Title
Specification Issues in Models of Population and Employment Growth

Authors
Boarnet, Marlon G.
Chalermpong, Saksith
Geho, Elizabeth

Publication Date
2001-12-01

Specification Issues in Models of Population and Employment Growth

Marlon G. Boarnet
Associate Professor
Departments of Urban and Regional Planning and Economics
and Institute of Transportation Studies

Saksith Chalermpong
Ph.D. Student, Transportation Science
and Institute of Transportation Studies

Elizabeth Geho
Ph.D. Student, Transportation Science
and Institute of Transportation Studies

University of California, Irvine
Irvine, CA 92697

Boarnet contact information:
Ph. 949-824-7695
Fax: 949-824-8566
Email: mgboarne@uci.edu

December, 2001

The University of California Transportation Center, funded by the U.S. and California Departments of Transportation, provided financial support for this research.

I. Introduction

Spatial econometric adaptations of population and employment growth models have been used to study the employment impacts of urban rail transit (Bollinger and Ihlanfeldt, 1997), the links between urban and rural development (Henry, Barkley, and Bao, 1997; Schmitt and Henry, 2000), and causality between intra-metropolitan population and employment location (Boarnet, 1994b). Yet the literature has so far given limited attention to two specification issues that are fundamental to the performance of spatial econometric population and employment growth models. First, the weight matrix, which defines how geographic units of observation relate to one another, must be defined *a priori*, and alternative versions of the weight matrix have rarely been consistently compared. Second, most recent population-employment growth models are lagged adjustment models, yet the estimated lag parameters often imply that the system does not adjust to a long-run equilibrium, violating one of the maintained hypotheses of the lagged adjustment approach. This paper analyzes those three specification issues, and provides insight into both the validity of various econometric practices that have been common in the recent literature and the stability of econometric population and employment growth models when typical assumptions and approaches are changed.

II. Background

Population and employment growth models have a long history in regional science and urban economics. Steinnes and Fisher (1974) first introduced a two-equation, intra-metropolitan population and employment growth model that, by including a spatial framework via potential variables, was quite advanced for its time. Since then, other authors have studied the determinants of growth within and across metropolitan areas (e.g. Grubb, 1982; Luce, 1994). A related literature divides metropolitan areas into two parts, central cities and suburban rings, and examines growth in the central and suburban areas, (e.g. Bradford and Kelejian, 1973; Palumbo, Sacks and Wasylenko, 1990; Leichenko, 2001). While both literatures pre-date the development of many of the recent tools of spatial data analysis, recent articles have increasingly used spatial analysis to both model and interpret how proximate geographic areas grow (e.g. Fingleton, 2001; Goffette-Nagot and Schmitt, 1999; Mulligan, Vias, and Glavac, 1999; Roberts, 2000; Wheeler, 2001). Other recent articles that have examined growth within both urban and rural areas have not emphasized spatial interactions but adopt simultaneous population-employment growth models that relate to several of the issues discussed in this paper (e.g. Deller, Tsai, Marcouiller, and English, 2001; Glavac, Vias, and Mulligan; 1998).

For small, geographically proximate observations, one motivation for an explicitly spatial econometric treatment is that the link between population and employment changes extends beyond commonly used geographic boundaries. Population changes in a county, municipality, census tract, or other intra-metropolitan geographic unit depend not only on employment changes within the same jurisdiction but also on employment changes in a labor market area that could extend beyond that jurisdiction. Similarly, employment changes depend on population changes in surrounding labor markets. This leads to a spatial structure in the econometric model, and the problem of spatial dependence across observations is likely more severe for smaller observations of the sort that are inherent in intra-metropolitan as opposed to inter-metropolitan models.

Boarnet (1994a) applied spatial econometrics to a Carlino and Mills (1987) lagged adjustment model of population and employment growth to handle this problem of spatial dependence across observations. The theoretical rationale was that interactions across observations (New Jersey municipalities in the case of Boarnet's 1994 study) are mediated by a commuting relationship. Thus the link between population and employment is best modeled as a dependence within labor-market areas (or commuter-sheds) which, given the small size of New Jersey municipalities, almost certainly were larger than any one municipal observation. We adapt that model here to study the issues related to the weight matrix, estimated lag parameters, and identification discussed above.

III. Model

The model used here is the one developed in Boarnet (1994a), which is the same model that is applied or adapted in Boarnet (1994b), Bollinger and Ihlanfeldt (1997), Henry, Barkley, and Bao (1997), Henry, Schmitt, Kristensen, Barkley, and Bao (1999), and Schmitt and Henry (2000). The model starts with the lagged adjustment model of population and employment growth presented in Carlino and Mills (1987). Because the research discussed in this paper uses census tracts as the geographic unit of observation, the description below will speak in terms of census tracts, but the concepts apply generally to any unit of geographic measurement.

Equilibrium tract population and employment are assumed to follow the relationships shown below.

$$POP^*_{i,t} = f(X_{i,t}, \overline{EMP}^*_{i,t}) \tag{1}$$

$$EMP^{*}_{i,t} = g(Y_{i,t}, \overline{POP}^{*}_{i,t}) \tag{2}$$

where
$POP^{*}_{i,t}$ = equilibrium population $EMP^{*}_{i,t}$ = equilibrium employment
$X_{i,t}$ = a matrix of characteristics that affect equilibrium employment
$Y_{i,t}$ = a matrix of characteristics that affect equilibrium population $\overline{POP}^{*}_{i,t}$ = equilibrium population in the labor market centered on census tract "*i*" in time "*t*" $\overline{EMP}^{*}_{i,t}$ = equilibrium employment in the labor market centered on census tract "*i*" in time "*t*""*i*" subscripts refer to census tracts
"*t*" subscripts refer to time periods

Following Carlino and Mills (1987), tract population and employment are assumed to adjust to their equilibrium values with a lag, as shown below.

$$POP\Delta_{i,t} = POP_{i,t} - POP_{i,t-1} = l_{p}(POP^{*}_{i,t} - POP_{i,t-1}) \tag{3}$$
$$EMP\Delta_{i,t} = EMP_{i,t} - EMP_{i,t-1} = l_{e}(EMP^{*}_{i,t} - EMP_{i,t-1}) \tag{4}$$

where l_p and l_e both take on values between 0 and 1.

Assuming that (1) and (2) are linear, with normally distributed error terms, and then substituting into (3) and (4) gives

$$POP\Delta = Xb_{1} + a_{1}\overline{POP}^{*} - l_{p}POP_{t-1} + u \tag{5}$$

$$EMP\Delta = Yb_{2} + a_{2}\overline{EMP}^{*} - l_{e}EMP_{t-1} + v \tag{6}$$

where the "*i*" and "*t*" subscripts have been dropped except in the case of POP_{t-1} and EMP_{t-1}, $ß_1$ and $ß_2$ are column vectors of parameters, a_1 and a_2 are scalar parameters, and u and v are normally distributed error terms.

