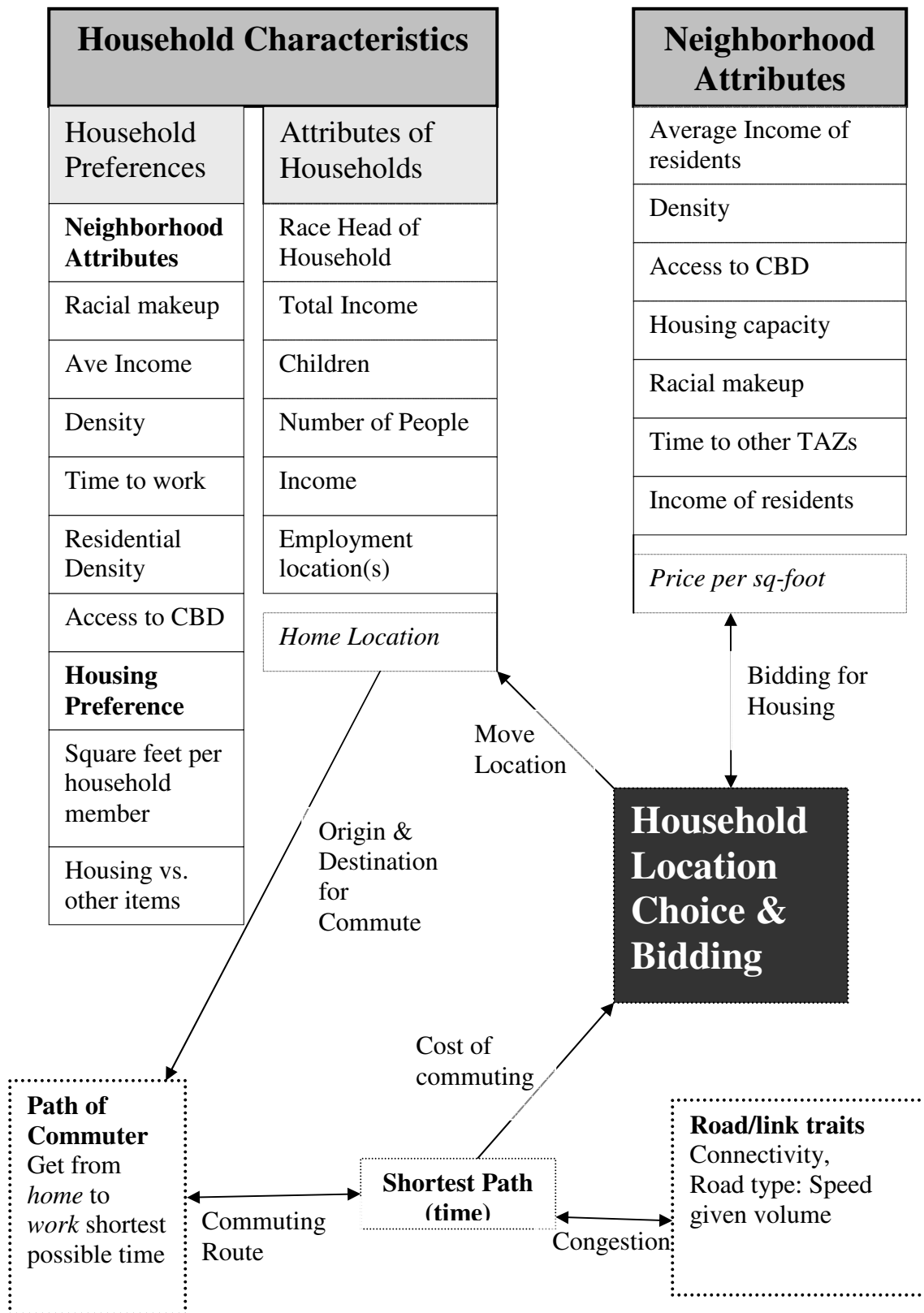
² Complex system for the purposes of this proposal is defined as a system in which simple rules create more complex structures through interaction of objects in the system.

The last of the three main elements is travel time to work (Fig 3.1-2). This is affected by the network distance and speed between home and work (note: in reality there are multiple destinations on many commuting tours). Assuming we are dealing with automobiles and roads the network speed results from the Level of Service (LOS) given utilization and the utilization rate. The utilization rate depends on the shortest path for commuter agents between origin and destinations in a given time frame. Given perfect information, travelers will change their path in order to achieve greater speed eventually reaching user-equilibrium.

A change in transportation time or preferences for commuting time may have far reaching effects for neighborhood demographics and housing price. A main effect that improved access to employment centers would have on a neighborhood is a higher price per sq-foot. This is because people will pay more to have a shorter commute all else being equal although very short trip times may have an actual disutility for a small percentage of commuters (Redmond and Mokhtarian 2001). With relative indifference between short trip times this should not be a large source of error in how households choose housing. The higher price would in theory change housing location decisions. That would result in a change in the demographic make up of a neighborhood creating an additional 'feedback' effect on choice of neighborhood (Fig. 3.1-2 below).

The housing price is 'demand' driven with the supply calculated from external factors (explained in detail below in section 3.2). The route choices and congestion levels are changed when residential choice changes, which can result in a different work-trip time for neighborhoods. The change in work-trip time then feeds back into residential location decisions because of the time to work preference. In order to model this particular interaction within the modeling framework a travel demand model is needed which is beyond the scope of this research. However traffic can and is changed exogenously in the simulation runs.

Figure 3.1-2 Illustration of Conceptual Model



In the conceptual model, several simplifying assumptions have been made.

Process simplifications

- Work location does not change.
- Income does not change.
- No change in attitudes or preferences over time.
- Search covers all neighborhoods.
- No real estate agents.
- Bidding process for all households in a neighborhood. Not between sellers and buyers of individual units.
- People and households stay constant, no ageing, death, moves out of household.
- Schools and Crime are not factored into the move decision (lack of data).
- There is no walking or transit use in the network.
- There is no negotiating among individuals within the household
- No future or adaptive planning by agents as they stay constant
- No social networks.
- Everyone drives alone.
- People have analytic utility functions for attributes and act on them maximizing utility.

Data simplification

- No developers building new housing, renovating, or building office parks.
- People only live in one place.
- No government subsidy.
- Work trips are pre determined.
- Cost in time only but no variance in travel time incorporated. No gas and car cost.
- Only rent.

- Housing is rented by the sq-foot not units.
- Neighborhood attributes are defined by TAZs or groups of TAZs.
- Homogeneity within neighborhoods including:
 - Housing quality.
 - Neighborhood's socio-economic level.

This conceptual model focuses on the interaction between three variables: housing price, location, and travel time. As stated previously there is not a travel demand model so a change in travel patterns has to be exogenously added into the software. In addition to the simplifying assumption listed above many social, political and economic forces are also not included in the model. The people, society, and city are simplified in order to create a framework that can be modeled on a computer. As with any model these kinds of assumptions are necessary.

3.2 Implementation Model

Subsection 3.2.1 Introduction of the Implementation Model

The model created for this research will simulate location choice, housing price, and the JTW. The model is a system meaning there can be feedback, non-linear effects, and many complex interactions between different variables.

The study site is Knox County, TN, which includes the city of Knoxville, and its road network. The reason this site is chosen as opposed to some other medium size location is a usable Dynasmart road network has been constructed and tested for this county and O-D traffic matrix constructed. Dynasmart is used to calculate travel times between TAZs.

Traffic can be added to Knox County by changing the O-D matrix externally. People and households will remain constant, meaning they will not age, divorce, or die to simplify the simulation. Crime and schools will not be a factor in move decisions as

there is a lack of reliable data on these values and an inability to predict new values simply based on demographic change. Because of correlations between income, schools, and crime levels the effect of income will undoubtedly be magnified within the model. The housing model alters residential location and housing price until a steady state is reached. This is defined as between 90 and 100 percent of available square feet rented in each TAZ for a given iteration of the model. At that point the run stops. The resulting outcome represents the ‘long-term’ price, residential location, and population change in a given area all else held constant. Decisions are made at the agent-level (either individual or household). Congestion levels can be changed by altering inputs to Dynasmart to test sensitivities of ‘Journey to Work’ (JTW) times. Additionally attitudes toward travel times can be altered separately or in combination with travel times.

The basis for move choice is modeled using preferences and is informed by empirical studies and/or preference surveys although it is difficult to ever find true preferences for individuals. Households are assigned different marginal utilities stochastically using a normal distribution around a mean. In some cases different types of households are created with different means depending on race or the presence of children. These are then put into a simulation in order to determine the value households put on a neighborhood. A detailed description of this method is provided later in this section (subsection 3.2.3).

Neighborhoods have limited housing stock and several characteristics that are important to consumers. Households will rent housing measured in square feet and there is no cost to moving. Housing supply is fixed for each TAZ in each model run. How the total housing square feet per TAZ, which is not provided by census data, is calculated will be explained below (subsection 3.2.3).

The commuting time can be modeled using agent-based user equilibrium or a one-shot assignment using Dynasmart. This information is then input into the urban location model using a C++ program. Both methods assume a perfect knowledge of

link speed and that people will adjust so that no one can increase travel time by changing links. Although this does not exist in reality, it is a reasonable assumption for commuting because people usually are aware of efficient ways to and from work. Because of computational and data limitations only automobiles are modeled. It is assumed all workers have a car. Because of variability in Dynasmart and real life, multiple runs were made for each travel scenario the results of which are averaged. Therefore, the cost of travel is purely the average commuting time between home and work. There is not a fixed dollar cost to travel rather travel time like other variables is evaluated using an individual utility function, with different values for each household, and can be sacrificed for an increase in other goods.

The O-D trip matrix depends on where people travel or ‘travel demand’. The O-D trip matrix is not determined within Dynasmart but comes from external inputs. Changes can be made manually or by simply multiplying present trips. As mentioned previously a sophisticated ‘travel demand model’ could be added but is beyond the scope and available data of this research. The model run produces link speed, which is used to find travel times between TAZs. That time is used to estimate commuting times of workers in the model.

Subsection 3.2.2 Initial Conditions of the Model

Neighborhoods are defined by TAZs; the initial conditions of each neighborhood are set to 1990 data. The household agents are created using *Public Use Micro Sample* (PUMS) and *Census Transportation Planning Package part 1* (CTPP 1) data (method explained in the data section 4.3). Additionally each ‘employed person’ is assigned to a job in a particular neighborhood. The job allocation match *Census Transportation Planning Package part 2* (CTPP 2) job data in the TAZs and the PUMS employment data for the individuals. The initial travel speed will be based on times to and from TAZs using Knox County travel data coming from Dynasmart. When this is accomplished, the location model will be run. This process will be iterated until a ‘stable’ result is derived. A stable result is defined as no TAZ having a vacancy rate

of more than 10 percent. Additionally TAZ results ideally will be within 50 percent and 200 percent of real census data. This run will serve as the baseline of comparison for different scenarios. Like many urban and transportation models because of the differences in data and process, achieving the same results will not validate the model or parameters. This is because of the concept of *equifinality* meaning that different parameters may produce the same results. That said making a baseline run does provide several benefits. It allows for checking of bugs or parameters specification that produce results or relationships between variables that are unrealistic. It makes scales of results more intuitive for the modeler and reader. It provides a known target for a baseline run. Although it does not validate a calibrated model it may provide additional realism to scenario runs.

Subsection 3.2.3 Location Choice Model

It is assumed there is homogeneity of neighborhood 'quality' within a given neighborhood. This is a necessary assumption because of data and computational limitations but is not always realistic. Except for travel time TAZs are used to define neighborhood boundaries; the travel time is measured in terms of “Super TAZs”. As previously stated these are groups of contiguous TAZs that have something in common. They are grouped in Dynasmart as a way of reducing computational work. Just as with the use of centroids, in more densely populated areas with geographically smaller TAZs, this would have only minor effect on travel times while the effect on mid-size and larger TAZs could be several minutes.

The functional form for the utility was chosen to be as realistic as possible while still allowing for an analytic solution for the optimal space for all households in each TAZ. Without such a solution the computation would be much longer and impossible to solve in a reasonable time limit. Household preferences are given by functional forms with parameters that determine the mean but have random deviations for each agent. The result is that no one household has the exact same preferences as another

one, but aggregately each preference has a known mean and distribution in the simulation.

The decision of which variables to include for neighborhood preferences is based on current knowledge but, as discussed previously in the literature review, the values of the preferences are not known with precision. The functional forms of the preference were set prior to simulation runs while the parameters used on the functions were adjusted by comparing run results with data.

The household agent **J** has an income **I_J**, minimum space requirement **q_j** (estimated to be about 140 sqft + 100 sqft per person in each household). This number is estimated from people per 'room' data in Knox county census data. A perceived utility **Φ** for each TAZ that accrues to household agents is independent of the amount of space occupied. Each neighborhood **n** has a price per square foot **p_n**. We construct a utility function of agent **J** for neighborhood **n** that is given by the product of his utility for space **S**, other goods **Y**, and **Φ** that represent the agent's preferences for neighborhood **n**.

$$U_{jn} = \Phi_{jn} S_j Y_{jn}^{\delta}$$

Here δ can be thought of as the relative strength of an agent's desire for goods vs. housing space. Since housing is a necessity, an agent is assumed to have zero utility until the minimum space requirements are met so **S** is written

$$S_j = Q - q_j$$

The utility for goods **Y** is just the utility of the money left over after the rent is paid for all other goods. Using the budget constraint **I_j = p_nQ + Y** the utility function for **Y** can be written

$$Y_{jn} = I_j - p_n Q$$

Substituting into the equation for U the utility agent j will have if he resides in a neighborhood n gives

$$U_{jn} = \Phi(Q - q_j)(I_j - p_n Q)^\delta$$

The amount of space Q maximizes the utility of agent j in a given neighborhood n can be calculated by the standard method of maximization of a function, (Φ has no dependence on Q).

$$Q_{jn}^{\max} = \frac{I_j + \delta p_n q_j}{p_n (1 + \delta)}$$

Therefore the fraction of an agent income spent in a neighborhood occupying the optimum space is $p_n Q_{jn}^{\max} / I_j$ and will depend only on the agent's tastes, income, and linearly on the price per square foot in the neighborhood. Since most people spend a fixed fraction of their income on housing (usually about 1/3) the neighborhood preference function Φ , can be viewed as a measure of how much space ΔQ an agent is willing to give up to live in a more expensive neighborhood or in money terms the price of the forgone space $p_n \Delta Q$. Substituting the value for the optimal space into the utility function we can also obtain the maximum utility that an agent j can obtain in the neighborhood n

$$U_{jn}^{\max} = \frac{\Phi_{jn} \delta^\delta}{p_n (1 + \delta)^{1+\delta}} (I_j + \delta p_n q_j)(I_j - p_n q_j)^\delta$$

This simplifies the problem of choosing where an agent will live since every neighborhood will have a single value for his utility, which will be determined by his income, preferences, and price per square foot in a neighborhood.

As stated previously in this section household agents are assigned different utilities by stochastically varying parameters with a normal distribution around a mean. The values of parameters used for the utility function are $\delta_j = N(3.5, .35)$ and the minimum space $q_j = (140 \text{ sqft} + 100 \text{ sqft per person}) N(1., 0.1)$

To model preferences, estimated functional dependencies and means (normalized to be near one) from studies are used. We adjusted values of the means and of the standard deviation of parameters to get the best fit for the existing populations when run to equilibrium.

The preference function Φ is assumed to be a product of preference functions for the following variables.

t_{nj} - the travel time from (Super) TAZ n to the work location of agent j

d_{nCBD} - the distance from the TAZ n to the CBD

M_n - the median income of TAZ n

P_n - the population density of TAZ combined with Super TAZ n

R_n - the racial composition of TAZ n

Many studies have shown that commuters value commuting time nonlinearly. We set t_{nj} to the 1.5 power. We also assume that there are some commuting times so long that an agent utility for that location is zero. This allows us to construct a utility function for a two-way trip with travel time t_{nj} from a TAZ n to the work location of agent j

$$\Phi_{jn}^{\text{traveltime}} = \left[2 - 2 \left(\frac{t_{nj}}{T_j} \right)^{1.5} \right]$$

The same form is used for the CBD attractor with a different constant for the distance (driving time) that is so large as to represent zero utility.

$$\Phi_{jn}^{\text{distanceCBD}} = \left[2 - 2 \left(\frac{d_{nCBD}}{D_j} \right)^{1.5} \right]$$

The functional dependence for the preference of agents to live in neighborhoods with higher median income is assumed to have a power law dependence on the median income divided by \$10,000.

$$\Phi_{nj}^{\text{income}} = [M_n / 10000]^\alpha$$

The preference for lower density is assumed to be a power of the logarithm of the inverse of people per square mile

$$\Phi_{jn}^{\text{density}} = [\log(100000/P_n)]^\beta$$

From previous surveys, although there was a broad range of preferences, whites tended to prefer majority white neighborhoods while blacks preferred more evenly split neighborhoods (Pettigrew 1973; Clark 1991). If they had to live in a segregated neighborhood both tended to choose to live in neighborhoods dominated by people of similar race or ethnicity. Only White and Black were used for race/ethnicity in the model although there were small numbers of other groups. The preference for both groups was based on percent black in the neighborhood. The preference people have for different neighborhoods with a given racial composition is represented by the linear difference between the fraction of black heads of households in a TAZ and the desired fraction r_j

We can now write the complete expression for preferences

$$\Phi_{nj} = [M_n / 10000]^\alpha [\log(100000/P_n)]^\beta [R_n - r_j] \left[2 - 2 \left(\frac{t_{nj}}{T_j} \right)^{1.5} \right] \left[2 - 2 \left(\frac{d_{nCBD}}{D_j} \right)^{1.5} \right]$$

As was done for the utility function, all parameters are multiplied by a normal random distribution to account for differences between people. The values chosen for these parameters for the preferences are

$$\alpha_j = N(0.63, 0.1)$$

$$\beta_j = N(0.2, 0.02)$$

$$r_j = N(.6, 0.3) \text{ for black and } N(0, 0.1) \text{ for white}$$

$$T_j = N(180, 18) \text{ min}$$

This model incorporated distance (driving time) to the CBD as a stand-in for city amenities. For this model 35 percent of singles and 15 percent of family households were assigned a mean that represented a strong preference to live near the CBD.

$D_j = N(120,24)min$ for those who prefer urban living

$D_j = N(360,72)min$ for everyone else

Program Flow:

Initialize: Estimate the available area measured in square feet for each TAZ A_n^{supply}

Guess at starting price p_n for each TAZ

Iterate: Calculate the max utility in all TAZ n for agent's j and then assign each agent to the TAZ that provides greatest utility. Sum the space Q_{jn} for all agents who pick TAZ n to get A_n^{demand} .

Then apply the rule if

If $A_n^{demand} > A_n^{supply}$ increase price p_n

If $A_n^{demand} < A_n^{supply}$ decrease price p_n

The price is not adjusted if the TAZ has between 0 to 10 percent vacancy rate. This makes the convergence algorithm more stable. The run is ended when all TAZs meet the criteria that

$$A_n^{supply} \geq A_n^{demand} > 0.9 A_n^{supply}$$

The final housing price is a result of households bidding for space in a TAZ and of the price/space trade off that working households make in this simulation.

