RSP 8(2): 101-115.

A SIMULTANEOUS ECONOMETRIC MODEL OF THE INTRAURBAN LOCATION OF EMPLOYMENT AND RESIDENCE

Donald N. Steinnes*

Introduction

Urban economics is perhaps the one branch of economics which has developed as much from other disciplines (e.g., sociology, geography, and city planning) as from economics. Nonetheless, the urban economics literature has been replete with sophisticated models and the use of advanced statistical estimating techniques which are the common tools of the social sciences. However, in many cases these efforts have ignored another important branch of economics - econometrics. Goldberger [5, p. 4] suggests that econometrics, "is not simply a matter of fitting curves to data, of 'measurement without theory'". Rather, econometrics is concerned with the quantifying, through statistical estimation, of relationships which are derived from economic theory. It could be contended that the interdisciplinary nature of urban economics has slowed the development of urban or spatial econometrics.

Though the spatial nature of most data used to estimate urban economic models leads to unique estimation problems,[1] the purpose in this paper is to explore certain more basic econometric aspects of one important part of urban economics - intraurban location. It is possible to differentiate three causal specifications for an intraurban econometric model, each with its own appropriate estimation procedure. In econometrics a simple model is an equation or system of equations (relationships) in which each endogenous (Y) variable is determined solely by exogenous (X) variables. In this context an exogenous variable is one not explained by the model whereas endogenous variables are explained by the model. Each equation of a simple model is appropriately estimated by ordinary least squares (OLS). A recursive model is one wherein each of the endogenous variables is determined successively. For instance, Y_1 is determined by only exogenous (X) variables, Y_2 is determined by Y_1 and exogenous variables, Y_3 is determined by Y_1, Y_2 and exogenous variables, etc. Each of the relationships in such a system may also be estimated by OLS. Finally, there is a simultaneous equation model in which each endogenous (Y) variable may affect and be affected by every other endogenous variable. Estimation of such a system, when possible, requires more advanced estimation procedures than OLS.[2]

*Associate Professor, Department of Economics, University of Minnesota-Duluth.

[1]These problems include spatial autocorrelation, which is analogous to serial correlation for time series data, and statistical gerrymandering which is a special type of data aggregation problem.

[2]Details on these procedures, as well as the theoretical rationale for them, can be found in many econometrics books (e.g., Goldberger [5] or Theil [21]).

Before exploring the econometrics of intraurban location models[3] it is necessary to review and summarize the range of such models which have been put forth in the literature. The emphasis of this review will be on the extent to which the models have been derived from economic theory since it is only such theoretical models which can be considered from an econometric point of view.

A basic distinction can be drawn between urban simulations and what are considered to be more theoretical models. Large scale urban simulations were introduced around 1960 by city planners and others with great expectations but have since passed out of fashion for a number of reasons.[4] From the point of view of this paper the shortcoming of these simulations, as well as those later developed by economists,[5] was their lack of a basis in economic theory. The most popular Lowry [12] model and its derivatives is based on a gravity-type of mechanism rather than on economic theory. Though not econometric models these simulations can be classified on a causal basis. In almost all cases such simulations have treated employment location as given or exogenous and proceeded to explain the distribution of population (residences). In some cases (e.g., Lowry [12])employment has been categorized as export (or base) and population serving (retail trade) resulting in a recursive type of model where population (Y_1) distribution is first found given export employment (X) and retail trade (Y_2) is then located given population (Y_1) and export employment (X). Finally, a sort of simultaneity may be introduced via feedbacks (e.g., on subsequent iterations population (Y_1) is distributed given retail trade (Y_2) location).

The more theoretical intraurban location models[6] proposed by economists have concentrated on explaining residential location while assuming employment to be exogenous. In most of these models an even more restrictive, and often criticized, assumption has been made--that all employment is located in the center of the city. Though based on this questionable assumption of exogenous employment, these models, nonetheless, are derived from economic theory and, hence, constitute legitimate simple econometric models which should be estimated using OLS.