Equations (5) and (6) cannot be estimated directly because the equilibrium labor market variables, $\overline{POP}^{*}$ and $\overline{EMP}^{*}$, are unobservable. The difficulty is twofold. First, the equilibrium values must be related to actual, observable values. Second, the census tract data must be aggregated into labor market variables. To relate equilibrium values to observable quantities, assume that the labor market variables, $\overline{POP}^{*}$ and $\overline{EMP}^{*}$, adjust to equilibrium according to the same lag process as in equations (3) and (4).

$$\overline{POP}^{*}_{i,t} = \overline{POP}_{i,t-1} + (1/\, l_p)(\overline{POP}_{i,t} - \overline{POP}_{i,t}) \tag{7}$$

$$\overline{EMP}^{*}_{i,t} = \overline{EMP}_{i,t-1} + (1/\, l_e)(\overline{EMP}_{i,t} - \overline{EMP}_{i,t}) \tag{8}$$

where overbars denote labor market values
asterisks denote an equilibrium value
other values are actual values

The actual (not equilibrium) labor market variables are then measured by potential variables which can be represented in matrix notation. Representing labor market variables as a weighted sum and substituting (7) and (8) into (5) and (6) gives

$$POP\Delta_t = X_{t-1}b_1 + a_1(\mathbf{I}+\mathbf{W})EMP_{t-1} + \frac{a_1}{l_e}(\mathbf{I}+\mathbf{W})EMP\Delta_t - l_p POP_{t-1} + u \tag{9}$$

$$EMP\Delta_t = Y_{t-1}b_2 + a_2(\mathbf{I}+\mathbf{W})POP_{t-1} + \frac{a_{21}}{l_p}(\mathbf{I}+\mathbf{W})POP\Delta_t - l_p EMP_{t-1} + v \tag{10}$$

where **I** is an (n x n) identity matrix
W is an (n x n) weight matrix
all other variables are matrices or vectors.

Following Carlino and Mills (1987), the variables in the X and Y matrices were lagged to the base year, "*t-1*", to help identify the system in (9) and (10).

This is what Rey and Boarnet (2000) call a spatial cross-regressive system. The spatial lag of employment change is an independent variable in the population change equation, and similarly the spatial lag of population change appears in the employment change regression. Because the first-stage regression in two stage least squares would have spatial lags of the dependent variable on the right-hand side, standard least squares routines will be biased and inconsistent. Boarnet (1994a) uses an application of the instrumental variables estimator proposed in Anselin (1980) to estimate the system in equations (9) and (10).

In this paper, we examine two specification issues that are common in the model described above.

*Weight (**W**) matrix and the specification of the population-employment interaction*

In the literature on population and employment growth models, the **W** matrix is often chosen to reflect commuting relationships between residents and job sites in geographic observations in the data set (e.g. Boarnet, 1994a; Bollinger and Ihlanfeldt, 1997; Henry, Barkley, and Bao, 1997). Various specifications of **W** have been used in the literature, including **W** matrices based on national commuting data for the United States (Boarnet, 1994a) and **W** matrices based on *a priori* concepts of uniformly sized commuter-sheds (Schmitt and Henry, 2000). Other candidates for W matrices include contiguity matrices commonly used in other applications and **W** matrices based on measures of commute flows between geographic areas within the data set. In this paper, we examine several alternative **W** matrices, comparing how different definitions of **W** change the results of parameter estimates.

Implied stability of the lagged adjustment process

Recent models of population and employment growth typically follow a lagged adjustment process as specified in, e.g., Carlino and Mills (1987). Such lagged adjustment models yield estimates of lag parameters that can be used to infer whether the adjustment process will converge to an equilibrium over time (see, e.g., Carlino and Mills, 1987). Many recent applications of small area population and employment growth models have yielded estimated lag parameters that do not give stable convergence to an equilibrium (e.g. Boarnet, 1994a; Mulligan, Vias, and Glavac, 1999; Schmitt and Henry, 2000).[1] Boarnet (1994b) speculated that this problem is linked to the fact that the independent variables in these models typically have little or no information on land use regulation. Variables that might measure constraints on intra-metropolitan location, including constraints imposed by local government regulation (e.g. zoning or land use controls), are rarely included in the models. Thus the equilibrium values of population and employment might be poorly measured due to omitted variables that relate to local land use regulation. If the omitted variables are correlated with census tract population and employment levels, this could bias the estimate of a lag parameter since the estimate is the coefficient on either a base year population or employment variable in equations (9) and (10).

In this paper, we examine whether including detailed information on various land use categories in both the population and employment growth regressions, as a proxy for local land use policies, can improve the reasonableness of the estimated lag parameters. Thus we examine the hypothesis that the poor performance of the lagged adjustment approach is due to an incomplete set of independent variables, rather than an inappropriateness of the lagged adjustment assumption. Importantly, for purposes of this paper the examination of the estimated lag parameters (a relatively technical question) and the validity of adding land use variables to the model (something that could have profound implications for policy analysis) are linked.

[1] See Mulligan, Vias, and Glavac (1999) for results of specification tests of lagged adjustment population and employment growth models with different time lags and a discussion of estimated lag parameters and model stability.

To examine these two specification issues, we estimated regressions that consistently examined two issues – (1) how estimates of lag parameters change when variables that measure land uses are added to the model and (2) how the model estimates vary across a range of weight matrices. First we discuss the various weight matrices tested.

IV. Alternative Weight Matrices

The form of the spatial interaction across geographic units of observation in equations (9) and (10) has a clear basis in theory. For intra-metropolitan models, Boarnet (1994a) suggested that the **W** matrix should reflect commuting patterns that tie any one small geographic unit to a larger labor market area.[2] One question examined here is whether the parameter estimates from equations (9) and (10) are sensitive to different definitions of **W**, and in general how the estimates from relatively simple and more complex **W** matrices compare.

We tested several **W** matrices, listed below roughly in order from the simplest to construct to the most complex. Each **W** matrix is an a priori definition of a labor market that aggregates census tracts into commuter-sheds based on differing implementations of a concept of labor market areas.

[2] Alternative ides of spatial interaction , not examined here, include spread and backwash effects that involve positive or negative spillovers of population or employment growth across geographic units of observation. For a discussion of alternative models, see Rey and Boarnet (2000), and for an examination of spread and backwash within the context of the model in equations (9) and (10), see Henry, Barkley, and Bao (1997).

Neighbor Matrices

The simplest proximity matrix is the 0,1 neighbor matrix. In this matrix, element w_{ij} equals one if tracts *i* and *j* border each other and zero otherwise. Tracts that meet only at a corner were defined as neighbor tracts. The diagonal elements are zero as a tract cannot be its own neighbor. The matrix was created from a visual inspection of the 1980 Orange County census tract maps. Two versions of the 0,1 neighbor matrix were created — one that is row normalized and one that is not.[3] This **W** matrix can be constructed from visual inspection of maps or automatically from a geographic information system (GIS), and requires no data other than a census tract boundary map.