The estimate of square feet in each TAZ is a guess, as census data does not provide this information, and tax records are not electronic in Knox County at this time. To improve our estimate for A_n^{supply} the available space in each TAZ is calculated (for the baseline conditions) using the prices p_n from a converged run and with the mean income, number of households, and people from the CTPP 1 file, assuming that all agents in TAZ behave like the mean. This improved estimate for the A_n^{supply} is used for all runs

Subsection 3.2.4 Implementation Tools

Dynasmart P

Dynasmart P will be used to calculate transportation time between neighborhoods. The agent-based one shot assignment function is used to calculate travel time between TAZ's over a two-hour rush hour. Dynasmart has several features that make it a good tool:

1. Not all traffic is generated from the centroids of neighborhoods instead links are used to generate origins of traffic. That will prevent artificially high congestion results.
2. In addition, it is possible to have different departure times within a given time period, i.e. not every one who leaves for work between 9:30 and 9:35 will leave at 9:30 am exactly increasing congestion in the simulation.

Handling data files in Dynasmart is difficult and it does not link well with GIS. The GIS map in Dynasmart can be exported to Trans-Cad with some difficulty.

A main feature that an agent-based approach offers is the ability to know 'who' is traveling 'where' and at 'what' time. This detailed knowledge is not attainable in any land use/transportation modeling simulation and could prove invaluable in the future when examining a myriad of urban issues from transportation to health.

SAS

SAS is used for data processing- PUMS, CTPP I and II; used to impute households using PUMS and neighborhood (TAZ's) data for initial conditions.

ArcGIS- This is used to find sq. miles of TAZs, display data, and calculate centroids of TAZs.

C++

This language is used to program the housing price and location simulation. Additionally it was used to transfer data from Dynasmart to the simulation in the proper format.

Chapter 4: Data

4.1 Knox County

The study area is Knox County, Tennessee. The data used for this model came from two main sources, census data, and data for use with the transportation software Dynasmart. The census data includes the CTPP I and CTPP II 1990. Additionally individual and household data comes from PUMS, 5 percent sample, which provides household and individual data for Knox County. The scale of analysis is TAZs, which are geographic areas akin to census tracts but about half the size. The reason for this is that in Dynasmart the O-D matrix was given at the scale of TAZs or Super TAZs. Boundary data came from shape files provided by the United States' Census. Data packaged with Dynasmart includes road network data and trips over a two-hour period. The extent of Dynasmart of the transportation network also includes part of Blount County to the south. This county was omitted as its inclusion created unrealistically long commuting times as most of the jobs in Blount are not covered by the road network in Dynasmart and could not be included. Aggregated demographic data was harnessed from CTPP I 1990 while work data came from CTPP II 1990. Although the data used is old the purpose of the model is to examine changes in location and JTW times when altering transportation or demographic variables and preferences not a prediction for a given date.

Knox County has a typical spatial structure for Eastern and Midwestern metro areas of the United States that includes a CBD with a large concentration of jobs with poorer minority residents living in the central city. It has not had the demographic influxes of emigration that many other large metro areas have experienced. Unemployment has been lower than the national average, 4.1 percent in 1990 and 2.5 percent in the year 2000. It is a service-based economy, the largest employer being the University of Tennessee main campus, which has 26,000 students and is centrally located in the City of Knoxville. Knox County is one of six counties in the MSA. It has more than 55 percent of the metro area's population and is not very different from other counties in the area with a slightly higher per-capita income than most

surrounding counties as well has a higher percentage of minorities (Table 4.1-1). The main urbanized areas extend into Blount County, not all of Knox County is urbanized (Fig 4.1-1). The largest employment center outside of Knox County is Oakridge to the northwest. Per capita income and educational attainment have typically been close to the national average and private automobile the predominant means of transportation. There are two main interstates, I-40 and I-75, crisscrossing the county with an additional bypass around the CBD. Black residents live in some of the poorest areas of the county particularly near the CBD of Knoxville. Overall the City of Knoxville is poorer and has more minorities than the county as a whole.

In 2004, 12.7 percent of residents lived in poverty in Knox County. In 1993 the poverty rate was 15.1 percent. Between 1993 and 2000, the poverty rate fell each year, reaching a low of 11.3 percent in 2000. Housing cost increases have not been as great as in the US as a whole but in most of the urban area there has been a marked increase in home values. About 22 percent of the housing in Knox County was built in the 1990s. About 65 percent of residents of Knox County own their homes but this figure is about 50 percent in the city of Knoxville. Median housing price was about \$100,000 in the year 2000. The higher rents were about \$1.5 a square foot in 2003. Whites and African Americans are the predominant ethnicities of Knox County and the metro area. There has been a marked increase of Asians and Hispanics in the 1990s in terms of percentage but the total numbers are still small (Table 4.1-2).

Figure 4.1-1Knoxville MSA 1990

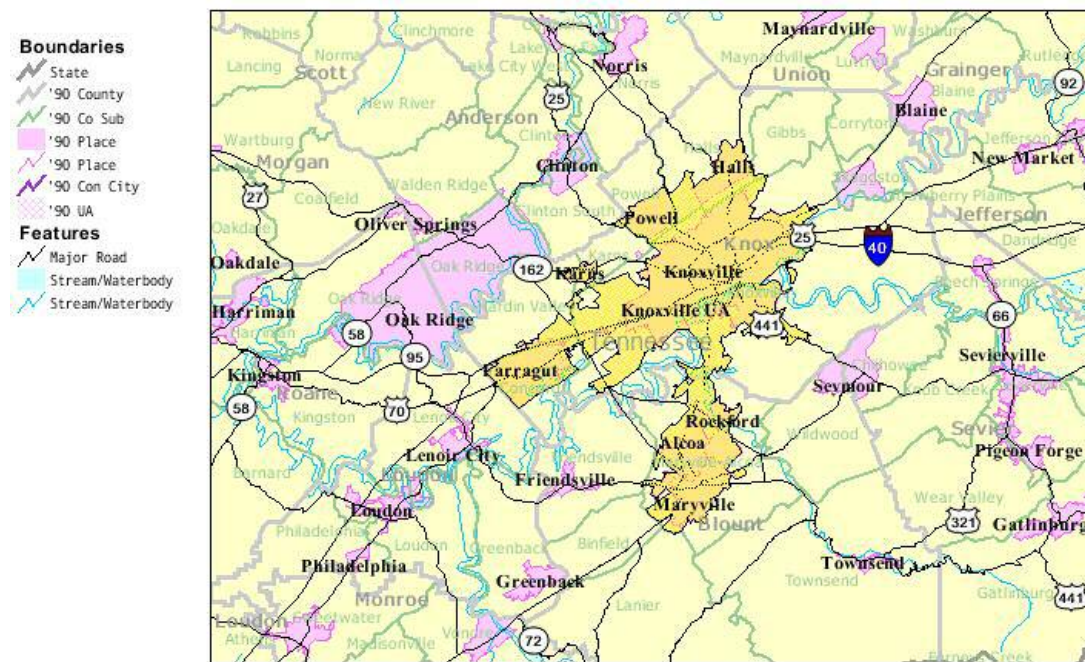


Table 4.1-1 POPULATION BY RACE (1990 and 2000)

Race	Knoxville MSA 2000	Knox County 2000	City of Knoxville 2000	Knox County 1990
Total	687,249	382,032	173,890	335,749
Population				
White	627,651	336,571	138,611	301,421
Black	39,691	32,987	28,171	29,603
Am. Indian	1,937	1,007	541	787
Asian	6,796	4,937	2,525	3,136
Other	3,426	1,902	1,257	601
*Hispanic	8,628	4,803	2,751	1,935

(Hispanics are an ethnicity and can be of any race.)

In 1990 there were 335,749 people in Knox county and 133,639 households (US Census STF 1 1990). This increased 14 percent to 382,032 people by the year 2000. The number of jobs increased 5,000 from 190,000 jobs in 1990 to 195,000 jobs in 2000.

Table 4.1-2 Average income (1999) and (1989) in 1999 dollars

Per Capita Income 1999		Per Capita Income 1989	
USA	\$21,587	USA	\$11,400
Tennessee	\$19,393	Tennessee	\$14,003
Knoxville MSA	\$20,538	Knoxville MSA	\$16,020
Knox County	\$21,875	Knox County	\$18,093
Anderson	\$19,009	Anderson	\$14,104
Blount	\$19,416	Blount	\$13,597
Loudon	\$21,061	Loudon	\$18,965
Sevier	\$18,064	Sevier	\$11,558
Union	\$13,375	Union	\$7,067

Source: US Census STF 3 1990 and 2000

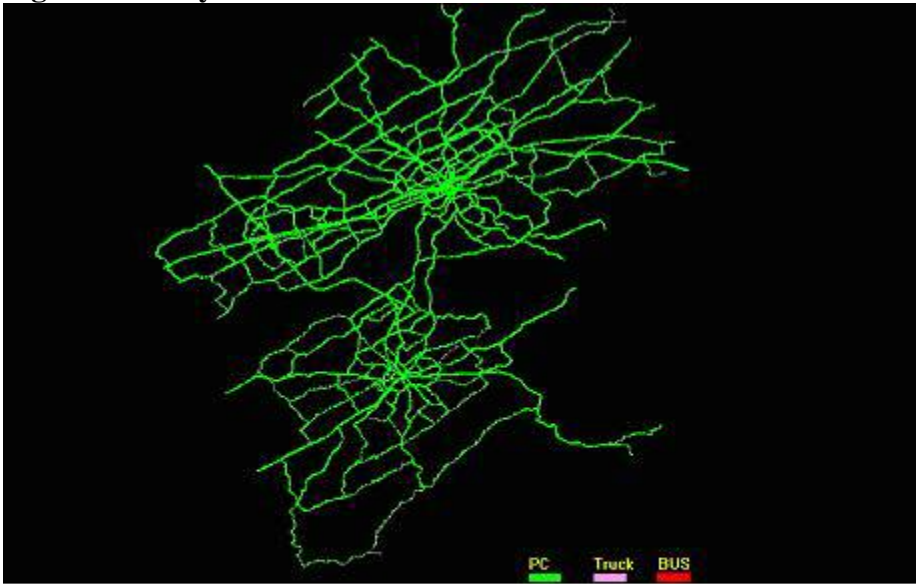
There was a significant increase in the numbers of people who drove alone to work from 77 percent in 1990 to 81 percent in 2000. Carpooling is the most frequent alternative to 'drive alone' with 8,300 people or 11 percent of residents carpooling to work in 2000, a reduction of 1,500 carpoolers from 1990. Public transit figures saw a 23% drop from 1990 rates with fewer than 1,200 Knoxville workers using buses, trolleys, and taxis as their primary means of transportation to work with about 5,000 people walking or working-at-home. There are 1.8 vehicles per household in Knox County, with a 1.5 vehicle average in Knoxville (Source: <http://www.knoxmpc.org/locldata/locdem02/locdem02.htm>).

4.2 Dynasmart

The Dynasmart traffic network includes freeways, primary and some secondary roads. It covers Knox county and the northern portion of Blount county (Fig 4.2-1). For this thesis only the Knox county transportation network is used. The purpose of Dynasmart is to test intelligent transportation systems by relaying information to driver agents and examine the changes in the transportation system. While that is beyond the scope of this paper what Dynasmart is used for is to calculate times resulting from utilization of the transportation network and resulting congestion levels. The network is typical of other transportation software using nodes, arcs, turn penalties and other common features. As stated earlier it is agent-based, which means individual automobiles are modeled through the transportation network resulting in congestion times. While there is not a travel-demand model, traffic can be increased in two ways, either by adjusting the multiplicative factor for the O-D matrix which will alter all numbers of trips by the same factor for all zones or by manually creating a O-D matrix one O-D pair at a time.

The number of vehicle trips over a 2-hour period is approximately 300,000 and represents a peak period travel time. This was the basis of travel times between zones and does not change during the simulation. It can include cars going in and out of Knox County but only travel times between Knox County TAZs are output. Dynasmart provides a plethora of information including a travel time for vehicles between origins and destinations. Because not all TAZs have a trip between each other some TAZ times are imputed. However, this causes minimal disruption as TAZs with large numbers of jobs invariably have trips between one another. The simulation was run several times so an average of O-D times is used to get an even time distribution. The O-D matrix is estimated from real data in the 1990s.

Figure 4.2-1 Dynasmart network traffic flows



4.3 Creating Individual Households from PUMS and CTPP data

To create households with job location, data is used from a combination of data sources including census PUMS and CTPP I, II data.

The procedure to combine the data for use in the simulation is as follows:

- Start with the Public Use Micro Sample (PUMS) 1990 5 percent sample. This data has approximately 13,000 individual and household records. Each record has a weight associated with it that indicates the number of people it represents. This data contains individual and household data for the metro area with detailed demographic variables.
- Eliminate households with no workers as transportation aspect of the model only works for people with jobs. The housing square feet is adjusted for the number of residents in the model.

- Change PUMS weights so correct numbers of demographic types are created as compared with numbers in the TAZs for the study area. This is determined by number of working families given in the CTPP I data.
- Assign each worker to a job location randomly based on probabilities from CTPP II data. For example if a TAZ has .5 percent of all jobs in an employment sector then the probability a person employed in that sector will work in that TAZ is .5 percent.
- Add income to thousands of households with less than \$5,000 of reported income to buy housing. Many households reported incomes that could not support it with food and housing. While in theory all income should be reported obviously many income sources were omitted in the Census PUMS data.
- Finally eliminate TAZs with fewer than 36 workers as the model has difficulty converging with such small numbers. Additionally accuracy is difficult to attain with so few households.

Not all data sources come at the same scale needed for the model and some elements did not combine together easily (Fig 4.3-1). The CTPP I 1990 was used to provide demographic data at the TAZ scale that range from hundreds of feet to miles. These are smaller than a census tract. Jobs come from CTPP II data. CTPP data covers two counties, Knox, the largest, and all of Blount to the south. The total population in the CTPP 1990 two county area is 421,718 with 167,074 households, 162,901 jobs, 305 TAZs, and about 1,000 sq miles. Individual data came from 1990 PUMS data for Knox County. After eliminating TAZs with less than 35 working households in Knox County the data was reduced to 196 TAZs and 98,294 households. The total population after fitting the PUMS data to working households was 248,087.

Table 4.3-1 Variable types and resolution

Type of Variable	Spatial Resolution	Temporal Resolution	Unit Measures	Source
Household Characteristics:				
<i>Household Traits</i>				
Head of Household demographics	n/a	n/a	Nominal	PUMS
Total household income	n/a	n/a	Dollars	PUMS
Number of children in household	n/a	n/a	Integer	PUMS
<i>Preference of Neighborhood</i>				
<i>Demographics:</i>				
Income of residents in the neighborhood	100s yards	n/a	Dollars	CTPP I
The race/ethnicity makeup	100s yards	n/a	Black/White	CTPP I
<i>Preference</i>				
<i>Geographical/Economic:</i>				
Residential Density	100s yards	n/a	Pop. Sq. Mile	simulation
Sq-foot per person	feet	n/a	Feet	simulation
Access to retail and entertainment (CBD)	100s yards	minutes	minutes	simulation
Commuting cost (time)	n/a	minutes	minutes	simulation
Neighborhood Characteristics:				
<i>Demographic Make up</i>				
Average Income of residents in the neighborhood	n/a	years	Dollars	CTPP I
The race/ethnicity makeup of neighborhood (percent Black/White)	n/a	years	Black/White	CTPP I
<i>Economic/Geographic</i>				
Price per sq-foot of housing	n/a	Years	Dollars	simulation
Residential Density	100 yards	years	People/mile	CTPP I/Tiger
Access to retail and entertainment/Dist CBD	miles	minutes	minutes	Dynasmart
Time to Work:				
Location of home	100s yards	years	TAZ	simulation
Location of work	100s yards	n/a	TAZ	Imputed CTPP II
Connectivity and congestion on the road network	feet	minutes	minutes	Dynasmart

Chapter 5: Calibration Results

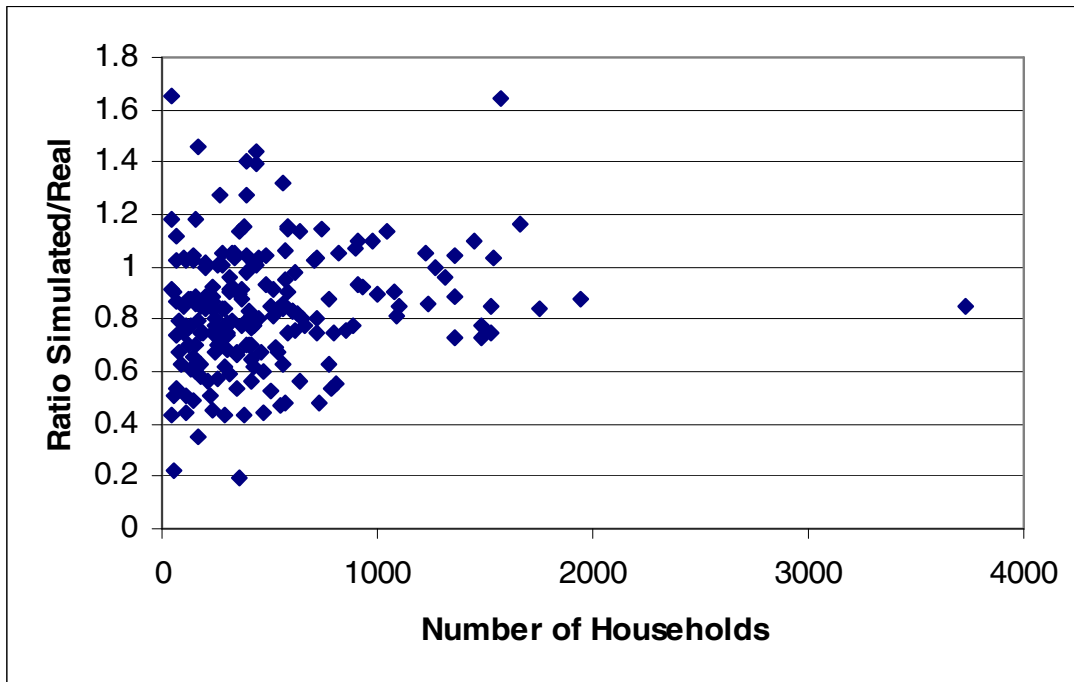
The results of this research are discussed in the next two chapters. In this chapter the results of the calibration or baseline run is shown. In order to calibrate the model to fit real data, agent preferences are adjusted so that working families move to neighborhoods consistent with ‘reality’ as measured by 1990 Census data. In the next chapter, several scenarios are run in which transportation variables are changed in order to measure their effect on location, housing price, population, and JTW time.

In order to create a base-line run, realistic data in Knox County is used as a comparison point for simulation runs. Agents are given preferences and data is normalized based on previous work and the modeler’s best estimation. As discussed in the literature review section, preferences are not measurable to the level of precision necessary to model agent parameters. For the purposes of the study (given the data issues and simplifications used) having individual TAZs match Census data within a factor of 2 is considered acceptable for this research. While this may be considered arbitrary there are several reasons to use it as a goal. It allows the modeler and reader to get a sense of proportion between runs if compared against something familiar and intuitive like Census data variables. If there are either bugs in the code or an extremely unrealistic decision making process programmed into the code it will show up in very bizarre results in the baseline run. Also it provides a reality check for the relationships between variables. Another reason is that on average TAZs with more residents should on average match ‘real’ data more precisely than small TAZs because there should be less variability of results. This can also be verified for model runs. That said even if results could be matched precisely with the data it would not indicate more accurate preferences since with simplifying assumptions of variables and processes one would expect simulation results to differ from reality even if the preferences are accurate. A last goal of the calibration runs is to minimize systematic bias as much as possible, because results in different simulation scenarios could be biased in a particular direction.