In general, models attempting to explain the intraurban location of employment have less of a theoretical basis than the models of residential location just discussed. In some models (e.g., Berry [2], Kain and Meyer [9] and Lowry [12]) residence, as a source of demand, is seen as a determinant of retail trade location but, in general, residence, as a source of labor, is not seen as determining employment location. An exception is a model tested by Moses and Williamson [14] where residence in a subarea is an exogenous variable determining the number of firms that will locate in a subarea.

---

[3] It is possible to classify such models as being of two types: those concerned with the location of employment (jobs) and those concerned with the location of residences (people) within an urban area.

[4] See Lee [11] or Goldstein and Moses [6] for further elaboration.

[5] For example, Ingram, Kain and Ginn [8] and Lowry [12].

[6] For example, Alonso [1], Mills [13], Muth [15], and Wingo [20].

To summarize, models explaining residential location assume employment exogenously determined while models explaining employment location assume residence to be exogenous. This suggests that there exists a contradiction[7] between the two types of models in their conception of the causal relationship between the housing (residence) and labor (employment) markets in an urban area. A reasonable contention then is that these two types of models cannot both be theoretically correct since they are contradictory. From an econometric standpoint this would mean that OLS estimations of such contradictory models will be suspect.

A model which resolves this conflict has been derived from economic theory by Steinnes and Fisher [19] which treats both residence and employment as endogenous as a result of assuming location decisions made in the housing and labor markets are interdependent. Given this model has a theoretical basis it can be viewed as a simultaneous equation econometric model and, as previously discussed, should be estimated using some simultaneous estimation procedures (e.g., two-stage-least-squares (TSLS))[8] rather than OLS.

It will be the purpose in the remainder of this paper to specify and estimate using TSLS for this simultaneous econometric model. Corresponding results using OLS will also be presented on the grounds that they represent the contradictory models previously discussed. The comparison of TSLS and OLS estimation results will lead to a conclusion that OLS estimation of the contradictory models lends false statistical validity to the predominant assumption in the literature--that employment location determines residential location but not vice versa. It will be shown that correct estimation (i.e., TSLS) of what is believed to be a theoretically sound simultaneous model yields opposite statistical evidence regarding causality. That is, residential location is more significant in determining employment location than vice versa. This is the same conclusion the author has reached in related studies [17, 18] of suburbanization using a different data base and model.

Specification of the Model

Assume that places of employment (firms) are classified into n "industries" and that individuals are classified into m race-occupation categories. In the

---

[7]This contradiction is most striking in the recent work of Fales and Moses [3]. Here the authors present, but do not derive, a model which they estimate using OLS. The model attempts to explain residential and employment location assuming employment is an exogenous determinant of residence at the same time that residence is an exogenous determinant of employment location.

[8]It should be noted that there is some precedent in the literature for the use of simultaneous estimation procedures. For example, Hill [7] uses TSLS to estimate a model devoid of theoretical foundation while Muth [15] employs the same technique on a model already estimated using OLS. However, Muth switches from OLS to TSLS without providing a theoretical derivation for the simultaneous model as he had done previous to the OLS estimations.

empirical work there are two types of industry (manufacturing and nonmanufacturing) and four race-occupation categories (white-white-collar, white-blue-collar, black-white-collar, and black-blue-collar),[9] so $n = 2$ and $m = 4$.

Assume that K (the empirical estimation is based on a sample of 100 subareas of metropolitan Chicago) geographic subareas of the metropolitan area are defined, including the city core as one such subarea, and that a scalar measure of commuting accessibility, $a_{kk'}$, between area k and area k' is available for all pairs. This accessibility coefficient is to be regarded as a generalized measure of the ease of traveling between the two areas. Then define the point accessibility of area k as:

$$(1) \quad a_k = \sum_{k' \neq k} a_{kk'}$$

which may be interpreted as an overall accessibility of area k to or from the other areas. Define, for any variable Y_k, whether vector or scalar, that takes on values for all areas, the "potential value."[10]

$$(2) \quad \bar{Y}_k = 1/a_k \sum_{k \neq k} a_{kk'} Y_k$$

Then the model consists of two vector equations (subscript k omitted):

$$(3) \quad R = f_R\,(E, \bar{E}, A_R, \bar{A}_I, \bar{X}_0, P, a_k) + U_1$$

$$(4) \quad E = f_E\,(R, \bar{R}, A_I, \bar{A}_R, X_0, \bar{P}, a_k) + U_2$$

where f_R and f_E denote linear functions with constant terms, and where U_1 and U_2 are random vectors of $m = 4$ and $n = 2$ elements, respectively.

