Residential Density Structure: An Analysis and Forecast With Evaluation

CARL N. SWERDLOFF, Highway Research Engineer, U. S. Bureau of Public Roads

This report is on a practical investigation of the residential density structure of a typical smaller sized urban area, Greensboro, N. C. Utilizing a rather extensive supply of land-use and transportation data for 1948 and 1960, the analysis followed two major lines of pursuit: (a) analysis of the existing 1948 residential density structure with particular emphasis on investigation of classical mathematical expressions of a distance-gradient nature, and (b) comparative analysis of the outcomes of several simplified attempts at forecasting the 1960 density structure.

•THE study of urban population density has absorbed the energies of professionals in a number of related disciplines for a great number of years. Ever since Clark (3) published his now universally familiar exponential decay formulation of the spatial arrangement of urban population densities, economists, ecologists, geographers, city planners, and others have been intrigued as to the actual universality of the hypothesis. Remaining unanswered are questions concerning the factors contributing to interregional variability in the expression's parameters and the temporal stability, or lack thereof, of the relationship for any particular urban region.

The onrush of urban transportation planning studies in the late 1950's brought with it the requirement to estimate the future activity distribution pattern in urban areas as a necessary prelude to estimating future transportation demands. This impetus of attention to the analysis and forecasting of urban activity structure logically carried along with it an increased concern and interest in the population density question.

Recent concern with population or residential density has primarily stemmed from an accounting requirement of the land-use or activity distribution process of the traditional transportation planning program. Having made the distribution of the resident population to individual geographic analysis zones, utilizing some sort of mathematical model or distributive device, the analyst usually relies on estimates of population density (a) to ascertain whether the distribution results imply unrealistic zonal residential densities, or (b) to compute the consumption of previously vacant land by the increment of residential growth in each zone. The latter process is of particular importance if the urban simulation process is performed in a finite number of incremental time periods requiring an updating of the relevant data files at the conclusion of each simulation period in preparation for the sequential increment. The density configuration has been integrated as an active and interrelated element of the distributive mechanism. In particular, reference is made (9) to the inclusion of individual density submodels as part of the total activity forecasting model system developed at the Delaware Valley Regional Planning Commission (DVRPC). Under this simulation system, the distribution pattern of activities is responsive to the existing density patterns in the region, just as the forecast density pattern is sensitive to the existing activity location pattern. There are numerous examples of activity distribution relationships which contain as

---

Paper sponsored by Committee on Land Use Evaluation and presented at the 46th Annual Meeting.

explanatory variables measures of activity density; however, unlike the DVRPC procedure, the values of such variables are normally exogenously determined rather than being forecast using separate relationships.

Activity density is useful in the trip generation analysis of a great many transportation studies. Trip generation, or the estimating of the total numbers of trips originating or terminating in each analysis zone, is frequently accomplished using a multiple-regression relationship equating trip production to a number of measurable characteristics of the analysis zone, one of them often being a measure of existing density. It can be inferred from trip generation analysis of this kind that the manner in which activities, for example households, arrange themselves will in and of itself influence the total volume of daily trip-making. The external economies associated with more dense activity arrangements undoubtedly are of some influence.

Transportation planners are currently interested in the question of how the community distributes its daily travel demands between available private and public transportation facilities, better known as the question of mode split. The activity density pattern has been observed to be of importance to this whole area of analysis. One could also refer to a substantial amount of professional speculation in the literature as to the future role, function, form, and viability of our great urban regions. These speculations, almost without exception, are heavily contingent upon the individual author's assessment of the levels of activity densities that the future populace is both willing and desirous of sustaining.

The question naturally arises as to current knowledge and technical competence in this area insofar as the existence of operational procedures or normative guides which can be utilized by those confronted with the overall task of estimating the future urban region is concerned. In fact the urban analyst and transportation planner will find little assistance with regard to the whole question of the future distribution of urban densities other than a number of less than completely satisfactory statements prognosticating the continued growth of the suburbs. This condition has prompted the present study, which attempts to investigate the household or residential density pattern existing in a small-sized urban region for two time periods, 12 years apart. The author has attempted to establish the conformity of the observed density structure to the previously mentioned universal formulation of urban population density, an exercise that may be of questionable payoff value but which arouses one's curiosity and is difficult to avoid. A considerable amount of multiple-regression analysis has also been undertaken. There should be some interest in the observed shifts in the parameters of these derived relationships over the 12-year interval. Finally, an attempt has been made to examine the relative order of accuracy associated with each of several forecasts of the residential density structure. The analysis and forecasting procedures examined have been purposely kept simple in an effort to maintain a balance with the kind of effort which a transportation planning study could realistically undertake in such a moderate-sized urban area.

RESIDENTIAL DENSITY: A DEFINITION OF TERMS

Residential density might at first appear to be a fairly clear and unambiguous quantity which should not require any extensive definition. However, a perusal of the literature reveals an extraordinary amount of confusion resulting from the avoidance in many cases of a rigorous definition of terms. Although this confusion cannot be settled here, it is necessary that the terms and concepts referred to in the remainder of this paper be defined at this point.

The notion of residential density refers to the ratio of some measure of the volume of residential activity per unit of land or space. From this very basic concept there arises a host of possibilities stemming primarily from the manner in which these two quantities are defined and measured (6). The numerator of the ratio presents far less difficulty than the denominator. The volume of residential activity can refer either to numbers of persons, households, or dwelling units. While there is room for ambiguity here, these quantities are familiar and have fairly strict definitions. For the purposes of this study, residential activity is always expressed in total dwelling unit terms. The

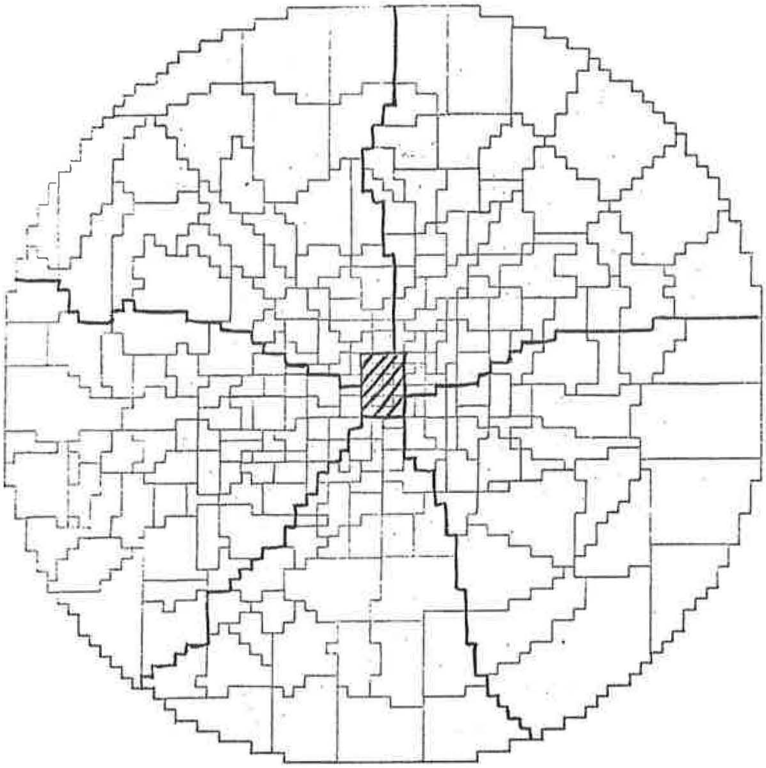

Figure 1. Study area showing zone and sector boundaries, Greensboro, North Carolina.

denominator of the ratio, on the other hand, has been the major source of definitional inconsistency. If we eliminate from consideration all third-dimensional possibilities, such as cubic feet of living space, we are still left with considerable room for variation using second dimension terms. A major breakdown here distinguishes between areal measures on the land itself, such as square feet, square miles, or acres, as opposed to areal quantities measured in man-made structures utilized by the resident population—for example, square feet of housing floor area. Discounting quantities of this latter nature and considering only areal units referenced to the earth's surface, one is confronted by a final major division which segregates residential density measures into what are commonly referred to as net and gross quantities. There is a fuzzy and shifting line which separates the two; however, the utility of each differs significantly. This point will be discussed later. In general, and for the purposes of this paper, the denominator of the gross residential ratio refers to the total area of the analysis unit, i. e., the area arrived at by planimetering its boundaries. Gross census tract dwelling unit density would then be calculated by dividing the total number of dwelling units in a census tract by the total area contained within the boundaries of the tract. Net residential density is a sharper measure than is gross density, and differs from the latter primarily as a result of a classification or stratification of the land uses contained within the boundaries of the analysis unit. For example, all land constituting a given census tract may be classified as either being used or vacant, and a net census tract dwelling unit density per square foot of used land may be computed. Used land may be further divided into residential and nonresidential usage terms and a dwelling unit density per square foot of residentially used land computed. In this paper, net residential density is defined as total number of dwelling units per unit of residentially used land, including street area.

