A Spatial Analysis of Population Change in the United States, 1990-2000

Matthew J. Moehr
mmoehr@ssc.wisc.edu
Department of Sociology and Rural Sociology
University of Wisconsin - Madison
350 Agriculture Hall
1450 Linden Drive
Madison WI 53706

The current study addresses the demographic and geographic aspects of population change in the United States during the 1990s. Census data and vital statistics are used to calculate natural increase rates and net migration rates for all 3070 counties in the contiguous United States. The components of population change are broken down into three race/ethnicity groups - black, Hispanic, and white - and a series of exploratory and regression analyses are conducted to examine the relationship between natural increase and net migration. The natural increase rates of blacks, Hispanics, and whites are each significant predictors of total net migration rates for a county. The data also provides evidence for spatial processes such as diffusion and regionalization in the patterns of population change.

BACKGROUND

Population distribution in America across the rural to urban continuum has been a focus of demography and population geography for decades. Studies tend to categorize some unit of geography – usually the county – into a hierarchy based on population. However, the hierarchical analysis is often used to reach spatial conclusions, which may or may not be valid given the spatial distribution of population. For instance, claims of “deconcentration”, “regionalization”, and “segregation” all imply a real-world, spatial arrangement of population, but space is often ignored as researchers aggregate areas into a hierarchy. This paper will attempt to use an explicitly spatial approach to measure change in population distribution among continental U.S. counties during the 1990s. First, a brief literature review on the themes of deconcentration, regionalization, and segregation is offered as a baseline for understanding the most commonly cited processes of population distribution. Then, a series of spatial analyses are conducted on the components of population change in the U.S. – net migration and natural increase – to test the operation of these three processes across space. The conclusions do not entirely contradict previous research, but this work suggests that there are important theoretical and empirical differences between hierarchical and spatial approaches to understanding population distribution.

Claims of population deconcentration over the last five decades must be couched in the long-run trend from 1790 to 1900 during which American population moved to frontier areas and then began to concentrate in urbanized areas from 1900 to the present (Otterstrom 2001). Much of the research which focuses on more recent decades relies on the concept of "equilibrium seeking" (Otterstrom 2001; Manson & Groop 2000) or detailing localized sub-processes within an ongoing urbanization trend (Elliott 1997). In either case, researchers usually test the implicit hypothesis that population is moving down a hierarchy from urban to suburban to rural counties. The evidence for population deconcentration has been mixed. The 1970s witnessed a higher rate of population retention and more in-migration for rural counties (Beale 1975) leading to the widespread use of the term "rural rebound" (cf. Johnson & Fuguitt 2000; Johnson, Nucci & Long 2005). The 1980s did not continue this trend but did not entirely revert to long-run urbanization trends, either (Johnson 1993). The 1990s resembled a less dramatic version of the 1970s for some age-groups and economic-sectors within rural areas (Plane, Henrie & Perry 2005; Johnson, Voss, Hammer, Fuguitt & McNiven 2005). Migration to rural areas was strongest in the early and middle part of the decade but was significantly lower after about 1997. Since the 2000 Census, post-censal population estimates have been used to suggest that there may be a slight upturn in migration to rural counties since the lull at the end of the 1990s (Johnson, Nucci & Long 2005).

For the most part, this approach to population change does not use counties as geographical units arranged in space; instead, the research is much more likely to treat a county as a category within a rural-urban continuum. For example, Plane, Henrie, and Perry (2005) begin with Ravenstein's "step migration" theory and delimit a hierarchy of counties with five sizes of metropolitan counties, one size of micropolitan county, and one size of non-CBSA counties. Using layer cake diagrams, the authors create very clear visual representations for the movement of population down the hierarchy (Plane, Henrie & Perry 2005: 15314-5). However, deconcentration may be well supported as a national trend within the rural to urban hierarchy, but it may not actually occur across the space adjacent to all urban areas in the country.