The endogenous variables, R and E, have previously been defined and given

---

[9] The white collar group comprised the following census occupations: professional, technical and kindred workers; managers, officials and proprietors, including farmers and farm managers; clerical and kindred workers; and sales workers. Blue collar designates workers in the following occupational groups: craftsmen, foremen and kindred workers; operatives and kindred workers; private household workers; service workers, except private household; and laborers, except mining, but including farm laborers and farm foremen. Each of these categories are subdivided into racial categories "white" and "black", where "black" denotes all non-white races. The number of employed residents in each subarea in 1960 for the four resulting categories will be denoted by the codewords WWC, WBC, BWC, and BBC. The employment variables, MFGEMP (manufacturing) and NONEMP (nonmanufacturing), measure the number of employees working in each of these industries in a subarea in 1960.

[10] Thus, potential in this study is defined as a kind of weighted average, whereas the more customary definition of potential is a weighted sum equal to $\bar{Y}_k$ times a_k.

codewords. Potential values for R and E are regarded as predetermined[11] and given codewords ending in PT (e.g., MFGMPT). Likewise, any potential variable will have a codeword ending in PT.

The vector of residential amenities, A_R, includes: LANDRE,[12] amount of land utilized for residential purposes; GHETTO, categorical variable indicating if an area is in or near ghetto; COLFAC, number of college faculty employed in an area (a proxy for the cultural attractiveness of an area); PARK60, number of park employees in an area (a proxy for recreational attractiveness of an area); ONLAKE, dummy variable indicating if an area is on Lake Michigan; RAPIDT, dummy variable indicating if area has mass transit (i.e., elevated or subway); and PROTAX, property tax rate of an area.

Included in A_1, the vector of industrial amenities, are: LANDEM, amount of land utilized for industrial purposes; TRZONE, dummy variable indicating if an area is in trucking zone; and INDPRK, number of acres of industrial parks in an area. Included in X_0 are activities whose location depends primarily on decisions made outside of the metropolitan area being studied. They are: HOSBED, number of beds in state and federal hospitals in an area; COLFAC, defined above; and EXOEMP, number of persons employed in various exogenous activity (e.g., military installations). The variable P denotes purchasing power and is measured in the empirical work by MEDINC, median income of residents in an area. a_k has been previously defined.

Equation (3) may be termed a "residential location" function that seeks to explain the distribution of residents by race-occupation categories over the geographic areas. Explanatory variables are of two types: those affecting the level of supply and demand for residences in the same area as that of residence being explained, and those affecting the demand for labor in neighboring or accessible areas (the barred or potential variables in this equation). Similarly,

---

[11] In this paper the potentials are assumed to be predetermined variables by analogy with linear combinations of lagged variables in time series analysis. It is beyond the scope of this paper to examine the questionable validity of this assumption. Such an examination would lead into a relatively new field of stochastic processes over space.

[12] The specification of LANDRE, amount of land used for residential purposes, and LANDEM, amount of industrial land used, as exogenous variables may certainly be questioned. A larger model was formulated which included additional equations explaining the determination of these two variables as endogenous variables, as well as the utilization of agricultural land, and land rent, as has been done in previous well known urban models. Such a formulation is undoubtedly superior from the viewpoint of economic theory over the present one, in which it is essentially assumed that the allocation of land is made prior to the allocation of people. The decision to proceed this way was made for the following reasons: (1) the present model is simpler, and is worth testing; and (2) the data base available was not adequate to test the more elaborate model, so in the empirical work it was necessary to assume the two land variables exogenous. The more elaborate model is presented in Steinnes [16].