Fixed Distance Matrix

The distance-based matrix is also a 0,1 matrix. Again all elements start at zero. A matrix element w_{ij} is changed to one if the distance between the centroids of tract *i* and tract *j* is less than a predetermined amount. For this paper the cut-off distance was set at 10 miles, an estimate of average commute distance based on average commute time in Orange County. The mean travel time to work for the Anaheim-Santa Ana-Garden Grove SMSA in 1980 was 23.6 minutes (Cenus of Population and Housing: Table P-9). Under the assumption that average travel speed is roughly 25 miles per hour, the average travel distance would be about 10 miles. Population centroids were obtained from the 1980 Master Area Reference File 2 (MARF 2). In the MARF 2 data, approximately half the tracts were split across municipal boundaries, and the centroids were reported for each portion of these splits tracts. For split tracts, the population centroid was created by weighting each centroid by the proportion of tract population in the corresponding section of the tract. The population-weighted center of the straight-line distance between each pair of census tract centroids was used as the centroid for a split tract. The distance-based **W** matrix incorporates information about commute flows, but has the possible disadvantage of making all commuter-sheds be equal-sized, ignoring any internal variation in commuting patterns within the study area. While more difficult to calculate than the contiguity matrix, the data and processing requirements for the distance-based **W** matrix are relatively manageable.

Weighted Inverse Distance-Based Matrix

This matrix uses the distance between tract centroids, as does the distance-based matrix described above. Instead of matrix elements that have values of zero or one, the elements in this **W** matrix are equal to $1/d^a$, where a is an exponent determined by the researcher and d is the distance between centroids. (The subscripts i and j have been suppressed, but the matrix elements are based on $d_{i,j}$, the distance between any two tracts i and j.) Labor market areas are potential variables, with tracts closer to any particular tract weighted more heavily. The magnitude of the damping coefficient, a, determines how quickly the labor market relation damps with distance. Here we adopt a value of a equal to 0.67, which Boarnet (1992, 1994a) estimated from national commuting data. Bollinger and Ihlanfeldt (1997) provide evidence that the regression results are not sensitive to changes in a ranging from 0.5 to 2.0 for their analysis of Atlanta census tracts.

[3] Row normalizing contiguity matrices is common. See Anselin (1988).

With the exception of choosing a damping parameter, a, this **W** matrix is no more complicated to construct than the distance-based **W** matrix discussed above. Like the distance-based matrix, the inverse distance **W** matrix imposes the same spatial form on commuter-sheds centered on all census tracts in the data set.

Tract-to-Tract Flow Matrix

A theoretically ideal **W** matrix would incorporate data about commute flows across all tracts, allowing the labor market area (or commuter-shed) centered on any one tract to be based on commute patterns between that tract and other tracts. Thus commuter-sheds would vary within the study region, in ways that incorporate variations in commuting patterns within the region. For example, residents on some locations might commute somewhat more than residents of other locations that are more proximate to jobs, and one would prefer that the **W** matrix capture that.

Such a **W** matrix can be constructed from data on tract-to-tract commute flows that are available for 1990 from the STP154 file from the United States Bureau of the Census. The STP154 file has estimates of the number of journey-to-work commutes that originate and terminate at each census tract within Orange County. Such data were not readily available for 1980, but to test the performance of a tract-to-tract commute flow matrix we used the STP154 data from 1990. The 1990 STP154 data were converted to 1980 tract boundaries. A commute flow matrix was based on the number of commutes between tracts, with elements $w_{i,j}$ being the number of commuters traveling between tracts i and j. We tested both row normalized and non-normalized versions of this matrix.

While the tract-to-tract flow matrix captures rich information about commute flows, and is closer to a theoretical ideal of a commuter-shed than any other **W** matrix, it brings two possible disadvantages. The STP154 data are difficult to manipulate, making this the most tedious to construct **W** matrix of the various matrices tested in this paper. Also, the Census Bureau warns that the tract-to-tract commute flows are in some instances estimated, and might include inaccuracies. Our own analysis verified that the STP154 data are noisy. We summed journey-to-work destinations by tract, and assuming that the number of commute journeys ending in a tract should equal the number of jobs in that tract, compared the number of journey-to-work destinations to the number of jobs in each tract. Job data for tracts are from the Southern California Association of Governments (SCAG). For the 484 tracts (based on 1990 tract maps), the difference between the employment implied by the STP154 data and the employment implied by SCAG data exceeded 30% of the SCAG estimate in 47% of the cases (226 tracts). This suggests that the advantage of the theoretically more precise commute-flow data could be outweighed by the noisiness of the STP154 file. Because our purpose here is in part to test various **W** matrices with different advantages and disadvantages, we include estimates of equations (9) and (10) using the two tract-to-tract **W** matrices (normalized and non-normalized) to compare the results to other matrices.

V. Data

The data used in this paper are from Orange County, California. Because a large portion of the data on explanatory variables was obtained from the United States Bureau of the Census, the paper will focus on explaining changes between the census years 1980 and 1990. Data were obtained at the census tract level. The 1990 data were converted to 1980 census tract boundaries for all aspects of the analysis. The conversions were made based on the census listing of how tracts had changed and a visual comparison of the 1992 Census Tract Edition of the Thomas Guide and the Block Statistics Maps for 1980 Census of Population and Housing Anaheim-Santa Ana-Garden Grove, CA SMSA PHC80-1-67.[4] Where there was not a clean conversion, the conversion was made based on an estimate of the percent of one tract located in another.

Compared to some previous research, the set of location specific characteristics, **X** and **Y** in equations (9) and (10), is limited. Our focus here is on alternative specifications of the **W** matrix and tests of the lagged adjustment model, and those two tasks get more attention than testing a broad range of amenity variables.[5] The independent variables in the model fall into four groups: (1) variables required by the spatial and lag structure of the model, (2) location specific characteristics (other than land use) that influence equilibrium tract population level (the vector $\mathbf{X}_{t-1}$ in equation 9) or equilibrium tract employment levels (the vector $\mathbf{Y}_{t-1}$ in equation 10), (3) selected land use variables used to examine whether more complete proxies for equilibrium population and employment levels improves the estimated lag parameters, and (4) 37 census defined place (CDP) dummy variables to proxy for characteristics associated with municipalities

4 Four tracts were dropped to the data set due to data inconsistencies that made conversion from 1990 to 1980 boundaries not possible for all data sources. The resulting data set has 415 census tracts based on 1980 tract boundaries.

5 For a similar approach, see Mulligan, Vias, and Glavac (1999) who tested different lag specifications with essentially no location-specific characteristic (or amenity) variables. Other authors, such as Deller, Tsai, Marcoullier, and English (2001) have focused on amenities with relatively less attention to spatial interactions and lag structure.

or coherent areas that were not incorporated municipalities.[6] The CDP dummy variables are intended to proxy for local characteristics and, for incorporated CDP's, municipal policies that influence population and employment growth. This includes school quality (many municipalities in Orange County have their own school district), municipal expenditure policy, local tax revenues, and crime rates that are not measured with other readily available data at the census tract level. Some tracts are in more than one CDP, in which case the dummy variable for each CDP that contains some of the tract is set equal to one. For other tracts that are wholly in one CDP, the dummy for that CDP only is set to one.

The dependent variables and the variables required by the structure of the model (group 1, above) are listed in Table 1. Data sources are also listed for each variable. Time subscripts are changed to correspond to the years of this study; "t-1" is 1980 and "t" is 1990.

(Table 1 somewhere around here.)

The independent variables in the population equation also include housing stock age, measured by the fraction of a tracts housing stock built before 1960 (% Pre 1960 Housing) and the fraction of the tract housing stock built before 1940 (% Pre 1940 Housing), the proportion of tract residents who were Hispanic in 1980, and the proportion

[6] Most of the census defined places (CDP's) in 1980 were municipalities, and several unincorporated CDP's had a unique identity and have since incorporated, although such incorporations did not necessarily strictly follow 1980 CDP boundaries.

of tract residents who were Black in 1980. All of those variables are from the 1980 Census of Population and Housing.