Regionalization processes may help to explain larger scale population shifts. Increased migration and natural increase have spurred population growth in more southern and coastal areas during the 1990s. Two explanations for regional shifts in population have been posited (Frey & Johnson 1998). First, the climatic and recreational benefits of Sun Belt and amenity-rich locations have been identified as major pull factors for retirees and other migrants (Cook & Mizer 1994; McGranahan 1999). Second, the economic restructuring of the 1970s and 1980s increased the push factors from farming and manufacturing dependent areas such as the Midwest and Great Plains (Noyelle & Stanback 1984). While this area of literature seems to make an explicitly spatial argument – migration flows from one region to another region – the definition of 'region' will obviously have important effects on the conclusions. Allowing for a gradual transition between 'regions' within in the U.S. instead of arbitrarily designating some places as part of the Sun Belt or Rust Belt, for example, seems like a preferable operationalization of geographic space.

Both deconcentration and regionalization trends are often cited in literature as promoting residential segregation based on race/ethnicity and socioeconomic status (see for example, Manson & Groop 2000; Massey & Denton 1985, 1987, 1993). The relation between deconcentration, regionalization, and racial/ethnic segregation was the motivation to include natural increase rates broken down into race/ethnic categories in the current analyses. Much of the research on residential segregation would suggest two reasons for the statistical relationship between natural increase rates and net migration rates. First, NIRs are different across race/ethnic categories, so in some ways, the NIRs act as surrogate indicators of the segregation present at the beginning of the 1990s. Second, NIRs also indicate the age-structure of the population, and thus, may be able to capture places attractive to young families, which would presumably create higher fertility rates and lower mortality rates.

For example, Logan, Stults, and Farley (2003) argue that Hispanics experienced an increase in residential segregation in America's metropolitan areas during the 2000s, even despite the fact that Hispanic migration into less segregated counties was higher than Hispanic migration into more segregated counties. The authors suggest the following mechanism to explain the paradox: 1) in-migration during the 1980s and early 1990s of Hispanics into less segregated counties, 2) higher population growth rates (high birth rates and low death rates) for Hispanics than other race/ethnicities, 3) resulting in higher average segregation ratios because Hispanics are living in counties with higher percentages of Hispanics at the end of the decade. Note here that the authors, following a general trend in residential segregation research, use statistics such as dissimilarity scores for individual Census Tracts, which are then averaged across a metropolitan area or across the entire nation to draw conclusions about the changes in residential segregation over time. This may not be the best method for examining the true amount of residential segregation across space.

Many conclusions about residential segregation tend to treat human individuals as merely agents of rational decision making (cf. Massey, Arango, Hugo, Kouaouci, Pellegrino & Taylor 1993 for a discussion of migration theories); however, there is a broader theoretical framework which attempts to incorporate social networks, family ties, and cultural factors into explanations of residential segregation. Especially relevant to the current analysis is work on the patterns in migration decisions exhibited by non-white individuals. For example, there is evidence that blacks are more likely than whites to return to a previous home after initial migration (Newbold 1997). In terms of regionalization, blacks have exhibited higher migration than whites into the Southeast since the 1970s after a period of very high black out-migration from 1940 to 1970 (Adelman, Morett & Tolnay 2000). Specifically related to the current focus on natural increase rates, studies have supported the conclusion that kinship ties are more important in determining migration decisions for single black women with children than is the black population in general (Johnson & Roseman 1990). Taken together, this body of literature would suggest that the natural increase rates for non-white groups have important predictive power for the overall migration rates of a particular county.

This is certainly not an exhaustive overview of the literature on population distribution, but it does serve to outline the three processes of deconcentration, regionalization, and segregation. Hopefully, it is now clear that much of the work described here has failed to incorporate the spatial distribution of population changes. This is especially unfortunate, because the use of space offers one possible way to integrate all three processes in an understanding of population distribution. The current paper is a first attempt to examine the interplay of deconcentration, regionalization, and segregation across space. Instead of assuming a hierarchical or arbitrary designation of counties in the United States, the spatial analyses below will only need to adhere to Tobler's First Law of Geography as a theoretical underpinning. Namely, "… everything is related to everything else, but near things are more related than distant things" (Tobler 1970: 236).

