CHANGES IN REGIONAL POPULATION PATTERNS
1990-2000: An Analysis of US Census Data

M.E. O'Kelly
and
M. W. Horner

Department of Geography
The Ohio State University
1036 Derby Hall, 154 North Oval Mall
Columbus, OH 43210-1361, USA
Email: okelly.1@osu.edu

June 30, 2001
1. Introduction

There has been much speculation throughout the 1990's as to where the most significant population growth and decline took place in the United States. With respect to metropolitan areas, many believe that suburban locations grew faster and at the expense of central city cores. These trends are discussed in the literature in terms of increased suburban subcenter development, so-called ‘edge city’ development, and development beyond the rural-urban fringe known as exurban development (Lucy and Phillips 1997). At larger scales, different regions experience vastly different population outcomes. The well-known Rust Belt to Sun Belt migration stream, which has sent mobile Americans southward in search of new residences, is but one example of a significant regional population shift contributing to change (Plane and Rogerson 1994).

With the release of the 2000 decennial census data, some of these important questions regarding regional population change may begin to be investigated. To this end, we analyze 1990 and 2000 US county population data in an effort to identify regional trends in population change over the past decade. Two measures of population accessibility are formulated and computed to illustrate how regional population change occurred during the period 1990-2000. To place these results in context, we also analyze data from past censuses 1940-1980 and report on historical county accessibility trends. These measures of population accessibility are then mapped and displayed for the continental United States.

Using population accessibility indices to investigate regional population change is advantageous since population accessibility indices are spatial measures of population, whereas simple change statistics do not take regional trends into account. Moreover,
accessibility indices provide an objective quantification of the opportunities relative to a
given location since the indices illustrate an area's potential opportunity within larger
urban and regional contexts. From a practical standpoint, these indices are easily
computed and visualized using commercial geographic information systems (GIS)
available today.

The paper is organized as follows. Section two develops a general conceptual
background of the study through a review of the literature, and then it presents the
mathematical formulations for our population accessibility analysis. Following the
formulations, the visual patterns of accessibility are presented and discussed. Lastly,
conclusions and suggestions for future research are provided.

2.1 Literature Review
Numerous researchers have analyzed aggregate spatial data in an attempt to discern
patterns in population phenomena. Indeed, researchers often use state or county level
population data to study changes in the distribution, composition or movement of
population. As such, this section reviews selected studies of aggregate population issues.

During the 1970’s, there was a debate as to whether increased nonmetropolitan
population growth was occurring as response to urban sprawl or because of other factors.
In light of this debate, seminal work by Vining and Strauss (1977) reviewed trends in US
population concentration. Using the Hoover concentration index, Vining and Strauss
(1977) showed empirically that the trend towards population concentration in prior
decades reversed in 1970, as population began to deconcentrate again. In an extension to
Vining and Strauss (1977), Lichter (1985) computes similar indices of population
concentration by race and region for the years 1950-1980. Later work by Fugitt and Beale (1996) also reviewed trends in US population, though they look at nonmetropolitan migration over the past thirty years. Fugitt and Beale (1996) note that the migration balance favored nonmetropolitan areas in the 1970's, though in the 1980's, the migration balance shifted such that metropolitan areas were favored by migrants. However, by the 1990's the balance was again in favor of non-metropolitan areas as the trend reversed itself.

The nature of the spatial distribution of migration flows themselves has been explored in recent work. Rogers and Sweeney (1998) use the Gini index and the coefficient of variation to measure the spatial focus of migration streams in the US. Spatial focus is simply the degree of concentration of the migration streams between places. For example, Florida, which for some time has been the preferred destination of migrants from all over the country, receives migrants from different places and in different quantities, but would still be considered a focal point for a larger than average proportion of US migrants. In other work, Rogers and Raymer (1998) consider the spatial focus of US interstate migration flows. They detail the differences in spatial migration patterns based on age, race, and other disaggregate characteristics by time period.