Land use data by census tract for 1990 were obtained from Aerial Information Systems.[7] The land use variable is total tract area (in acres) in a given use. The land use variables are defined in Table 2. When the land use variables are added to the system of equations, the following variables were added to the population change equation: lu1110 (single family residential), lu 1120 (multi-family residential), lu1140 (mixed residential), lu2000 (agricultural) and lu3000 (vacant). In the employment change equation, the following land use variables are used: lu1210 (general office use), lu1220 (retail stores and commercial services), lu1230 (other commercial), lu1240 (public facilities), lu2000 (agricultural) and lu3000 (vacant). In Orange County during the 1980s, much agricultural land was converted to residential or employment-generating land uses, so the amount of agricultural and vacant land is intended to measure the amount of developable land available in a tract.

(Table 2 somewhere around here.)

Dummy variables for census places are included to control for unobservable characteristics that are homogeneous within places. Because many place boundaries coincide with the boundaries for incorporated areas, these variables may be used as proxies for city amenities or disamenities such as tax rates, school quality or safety that vary largely or completed according to the city of residence. The place variables were

[7] We are grateful to Rena Sivitanidou for providing these data.

dummy variables equal to one if a portion of the tract was in the place and zero otherwise. The place dummy variables are listed in Table 3.

(Table 3 somewhere around here.)

VI. Results

The regressions in equations (9) and (10) were estimated with each of the six **W** matrices described in Section V. In Tables 4 and 5, we give results of estimating equations (9) and (10) without the land use variables listed in Table 2. In Tables 6 and 7, we shows the results of estimating the regressions with the land use variables added. Recall that the land use variables are intended to proxy for land use regulatory policy and/or local attitudes toward population or employment generating development. The local regulatory environment is hypothesized to be a key variable that was omitted from previous work due to lack of available data.[8]

The regressions in Tables 4-7 were estimated with an instrumental variables technique suggested by Anselin (1980, 1988). See Boarnet (1994a) for a description of the instrumental variables technique in the context of the models used here. In Table 4, the instruments for the endogenous EMP? in (**I**+**W**)EMP? in the population change equation are (**I**+**W**)POP80 and EMP80 (1980 tract employment). In Table 5, the

[8] Some readers might think that the CDP dummy variables can proxy for land use regulatory decisions and/or attitudes toward growth. Yet many municipalities have different land use regulatory policies for different areas within their jurisdiction, and a CDP dummy variable will not capture any variation in policy within a municipality or census defined place. Also, our concern is in adding information that we hypothesize is related to the regulatory environment and that improves the performance of the lagged adjustment model. The results suggest that the land use variables improve the performance of the lagged adjustment model, yielding lag parameters of the appropriate sign, while including the CDP dummy variables without land use variables does not give lag parameters of the appropriate sign.

instruments for POP? in (**I**+**W**)POP? are (**I**+**W**)EMP80, POP80 (1980 tract population), the proportion of tract population Hispanic, proportion of tract population Black, % Pre 1960 Housing, and % Pre 1940 Housing. In Table 6, the instruments for EMP? in (**I**+**W**)EMP? include the instruments used in Table 4 plus lu1210, lu1220, lu1230, lu1240, lu1310, lu1320, and lu1340. In Table 7, the instruments for POP? in (**I**+**W**)POP? include the instruments used in Table 5 plus lu1110, lu1120, and lu1140.

(Tables 4-7 somewhere around here.)

Looking first at Tables 4 and 5, there is some variation in the significance of the independent variables depending on the choice of **W** matrix. For population change, five of the six specifications in Table 4 show that tracts with higher proportions of 1980 population that is hispanic had larger population growth in the decade. The coefficient on proportion hispanic is not significant when the row normalized tract-to-tract commute flow **W** matrix is used. The sign of (I+W)EMP80 varies across different **W** matrices, and (I+W)EMP80 is only significant (with positive sign) when the inverse distance matrix is used. Yet the pattern of the coefficients on independent variables generally shows more stability than variation across **W** matrices.

The contemporaneous interaction between population and employment growth depends on the sign and significance of the variable (**I**+**W**)EMP? in the population change equation and the sign and significance of the variable (**I**+**W**)POP? in the employment change equation.[9] The (**I**+**W**)EMP? variable is significantly positive only in the two specifications that used the non-normalized contiguity **W** matrix in Table 4. In the other specifications, (**I**+**W**)EMP? is not significant.[10]

The lag parameter, $?_p$ in equation (9), is the negative of the estimated coefficient on 1980 population in the population change equation in Table 4. The parameter is significantly negative (the estimated coefficient is significantly positive) for all of the six **W** matrices tested in Table 4, implying that the estimated $?_p$ is not within the required range, between 0 and 1.

For the employment change equation in Table 5, changes in population in the surrounding labor market, (**I**+**W**)POP?, are only significant using the normalized contiguity **W** matrix. The term (**I**+**W**)POP80 is significantly negative using the inverse distance matrix. The lag parameter, the negative of the coefficient on EMP80, is significantly positive for all **W** matrices except the non-normalized tract-to-tract commute flow matrix.

To summarize, the regressions in Tables 4 and 5 suggest the following. Census tract employment changes from 1980 to 1990 do not appear to depend on changes in population in surrounding labor market areas, with the exception of the specification with the normalized contiguity matrix, which suggests that employment changes depend positively on population changes in surrounding labor markets. Changes in tract population growth depend positively on changes in employment in surrounding labor

---

9 There is a large literature that interprets whether population growth follows employment growth or vice versa within metropolitan areas or whether the two are simultaneous. For a discussion see, e.g., Steinnes and Fisher (1974), Steinnes (1977), Boarnet (1994a), and Deitz (1998).

10 The coefficients on the census defined place dummy variables are not reported in Tables 4 through 7 for brevity.

market areas for the un-normalized **W** matrix based on geographic contiguity, but other **W** matrices show no dependence of tract population change on labor market employment change.[11] The location-specific characteristics are sparse in Tables 4 and 5, and while some variables, such as the proportion Hispanic in variable in Table 4, showed some sensitivity to the choice of **W** matrix, in general the sign and significance pattern on the location-specific characteristics was somewhat stable across **W** matrices. The estimated lag parameters are consistently of the wrong sign (while the significance varies), implying that the equations estimated in Tables 4 and 5 do not meet the conditions for dynamic stability of the lagged adjustment model.