METHODS

The data used for the current analysis is drawn from the 1990 and 2000 U.S. Census and from vital statistics provided by the National Center of Health Statistics (Voss, McNiven, Hammer, Johnson & Fuguitt, 2003). The exact methodology used to compute county-level birth, death, and net migration statistics is detailed in a working paper from the Center for Demography and Ecology at the University of Wisconsin-Madison (Voss, McNiven, Hammer, Johnson & Fuguitt 2004). There are three important considerations to keep in mind for all analyses drawn from this dataset. First, all counties in Alaska and Hawaii have been removed from the working dataset because the spatial analyses require counties to share borders. Second, there were several county boundary changes between 1990 and 2000. For the most part these changes were dealt with by aggregating birth, death, and net migration statistics into larger geographic areas which included all territory *ever* included in the county. While this approach sacrifices some precision, it allows for the comparison of identical geographic units in 1990 and 2000. These two changes resulted in a final database of 3070 counties in the contiguous United States.

A third consideration of the data is the change in the race/ethnic groups between the 1990 and 2000 U.S. Census. The main problem is data consistency stemming from the ability of respondents to select multiple race/ethnic categories on the 2000 version. The data used here make use of proportional distribution methods for all respondents that selected more than one race/ethnic group in 2000 in order to maintain comparability in the populations across the decade. Briefly, a person that selected Asian and black on the 2000 Census would be proportionally distributed ½ to Asian and ½ to black populations within the county; however, this procedure is not used to distribute those people that selected Hispanic as an ethnic origin. In the analyses that follow, Hispanic is defined as a person of any race that indicated Hispanic origin, white is defined as non-Hispanic whites, and black is defined as non-Hispanic blacks. Because of relatively small populations in many counties and increased problems with proportional distribution, additional race/ethnic groups – for example Asians, Native Americans, and others – are excluded from the present study.

Exploratory spatial data analysis was carried out using the software package GeoDa[1]. Further exploratory techniques were utilized within Geographically Weighted

[1] GeoDa software and documentation available at: http://sal.agecon.uiuc.edu/geoda_main.php. See for example, Anselin, Luc, Ibnu Syabri, and Youngihn Kho. 2004. "GeoDa: An introduction to spatial data analysis."

Regression[2] (GWR) software. The standard regression models were estimated using Ordinary Least Squares methods, and all spatial regression models were estimated using Maximum Likelihood methods – both within GeoDa. Throughout, the ESRI ArcGIS[3] suite version 9.1 was used to map and display the outputs of the spatial data analysis.

RESULTS & ANALYSES

This section contains five sub-sections which address the main steps of the spatial analyses: exploratory analysis of the dependent variable, exploratory and GWR analysis of the explanatory variables, trend surface analysis, specification of a standard linear regression model, and the specification of spatial regression models.

Net Migration Rates as Dependent Variable

Figure 1 displays the county-level net migration rates in six equal categories with lighter colors representing more out-migration and darker colors signifying more in-migration. Conducting a cursory visual assessment should yield an immediate conclusion: there is a very dramatic region of out-migration in the center of the country stretching from western Texas to the Canadian border. West of this area many counties register in the very highest in-migration categories, and east of the this area there is strong clustering of in-migration counties around major metropolitan areas such as Atlanta, GA, Orlando, FL, and San Antonio, TX. Of course it is important to keep in mind the differences in county size in the western and eastern halves of the United States, which may be contributing to the visual pattern of widespread in-migration in the west and metropolitan clusters of in-migration in the east. The Moran's I statistic reported in

