SPATIAL ECONOMETRIC MODELING OF COMMUTING TIMES

THESIS

Presented to the Graduate Council of
Texas State University-San Marcos
in Partial Fulfillment
of the Requirements

for the Degree

Master of Business Administration

by

Dustin K. Kirby

San Marcos, Texas
May 2008

SPATIAL ECONOMETRIC MODELING OF COMMUTING TIMES

Committee Members Approved:

James LeSage, Chair

Eric Blankmeyer

David Beckworth

Approved:

J. Michael Willoughby
Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to thank my thesis chair, Dr. James LeSage, for all his assistance throughout this project. I have benefited tremendously from his knowledge and his guidance has been instrumental in my academic development.

I'd also like to thank my wife, Angela Kirby, whose sacrifice and support has made this entire endeavor possible. Many other family members have also been an important source of support during my studies. Thanks to my parents, Gene and Sharon Kirby, to my grandparents, John and Elaine Atkinson, and to my brother Neil and my sister Laura Kirby.

Finally, I'd like to thank all the other faculty members of the McCoy College of Business Administration for their guidance and encouragement throughout my master's studies. A special thanks to Dr. Gail Zank, Dr. Janet Parish, and Dr. Vishag Badrinarayanan, who have each been wonderful mentors.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

Research by Reschovsky (2004) has shown that the number of people with travel to work times of greater than 45 minutes increased from 12.5 percent of the population in 1990, to approximately 15.4 percent by the year 2000. After standardizing on the percentage of workers with at least 45 minute one-way commutes in 1990, this translates to an increase in long commutes of 10.39 percent over this period. Reschovky's work is supported by the findings of McGuckin and Srinivasan (2003) who analyzed the distribution of workers' commuting times in 1980, 1990 and 2000. Some of their findings are presented in Table 1.1, which gives the percentage of workers with travel to work times of less than 15 minutes, 15 to 44 minutes, and 45 minutes or more across the ten largest metropolitan areas (MSA), as well as across an aggregate of the largest 49 MSAs in the United States. It should be noted that the figures presented by McGuckin and Srinivasan, as well as those by Reschovsky, do not include individuals that work at home, which accounted for 3 percent of workers in 1990 and 3.3 percent in 2000. Nevertheless, a clear break with previous trend is visible during the period from 1990 to 2000. Whereas, from 1980 to 1990, the percentage of workers in the ten largest metro areas with one-way commutes of 45 minutes or more tended to increase only slightly and in some cases actually decreased, from 1990 to 2000, the percentage of workers with long commutes increased across each of the ten largest metro areas. These results can be extended to the average of the largest 49 metro areas, where we find that the percentage of workers with at least a 45 minute travel time to work increased from 14.5 to 15.3, over the period from 1980 to 1990, but then increased to 18.5 percent by 2000. These findings suggest that there was a substantial break in commuting time trend over this period, the causes of which will be the focus of this thesis.

Table 1.1: The percentage distribution of workers' commuting times 1980, 1990 and 2000

MSA Name	< 15			15 - 44			45 or more		
	1980	1990	2000	1980	1990	2000	1980	1990	2000
New York	23.7	23.3	20.2	49.4	50.7	50.1	26.9	26.1	29.7
Los Angeles	27.7	24.5	22.4	58.5	57.5	57.5	13.9	18.0	20.1
Chicago	25.1	23.7	21.3	53.7	54.6	53.8	21.2	21.6	24.9
Washington, DC	20.9	20.0	17.7	59.2	58.5	57.0	19.9	21.5	25.2
San Francisco	27.4	25.9	22.2	57.7	60.6	56.0	14.9	13.5	21.8
Philadelphia	27.9	27.8	24.7	56.1	57.3	56.3	16.0	14.9	19.0
Boston	34.4	30.3	25.6	53.4	55.2	54.8	12.2	14.5	19.7
Detroit	27.5	26.9	24.5	61.7	61.5	60.3	10.8	11.6	15.2
Dallas	27.0	24.6	22.4	61.6	62.0	59.9	11.3	13.4	17.7
Houston	23.3	22.7	21.0	58.6	60.1	58.9	18.0	17.3	20.1
49 MSAs	27.9	26.2	23.5	57.7	58.5	58.1	14.5	15.3	18.5
Source: McGuckin and Srinivasan (2003)									

Our study, however, is not the first to address the causes of this break in commuting time trend between 1990 and 2000. This was also the subject of an analysis by Gordon, Lee, and Richardson (2004), which looked at long commuting times using data from the National Personal Transportation Study and National Household Travel Survey. They very astutely recognized that growth in the number of workers with long commutes during the 1995 to 2000 period coincided with a period of rapidly rising household income levels. They concluded that the income shock of the late 1990s led to a situation where the land and housing market flexibility, which had previously helped to control longer commuting times, failed to respond rapidly enough to new housing location decisions. They further argued the affluence of the late 1990s led households to consume larger homes and more land at distant exurban locations, which overwhelmed the urban growth accommodation system that had worked well in the past. This assertion is predicated upon urban economics literature on the dynamic nature of cities and how flexible land markets have the ability to offset congestion caused by urban growth. More specifically, research by Gordon, Kumar, and Richardson (1989), as well as Crane and Chatham (2003), suggests that suburbanization enhances suburb-to-suburb commuting tendancies, which results in fewer long commutes than would be observed in a more traditional suburb-to-central business district commuting pattern. Given these implications, one could interpret Gordon, Lee, and Richardon's argument as evidence that the suburbanization of employment and residential location decisions in the U.S. has increased to the point that location of the manufacturing and services employment in exurban and rural areas is a consequence of the continued weakening of the agglomeration economies that shaped the now outdated downtown-oriented city.

In addition to their conclusions, we must note several important features of the empirical analysis performed by Gordon, Lee and Richardson. Their analysis utilized a cross-sectional sample, from 1990 and 2000, of Census Bureau Urbanized Areas corresponding to MSAs/CMSAs with a population of over 500,000 in the year 2000. The use of MSA and CMSA data, which represent aggregations of counties, has the well-known shortcoming of obscuring a great deal of variation in travel to work times that exist for those residing at various locations throughout the metropolitan area. Furthermore, such a metropolitan area analysis ignores commuting time patterns occurring in more rural locations which could account for a substantial portion of American travel times to work. To remedy some of these issues, in our study we will model the number of workers in each continental United States' census tract with travel to work times of 45 minutes or more in 1990 and 2000 as a function of tract-level characteristics of the resident population. One contribution of this study is the fact that our analysis takes place at the fine spatial scale of census tracts. Another is the spatial regression methodology employed that allows us quantify the direct (own-tract) influence of tract-level characteristics on longer travel to work times versus indirect (spatial spillover) effects that influence commuting times of residents in nearby tracts.

In the chapters that follow, we will begin by setting forth a simplified example that motivates the inclusion of spatial spillovers into our commuting time model. This example will also serve as an introduction to spatial econometric techniques to be used in our study. We will then turn our attention to research methodology which will include a discussion of our spatial durbin commuting time model and its associated direct, indirect, and total effects estimates. In Chapter 4, we will specify the various demographic, housing location, and control variables used in our study and establish relationships between each of these variables and the number of workers with at least 45 minute travel times to work in each of 1990 and 2000. These relationships will then be used to identify significant changes that took place over this time period. We will conclude our analysis by examining these differences and drawing conclusions on the underlying causes of the break in commuting time trend.

CHAPTER 2

A SIMPLIFIED EXAMPLE

2.1 Spatial spillovers in commuting time models

To motivate an analysis of the effects of spatial dependence on commuting times, we will begin our study with a simplified example. This example will serve as an introduction to spatial econometric models that will be used throughout the remainder of our analysis.

Figure 2.1 shows seven regions located from west to east along a single highway.

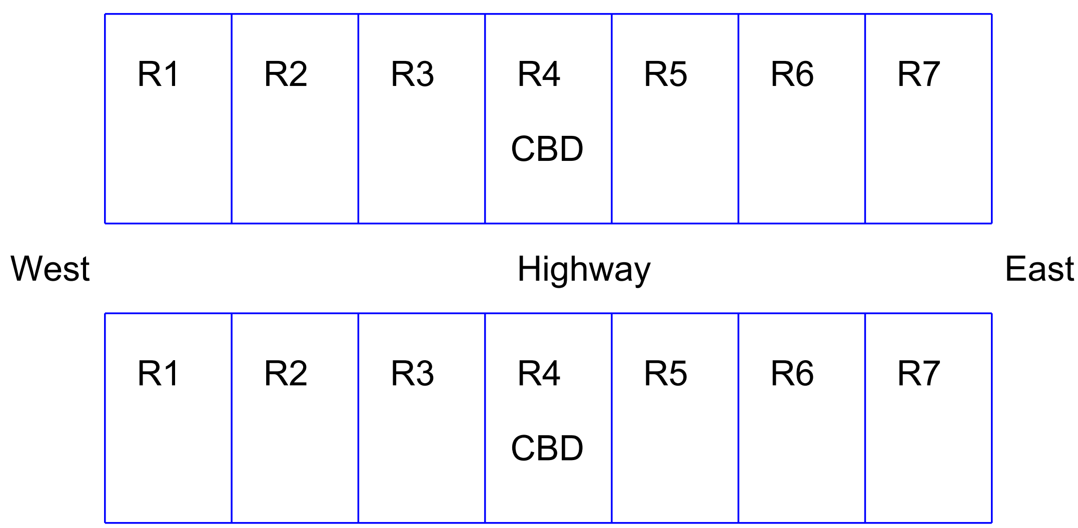

Figure 2.1: Seven regions along a highway

For the purpose of this example, we will consider these seven regions to constitute a single metropolitan area, with region #4 being the central business district. Since the entire metro area contains only a single roadway, all commuters will share this route to and from the central business district (CBD).

Since all commuters share this single roadway, the average commuting time to the CBD from region #2 will clearly be impacted by the average commuting time from region #3 to the CBD, as increased congestion occurring in region #3 will add to the commuting times of those traveling from region #2. Research by Arnott, de Palma, and Lindsey (1990) also suggests that average commuting times in region #2 could be impacted by commuting times in region #1, as an increase in region #1 commuting times resulting from additional commuters in that region will place greater demands on the roadway connecting region #2 to the CBD.

Given the dependence associated with commuting times across our metropolitan area, any ordinary least-squares regression attempting to explain variation in regional commuting times would be both biased and inconsistent as it represents a violation of the Gauss-Markov assumption that explanatory variables are fixed in repeated sampling [LeSage and Pace (2004)]. Therefore, any valid attempts to explain variation in commuting times among our seven regions must account for the area's commuting time dependence structure.

One such method of incorporating dependence is the spatial autoregressive process shown in (2.1).

$$y = \alpha\iota_n + \rho W y + \varepsilon \tag{2.1}$$

For the case of our simplified example, y would represent a 7×1 vector containing our dependent variable, commuting times, for the seven regions and W would represent a 7×7 spatial weight matrix. Accordingly, ρ would be a scalar parameter that relates the product of Wy to our dependent variable. The 7×1 constant term vector ι_n and the corresponding parameter α are included for instances where our commuting time vector y does not have a mean value of zero. The 7×1 vector ε is included to account for disturbances, which are assumed to have a normal distribution with a mean of zero.

The aforementioned W matrix allows us to incorporate a dependence structure into our model by quantifying the connectivity among regions in our metro area. For the purpose of this example, our W matrix will only take into account first-order neighbors, that is, those regions which share a border with the given region.

To construct our W matrix, we begin with a 7×7 binary indicator matrix which we shall call matrix P. Row n of matrix P corresponds to region #n of our metro area, that is, row 2 corresponds to region #2 and so forth. Within a given row, a 1 is inserted to indicate neighbors of the given region. For example, row 5 of matrix P would contain a 1 in column 4 and column 6, to indicate that region #5 borders regions #4 and #6. All non-neighbor positions within our matrix are then set to 0. It is important to note that the main diagonal of matrix P is set to zero to prevent a region from neighboring itself. Binary indicator matrix P is shown in (2.2).

$$P = \begin{array}{cc} & \begin{array}{ccccccc} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{array} \\ \begin{array}{c} R1 \\ R2 \\ R3 \\ R3 \\ R5 \\ R6 \\ R7 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array} \tag{2.2}$$

With the binary indicator matrix complete, we can then normalize matrix P to have row-sums of unity by dividing the entries of each row by the number of neighbors associated with that particular region. The result is a spatial weight matrix, previously referred to as matrix W and shown in (2.3).