TEST CONDITIONS AND METHODOLOGY

The Study Region

The Greensboro, North Carolina, metropolitan region served as the locale for this study. A rather complete and detailed data supply was available at a detailed geographic level for the years 1948 and 1960 (10, pp. 5-7; 2). The primary data source contained measures of total dwelling units, land area measures by use and unusable land, assessed land value, proximity to a variety of urban activities, and to the central business district (CBD) all coded to 3,980 thousand-foot-square grid cells which covered the circular study area (approximately 8 miles in radius) centering about Greensboro's CBD. Additional data consisting of total employment and an index of accessibility to total employment were also available on a travel analysis zone level. Figure 1 shows the entire study area structured into 249 analysis zones. Average family income was not present in these original data sources, but was available for both 1950 and 1960 from the U. S. Bureau of the Census (11, 12). Each zone was assumed to exhibit an average family income equal to the mean for the census tract into which it fell.

The study area was aggregated further into 5 sectors radiating out from the center of Greensboro and into circular rings each one mile wide. The sector boundaries of the study area (which were forced to analysis zone boundaries) are shown superimposed in Figure 1. The first ring (shaded area in Fig. 1), which was one-half mile in radius, circumscribed the central core of Greensboro City. The primary areal analysis unit utilized in this study was the district defined as that area contained within the intersection of successive sector and ring boundaries. The study area could have been structured into "driving time to the CBD" time increment rings as opposed to distance increments. However, earlier work (10) with this same data indicated little advantage to either. Therefore, distance units were selected primarily for correspondence with the bulk of earlier structural density analysis reported on in the literature.

Excluding the central core ring, the remaining 8-mile rings and 5 sectors totaled 40 districts. A primary reason for selecting the district as the basic analysis unit is that it most nearly approximated, in average resident population terms (the average district in Greensboro had 660 dwelling units in 1948), the typical traffic analysis zone used in traffic simulation analysis. The analysis zone had previously been shown to be too fine in tests of residential location models (10) and was therefore judged to be inappropriate for net density analysis.

In order that some justification for this rather coarse aggregation level could be provided, a one-way analysis of variance on the 1948 net residential density was performed. Table 1 gives the results of the analysis and the finding of statistically significant between-column variance, indicating that the district aggregation of zones did not mask out the prevailing zonal net density variability and was therefore not an inappropriate analysis unit. The reader may have noticed that the central core district and its composite analysis zones were not included in the analysis of variance. In fact, the central core district was removed from all analysis in this study. The density quantities computed for the central core district consistently deviated substantially and quite illogically from what would be expected from the findings for the remainder of the study area. The central core area is traditionally quite distinct in residential terms and for that reason is often treated as such in transportation simulation analysis. Beyond this intrinsic difference, the quality of the residential land-use and dwelling unit data is often less reliable than that for the remainder of the urban region and could account for further difficulty.

The sectors and rings have an historical analytic attraction and were investigated as additional levels of geographic analysis of residential density structure. A two-way analysis of variance was also performed on the same 1948 zonal net density values stratified by ring and sector. Table 2 gives the results of this analysis, revealing that significant between-ring variability existed but not significant inter-sector variability. Apparently distance out from the region's center is a more appropriate indicator of prevailing residential density than is the angular direction with respect to some reference axis. Based on these results, the sector was abandoned as a potentially fruitful analysis unit.

TABLE 1

ANALYSES OF VARIANCE RESULTS FOR INTERDISTRICT NET RESIDENTIAL DENSITY

Source of Variation	Sum of Squares	Degrees of Freedom	Estimated Variance
Between districts	6,860.1	39	175.9
Within districts	11,494.4	191	60.2

$F = \frac{175.9}{60.2} = 2.92$ (significant at 0.001 level)

TABLE 2

ANALYSES OF VARIANCE RESULTS FOR INTER-SECTOR AND INTER-RING NET DENSITY

Source of Variation	Sum of Squares	Degrees of Freedom	Estimated Variance
Between rings	703.1	7	100.4
Between sectors	14.8	4	3.7
Residual	98.8	29	3.4

F for column means $= \frac{100.4}{3.4} = 29.5$ (significant at 0.001 level)

F for row means $= \frac{3.7}{3.4} = 1.1$ (not significant)

In summary, all reported analysis of the distribution of gross and net residential density is for the district and ring units with the central Greensboro core area having been removed from consideration.

The Analysis Methodology

The strategy followed in this study centered about an investigation of the residential density structure (and the observed change in this structure over a 12-year period) of what was considered a fairly typical small-sized urban region. The study area sustained a 52 percent increase in numbers of dwelling units over the 12-year period, a rate of growth which is well above the average for the nation as a whole. Density-distance gradients were developed for both the 1948 and 1960 regions using the least squares criterion. Multiple-regression relationships for net density were calibrated for both 1948 and 1960. The analysis concluded with investigations of expected error in forecasting density. The following summarizes the primary objective of the study: (a) to investigate the appropriateness of several simple techniques which could be undertaken by a small planning study and staff with a minimal data supply; (b) to provide comparative quantitative measures of forecasting accuracy for each procedure or method investigated; and (c) to present some indication of any apparent advantage or disadvantage in selecting between gross or net residential density as the unit of measurement. The absence of any substantial amount of material in the literature on forecasting trends of residential density patterns coupled with the availability of data for only two time periods severely limited the selection of even simple forecasting techniques.

Because of personal bias, the bulk of the analysis concentrated on net residential density. Two district multiple-regression relationships were developed for the 1948 condition and were tested as valid forecasting devices. The initial regression formulation was modeled after the general form of the SPACEC I submodel of the previously mentioned Delaware Valley Regional Planning Commission's Activities Allocation Model system. The second regression relationship examined was of the usual multiple linear form. All calibration and forecasting errors are reported in coefficient of determination (R^2) terms[1] and therefore maintain comparability for cross-comparisons.

Gross Residential Density Analysis

Gross residential density has been defined as total numbers of dwelling units per unit of total land and thus avoids any consideration of the actual usage of the total land stock. This probably accounts for the historical orientation of previous density analysis to gross density measures. However, this simplicity is not achieved without a price, namely, a rather superficial measure of the individual household's consumption of land. Gross residential density (D_G) is defined as

---

[1] $R^2 = \frac{\text{original variance} - \text{explained variance}}{\text{original variance}}$; where estimates by particular techniques are transformations of density (e.g., logarithmic), they have been converted to density prior to the computation of residual error.