5.1 Population and Household Fit of Simulation Results

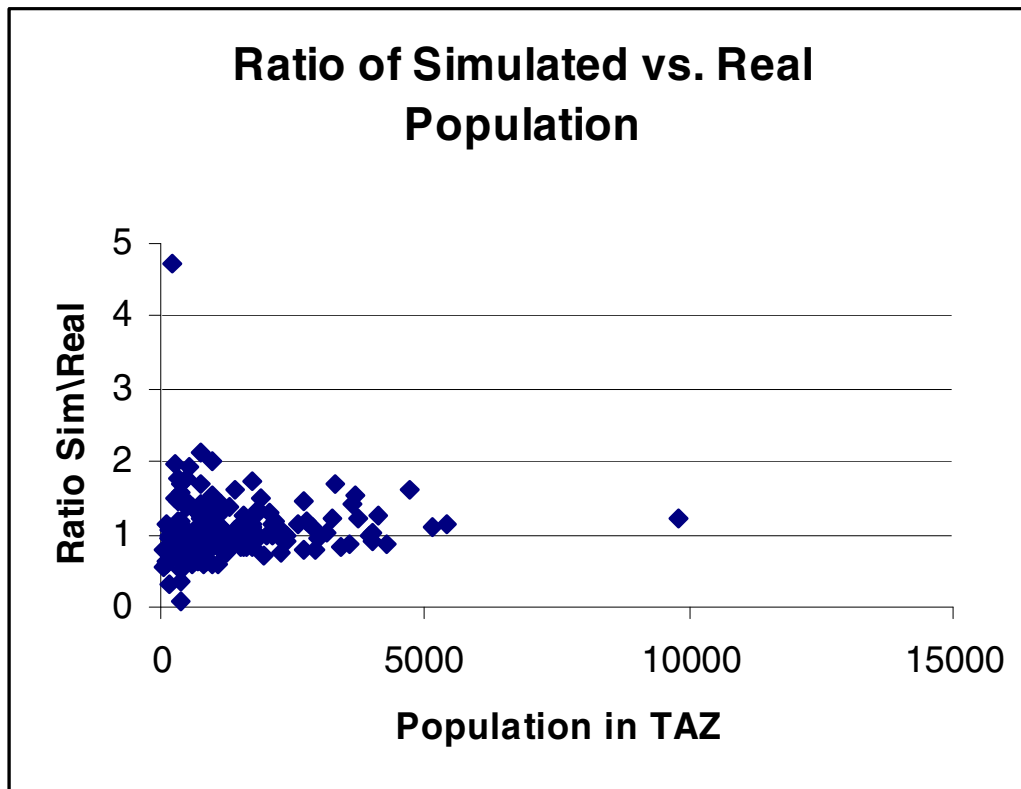
All TAZs converged during the base-line run. The criteria for convergence requires a vacancy rate of less than 10 percent for all TAZs, this is explained in detail in Chapter 3. As shown in Fig 5.1-1, TAZs with more than 800 people had a fit typically within 25 percent of real data which is well within the target range. The deviations reflect the error in the approximation for the space available in each TAZ. Several TAZs missed the factor-of-two target for this variable but those had small numbers of households.

Figure 5.1-1 Ratio households: simulated vs. real data by numbers of households for individual TAZs



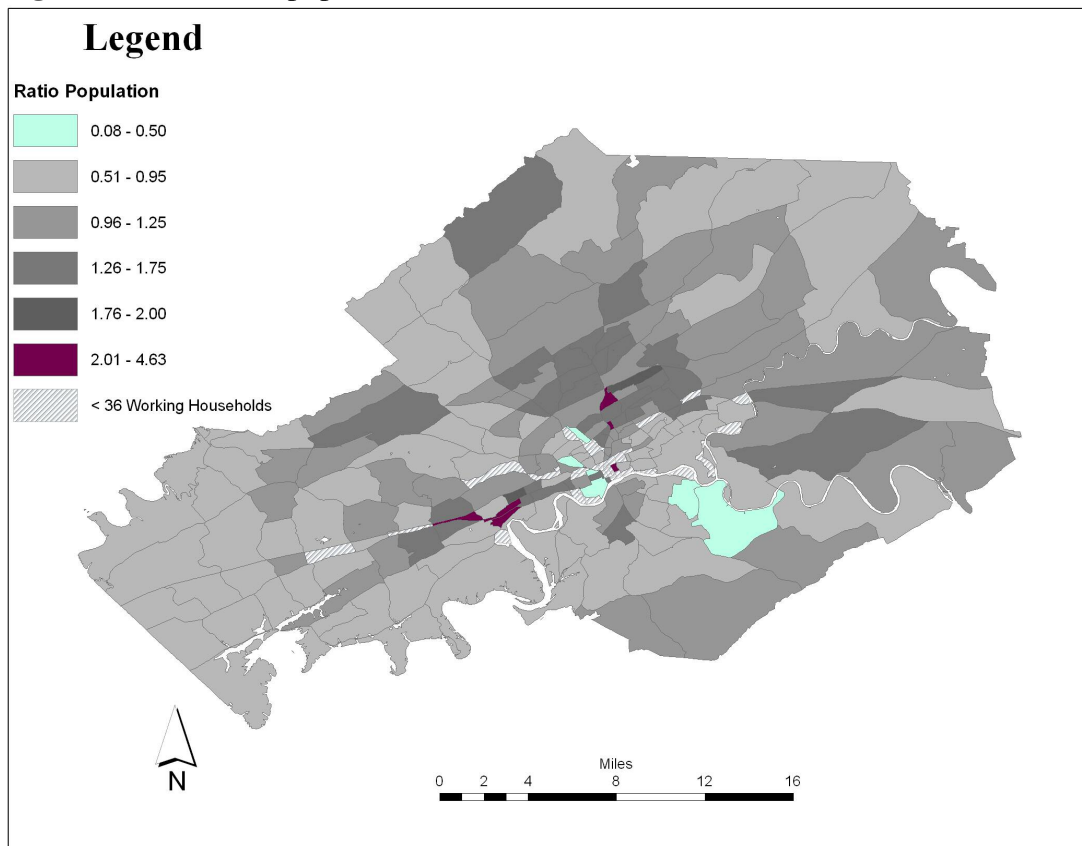
Examining the population graph (Fig 5.1-2) a noticeably larger dispersal exists for the population fit than households. The model was programmed to converge on the correct number of households not people so this is not a surprise. This is explained by households of different sizes than the size of households in real life moving into a TAZ. Both variables fit better in TAZs with large populations indicating a sound process is used in the simulation.

Figure 5.1-2 Ratio population: simulated vs. real data by numbers of households for individual TAZs



The geographic distribution of error in the population run did not show systematic bias except the large errors tended to be in and around the CBD (Fig. 5.1.3). These TAZs are more likely to miss the factor of two targets than the other areas. One of the TAZs that missed the target was the University of Tennessee TAZ. This is expected as commuting times are based on having one commute and having a car. Students without cars and multiple commute trips (classes) are not modeled properly in the simulations, for example raising the value of being close to the University. Low numbers of people in some TAZs could be richer households buying lots of square feet reducing population. Other TAZs showed a mix of results too high and low but within the target.

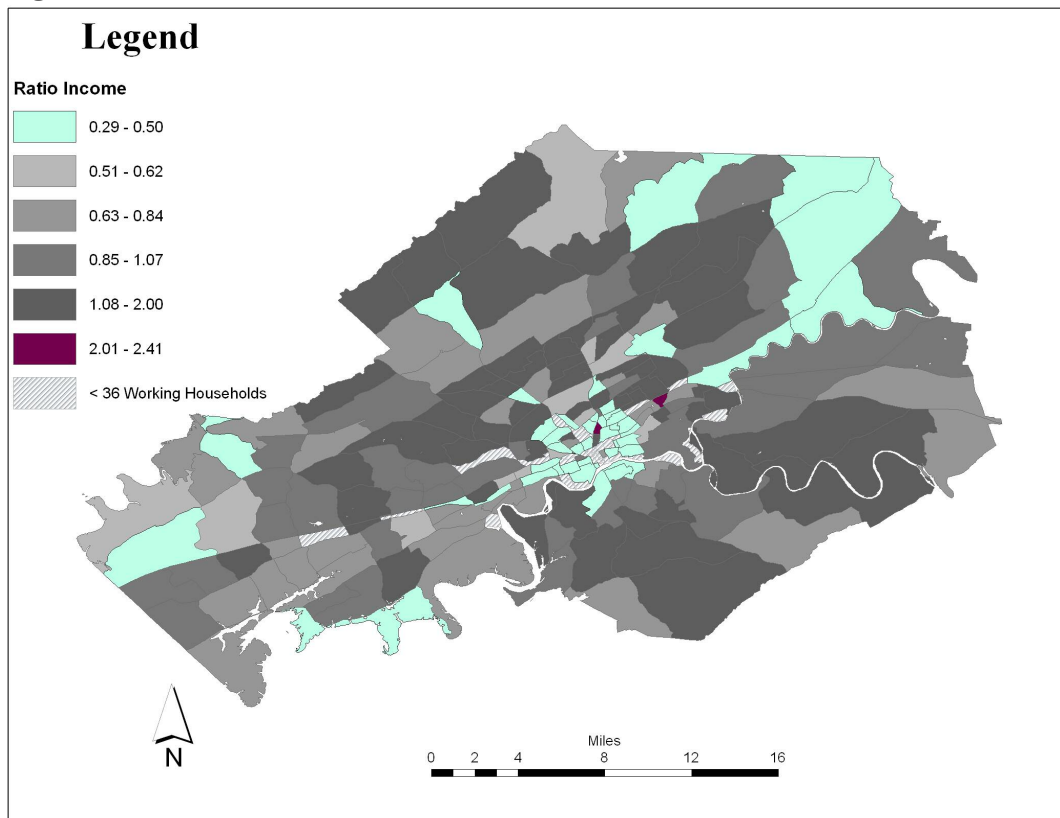
Figure 5.1-3 Ratio of population: baseline run vs. real data



5.2 Income Fit of the Simulation

The income fit was not as good as housing and population but it still was mostly within the target for this simulation. The larger TAZs had the better fit to real data. Examining the ratio map of the simulated versus real income outcome the geographic pattern reveals a tendency for the poor areas in the central city to be less than 50 percent of real income levels (Fig 5.2-1). Also some outer TAZs particularly in the East were lower than the real data. Some wealthier TAZs in the west also fell below the target range. Only two TAZs had a simulated income twice as high as the real data and both were located near the CBD. Most of the county showed mixed results higher or lower than the true value but within the target range. These results, while acceptable, do show some systematic bias that could not be eliminated.

Figure 5.2-1 Ratio of income: baseline run vs. real data



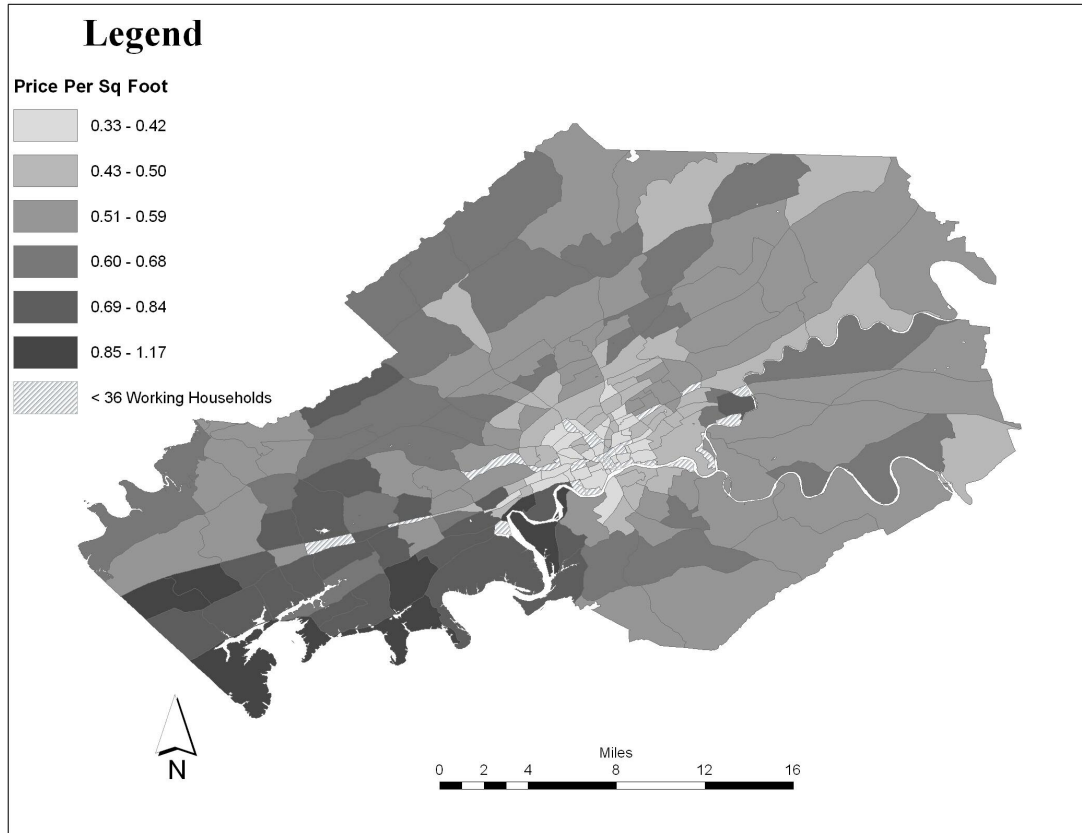
5.3 Price per Square Foot comparisons

The price per square foot is not in the 1990 Census data. The mean calculated for each TAZ is for rent per month per square foot measured in dollars. In the simulation the price varied between 35 cents to \$1.22 per square foot. This number on average would increase by about 50 percent by 2000. Spot checking online rental prices gave a bound on upper and lower rents.

Examining the map of price per square foot (Fig 5.3-1) sheds more light on the results. There is a fairly tight relationship between income (Fig 5.2-1 above) and

property values, such as in the western parts of the county that are richer and also have good access to jobs and the interstates. The close-in areas with high poverty levels and some outer suburban TAZs with low accessibility and low average incomes have the lowest price levels. These results seem reasonable and fit with expected household agent behavior³.

Figure 5.3-1 Price per square foot baseline run



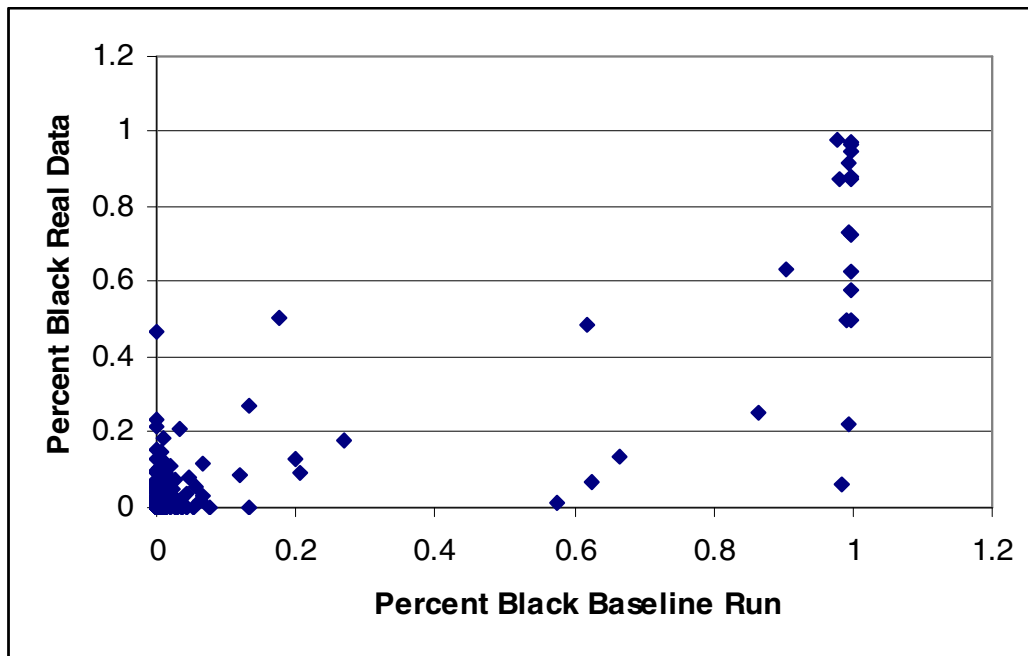
5.4 Racial segregation comparisons

A racial preference for percent black is included in the simulation. It was not expected to give a very good fit between the simulated and real values. This is because the residents are free to move in the simulation without cost and the output represents equilibrium so its results might more accurately reflect what will occur

³ These are housing not land costs so they do not contradict land rent gradient research. Many inner city areas have low cost housing in the United States.

given the present state of income and race preference over the long-term. Because some neighborhoods are in a state of transition, the results will not replicate the present state. The graph of real versus simulated race (Fig. 5.4-1) shows a very weak fit between the simulation and real data.

Figure 5.4-1 A comparison of percent black in TAZs: real vs. baseline run



Both real and simulated percent black outputs are mapped using the natural break categories calculated for the real data and tell a different story from the graphs above (Fig. 5.4-2 and 5.4-3). The concentration of poor black TAZs remains in the inner city although with more intensity than the present state. The outlying areas have a more dispersed black population in the simulation. The concentration of black residents radiating to the east and west of the CBD disappears being replaced by a small cluster in the far eastern of the county and less overall segregation outside the CBD. The black preference exerts a powerful demographic affect on the CBD and should be included in the model for accuracy in this area. The fit is the worst of all the variables but not worse than expected.