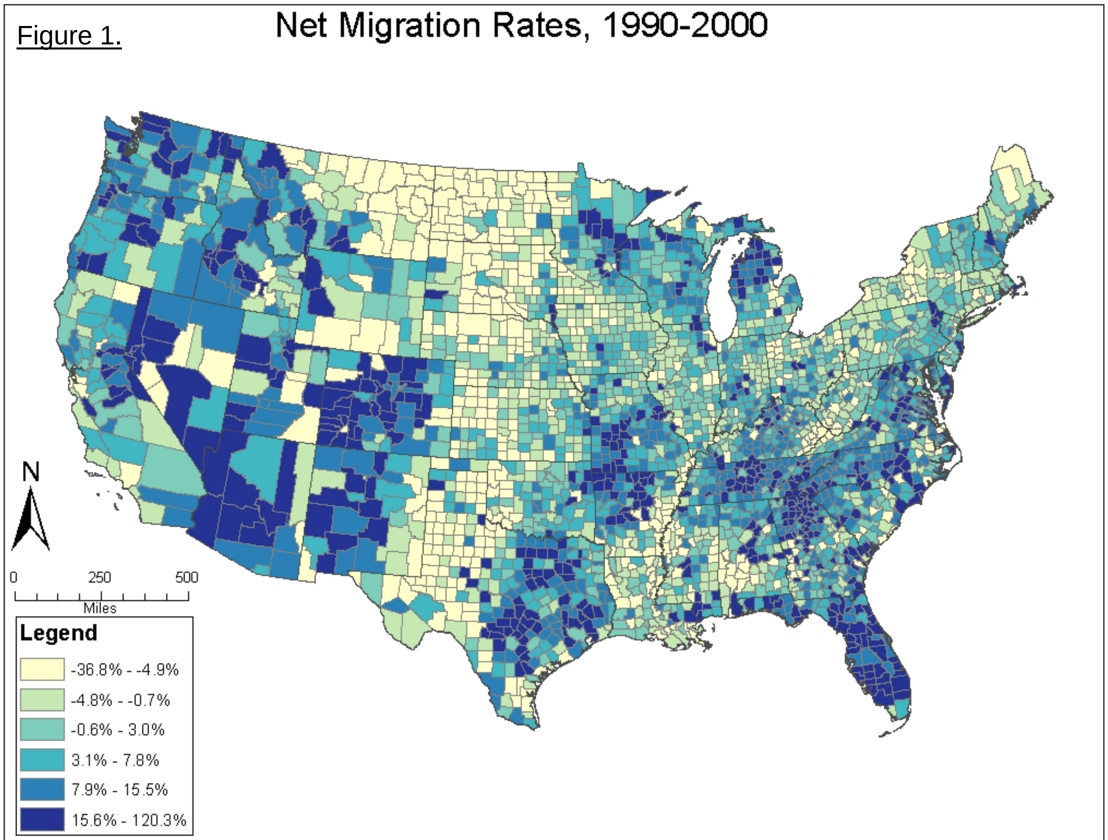
[2] Geographically Weighted Regression (GWR) software and documentation available at: http://www.nuim.ie/ncg/GWR/index.htm. See for example, Charlton, Martin, Stewart Fotheringham, and Chris Brunsdon. 2003. "GWR 3: Software for geographically weighted regression."
[3] More information about ArcGIS suite version 9.1 is available from the ESRI website: http://www.esri.com/software/arcgis/.

Table 1 confirms that there is a strong positive autocorrelation among county-level net migration rates[4].

Natural Increase Rates as Explanatory Variables

County-level natural increase rates show significant autocorrelation, as is shown by the Moran's I reported in Table 1. More importantly, though, the natural increase rate

Figure 1.



[4] The Moran's I statistic is the most commonly used measure of spatial autocorrelation (Cliff and Ord 1981). For target county j and all other counties i, it takes the form

$$I_j = \Sigma_i \boldsymbol{W} x_i x_j / x_j^2$$

where $\boldsymbol{W}$ is a weight matrix that defines the neighborhood of each county, and the x's represent the deviation from the global mean net migration rate for either the target county or all counties. Summing across all possible target counties, i.e. summing across all j's, results in the global Moran's I value for all counties. The Moran's I value can generally be interpreted like a Pearson's correlation coefficient between the net migration rates of a target county and the average net migration rates in its neighborhood. Therefore, a positive Moran's I signifies clusters of similar values (high-high or low-low) and a negative Moran's I signifies the juxtaposition of dissimilar values (resembling a checkerboard pattern). The for each of the three race/ethnicity groups is also highly autocorrelated. This suggests that there are clusters of counties throughout the United States with above average (and below average) natural increase rates for each race/ethnicity group. These clusters provide evidence that natural increase rates are not the same across all counties, and the differences in natural increase rates may help to predict differences in net migration rates. There are two possible interpretations of the natural increase rates as explanatory variables, and each argues for a slightly different interpretation of the segregation processes of population distribution. Although not strictly used as a hypothesis testing instrument, GWR software allows for the visualization of the three race/ethnic groups' natural increase rates as predictors of net migration rates across the country. Because of the way the software works, it is possible to see if different natural increase rates are better or worse predictors in different areas of the country. It will be explained how this can be used for evidence to decide between the two possible interpretations of natural increase rates.