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{2.3}$$

By constructing a spatial weight matrix in such a manner, the product of matrix W and the y vector yields a 7×1 vector containing an average of neighbors commuting times. This product is referred to as a spatial lag vector and is shown in (2.4). The spatial lag vector, in turn, allows us to incorporate spatial dependence among observations into our model and ultimately derive accurate parameter estimates.

$$Wy = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} = \begin{pmatrix} y_2 \\ 1/2y_1 + 1/2y_3 \\ 1/2y_2 + 1/2y_4 \\ 1/2y_3 + 1/2y_5 \\ 1/2y_4 + 1/2y_6 \\ 1/2y_5 + 1/2y_7 \\ y_6 \end{pmatrix} \tag{2.4}$$

The previously discussed spatial autoregressive process can be introduced to a standard regression model to produce a spatial autoregressive (SAR) model shown below in (2.5), along with the associated data generating process shown in (2.6).

$$y = \rho W y + X\beta + \varepsilon \tag{2.5}$$

$$y = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1}\varepsilon \tag{2.6}$$

Matrix X is composed of explanatory variables, which for the purpose of our example will be a 7×2 matrix containing vectors for each region's distance to the CBD and population density measure. Although not utilized in our example, matrix X could also incorporate a constant term vector akin to ι_n discussed in (2.1). X is shown in (2.7).

$$X = \begin{array}{cl} \begin{array}{cc} \text{Density} & \text{Distance} \end{array} & \\ \begin{pmatrix} 10 & 30 \\ 20 & 20 \\ 30 & 10 \\ 50 & 0 \\ 30 & 10 \\ 20 & 20 \\ 10 & 30 \end{pmatrix} & \begin{array}{c} \text{ex-urban areas} \\ \text{far suburbs} \\ \text{near suburbs} \\ \text{CBD} \\ \text{near suburbs} \\ \text{far suburbs} \\ \text{ex-urban areas} \end{array} \end{array} \tag{2.7}$$

Vector β is composed of the associated regression parameters, which for our example will be a 2×1 vector. As previously discussed, Wy represents a 7×1 spatial lag vector composed of the average commuting times of each region's neighbors. The scalar parameter ρ indicates the strength of the spatial lag vector's impact on our regression results, alternatively referred to as spatial dependence. It is worthwhile to note that when ρ takes on the value of zero, ρWy collapses to a zero vector and the SAR model simplifies to an ordinary least-squares regression. As in (2.1), ε represents a disturbance vector, for our example a 7×1, assumed to have a normal distribution with a mean of zero.

We should also point out that our discussion thus far has revealed no reason why the parameter ρ should be restricted to a specific range. However, it follows from (2.6) that the variance-covariance of our SAR model regression is

$E[(I_n - \rho W)^{-1} \varepsilon \varepsilon' (I_n - \rho W)^{-1\prime}]$. Thus, $(I_n - \rho W)^{-1}$ must be invertible and the variance-covariance matrix, represented by $(I_n{-}\rho W)^{-1}(I_n{-}\rho W)^{-1}$, must be positive-definite, a requirement of any regression model. Since the W matrix has row-sums of unity, it is a sufficient condition for a positive-definite variance-covariance matrix that $-1 < \rho < 1$. Therefore, all further analyses will limit ρ to this range.

2.2 Simultaneous feedback and steady-state equilibrium

In our example, we noted that growth in the number of commuters in region #1 could place an additional burden on the metro area's single roadway and thereby impact the commuting times of those in region #2. It seems logical that the increased commuting times of those in region #2 could then spillover to those commuting to the CBD from region #3. Given the SAR model data generating process in (2.6), we can use the infinite series representation for the matrix inverse $(I_n - \rho W)^{-1}$ set forth by Debreu and Herstein (1953), which is shown in (2.8), to quantify some of the spillover by expanding the SAR model data generation process as shown in (2.9).

$$\begin{aligned} (I_n - \rho W)^{-1} &= I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 ... \\ y &= (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1} \varepsilon \end{aligned} \tag{2.8}$$

$$\begin{aligned} y &= X\beta + \rho W X\beta + \rho^2 W^2 X\beta + ... \\ &+ \varepsilon + \rho W \varepsilon + \rho^2 W^2 \varepsilon + \rho^3 W^3 \varepsilon + ... \end{aligned} \tag{2.9}$$

Whereas the original W matrix reflected first-order contiguous neighbors, here the matrix W^2 will reflect second-order contiguous neighbors, that is neighbors of neighbors, the W^3 matrix will reflect third-order contiguous neighbors, and so forth.

It is also important to note that implicit in the SAR data generating process is the simultaneous nature of the data generation process, this follows as the expected value of each region's commuting time will depend on the mean value of $X\beta$ plus a linear combination of neighboring commuting times scaled by the ρ parameter.

When considering the simultaneous feedback implicit in our model, the point could be made that feedback is a latent process and, as such, may take time to manifest itself. To consider this possibility, we will relax the cross-sectional design of our previous equation and incorporate a space-time lag variable Wy_{t-1} into our model. In this model, y_t will denote commuting time to work at time t, which will depend upon current own-region explanatory variables, assumed to be stable over time which we reflect by setting $X_t = X$, and an average of neighboring regions' commute times in period $t-1$. This model is shown in (2.10).

$$y_t = \rho W y_{t-1} + X\beta + \varepsilon_t \tag{2.10}$$

Since $y_{t-1} = \rho W y_{t-2} + X\beta + \varepsilon_{t-1}$ and $y_{t-2} = \rho W y_{t-3} + X\beta + \varepsilon_{t-2}$, we can incorporate q prior periods into our model through recursive substitution of past values. The resulting equation is shown in (2.11) and (2.12).

$$y_t = (I_n + \rho W + \rho^2 W^2 + ... + \rho^q W^q) X\beta + \rho^q W^q y_{t-q} + u \quad (2.11)$$

$$u = \varepsilon_t + \rho W \varepsilon_{t-1} + \rho^2 W^2 \varepsilon_{t-2} + ... + \rho^{q-1} W^{q-1} \varepsilon_{t-(q-1)} \quad (2.12)$$

Since $-1 < \rho < 1$, as $q \to \infty$, then $\rho^q W^q y_{t-q}$ approaches a zero vector as ρ decays the influence of higher order neighbors. Further, from (2.12), $E(u) = 0$ since $E(\varepsilon_{t-i}) = 0$ for $i = 0, 1, 2, ..., q$. It can then be shown that the limit as $q \to \infty$, which represents the steady-state equilibrium of this space-time model, is in fact equal to our original SAR data generating process. This simplification is shown in (2.13).

$$\begin{aligned} \lim_{q \to \infty} E(y_t) &= (I_n + \rho W + \rho^2 W^2 + ... + \rho^q W^q) X\beta + \rho^q W^q y_{t-q} + u \quad (2.13) \\ &= (I - \rho W)^{-1} X\beta \end{aligned}$$

Therefore, we will consider changes in regions' commute times to exert simultaneous feedback upon one another as their impact extends to higher order neighbors. Although changes in commute times will ultimately alter other regions' outcomes, we will view this feedback as occuring instantaneously and interpret this as a shift to a new steady-state equilibrium consistent with our static cross-sectional model.

2.3 Demonstration

With an understanding of the simultaneous feedback inherent in our static cross-sectional model, we can now demonstrate the impact of changes in region-level characteristics on commuting times across our metro area. Recall from (2.9), the SAR data generating process for our simple seven-region example, where y, our dependent variable, is a 7×1 vector composed of each region's commuting time, W is a 7×7 row-stochastic spatial weight matrix, and ρ is a scalar parameter indicating the strength of spatial dependence across the metro area. In this example, X represents a 7×2 explanatory variable matrix composed of vectors corresponding to each region's distance to the CBD and an arbitrary, albeit intuitively appealing, population density measure. Further, β represents a 2×1 vector containing estimated coefficients that relate our explanatory variables to our dependent variable, a vector of each regions average commuting time. Finally, ε is a $7{\times}1$ disturbance vector assumed to have a normal distribution with a mean of zero.

In the spirit of simplicity, we will assume the true value of ρ across our metro area is 0.7 and that the true values of β_1 and β_2, the coefficients that relate distance and population density to average commuting time, are 0.5 and 0.1, respectively. Given matrices W and X, we can use these parameters to generate a vector of average commuting times across our metro area where we set the statistical disturbances equal to their mean value of zero. This vector will be referred to as $y_{initial}$ and is shown in (2.14).

$$y_{initial} = \begin{pmatrix} 42.1222 \\ 37.3175 \\ 30.2134 \\ 26.1494 \\ 30.2134 \\ 37.3175 \\ 42.1222 \end{pmatrix} \tag{2.14}$$

To be consistent, we will consider these average commuting times to represent minutes. We note that as we move away from the CBD average commuting times increase and are symmetric across our metro area, reflecting the symmetry of the distance and population density vectors composing matrix X.

Now, lets assume that the population density in region #2 doubles in measure, from 20 to 40, and thereby alters our explanatory variable matrix X. Holding ρ and β constant, we can generate a new vector of average commuting times for our metro area. This vector will be referred to as y_{new} and is shown in (2.15).

$$y_{new} = \begin{pmatrix} 44.4099 \\ 40.5856 \\ 31.5488 \\ 26.6968 \\ 30.4421 \\ 37.4235 \\ 42.1964 \end{pmatrix} \tag{2.15}$$

Notice that as in our previous results, average commuting times increase as we move away from the CBD. However, as a result of the increase in region #2's population density, average commuting times increase in all regions, but at a greater rate as we move towards region #1. By comparing these estimates to our original results, other implications of the population density increase become apparent. A comparison table of these results along with net change in average commuting time for each region is shown in Table 2.1.

Table 2.1: Predictions of the impact from changing population density in $R2$

Scenario	$y_{initial}$	y_{new}	Difference: $y_{new} - y_{initial}$
Regions			
$R1$:	42.1	44.4	+2.2
$R2$:	37.3	40.5	+3.2
$R3$:	30.2	31.5	+1.3
$R4$: CBD	26.1	26.6	+0.5
$R5$:	30.2	30.4	+0.2
$R6$:	37.3	37.4	+0.1
$R7$:	42.1	42.1	+0.0

It is quickly noticed that although population density only increased in region #2, average commuting times increased across all regions in the metro area. With regard to these net changes, region #2 experienced the largest change in average commuting time, followed by its neighboring regions #1 and #3, respectively. Perhaps the most interesting feature of the population density increase in region #2 is that the effects of this increase were felt on the opposite side of the CBD in regions #5, #6, and #7, which experienced increases in average commute time that decay with distance from region #2. Even though residents of regions #5, #6, and #7 do not utilize the same portion of the roadway when commuting to the CBD or even neighbor region #2, the new steady-state equilibrium brings to them an increase in average commuting time.

It is also of interest that the cumulative indirect effects (spillovers) can be found by adding up the increased commute times across all other regions (excluding the own-region change in commuting time) This equals $2.2 + 1.3 + 0.5 + 0.2 + 0.1 + 0 = 4.3$ minutes, which is larger than the direct (own-region) effect of 3.2 minutes.

2.4 Implications

The nature of the results obtained in our previous demonstration were not unique to the parameter values or regional characteristics selected for our metro area. Rather, they stem from feedback exerted among higher order neighbors in a spatially dependent structure as the system instantaneously shifts to a new steady-state equilibrium.

This has implications not only for our simplified seven region metro area, but for any area experiencing spatial dependence as a result of shared roadways or other regional characteristics that we have not yet explored.

In the analyses that follow, spatial econometric techniques will be used to examine the impact of spatial dependence on commuting times, whereby our previous example of dependence among neighboring regions is extended to a cross-section of neighboring census tracts in the continental United States. Methods for introducing spatial dependence will be discussed and we will develop several summary measures for the own-tract (direct) as well as the spillover (indirect) impact of our explanatory variables.