In Tables 6 and 7, we add the land use variables to the regressions from Tables 4 and 5, to examine whether additional variables that might proxy for previously unmeasured aspects of equilibrium tract population and employment improve the performance of the lagged adjustment model by yielding lag parameters that are consistent with dynamic stability. The six **W** matrices from Tables 4 and 5 are repeated in Tables 6 and 7. In the population change equation, census tracts with larger proportions of their 1980 population black or Hispanic had more population growth. In Table 6, this result does not vary depending on the choice of **W** matrix. One explanation for the positive coefficient on proportion black and proportion Hispanic in the population

[11] Depending on the **W** matrix, this result differs from recent evidence on the causality between intra-metropolitan population and employment change. Boarnet (1994) found that employment changes in New Jersey municipalities from 1980 to 1988 depend on surrounding population changes, but not vice versa. Dietz (1998) found that employment levels in Boston census tracts in 1990 depended on population in surrounding areas, but not vice versa. Table 7, presented below, also gives no evidence that employment changes in Orange County census tracts in the 1980s depend on surrounding population changes, except for the row normalized tract-to-tract flow matrix in Table 7, which shows a positive association between tract employment change and surrounding population change. Mulligan, Vias, and Glavac (1999) found that the pattern of population-employment interaction varied depending on the time period studied, which when combined with the results here suggests that causality between population and employment location and changes might vary across time periods and possibly across metropolitan areas, in addition to being sensitive to the specification of the **W** matrix.

change equation is the growth of minority and immigrant populations in Orange County during the decade, with that growth occurring disproportionately in areas that had some minority concentrations as of 1980. The age of the housing stock (percent of housing built before 1940 and 1960) was not a significant predictor of population growth in Table 6, with the exception of the significant positive coefficient on *% Pre 1960 Housing* for the normalized tract-to-tract flow matrix. The coefficient on (**I**+**W**)EMP80 is significantly positive for two **W** matrices – the inverse distance matrix and the non-normalized contiguity matrix. The coefficient on (**I**+**W**)EMP? is significantly negative for the non-normalized contiguity **W** matrix, significantly positive for the normalized tract-to-tract flow **W** matrix, and insignificant for the other **W** matrices. The estimated lag parameter $?_p$ is consistently positive, with a value close to 0.35, for all **W** matrices.[12] Four land use variables – the number of acres in the tract devoted to single family residential, multi-family residential, mixed residential, and agriculture – are significantly positive regardless of the choice of **W** matrix. There is a correlation between residential land uses and population growth. During the 1980s, Orange County grew rapidly, and agricultural land was often developed, explaining the correlation between agricultural land use and population growth. Note that the R-squared of the regression is substantially higher than in the specification without land use variables in Table 4.

In the employment regressions in Table 7, (**I**+**W**)POP? is significantly positive for the normalized tract-to-tract flow matrix, but insignificant in the specifications that use other **W** matrices. The term (**I**+**W**)POP80 is always insignificant in the employment change regressions in Table 7. The lag parameter $?_e$ is significantly positive, with magnitude close to 0.6, in all specifications in Table 7, implying that, on average, about 60% of the gap between equilibrium and actual tract employment was closed during the 1980s. Six land use variables are positive in all six employment change regressions in Table 7 – the number of acres in the tract in: general office use, retail stores and commercial services, other commercial, public facilities, light industrial, and heavy industrial. In addition to those variables, the amount of vacant land is significantly negative in the specification that uses the normalized contiguity matrix, and both the amount of vacant land and the amount of agricultural land are significantly negative in the specification that uses the normalized tract-to-tract flow matrix.

Overall, in Tables 6 and 7 the coefficients on the land use variables are stable across different **W** matrices with only minor variations, and the lag parameters do not vary across specifications with different **W** matrices. The lag parameters have positive values in Tables 6 and 7, suggesting that adding land use variables better measures the equilibrium levels and also suggesting that the lagged adjustment model is valid and dynamically stable and that past results implying otherwise might have reflected incomplete measures of equilibrium population and employment rather than a shortcoming in the lagged adjustment approach. The link between contemporaneous population and employment changes depends on the choice of **W** matrix.

In the bottom row of each column in Tables 4 through 7, we report the results of an overidentification test for instrument validity. Overidentification tests have been used in the past in population and employment growth models to examine the validity of lagging the independent variables to a base year (e.g., Boarnet, 1994a). Here we examine

---

[12] The value of 0.35 for $?_p$ implies that, on average, 35% of the gap between equilibrium and actual census tract population was closed during the 1980s.

both that concern and the exogeneity of the land use variables. The land use variables that were added to the regressions in Tables 6 and 7 give the number of acres of land in a particular use in 1990. This raises some concern, since these variables are at the end of the period rather than the beginning of the 1980 to 1990 study period. Because it is possible that the amount of land in, for example, residential uses could be correlated with the error term in the employment growth equation, we use the overidentification test to examine the validity of the instrumental variables technique with the land use variables in the model.

The test is calculated by multiplying the R-squared from a regression of the second-stage residuals on all included and excluded pre-determined and exogenous variables in the system by the number of observations. The statistic is distributed chi-squared with degrees of freedom equal to the number of excluded instruments for the endogenous variable in each regression. The critical chi-squared values (at the five percent level) for rejecting a null hypothesis of valid instruments are 5.99 for Table 4 (two degrees of freedom because (**I**+**W**)EMP? is instrumented by two variables in that equation), 12.59 for Table 5 (six degrees of freedom, based on the six instruments for (**I**+**W**)POP? in that equation), and 16.92 for Tables 6 and 7 (nine degrees of freedom).[13] See, e.g., Angrist and Krueger (1989 and 1994) for examples and discussions of the overidentification test.

For the specifications without land use variables, the overidentification statistic fails to reject the null hypothesis of valid instruments (at the five percent level) for four of the six population change regressions in Table 4 (the null is rejected for the normalized and non-normalized contiguity matrices) and for four of the six employment change regressions in Table 5 (the null is rejected for the non-normalized contiguity matrix and the non-normalized tract-to-tract flow matrix). For the specifications with land use variables, the null of valid instruments is not rejected in four of the six population change regressions in Table 6 (the null is rejected for the 10 mile labor market area and non-normalized tract-to-tract flow matrices in Table 6). The null hypothesis of valid instruments is rejected in all of the employment change regressions in Table 7. These overidentification statistics give some evidence that the specifications with land use variables, measured at the end of the study period, are not as consistently valid as the specifications that include no land use variables and hence have independent variables that are measured only at the beginning of the study period.

Land use data were not available for 1980, and hence for this study 1990 land use data were the only data available. Since the development of GIS programs, land use inventories have become increasingly common, and future studies are more likely to have available land use data for the beginning of the study period. Thus it is encouraging that the overidentification statistics performed better when all independent variables were measured at the beginning of the study period.

VIII. Interpretation and Conclusion

In this paper, we examined two specification issues related to population-employment growth models – the definition of the weight matrix and the performance of

[13] The overidentification test examines both the validity of the instruments and the appropriateness of the specification, so rejecting the null hypothesis cannot definitively establish the form of the specification error.

the lagged adjustment model. Our results suggest that, when data such as the land use measures used here are available to measure or proxy for policy toward growth, the lag parameters are typically of the appropriate sign. The overidentification tests suggest that such data should be for the beginning of the study period rather than the end of the study period. Given that data on land use inventories, zoning regulations, and other measures of land use policy will likely be increasingly available at fine levels of geographic detail over the next several years, this bodes well for the ability to apply population-employment growth models in the future.

The nature of the contemporaneous interaction between population and employment growth was sensitive to the specification of the weight matrix, while the coefficients on many of the other independent variables (the location-specific characteristics discussed in Section VI) were not sensitive to the specification of the weight matrix. The location-specific variables, such as proportion black or Hispanic or the proportion of housing built before 1940 and before 1960 in the population change equation, appear to be orthogonal to the population-employment change interactions that are modeled in the weight matrix. This is a potentially important insight. Some recent articles have examined the role of location-specific characteristics in population and employment growth without using a spatial model to examine the interaction between population and employment growth (e.g. Deller, Tsai, Marcouiller, and English, 2001). The results here suggest that such an approach might be valid, or at least that the specification of the population-employment growth interaction might not affect hypotheses about the influence of other, location-specific, variables.