Table 1. Descriptive Statistics.

	Global Mean	Minimum	Maximum	Moran's I
Net Migration Rate	0.0540	-0.3684	1.2034	0.3455***
Total NIR	0.0342	-0.3513	1.5701	0.1532***
Black NIR	0.0423	-10.5000	1.0000	0.0350**
Hispanic NIR	0.1340	-16.0000	1.0000	0.0601**
White NIR	0.0127	-0.1897	0.2383	0.3102***

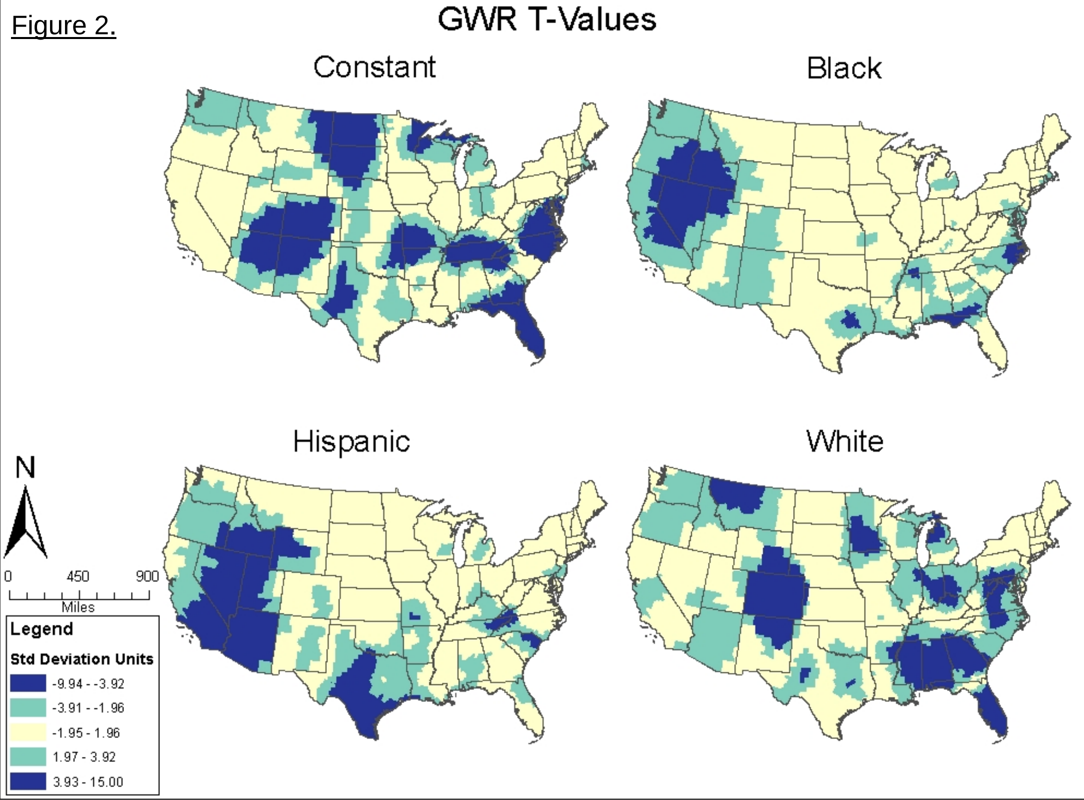
* p<.05, ** p<.01, *** p<.001

significance test for the Moran's I statistic represents the chance that the given spatial pattern was drawn from random distribution across counties of the observed net migration rates.