CHAPTER 3

MODEL SPECIFICATIONS

3.1 Commuting time as a spatial durbin model

In additional to the spatial lag variable, it seems plausible that neighboring region characteristics, such as percentage of the population utilizing mass transportation, could play a significant role in explaining variation in a given region's commuting times. Thus, using census tracts as our observational units, we will employ a variation of the SAR model introduced in (2.5) in our analyses. Our model will incorporate neighboring regions' characteristics to account for any influence they may exert on their neighbors' commuting times. We will accomplish this by including the matrix product WX in our model, along with an $k \times 1$ parameter vector θ that will indicate the strength of this relationship for each of the k explanatory variables in matrix X. The result is a spatial durbin model (SDM), shown in (3.1), where the constant term vector ι_n has been removed from the explanatory variable matrix X.

$$\begin{aligned} y &= \rho W y + \alpha\iota + X\beta + WX\theta + \varepsilon \qquad (3.1) \\ \varepsilon &\sim N(0_{nx1}, \sigma^2 I_n) \end{aligned}$$

By using the SDM model, we allow commuting times to depend upon both own-region explanatory variables from the X matrix, as well as an average of the same explanatory variables from neighboring regions. In terms of our notation, since commuting time y_i depends upon its neighbor's commuting time, y_j, which in turn depends upon region j's explanatory variables x_{jr}, then changes in county j's explanatory variables will affect the dependent variable y_i. Using our simplified example, the average commuting time in region # 2, may depend not only upon the percentage of region # 2 residents that use mass transportation, but also an average of the percentage of region #1 and region # 3 residents that use mass transportation. We are able to isolate and quantify this effect by introducing WX as an explanatory variable in our model.

3.2 Interpretation of the commuting time model

Caution must be taken when interpreting the parameter vector β, as previous studies have shown that in cases where $\rho \neq 0$, interpretation of the SDM model differs from that of its standard non-spatial regression counterpart (Pace and LeSage, 2006). For ordinary regression models, the rth parameter from the vector β, β_r, is interpreted as representing the partial derivative of y with respect to change in the rth explanatory variable from the matrix X, which we will denote as x_r. In the case of a standard least-squares regression where the Gauss-Markov assumptions are satisfied, $y = \sum_{r=1}^{k} x_r \beta_r + \varepsilon$, and the partial derivative of y_i with respect to x_{ir} takes on the form of $\partial y_i / \partial x_{ir} = \beta_r$ for all i, r; and

$\partial y_i/\partial x_{jr} = 0$, for $j \neq i$ and all variables r. Operating under such assumptions, we find that $E(y_i) = \sum_{r=1}^{k} x_r\beta_r$. This interpretation hinges upon the independence assumption though as observation i depends only upon predetermined, exogenous variables.

However, when incorporating information from neighboring tracts, this interpretation no longer holds as this type of model expands the information set. That is, tract i's observed value is dependent on tract j, which is in turn dependent on tract i. To better quantify this interplay, consider the model from (3.1) expressed as:

$$\begin{aligned}
(I_n - \rho W)^{-1} y &= X\beta + WX\theta + \iota_n\alpha + \varepsilon \\
y &= \sum_{r=1}^{k} S_r(W)x_r + V(W)\iota_n\alpha + V(W)\varepsilon \\
S_r(W) &= V(W)(I_n\beta_r + W\theta_r) \\
V(W) &= (I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 ...
\end{aligned} \tag{3.2}$$

To better illustrate the role of $S_r(W)$, consider the expansion of the data generating process from (3.2) as shown in (3.3) [Pace and LeSage (2007)].

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{r+1}^{k} \begin{pmatrix} S_r(W)_{11} & S_r(W)_{12} & \dots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & & \\ \vdots & \vdots & \ddots & \\ S_r(W)_{n1} & S_r(W)_{n2} & \dots & S_r(W)_{nn} \end{pmatrix} \begin{pmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{nr} \end{pmatrix} \tag{3.3}$$

$$+ \quad V(W)\iota_n\alpha + V(W)\varepsilon$$

In the case where y_i is the only dependent variable observation, the data generating process simplifies as shown in (3.4).

$$\begin{aligned} y_i &= \sum_{r=1}^{k}[S_r(W)_{i1}x_{1r} + S_r(W)_{i2}x_{2r} + \ldots + S_r(W)_{in}x_{nr}] \\ &+ V(W)\iota_n\alpha + V(W)\varepsilon \end{aligned} \tag{3.4}$$

From (3.4), we can calculate the derivative of y_i with respect x_{jr}, which is shown in (3.5), and the derivative of y_i with respect to x_{ir}, as shown in (3.6).

$$\frac{\partial y_i}{\partial x_{jr}} = S_r(W)_{ij} \tag{3.5}$$

$$\frac{\partial y_i}{\partial x_{ir}} = S_r(W)_{ii} \tag{3.6}$$

Thus, for a set of spatially dependent observations and its associated explanatory variables, it is unlikely that the derivative of y_i with respect to x_{ir} would equal β_r, and that the derivative of y_i with respect to x_{jr}, for $j \neq i$, would equal 0. In other words, by changing the value of a single observation's explanatory variable, we alter all other observations' commuting times. Such a structure allows changes in tract characteristics to set forth a chain reaction, whereby the resulting changes in commuting times are extended throughout both lower and higher-order neighbors. This feature follows directly from our incorporation of an average of neighboring region characteristics in the commuting time model.

At this time, we should also discuss interpretation of the derivative of y_i with respect to x_{ir}, as shown in (3.6). Although $S_r(W)_{ii}$ expresses the impact on y_i derived from direct influences, that is changes in tract i's characteristics, it also includes changes derived from indirect influences These indirect influences are the result of impacts being passed to neighboring tracts, and eventually passed back to the census tract where they originated. In our study, the magnitude of such impacts will depend upon: (1) the position of the census tract in the spatial structure, (2) the degree of connectivity among tracts as quantified by the W matrix, (3) the ρ parameter which measures the strength of spatial dependence, and (4) the magnitude of the coefficient estimates for β and θ.

Given the differing impact of characteristic changes on other tracts' commuting times, it will be useful to separate these impacts into those originating from within the census tract itself and those emerging indirectly from neighboring tracts. Pace and LeSage (2006) have set forth the following scalar summary measures which we will use to average these impacts across all census tracts in the sample.

The *Average Direct effect* provides a summary measure that represents an average over all census tracts of the impacts on commuting times arising from own-tract changes in variable r. For example, if there is an increase in the number of tract i residents utilizing mass transportation, what will be the average impact on tract i's commuting time. It is important to note that this measure will include feedback effects resulting from the change in tract i's mass transportation rate on neighboring tracts' commuting times.

Another summary measure to be used in our analyses is the *Average Total effect*. Unlike other measures, the Average Total effect has two interpretations. An example of the first interpretation would be if all tracts see a commensurate increase in their mass transportation utilization rate, what will be the impact on the typical tract's commuting time. Whereas the second interpretation is illustrated by the example, if one census tract increases its mass transportation utilization, on average, what will be the cumulative impact on neighboring tracts' commuting times.

Finally, the *Average Indirect effect* is defined as the difference between the Average Total effect and the Average Direct effect. We can think of this measure as the average impact on a single tract's commuting time if all neighboring tracts saw an approximately equivalent increase in their mass transportation utilization rates.

Similar to the interpretation of standard least-squares parameter estimates, we must emphasize that these summary measures represent an average over all observations in our sample, from which micro-level impacts may vary. Although it is possible to examine the precise impact on tract i's commuting time as a result of changes in tract j's explanatory variables, such analyses would require analyzing vectors containing more than 61,000 observations for each of the 61,000 census tracts.

To make inferences regarding the statistical significance of the direct, indirect, and total effects estimates, we will utilize a Bayesian Markov Chain Monte Carlo estimation method set forth by LeSage (1997) which will provide a large number of draws for the model parameters, ρ, β, θ, α, and σ^2. Gelfand and Smith (1990) have shown that valid inferences can be obtained by using these draws in the non-linear functions of these parameters representing the partial derivatives in which we are interested. However, this necessitates evaluation and storage of these draws, against which we will apply variance calculations to obtain our posterior estimates.

CHAPTER 4

EMPIRICAL RESULTS

4.1 Data set and model specifications

Our analyses will utilize a data set obtainable via the U.S. Census Bureau which contains commuting time categories and explanatory variables for the years 1990 and 2000. We use all census tracts in the contiguous U.S. states plus the District of Columbia, a sample size of 64,111 tracts for the year 2000 and 63,932 tracts for 1990. The sample was constructed starting with 64,642 year 2000 tracts (64,326 year 1990) that had non-zero population and area. These samples were further restricted to tracts with a non-zero number of workers over 16 years of age, a non-zero number of workers traveling to work by car, truck, or van, a non-zero number of workers with travel to work times of less than 25 minutes or work at home, a non-zero number of workers with travel time to work of 25-44 minutes, and a non-zero number of workers with travel time to work more than 45 minutes. This resulted in sample sizes of 64,111 tracts for 2000 and 63,932 for 1990. The tracts were harmonized to the year 2000 census boundaries by Geolytics, Inc. Imposing these restrictions resulted in a small amount of sample censoring, 541 of the 64,642 year 2000 tracts and 394 of the 64,326 year 1990 tracts.

Although several commuting time categories are offered in the data set, we will use the natural log of the number of people with commute times in excess of 45 minutes as our dependent variable. Using this category will allow us to forego analyses of less intriguing commuting-time categories and focus only on the census tract characteristics which are positively or negatively associated with longer commutes. Furthermore, we believe that people with longer commute times should be more responsive to changes in the data set's explanatory variables.

Since more than 40 explanatory variables will be used in our model, to ease analysis we have grouped the variables into three distinct suites: demographic characteristics, housing location characteristics, and control variables. The demographic suite will contain variables related to the number of employed males and females in each census tract, as well as the number of people within various age and sex categories. Housing location characteristics will include variables related to the number of people in each tract that exhibit various metro residence preferences relative to their workplace. For example, the *Work in city* variable corresponds to the number of workers in the census tract that live in the metro area and work in the city. Other variables in the housing location suite relate to households' tenure in their current residence as well as the type of metro setting from which people have relocated in the past five years. The control suite consists of other variables unrelated to either demographic or household location decisions, which could also have a significant impact on the number of people in each census tract with long commutes. Examples of variables included in this suite are *Income*, *Population*, and several variables related tract members' preferred mode of transportation to work. A detailed list of all variables along with their definitions will be given in the coming sections.

In Section 4.2, SDM effects estimates will be presented for our suite of demographic variables in 1990 as well as 2000. These results will consist of the direct, indirect, and total effects associated with this suite of variables. Section 4.3 will then present effects estimates associated with our household location suite while section 4.4 will offer results associated with our control suite of variables. After the direct, indirect, and total effects estimates have been presented for each of the suites, section 4.5 will examine how these effects estimates changed between 1990 and 2000. Although effects estimates for the years 1990 and 2000 will help us understand the relationship that exists between our explanatory variables and the number of people with long commutes in each of those years, it is the changes in these relationships that will best explain the break in commuting trend that occurred during this period.

4.2 Demographic suite effects estimates

Table 4.1 contains a complete list of the demographic variables and definitions used in our model. In the analyses that follow we will examine the direct, total, and indirect effects associated with our suite of demographic variables for the years 1990 and 2000. Since we will be using the logged values of our dependent and explanatory variables, we will interpret the effects estimates as indicating how a percentage change in our explanatory variables will impact the percentage of people in the same or neighboring tracts with commutes in excess of 45 minutes. For example, if the direct effects estimate associated with *Employed females* is 0.5, we can infer that, all other things being equal, a 10 percent increase in the number of employed females in a census tract will result in, on average, a 5 percent increase in the number of people in that tract with commute times of at least 45 minutes. Likewise, an indirect effects *Employed females* estimate of 0.5 would indicate that a 10 percent employed females increase in a given tract tends to be associated with a 5 percent increase in the number of people with commute times longer than 45 minutes in neighboring tracts. Similarly, a total effects estimate of 0.5 associated with *Employed females* would imply that a 10 percent increase in the number of employed females in a tract tends to result in a 5 percent increase in the number of people in own and other tracts with long commutes. It must be stressed that all estimates represent an average across all census tracts in our data set, where all other variables are held constant. It is likely that actual commuting time changes in a specific tract will deviate from these estimated values.