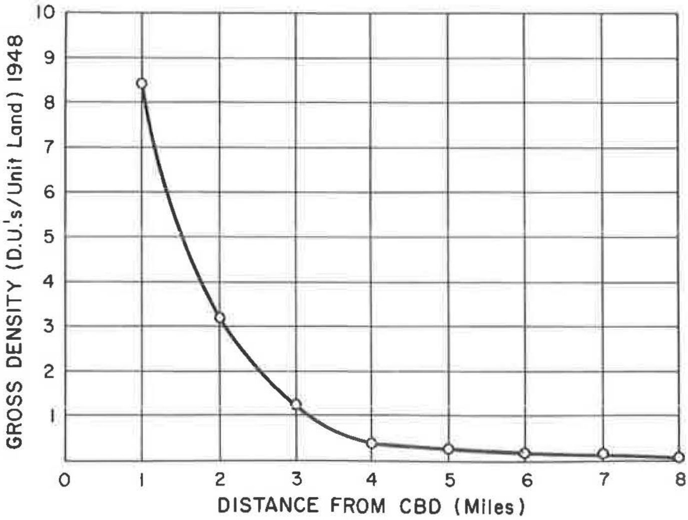

Figure 2. Gross density gradient (1948).

$$D_G = \frac{\text{D. U. 's}}{A} \qquad (1)$$

where

D. U. 's = total number of dwelling units in the analysis unit, and
A = total land area of the analysis unit.

That is, a general expression of gross density for any geographic unit i may be expressed as follows:

$$D_{G_{it}} = (c)\ \text{D. U. 's}_{it} \qquad (2)$$

where

$D_{G_{it}}$ = gross density in analysis unit i at time t,

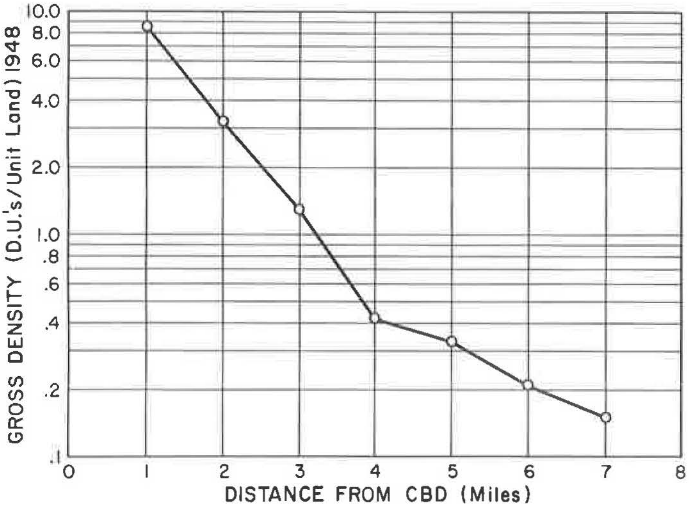

Figure 3. Semilogarithmic gross density plot (1948).

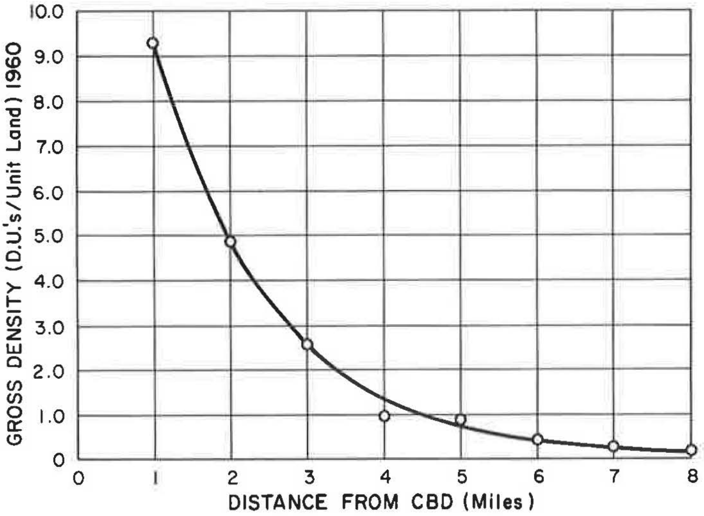

Figure 4. Gross density gradient (1960).

$D.U.'s_{it}$ = number of dwelling units in unit i at time t, and

c = constant equal to $\frac{1}{A}$.

Gross density is then proportional to the dwelling unit stock in the analysis unit, and as such provides little gain in a time series analysis over a simple accounting of the fluctuations in the dwelling unit stock. It provides little information concerning the actual living compactness of the population.

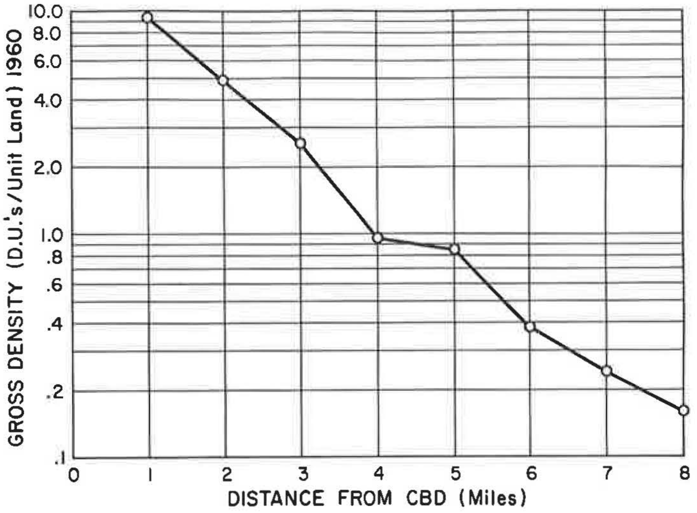

Figure 5. Semilogarithmic gross density plot (1960).

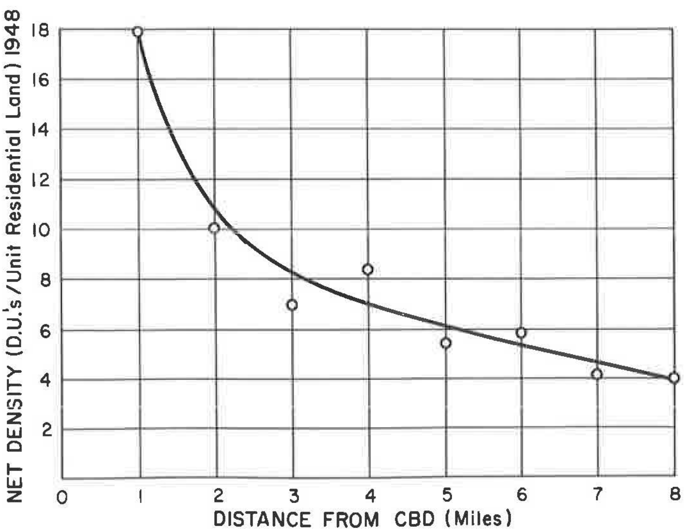

Figure 6. Net density gradient (1948).

The gross residential density of each of the 8-mile wide rings was computed for the study area for both 1948 and 1960.[2] The 1948 results were then plotted on regular graph paper as a function of the distance of the ring from the CBD of Greensboro in miles (Fig. 2). Clearly, a nonlinear relationship is in evidence. Figure 3 shows a replotting of the same data on semilogarithmic paper. A straight line fit in Figure 3 would give evidence of a negative exponential relationship. A definite straight line tendency is observed. A simple linear regression line fit to the points in Figure 3 resulted in the following:

$$\ln D_{G.48} = 2.43 - 0.648X \qquad (3)$$

where

X = miles from the CBD, and
$D_{G.48}$ = gross residential density (1948).

Transforming this regression equation to its antilog form yields:

$$D_{G.48} = 11.36\, e^{-0.648X} \qquad (4)$$

which is in the general negative exponential form. Note that the least squares fit obtained for the dependent variable in logarithmic form will not necessarily yield the best equation in terms of minimum residual variance when the relationship is solved for the dependent variable in antilogarithmic terms.