Natural Increase Rates as Socioeconomic Prosperity Indicator. High natural increase rates are direct indications of a county with a high birth rate and a low death rate. This is a simple demographic fact, but the underlying causes of high birth rates and/or low death rates are less clear. One possibility is that a high birth rate may be an indicator of a county's health care, education, or economic infrastructure. Also, a low death rate may arise from a relatively lower crime rate, poverty rate, or better access to health care. Taken as a whole, a higher natural increase rate could be interpreted in this manner as an indicator of a more socioeconomically prosperous county. This interpretation of natural increase rates draws from the theoretical background offered by economic explanations of net migration, such that a county blessed with beneficial characteristics would be hypothesized to display a relatively higher natural increase rate for all three race/ethnic groups and in most cases also a higher net migration rate. If natural increase rates do indeed indicate socioeconomic prosperity, the GWR output should show that all three race/ethnic group natural increase rates vary together as predictors of net migration rates.

Natural Increase Rates as Age-Race Structure Indicator. As an alternative, natural increase rates could be interpreted as an outcome of residential sorting based on age and race characteristics. To be more precise, a county's natural increase rate may be understood as a function of its inhabitants. If, for example, a county has historically attracted and/or retained fertile-aged men and women of a particular race/ethnic group then that county will be more likely to have a high natural increase rate for that race/ethnic group. If, on the other hand, a county has lost fertile-aged adults and gained older adults – most likely in the form of retirees – within a certain race/ethnic group, then that county would be much less likely to have a high natural increase rate for the 1990s.

Figure 2.