Studies have also assessed other aggregate aspects of migration. Plane (1999) develops the concept of migration drift, which is a summary of the average distance and direction moved by migrants in a country during a given time period. Plane points out that the advantage of the migration drift statistic is that it only summarizes people's movement, thereby ignoring other aspects of population change. Essentially, this feature of the migration drift statistic renders it more advantageous than the population centroid
calculation, which is sensitive to the other components of population change. Interestingly, Plane finds that during the early 1990's the predominant east to west migration trend actually reversed itself, due in part to large amounts of out-migration from California. Other work by Plane (1999) accounts for the effects of migrants’ personal income on state economies. Plane found that the states receiving the largest gains in absolute dollars from in-migrants (1993-1994) were Florida ($3 billion), Arizona ($1.3 billion), North Carolina ($1.2 billion) and Georgia ($1.1 billion).

To summarize, the aggregate trends reported in these research papers are of great interest to geographers, regional scientists and demographers interested in spatial population phenomena. Additionally, such studies may be used to inform research efforts at other scales. For example, Beyers and Nelson (2000) investigate macro demographic trends at more localized scales. In a case study based on interviews of citizens residing in four nonmetropolitan western counties, Beyers and Nelson discuss linkages between economic factors such as non-farm income, resource dependency, etc. and local population growth. The next section presents the measures of accessibility to be used in our analysis.

2.2 Accessibility Indices

Accessibility or the relative potential of a given location is an important topic in urban and regional research (Craig 1987, Taaffe et al. 1996). The problem of measuring a given location's accessibility is one of determining the magnitude of opportunities within some specified distance or threshold of the location. Thus, when regional accessibility is
to be calculated, nearby population, income, or other measure of demand may serve as a proxy for the opportunity accessible to the location.

Operationally, a set of areas and their populations are easily input into an accessibility model such as the one used in the seminal work of Harris (1956) to measure US market potential. To demonstrate his model, we let $A_i$ be the accessibility at area $i$ and $P_j$ be the population at area $j$. Further, we let $d_{ij}$ define a matrix of straight-line distances between area centroids. The model is:

$$A_i = \sum_{j \neq i} P_j \ d_{ij}^{-1}$$

(1)

In the model above, distance between locations negatively impacts an area's accessibility. That is to say, if one holds population constant, more proximal locations contribute more to an area's accessibility than do distant ones. Indeed, the most accessible places are those that are near many other large centers of population.

Unfortunately, the model of potential in equation (1) is formulated such that an area’s own population is not counted into its accessibility score. This is a necessary exclusion because when $i=j$, $d_{ij}=0$, which would lead to an invalid divide by zero in equation (1). Although Harris’ model appears to be a plausible approach to measuring locational accessibility, in practice, a model of the form in equation (1), where an area’s own population is excluded from its accessibility score, may produce ‘donuts’ in the map pattern. For example, when these accessibility scores are visualized, it may be the case where highly accessible central counties actually appear to be less accessible than their peripheral neighbors. We point out that is possible to circumvent this property of equation (1) by finding nonzero values for $d_{ii}$. This might entail assigning each area’s
nearest neighbor distance to the diagonal of $d_{ij}$ (Plane and Rogerson 1994). However, a
different approach to accounting for an area’s own population in the accessibility score is
to model the deterrent effect of distance using the exponential function. Consider the
following equation:

$$A_i = \sum_j P_j \exp(-\beta d_{ij})$$

(2)

In equation (2), the exponential function in used to model the deterrent effect of distance
on an area’s accessibility score. Since $\exp(0) = 1$, the case of $d_{ii} = 0$ is easily handled,
therefore an area’s entire population is counted into its own accessibility statistic. The
decay of the exponential function in equation (2) is governed by the parameter $\beta$. The
parameter $\beta$ is shown with a negative sign by convention, as it is indicative of the
deterrent effect of distance (Fotheringham and O’Kelly 1989). To operationalize
equation (2) one need only choose a suitable value for $\beta$. So, suppose that at a distance,
$Q$ from area $i$, we want exactly half, or 0.5 of the $j^{th}$ area’s population to be counted into
$i^{th}$ area’s accessibility score. Then,

$$\exp(-\beta Q) = 0.5$$

(3)

Taking the natural logarithm of both sides yields:

$$\ln(\exp(-\beta Q)) = \ln(0.5)$$

(3a)

Rearranging terms in 3a and solving for $\beta$ explicitly yields:

$$\beta = -\frac{\ln(0.5)}{Q}$$

(3b)