Equation (4) may be referred to as the "employment location" function that explains the distribution of employment by industrial type by means of exogenous variables that are assumed to affect demand for labor in the area, and also includes potential variables that are assumed to affect the supply and demand for residences in accessible areas.

The model thus presented as Equations (3) and (4) has been derived in Steinnes and Fisher [19] from conventional theoretic postulates that involve production functions, utility functions, and transportation costs. The derivation also is based on the supply and demand relations for labor and residences in each subarea referred to above in defining Equations (3) and (4) as the "residential" and "employment" functions.

Estimation Results

In this section estimation results will be presented for the model (Model 1) specified in the last section as Equations (3) and (4) using TSLS, an appropriate (since the equations are overidentified by the order condition) simultaneous estimation procedure. These results, in Tables 1 and 3, will be compared to the corresponding OLS results in Tables 2 and 4. These OLS results would represent the appropriate estimation of two separate, simple (though theoretically contradictory) models (Table 2 being a model (Model 2) explaining residential location assuming employment exogenous and Table 4 a model (Model 3) explaining employment location assuming residence exogenous) that have been used to explain intraurban location.[13] From an econometric standpoint these simple models would be appropriately estimated using OLS were it not for the theoretical contradiction of the models which invalidates such OLS estimations. One goal of this paper will be to point out differences in the TSLS and OLS results so as to illustrate what false conclusions can be arrived at using OLS (i.e., accepting contradictory models).

As indicated in the previous section, Equations (3) and (4) apply to subareas within an urban area. The estimations presented in Tables 1 through 4 were made using data for 1960 for a sample of 100 scattered corporate suburbs and Community Areas[14] (neighborhood size areas defined for the city of Chicago) of the Chicago SMSA (Standard Metropolitan Statistical Area).

---

[13]Model 2 can be considered a representation of the Muth [15] model though in his model residential density is the dependent variable. Using density as the dependent variable means the estimated coefficients measure the effects of independent variables on density. However, the coefficients in the estimations of Equation (3) have a similar interpretation since LANDRE has been included as one of the independent variables. Model 3 can be considered a representation of models such as Moses and Williamson [14] and the models of retail trade discussed in the Introduction.

[14]Areas defined in Kitagawa and Taeuber [10].

TABLE 1: Equation (3) Estimation Results Using Two Stage Least Squares

Dependent Variables:	White Residents		Black Residents	
	white collar	blue collar	white collar	blue collar
	R_1 (persons)	R_2 (persons)	R_3 (persons)	R_4 (persons)
Right-hand Variables:	Regression Coefficients (t values in parentheses)			
MFGEMP E_1[a]	-.12	.17	-.04	-.07
(persons)	(-.68)	(1.18)	(-1.35)	(-.99)
NONEMP E_2[a]	-.24	-.03	.03	.09
(persons)	(-1.59)	(-.26)	(1.23)	(1.37)
MFEMPT $\overline{E}_1$[a]	.43	.86	-.004	-.35
(persons)	(.47)	(1.15)	(-.13)	(-.92)
NMEMPT $\overline{E}_2$[a]	.69	.20	-.05	.17
(persons)	(.52)	(.18)	-.22)	(.31)
LANDRE (000 sq. ft.)	.15 (4.87	.11 (4.43)	.01 (1.33)	.01 (.97)
GHETTO[b]	-2731 (-3.46)	-2285 (-3.55)	722 (5.62)	2057 (6.33)
COLFAC (persons)	3.89 (.37	-20 (-2.27)	.40 (.23)	.92 (.21)
MEDINC (dollars)	-.39 (-1.24)	-.64 (-2.52)	-.04 (0.88)	-.06 (-.48)
PARK60 (persons	193 (2.42)	162 (2.50)	-7.04 (-.55)	-16 (-.47)
ONLAKE (1 if on lake, 0 otherwise)	745 (.31)	700 .35)	- 437 (-1.10)	-2299 (-2.30)
RAPIDT (1 if transit, 0 otherwise)	4274 (1.53)	3625 (1.59)	436 (.96)	1947 (1.69)
PROTAX (per $1000 value)	3190 (2.44)	2293 (2.15)	-1.65 (-.77)	-7.62 (-1.41)