Effects estimates for each of the explanatory variables will be accompanied by standard deviation, $t-$statistic, and $t-$probability measures to quantify the statistical significance of our estimates. These measures were obtained by using a Markov Chain Monte Carlo estimation procedure with 5000 draws to calculate the standard deviation from the posterior distribution of the effects estimates. For cases where the probability level associated with a variable is not significant we will be unable to reject the null hypothesis that the effects estimate is in fact different from zero. Accordingly, we will be unable to conclude that the variable in question has a significant direct/indirect/total impact on the number of people with commuting times of 45 minutes or greater.

Table 4.1: Definition of demographic suite variables

Labels used in tables	Definition
Employed females	Employed civilian females 16+ years old
Employed males	Employed civilian males 16+ years old
Females 10-14	Females 10-14 years old
Females 15-19	Females 15-19 years old
Females 20-24	Females 20-24 years old
Females 25-29	Females 25-29 years old
Females 30-34	Females 30-34 years old
Females 35-44	Females 35-44 years old
Females 45-54	Females 45-54 years old
Females 55-64	Females 55-64 years old
Males 10-14	Males 10-14 years old
Males 15-19	Males 15-19 years old
Males 20-24	Males 20-24 years old
Males 25-29	Males 25-29 years old
Males 30-34	Males 30-34 years old
Males 35-44	Males 35-44 years old
Males 45-54	Males 45-54 years old
Males 55-64	Males 55-64 years old

4.2.1 Demographic suite direct effects estimates

Table 4.2 contains direct effects estimates that quantify the impact of our demographic variables on excessive commuting times for the years 1990 and 2000. To ease analysis, the direct effects estimates have been placed in ascending order with the greatest negative value occurring in the first row and the greatest positive value occurring in the last row of the table.

An examination of Table 4.2 reveals the *2000 Males 20-24* variable, with a value of approximately -0.0149, has the largest negative direct effects estimate of all the variables within the demographic suite. This value indicates that for the year 2000, a 100 percent increase in the number of males aged 20-24 in a census tract would tend to be associated with a 1.49 percent decrease in the number of workers with at least a 45 minute commute in that tract. This variable is followed by *2000 Females 15-19* and *2000 Males 45-54*, with direct effects estimates of approximately -0.0097 and -0.0067, respectively. However, the $t-$probability estimates associated with these variables would not allow us to conclude that the associated direct effects estimates are significantly different from zero at the 95 percent confidence level.

Among the variables with positive direct effects estimates, the *2000 Employed males* variable exhibits the greatest magnitude in the demographic suite with a value of 0.3764. The next largest positive value is associated with the *1990 Employed males* variable

Table 4.2: Sorted demographic direct effects estimates for 1990 and 2000

Variable	Direct Effect	Standard Deviation	$t-$stat	$t-$prob
2000 Males 20-24	-0.0149	0.0056	-2.6603	0.0078
2000 Females 15-19	-0.0097	0.0061	-1.5867	0.1126
2000 Males 45-54	-0.0067	0.0104	-0.6434	0.5200
1990 Males 20-24	-0.0064	0.0072	-0.8995	0.3684
1990 Females 15-19	-0.0054	0.0072	-0.7512	0.4525
1990 Males 15-19	-0.0051	0.0072	-0.7053	0.4807
1990 Females 20-24	-0.0004	0.0080	-0.0498	0.9603
2000 Males 30-34	0.0002	0.0074	0.0260	0.9793
2000 Females 10-14	0.0018	0.0062	0.2968	0.7666
2000 Males 10-14	0.0019	0.0062	0.3038	0.7613
2000 Males 15-19	0.0019	0.0061	0.3139	0.7536
1990 Females 45-54	0.0041	0.0100	0.4094	0.6822
1990 Females 10-14	0.0044	0.0067	0.6557	0.5120
1990 Males 10-14	0.0074	0.0067	1.1145	0.2651
1990 Males 55-64	0.0108	0.0091	1.1914	0.2335
2000 Males 25-29	0.0117	0.0064	1.8398	0.0658
2000 Males 55-64	0.0128	0.0082	1.5696	0.1165
1990 Females 55-64	0.0141	0.0068	2.0847	0.0371
2000 Females 20-24	0.0143	0.0058	2.4560	0.0141
1990 Males 30-34	0.0167	0.0088	1.8907	0.0587
2000 Females 45-54	0.0263	0.0112	2.3554	0.0185
2000 Females 55-64	0.0285	0.0083	3.4351	0.0006
1990 Males 25-29	0.0328	0.0080	4.0961	0.0000
2000 Females 25-29	0.0356	0.0070	5.0633	0.0000
2000 Males 35-44	0.0436	0.0105	4.1419	0.0000
1990 Females 30-34	0.0442	0.0096	4.6188	0.0000
2000 Females 30-34	0.0478	0.0081	5.9314	0.0000
1990 Males 35-44	0.0519	0.0116	4.4827	0.0000
1990 Females 35-44	0.0569	0.0127	4.4672	0.0000
1990 Females 25-29	0.0658	0.0088	7.5098	0.0000
1990 Males 45-54	0.0739	0.0084	8.7706	0.0000
2000 Females 35-44	0.0931	0.0119	7.8019	0.0000
2000 Employed females	0.1061	0.0143	7.4177	0.0000
1990 Employed females	0.1306	0.0275	4.7510	0.0000
1990 Employed males	0.2742	0.0328	8.3598	0.0000
2000 Employed males	0.3764	0.0142	26.5969	0.0000

with a value of approximately 0.2742. These values suggest that over the decade from 1990 to 2000, the number of employed males within a census tract may have had an increasing influence on the number of tract residents with 45 minute commutes. The *1990 Employed females* variable exhibited the next greatest positive direct effects estimate, though the value of this estimate was less that half that of *1990 Employed males.* Unlike the variables associated with negative direct effects estimates, of which only one was significant at the 95 percent confidence level, nearly half of the demographic suite direct effects estimates were positive and significant. This finding, in conjunction with the magnitude of the positive estimates, may suggest that the demographic suite as a whole tended to be associated with increases of own-tract commute times of at least 45 minutes.

4.2.2 Demographic suite indirect effects estimates

Table 4.3 presents indirect effects estimates for the demographic suite of variables in both 1990 and 2000. As in the previous table, these results have been sorted with the greatest negative value occurring in the first row and the greatest positive value occurring in the final row of the table.

An examination of the table reveals the *1990 Employed males* variable exhibits the largest negative magnitude with a value of approximately -0.9802. This estimate indicates that, in 1990, a ceteris paribus 10 percent increase in the number of employed males in a tract tended to be associated with a 9.8 percent decrease in the number of people with long commutes in neighboring tracts. The *2000 Employed females* and *1990 Males 55-64*

Table 4.3: Sorted demographic indirect effects estimates for 1990 and 2000

Variable	Indirect Effect	Standard Deviation	$t-$stat	$t-$prob
1990 Employed males	-0.9802	0.1382	-7.0919	0.0000
2000 Employed females	-0.2845	0.0609	-4.6741	0.0000
1990 Males 55-64	-0.1925	0.0522	-3.6855	0.0002
2000 Employed males	-0.1551	0.0652	-2.3795	0.0173
1990 Females 20-24	-0.1272	0.0440	-2.8902	0.0039
2000 Males 45-54	-0.1205	0.0607	-1.9864	0.0470
1990 Males 30-34	-0.0766	0.0511	-1.4992	0.1338
2000 Females 15-19	-0.0662	0.0343	-1.9304	0.0536
1990 Females 45-54	-0.0652	0.0580	-1.1240	0.2610
2000 Males 30-34	-0.0583	0.0429	-1.3602	0.1738
2000 Females 10-14	-0.0560	0.0345	-1.6225	0.1047
1990 Females 35-44	-0.0413	0.0724	-0.5696	0.5690
2000 Females 45-54	-0.0365	0.0634	-0.5755	0.5649
2000 Males 10-14	-0.0233	0.0350	-0.6656	0.5057
1990 Females 10-14	-0.0154	0.0381	-0.4055	0.6851
2000 Females 25-29	0.0007	0.0401	0.0175	0.9860
1990 Males 10-14	0.0085	0.0379	0.2253	0.8217
2000 Males 15-19	0.0139	0.0341	0.4068	0.6842
2000 Males 20-24	0.0182	0.0320	0.5682	0.5699
1990 Females 55-64	0.0231	0.0343	0.6736	0.5006
2000 Females 55-64	0.0323	0.0448	0.7207	0.4711
1990 Males 15-19	0.0334	0.0421	0.7935	0.4275
1990 Males 20-24	0.0444	0.0399	1.1129	0.2657
2000 Males 25-29	0.0581	0.0372	1.5637	0.1179
1990 Females 15-19	0.0662	0.0417	1.5868	0.1126
1990 Males 35-44	0.0673	0.0654	1.0280	0.3039
1990 Males 25-29	0.0902	0.0462	1.9535	0.0508
1990 Females 30-34	0.1013	0.0551	1.8387	0.0660
2000 Females 20-24	0.1059	0.0323	3.2789	0.0010
2000 Males 55-64	0.1217	0.0484	2.5143	0.0119
2000 Females 30-34	0.1221	0.0465	2.6251	0.0087
2000 Males 35-44	0.1239	0.0584	2.1230	0.0338
1990 Females 25-29	0.1805	0.0495	3.6454	0.0003
1990 Males 45-54	0.3406	0.0457	7.4530	0.0000
2000 Females 35-44	0.4170	0.0678	6.1514	0.0000
1990 Employed females	0.6305	0.1197	5.2679	0.0000

variables exhibited the next largest negative magnitudes with values of -0.2845 and -0.1925, respectively. Although increases in either of these variables wouldn't confer the same decrease in the number of people in neighboring tracts with lengthy commutes as *1990 Employed males*, it is worth noting that these variables would still be associated with a significant decrease in the number of neighboring tract residents with long commutes.

Among the demographic variables with positive indirect effects estimates, the *1990 Employed females* variable exhibits the largest positive magnitude with a value of 0.6305. Such a value indicates that, as of 1990, a ceteris paribus 10 percent increase in the number of employed females in a tract tended to be associated with a 6.3 percent increase in the number of people in neighboring tracts with long commutes. Similarly, the values associated with the *2000 Females 35-44* and *1990 Males 45-54* variables, 0.4170 and 0.3406 respectively, indicate that increases in either of these variables would have a positive and significant impact on the number of people in neighboring tracts with long commutes. Again, it is worth noting that the number of demographic variables exhibiting positive and significant indirect effects estimates is nearly double that of the number of variables with negative and significant estimates. Also of interest, we find that the magnitudes of the greatest indirect effects comfortably exceed the magnitudes of the greatest direct effects estimates.

4.2.3 Demographic suite total effects estimates

Demographic suite total effects estimates for 1990 and 2000 are shown in Table 4.4. Again, the estimates have been placed in ascending order with the greatest negative magnitude occurring in the first row and the greatest positive magnitude occurring in the final row of the table.