Figure 5.4-2 Percent black Knox County 1990

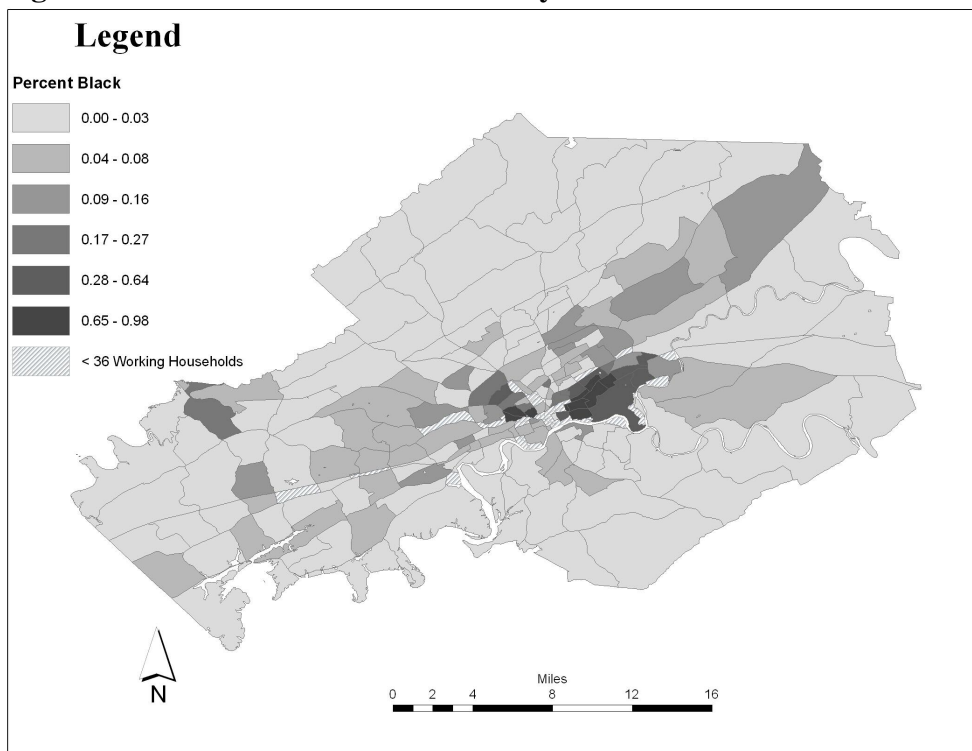
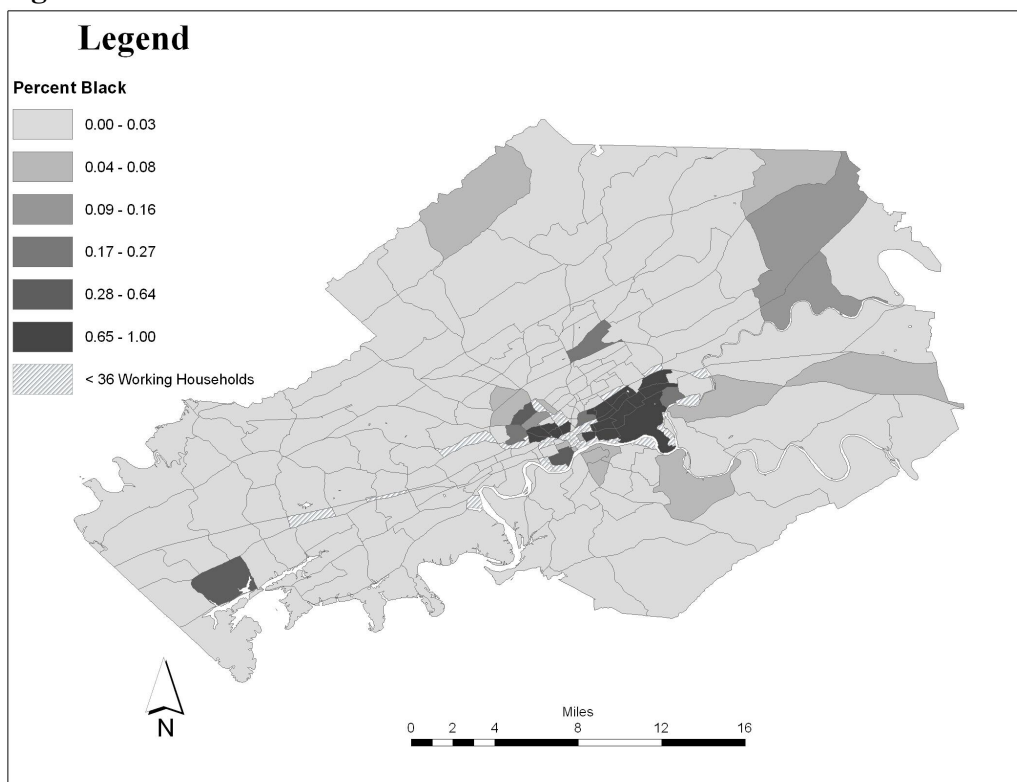
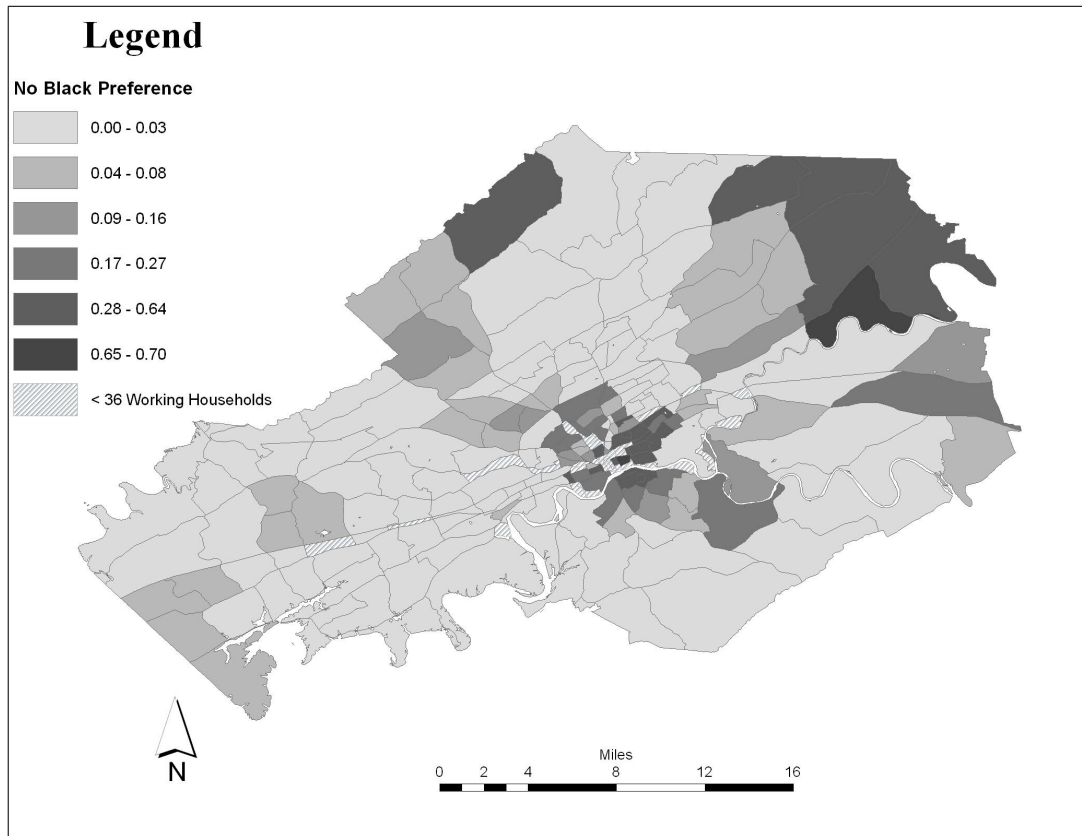


Figure 5.4-3 Percent black baseline run



To examine how strong the black preference is, the simulation run was done with baseline preference minus the black preference (Figure 5.4-4).

Figure 5.4-4 Percent black with no black preference included for household agents

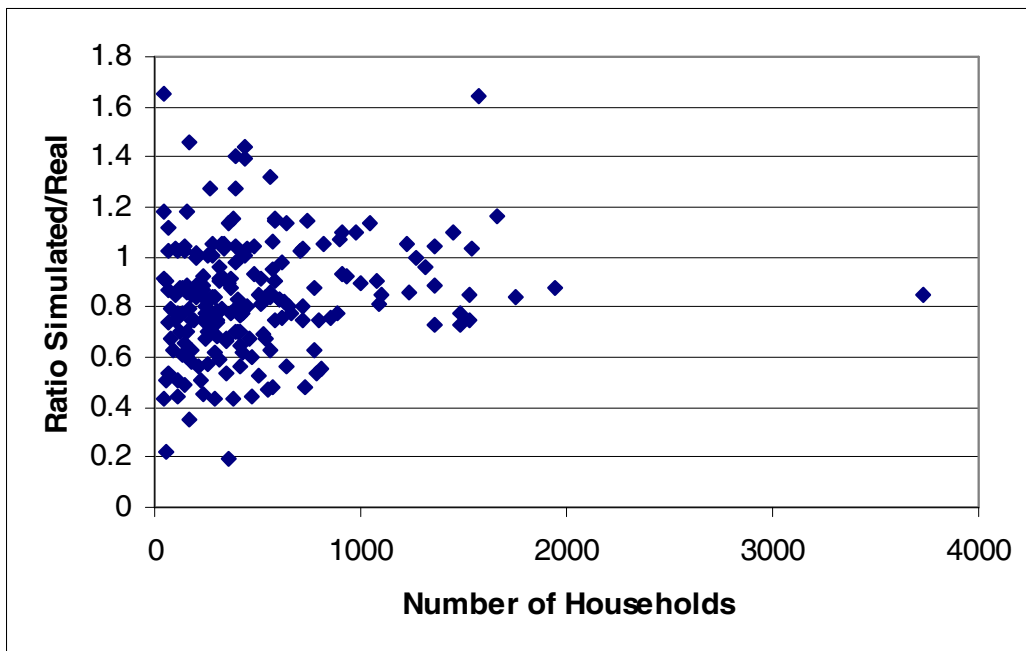


This map has the same color categories as the maps above. The black central city areas were recreated with the less intensity than the real and baseline simulation. There is also a large concentration in the eastern and northern section of the county that has low incomes and moderate accessibility levels. The main difference is that there are more blacks in the outer parts of the county and blacks are spread more evenly throughout the central city than in the baseline simulation. This simulation result shows segregation to be a combination of income and race but all black neighborhoods are not possible without racial biases.

5.5 Journey to Work Comparisons

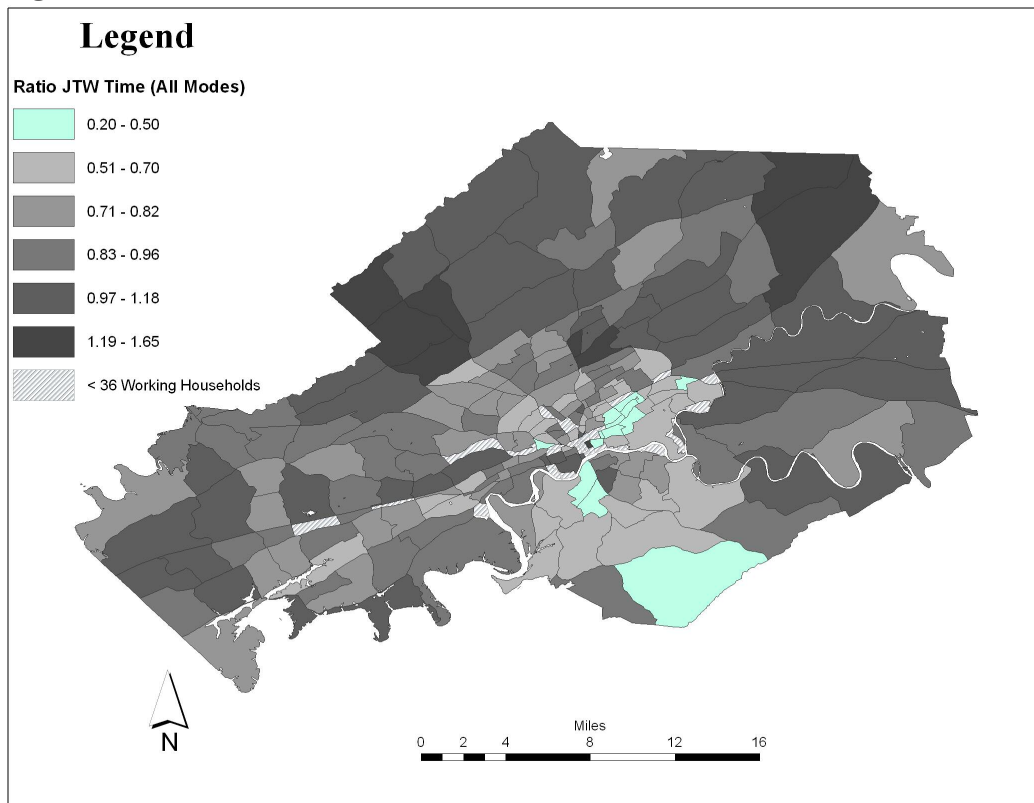
The average travel time per commuter was 17.85 minutes in the base-line run versus 20.91 minutes in the real data for all modes. Examining Fig 5.5-1 the JTW times seem to match reasonably well, although on average times were about 25 percent longer per TAZ. There was a reduction of this error for the larger TAZs. The fit of most of the TAZs were within the factor of two criteria. It performed similarly to population data.

Figure 5.5-1 Ratio of baseline vs. real JTW times (all modes) by number of households per TAZ



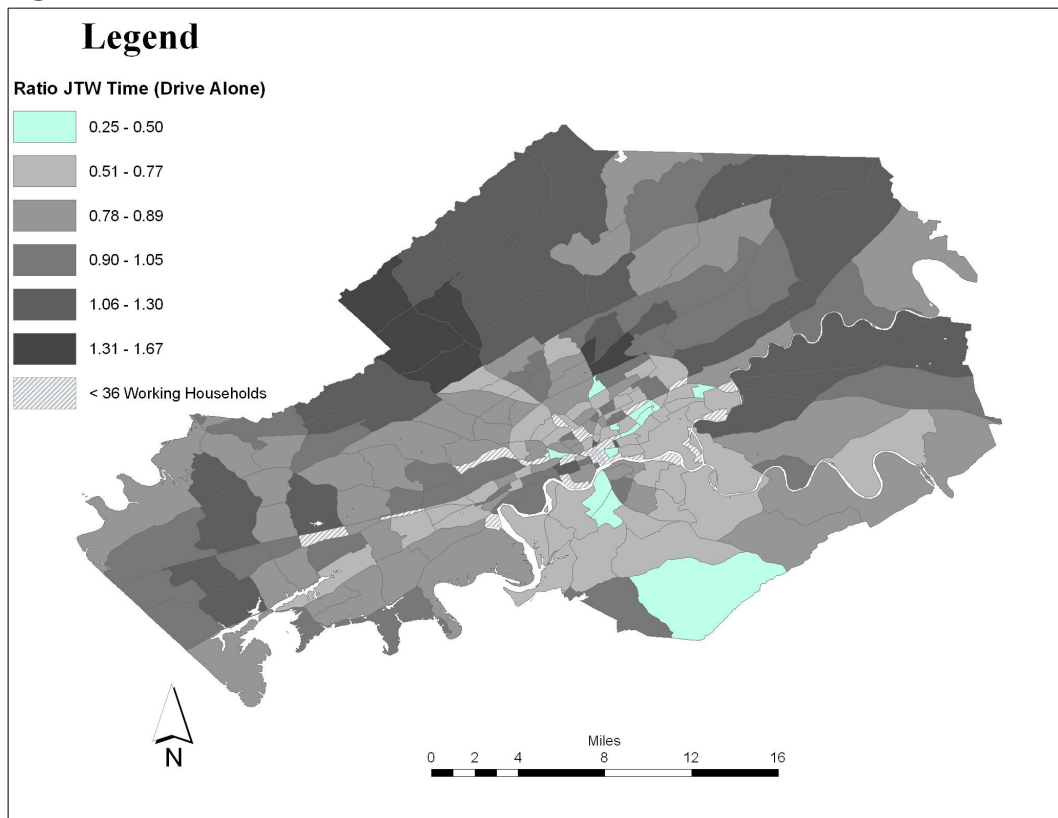
The northern and northeast section of Knox County had the largest overestimation of travel times although nothing double the original times (Fig 5.5-2). This could result from high congestion levels through the more sparsely populated areas that resulted in longer than reported travel times. The largest underestimates were in the central city particularly in the black areas resulting in part from the use of cars by everyone in the simulation.

Figure 5.5-2 Ratio JTW time: baseline vs. real (all modes)



The JTW time drive alone map (Fig 5.5-3) shows similar results to the all modes comparison map, although the underestimation near the CBD did not encompass as many TAZs, and was not as great, as the all modes comparison. This seems to bolster the car explanation for underestimations in the above runs. The overestimations were also similar.

Figure 5.5-3 Ratio JTW time: baseline vs. real (drive alone)



There are several factors that make analysis of JTW times both complicated and difficult. First, in the simulation everyone drove alone so all other modes such as carpooling and walking, which made up about 23 percent of all work trips, were allocated to auto travel. Also, when a commuter reached their work place TAZ centroid the journey is considered complete. These two factors would on average underestimate work times compared to real data. The journey to work data CTPP is considered to be an underestimate of real times as they tend to be best case estimates when self-reported. Additionally, simulation times were based on congested times while there are many non-peak work journeys. This would tend to overestimate simulation times.

Comparing drive alone times may seem a quick solution but 23 percent of the work trips do not use this mode, this is higher in poorer TAZs. Both comparisons are made to try and get an accurate picture of the fit between the two times. As there is no cost

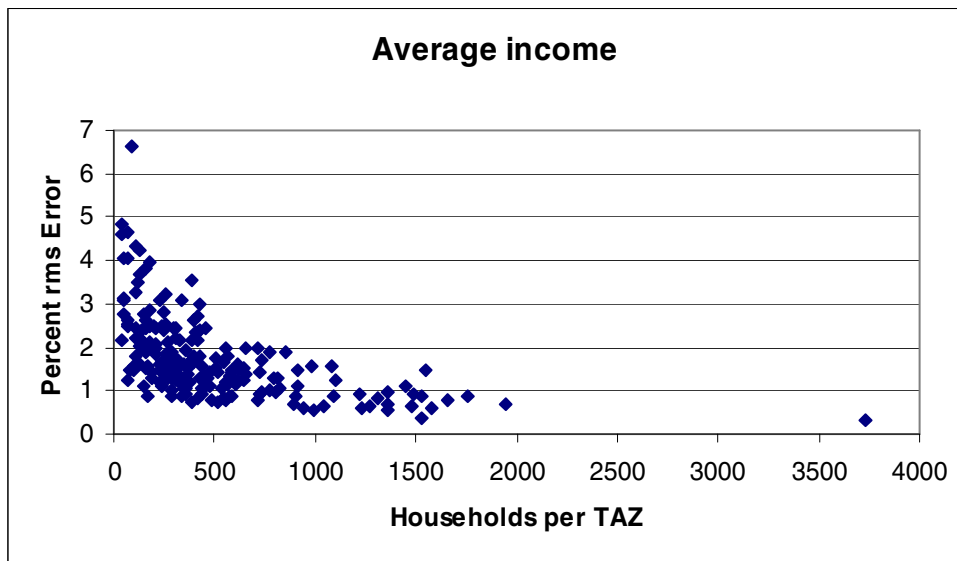
to moving residential location and jobs are fixed in the simulation making the results represent long-term adjustment to work location. A last issue is that travel times are grouped into *Super* TAZs that may alter travel times in some cases at the outer TAZ that are large in area WHAT DOES THIS MEAN?. This could be a significant time change but as explained in the method section this was necessary to get trips between almost all the TAZs.

5.6 Housing Model variability due to random variation

As stated in the method section the preference for housing and neighborhood attributes are assigned to agent households using random numbers. To examine the effect that using different random numbers has on the simulation ten runs were made using different seeds, meaning there were different random numbers used in the scenario. To examine the variability between runs, a *root mean squared* (rms) error was calculated and divided by the original values to give the rms error in terms of percentage for each TAZ. This was calculated for variables that will be used in analyzing the scenario runs including: average income per TAZ, price per sq foot per TAZ, people per TAZ, and JTW time.