[a]Treated as simultaneously determined

[b]4 = over 50% nonwhite in both 1950 and 1960; 3 = 50% or more increase in nonwhites 1950-1960; 2 = areas adjacent to areas classified 3 or 4; 1 = areas not classified 3 or 4 which adjacent to areas classified 2; 0 = all other areas.

TABLE 2: Equation (3) Estimation Results Using Ordinary Least Squares

Dependent Variables:	White Residents		Black Residents	
	white collar	blue collar	white collar	blue collar
	R_1 (persons)	R_2 (persons)	R_3 (persons)	R_4 (Persons)
Right-hand Variables:	Regression Coefficents (t values in parentheses)			
MFGEMP E_1 (persons)	.16 (1.81)	.23 (2.88)	-.05 (-3.11)	-.14 (-3.56)
NONEMP E_2 (persons)	-.21 (-2.11)	-.09 (-.99)	.02 (1.34)	.08 (1.80)
MFEMPT $\bar{E}_1$ (persons)	1.15 (1.24)	1.85 (2.28)	-.12 (-.75)	-.70 (-1.72)
NMEMPT $\bar{E}_2$ (persons)	-1.09 (-.72)	-.59 (-.44)	.20 (.73)	1.03 (1.56)
LANDRE (000 sq. ft.)	.13 (4.95)	.12 (4.94)	.01 (1.55)	.02 (1.42)
GHETTO[a]	-2484 (-3.53)	-2262 (-3.66)	660 (5.32)	1890 (6.14)
COLFAC (persons)	7.31 (.74)	-16 (-1.84)	-.02 (-.01)	-1.12 (-.26)
MEDINC (dollars)	-.30 (-1.08)	-.70 (-2.93)	-.04 (-.89)	-.07 (-.57)
PARK60 (persons)	133 (2.00)	135 (2.29)	-5.4 (-.46)	-3.4 (-.11)
ONLAKE (1 if on lake, 0 otherwise)	2012 (.95	1062 (.57)	-459 (-1.22)	-2503 (-2.69)
RAPIDT (1 if transit, 0 otherwise)	.898 (.79)	2102 (1.00)	541 (1.28)	2561 (2.45)
PROTAX (per $1000 value)	3659 (3.28)	2207 (2.25)	-154 (-.78)	-858 (-1.76)
R^2	.65	.68	.53	.66

[a]See Table 1.

TABLE 3: Equation (4) Estimation Results Using Two Stage Least Squares

Dependent Variables:	Manufacturing Employment E_1 (persons)	Nonmanufacturing Employment E_2 (persons)
Right-hand Variables:	Regression Coefficients (t values in parentheses)	
WWC R_1[a] (persons)	.37 (.97)	.41 (.87)
WBC R_2[a] (persons)	.12 (.25)	-.82 (-1.44)
BWC R_3[a] persons	-7.67 (-1.64)	- 15.8 (-2.73)
BBC R_4[a] (persons)	3.07 (1.85)	4.46 (2.17)
WWCPT $\overline{R}_1$ (persons)	-3.61 (-1.03)	8.48 (1.96)
WBCPT $\overline{R}_2$ (persons)	9.44 (2.49)	-6.29 (-1.34)
BWCPT $\overline{R}_3$ (persons)	- 122 (-2.03)	-202 (-2.70)
BBCPT $\overline{R}_4$ (persons)	48 (1.91)	80 2.53)
LANDEM (000 sq. ft.)	.13 (1.29)	.23 (1.75)
CHETPT[b]	-8960 (-.58)	40011 (2.08)
ONLKPT[b]	75026 (1.57)	74986 (-1.27)
PARKPT (persons)	-3439 (-1.94)	2935 (1.34)
HOSBED (beds)	1.91 (1.85)	1.23 (.97)

[a]Treated as simultaneously determined.