A second approach to measuring regional population accessibility is to impose
some threshold to delimit which areas may count into the area's statistic (Plane and
Rogerson 1994). Notice the model in equation (2) imposes no strict limitations on an
area's accessibility score beyond those implied by the deterrent effect of distance. To remove this property from equation (3), we modify the formulas above such that only the population within some pre-specified distance, \( S \) is counted into the area's accessibility statistic. This yields the following formulation:

\[
A_i' = \sum_j P_j \quad \forall j \ni d_{ij} \leq S
\]  

(4)

Given this formulation, there are two subtle differences between the models in equation (2) and (4) that should be pointed out. First, because equation (2) uses continuous distance, it will always produce a more generalized map pattern than equation (4), unless one chooses a relatively large value for \( S \) in equation 4, or a very large value for \( \beta \) in equation 2. Second, a possible advantage of the model in equation (4) is that \( A_i' \) is calculated in terms of population, whereas \( A_i \) calculated in equation (2) is an index with a slightly less direct interpretation: it measures weighted population, where the weights are governed by the exponential distance decay parameter. This kernel smoothing technique is familiar to geographers, and is more or less a standard operation in commercial GIS packages. On the other hand, if we let \( S=50 \) in equation (4), the population accessible within 50 miles of the location currently studied is returned. However, both models provide useful measurements of accessibility, especially when there is an opportunity to investigate changes in them over time. The next section applies these two concepts of accessibility to population data from the 2000 and past censes.
3.1 Analysis

Data describing US county population from the 1990 and 2000 decennial census were obtained from the US census web site (www.census.gov). These data were imported into a GIS layer containing all of the US counties using TransCAD 3.2. For purposes of the study, only the counties of the continental US were included in the database.

A statistic familiar to demographers and geographers interested in population issues is that of population change. **Figure 1** shows the map of aspatial population change for the years 1990 and 2000. Darker shades are used to illustrate areas that have experienced greater positive population change. Although the map is quite varied in terms of spatial patterns, there are some trends to be pointed out. First, both the western and southeastern United States were among the fastest growing regions during the past decade. The five fastest growing counties during this time period were Douglas CO (191%), Forsyth GA (123%), Elbert CO (106%), Henry GA (103%) and Park CO (102%). Conversely, there were several regions that experienced flat or even negative population growth. Examples of these regions are found in the western plains states of Montana stretching southward to Texas, and the Appalachian core region extending from West Virginia into parts of eastern Ohio, western Pennsylvania and southern New York. Several of the counties along the Mississippi River in Louisiana and Arkansas also had little or no population growth.

3.2 County Accessibility Index Results

**Figures 2 and 3** show the maps of accessibility potentials as calculated by equation (2). For both years, we let $Q=100$ (see equation 3b), thereby counting half of the $j^{th}$ county’s
population at a distance of 100 miles away into the $i^{th}$ county’s accessibility score. Each county’s accessibility score is divided by an arbitrary scalar (1,000,000) to make them more manageable. The maps reveal very broad regional trends of accessibility, with the highest areas concentrated in the Northeast and Midwestern US. As one moves west, one generally finds declining accessibility until California is reached. Incidentally, the five most accessible counties in 1990 were Hudson NJ (1), New York NY (2), Essex NJ (3), Union NJ (4) and Kings NY (5). These five counties were also the most accessible in 2000.

Historically speaking, the spatial patterns of accessibility in both figures are not dissimilar to those found in Harris’ (1954) original work. Using data on county population from past censes, we are able to discern trends in the data and comment on long-term accessibility patterns. Just as was done for the 1990 and 2000 county population statistics, the exponential accessibility formula with $Q=100$ (see equation 3b) is calculated for each decade 1940 to 1980. Table 1 presents a summary statistic for each decade’s average county accessibility score. Depending on how the statistic is calculated, one finds quite differing trends in average accessibility for the United States. If all counties’ population and accessibility are each summed, and then the total accessibility is divided by the total population, one finds average accessibility ($\Sigma A_i / \Sigma P_i$). This quantity actually declines in our table due to total population growing faster than total accessibility over the period 1940-2000. Conversely, if each county’s accessibility score is divided by its population, and this ratio is added up for all counties, average accessibility generally increases from 1940-2000 ($\Sigma (A_i / P_i)$). The only anomaly in the trend is the flat accessibility from 1970 to 1980 (both scores were about 1.73). This
flatness in accessibility growth is perhaps explained by population deconcentration throughout the 1970’s as discussed in Vining and Strauss (1977).