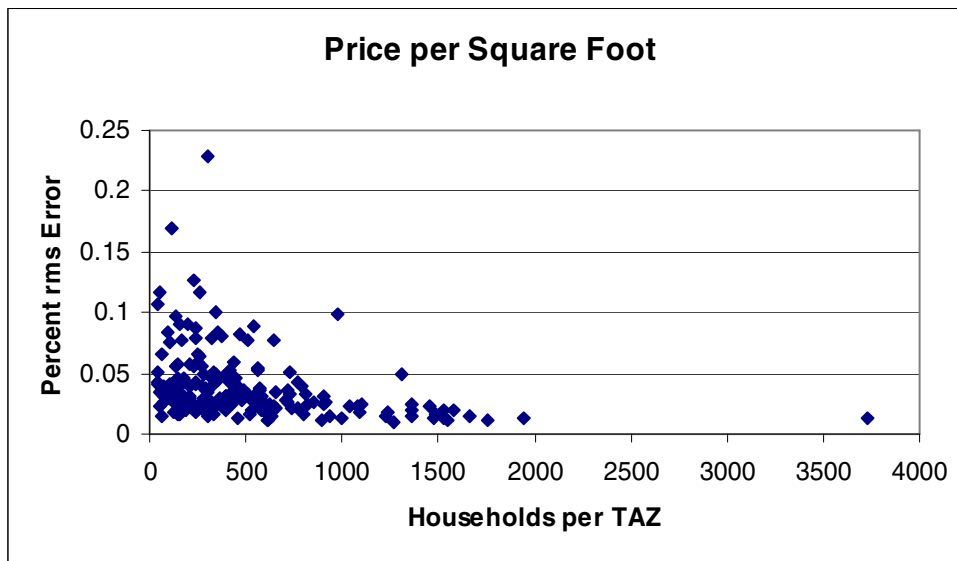
The rms error for average income was the largest out of all the four variables. However, no TAZ had a larger rms error than 6.5 percent (Figure 5.6-1). Like the calibration runs the smaller TAZs, less than 500 people, had much larger errors than the larger TAZs.

Figure 5.6-1 Root mean square error for average income between baseline simulation runs



The distribution of rms errors for the other variables had the same pattern as households. The rms error for population was about half that of average income (not shown). Most of the rms errors were less than .2 percent of the TAZ value. The maximum rms error is 6 percent with TAZs over 500 people having a 1 percent rms error or less. The price per square foot rms error is the smallest of the four variables with nothing larger than 0.25 percent (Figure 5.6-2). JTW time also had a small rms error.

Figure 5.6-2 Root mean square error for price per square foot between baseline simulation runs



These runs show that scenario results only alter slightly due to randomness. As only large changes in results are considered informative in this model and the error is small only single runs are compared.

Chapter 6: Scenario Runs

6.1 Scenario Runs for JTW variables

Several scenarios are run to examine the effect changing transportation variables will have on Knox County. The first runs examine how the alteration of preferences changes outcomes of housing price, income, population, and JTW times. This is accomplished by first excluding people's travel time utility then running a scenario only including utility for travel time and housing space. Next, travel time in-between TAZs are altered by increasing traffic from the original run by 50 percent as well as doubling travel times between all TAZs. In the final scenarios a travel and space only preference is combined with increasing travel times. This will provide an upper bound on the importance of transportation for residential location decisions in the model.

Like the baseline run, the demographics of the neighborhoods are not altered to reflect new realities of population movements. This is done so the initial effect of transportation change can be compared with the baseline run and with each other. If the demographics were changed, many more iterations would be required and the final results would be affected by many different variables not just the one changed for the scenario run. However, where demographic and density preference are eliminated the simulation represents a close approximation to a steady state.

6.2 Removal of Travel Preference

In this scenario, the travel time preference is removed from the location choice model. This will give an indication of the importance of transportation in the spatial arrangement of residence in Knox County. The same data and preference values, except travel preference, as the baseline run are used for the model run. Examining table 6.2-1 we can see the effect of this change on price per square foot. The overall minimum, maximum, and spread of values are similar to the baseline run about 70 to

75 cents. The median and standard deviation of TAZs of both runs are very similar as well.

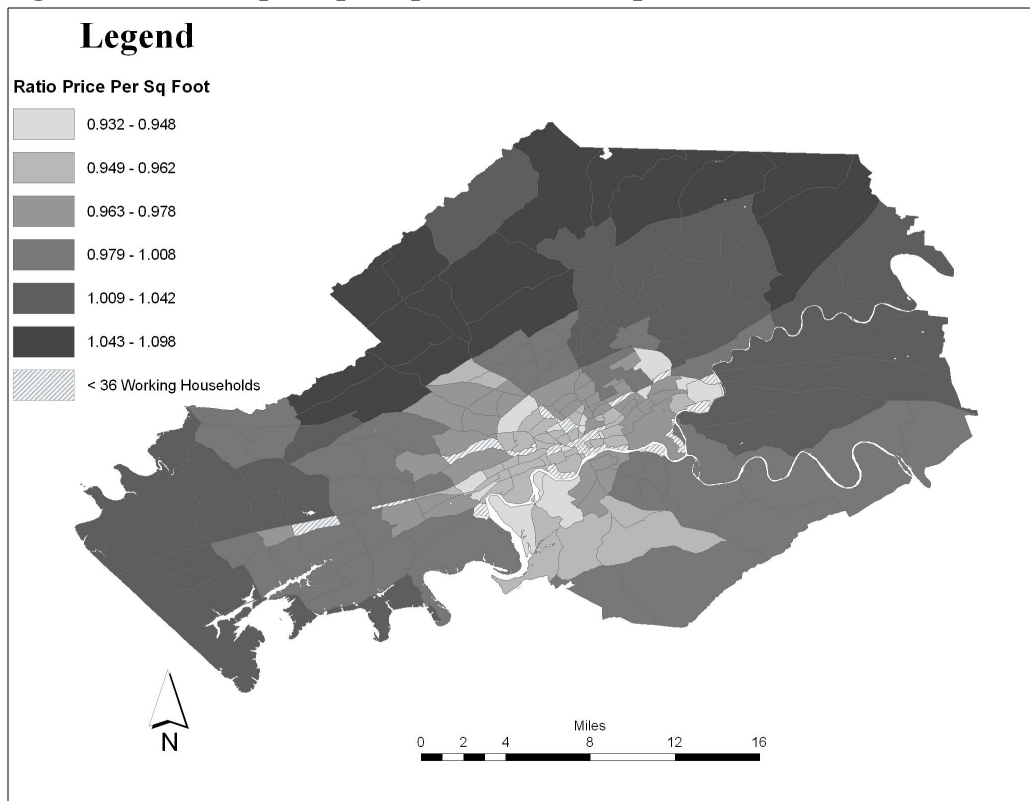
Table 6.2-1 Comparison baseline vs. no-travel preference run (change)

	PPSQ Foot (Standard)	PPSQ Foot (No Trav Pref)	Ave HH Income (Standard)	Ave Income (No Travel Pref)	JTW (Standard)	JTW times (No Travel Pref)
Median	0.53	0.55	29,323	27,649	24.0	24.0
S.D.	0.13	0.14	19,635	20,526	13.0	13.0
Ave. All	n/a	n/a	Exogenous	Exogenous	17.8	28.1
Min	0.35	0.33	2970	3,015	2.0	2.0
Max	1.22	1.17	99,339	98,858	65.0	65

Subsection 6.2.1 Preference change affect on price per square foot

The next step is to examine the difference in spatial distribution of the baseline simulation versus ‘no-travel preference’ runs. The main changes to price per square foot are reduced housing prices of TAZs in the central area of the county but nothing greater than 8 percent (Fig 6.2-1). Also prices increased modestly in TAZs about 4 miles from the CBD to the county border, although this effect is more pronounced in the eastern and northern parts of the county. Not surprisingly the largest positive change is in the northern TAZs, which had the longest JTW times and correspondingly bad accessibility in the baseline run. The magnitude of the change is not as large as expected, if demographic changes and new developments were included as a feedback in the model the effect could be much greater.

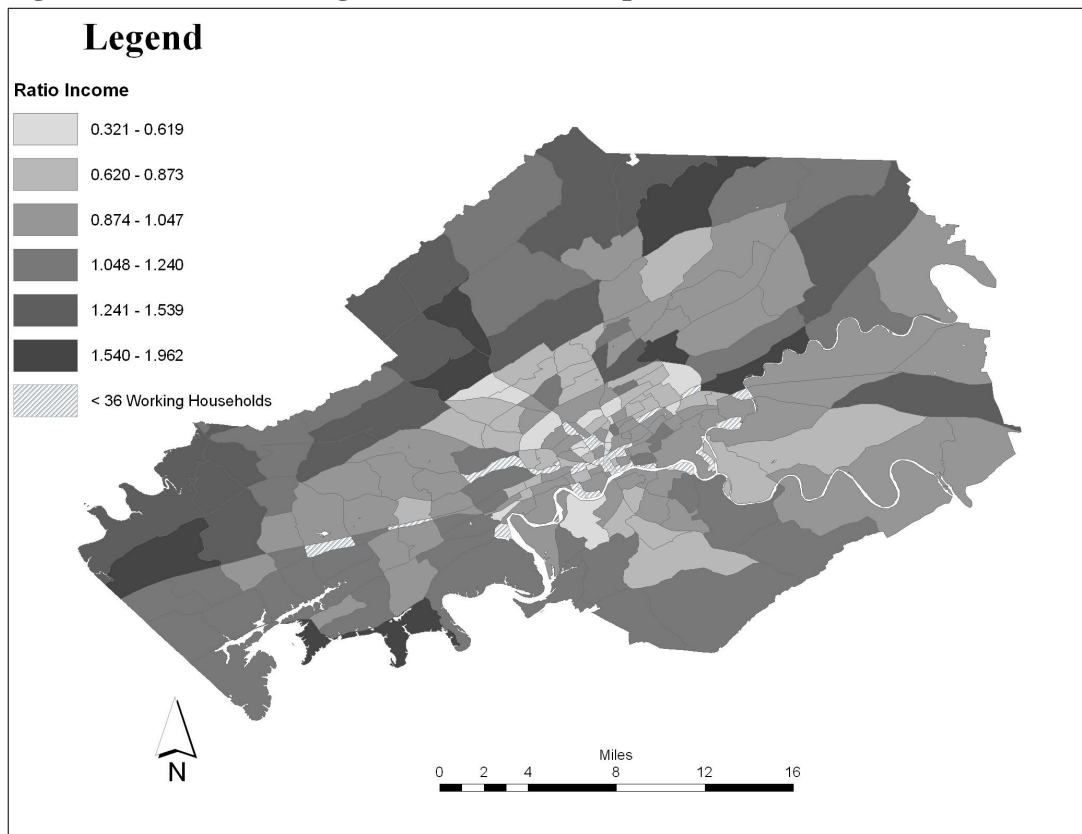
Figure 6.2-1 Ratio price per sq foot: no-travel preference vs. baseline



Subsection 6.2.2 Average income results for no travel preference run

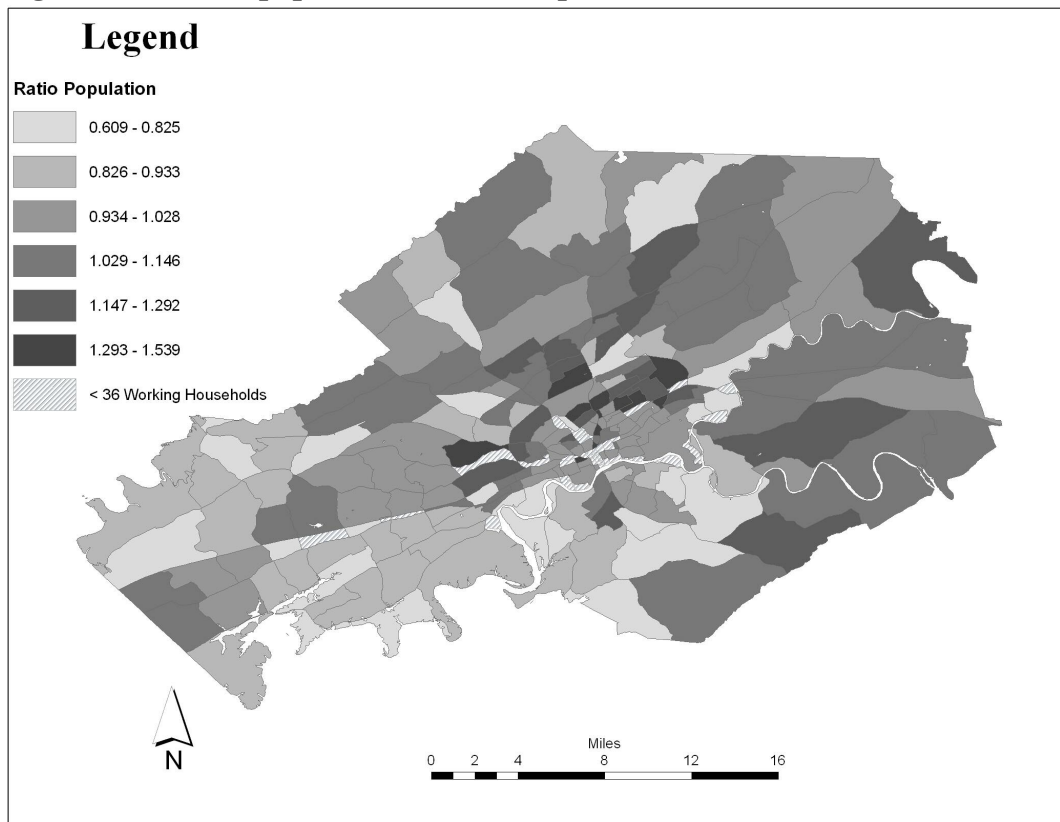
Next the change in household income is examined. The values of the maximum, minimum, median, and average income of TAZs are almost identical. This was also true of the standard deviation (Table 6.2-1). The average income did not mirror the changes in the price per square foot run geographically (Fig. 6.2-1 above Fig 6.2-2). Like price per square foot, the negative change was larger in the CBD but this change is not uniform. For instance the predominately black area in the CBD increased in income. Also, there were decreases in income in some of the southeastern areas, which were poorer but had better access than the north. While lower access TAZs with other ‘positive’ attributes (particularly in the west which were closer to the CBD) increased in absolute and relative income levels.

Figure 6.2-2 Ratio average income: no-travel preference vs. baseline



The map of population change shows a mix of results. Just north of the CBD tended to increase while these areas both gained and lost average income (Fig. 6.2-3). The eastern and southern parts of the county also had mixed population and income change; although some of the largest changes were TAZs sandwiched between the CBD and the border TAZs. Conversely the TAZs with the greatest average income away from the CBD, almost uniformly, had population loss but with income gain, particularly in the western parts of the county. This indicates wealthier residents displacing poorer residents because of neighborhood quality. These residents then moved either north, east, or into the black areas of the CBD, often raising the average income levels.

Figure 6.2-3 Ratio population: no-travel preference vs. baseline



Subsection 6.2.3 JTW Times No Travel Preference

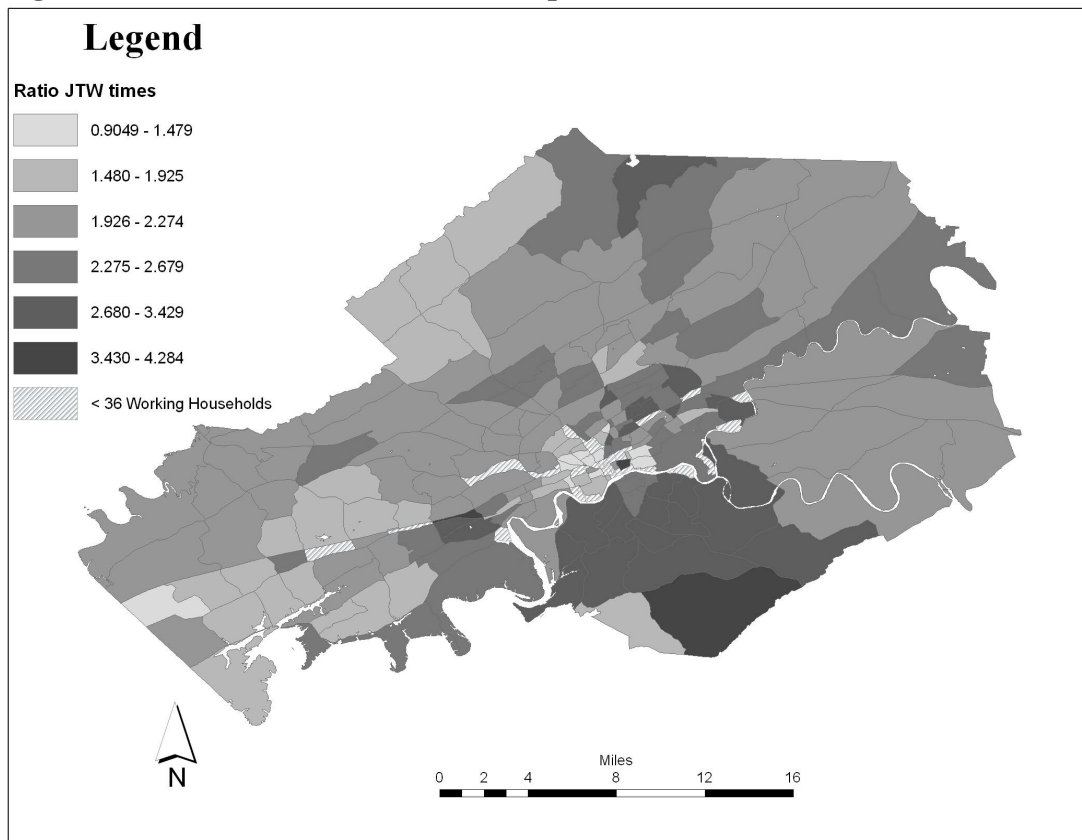
Not surprisingly there is a large increase in the average JTW times for all commuters from 17 to 28 minutes (Table 6.2-1). Individual TAZs all increase their average travel times except near the CBD which has good accessibility. The largest increase is in the south of Knox County.

Households on average are moving to neighborhoods with less accessibility to their jobs. They have longer commutes than they did in their old neighborhoods illustrated by increased average times and traveling further than the previous residents in their new neighborhoods (Figure 6.2-4). The results show that the JTW preference will have an important but uneven affect on urban location. Price and income do show a similar pattern both declining near the center of the county in general but population levels fluctuated by TAZ (Figure 6.2-1 and 6.2-2 and 6.2-3). However, price barely

changes while income alters by a factor of 2 or more. The largest effect is on the JTW times, as one might expect, with income having the second largest change by TAZ. In a sense we have turned the county in to a retirement community with only demographics playing a large role in location decisions.

This shows with no demographic or developer feedbacks that much of the income and price differences we observe are the result of neighborhood quality within small urban areas. Unexpectedly eliminating transportation preference exacerbated already existing income variations while keeping existing housing price at similar values. The desire to live near work by itself exerts a large integrating affect on average neighborhood income levels. This finding does not negate the possibility of increased accessibility creating, through a feedback effect of changing demographics and development, very wealthy neighborhoods.

Figure 6.2-4 Ratio JTW time: no-travel preference vs. baseline



6.3 Only Housing and Space Preference Scenario

In the next simulation run, demographic elements like income, race differences, and the tendency to dislike the dense urban environment are eliminated. This leaves only the travel time to work and housing space preference of household agents. Like the previous scenario it will produce a bound on the importance of transportation but in this case its importance is exaggerated (in theory) by only paring it with the housing space preference. The results will reach a steady state in a way the previous runs do not as demographic feedback from location change (not put in the model) will not affect household agent location decisions.