[b]See Table 1.

TABLE 4: Equation (4) Estimation Results Using Ordinary Least Squares

Dependent Variables:	Manufacturing Employment E_1 (persons)	Nonmanufacturing Employment E_2 (persons)
Right-hand Variables:	Regression Coefficients (t values in parentheses)	
WWC R_1 (persons)	-.04 (-.15)	.11 (.48)
WBC R_2 (persons)	.53 (2.39)	-.19 (-.77)
BWC R_3 (persons)	-3.55 (-1.39)	-5.81 (-2.05)
BBC R_4 (persons)	1.22 (1.23)	1.52 (1.39)
WWCPT $\overline{R}_1$ (persons)	-4.16 (-1.30)	6.69 (1.88)
WBCPT $\overline{R}_2$ (persons)	10 (3.07)	-3.65 (-.99)
BWCPT $\overline{R}_3$ (persons)	(-51 (-1.25)	-109 (-2.41)
BBCPT $\overline{R}_4$ (persons)	19 (1.07)	46 (2.31)
LANDEM (000 sq. ft.)	.09 (1.15)	12 (1.40)
GHETPT (See Table 1)	6080 (-.42	30673 (1.89)
ONLKPT (see Table 1)	94784 (2.23)	-30319 (-.64)
PARKPT (persons)	-3276 (-2.00)	3536 (1.94)
HOSBED (beds)	2.13 (2.22)	1.37 (1.37)
R^2	.68	.59

In order to concentrate attention on the differences between the three models proposed no attempt will be made to discuss each of the independent variables and their significances.[15] In fact, it should be noted that the estimations presented included other independent variables, as indicated in the last section, but those that did not prove significant in testing have not been included in the tables. The overall fit, as measured by R^2, is about the same in Tables 2 and 4 and is acceptable for an urban cross-section study.

There is one empirical finding in Tables 3 and 4 which should be contrasted with a finding in Tables 1 and 2. As mentioned, in Steinnes and Fisher [19] it was concluded from estimation of Equation (3) that racial differences existed with regards to the determinants of residential location because coefficients of independent variables tended to have different signs for the two races. The estimation of Equation (4) (i.e., Tables 3 and 4) exhibit an opposite result. That is, R (and $\bar{R}$) as independent variables do not have opposite signs for the two races. In fact, the difference tends to be occupational. White residence variables (R_1, R_2, $\bar{R}_1$, and $\bar{R}_2$) in Tables 3 and 4 exhibit opposite signs for the two equations (E_1 and E_2) with the exception of R_1 in E_1 equation of Table 3. Furthermore, R_1 and $\bar{R}_1$ being positive for E_2 while R_2 and $\bar{R}_2$ are positive for E_1 can be explained, in part, by the relative concentration of respective occupations in these industries. In other words, there is a tendency for manufacturing to be located in or near blue-collar (white) residential area while nonmanufacturing tends to be in or near white-collar (white) residential areas. For blacks there are occupational differences in the effect on employment location but the signs are the same for E_1 and E_2. The positive effect of R_4 and $\bar{R}_4$ on E_1 can again be explained on the basis of the relative concentration of blue-collar occupations in manufacturing. R_4 and $\bar{R}_4$ being positive in the E_2 equation may be explained by the fact that E_2 includes service industries which employ the blue-collar occupations in which blacks are concentrated. While a fuller investigation is necessary, it can be conjectured that employment location exhibits occupational, not racial, differences while residential location exhibits racial, not occupational, differences.[16]

This much aside, the next, and more important, task is to compare the three models proposed and estimated in Tables 1 through 4. First it should be noted that similarities exist between the OLS and TSLS estimations of Equations (3) and (4) in that the signs for the independent variables are, in most cases, the same using either method of estimation. Where there is a difference (e.g.,

---

[15]The estimation results in Tables 1 and 2 were analyzed in Steinnes and Fisher [19]. One of the important results was the tendency for the independent variables to have racial, and not occupational, differences in signs. This conclusion being that failure to allow for race will lead to false conclusions regarding the determinants of residential location. The same conclusion was sustained in subsequent testing of the model by Fisher and Fisher [4] using a different data base.