Overall, these results suggest that econometric models of population and employment growth can be tailored to fit the questions that the researcher wishes to answer. If the question bears on the lagged adjustment model, such as an estimate of the speed of adjustment to equilibrium, the researcher should gather data that measure attitudes or policies toward land use regulation, such as the land use data used in this paper. Such data should ideally be measured at the beginning of the study period. If the focus is on the role of location-specific amenities other than land use, a rich set of such variables should be gathered but a relatively simple implementation of the population-employment growth interaction might suffice. For example, for intra-metropolitan models, the ten-mile distance weight matrix (or similar fixed distance matrices) is simple to construct and might suffice if the focus is the role of locational amenities, rather than the interaction between population and employment. If the researcher is testing hypotheses about the interaction between population and employment growth, including the classic question of whether jobs follow people or people follow jobs in urban decentralization, the specification of the weight matrix is somewhat more crucial. The results of hypothesis tests about the interaction between population and employment changes within labor market areas are quite sensitive to the choice of a weight matrix that specifies the labor market area. On *a priori* grounds, one should prefer a weight matrix that is based on data about the spatial extent of labor markets within urban areas, such as the commuting data used for several of the matrices in this paper.

The results in this paper give some insights about the choice of a weight matrix. Note that the contiguity matrices apply the most crude definition of a labor market area, and also give results, in terms of population and employment interactions, that are not consistent with past literature. For example, in Table 6, the non-normalized contiguity matrix shows that population change depends negatively on employment change in a surrounding labor market area, while much recent literature has shown no statistically significant dependence of population change on employment changes for small areas within metropolitan regions (e.g. Boarnet, 1994a; Bollinger and Ihlanfeldt, 1997; Deitz, 1998). Also note that, in terms of the overidentification tests, the normalized tract-to-tract commute flow matrix is slightly better than the non-normalized commute flow matrix, the inverse distance matrix performs slightly better than the ten mile distance-based matrix, and the normalized commute flow matrix might be slightly preferred over the ten-mile distance based matrix.[14] Finally, note that the normalized commute flow matrix, in Tables 6 and 7, gives the result that population change depends on employment changes in a surrounding labor market area and that employment change depends on population changes in a surrounding labor market area. While this does not agree with recent articles (e.g. Boarnet, 1994b, Dietz, 1998) that suggest that employment change depends on surrounding population growth but not vice versa, the simultaneous dependence of population and employment change in Tables 6 and 7 is likely the most reasonable agreement with past research among the specifications in Tables 4-7, and so the normalized commute flow matrix might also be preferred for that reason.

In summary, the results in this paper suggest that the questions being asked of a population-employment growth model will, in part, determine which specification issues should receive the most attention. Recent advances in data availability give the promise of implementing theses models with increasing frequency, allowing more ability to test

[14] One should interpret this with some caution, since the differences in overidentification statistics across some specifications is small.

different model specifications with data from a number of metropolitan areas and rural regions. Overall, the lagged adjustment performance of the model appears to be an issue that can be addressed by improved data availability, and the evidence here suggests that weight matrices based on commute flow data that are specific to the study area should be preferred if questions about population-employment interactions are central to the analysis.

Bibliography

Angrist, John and Alan B. Krueger. 1989. *Why do World War II Veterans Earn More than Nonveterans?* Working Paper No. 2991, National Bureau of Economic Research, Cambridge, MA.

Angrist, John and Alan B. Krueger. 1994. Why do World War II Veterans Earn More than Nonveterans? *Journal of Labor Economics* 12(1): 74-97.

Anselin, Luc. 1980. *Estimation Methods for Spatial Autoregressive Structures: A Study in Spatial Econometrics*. Cornell University Program in Urban and Regional Studies, Regional Science Dissertation and Monograph Series Number 8. Ithaca, NY: Cornell University, Program in Urban and Regional Studies.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Boarnet, Marlon G. 1992. Intra-Metropolitan Growth Patterns: The Nature and Causes of Population and Employment Changes within an Urban Area. Unpublished Ph.D. dissertation, Princeton University, Princeton, NJ.

Boarnet, Marlon G. 1994a. "An Empirical Model of Intrametropolitan Population and Employment Growth." *Papers in Regional Science: The Journal of the RSAI* 73,2: 135-152.

Boarnet, Marlon G. 1994b. "The Monocentric Model and Employment Location." *Journal of Urban Economics* 36: 79-97.

Bollinger, C.R., and Keith Ihlanfeldt. 1997. "The Impact of Rapid Rail Transit on Economic Development: The Case of Atlanta's MARTA". *Journal of Urban Economics*, Vol. 42 (No. 2): 179-204.

Bradford, David F. and Harry H. Kelejian. 1973. An Econometric Model of the Flight to the Suburbs. *Journal of Political Economy* 81: 566-589.

Carlino, Gerald A. and Edwin S. Mills. 1987. The Determinants of County Growth. *Journal of Regional Science* 27: 39-54.

Deitz, Richard. 1998. A Joint Model of Residential and Employment Location in Urban Areas. *Journal of Urban Economics* 44: 197-215.

Deller, Steven C., Tsung-Hsiu (Sue) Tsai, David W. Marcouiller, and Donald B.K. English. 2001. The Role of Amenities and Quality of Life in Rural Economic Growth. *American Journal of Agricultural Economics* 83(2): 352-365.

Fingleton, Bernard. 2001. Equilibrium and Economic Growth: Spatial Econometric Models and Simulations. *Journal of Regional Science* 41(1): 117-147.

Glavac, Sonya M., Alexander C. Vias, and Gordon F. Mulligan. 1998. Population and Employment Interactions in the Growth of United States Micropolitan Centers. *Urban Geography* 19(7): 632-656.

Goffette-Nagot, F., and B. Schmitt. 1999. Agglomeration Economies and Spatial Configurations in Rural Areas. *Environment and Planning A* 31: 1239-1257.

Grubb, W. Norton. 1982. The Dynamic Implications of the Tiebout Hypothesis: The Changing Composition of Boston Communities, 1960-1970. *Public Finance Quarterly* 10: 17-38.

Henry, Mark S., David L. Barkley, and Shuming M. Bao. 1997. The Hinterland's Stake in Metropolitan Growth: Evidence from Selected Southern Regions. *Journal of Regional Science* 37(3): 479-501.

Henry, Mark S., Bertrand Schmitt, Knud Kristensen, David L. Barkley, and Shuming Bao. 1999. Extending Carlino-Mills Models to Examine Urban Size and Growth Impacts on Proximate Rural Areas. *Growth and Change* 30: 526-548.

Leichenko, Robin M. 2001. Growth and Change in U.S. Cities and Suburbs. *Growth and Change* 32(3): 326-354.

Luce, Thomas F. 1994. Local Taxes, Public Services, and the Intrametropolitan Location of Firms and Households. *Public Finance Quarterly* 22(2): 139-167.