Subsection 6.3.1 Price per Square Foot - Travel and Housing Space Preference

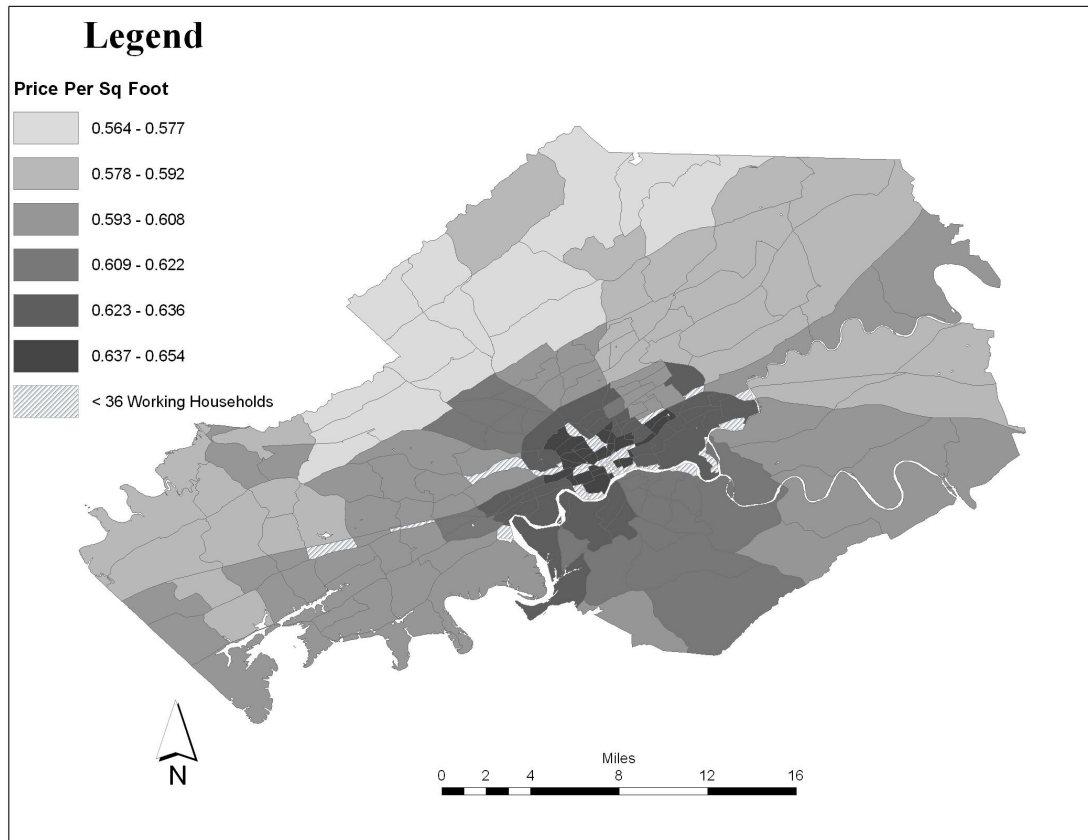
In contrast to the removal of travel preference when only housing space and travel times are included there is a large change in housing price and average income levels. First examining the comparison of price per square foot to the baseline run we see a significant change (Table 6.3-1). The price range is 56 to 69 cents per square foot with a standard deviation of only 2 cents. This is 21 cents higher than the baseline minimum but a 57 cent lower maximum. This indicates a marked flattening of rents between neighborhoods in this scenario, undoubtedly because of the removal of criteria that would differentiate price between TAZs. Despite this the median TAZ rental price is higher for this simulation run.

Table 6.3-1 Comparison Travel Space Preference Only Runs with the Base-line

	PPSQ Foot Baseline	PPSQ Foot T&SPrf	Ave Income Baseline	Ave Income T&S Pref	JTW Standard	JTW T&SPrf
Median TAZ	.53	.60	29,323	43,544	24	14
S.D. TAZ	.13	.02	19,635	13,071	13	7
Ave. All HH	n/a	n/a	Exog.	Exog.	17.85	17.35
Min	.35	.56	2970	6977	2	6
Max	1.22	.65	99,339	76,971	65	43

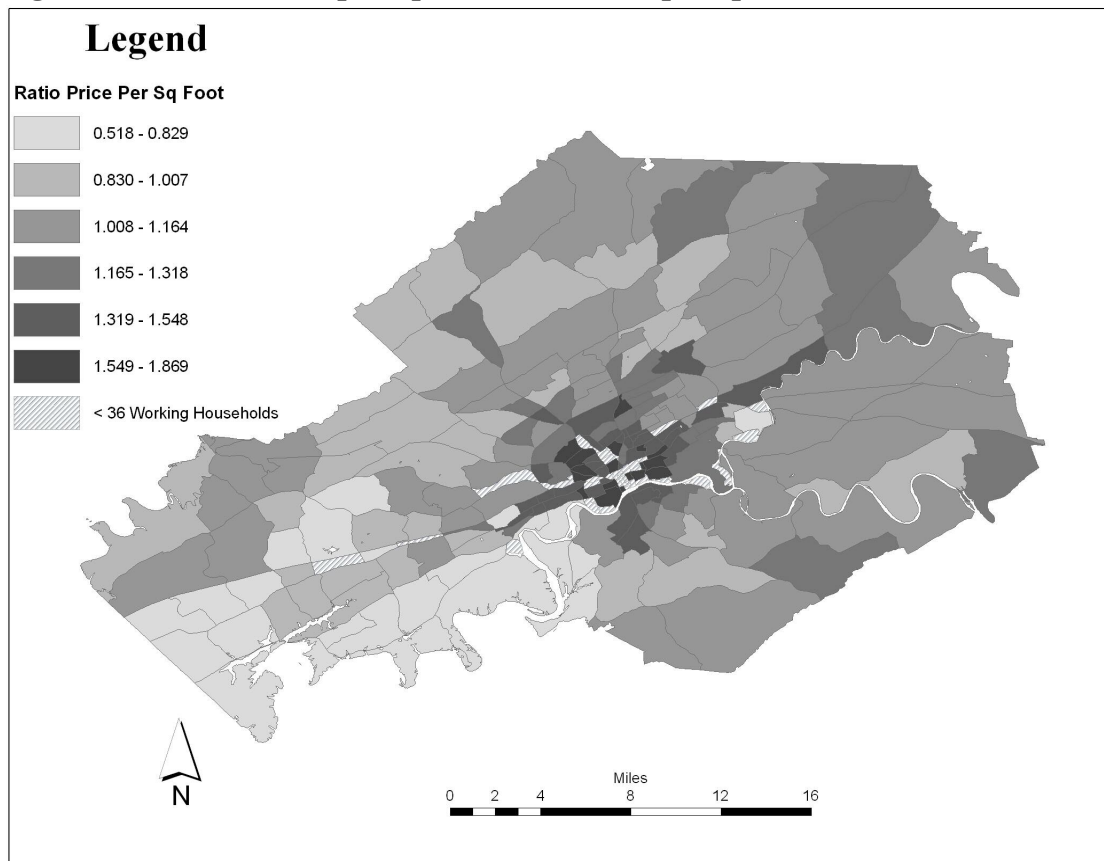
Looking at the map of price comparisons (Fig 6.3-1) the prices, which have flattened, have a reversed pattern from the baseline run. The CBD has the highest price per square foot while the outer areas notably in the north have the lowest values indicating an inferior accessibility to jobs. This map pattern looks more like New York than Knox County.

Figure 6.3-1 Price per sq foot only travel and space preference



The TAZs that lost the most value under this scenario are the highest priced areas in the southwest of the County. Their position as having high average income with low density no longer positively affects their housing price (Fig 6.3-2). The biggest increase is right downtown while north, south, and the east of the county have a moderate increase in price over the baseline run. Changes range from 52 percent of original value to an 87 percent increase. In a majority of TAZs housing price value experience a moderate increase. Areas that have poor accessibility but have prices that increase show the importance of demographics in the location decisions in those neighborhoods in the baseline run.

Figure 6.3-2 Ratio Price per sq foot: travel and space preference vs. baseline run



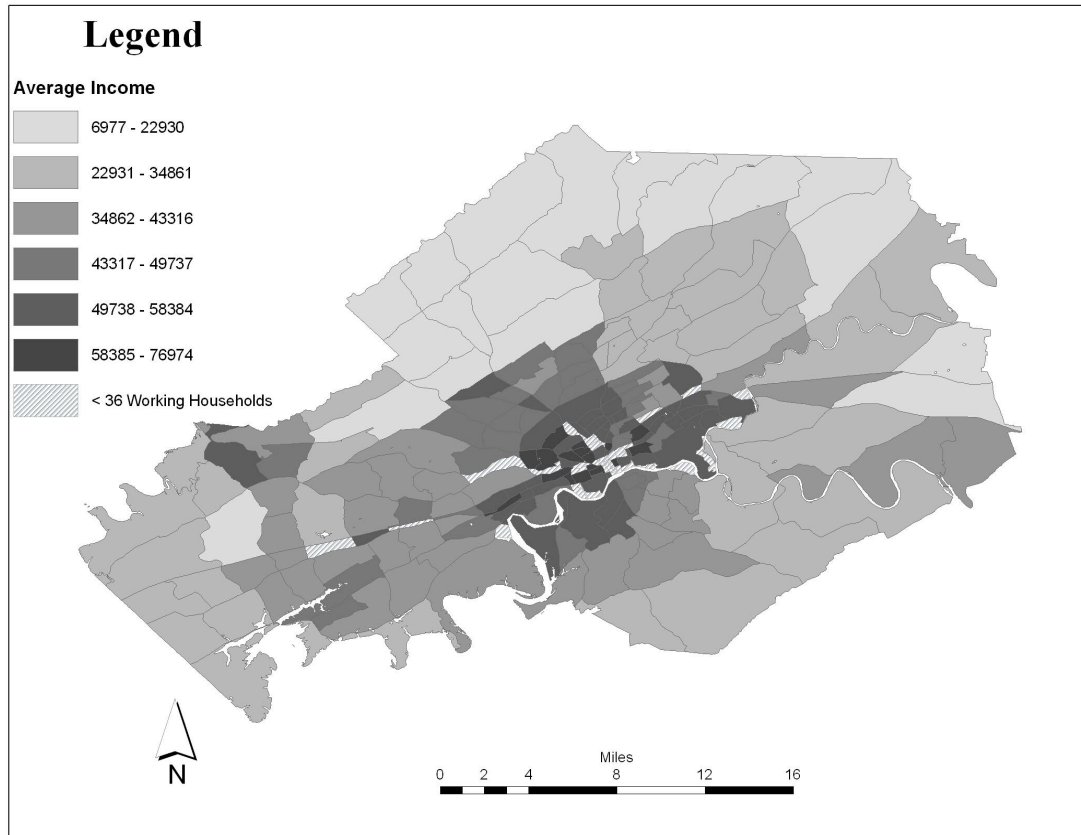
Subsection 6.3.2 Average Income - Only Travel and Space Preference

The income level had a similar overall pattern as price per square foot. With a higher minimum income and a \$60,000 lower maximum income than the baseline run (Table 6.3-1). The median income per TAZ was \$41,427 for this run vs. \$29,171 for the baseline run (Table 6.3-1). The standard deviation 11,831 was a little more than half the baseline run. The income pattern is not a pronounced donut shape instead there is a hodgepodge of higher income levels in the CBD and the western county near the freeway and job clusters (Fig 6.3-3). The CBD contains several high-income clusters. The higher income levels extend to the south and west a few miles. The far outer northern and eastern TAZs have the lowest income levels.

The reason for the flattening of TAZ income difference is that the rich can concentrate in TAZs where jobs are plentiful not because high-income residents live

in a given TAZ. Without the income preference they do not concentrate nearly as much in high-income neighborhoods and consequently do not bid up prices. Poorer households do not repel other households raising the minimum income levels.

Figure 6.3-3 Average Income only travel and space preference

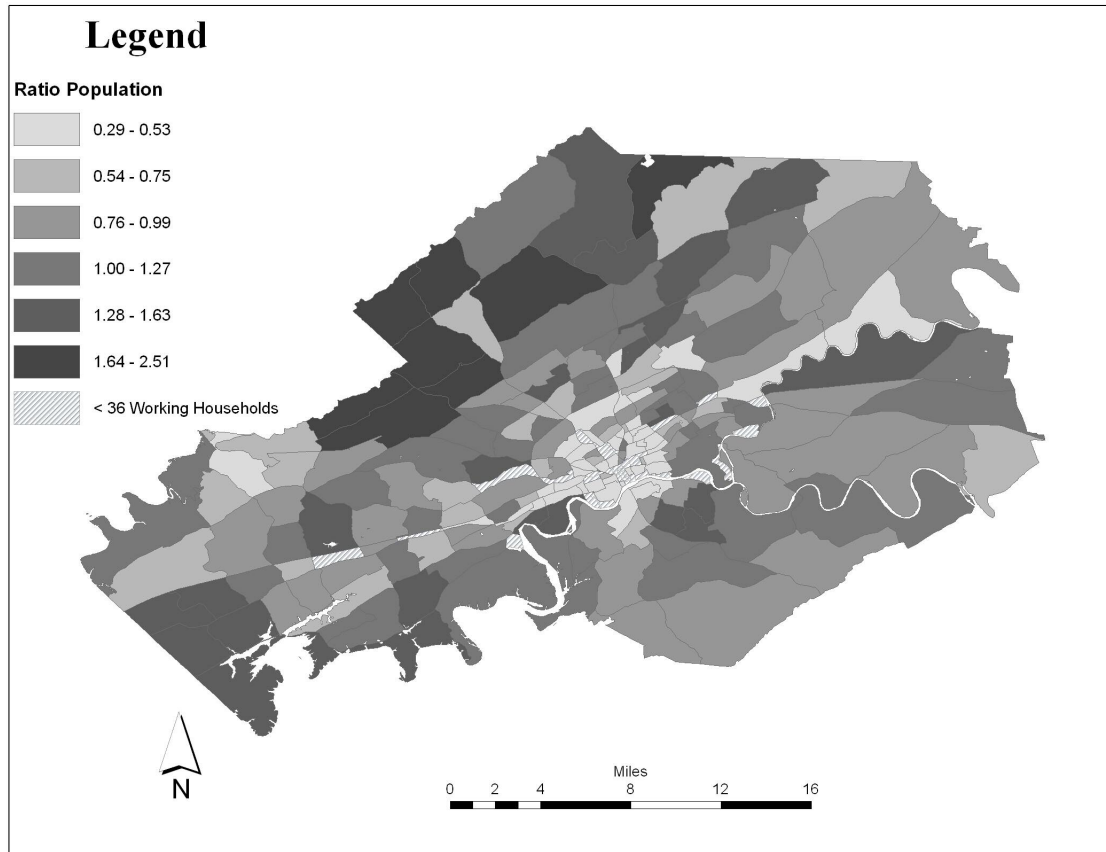


The change in the distribution of population offers more insight into the effects of changing preferences on Knox County. As the richer people moved to the center they rented larger amounts of square feet per person than the original inhabitants (Fig 6.3-3). This resulted in the CBD suffering a 50 to 70 percent reduction in population. This forced poor residents out of the CBD.

There generally is an increase of population in the outer parts of the county, only the CBD showed a clear pattern of population decline (Fig 6.3-4). In the CBD richer residents bought more square feet per person reducing the population levels despite being a more desirable place to live. The east and most of the north gained in population. This is indicative of household agents making different choices in the

trade off for JTW time and housing space. Often in TAZs with similar attributes, sometimes adjacent to each other, one will gain while the other loses population. The result is a pattern that did not follow the change of either average income or price per square foot.

Figure 6.3-4 Ratio people per TAZ: only travel and space preference vs. baseline



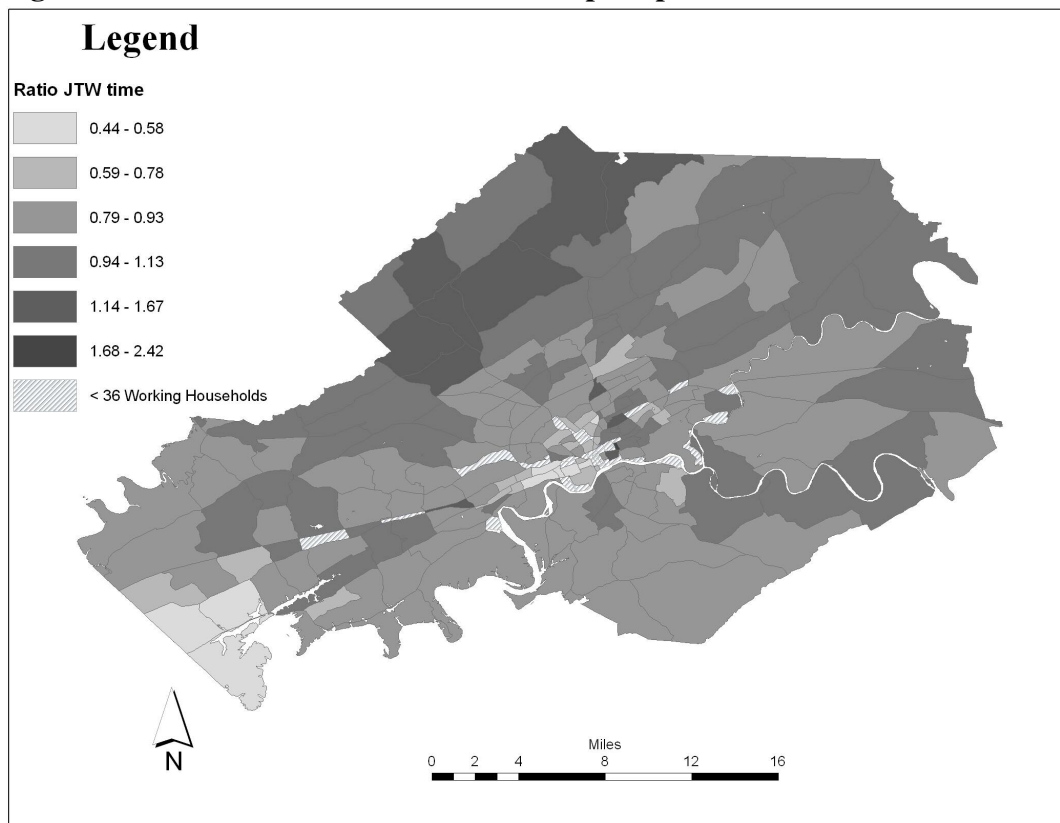
Subsection 6.3.3 JTW Time Results - Only JTW Time and Housing Space Preferences

The average time to work for the only travel and space preference scenario is about the same as the baseline run, 17.85 vs. 17.35 minutes each (Table 6.3-5). However, the median and standard deviation for this scenario are about ½ the baseline run. With the churning of population and the fact that travel time accounted for a much larger proportion of location decisions this overall result is unexpected. However, with the space preference still in operation and poor people being moved to outer suburbs away from jobs this is a logical outcome. In real life housing developments would be built to meet increased demand changing the JTW times. These results seem

to indicate increasing travel preferences importance will not, at least initially, affect travel times as much as the no-travel preference run. This possibility will be investigated later in section 6 below.

The geographic distribution of JTW times shows a weak donut shaped pattern (Fig. 6.3-5). This was the second greatest geographic change by TAZ, just behind population although as stated above the average travel time by county was nearly the same (Table 6.3-1). Originally it was expected to see a larger relative effect, many more declines by TAZ, and a larger change in the JTW time at the county scale but the space preference and bidding dampened this effect.

Figure 6.3-5 Ratio JTW time: Travel and space preference vs. baseline



These results show that different processes can attain similar travel times. Just the inclusion of the housing space preference is enough to give similar JTW times as the baseline run assuming no new development.