[16]This conjecture is reinforced by the findings in the recent study by Fisher and Fisher [4] referred to earlier.

NMEMPT ($\bar{E}_2$) in WWC (R_1) equation of Tables 1 and 2) the variable is insignificant.

Given the independent variables do not change signs in the OLS and TSLS estimations, the next comparison is their significance (t value) using the two procedures. Taking all the independent variables in Tables 1 and 2, except E and $\bar{E}$, and in Tables 3 and 4, except R and $\bar{R}$, it can be seen that their significance is, for the most part, the same using OLS and TSLS. Hence, it remains to examine the significance of the independent variables R, $\bar{R}$, E, and $\bar{E}$ in Tables 1 through 4, which will be the crucial comparison of the OLS and TSLS results.

Looking at Model 2 (Table 2) and Model 3 (Table 4) in terms of the significance of the R, $\bar{R}$, E, and $\bar{E}$ variables it might be concluded that Model 2 is superior. This conclusion would follow from the fact that in Table 2 the variables E and $\bar{E}$ are significant (i.e., $|t \text{ value}| > 1.5$) 56 percent of the time (in nine of 16 instances) whereas in Table 4 the variables R and $\bar{R}$ are only significant 38 percent of the time (in six of 16 instances). While such an argument would provide some statistical justification for the widespread use of Model 2 in the intraurban literature, the models nonetheless contradict one another theoretically. What, then, can be learned by examining the TSLS results for Model 1 in Tables 1 and 3? It has been contended that Model 1 is superior to Models 2 and 3 on theoretical grounds and, hence, if the OLS and TSLS results differ reliance should be put on TSLS since the OLS results may be spurious.

Turning to the R equations and again using $|t \text{ value}| > 1.5$ as a measure of significance for E and $\bar{E}$ in Table 1 it is found that these variables are only significant six percent of the time (in one of 16 instances) whereas in Table 2 they were significant 56 percent of the time. This suggests that using OLS (or Model 2) tends to overestimate the significance of employment location as a determinant of residential intraurban location. On the other hand, an examination of Table 3 reveals that R and $\bar{R}$ are significant ($|t \text{ value}| > 1.5$) 63 percent of the time (in ten of 16 instances) whereas in Table 4 they were significant only 38 percent of the time. This suggests that using OLS (or Model 3) tends to underestimate the significance of residential location as a determinant of employment of employment location.

To summarize, what the results demonstrate is that failing to understand or realize the causal contraction of Models 2 and 3 one may on statistical grounds (using OLS) conclude that Model 2 is superior to Model 3. This could explain, in part, the predominance of Model 2, especially the forms which are descriptive in nature. Another purpose of this paper has been to answer the question, "Given Model 1 is the correct one, which is supported by its theoretical foundation and the causal contradiction in the alternative models (Model 2 and Model 3), and hence, TSLS is a proper estimation procedure, what false conclusions will be arrived at using OLS?" First, that employment is an important determinant of residential location and, second, that residence is not an important determinant of employment location. This coincides with the prevailing assumption made in the modeling of intraurban location. It has been concluded here by looking at the TSLS results, that, in fact, the opposite is true. That is, residential location is an important determinant of employment location but not vice versa. As shown in Steinnes and Fisher [19], residential location is determined primarily by racial considerations which most previous models have failed to incorporate because they only attempted to explain total residents (or population) in a subarea without disaggregating by occupation and race.

Conclusion

In this paper an attempt has been made to inject into urban economics some basic tenets of econometrics which already permeate the other branches of economics. The cornerstone of any econometric model, a basis in economic theory, was found to be lacking in many models of intraurban location. Furthermore, a discussion of various types of econometric (i.e., simple, recursive, and simultaneous) models and their appropriate estimation techniques revealed a contradiction between certain urban economic models as to their specification of the causal relationship between the intraurban location of residents and employment. In such a situation estimation results of the contradictory models may be spurious.