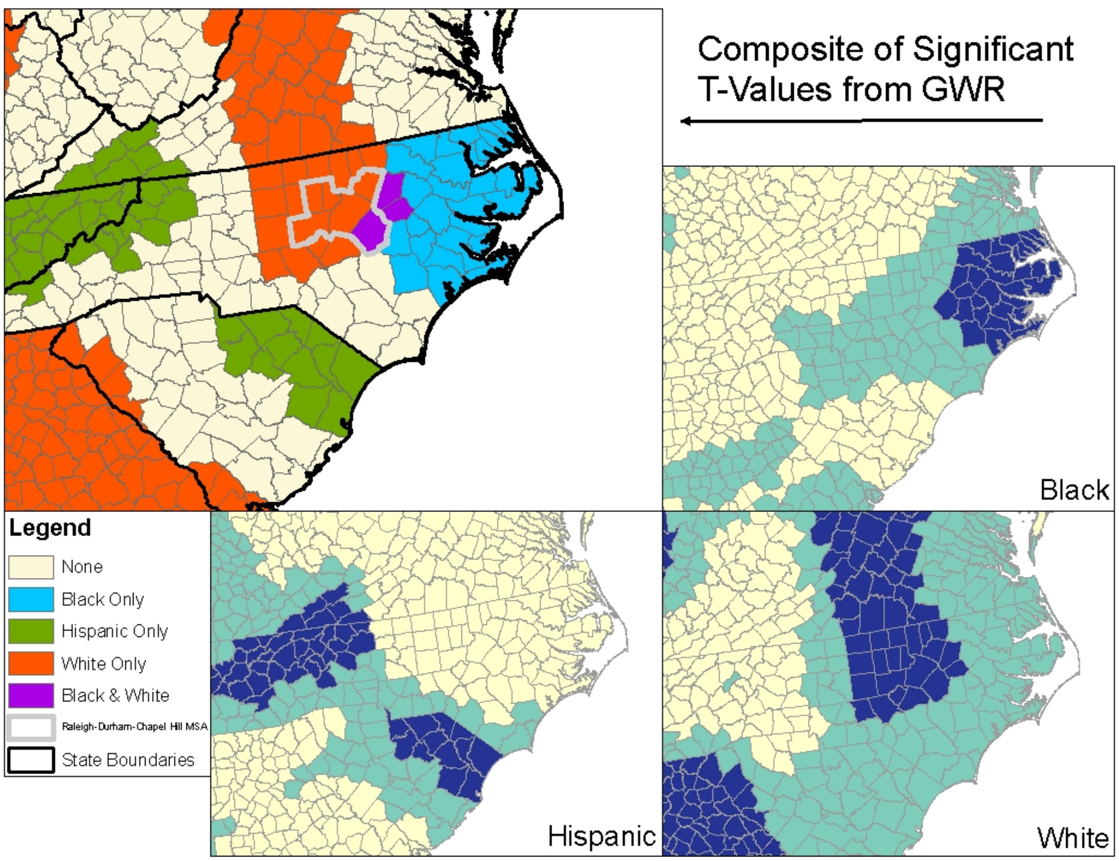


This interpretation is closely aligned with the theories of racial/ethnic residential segregation outlined above, which suggested that natural increase rates could be used as indicators of the race/ethnic composition and the age-structure of a county. Given this interpretation, it would be hypothesized that net migration rates would be most closely related to the natural increase rate of the largest and most mobile race/ethnic group in the county. If natural increase rates are capturing the effect of the age-race structure in a county, the GWR output should show different "hot" and "cool" areas for each race/ethnic group coefficient. One specific prediction, for example, would be a highly significant coefficient for the Hispanic NIR in a band from California to Texas while the two other groups would not show up as significant in these areas.

Figure 2 displays four maps of the significance level for each variable in the GWR model: constant, black NIR, Hispanic NIR, and white NIR. As can be clearly seen, there is very little overlap between each of the three race/ethnic groups. I therefore conclude that the second interpretation of natural increase rates as indicators of the age-race structure is the more valid approach. In other words, a county's historical trends in the migration and retention of fertile aged adults of certain race/ethnic groups is a significant predictor of further net migration rates for the county.

One great example of this is the Virginia-North Carolina-South Carolina[5] area shown in Figure 3. Almost no county falls within the darkest region in more than one of these maps, and the composite map shows distinct regions where the NMRs would be best predicted based on a different set of NIRs. For example, west of Raleigh, NC the

Figure 3.



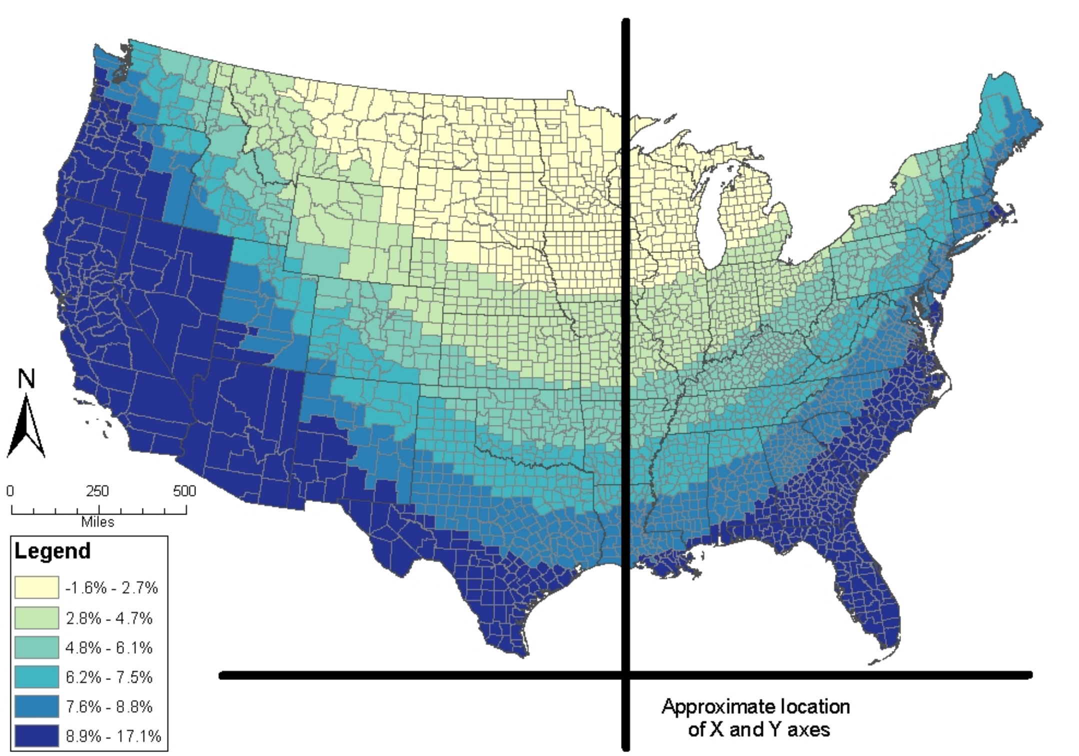
[5] This example was inspired by the ethnographic dissertation research of Helen B. Marrow, in which she describes the interaction between African Americans and new Hispanic immigrants in rural South Carolina. For more information see: http://www.rprconline.org/Fellowships/Marrow_Abstract.pdf.