Since total effects are by definition the sum of the direct and indirect effects, we would expect to see a great deal of similarity between the demographic suite total effects estimates and the direct and indirect effects which we have already discussed. This is indeed the case, however, since the magnitudes associated with the indirect effects greatly exceed those of the direct effects, the order of total effects magnitudes tends to be much more reflective of the indirect effects estimates. For example, we see that the demographic variable with the greatest negative magnitude total effects estimate is *1990 Employed males*, just as it was in the indirect effects. However, if we recall, the *1990 Employed males* variable had the second largest direct effect estimate. Thus, although the increases in the number of employed males in a census tract tended to increase the number of people in that tract with long commutes, the associated decrease in neighboring tract residents with long commutes is so much greater that, in total, increases in this variable are associated with a significant decrease in the total number of people with at least 45 minute travel times to work. To be more specific, the total effects estimate of -0.7059 indicates that, in 1990, a 10 percent increase in the number of employed males in a census tract tended to be associated with a more than 7 percent decrease in the number of people with

Table 4.4: Sorted demographic total effects estimates for 1990 and 2000

Variable	Total Effect	Standard Deviation	$t-$stat	$t-$prob
1990 Employed males	-0.7059	0.1710	-4.1279	0.0000
1990 Males 55-64	-0.1817	0.0613	-2.9643	0.0030
2000 Employed Females	-0.1783	0.0752	-2.3726	0.0177
1990 Females 20-24	-0.1276	0.0520	-2.4550	0.0141
2000 Males 45-54	-0.1272	0.0711	-1.7896	0.0735
2000 Females 15-19	-0.0758	0.0404	-1.8784	0.0603
1990 Females 45-54	-0.0611	0.0679	-0.8993	0.3685
1990 Males 30-34	-0.0600	0.0599	-1.0006	0.3170
2000 Males 30-34	-0.0581	0.0502	-1.1572	0.2472
2000 Females 10-14	-0.0542	0.0407	-1.3314	0.1831
2000 Males 10-14	-0.0214	0.0412	-0.5205	0.6027
1990 Females 10-14	-0.0111	0.0448	-0.2469	0.8050
2000 Females 45-54	-0.0102	0.0746	-0.1366	0.8914
2000 Males 20-24	0.0033	0.0376	0.0876	0.9302
1990 Females 35-44	0.0157	0.0852	0.1840	0.8541
2000 Males 15-19	0.0158	0.0402	0.3926	0.6946
1990 Males 10-14	0.0160	0.0446	0.3582	0.7202
1990 Males 15-19	0.0283	0.0493	0.5744	0.5657
2000 Females 25-29	0.0363	0.0471	0.7704	0.4411
1990 Females 55-64	0.0373	0.0411	0.9065	0.3647
1990 Males 20-24	0.0379	0.0470	0.8065	0.4200
2000 Females 55-64	0.0608	0.0531	1.1450	0.2522
1990 Females 15-19	0.0608	0.0489	1.2423	0.2141
2000 Males 25-29	0.0698	0.0435	1.6042	0.1087
1990 Males 35-44	0.1192	0.0770	1.5476	0.1217
2000 Females 20-24	0.1201	0.0381	3.1534	0.0016
1990 Males 25-29	0.1230	0.0542	2.2699	0.0232
2000 Males 55-64	0.1346	0.0566	2.3778	0.0174
1990 Females 30-34	0.1455	0.0647	2.2501	0.0244
2000 Males 35-44	0.1675	0.0689	2.4312	0.0151
2000 Females 30-34	0.1699	0.0546	3.1129	0.0019
2000 Employed males	0.2213	0.0793	2.7896	0.0053
1990 Females 25-29	0.2463	0.0583	4.2264	0.0000
1990 Males 45-54	0.4145	0.0541	7.6580	0.0000
2000 Females 35-44	0.5101	0.0797	6.3984	0.0000
1990 Employed females	0.7611	0.1472	5.1714	0.0000

long commutes in own and neighboring tracts. Other demographic variables for which the total effects estimates were negative and significant at the 95 percent confidence level include *1990 Males 55-64*, *2000 Employed females*, and *1990 Females 20-24.*

Much as we saw with the negative total effects estimates, the demographic variables exhibiting positive total effects are very similar in order of magnitude to our indirect estimates. The total effects estimate of 0.7611 associated with the *1990 Employed females* variable indicates that, in 1990, a 10 percent increase in the number of employed females in a tract tended to be associated with a 7.6 percent increase in the number of people in own and neighboring tracts with long commutes. However, given the positive and significant direct and indirect estimates, we could conclude that such an increase would tend to be associated with increases in long commutes in both the origin tract and its neighbors. The *1990 Employed females* variable is followed by *2000 Females 35-44*, *1990 Males 45-54*, and *1990 Females 25-29* variables in order of positive magnitude, all of which are significant at the 99 percent confidence level. Again, it is worth noting that the demographic variables with positive and significant total effects estimates outnumber those that are negative and significant at the 95 percent confidence level at a ratio of more than 2 to 1.

4.3 Housing location suite effects estimates

A complete list of the housing location variables and definitions used in our model is presented in Table 4.5. In the analyses that follow, we will examine the direct, indirect,

Table 4.5: Definition of housing location suite variables

Labels used in tables	Definition
Work at home	Workers 16+ that work at home
Work in city	Workers 16+ in the metro area, work in the city
Work in suburbs	Workers 16+ in the metro area, work in suburbs
Work in ocity	Workers 16+ in the metro area, work in the city, another metro area
Work in osuburbs	Workers 16+ in the metro area, work in suburbs, another metro area
Work in non-metro	Workers 16+ that work in non-metro areas
Moved from city	Persons 5+ moved from central city, same metro area last 5 years
Moved from ocity	Persons 5+ moved from central city, another metro area last 5 years
Moved from nmetro	Persons 5+ moved from non-metro area last 5 years
Moved from suburbs	Persons 5+ moved from suburbs, the same metro area last 5 years
Moved in 1 yr	Housing units where the householder moved in during the last year
Moved in 2-5 yrs	Housing units where hhdr. moved in during the last 2-5 years
Moved in 6-10 yrs	Housing units where hhdr. moved in during the last 6-10 years
Moved in 11-19 yrs	Housing units where hhdr. moved in during the last 11-19 years
Moved in 20-29 yrs	Housing units where hhdr. moved in during the last 20-29 years
Moved in 30+ yrs	Housing units where hhdr. moved in over 30 years ago

and total effects associated with this suite of variables for the years 1990 and 2000. As in our previous analyses, since we will be using logged values of the dependent/explanatory variables, all estimates can be interpreted as elasticity measures. These effects estimates will again be accompanied by standard deviation, $t-$statistic, and $t-$probability measures obtained from a Markov Chain Monte Carlo estimation procedure using 5000 draws.

4.3.1 Housing location suite direct effects estimates

Table 4.6 contains direct effects estimates that quantify the impact of our housing location variables on excessive commuting times for each of the years 1990 and 2000. To ease analysis, the direct effects estimates have been placed in ascending order with the greatest negative value given in the first row and the greatest positive value given in the final row of the table.

Table 4.6: Sorted housing location direct effects estimates for 1990 and 2000

Variable	Direct Effect	Standard Deviation	$t-$stat	$t-$prob
1990 Moved in 30+ yrs	-0.0253	0.0029	-8.8647	0.0000
1990 Moved from city	-0.0253	0.0043	-5.8750	0.0000
2000 Moved from city	-0.0250	0.0035	-7.1001	0.0000
2000 Moved in 30+ yrs	-0.0239	0.0025	-9.6333	0.0000
1990 Moved from suburbs	-0.0239	0.0039	-6.1265	0.0000
1990 Moved from ocity	-0.0211	0.0034	-6.1384	0.0000
2000 Moved from ocity	-0.0186	0.0031	-6.0043	0.0000
2000 Moved from suburbs	-0.0161	0.0036	-4.4572	0.0000
1990 Moved from nmetro	-0.0112	0.0027	-4.1102	0.0000
1990 Moved in 20-29 yrs	-0.0071	0.0038	-1.8668	0.0619
2000 Moved in 6-10 yrs	-0.0031	0.0066	-0.4722	0.6368
2000 Work at home	-0.0026	0.0025	-1.0191	0.3081
2000 Moved from non-metro	-0.0013	0.0022	-0.6174	0.5370
1990 Work at home	0.0016	0.0027	0.5803	0.5617
2000 Moved in 20-30 yrs	0.0062	0.0037	1.6988	0.0894
1990 Moved in 6-10 yrs	0.0074	0.0071	1.0333	0.3015
2000 Moved in 11-19 yrs	0.0092	0.0054	1.6907	0.0909
1990 Moved in 11-19 yrs	0.0134	0.0059	2.2828	0.0224
2000 Work in non-metro	0.0148	0.0018	8.1525	0.0000
2000 Moved in 1 yr	0.0164	0.0074	2.2080	0.0272
2000 Work in suburbs	0.0201	0.0041	4.9338	0.0000
1990 Work in non-metro	0.0247	0.0021	11.5062	0.0000
1990 Moved in 1 yr	0.0299	0.0082	3.6657	0.0002
1990 Work in suburbs	0.0337	0.0043	7.9028	0.0000
2000 Work in ocity	0.0484	0.0019	25.4953	0.0000
2000 Work in osuburbs	0.0496	0.0018	27.2445	0.0000
1990 Work in city	0.0518	0.0053	9.7863	0.0000
1990 Work in ocity	0.0519	0.0021	24.2134	0.0000
2000 Work in city	0.0529	0.0046	11.6133	0.0000
1990 Work in osuburbs	0.0590	0.0020	28.8100	0.0000
1990 Moved in 2-5 yrs	0.0814	0.0101	8.0250	0.0000
2000 Moved in 2-5 yrs	0.0828	0.0098	8.4725	0.0000

An examination of Table 4.6 reveals that the *1990 Moved in 30+ years* and *1990 Moved from city* variables exhibit the greatest negative magnitudes with a value of -0.0253. The first result indicates that in 1990 a 100 percent increase in the number of people within a census tract that moved into their current residence over 30 years ago tended to be associated with a 2.5 percent decrease in the number of people in that tract with commute times of greater than 45 minutes. Similarly, the second result suggests that, in 1990, a 100 percent increase in the number of tract residents that relocated from the central city of the same metro area within the past five years tended to be associated with a 2.5 percent decrease in the number of people in that tract with long commutes. A number of other housing location variables exhibited comparable direct effects estimates which were significantly negative. These variables include *2000 Moved from city*, *2000 Moved in 30+ years*, and *1990 Moved from suburbs.*

Table 4.6 also offers a number of variables which exhibited positive and significant direct effects estimates during this period. Among them, the *2000 Moved in 2-5 years* variable displayed the greatest positive magnitude with a value of approximately 0.0828. This estimate indicates that, all other things being equal, a doubling of the number of people within a census tract that moved-in between two and five years ago tended to be associated with an 8.28 percent increase in the number of people in that tract with commute times of at least 45 minutes. The *1990 Moved in 2-5 years* variable exhibited a similar relationship with a value of approximately 0.0814. Thus, we could infer that, in 1990, a 100 percent increase in the number of people within a tract that moved-in two to five years ago tended to be associated with an 8.14 percent increase in own-tract residents with long commutes. The *2000 Work in city* and *1990 Work in city* variables also displayed approximately equivalent direct effects estimates which were relatively large in magnitude. Other housing location suite members with large positive direct effects estimates include the *1990 Work in osuburbs* and *1990 Work in ocity* variables.

4.3.2 Housing location suite indirect effects estimates

Table 4.7 presents indirect effects estimates for the housing location suite of variables in both 1990 and 2000. As in previous tables, these results have been sorted with the greatest negative value occurring in the first row and the greatest positive value occurring in the final row of the table.