The R^2 for Eq. 4 was computed as 0.886. This same relationship, calibrated on the ring gross density values, was then examined as a fit of the gross density values at the district level. Solutions to Eq. 4 for the districts yielded an R^2 of 0.834. A reduced

---

[2] The coding of the land-use data from the 1,000-ft-sq grid file was in units of ninth's of development of the total area of the grid for the particular use category. For this reason all of the density analysis of this paper is in dwelling units per ninth of 1,000-ft-sq grid. This rather awkward dimension does not, of course, affect any of the structural analysis or measures of calibration and forecasting accuracy. Any of the density values reported in this paper can be converted to D.U.'s per acre by multiplying by the constant 0.392, or to D.U.'s per sq mi by multiplying by the constant 250.9.

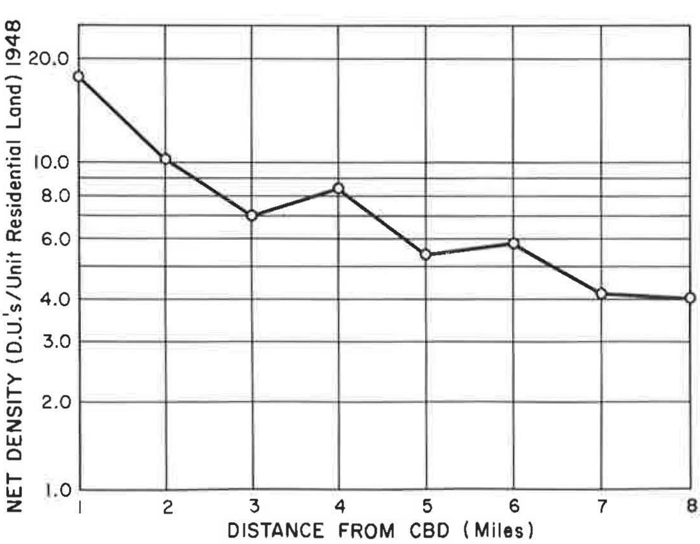

Figure 7. Semilogarithmic plot of 1948 net densities.

R^2 is to be expected if only because of the disaggregation and the resultant introduction of greater variability.

The ring gross densities for 1960 were then plotted on regular graph paper (Fig. 4). The nonlinear relationship suggested by the 1948 plot is again present in Figure 4. Replotting the 1960 points on semilogarithmic paper (Fig. 5) established the following least squares relationship:

$$\ln D_{G.60} = 2.75 - 0.585X \qquad (5)$$

which transforms to

$$D_{G.60} = 15.58\, e^{-0.585X} \qquad (6)$$

R^2's for Eq. 6 for both 1960 ring and district analysis levels are 0.989 and 0.923 respectively.

The marginal shifting of parameters observed between the 1948 and 1960 gradients suggested a test of the utility of the 1948 relationship as a predictor of 1960 densities. Solutions of the 1948 equation were then used as estimates for 1960 again at both a ring and district level. The resultant R^2's were then computed as 0.784 and 0.743 respectively.

Negative exponential relationships of the general form

$$D_G = a\,(X)^b$$

were also investigated as potentially useful gross density gradients. While the data did plot in a linear fashion on log-log paper, the calibration and forecast R^2's associated with these relationships were consistently below those previously reported.

Net Residential Density Analysis

Those most concerned with the residential density structure of urban areas are fundamentally pursuing indications or measures of the living compactness of households. A substantial amount of discussion exists in the literature, of fairly recent origin, which is directed toward the theoretical workings of household space consumption and residential location processes. Residential land consumption is treated as a resolution of an economic equilibrium between demand and supply. Viewed as one of many economic transactions engaged in by the urban household, the selection of a residential site is

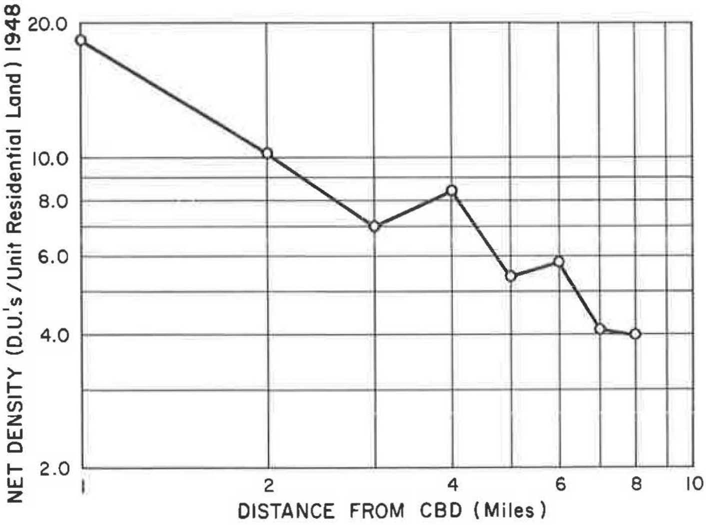

Figure 8. Logarithmic plot of 1948 net density.

determined by the economic condition of the household, its preference pattern in terms of trade-offs with other commodities, the state of the housing market and its relationship to the transportation system (7, 13, and 14). The household is provided with an income which it must allocate in the purchase of goods and services in such a way as to achieve as much total satisfaction as possible. For simplicity, let us assume that all household expenditures fall into three general classifications: transportation, housing and other. If we assume further that the "other" purchases absorb a fixed proportion of total income, the urban household faces the problem of purchasing housing and transportation such that composite satisfaction is maximized and total purchases do not exceed a fixed amount. As a first solution the household head might elect to buy housing where land costs are cheapest, thereby getting the most space for his money; however, it is likely that this location is remote from the remaining activities of the urban area with which he must interact, thereby leaving him with an extravagent transportation bill. On the other hand, he might elect to locate where transportation service is best but where housing cost is so high that to stay within his fixed expenditure allowance he is constrained to the purchase of an undesirably small housing package. Contained within this total theoretical framework is a causal relationship between the land value distribution in the region and transportation service. Areas which are highly accessible are most desirable and can therefore command a higher price. It is primarily this final consideration which directly links the urban transportation planner's decisions with the course of urban development (15, pp. 256-257).

This somewhat tangential discussion has been made to show the appropriateness of net as opposed to gross density data and analysis. Net density analysis can contribute to as well as draw upon this theoretical framework. Gross density techniques, with their vague tie to land consumption, cannot so contribute. Conceivably, analytic tools will be forthcoming, incorporating these theoretical relationships, which will provide the transportation planner with direct assignments of the form and composition of marginal development to the areas of influence about proposed transportation routes or improvements thereto.

Net residential density was computed for each district and ring in the study region by totaling the dwelling units and dividing by the total area of land existing in residential use. A much more desirable procedure for computing average net density would have been to average the density of each individual dwelling; however, this requires consumed land on an individual dwelling basis. The computed average net density must be

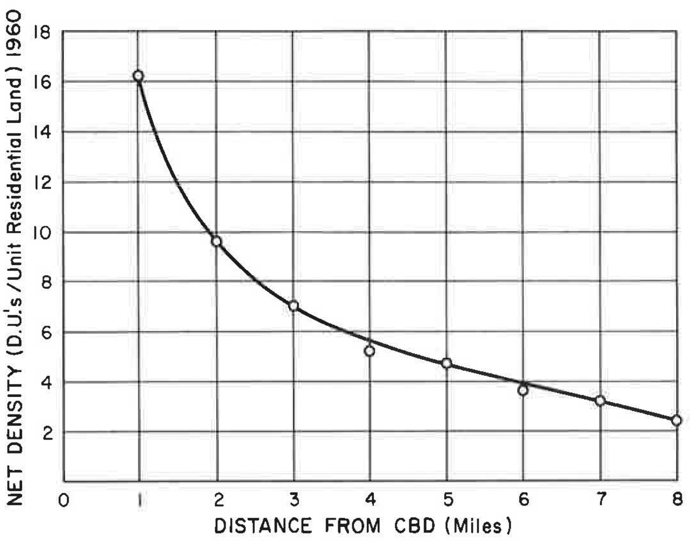

Figure 9. Net density gradient (1960).

treated as representative of the average condition in the analysis unit. Its representativeness is dependent on the variability of the individual dwelling densities within the unit. It should also be noted that unlike the gross density measure, net density is not monotonically related to the total dwelling stock; it can rise or fall both with increases or decreases in the contained dwelling unit total.