6.4 Increased Traffic 50 Percent and Double Travel Time Runs

As the simulation covers only one county with relatively low congested travel times the results from the previous runs may shift radically if travel times mimic larger urban areas. To test these results two kinds of travel times are used. First traffic was increased to 150 percent of the original levels in Dynasmart resulting in new travel times. This will be concentrated on heavily used links. Second the travel time was multiplied by two for the time it takes to travel between TAZs. Both approaches may not behave exactly like real-life traffic increases but they offer two different ways of testing the results of increasing travel times.

Subsection 6.4.1 Price per square foot 50% more traffic and double travel times

The increased price per square foot in the increased traffic scenario has a similar pattern to the baseline run. The minimum, maximum and standard deviation are virtually the same for the 50% more traffic scenario and just a little higher for double travel time scenario (Table 6.4-1).

Table 6.4-1 Price per sq foot for increased traffic runs

	Baseline run	50% more traffic	Double travel times
Median TAZ	0.53	0.54	0.57
S.D. TAZ	0.13	0.13	0.13
Min	0.35	0.36	0.37
Max	1.22	1.24	1.32

The geographic distribution (not shown) has slightly higher prices in the high accessibility areas. This is expected although the changes are rather small, nothing more than 18 percent for any individual TAZ.

Subsection 6.4.2 Average Income and Population Change for Increased Travel Time Scenarios

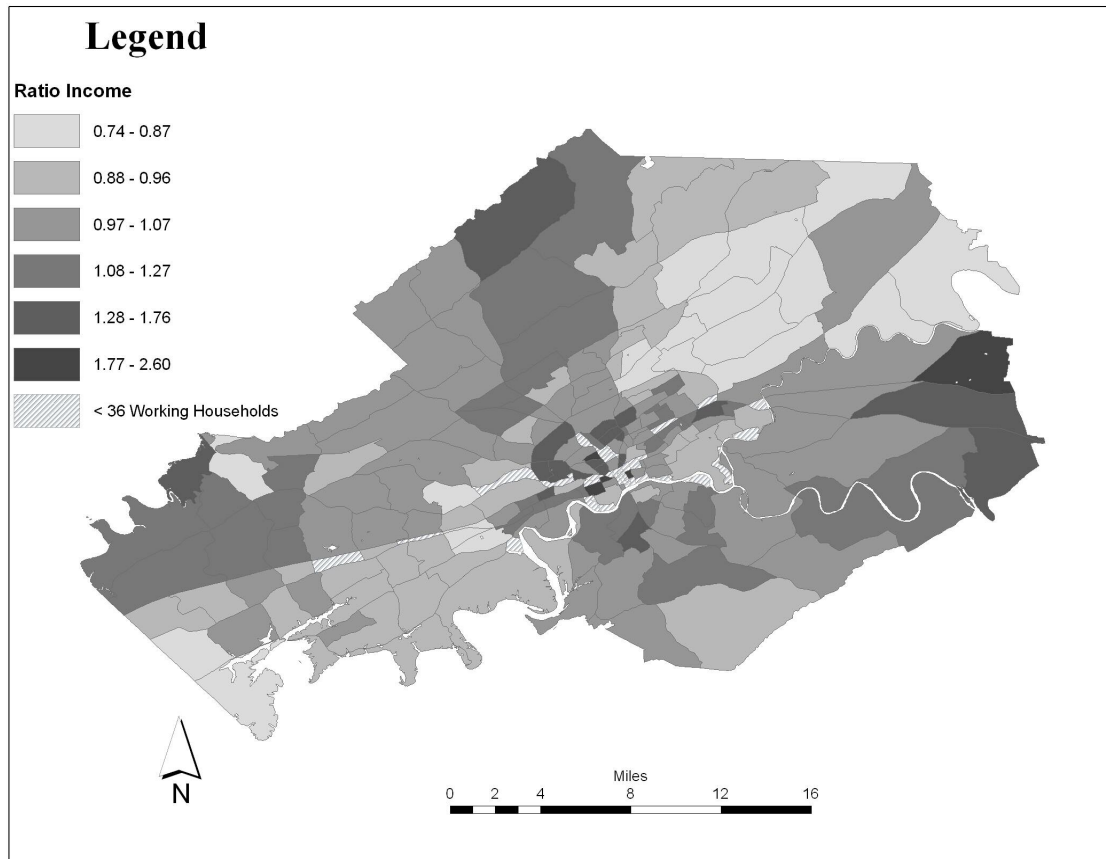
Increased traffic scenarios displayed similar change for average income as the price per square foot. The maximums were both about 3,000 dollars lower with medians and minimum the same as the baseline run (Table 6.4-2). The standard deviation was about the same for the 50 percent run but \$1,700 lower for the double travel time run. The categories were similar between all three runs with minor variations.

Table 6.4-2 Average income results for increased traffic run

	Baseline run	50% more traffic	Double travel times
Median	29,323	29,521	30,614
S.D.	19,635	19,279	17,943
Min	2970	2,940	2,966
Max	99,339	96,814	96,869

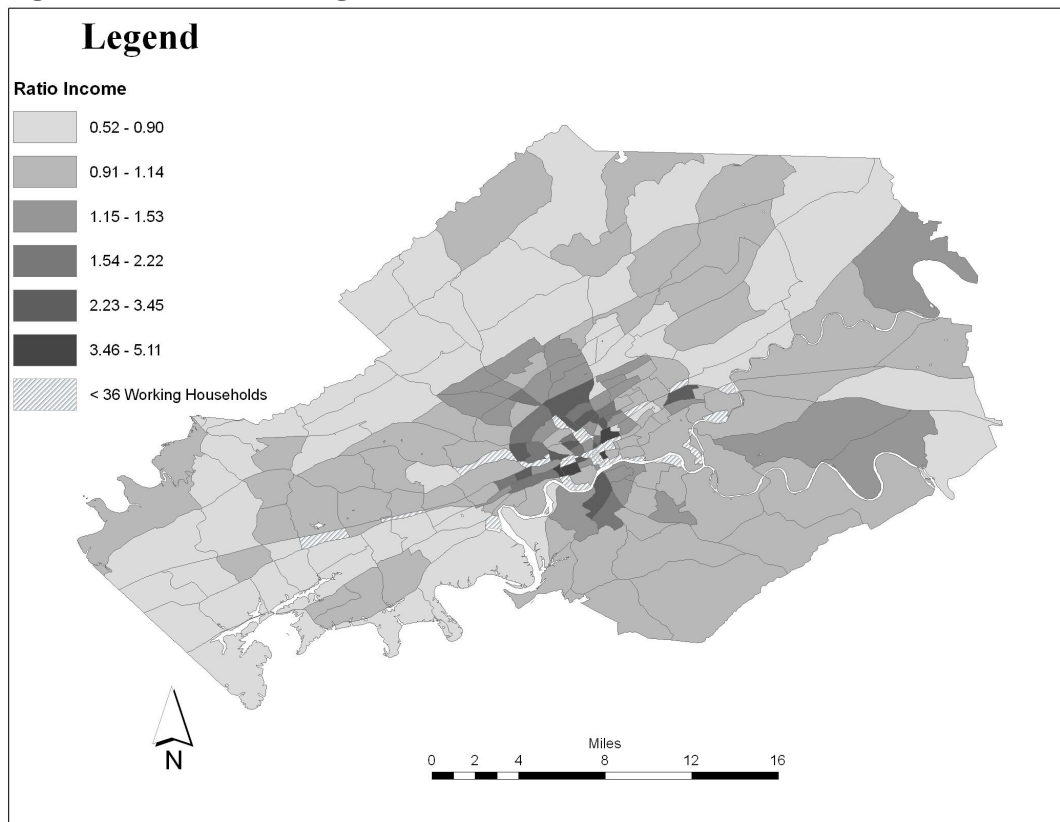
The increase traffic 50% scenario produced increases in average income TAZs in similar areas as the price per square foot. The CBD itself has more mixed results than the housing price change with some TAZs having higher average incomes while others have lower values than the baseline run (Fig 6.4-1). The variation is larger than price per square foot with a maximum decrease of 46 percent and an increase of 260 percent over the baseline run. Most of the TAZs fall in the two middle categories with a ratio of .97 to 1.27 compared to original average income results. The increase in average income in the CBD outcome is expected but not the increase in the outer parts of the county that have low accessibility. This may, in part, result from being insulated from changes to congestion levels, as there is little traffic to begin with.

Figure 6.4-1 Ratio average income: 150% traffic vs. baseline



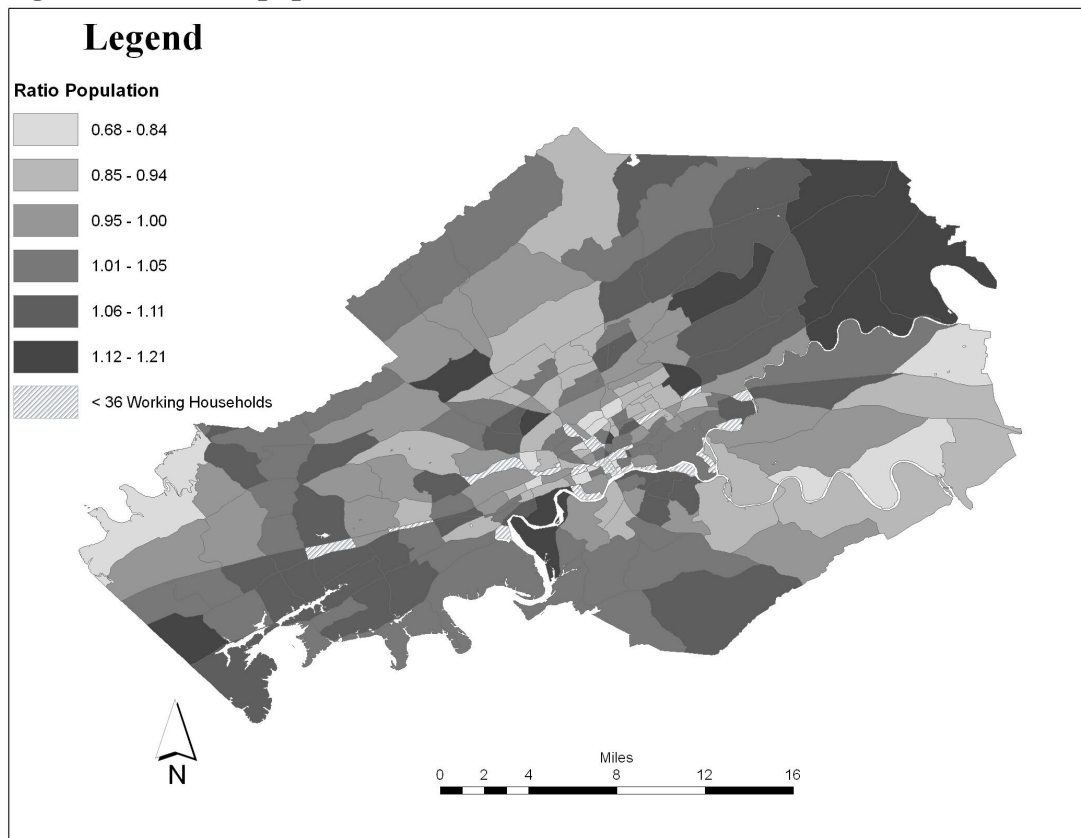
The double travel time results for average income shows an increase in and around the CBD (Fig 6.4-2). Away from the CBD incomes stayed the same or were reduced a maximum of 50 percent. The income effect is larger than the 150 percent traffic scenario ranging from minus 48% to plus 510%. The pattern did not match the 150 percent traffic distribution having a much stronger positive correlation near the CBD. The outer county stayed the same or dropped in income.

Figure 6.4-2 Ratio average income: 2*travel time vs. baseline



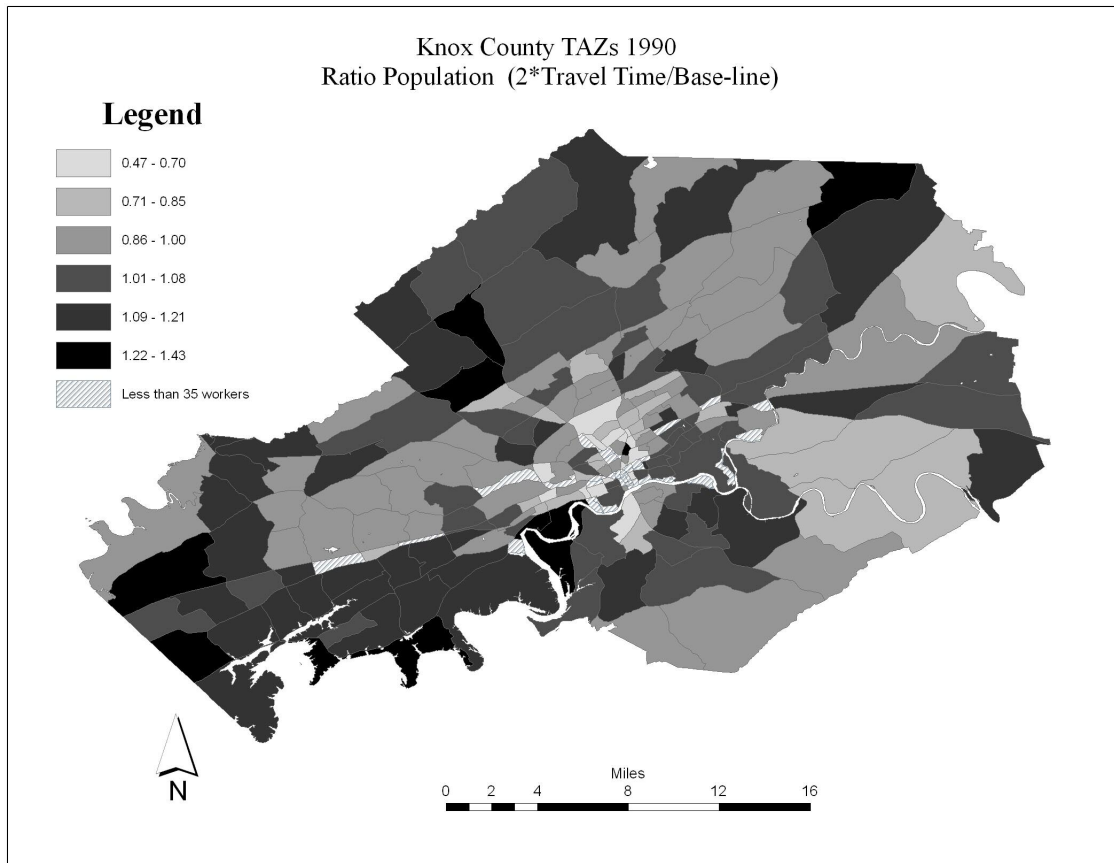
Looking at the population-change maps adds a bit of clarity. In the 150 percent traffic scenario, the more affluent residents displace poor ones in the CBD in the central and northern section while the black area gains population. Those areas that have increased income also correspond to higher rental prices but with fewer people (Fig 6.4-3). The population shifts to the south and southwest close to jobs. However, many of the largest shifts are to the east with some of the worst access but cheap rents. Some of these areas have desirable demographic characteristics others do not, such as in the northern parts of the CBD.

Figure 6.4-3 Ratio population: 150% traffic vs. baseline



The double travel time also had a mixed relationship. The north simply increased in population with reduced rental price resulting from poor access (Fig 6.4-4). The northern CBD lost population as more wealthy households moved in acquiring more space per person than previous residents. To the southwest and south generally TAZs that increased average income also gained population and increased housing cost. The far eastern part of the county had poorer residents move into those areas. The scale of change was about 20 percent larger than the 150% traffic scenario. Overall the population change did not match income or rental price further illustrating that demographically similar household agents are making different household location decisions.

Figure 6.4-4 Ratio population: 2*travel time vs. baseline run



Subsection 6.4.3 JTW Times for Increase Traffic Scenarios

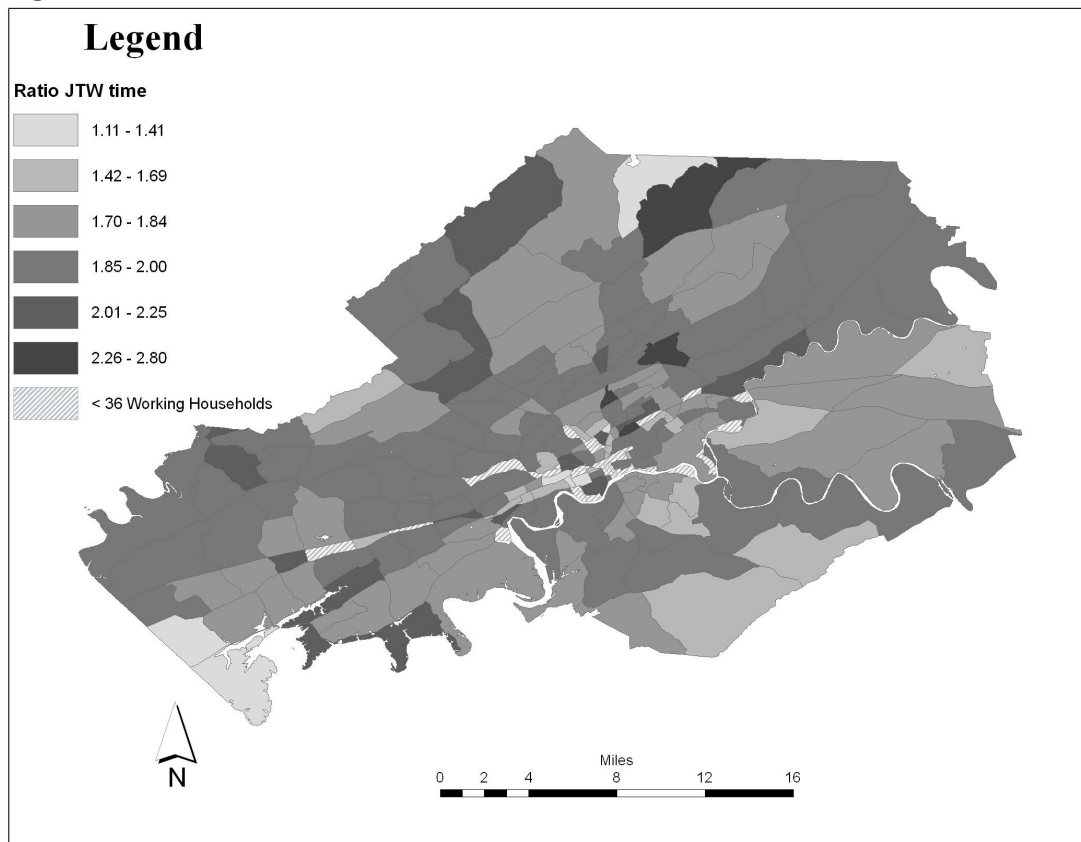
The 150% traffic scenario produces relatively predictable JTW time outcomes. The average for all households is about 7 minutes longer increasing from 17.85 to 24.58 minutes. The overall average for double travel times is 32.96 minutes about 80% longer than the baseline run (Table 6.4-3). The double travel time scenario is about twice as high as the baseline for median, minimum, maximum, and the standard deviation of TAZ travel times.

Table 6.4-3 JTW travel time comparison with different traffic levels

	Baseline run	50% more traffic	Double travel times
Median TAZ	24	34	48
S.D. TAZ	13	15	25
Ave. All HH	17.85	24.58	32.96
Min	2	3	4
Max	65	79	130

The geographic distribution of JTW times for the double travel time run is not the same as the baseline run with proportionally larger increases to the north and far west of the County (Fig 6.4.-5). Times in the northwest and the northern CBD increased the greatest amounts. Very few TAZs increased by more than a factor of two over the baseline indicating households are changing their locations to lessen the overall effect of travel time increases.