An examination of Table 4.7 reveals that the *2000 Moved in 1 yr* variable offers the largest negative indirect effects estimate with a value of approximately -0.1782. This estimate indicates that in 2000 a 10 percent increase in the number of tract residents that moved-in within the last year tended to be associated with a 1.78 percent increase in the number of people in neighboring tracts with at least a 45 minute commute. This result is followed in magnitude by the *1990 Moved from ocity* and *1990 Moved in 30+ years* variables with values of approximately -0.1721 and -0.1628, respectively. These results are then followed by the *2000 Moved from ocity* variable which was relatively close to its 1990 value. Although relatively large estimates associated with the same variable in both 1990 and 2000 may offer some insight as to the underlying causes of long commute

Table 4.7: Sorted housing location indirect effects estimates for 1990 and 2000

Variable	Indirect Effect	Standard Deviation	$t-$stat	$t-$prob
2000 Moved in 1 yr	-0.1782	0.0371	-4.8073	0.0000
1990 Moved from ocity	-0.1721	0.0147	-11.6713	0.0000
1990 Moved in 30+ yrs	-0.1628	0.0122	-13.3154	0.0000
2000 Moved from ocity	-0.1624	0.0139	-11.6718	0.0000
2000 Moved in 6-10 yrs	-0.1490	0.0350	-4.2603	0.0000
2000 Moved from city	-0.1487	0.0129	-11.4953	0.0000
2000 Moved in 30+ yrs	-0.1391	0.0103	-13.4765	0.0000
1990 Moved from city	-0.1334	0.0161	-8.2606	0.0000
1990 Work in non-metro	-0.0980	0.0082	-11.9873	0.0000
2000 Work in non-metro	-0.0890	0.0071	-12.6174	0.0000
1990 Moved from nmetro	-0.0890	0.0123	-7.2418	0.0000
1990 Work in suburbs	-0.0409	0.0157	-2.6035	0.0092
1990 Moved in 20-29 yrs	-0.0301	0.0181	-1.6632	0.0963
2000 Moved from suburbs	-0.0269	0.0135	-1.9915	0.0464
2000 Moved from nmetro	-0.0232	0.0097	-2.3936	0.0167
1990 Moved in 6-10 yrs	-0.0226	0.0367	-0.6161	0.5378
2000 Work in suburbs	-0.0167	0.0133	-1.2532	0.2101
1990 Moved in 11-19 yrs	-0.0118	0.0277	-0.4278	0.6688
1990 Moved from suburbs	-0.0109	0.0158	-0.6906	0.4898
1990 Work in ocity	0.0092	0.0085	1.0757	0.2820
2000 Work in ocity	0.0163	0.0079	2.0645	0.0390
2000 Moved in 20-30 yrs	0.0235	0.0184	1.2813	0.2001
2000 Work in osuburbs	0.0273	0.0063	4.3733	0.0000
1990 Moved in 1 yr	0.0276	0.0407	0.6786	0.4974
2000 Moved in 11-19 yrs	0.0292	0.0270	1.0812	0.2796
1990 Work in osuburbs	0.0511	0.0071	7.2377	0.0000
1990 Work in city	0.0575	0.0160	3.5875	0.0003
2000 Work at home	0.0598	0.0125	4.7888	0.0000
1990 Work at home	0.0658	0.0129	5.1038	0.0000
1990 Moved in 2-5 yrs	0.0737	0.0501	1.4697	0.1417
2000 Work in city	0.0975	0.0130	7.4928	0.0000
2000 Moved in 2-5 yrs	0.3457	0.0519	6.6605	0.0000

times, since the estimates' values changed little between these years it wouldn't help explain the break in commuting time trend.

Table 4.7 also provides a number of housing location variables with positive indirect effects estimates over this period. Among them, the *2000 Moved in 2-5 years* variable offers the largest positive estimate at a value of approximately 0.3457. Such an estimate indicates that in the year 2000, a ceteris paribus 10 percent increase in the number of tract residents that moved in within the last two to five years tended to be associated with a 3.457 percent increase in the number of residents in neighboring tracts with long commutes. Following the *2000 Moved in 2-5 years* variable, the indirect effects estimates exhibit a pronounced decrease in magnitude. The *2000 Work in city* variable has the second largest magnitude, but is only estimated at 0.0975. This result is followed by the *1990 Moved in 2-5 yrs* variable, however, this estimate is not significant at even the 90 percent confidence level.

4.3.3 Housing location suite total effects estimates

Housing location suite total effect estimates for 1990 and 2000 are shown in Table 4.8. Again, the estimates have been placed in ascending order with greatest negative magnitude occurring in the first row and the greatest positive magnitude occurring in the final row of the table.

As we saw with our demographic variables, since the total effects estimates are the sum of the direct and indirect effects, there should exist a great deal of similarity between

Table 4.8: Sorted housing location total effects estimates for 1990 and 2000

Variable	Total Effect	Standard Deviation	$t-$stat	$t-$prob
1990 Moved from ocity	-0.1932	0.0182	-10.6234	0.0000
1990 Moved in 30+ yrs	-0.1881	0.0151	-12.4735	0.0000
2000 Moved from ocity	-0.1811	0.0170	-10.6386	0.0000
2000 Moved from city	-0.1737	0.0165	-10.5534	0.0000
2000 Moved in 30+ yrs	-0.1629	0.0128	-12.7328	0.0000
2000 Moved in 1 yr	-0.1618	0.0445	-3.6339	0.0003
1990 Moved from city	-0.1586	0.0204	-7.7590	0.0000
2000 Moved in 5-10yrs	-0.1521	0.0416	-3.6607	0.0003
1990 Moved from nmetro	-0.1002	0.0150	-6.6738	0.0000
2000 Work in non-metro	-0.0742	0.0089	-8.3596	0.0000
1990 Work in non-metro	-0.0733	0.0103	-7.0957	0.0000
2000 Moved from suburbs	-0.0430	0.0171	-2.5104	0.0121
1990 Moved in 20-29 yrs	-0.0372	0.0219	-1.6986	0.0894
1990 Moved from suburbs	-0.0348	0.0197	-1.7697	0.0768
2000 Moved from nmetro	-0.0245	0.0119	-2.0701	0.0384
1990 Moved in 6-10 yrs	-0.0152	0.0438	-0.3479	0.7279
1990 Work in suburbs	-0.0072	0.0200	-0.3594	0.7193
1990 Moved in 11-19 yrs	0.0016	0.0336	0.0471	0.9624
2000 Work in suburbs	0.0034	0.0174	0.1959	0.8447
2000 Moved in 20-30 yrs	0.0298	0.0221	1.3508	0.1768
2000 Moved in 11-19 yrs	0.0383	0.0324	1.1834	0.2367
2000 Work at home	0.0572	0.0150	3.8087	0.0001
1990 Moved in 1 yr	0.0575	0.0489	1.1772	0.2391
1990 Work in ocity	0.0611	0.0107	5.7140	0.0000
2000 Work in ocity	0.0646	0.0098	6.6143	0.0000
1990 Work at home	0.0674	0.0156	4.3186	0.0000
2000 Work in osuburbs	0.0769	0.0081	9.5296	0.0000
1990 Work in city	0.1093	0.0213	5.1279	0.0000
1990 Work in osuburbs	0.1101	0.0091	12.0915	0.0000
2000 Work in city	0.1504	0.0176	8.5605	0.0000
1990 Moved in 2-5 yrs	0.1550	0.0603	2.5724	0.0101
2000 Moved in 2-5 yrs	0.4286	0.0617	6.9477	0.0000

the total effects estimates and our previous housing location findings. This is indeed the case as the *1990 Moved from ocity* variable exhibits the greatest negative magnitude total effects estimate with a value of approximately -0.1932. If we recall, this same variable was associated with the second greatest indirect effects estimate. Nevertheless, the total effects estimate of -0.1932 indicates that in the year 1990, a 10 percent increase in the number of tract residents that moved from the central city of another metro area in the last five years, tended to be associated with more than a 1.9 percent decrease in the number of workers in own and neighboring tracts with long commutes. This estimate is followed in magnitude by the *1990 Moved in 30+ years* and *2000 Moved from ocity* variables with estimates of -0.1881 and -0.1811, respectively. It is worth noting that for each of these variables, the direct and indirect effects estimates were significantly negative and relatively large in magnitude.

An examination of Table 4.8 also reveals that the *2000 Moved in 2-5 years* variable again exhibits the largest positive magnitude. The total effects estimate of 0.4286 associated with this variable indicates that in the year 2000, a ceteris paribus 10 percent increase in the number of tract residents that moved-in within the past two to five years tended to be associated with more than a 4.2 percent increase in the number of people in own and neighboring tracts with one-way commutes of greater than 45 minutes. From this variable there is again a pronounced decrease in the magnitude of total effects estimates, with the *1990 Moved in 2-5 years* and *Work in city* variables exhibiting estimates of 0.1550 and 0.1504, respectively. Of these two, the *1990 Moved in 2-5 years* isn't significant at the

Table 4.9: Definition of control suite variables

Labels used in tables	Definition
Pop	1990 or 2000 population (logged)
Land	Land area in square miles
Water	Water area in square miles
Public trans	Workers 16+ using public transportation
Walking	Workers 16+ walking (or other means)
No car	Households with no car
Three cars	Households with 3 cars
Income	Median household income
College	persons 25+ with college or graduate/professional degrees
Renters	Total renter-occupied housing units

99 percent confidence level. It is also worth noting that the number of housing location variables with negative and significant total effects estimates slightly outnumbers the number of variables whose estimates are positive and significant.

4.4 Control suite effects estimates

A complete list of the control variables and definitions used in our model is presented in Table 4.8. In the analyses that follow we will examine the direct, indirect, and total effects associated with this suite of variables for the years 1990 and 2000. Although in previous analyses we were able to interpret all estimates as elasticity measures, we must exercise caution when interpreting the *Land*, *Water*, and *Income* variables as they don't represent logged values. All control variable effects estimates will again be accompanied by standard deviation, $t-$statistic, and $t-$probability measures obtained from a Markov Chain Monte Carlo estimation procedure using 5000 draws.

Table 4.10: Sorted control direct effects estimates for 1990 and 2000

Variable	Direct Effect	Standard Deviation	$t-$stat	$t-$prob
1990 Renters	-0.0727	0.0066	-11.0776	0.0000
2000 College	-0.0722	0.0049	-14.6399	0.0000
1990 College	-0.0605	0.0050	-12.0589	0.0000
1990 Income	-0.0601	0.0081	-7.3789	0.0000
2000 Renters	-0.0533	0.0053	-9.9838	0.0000
2000 Income	-0.0463	0.0073	-6.3095	0.0000
1990 Walking	-0.0367	0.0034	-10.8444	0.0000
2000 Walking	-0.0228	0.0026	-8.8755	0.0000
2000 Three cars	-0.0010	0.0046	-0.2199	0.8260
1990 Three cars	0.0004	0.0050	0.0774	0.9383
2000 No car	0.0109	0.0037	2.9320	0.0034
1990 No car	0.0245	0.0039	6.3330	0.0000
1990 Water	0.0333	0.0046	7.2438	0.0000
2000 Water	0.0393	0.0042	9.4640	0.0000
1990 Land	0.0486	0.0024	20.0071	0.0000
2000 Public trans	0.0530	0.0018	29.9574	0.0000
2000 Land	0.0533	0.0022	24.0594	0.0000
1990 Public trans	0.0563	0.0020	27.8689	0.0000
1990 Pop	0.0980	0.0203	4.8356	0.0000
2000 Pop	0.1053	0.0176	5.9993	0.0000

4.4.1 Control suite direct effects estimates

Table 4.10 contains direct effects estimates that quantify the impact of our control variables on excessive commuting times for the years 1990 and 2000. To ease analysis, the direct effects estimates have been placed in ascending order with the greatest negative value given in the first row and the greatest positive value appearing in the final row of the table.

An examination of Table 4.10 reveals that a number of control variables displayed negative direct effects estimates during this period. The *1990 Renters* variable exhibits the greatest negative magnitude estimate with a value of approximately -0.0727. This estimate indicates that, in 1990, a ceteris paribus 100 percent increase in the number of renters in a census tract tended to be associated with a 7.27 percent decrease in the number of tract residents with a commute time of more than 45 minutes. This result is followed in magnitude by the *2000 College* and *1990 College* variables with direct effects estimates of -0.0722 and -0.0605, respectively. Again, though the relatively large magnitude of these variables in 1990 and 2000 may help us understand how they relate to longer commutes, since they don't differ dramatically from one another, it doesn't explain the break in commuting time trend.

Several other variables from this suite offer positive and significant direct effects estimates during 1990 and 2000. Among them, the *2000 Pop* variable exhibits the largest positive estimate with a value of approximately 0.1053. This result indicates that, in the year 2000, a 10 percent population increase within a given tract tended to be associated with just over a one percent increase in the number of workers in that tract with long commutes. This result was followed in magnitude by the *1990 Pop* and *1990 Public trans* variables with estimates of 0.0980 and 0.0563, respectively. It is worth noting that a slightly greater number of our control variables displayed significantly positive direct effects estimates than those that were negative and significant.