Mulligan, Gordon F., Alexander C. Vias and Sonya M. Glavac. 1999. Initial Diagnostics of a Regional Adjustment Model. *Environment and Planning A* 31: 855-876.

Palumbo, G., S. Sacks, and M. Wasylenko. 1990. Population Decentralization within Metropolitan Areas: 1970-1980. *Journal of Urban Economics* 27: 151-167.

Rey, Serge and Marlon Boarnet. 2000. A Taxonomy of Spatial Econometric Models for Simultaneous Equations Systems. In *Advances in Spatial Econometrics* (Luc Anselin and Raymond Florax, eds.) New York: Springer-Verlag.

Roberts, Deborah. 2000. The Spatial Diffusion of Secondary Impacts: Rural-Urban Spillovers in Grampian, Scotland. *Land Economics* 76(3): 395-412.

Schmitt, Bertrand and Mark S. Henry. 2000. Size and Growth of Urban Centers in French Labor Market Areas: Consequences for Rural Population and Employment. *Regional Science and Urban Economics* 30: 1-21.

Steinnes, Donald N. 1977. Causality and Intraurban Location. *Journal of Urban Economics* 4: 69-79.

Steinnes, Donald N. and Walter D. Fisher. 1974. An Econometric Model of Intraurban Location. *Journal of Regional Science* 14: 65-80.

Wheeler, Christopher H. 2001. A Note on the Spatial Correlation Structure of County-Level Growth in the U.S. *Journal of Regional Science* 41(3): 433-449.

Table 1: Variable Definitions and Data Sources, Variables Required by the Structure of the Model

Population Change (Equation 9)	**Data Source**	**Employment Change (Equation 10)**	**Data Source**
Dependent variable = POP? (population change, 1980 to 1990)	1980 Census of Population and Housing	Dependent Variable = EMP? (employment change, 1980 to 1990)	Southern California Association of Governments
1980 Population (coefficient is $?_p$)	Census	1980 Employment (coefficient is $?_e$)	
(I+W)EMP? (instrumented)	Constructed from census data and alternative W matrices	(I+W)POP? (instrumented)	Constructed from census data and alternative W matrices
(I+W)EMP80	Constructed from census data and alternative W matrices	(I+W)POP80	Constructed from census data and alternative W matrices

Table 2: Land Use Variables

Variable Name	**Description**
lu1110	Single Family Residential
lu1120	Multi-Family Residential
lu1140	Mixed Residential
lu1210	General Office Use
lu1220	Retail Stores and Commercial Services
lu1230	Other Commercial
lu1240	Public Facilities
lu1310	Light Industrial
lu1320	Heavy Industrial
lu1340	Wholesaling and Warehousing
lu2000	Agriculture
lu3000	Vacant

Source: Aerial Information Systems. Data converted to 1980 tracts using Access. Variable value is total tract area, in acres, in given use.

Table 3: Place Dummy Variables

Variable Name	Place	Variable Name	Place
pl0070	Anaheim	pl1615	Los Alamitos
pl0325	Brea	pl1786	Mission Viejo
pl0335	Buena Park	pl1915	Newport Beach
pl0398	Capistrano Beach	pl2015	Orange
pl0625	Costa Mesa	pl2195	Placentia
pl0685	Cypress	pl2411	Rossmoor
pl0705	Dana Point	pl2470	San Clemente
pl0903	El Toro	pl2519	San Juan Capistrano
pl0904	El Toro Station	pl2570	Santa Ana
pl1065	Fountain Valley	pl2650	Seal Beach
pl1095	Fullerton	pl2735	South Laguna
pl1110	Garden Grove	pl2800	Stanton
pl1300	Huntington Beach	pl2965	Tustin
pl1347	Irvine	pl2967	Tustin Foothills
pl1420	Laguna Beach	pl3009	Villa Park
pl1423	Laguna Hills	pl3085	Westminster
pl1424	Laguna Niguel	pl3169	Yorba Linda
pl1428	La Habra	pl9999	Unincorporated
pl1477	La Palma		

Source: Census MARF80 data file. Places are 1980 places.

Table 4: Population Change Equation, without land use variables

Population change 1990-1980	Contiguity matrix non-normalized		Contiguity matrix row normalized		10 mile labor market area		Inverse Distance W matrix		Tract-to-Tract commute flows		Tract-to-Tract flows normalized	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
(**I**+**W**)EMP? *	**0.093739**	**2.00**	0.24486	0.97	0.0167616	1.47	-0.051974	-1.86	0.0000643	0.95	-0.4262051	-0.42
1980 Population	**0.256773**	**3.48**	**0.21434**	**2.90**	**0.2376649**	**3.42**	**0.22448**	**3.21**	**0.2283869**	**3.26**	**0.270492**	**2.20**
Proportion Hispanic	**2619.96**	**2.55**	**2635.39**	**2.61**	**2572.043**	**2.59**	**2697.474**	**2.73**	**2819.487**	**2.87**	3556.687	1.64
Proportion Black	-7957.374	-1.22	-5396.3	-0.81	-7805.392	-1.25	-6974.002	-1.12	-6228.215	-0.97	-8890.317	-1.07
(**I**+**W**)EMP80	-0.0104064	-0.98	-0.0363	-0.87	-0.0026292	-0.66	**0.024343**	**2.45**	-0.0000425	-0.96	0.0308469	0.31
% pre 1960 housing	-253.2189	-0.36	-206.24	-0.27	-599.8127	-0.87	-1243.817	-1.70	-549.2514	-0.81	-1144.702	-0.66
% pre 1940 housing	2025.794	1.05	1393.13	0.69	2516.892	1.34	3049.6	1.63	2165.866	1.18	3360.412	0.95
Constant	**-1276.835**	**-2.21**	-942.73	-1.68	**-2466.589**	**-2.79**	-1398.589	-0.84	-708.7609	-1.42	-320.3486	-0.28
Number of obs	415		415		415		415		414		414	
R-squared	0.3259		0.3606		0.3857		0.3846		0.3846		0.216	
Adj R-squared	0.2457		0.2846		0.3127		0.3114		0.3112		0.1225	
Overident. Stat.	15.4795		12.699		1.162		0.913		1.9458		5.5476	

* Instrumented

Notes: Coefficients on census defined place dummy variables are not reported. Significant coefficients (at five percent, two-tailed level) are shown in bold.

Table 5: Employment Change Equation, without land use variables

Employment change 1990-1980	Contiguity matrix non-normalized		Contiguity matrix row normalized		10 mile labor market area		Inverse Distance Weight matrix		Tract-to-Tract commute flows		Tract-to-Tract flows normalized	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
(**I**+**W**)POP? *	0.026157	1.15	**0.440024**	**2.43**	0.0049622	0.83	0.024587	1.45	0.0000134	0.39	0.3289731	1.54
(**I**+**W**)POP80	0.009362	0.61	-0.059418	-0.69	-0.0014919	-1.39	**-0.00785**	**-2.13**	6.51E-06	0.33	-0.057219	-0.67
1980 Employment	**0.062359**	**2.28**	**0.0691933**	**2.38**	**0.0706614**	**2.59**	**0.072663**	**2.67**	0.0151274	0.40	**0.0830045**	**2.87**
Constant	-378.651	-0.77	-123.7939	-0.21	548.2756	0.63	1244.201	0.77	200.4799	0.65	-166.8091	-0.25
Number of obs	415		415		415		415		414		414	
R-squared	0.1374		0.0137		0.1506		0.1567		0.1602		0.0291	
Adj R-squared	0.0451		-0.0918		0.0598		0.0665		0.0701		-0.075	
Overident. Stat.	19.2145		6.889		10.3335		7.4285		18.1332		2.691	

* Instrumented

Notes: Coefficients on census defined place dummy variables are not reported. Significant coefficients (at five percent, two-tailed level) are shown in bold.