Figure 6 shows a plot of the computed ring net densities for 1948 as a function of distance to the CBD. The general conformity in shape with the equivalent gross density plot is evident. Figure 7 reveals the general linear relationship obtained by a replotting on semilogarithmic paper. However, as suggested partially by the evidence of nonlinearity in the plot in Figure 7 and from Kramer's work (4), the net residential density

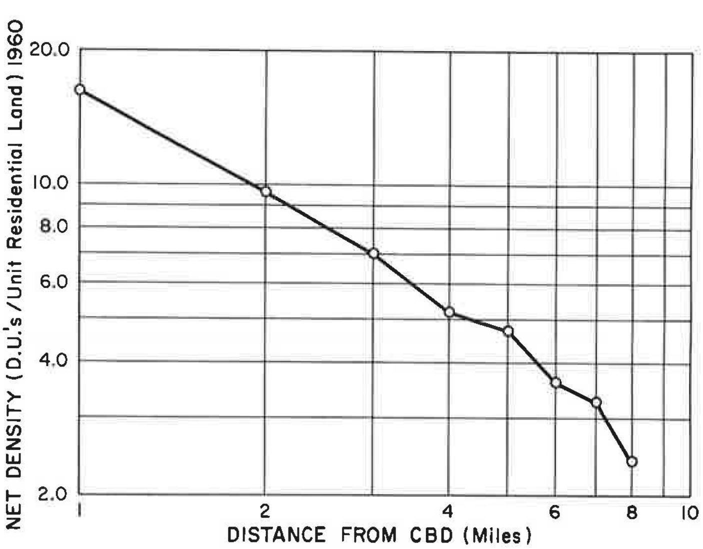

Figure 10. Logarithmic plot of 1960 net density.

data were plotted on log-log paper (Fig. 8). Least squares fits were computed for both scatter diagrams, Figures 7 and 8, and R^2 values computed. The doubly logarithmic relationship proved to be a better linear fit.

The linear equation fit to the 1948 net ring densities was

$$\ln D_{N.48} = 2.850 - 0.688 \ln X \tag{7}$$

which in nonlogarithmic form is

$$D_{N.48} = 17.29 (X)^{-0.688} \tag{8}$$

The R^2 associated with Eq. 8 was 0.957. Eq. 8 was then examined as an estimator of the 1948 district net densities and yielded an R^2 of 0.835.

Under the assumption of stability in the net density structure of the test region over the 12-year period, Eq. 8 was tested as a valid predictor of the 1960 net densities at both the ring and district level. The resulting R^2's were respectively computed to be 0.927 and 0.844.

To complete this particular line of investigation, the 1960 net ring densities were plotted first on regular graph paper and then on log-log paper (Figs. 9 and 10). The least squares regression fit to Figure 10 resulted in the following relationship:

$$\ln D_{N.60} = 2.855 - 0.876X \tag{9}$$

or

$$D_{N.60} = 17.4 (X)^{-0.876} \tag{10}$$

with R^2's of 0.986 and 0.934 at the ring and district aggregation levels.

The relative success of these investigations suggested the testing of the following less involved procedure; ring densities in 1960 were estimated to remain exactly as they were computed to be in 1948. This simplifying assumption implies that the added dwelling units over the test period consumed, on the average, the same amount of land as the average dwelling unit existing in the ring in 1948. The computed R^2 for the 1960 net ring densities was 0.849. Carrying this procedure down to the districts, incremental dwelling growth in each district was assumed to locate at the same average 1948 net density as for the particular ring to which it fell. Implicit in this trial is that the intraring net density variability is diminishing over time with each district's net density approaching its ring average. The percent of 1960 net district variance explained utilizing this technique was 0.533.

Finally, each district was assumed to maintain constant average net density from 1948 to 1960, the 1948 values then serving as 1960 estimates. An R^2 of 0.640 was computed for this case.

The concluding analysis of the net residential density pattern involved the development of multiple-regression equations utilizing as independent explanatory variables selected data items from the rather extensive list available. However, because the majority of these data were already available at the analysis zone level, the decision was made to calibrate the net density multiple regressions at this level, and to utilize them as estimators for both districts and rings. While this procedure violates strict regression procedure, the errors introduced were thought not to be severe, partially relying on the results of the interdistrict analysis of variance reported on earlier which revealed that the intradistrict variability was minor in relation to the interdistrict variance. Additionally, this approximating procedure required that the dependent variables for each of the regressions be an intensive quantity, and thereby independent of the size of the analysis observation unit.

The functional form of the DVRPC's density submodel, SPACEC I, was investigated as representative of the study region's density pattern. In a much simplified form the relationship is

$$D_N = a e^{\Sigma b_i X_i} \tag{11}$$

where

X_i = independent variable i,
b_i = the coefficient of variable i,
a = constant, and
D_N = net residential density.

This relationship transforms by logarithmic conversion to a standard linear multiple-regression relationship:

$$\ln D_N = \ln a + \Sigma b_i X_i \tag{12}$$

Using a stepwise regression program, least squares relationships of the form of Eq. 12 were developed for both the 1948 and 1960 net densities. A considerable number of trials were attempted before two final relations were accepted which were logically sound and which contained only statistically significant explanatory variables. The 1948 equation computed was

$$\ln D_{N.48} = 1.534 + \underset{(1.92)}{0.005X_1} + \underset{(2.12)}{0.017X_2} + \underset{(6.62)}{0.109X_3} \tag{13}$$

or

$$D_{N.48} = (4.64)\, e^{(0.005X_1 + 0.017X_2 + 0.109X_3)}$$

where

X_1 = land value 1948 (\$/sq ft),
X_2 = percent developed land in industrial use (1948), and
X_3 = gross residential density (1948) = $D_{G.48}$.

The numbers in parenthesis below each coefficient are the regression "t" values. Eq. 13 was then used to estimate the 1948 ring and district net densities by a simple substitution of the appropriate values for the independent variables. The R^2's computed for the rings and districts were 0.802 and 0.714.

The stability of the relationship developed for 1948 was investigated by using Eq. 13 as a predictor for 1960, substituting 1960 values for the explanatory variables. Values of variable X_1, land value, did not change inasmuch as these data were only available for 1948. The R^2's resultant from this predictive effort were 0.902 and 0.802, respectively, for the rings and districts. Solutions to Eq. 13 were transformed to non-logarithmic form prior to the calculation of residual errors.

A least squares regression of the general form of Eq. 12 was then made for the 1960 net ring density distribution. The measures of accuracy for this relationship could then be used to evaluate how well the 1948 relationship held up. In addition, the changes in the variable makeup of this new relationship might provide some interesting comparisons with the 1948 equation. The 1960 least squares relationship was

$$\ln D_{N.60} = 0.086 + \underset{(6.83)}{0.123X_1} + \underset{(8.21)}{0.542X_2} \tag{14}$$

or

$$D_{N.60} = (1.9)\, e^{0.123X_1 + 0.542X_2}$$

where

$D_{N.60}$ = net residential density (1960),
X_1 = gross residential density (1960), and

X_2 = logarithm of net residential density (1948) = ln $D_{N.48}$.

The ring and district R^2's were 0.805 and 0.832.

Multiple-regression estimates were then developed for a nontransformed dependent variable. Again a stepwise procedure was used in testing a number of independent variable combinations before two relationships, for 1948 and 1960, were accepted. The two equations were

$$D_{N.48} = 6.25 + \underset{(4.04)}{0.085X_1} + \underset{(2.96)}{0.206X_2} + \underset{(6.49)}{0.922X_3} \quad (15)$$

where

X_1 = land value - (\$/sq ft) 1948,
X_2 = percent of developed land in industrial use (1948), and
X_3 = gross residential density (1948) = $D_{G.48}$;

and

$$D_{N.60} = 1.960 + \underset{(2.00)}{0.012X_1} + \underset{(3.12)}{0.053X_2} + \underset{(28.47)}{1.082X_3} + \underset{(7.00)}{0.119X_4} \quad (16)$$

where

X_1 = land value (\$/sq ft) 1948,
X_2 = percent developed land in industrial use (1960),
X_3 = gross residential density (1960) = $D_{G.60}$, and
X_4 = net residential density (1948) = $D_{N.48}$.