Table 4.11: Sorted control indirect effects estimates for 1990 and 2000

Variable	Indirect Effect	Standard Deviation	$t-$stat	$t-$prob
2000 Renters	-0.1641	0.0226	-7.2510	0.0000
1990 Renters	-0.1519	0.0274	-5.5393	0.0000
1990 Walking	-0.1372	0.0151	-9.0810	0.0000
2000 Walking	-0.0823	0.0130	-6.3418	0.0000
2000 Three cars	-0.0815	0.0177	-4.6019	0.0000
2000 College	-0.0730	0.0192	-3.7928	0.0001
1990 Three cars	-0.0485	0.0180	-2.6885	0.0072
1990 College	-0.0359	0.0204	-1.7568	0.0790
2000 Water	0.0236	0.0135	1.7494	0.0802
1990 Water	0.0354	0.0151	2.3453	0.0190
1990 Land	0.0927	0.0086	10.8247	0.0000
2000 Land	0.1026	0.0075	13.6583	0.0000
2000 No car	0.1097	0.0187	5.8713	0.0000
1990 Public trans	0.1344	0.0077	17.4669	0.0000
2000 Public trans	0.1556	0.0069	22.5220	0.0000
1990 No car	0.1970	0.0178	11.0396	0.0000
2000 Income	0.1979	0.0358	5.5268	0.0000
1990 Income	0.3097	0.0403	7.6851	0.0000
2000 Pop	0.3283	0.0765	4.2894	0.0000
1990 Pop	0.4262	0.0936	4.5519	0.0000

4.4.2 Control suite indirect effects estimates

Table 4.11 presents indirect effects estimates for the control suite of variables in both 1990 and 2000. As in previous tables, these results have been sorted with the greatest negative value occurring in the first row and the greatest positive value occurring in the final row of the table.

Among the control suite members with negative indirect effects estimates, the *2000 Renters* variable exhibits the largest negative magnitude with a value of approximately -0.1641. Such a result implies that, for the year 2000, a 10 percent increase in the number of renters within a given census tract tended to be associated with a 1.641 percent decrease in the number of people in neighboring tracts with long commutes. This variable is followed in magnitude by the *1990 Renters* and *1990 Walkers* variables with estimates of -0.1519 and -0.1372, respectively. The latter estimate indicates that as of 1990, a 10 percent increase in the number of people that walked to work resulted in, on average, more than a 1.3 percent decrease in the number of people in neighboring tracts with 45 minute commutes.

As was the case in our direct effects results, it appears a greater number of our control variables exhibited positive and significant effects estimates. The *1990 Pop* variable showed the greatest positive estimate with a value of approximately 0.4262. Such a result indicates that for the year 2000, a 10 percent increase in tract population was associated with more than a 4.2 percent increase in people of neighboring tracts with long commutes. This result was followed in magnitude by the *2000 Pop* and *1990 Income* variables. It is worth pointing out that for all of our variable suites, the indirect effects tended to be larger in magnitude than the direct effects. This follows from the cumulative nature of the indirect effects, much like we saw in our applied example. As a result, total effects estimates have a greater propensity to mirror their indirect counterparts. This finding reemphasizes the strong impact that congestion can have on commuting times and motivates the inclusion of spillover effects in commuting time models.

Table 4.12: Sorted control total effects estimates for 1990 and 2000

Variable	Total Effect	Standard Deviation	$t-$stat	$t-$prob
1990 Renters	-0.2246	0.0340	-6.6093	0.0000
2000 Renters	-0.2174	0.0280	-7.7731	0.0000
1990 Walking	-0.1738	0.0185	-9.4037	0.0000
2000 College	-0.1452	0.0242	-6.0072	0.0000
2000 Walking	-0.1051	0.0155	-6.7598	0.0000
1990 College	-0.0964	0.0255	-3.7869	0.0002
2000 Three cars	-0.0825	0.0223	-3.7008	0.0002
1990 Three cars	-0.0481	0.0230	-2.0867	0.0369
2000 Water	0.0630	0.0177	3.5652	0.0004
1990 Water	0.0687	0.0197	3.4887	0.0005
2000 No car	0.1206	0.0224	5.3830	0.0000
1990 Land	0.1412	0.0110	12.8540	0.0000
2000 Income	0.1516	0.0431	3.5140	0.0004
2000 Land	0.1559	0.0097	16.0252	0.0000
1990 Public trans	0.1907	0.0097	19.6292	0.0000
2000 Public trans	0.2086	0.0087	24.0391	0.0000
1990 No car	0.2215	0.0217	10.1999	0.0000
1990 Income	0.2497	0.0484	5.1529	0.0000
2000 Pop	0.4336	0.0941	4.6085	0.0000
1990 Pop	0.5242	0.1139	4.6024	0.0000

4.4.3 Control suite total effects estimates

Control suite total effects estimates for 1990 and 2000 are shown in Table 4.12. Again, the estimates have been placed in ascending order with the greatest negative magnitude in the first row and the greatest positive magnitude placed in the final row of the table.

As previously mentioned, the order of the control suite total effects estimates is nearly identical to that of the indirect estimates. As such, the *1990 Renters* variable again exhibits the greatest negative magnitude with a value of approximately -0.2246. Accordingly, for the year 1990, we would expect a 10 percent increase in the number of renters in a tract to result in a 2.246 percent decrease in the number of residents in own and neighboring tracts with long commutes. This variable was followed in order of negative magnitude by the *2000 Renters* and *1990 Walkers* variables, with effects estimates of -0.2174 and -0.1738, respectively.

The control variables with positive total effects estimates also mirror their indirect counterparts, with the *1990 Pop* variable exhibiting the greatest positive magnitude at a value of approximately 0.5242. This estimate indicates that for the year 1990, a 10 percent increase in a census tract's population tended to result in more than a 5 percent increase in the number of people in own and neighboring tracts with long commutes. The *2000 Pop* and *1990 Income* variables followed in order of positive magnitude. Perhaps the most unusual result associated with the control variable total effects estimates is the fact that all of the estimates were significantly different from zero at the 95 percent confidence level. This differs from our demographic and housing location suites which each contained a number of total effects estimates which were not statistically different from zero.

Before concluding our discussion of effects estimates, we should reiterate several important principles that have guided our analyses thus far. The purpose of determining effects estimates was to establish relationships among our explanatory variables and commuting times for both 1990 and 2000. Although some of these relationships may provide insight as to the underlying causes of long commute times, the individual estimates alone are of no use in determining what caused the break in commuting time trend that occurred over this decade. What is of interest, however, are how these estimates changed from 1990 to 2000. For instance, the total effects estimate for the *1990 Income* variable was 0.2497, whereas its counterpart *2000 Income* had a value substantially less at 0.1516. Such a result, if significant, indicates that increases in household income in the year 2000 had less of an impact on long commutes in own and neighboring tracts than it did in 1990. This finding could potentially be at odds with the results reported by Gordon, Lee and Richardson, who concluded that rising household income levels played a role in commuting time increases over this period. Such scenarios will be explored in greater detail in the coming sections where differences between the two time periods are formally analyzed.

4.5 Changes in effects estimates between 1990 and 2000

To determine which dependent variables could potentially account for the break in commuting time trend, the direct, indirect, and total effects estimates for 1990 and 2000 were subjected to a $t-$test for significant differences. For each variable, this was done by subtracting the 1990 effects estimate from the 2000 estimate, and then dividing the difference by the average of the years' standard deviations. The result of this calculation is a $t-$statistic which can be used to determine if the difference between the 1990 and 2000 effects estimates is significantly different from zero. Significant differences will suggest that the variable's relationship with long commute times did change over this period, which will be of use in determining the most important drivers of the change in

Table 4.13: Significant changes in direct effects estimates over the 1990 to 2000 period

Variable	Direct Effect 1990	Direct Effect 2000	Change 2000-1990	$t-$stat	$t-$prob
Part I: Significant Negative Changes in Direct Effects					
Males 45-54	0.0739	-0.0067	-0.0806	-8.5745	0.0000 *
Females 25-29	0.0658	0.0356	-0.0302	-3.8228	0.0001
Males 25-29	0.0328	0.0117	-0.0211	-2.9306	0.0034
Work in suburbs	0.0337	0.0201	-0.0136	-3.2381	0.0012
No car	0.0245	0.0109	-0.0136	-3.5789	0.0003
College	-0.0605	-0.0722	-0.0117	-2.3636	0.0181
Work in non-metro	0.0247	0.0148	-0.0099	-5.0769	0.0000
Work in osuburbs	0.0590	0.0496	-0.0094	-4.9474	0.0000
Part II: Significant Positive Changes in Direct Effects					
Moved from nmetro	-0.0112	-0.0013	0.0099	4.0408	0.0001
Moved in 20-30 yrs	-0.0071	0.0062	0.0133	3.5467	0.0004 *
Walking	-0.0367	-0.0228	0.0139	4.6333	0.0000
Renters	-0.0727	-0.0533	0.0194	3.2605	0.0011
Females 35-44	0.0569	0.0931	0.0362	2.9431	0.0033
Employed males	0.2742	0.3764	0.1022	4.3489	0.0000
* indicates a change in sign of the effect between 1990 and 2000					

commuting time trend.

Significant changes in direct effects estimates are presented in Table 4.13. Effects estimates and their differences are accompanied by a $t-$statistic and the associated p-level. To ease analysis, results have been sorted from the greatest negative magnitude difference to the greatest positive difference. The results have been further separated, with negative differences given in Part I and positive differences in Part II.

Table 4.13 reveals that the variable with the greatest negative difference in direct effects estimates was *Males 45-54*, where we see a large positive effect from the 1990 model become a small negative effect in 2000. This represents a situation where the positive (adverse) effect from 1990, which led to increases in long commutes, turned to a negative (beneficial) effect that would lead to a smaller number of own-tract residents with long commute times.

Within Part I of Table 4.13, we find a number of negative differences which were the result of a positive 1990 direct effects estimate, followed by a smaller positive direct effects estimate in 2000. This indicates that the associated variables experienced a significant decrease in their adverse own-tract impact on commuting times over this period. Variables in this category include, in order of negative magnitude, *Females 25-29*, *Males 25-29*, *Work in suburbs*, *No car*, *Work in non-metro*, and *Work in suburbs.* It is important to note that although each of these variables experienced a decrease in direct effects, their impact on long commuting times was still adverse in both periods.

The *College* variable also exhibited a significant decrease in direct effects estimates between 1990 and 2000. However, since the effects estimate was negative in both time periods, this change suggests that the number of people over the age of 25 with college degrees had an increasing beneficial impact on the number of people with long commutes.

From Part II of Table 4.13, we find that several variables exhibited significant positive changes in direct effects estimate over this period. Among them, the *Employed males* variable experienced the largest positive change in direct effects estimate between 1990 and 2000. Given the positive direct effects estimates associated with this variable in both years, the result indicates that the number of employed males in a census tract had an increasing adverse impact on the number of tract residents with long commutes. This result was followed by the *Females 35-44* variable which also exhibited positive direct effects estimates in each year. However, the magnitude of this adverse difference was only a third that of *Employed males*. On the other hand, the *Renters*, *Walking*, and *Moved from non-metro* variables each experienced a positive change in direct effects estimates over this time, but since each of their effects estimates was negative, we would conclude that the beneficial impact of these variables diminished over the decade from 1990 to 2000. We should also note the positive change in direct effects estimates associated with the *Moved in 20-30 years* variable, which was the result of a slightly negative 1990 direct effects estimate turning slightly positive by 2000.

Table 4.14 presents significant changes in indirect effects estimates over the period from 1990 to 2000, which have been placed in the same format as our previous table. An examination of the table reveals the changes in indirect effects were much larger than that of our direct effects. For example, even the smallest negative difference in indirect effects (-0.0873) is greater in magnitude than the largest negative difference in direct effects estimates (-0.0806). This result is similar to our previous analyses of direct and indirect effects and underscores the importance of incorporating spatial spillovers into commuting time models.