In order to resolve this contradiction a simultaneous econometric model with a theoretical basis is presented. An attempt is then made to compare estimation of this simultaneous model (using TSLS) with OLS estimation of the contradictory models. The main conclusion is that failing to allow for the simultaneity of residence and employment will lead to overestimating the significance of employment as a determinant of residential location. It is further suggested that the study of intraurban location has, perhaps, overlooked or dismissed the importance of residential location as a determinant of employment location.

While the model presented is admittedly static, the results suggest that the future formulation and testing of dynamic models of intraurban location should allow for the interdependence of employment and residence. Furthermore, the development of such dynamic models, which the author has already begun work on [17, 18], will allow for a more definitive answer as to the causal relationship between employment and residential intraurban location.

REFERENCES

1. Alonso, W. "A Theory of the Urban Land Market," Papers and Proceedings of the Regional Science Association, 6 (1960), 149-158.

2. Berry, B. Commercial Structure and Commercial Blight, University of Chicago, Department of Geography, Research Paper No. 85, 1963.

3. Fales, R. and L. Moses. "Land-Use Theory and the Spatial Structure of the Nineteenth-Century City," Papers and Proceedings of the Regional Science Association, 28 (1972), 49-80.

4. Fisher, W. and M. Fisher. "The Spatial Allocation of Employment and Residence Within a Metropolitan Area," Journal of Regional Science, 15 (1975), 261-276.

5. Goldberger, A. Econometric Theory, New York: John Wiley Press, 1964.

6. Goldstein, G. and L. Moses. "A Survey of Urban Economics," Journal of Economic Literature, 11 (1973), 471-515.

7. Hill, D. "A Growth Allocation Model for the Boston Region," Journal of the American Institute of Planners, 31 (1965), 111-120.

8. Ingram, G., J. Kain and J. Ginn. The Detroit Prototype of the NBER Urban Simulation Model, New York: Columbia University Press, 1972.

9. Kain, J. and J. Meyer. "A First Approximation to the Rand Model for the Study of Urban Transportation," Rand Corporation Report RM-2878-FF, Los Angeles, 1961.

10. Kitagawa, E. and K. Taeuber, eds. Local Community Fact Book, Chicago Metropolitan Area, 1960, Chicago: University of Chicago Press, 1962.

11. Lee, D. B., Jr. "Requiem for Large Scale Models," Regional Institute for Urban and Regional Development, Berkeley: University of California, April 1972.

12. Lowry, I. "A Model of Metropolis," Rand Corporation Report RM-4035-RC, Los Angeles, 1964.

13. Mills, E. "An Aggregative Model of Resource Allocation in a Metropolitan Area," American Economic Review, 17 (1967), 197-210.

14. Moses, L. and H. Williamson, Jr. "The Location of Economic Activity in Cities," Papers and Proceedings of the American Economic Association, 57 (1967), 211-222.

15. Muth, R. Cities and Housing: The Spatial Pattern of Urban Residential Land Use, Chicago: University of Chicago Press, 1969.

16. Steinnes, D. "An Econometric Model of Residential Location," unpublished Ph.D. thesis, Northwestern University, 1972.

17. ___________. "Causality and Intraurban Location," Journal of Urban Economics, 4 (1977), 69-79.

18. ___________. "The Future F-test: A Statistical Analysis of the Suburbanization Process," presented at the Western Regional Science Association meeting, Tucson, Arizona, 1977.

19. ___________, and W. Fisher. "An Econometric Model of Intraurban Location," Journal of Regional Science, 14 (1974), 65-80.

20. Wingo, L. "An Economic Model of the Utilization of Urban Land for Residential Purposes," Papers and Proceedings of the Regional Science Association, 7 (1961), 191-205.

21. Theil, H. Principles of Econometrics, New York: John Wiley Press, 1971.