The R^2 values for Eq. 15 were 0.755 and 0.561 at the ring and district. Eq. 15 was then tested as a predictor of 1960 ring district and densities. R^2 values of 0.264 and 0.151 were determined for solutions of Eq. 15, substituting where possible 1960 values for the independent variables. Eq. 16 was then solved and ring and district R^2 values of 0.963 and 0.938 computed.

The independent variables and the signs of the coefficient in Eq. 15 appear logical and causatively related to the quantity being estimated. The positive coefficient of assessed land cost reflects the economic supply and demand process at work. The higher the cost of land, the less the individual family can afford to consume, and the resultant increase in net density (7). The positive coefficient of percent industrial land is probably a reflection of the tendency for low income families to settle in the marginal residential areas which are often characterized by a heavy mixture of industrial development. This result is somewhat at odds with that found by Muth (7) who states that his finding of a negative relationship between net population density and proximity to local manufacturing centers is probably due to a net decline in housing price resulting from a generally undesirable neighborhood effect overcoming a coincident positive pricing effect associated with the increased accessibility of such areas.[3] Muth's findings suggest at least two alternative explanations of the positive relationship in Eqs. 15 and 16:

1. The transportation network and manufacturing sites are so located in the study area as to afford those areas of substantial manufacturing activity a decided accessibility advantage. This advantage would then be reflected in inflated housing pricing, enough to overcome any deteriorating effect of an unfavorable environment; and
2. The areas of manufacturing concentration have substantially remained in the older sections of the city where surrounding housing is traditionally of a higher density than in newer developing residential areas.

---

[3] Preliminary calibration results of the SPACEC I model by DVRPC have shown a similar positive relation between residential density and industrial activity. See internal staff memorandum of February 2, 1966, titled SPACEC I Parameters.

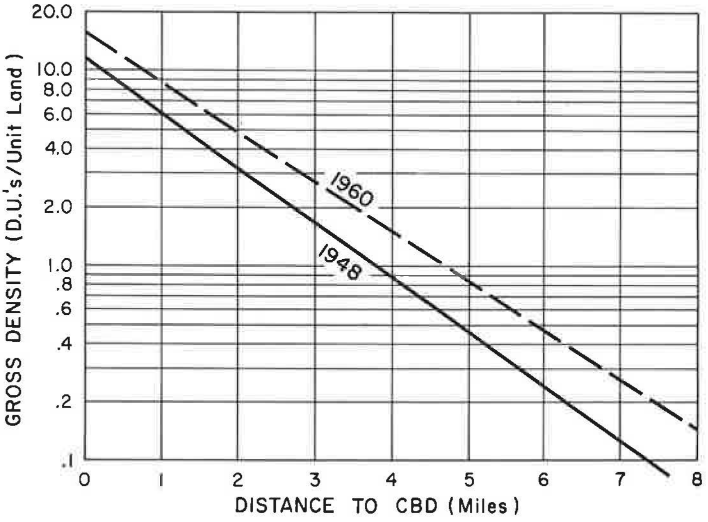

Figure 11. Calibrated gross density gradients (1948 and 1960).

These propositions are partially supported by the entries in the analysis zone simple correlation matrix. The "accessibility to employment" and "percent of developed land in industrial use" variables exhibit a moderate positive correlation of 0. 37, substantiating to some degree the first possible explanation. The employment accessibility index is a regional measure such that the built-in collinearity in the two variables is probably not overly biasing. However, recent findings (5) tend to suggest that the urban worker is giving only secondary consideration to access to the workplace in selecting his place of residence. As the mobility of labor increases, along with the eventual shortening of the work week, it is unlikely that this trend will reverse. It is therefore doubtful that the pricing effect of superior accessibility would be so substantial as to overcome the general nuisance effect of proximate industrial activity. This

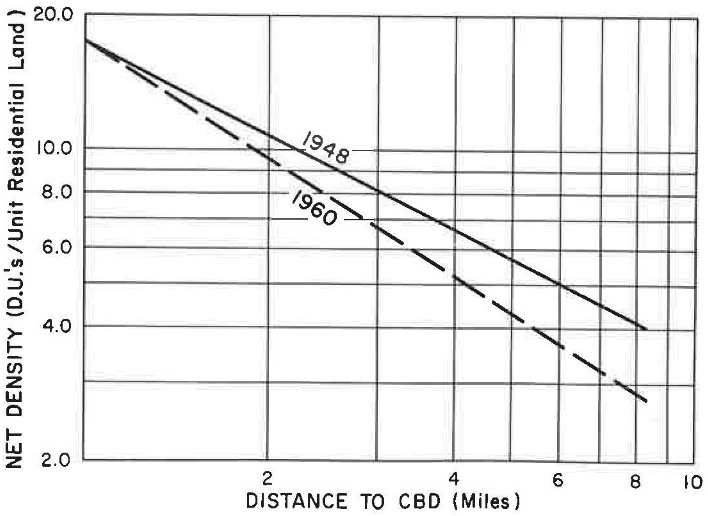

Figure 12. Calibrated net density gradients (1948 and 1960).

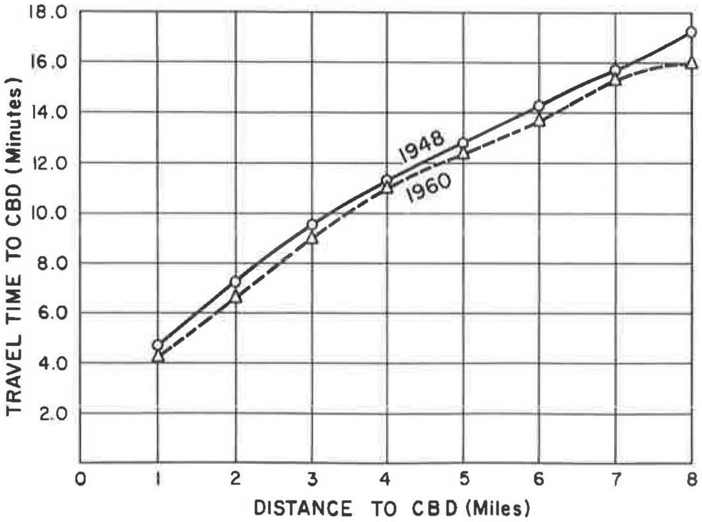

Figure 13. Travel times to the CBD (1948 and 1960).

is especially true for a smaller sized urban region where the influence of travel impedence as a locational factor has been observed to be below that found in larger urban areas.

There was no simple data item available which specifically measured the age of development in an area; however, the variable "percent of total land area not in use" was thought to be highly correlated with such an index. A negative correlation of 0. 38 was found for this variable and "percent of developed land in industrial use. " This relationship tends to support the second extended explanation for the positive correlations found between net density and degree of industrialization. However, these arguments must remain inconclusive, requiring a much more detailed analysis with additional data.

The strong positive correlation between gross and net density is revealed by the coefficient of the last independent variable in Eq. 15. As greater numbers of families locate in a given area, it is expected that increased demand for the relatively fixed stock of available land will result in increased land costs and increased densities.

The independent variable set in Eq. 16 is subject to the exact same interpretation as Eq. 15 with the addition of a fourth variable, net 1948 residential density. A substantial proportion of the dwelling stock in 1960 existed in 1948 in spite of the significant growth over this period. A strong serial correlation in the two net density variables is clearly evident, the correlation being 0. 53. However, the strong explanative relation between the two measures may also reflect a general inertia characteristic of new residential development. That is, the existing density pattern in an area influences the density characteristics of incremental dwelling development so that there is a tendency to avoid any great contrasts over relatively small geographic areas. This correlation, if it in fact exists, would probably be dependent on the degree of development existing in the area and the rate of growth. Further investigation of this point would require density data for the dwelling units locating over the study period, data which were not available for this study.