Figure 6.4-5 Ratio JTW: 2*travel time vs. baseline



It must be reiterated that these runs are not the end state. This measures the effects of what people will do given original demographics. So long-term changes may be magnified or reversed by feedback effects. These results show preference for density and desirable demographics still exert a very powerful effect even given radically altered transportation realities in a mid size city. These results are showing complex outcomes resulting from just a few preferences. Like the no-travel preference run, the major shift at the TAZ level was in income and JTW times.

The next runs will combine travel time with only travel and space preference so the effects of travel time can be isolated and these runs do run to equilibrium. Also this will create an upper bound on the effect of increasing travel times on the city.

6.5 Increased Travel Times Combined with Only Travel and Space Preference

These scenarios are run with only travel and housing space preference along with an increase in transportation times. This combination will allow a better understanding of the long-term effects of transportation change because they run to equilibrium and there is only one other preference, space, which affects location choice.

Subsection 6.5.1 Price per Square Foot Change

The price per square foot produces similar results to the original travel and a space preference scenario, which is the baseline run for this section. The increase traffic 50% run only changes slightly from the (new) baseline run (Table 6.5-1). The only real difference is the standard deviation was double the baseline run but still only 4 cents. The double travel times showed a larger housing price spread, 30 versus 9 cents, and a standard deviation 3 times higher than the baseline run (Table 6.5-1). The minimum was 7 cents lower but the maximum 14 cents higher than the baseline run. The relative changes were similar to the baseline run and original increased traffic scenarios.

Table 6.5-1 Price per Square foot comparisons: travel and space preferences with increased travel times

	Baseline (T & S preference)	50% more traffic	Double travel times
Median	.60	.60	.60
S.D.	.02	.04	.7
Min	.56	.56	.49
Max	.65	.69	.79

Subsection 6.5.2 Average Income Increased Traffic and Only Travel and Space Preference

Average income again mirrors the results for the new baseline run but with increased maximum income levels (Table 6.5-2). The standard deviation and median are a little higher for both increased travel time scenarios. Again there are very similar relative values as compared with original baseline and increased travel time scenarios. All the values are within a few percentages.

Table 6.5-2 Average income comparisons: travel and space preference with increased travel times

	Baseline run (T & S preference)	50% more traffic	Double travel times
Median	43,544	42,630	43,262
S.D.	13,071	13,300	13,355
Min	6,977	5,459	6,839
Max	76,971	79,924	82,952

Subsection 6.5.3 JTW times increased traffic scenarios with only travel and space preference

The results of JTW times produce similar averages between original increased traffic and these runs (Table 6.5-3 and Table 6.4-3). When compared to their respective baseline runs, the increased traffic scenarios have a similar relative difference for the median, minimum, and maximums. This is surprising as with only housing space to consider outside of travel time one would expect a greater adjustment to increased travel times by households in the model. As this simulation runs to steady state it shows that the previous runs might stay the same even if demographics were adjusted dynamically during the runs. It also underscores the importance of simply having a variable to compete for household resources.

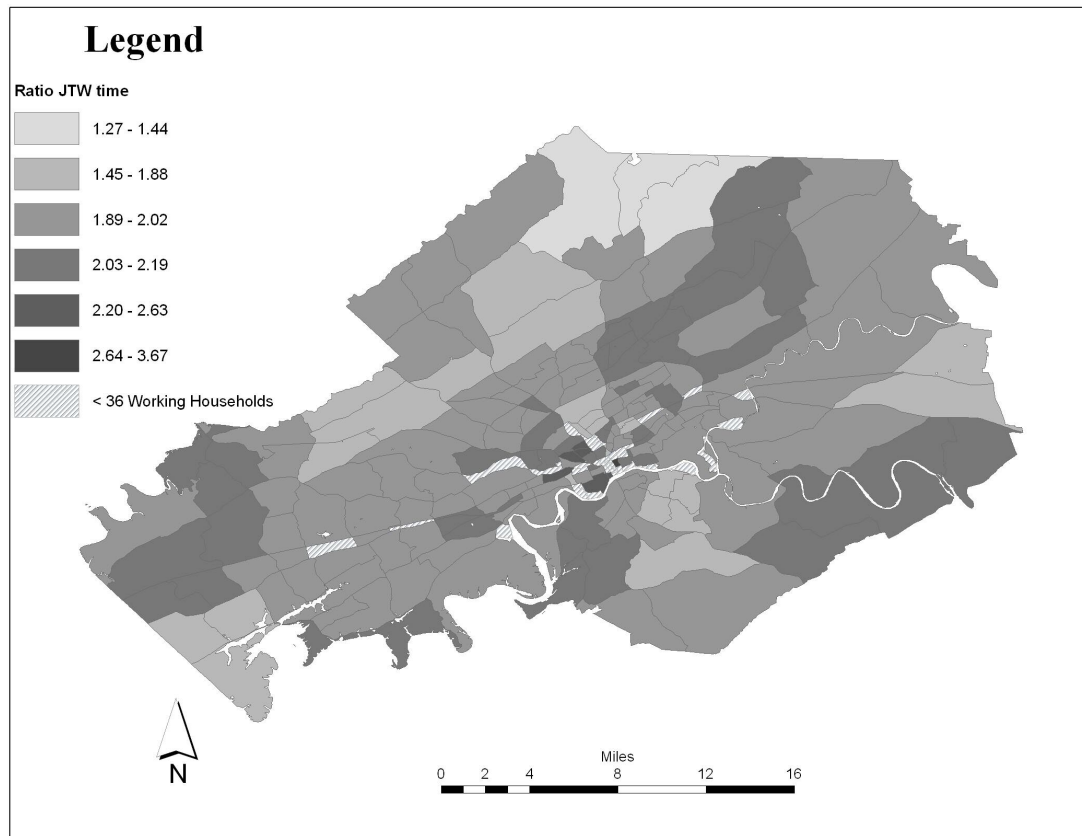
.

Table 6.5-3 JTW time comparisons: travel and space preferences with increased travel times

	Baseline run (T & S preference)	50% more traffic	Double travel times
Median	14	21	27
S.D.	7	7	13
Ave All HH	17.53	22.97	33.47
Min	6	7	11
Max	43	43	84

The relative change in the ‘double travel time only travel and space preference’ are similar to the original double travel time increase (Fig 6.5-2 Fig 6.4-5). These results again indicate that the urban structure is important to individual ‘neighborhood’ commuting times. This is true even if the model is run to an equilibrium state. However urban development not modeled would presumably alter urban structure to fit changing tastes of urban residents and ultimately neighborhood JTW times.

Figure 6.5-1 Ratio JTW time: time and space preference (2*travel time vs. baseline)



The last three runs point to the importance of the value of the transportation preference. The no-travel preference showed a large effect on overall JTW times. In the scenarios where the preference was only included with travel space the overall effect was minimal. Even when increased travel times were combined with only travel and space preference the relative behavior is the same. The next section will examine the sensitivity of overall JTW times when only travel preference value is changed for household agents.

6.6 Measuring the Effect of Changing Travel Preference

As a further measure of the effect of travel time preference on JTW times two additional runs were made multiplying the importance of original travel preference by 0.2 and 2 in the baseline run. This gives four travel preferences 0, 0.2, 1, and 2 times the baseline run (Table 6.6-1). One can see that the 0 and 0.2 preferences both have a large impact on travel times while the 2 times and original JTW time preference have

very similar travel times. The change just between the 0.2 and 0 travel preferences is about 5 minutes, 5 times larger than the difference between the original preference and double travel preference runs. This indicates the powerful effect of reducing this preference on JTW times particularly as the values get very small. This further indicates the importance of having a preference even if small on agent behavior in a complex system. As Knoxville has shorter travel times the same run in a place with high congestion such as San Francisco or Los Angeles might show different results.

Table 6.6-1 JTW times and Travel Preference Change

Preference Level	JTW Time Minutes
0. * Travel Preference	28.17
0.2* Travel Preference	23.15
1 * Travel Preference	17.85
2 * Travel Preference	16.46

Chapter 7: Summary

This model has been designed so it can be implemented with available data sources and can be run on existing desktops. This is the first attempt that the author is aware of that integrates housing choice and the transportation network with this degree of specificity. In addition, it is one of the few residential simulations that use individual household preferences as the basis for location choice. It is the only one that combines agents and utility functions in a systems modeling approach run on real data for an urban area.

There are several important findings of this model. The first goal was to test the ability of agent models using separable utility functions to model housing price, JTW times, and location choice. It is found that it can replicate real results and mimics some of the process of household decisions. While there was some systematic bias such as income levels in the CBD this was done successfully. Additionally the scenarios provide sensible relationships of variable change giving more credence to the model. While inferential statistical methods are valuable, this work frees quantitative analysis from that approach which relies on unrealistic assumptions of independent variables, no feedback, and simplified spatial and behavioral interaction, while still having a check of model parameters using empirical data. While this model has much in common with pure analytical models, this method can account for multiple feedback effects and individual disaggregated behaviors. It must be reiterated that direction and magnitude of results are the important outcomes of this model not slight variations. This is the case with many social and physical models.

The second purpose of the research is examining JTW times and preferences using simulations. A significant finding is that a dramatic reduction in the JTW time preference will exert a much greater influence on JTW time than a large increase in this preference in Knox County. This would indicate that a change in attitudes toward commuting by itself would, in practice, only have a large effect on JTW times if residents lessened their disutility for commuting. Along the same lines only the travel

and space preference produced the same average times in the county although there was significant change by TAZ and population movements. Even when combined with travel time increases the relative overall time changes were similar. As the housing space was inferred from housing data statistics and calibrated to fit census data its importance is probably not extremely overestimated pointing to the importance of having any preference competing for resources in this kind of complex system. A last interesting result is the lack of effect of all but the no-travel preference on price per square foot which flattened prices in the county and by TAZ. The other scenarios increased travel time. Travel and space preferences, and the combination of the two, had large effects on population, income, and JTW times for TAZs but very little impact on price.

There are some limitations in the model some of which may be addressed, others intrinsic to the approach taken in this research. First, as stated in the introduction its greatest strength is its greatest weakness: the precision of the utility function. It is unlikely that households are as precise in their evaluations as to think in such mathematical terms. The way the present model is set up only square feet can be purchased; this probably cannot be extended to discrete housing units. Lastly it is difficult to time step from year to year as the model runs to long-term equilibrium.

Limitations which can be addressed are data limitations like housing price, crime, school quality, generating households, and better transportation data. Additions to the model could include a lack of perfect information, a moving location cost, more transportation modes, tenure choice, differing working schedules, adding non-workers, inclusion of neighborhood attachment, life course effects like a change in family size. These would create more realistic urban processes than the present model. Other components could be added to the model such as a travel demand model or urban housing development.

The scenario runs have focused on JTW preference and times. The model can be used to examine other aspects of urban development such as residential segregation, or

housing market dynamics with changing incomes or housing supply. The model can be applied to larger scenarios as data and computational power is increased. More exotic applications could use results to model disease spread, which would require a travel demand model taking advantage of Dynasmart's ability to track individual movements through the road network. Evacuations can also be modeled using this approach. Ultimately this modeling approach can provide many benefits in urban development.

BIBLIOGRAPHY

- Abraham, J. E., & Hunt, J. D. (1997). *Specification and estimation of nested logit model of home, workplaces, and commuter model choices by multiple-worker households*. Washington D.C.
- Alonso, W. (1964). *Location and Land Use*. Cambridge: Harvard University Press.
- Anas, A., Arnott, R., Small, K. (1998). Urban Spatial Structure. *Journal of Economic Literature*, 36(3), 1426-1464.
- Andres, D. (2001). *Suburban Nation*. New York: North Point Press.
- Barrett, C., Berkigler, K., Smith L., Loose, V., Beckman, R., Davis, J. Roberts, D., Williams, M. (1995). An operational description of TRANSIMS. Los Alamos: Los Alamos National Laboratory
- Batty, M., Torrens, P.M. (2001). Modeling complexity: the limits to prediction, *CASA Working Paper* London: CASA.
- Baxandall, R., Ewen, E. (2000). *Windows: How the Suburbs Happened*. New York: Basic Books.
- Ben-Akiva, M., Bowman, J. (1998). Integration of an activity-based model system and a residential location model. *Urban Studies*, 37(7), 1231-1253.
- Ben-Akiva, M., Bielaire, M., Koutsopoulos, H., Yang, Q. (1998). Investigation of route Guidance generation issues by simulation with DynaMIT. In Ceder (Ed.), *Transportation and Traffic Theory, Proceedings of the 14 International symposium on transportation and Traffic Theory*. New York: Pergamon Press.
- Benenson, I. (2004). Agent-Based Modeling: From Individual Residential Choice to Urban Residential Dynamics. In M. F. Goodchild, Janelle, D. G. (Ed.), *Spatially Integrated Social Science: Examples in Best Practice* (pp. 67-95). New York: Oxford University Press.
- Benenson, I., Torrens, P.M. (2004). Geosimulation: object-based modeling of urban phenomena. *Computers, Environment and Urban Systems*, 28, 1-8
- BTS. (1990). 1990 census transportation planning package: Tennessee, urban parts 1,2: Department of Transportation

- Burgess, E. W. (1925). The growth of the city: An introduction to a research project. In R. Park, Burgess, E.W., and McKenzie. R.D. (Ed.), *The City* (pp. 47-62). Chicago: University of Chicago Press.
- Census. (2003). American Community Survey: US Census.
- Clark, W. A. V. (1986). Residential Segregation in American Cities. *Population Research and Policy Review*, 5, 95-127.
- Clark, W. A. V. (1989). Residential Segregation in American Cities: Common Ground Differences in Interpretation. *Population Research and Policy Review*, 7, 113-121.
- Clark, W. A. V. (1991). Residential preference and Neighborhood Racial Segregation: A test of the Schelling Segregation Model. *Demography*, 28, 1-19.
- Cropper, L., Gordon, P. (1991). Wasteful commuting: A reexamination. *Journal of Urban Economics*, 29, 2-13.
- Danielsen, K., Lang, R., Fulton, W. (1999). Retracting Suburbia: Smart Growth and the Future of Housing. *Housing Policy Debate*, 10, 513-540.
- Davis, M. (1990). *City of quarts: excavating the future in Los Angeles*. New York: Vintage.
- Emerson, M., Chai, K., Yancey, G. (2001). Does Race Matter in Residential Segregation? Exploring the Preferences of White Americans. *American Sociological Review*, 66(6), 922-935.
- Farley, R. (1977). Residential Segregation in Urbanized Areas of the United States in 1970. *Demography*, 14, 497-518.
- Forrester, J. (1969). *Urban Dynamics*. Cambridge: M.I.T. Press.
- Giuliano, G., Small, K. (1991). Subcenters in the Los Angeles Region. *Regional Science and Urban Economics*, 21(2), 163-182.
- Hamilton, B., Roell, A. (1982). Wasteful Commuting. *The Journal of Political Economy*, 90(5), 1035-1053.
- Hamilton, B. (1989). Wasteful Commuting Again. *The Journal of Political Economy*, 97(6), 1497-1504.
- Harris, C., Ullman, L. (1945). The Nature of Cities. *Annals of the American Academy*

- of Political and Social Science*, 242, 7-17.
- Harvey, D. (1973). *Social Justice and the City*. Baltimore: Johns Hopkins press.
- Heikkila, E., Gordon, P., Kim J. I., Peiser R. B., Richardson H. W., Dale-Johnson D. (1989). What happened to the CBD-distance gradient?: Land values in a polycentric city. *Environment and Planning A*, 21(2), 221-232.
- Hoyt, H. (1939). *The Structure and Growth of Residential Neighborhoods in American Cities*. Washington D.C.: Federal Housing Administration.
- Hu, T., and Mahmassani, H. (1995). Evolution of Network Flows under Real-Time Information: Day to day dynamic simulation assignment framework. *TRB*, 1493, 46-56.
- Kadanoff, L. P. (1971). From Simulation Model to Public Policy: An Examination of Forrester's Urban Dynamics. *American Scientist*, 60(1), 74-79.
- Lavine, J. (2005). *Zoned Out Regulation, Markets, and Choices in Transportation and Metropolitan Land Use*. Washington D.C.: RTF Press.
- Massy, D., Denton, N. (1988). The Dimensions of Residential Segregation. *Social Forces*, 67(2), 281-315.
- Massy, D., Denton, N. (1993). *American Apartheid, Segregation and the Making of the Underclass*. Cambridge: Harvard University Press.
- Mokhtarian, P., Salomon, I. (1999). Travel for the Fun of It. *Access*, 15, 26-31.
- Molotch, H. (1976). The City as a Growth Machine: Toward a Political Economy of Place. *American Journal of Sociology*, 82(September), 309-332.
- Muth, R. F. (1969). *Cities and Housing*. Chicago: University of Chicago Press.
- Myers, D., Gearin, E. (2001). Current Preferences and Future Demand for Denser Residential Environments. *Housing Policy Debate*, 12, 633-659
- Nagel, K., Rickert, M. (2001). Parallel implementation of the TRANSIMS micro-simulation. *Parallel Computing*, 27(12), 1611-1639.
- Openshaw, S. (1978). An empirical study of some zone design criteria. *Environment and Planning A*, 9, 169.
- Owen, L. E., Zhang, Y., Rao, L., McHale, G. (2000). *Traffic Flow Simulation Using Corsim*. Paper presented at the Proceedings of the 2000 Winter Simulation Conference.

- Quadstone. (2000). Paramics modeler V3.0 reference manual. Edinburgh, U.K.
- Reardon, S., O'Sullivan, D. (2004). Measures of spatial segregation. *Sociological Methodology*, 34, 121-164.
- Redmond, L., Mokhtarian P. (2001). The positive utility of the commute: modeling ideal commute time and relative desired commute amount. *Transportation*, 28(2), 179-205.
- Sasaki, Y., Box, P. (2003). Agent-Based Verification of von Thunen's Location Theory. *JASS*, 6(2).
- Sheppard, E. (2001). Quantitative geography: representations, practices, and possibilities. *Environment and Planning D*, 19, 535-554.
- Taeuber, K. (1965). The Effect of Income Redistribution on Racial Residential Segregation. *Urban Affairs Quarterly*, 4, 5-14.
- Timmermans, H. (2003). The Saga of Integrated Land Use-Transport Modeling: How Many More Dreams Before we Wake up, *10th International Conference on Travel Behavior Research Lucerne*. Lucerne.
- Torrens, P. M. (2001). New Tools for Simulating Housing Choices, *Housing and the New Economy*. Washington D.C. : Fannie Mae.
- Waddell, P. (1993). Exogenous Workplace Choice in Residential Location Models: Is the Assumption Valid. *Geographical Analysis*, 25, 65-81.
- Waddell, P., Ulfarsson, G. (2003). Dynamic Simulation of Real Estate Development and Land Prices within an Integrated Land Use and Transportation Model System." *TRB Annual Meeting*
- White, M. (1988). Urban Commuting Journeys Are Not "Wasteful. *The Journal of Political Economy*, 96(5), 1097-1110.
- Zubrinisky, C., Bobo, L. (1996). Prismatic Metropolis: Race and Residential Segregation in the City of the Angels. *Social Forces*, 25, 335-374.