The largest negative difference in indirect effects estimates was associated with the *Employed females* variable, which went from having a positive effects estimate in 1990 to having a negative indirect estimate in 2000. This suggests that although the number of employed females in a tract had an adverse impact on neighboring tract commute

Table 4.14: Significant changes in indirect effects estimates over the 1990 to 2000 period

Variable	Indirect Effect 1990	Indirect Effect 2000	Change 2000-1990	$t-$stat	$t-$prob
Part I: Significant Negative Changes in Indirect Effects					
Employed Females	0.6305	-0.2845	-0.9150	-10.1329	0.0000 *
Males 45-54	0.3406	-0.1205	-0.4611	-8.6673	0.0000 *
Moved in 1 yr	0.0276	-0.1782	-0.2058	-5.2905	0.0000 *
Females 25-29	0.1805	0.0007	-0.1798	-4.0134	0.0001
Females 15-19	0.0662	-0.0662	-0.1324	-3.4842	0.0005 *
Moved in 6-10 yrs	-0.0226	-0.1490	-0.1264	-3.5258	0.0004
Income	0.3097	0.1979	-0.1118	-2.9382	0.0033
No car	0.1970	0.1097	-0.0873	-4.7836	0.0000
Part II: Significant Positive Changes in Indirect Effects					
Moved in 20-30 yrs	-0.0301	0.0235	0.0536	2.9370	0.0033 *
Walking	-0.1372	-0.0823	0.0549	3.9075	0.0001
Moved from nmetro	-0.0890	-0.0232	0.0658	5.9818	0.0000
Females 20-24	-0.1272	0.1059	0.2331	6.1101	0.0000 *
Moved in 2-5 yrs	0.0737	0.3457	0.2720	5.3333	0.0000
Males 55-64	-0.1925	0.1217	0.3142	6.2465	0.0000 *
Females 35-44	-0.0413	0.4170	0.4583	6.5378	0.0000 *
Employed males	-0.9802	-0.1551	0.8251	8.1131	0.0000
* indicates a change in sign of the effect between 1990 and 2000					

times in 1990, by 2000, increases in the number of employed females within a tract actually had a beneficial impact on long commutes in neighboring tracts. This result is of particular interest as the estimates associated with the *Employed males* variable exhibited the greatest positive increase in indirect effects estimates, somewhat inversely mirroring that of the *Employed females.* That is, the number of employed males in a tract went from having a large beneficial spillover impact, in 1990, to having a much smaller beneficial spillover in 2000.

The second greatest negative difference of indirect effects estimates was associated with the *Males 45-54* variable which also went from having a positive effects estimate of 0.3406, in 1990, to having a negative estimate of approximately -0.1205 in 2000. Again, this indicates that the *Males 45-54* variable went from having an adverse spillover in 1990 to having a beneficial externality by 2000. Much as we saw with *Employed females* and *Employed males*, the negative difference in the *Males 45-54* variable is mirrored by a positive difference in the *Females 35-44* variable, which went from having a slightly negative indirect effects estimate in 1990, to having a relatively large positive effects estimate in 2000. Thus, the number of females aged 35-44 in a census tract went from having a beneficial commuting time spillover in 1990, to having a decidedly adverse impact on neighboring tracts by 2000.

Table 4.15 shows significant changes in total effects estimates for 1990 and 2000, again sorted from greatest negative magnitude to greatest positive magnitude. Since the total effects are, by definition, the sum of the direct and indirect effects, we would expect

Table 4.15: Significant changes in total effects estimates over the 1990 to 2000 period

Variable	Total Effect 1990	Total Effect 2000	Change 2000-1990	$t-$stat	$t-$prob
Part I: Significant Negative Changes in Total Effects					
Employed females	0.7611	-0.1783	-0.9394	-8.4478	0.0000 *
Males 45-54	0.4145	-0.1272	-0.5417	-8.6534	0.0000 *
Moved in 1 yr	0.0575	-0.1618	-0.2193	-4.6959	0.0000 *
Females 25-29	0.2463	0.0363	-0.2100	-3.9848	0.0001
Moved in 6-10 yrs	-0.0152	-0.1521	-0.1369	-3.2061	0.0013
Females 15-19	0.0608	-0.0758	-0.1366	-3.0594	0.0022 *
No Car	0.2215	0.1206	-0.1009	-4.5760	0.0000
Work in osuburbs	0.1101	0.0769	-0.0332	-3.8605	0.0001
Part II: Significant Positive Changes in Total Effects					
Move in 20-30 yrs	-0.0372	0.0298	0.0670	3.0455	0.0023 *
Walking	-0.1738	-0.1051	0.0687	4.0412	0.0001
Moved from non-metro	-0.1002	-0.0245	0.0757	5.6283	0.0000
Females 20-24	-0.1276	0.1201	0.2477	5.4983	0.0000 *
Moved in 2-5 yrs	0.1550	0.4286	0.2736	4.4852	0.0000
Males 55-64	-0.1817	0.1346	0.3163	5.3656	0.0000 *
Females 35-44	0.0157	0.5101	0.4944	5.9964	0.0000
Employed males	-0.7059	0.2213	0.9272	7.4087	0.0000 *
* indicates a change in sign of the effect between 1990 and 2000					

to see some similarities to our previous findings Furthermore, since the total effects represent both the impact on own and neighboring tract residents, these results may be most useful in assessing which variables were responsible for the overall break in commuting time trend.

From Table 4.15 we find that the *Employed females* variable experienced the greatest negative change in magnitude, whereas the *Employed males* variable exhibited the greatest positive change in total effects estimates over this period. This result is very similar to our indirect effects finding as these differences are again almost equal and offsetting in terms of magnitude. The second greatest negative difference in total effects estimates was associated with the *Males 45-54* variable, whereas the second greatest positive difference in estimates was associated with the *Females 35-44* variable. Although the magnitude of these differences was again almost equal and offsetting, we should note that the *Females 35-44* variable exhibited positive total effects estimates in both 1990 and 2000, whereas *Males 45-54* went from having a positive total effects estimate in 1990 to a negative estimate in 2000. This suggests that *Females 35-44* experienced an increasing adverse impact on commuting times in own and neighboring tracts, while *Males 45-54* went from having an adverse impact in 1990 to having a beneficial impact on own and neighboring tract long commutes in 2000.

Continuing our examination of changes in total effects, we find that each of the *Moved in 1 yr*, *Females 25-29*, *Moved in 6-10 yrs*, *Females 15-19*, *No car*, and *Work in osuburbs* variables experienced negative differences in total effects estimates between 1990 and 2000. Although several members of our housing location suite appear in this group, we should note that that magnitude of these differences was only a fraction of either *Employed females* or *Males 45-54*.

Other variables such as *Males 55-64*, *Moved in 2-5 yrs*, *Females 20-24*, *Moved from non-metro*, *Walking*, and *Moved in 20-30 yrs* experienced positive changes in total effects estimates over this period. Of particular interest is the *Moved in 2-5 yrs* variable since according to Gordon, Lee, and Richardson, rapid changes in income would have allowed households to purchase homes at more distant exurban locations during this period. The increasingly adverse own and neighboring tract impact associated with this variable may lend some support to their argument, however, we should note that the magnitude of this difference (0.3163) is substantially less than that of either *Females 35-44* (0.4944) or *Employed males* (0.9272).

CHAPTER 5

CONCLUSION

After establishing relationships between our explanatory variables and the number of workers with travel to work times of at least 45 minutes for each of the years 1990 and 2000, we examined significant differences in direct, indirect, and total effects over this period. An analysis of these results has led us to several conclusions. First, the indirect effects associated with our explanatory variables, as well as the changes they experienced over the 1990 to 2000 period, were greater than that of the direct effects. This emphasizes the important role that spatial spillovers play in the determination of commuting time systems. By ignoring this spillover, studies that rely on non-spatial regressions will result in estimates that are both biased and inconsistent. Second, in our analyses of the significant differences in effects estimates over the 1990 to 2000 period, we found that the greatest changes were associated with demographic suite variables which reflect the age and gender distribution of census tract residents. Differences in total effects estimates, which quantify the impact of our explanatory variables on commute times in both the own and neighboring tracts, suggest the most important changes that took place over this period were associated with the number of males and females living in a census tract. The change in total effect response elasticity of workers having long commutes for the number of employed females residing in a tract was approximately -0.939 over this period. In contrast, the change in total effects response elasticity for employed males was approximately 0.927. The next largest changes in total effects response elasticity were also associated with age and gender related variables. The number of males aged 45 to 54 in a tract conferred a response elasticity of approximately -0.5417 and the number of females aged 35 to 44 was found to have a total effects response elasticity of approximately 0.4944. The magnitude of these differences relative to other explanatory variables used in our model indicate that changes in demographic trends relating to age and gender characteristics are most important in explaining the increase in long travel to work times across the continental United States.

It is important to note that several model variables associated with recent location changes of tract residents did experience significant differences in terms of total effects response elasticity over this period. Variables relating to householders that moved into their residence in the past year, moved-in 2 to 5 years ago, moved-in 6 to 10 years prior, as well as those that moved from a non-metro area in the past five years all experienced a statistically significant change in total effects estimate. However, the magnitudes of these changes were far less than that of the aforementioned demographic variables. For instance, the moved-in two to five years ago variable experienced the greatest change in magnitude of all the housing location variables, however, this change was only slightly greater than half that of the change associated with the number of females aged 35 to 44, which was just the fourth greatest demographic change.

Perhaps the most interesting result in our examination of the significant changes in total effects estimates was the lack of a statistically significant difference in the income variable. This may serve as evidence against Gordon, Lee, and Richardson's (2004) argument that the increase in long commute times was the result of a rapid rise in household income over the 1995 to 2000 period, which led to an increased demand for larger homes at more distant locations from the central business district. They argued that it was the inability of manufacturing and services employment to re-locate to exurban and rural locations during this period that ultimately caused the increase in travel to work times.

From our results, we propose it was in fact rapid changes in demographic patterns and the associated location decisions, rather than income, that was responsible for the break in commuting time trend. Given the roughly equivalent, but opposite differences in total effect elasticities associated with the number of employed females and employed males in a tract, along with significant differences in several pertinent housing location variables, it seems plausible that households may have relocated during this period to achieve a more equitable commuting time distribution among its working female and male members. Further, the total effects differences in other age and gender variables may also suggest that changing demographic patterns had an important impact on commuting times during this period. Collectively, changes such as these may have substantially altered residents' locations relative to place of work in a manner that could not be matched by relocating manufacturing and services employers, thereby increasing the number of own and neighboring tract residents with travel to work times of 45 minutes or more.

References

Arnott, R. A. de Palma and R. Lindsey (1990). "Economics of a Bottleneck," *Journal of Urban Economics*, 27, 111-130.

Crane, R. and D. Chatman (2003). "Traffic and Sprawl: Evidence from U.S. Commuting, 1985-1997," Planning and Markets, 6, no. 3 (an online journal http://www-pam.usc.edu/volume6/v6i1a3s1.html).

Gelfand, A. E. and A. F. M. Smith (1990). "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association* 85, 398-409.

Gordon, P. and A. Kumar and H. W. Richardson (1989). "The Influence of Metropolitan Structure on Commuting Time," *Journal of Urban Economics* 26, 138-151.

Gordon, P. B. Lee and H. W. Richardson (2004). "Travel Trends in U.S. Cities: Explaining the 2000 Census Commuting Results," Lusk Center for Real Estate, University of Southern California Working Paper.

LeSage, J. P. (1997). "Bayesian Estimation of Spatial Autoregressive Models," *International Regional Science Review*, 20, 113-129.

LeSage, J. P. and R. K. Pace, (2004). "Issue Introduction: Spatial Statistics and Real Estate," *Journal of Real Estate Finance and Economics*, 29:2, 147-148.

McGuckin, N. and N. Srinivasan (2003). "Journey to Work Trends in the United States and its Major Metropolitan Areas 1960 - 2000," US Department of Transportation Federal Highway Administration Office of Planning, Publication No. FHWA-EP-03-058.

Pace, R. K. and J. P. LeSage (2008). "Biases of OLS and Spatial Lag Models in the Presence of an Omitted Variable and Spatially Dependent Variables," paper submitted to *Letters in Spatial Resource Sciences*.

Pace, R. K. and J. P. LeSage (2006). "Interpreting Spatial Econometric Models," forthcoming in *Geographical Analysis*.

Reschovsky, Clara (2004). "Travel Time to Work: 1990 and 2000," in *Journey to Work: 2000 Census 2000 Brief*, March 2004 census.gov/prod/2004pubs/c2kbr-33.pdf

VITA

Dustin K. Kirby was born in Houston, Texas, on May 6, 1980, to Gene and Sharon Kirby. After graduating from Eldorado High School, in Albuquerque, New Mexico, he enlisted in the United States Navy where he served as an electronics technician at Naval Air Weapons Station China Lake, California, as well as Naval Air Station Sigonella, Sicily, Italy. In August 2004, upon completion of his enlistment, he enrolled at Texas State University-San Marcos. He received the degree of Bachelor of Science from Texas State in December 2006 with a major in mathematics and a minor in economics. In January 2007, he entered the Graduate College of Texas State University-San Marcos.

Permanent Address: 1729 Gaylord Drive

Austin, TX 78728

This thesis was typed by Dustin K. Kirby.