Observations on the Change in Density Structure

It may be both interesting and informative to briefly examine the actual shifting in the residential density structure of the study area over the 12-year period of analysis.

Figure 11 contains the two density distance gradients (Eqs. 3 and 5) fit to the ring gross density observations for both 1948 and 1960. Two obvious changes have occurred in the gross density configuration manifest in the deviations of the two linear gradients.

Gross density has consistently increased in each ring. This is to be expected in light of the definition of gross density and the occurrence of a 52 percent growth in total dwellings in the study region over the test period. A decline in gross density could only have occurred under the condition of an absolute loss in total dwellings, a remote possibility in a region experiencing such a rapid expansion. Also the gradient has flattened out slightly, indicative of a less compact population distribution in 1960 than in 1948, and characteristic of a suburbanizing region. Berry et al (1) have found this phenomenon to be generally true for western cities and observed that density gradients tend to decline over time for a given city and tend to be flatter for larger sized cities.

The net residential gradients for both 1948 and 1960 are shown in Figure 12. In contrast to the gross density gradients, note the consistent decline in net density from 1948 to 1960 regardless of distance from the CBD. This observation is not simply explained in terms of absolute population growth and requires an extensive economic analysis of existing market conditions and consumer preferences. Clearly though, a major factor contributing to this overall density decline could be the substantial improvement in the transportation service in the region. Figure 13 shows average over-the-road travel time to the CBD for each ring in both 1948 and 1960. Highway service to the CBD has apparently shown consistent improvement over the 12-year span. The reduced travel costs associated with such improvements can provoke profound shifts in the locational equilibrium position of households. Reduced transportation costs can provide for decreasing land rent and also produce income effects which probably will increase the household's housing expenditure. Combining these two effects likely results in the consumption of more living space per household, perhaps explaining in part the results observed in Figure 12. However, this trend is not necessarily irreversible in spite of continued transportation improvements. As net densities continue to decline, the marginal worth of increased space necessarily falls with the distinct possibility of it reaching a point where it no longer is to the household's benefit to consume more (7, pp. 28-29). Housing space may in fact become an inferior good at some point (different for each household or household group). An interesting recent finding may provide some empirical justification for this prognostication. Lansing (5) has found that a majority of households unsatisfied with their present lot sizes prefer larger lots up to 3/10 of an acre. On the other hand, the majority of households living on lots larger in size than one-half in acre, and expressing dissatisfaction, would prefer smaller sized lots.

DISCUSSION OF RESULTS

Test Findings

Table 3 contains the results of all the analysis of gross residential density. The R^2 entries in the table reveal that the negative exponential formulation, equating gross density with distance outward from the center of the urban area, provides an effective description of the existing pattern. As expected, the results at the higher level of geographic aggregation show less residual error; however, considering that the reported district errors result from the application of the relationship calibrated to the ring values, it appears that the intra-ring variance is relatively minor in comparison with the inter-ring gross density variability. Using the exponential relationship fitted to the 1948 distribution as a forecasting device for 1960 proved to be moderately successful. The drop in explained variance from 1948 to 1960 was on the order of 11 percent at both the ring and district level. As was pointed out previously, the distance gradient fit to the 1960 gross density distribution was flatter than that for 1948 and, as shown by the R^2 values, accounted for approximately 11 percent more variance at both the ring and district aggregation levels. It should be noted that the 1948 gross density distance exponent of 0.648 is probably low for the size of the study area in comparison with the findings of Muth (7, p. 221).

The comparative results for the net density analysis are contained in Table 4. The procedures utilized in the net densities analysis fall under four general headings, corresponding to the four major sections in Table 4: distance gradients, multiple regressions (with both transformed and untransformed dependent variables), and assumed stability of 1948 values to 1960.

TABLE 3

CALIBRATION AND FORECAST R^2'S FOR GROSS DENSITY DISTANCE GRADIENTS

Equation	1948		1960	
	Ring	District	Ring	District
$D_G = 11.36\ e^{-0.648X}$	0.886	0.834	0.784*	0.743*
$D_G = 15.58\ e^{-0.585X}$	—	—	0.989	0.923

*Forecast results.

The distance gradients fit to the 1948 net and gross density distributions show significant R^2's at both the ring and district level. However, the reduction in accuracy at the district level is substantially greater in the net density case than was found in the gross density analysis. On the other hand, the stability of the 1948 net density gradient is considerably above that found for gross density as indicated by the comparative R^2 values found in utilizing the 1948 gradient as a projection tool to 1960. There was only a minor falling off in explained ring net density variance in 1960 and surprisingly a 10 percent increase in explained interdistrict variance. The latter result points to some of the peculiarities associated with the use of linear estimating procedures in fitting essentially nonlinear relationships by the expedient of logarithmic transformation. The R^2's calculated for the 1960 net density gradient are quite high and match almost exactly the equivalent gross density results. Once again the increase in the slope of the net density distance gradient from 1948 to 1960 is noted. Overall, the density-distance gradient formulations provided comparable accuracy at both aggregation levels for both net and gross values, with the single outstanding result being the superiority of the 1948 net gradient as a predictor of 1960 conditions.

The multiple-regression R^2's developed for the net density distributions are also presented in Table 4. For nonlinear formulation, the results indicate moderate explanatory success in calibration for 1948. The residual variance is greater than for the distance gradient trial at both the ring and district levels, there being a considerable decrease in explained variance for the ring analysis. Quite unexpectedly the equation produced higher R^2's in a projection role than it did in calibration. In fact, solutions to the equation with 1960 values of the independent variables (with the exception of land value) resulted in less residual error after transformation to nonlogarithmic form than the equation calibrated to the 1960 data.

TABLE 4

CALIBRATION AND FORECAST R^2'S FOR NET DENSITY ANALYSIS

Analysis Procedure	Model Form	1948		1960	
		Ring	District	Ring	District
Distance gradient	$D_n = (17.3)\ X^{-0.688}$	0.957	0.761	0.927*	0.844*
	$D_n = (17.4)\ X^{-0.876}$	—	—	0.986	0.934
Log linear multiple regression	$\ln D_{n.48} = 1.534 + 0.005$ land value $+ 0.017$ % industrial land $+ 0.109$ gross density_{48}	0.802	0.714	0.902*	0.820*
	$\ln D_{n.60} = 0.086 + 0.123$ gross density_{48} $+ 0.542 \ln D_{n.48}$	—	—	0.805	0.832
Linear multiple regression	$D_{n.48} = 6.257 + 0.085$ land value $+ 0.206$ % industrial land $+ 0.922$ gross density_{48}	0.755 —	0.561 —	0.264*	0.151*
	$D_{n.60} = 1.960 + 0.012$ land value $+ 0.053$ % industrial land $+ 1.082$ gross density_{60} $+ 0.119\ D_{n.48}$	—	—	0.963	0.938
Assumed no change in net densities	Assume 1960 ring densities same as 1948 ring densities	—	—	0.849*	0.533*
	Assume 1960 district densities same as 1948 district densities	—	—	—	0.640*

*Forecast results.

The results in Table 4 for the linear regression estimating equation without exception show that the accuracy of the 1948 calibrated relationship is inferior, at both aggregation levels, to those found for either the nonlinear regression or the distance gradient formulations. This is true both with respect to calibration and projection. These results are particularly interesting considering that the exact same explanatory variables compose both the linear and nonlinear 1948 regressions. However, the calibration R^2's for the 1960 linear regressions are quite high, comparable to those obtained for the distance gradient and considerably better than those resulting from the nonlinear regression relationship. The strong serial correlation between 1948 and 1960 net densities most probably accounts for the sharp increase in explanatory accuracy of the 1960 equation as compared to that for 1948, inasmuch as the two equations have precisely the same independent variable composition with the exception of the inclusion of 1948 net density in the 1960 equation.