Returning to the most recent decade, the latest two years' potential scores may also be compared simply by calculating the percent change in relative accessibility indices from 1990 to 2000. When this operation is performed, one roughly finds the inverse map to those maps appearing in Figures 2 and 3. Figure 4 shows that the areas experiencing the greatest percent gains in population potential were relatively inaccessible areas in both 1990 and 2000. On the other hand, the already accessible urban regions, particularly in the northeast maintained their stature from 1990-2000.

Table 2 reports that the counties experiencing the largest percent gains in relative accessibility ($A_i$) from 1990-2000 were Maricopa AZ (30.27%), Pinal AZ (29.87%), Gila AZ (29.41%), Pima AZ (28.48%) and Yavapai AZ (28.28%). If we consider the data from 1940 and calculate the percent change in accessibility from 1940 to 2000, we find a more varied set of fast-growing regions. Broward FL had the largest percent change in accessibility (721.46%), while other counties in Florida and Arizona rounded out the top five counties in Table 2. Table 2 also shows that the counties around New York City have been among the most accessible over the last 60 years. The percent change in accessibility (1940-2000) is mapped for the entire US in Figure 5. Similar to the changes in accessibility over the last decade, the trends in accessibility over the last 60 years may be characterized by increasing levels of potential in the southeast and western United States.
3.2 County Population Accessible Within 50 miles Results

To differentiate among the broad regional trends depicted in Figures 2 and 3, the model in equation (4) is calculated using $S=50$. Figure 6 shows the map of population accessible within 50 miles for 1990 (equation 4). Most of the major metropolitan areas in the United States appear as dark clusters of counties. The most interconnected urbanized areas, known as the megalopolises, appear as large expanses of darkly shaded counties in the northeast, midwest and southwestern United States. Equation (4) is calculated for the 2000 population data and mapped in Figure 7. In general, the map pattern of the 1990 statistic persists for the 2000 statistic. As shown in Table 2, the top 5 counties for largest population within 50 miles were the same for 1990 and 2000 (Middlesex NJ, Somerset NJ, Nassau NY, Bergen NJ, New York NY). However, when comparing individual counties from map to map, many times it is the case that metropolitan areas have expanded in terms of the number of counties that comprise them. For example, looking at the Atlanta area for 1990, one can see that the number of counties that are visually a part of the metropolitan area actually increased in 2000.

We may also look at $A_i^s$ from a historical perspective. Returning to Table 1, it is noted that the trends in county population within 50 miles virtually mirror the trends in the accessibility index ($A_i$). For the average of sums approach, we find that the summary statistic of population within 50 miles decreases over time, while for the sums of ratios approach, we find that the summary statistic of population within 50 miles increases over time.

These trends may be visualized for the entire set of counties simply by subtracting the 1990 population within 50 miles from the 2000 population within 50 miles. From this
operation, we know how many people within 50 miles a county added, or lost in some cases. These values are mapped in Figure 8 and reveal some very interesting patterns.

First, one notices that several regions have lost population as evidenced by the Appalachian region and some of the plains states. Conversely, the coastal areas of the mid Atlantic States and the southeastern United States systematically added population. Similar gains in regional population were made for the Midwest, Florida and California.

As listed in Table 2, the counties adding the largest number of persons within 50 miles from 1990 to 2000 were Morris NJ (1.43 million), Richmond NY (1.39m), Somerset NJ (1.38m), Bergen NJ (1.37m), Union NJ (1.36m). Using the historical data and performing the same calculations, we found that that the counties of southern California and Nassau NY added the most people within 50 miles from the period 1940-2000, as shown in Table 2. The difference in county population within 50 miles is also shown in Figure 9.