Table 6: Population Change Equation, with land use variables

Population change 1990-1980	Contiguity matrix row non-normalized		Contiguity matrix row normalized		10 mile labor market area		Inverse Distance W matrix		Tract-to-Tract commute flows		Tract-to-Tract flows normalized	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
(**I**+**W**)EMP? [*]	**-0.034157**	**-2.65**	0.0288016	0.79	0.00337	0.65	-0.007422	-0.61	0.00005	1.52	**0.30162**	**2.16**
1980 Population	**-0.343061**	**-7.63**	**-0.344166**	**-7.64**	**-0.352973**	**-7.91**	**-0.354728**	**-7.95**	**-0.34278**	**-7.61**	**-0.34394**	**-6.50**
Proportion Hispanic	**3143.001**	**6.22**	**3096.254**	**6.11**	**3041.70**	**5.95**	**3044.14**	**5.99**	**3194.97**	**6.29**	**2512.82**	**3.78**
Proportion Black	**10998.27**	**3.41**	**11116.94**	**3.42**	**10874.05**	**3.37**	**11107.89**	**3.45**	**11285.78**	**3.39**	**12326.09**	**3.19**
(**I**+**W**)EMP80	**0.01191**	**2.91**	0.01014	0.79	0.00028	0.15	**0.00925**	**2.00**	-0.00003	-1.21	-0.02126	-1.03
% pre 1960 housing	525.25	1.51	609.06	1.74	476.39	1.33	325.56	0.87	487.96	1.38	**972.26**	**2.12**
% pre 1940 housing	1055.55	1.11	947.28	0.99	1302.03	1.35	1405.06	1.46	1071.35	1.13	230.94	0.20
lu1110	**5.7800**	**17.09**	**5.8224**	**17.15**	**5.8668**	**17.33**	**5.9032**	**17.41**	**5.8025**	**17.09**	**5.6682**	**13.97**
lu1120	**15.0057**	**15.43**	**15.0308**	**15.44**	**15.0884**	**15.51**	**15.0947**	**15.53**	**14.8830**	**15.10**	**14.0277**	**11.21**
lu1140	**16.9417**	**2.35**	**16.8132**	**2.33**	**18.0771**	**2.50**	**18.2967**	**2.53**	**17.1275**	**2.37**	**18.3821**	**2.15**
lu2000	**1.4577**	**6.94**	**1.4405**	**6.82**	**1.3493**	**6.40**	**1.3490**	**6.41**	**1.4636**	**6.89**	**1.6466**	**6.07**
lu3000	0.20987	1.63	0.21564	1.67	0.21164	1.64	0.19627	1.51	0.20887	1.62	0.18049	1.18
Constant	**-551.295**	**-2.05**	**-678.8311**	**-2.60**	**-1313.943**	**-2.93**	**-2051.389**	**-2.51**	**-568.5665**	**-2.21**	**-936.374**	**-2.79**
Number of obs	415		415		415		415		414		414	
R-squared	0.8438		0.8435		0.8433		0.8437		0.8428		0.7826	
Adj R-squared	0.8228		0.8225		0.8223		0.8227		0.8217		0.7533	
Overident. Stat.	14.4005		12.4085		18.2185		14.4835		18.63		5.5062	

[*] Instrumented

Notes: Coefficients on census defined place dummy variables are not reported. Significant coefficients (at five percent, two-tailed level) are shown in bold.

Table 7: Employment Change Equation, with land use variables

Employment change 1990-1980	Contiguity matrix non-normalized		Contiguity matrix row normalized		10 mile labor market area		Inverse Distance Weight matrix		Tract-to-Tract commute flows		Tract-to-Tract flows normalized	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
(**I**+**W**)POP? *	-0.01055	-0.79	0.04854	0.87	-0.00204	-0.50	0.00635	0.55	-0.00002	-0.60	**0.11225**	**2.09**
(**I**+**W**)POP80	-0.00843	-0.78	0.04465	1.10	-0.00055	-0.73	-0.00500	-1.95	0.00001	0.67	0.00161	0.04
1980 Employment	**-0.602315**	**-13.60**	**-0.5994637**	**-13.58**	**-0.6019933**	**-13.60**	**-0.600693**	**-13.59**	**-0.5898267**	**-13.01**	**-0.582877**	**-12.97**
lu1210	**20.9106**	**8.93**	**21.1851**	**9.03**	**20.6109**	**8.76**	**20.9078**	**8.95**	**20.7365**	**7.34**	**20.7253**	**8.75**
lu1220	**11.7312**	**3.99**	**9.9535**	**3.37**	**11.1178**	**3.88**	**11.4449**	**3.99**	**10.5194**	**3.60**	**10.0849**	**3.49**
lu1230	**14.5123**	**5.37**	**14.3524**	**5.34**	**14.8802**	**5.49**	**14.6136**	**5.39**	**14.1730**	**5.25**	**14.0942**	**5.23**
lu1240	**44.5517**	**5.71**	**41.4688**	**5.28**	**43.7299**	**5.64**	**42.8619**	**5.54**	**43.4047**	**5.56**	**40.2019**	**5.12**
lu1310	**18.8777**	**12.46**	**18.6545**	**12.22**	**18.7897**	**12.35**	**18.6350**	**12.29**	**18.3939**	**11.70**	**17.9902**	**11.66**
lu1320	**32.6074**	**2.25**	**32.8489**	**2.27**	**32.9548**	**2.28**	**32.4366**	**2.25**	**31.4103**	**2.10**	**39.9358**	**2.69**
lu1340	-7.14453	-1.55	-7.24183	-1.58	-7.05655	-1.55	-6.68308	-1.47	-5.94007	-1.29	-6.08219	-1.33
lu2000	-0.31963	-1.10	-0.60872	-1.85	-0.39097	-1.43	-0.39164	-1.43	-0.40428	-1.46	**-0.78150**	**-2.38**
lu3000	-0.20649	-1.28	**-0.37151**	**-2.03**	-0.25192	-1.60	-0.25165	-1.60	-0.23484	-1.49	**-0.46075**	**-2.46**
Constant	507.81	1.38	-195.36	-0.54	1145.09	1.84	1975.59	1.71	204.73	0.86	-92.07	-0.21
Number of obs	415		415		415		415		414		414	
R-squared	0.5866		0.5908		0.588		0.5902		0.576		0.577	
Adj R-squared	0.5311		0.5359		0.5327		0.5352		0.5189		0.5201	
Overident. Stat.	45.401		35.6485		33.9055		30.2535		34.6932		30.3048	

* Instrumented

Notes: Coefficients on census defined place dummy variables are not reported. Significant coefficients (at five percent, two-tailed level) are shown in bold.