The final two entries in Table 4 testify somewhat to the point of temporal stability in aggregate net density patterns. These last two sets of R^2's coincide with a forecast of 1960 ring and district densities under the assumption of no change in average ring densities over the 12-year period. It is apparent that considerable net density variability can be explained as a carry-over from the base time period. However, the accuracy of this forecasting procedure falls off considerably in going to lower levels of geographic aggregation. Even when account is taken of the intra-ring variability and 1960 district densities are assumed to remain as they were in 1948, only 64 percent of the variance is accounted for. While the simple forecasting technique of projecting no change in the density distribution is effective, it does not do as well as the assumption of stability in a density gradient relationship. However, by introducing the possibility of simulating the effects of temporal changes in the character or nature of the urban region, as is the case in the development of properly structured regression equations, considerable improvement can be expected insofar as accounting for density variability. This appears to be even more true as the level of aggregation falls. It is well to reiterate at this point that the regression relationships developed in this study were calibrated at a lower level of aggregation than at which the indices of accuracy were calculated and reported on in Table 4. It is quite likely that had the regressions been developed on district data, higher calibration and projection R^2's would have been obtained.

The comparative results shown in Tables 3 and 4 indicate that the analysis and projection of net residential patterns can be made at least as accurately as for gross density, although personel bias of the author resulted in only limited investigation of gross density. The results contained in Table 4 substantiate that considerable success can be expected in the projection of urban net density configurations through the development of simple distance gradients. Conclusions concerning the development of multiple-regression relationships are difficult to construct from the somewhat inconsistent results obtained. While the nonlinear regression formulation proved quite superior to the linear equation in calibration to 1948 conditions and in projection to 1960, the reverse was found to be true with respect to the 1960 calibrated relationships. In any case it is apparent the significant regression relationships can be developed which contain explanatory variables with rational causative justification. It is unfortunate that the family income measure did not enter as a significant variable in any of the regression relationships in light of the apparent theoretic importance of this factor in the explanation of urban settlement. Unfortunately, the income data available for this analysis were census tract medians, too aggregated for the analysis zone level at which the remaining data were available and at which the actual regression calibrations were conducted. Perhaps zonal household income would have contributed to the explanatory relationships.

SUMMARY

There are some major points to be made concerning the analysis of small-sized urban residential structure resultant from this one limited test. To the analyst or planner concerned with developing a single best estimate of the future population density distribution, it is apparent that it is worth the limited amount of added effort to develop a best-fitting distance gradient as opposed to simply projecting base year conditions blindly into the future on a small-area basis. The utility of distance gradients as effective representations of the density surface quite likely diminishes as the geographic analysis unit becomes finer. At gross levels of analysis, residential density patterns are apparently well correlated with distance outward from the region's center; however, there exists an underlying pattern of small-area heterogeneity superimposed upon this gross pattern of exponential decay.[4] Accurate simulation of this lower level variability will likely depend on the development of causal relationships incorporating many of the notions currently contained in location theory. Additionally, distance gradients are quite useless in reproducing the likely fluctuations in residential development compactness resulting from alternations in one or a number of key policies or planning standards or from shifts in the socioeconomic character of the population. Only through the development of sound and logical models which simulate these interrelationships can such planning flexibility be established.

Finally, through the exchange medium of land value the urban transportation planner is able to contribute actively in the total effort aimed at bringing order and efficiency to the urban space. By providing and depriving transportation access spatially, heterogeneity is induced in the land value surface. The transportation planner thus participates directly in the alteration of the residential density configuration, a significant parameter of the urban mechanism.

CONCLUSIONS

It should be apparent that the results in this paper cannot be viewed as being conclusive coming from a single analysis of a particular urban area over a single time period. However, these results will hopefully contribute to and advance existing analytic and forecasting facility with respect to the residential density structure of small- and medium-sized urban areas. The following are then a brief listing of several of the most important conclusions to be drawn from this analysis:

1. Future analytic work in the general area of urban residential structure should concentrate on net as opposed to gross density measures. Net density is a much richer and more exact unit directly compatible with a substantial body of existing theory and apparently is as conducive, if not more so, to meaningful analysis and projection as gross density.
2. The analysis district, the basic unit of analysis for this study, having an average internal population of 2, 000 persons and thereby being comparable to the familiar urban transportation travel analysis zone, is a useful level of aggregation for studying urban residential density structure and does not subsume the most significant variability in net density within the urban region.
3. Density-distance gradients are useful tools in analyzing the density structure of the urban area and can also serve as appropriate projection devices. However, it is clear that density gradients are not static, suggesting that additional research be devoted towards developing rational explanations of, and procedures for estimating, these parametric shifts. Such knowledge would greatly improve the forecasting potential of distance-gradient relationships.
4. The development of accurate models of net residential density, which are logically structured in terms of existing theories of economic equilibrium and activity location, should be actively pursued. Results obtained in the present study are encouraging in this regard in spite of obvious informational deficiencies.
5. It is probable that the central core areas should be treated separately from the remainder of the urban region in the development of simple models of residential density. Also, considerable distortion can be introduced by the inclusion of substantially

---

[4] Witness the recent development of high-density high-rise apartment developments in what have traditionally been areas solely developed to typical suburban single-family dwelling densities.

rural areas, which are not expected to sustain significant urbanization, in the development of residential density relationships.

6. Future residential density configurations should not serve merely as exogeneously determined constraints to simulation models of residential location. If density patterns can be functionally related to socioeconomic characteristics, then it would appear that the future density structure should be responsive to, as well as influence, forecasts of the location of the urban area's activities.

REFERENCES

1. Berry, B. J. L., Simmons, J. W., and Tennant, D. J. Urban Population Densities: Structure and Change. Geographical Rev., July 1963.
2. Chapin, F. S., Jr., and Weiss, S. F. Factors Influencing Land Development. Univ. of North Carolina, April 1962.
3. Clark, C. Urban Population Densities. Jour. of the Royal Statistical Soc., Vol. 114, 1951.
4. Kramer, C. Population Density Patterns. CATS Res. News, Vol. 2, No. 1, Jan. 17, 1958.
5. Lansing, J. B. Residential Location and Urban Mobility: The Second Wave of Interviews. The Univ. of Michigan, Jan. 1966.
6. Ludlow, W. H. Urban Densities and Their Costs: An Exploration Into the Economics of Population Densities and Urban Patterns. In Urban Redevelopment: Problems and Practices (Coleman Woodbury, ed.), Univ. of Chicago Press, 1953.
7. Muth, R. F. The Spatial Structure of the Housing Market. Papers and Proc., Regional Science Assoc., Vol. 7, 1961.
8. Muth, R. F. The Variation of Population Density and Its Components in South Chicago. Papers and Proc., Regional Science Assoc., Vol. 15, 1965.
9. Seidman, D. R. Report on the Activities Allocation Model. P. J. Papers No. 22, Nov. 17, 1964.
10. Swerdloff, C. N., and Stowers, J. R. A Test of Some First Generation Land Use Models. Highway Research Record 126, pp. 38-59, 1966.
11. U. S. Bureau of the Census. Census of Population, 1950, Bull. P-D22.
12. U. S. Bureau of the Census. Census of Population and Housing, Final Report, PHC(1)-57, 1960.
13. Wingo, L., Jr. An Economic Model of the Utilization of Urban Land for Residential Purposes. Papers and Proc., Regional Science Assoc., Vol. 7, 1961.
14. Wingo, L., Jr. Transportation and Urban Land. Resources for the Future, Inc., Washington, 1961.
15. Wingo, L., Jr., and Perloff, H. S. The Washington Transportation Plan: Technics or Politics? Papers and Proc., Regional Science Assoc., Vol. 7, 1961.