Striking is the degree to which many regions of the US lost population from 1940 to 2000. The Plains States, the Mississippi River Basin and Appalachian region all appear to have lost population. In contrast, regions such as the Piedmont stretching from North Carolina to Georgia, the east Lakes of the Midwest, and Northeast all gained population. Similarly, Figure 9 also shows that California and Florida made significant gains in population within 50 miles from 1940-2000.
4. Conclusions

Using newly released data from the 2000 decennial census and prior years’ censes, we have illustrated regional population trends in terms of county-based accessibility measures. Our results based on both accessibility approaches show that much has changed in terms of regional population patterns during the past decades.

Perhaps our most striking findings are those illustrating the US regions that have lost population. Clearly, regions such as the plains states and Appalachia lost out in terms of population growth during the past decade. Possibly accounting for this phenomenon, Frey (1993) suggests that recent patterns of growth and decline reflect shifts in industrial structure, and favor areas with diversified economies consisting of higher order service and information processing industries. Anecdotally, these characteristics are the norm for the economies of the urban southeastern United States. Conversely, Frey’s assertion helps to explain why areas such as the plains states and Appalachia, which have less diversified economies, are not experiencing the same level of growth as other regions.

In short, we would anticipate much more research on the spatial aspects of population change. The new census data should facilitate further research into many substantive areas, including the regional growth of the southeast and Atlanta, the population decline of Appalachia, the expansion of metropolitan areas and other interesting topics.
5. References


Table 1: Summary of Historical Accessibility Indices

### Average of Sums Approach

<table>
<thead>
<tr>
<th>Year</th>
<th>Population (millions)</th>
<th>Total of US Accessibility scores (Ai)*</th>
<th>Average Ai Index (per capita)</th>
<th>Total of US Population within 50 miles (Ais)*</th>
<th>Average Ai Index (per capita)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940</td>
<td>131.67</td>
<td>19,469.77</td>
<td>147.87</td>
<td>1,714.24</td>
<td>13.02</td>
</tr>
<tr>
<td>1950</td>
<td>150.70</td>
<td>21,660.63</td>
<td>143.74</td>
<td>1,923.70</td>
<td>12.77</td>
</tr>
<tr>
<td>1960</td>
<td>178.46</td>
<td>24,762.16</td>
<td>138.75</td>
<td>2,210.92</td>
<td>12.39</td>
</tr>
<tr>
<td>1970</td>
<td>202.14</td>
<td>27,452.20</td>
<td>135.81</td>
<td>2,468.18</td>
<td>12.21</td>
</tr>
<tr>
<td>1980</td>
<td>225.18</td>
<td>29,852.98</td>
<td>132.57</td>
<td>2,667.41</td>
<td>11.86</td>
</tr>
<tr>
<td>1990</td>
<td>247.05</td>
<td>31,671.76</td>
<td>128.20</td>
<td>2,855.68</td>
<td>11.56</td>
</tr>
<tr>
<td>2000</td>
<td>279.59</td>
<td>35,333.00</td>
<td>126.37</td>
<td>3,194.28</td>
<td>11.42</td>
</tr>
</tbody>
</table>

### Sums of Ratios Approach

<table>
<thead>
<tr>
<th>Year</th>
<th>Population (millions)</th>
<th>Average Ai Index*</th>
<th>Average Ai Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940</td>
<td>131.67</td>
<td>1.14</td>
<td>68,988.58</td>
</tr>
<tr>
<td>1950</td>
<td>150.70</td>
<td>1.32</td>
<td>77,537.06</td>
</tr>
<tr>
<td>1960</td>
<td>178.46</td>
<td>1.56</td>
<td>86,052.26</td>
</tr>
<tr>
<td>1970</td>
<td>202.14</td>
<td>1.73</td>
<td>90,704.83</td>
</tr>
<tr>
<td>1980</td>
<td>225.18</td>
<td>1.73</td>
<td>89,523.44</td>
</tr>
<tr>
<td>1990</td>
<td>247.05</td>
<td>1.87</td>
<td>92,409.56</td>
</tr>
<tr>
<td>2000</td>
<td>279.59</td>
<td>2.00</td>
<td>93,601.01</td>
</tr>
</tbody>
</table>

*Scores are scaled by 1x10^-6

Table 2: Historic Accessibility Trends for Specific Counties

#### Top 5 accessible counties based on Ai

|------|------|------|------|------|------|------|------|---------------------------|

#### Top 5 accessible counties based on Ais

|------|------|------|------|------|------|------|------|----------------------------------------------------------|