white NIRs are significant predictors and east of Raleigh black NIRs are significant predictors of net migration, and there are only three counties of overlap. This suggests that there are important differences in the relationship between the natural increase rates of certain race/ethnic groups in a county and the overall net migration rate into that county.

<u>Trend Surface Analysis</u>

A trend surface analysis was conducted in order to measure and control for spatial drift during the 1990s in net migration patterns across the country. This "drift" can be thought of as the regionalization process described above: an increase in retirees migrating to the southwest and Florida combined with out-migration from the upper Great Plains. To find the spatial drift in county net migration rates, a model was

Figure 4. Trend Surface Predicted NMR Values



constructed which used the X and Y coordinates of each county's centroid as the set of explanatory variables. The complete trend surface model included coefficients for X, Y, X^2, Y^2, and X*Y; however, because of where the X- and Y-axis fell in relation to the United States map (see Figure 4), only three of the explanatory variables were deemed to have logical interpretations. An example of an illogical coefficient is X*Y, which represents the effect of county location being in *either* the northwest or the northeast, and since these counties are not a continuous region within the United States, it does not seem useful to construct a measure of the spatial drift towards these areas.

In the reduced model, the negative coefficient for X signifies that counties further to the west have a higher net migration rate. Likewise Y has a negative coefficient signifying a more southerly location predicts a higher net migration rate. Since X^2 has a positive coefficient, there is statistical evidence to support the visual inference such that counties further away from the center of the United States have higher net migration rates. Table 2 compares the complete trend surface model with the model containing

Table 2. Regression Coefficients of Trend Surface Model.

	Full Surface Model	Reduced Surface Model
Constant	0.0688**	0.1236***
X	0.6163***	-0.0697**
Y	0.0692	-0.4847***
X^2	0.1931***	0.2261***
Y^2	-0.1340	
X*Y	-0.3441***	
Adjusted R^2	0.0719	0.0558
AIC	-3855.19	-3804.43
SC	-3819.01	-3780.31

* p<.05, ** p<.01, *** p<.001

only "logical" coefficients, and it should be noted that the model selection criteria all point to the complete model. This difference disappears when the trend surface models are combined with natural increase rates as explanatory variables[6], and therefore, the reduced trend surface model will be used throughout to aid in the interpretability of the spatial models.

Standard Regression Model

An ordinary least squares estimation of a linear regression model was conducted to understand the global relationships between natural increase rates and net migration rates. Coefficients and model fit statistics are summarized in the first column of Table 4, which can be found in the next sub-section. Of particular note with this model is the generally low adjusted-R^2 value of 0.1085, indicating that there is quite a bit of variation in net migration rates not captured by natural increase rates.

To obtain spatial diagnostic statistics, a county's neighborhood is defined as a second-order queen contiguity matrix[7]. This means that each county has a neighborhood including every county it shares an immediate border with, and additionally, all counties that share a border with this first "ring" of neighbors (i.e. the neighbors' neighbors). As

[6] All OLS and spatial models were run with both the full surface model, reduced surface model, and no surface model. The addition of the reduced surface model improved the predictions of NMRs the most in the OLS, spatial lag, and spatial error models. However, neither the full or reduced surface model contained any significant variables within the full SARMA model described below.
[7] Within the field of spatial analysis, the specification of the neighborhood structure is often given little theoretical or empirical attention, and instead is often picked in an arbitrary or pragmatic way. This is an unfortunate reality because the geographic "neighborhood" for social phenomenon is usually not arbitrary. I am also vulnerable to this critique of spatial analysis, but I would offer that the social phenomenon of net migration has often been shown to operate within a neighborhood of immediately adjacent counties. For example, the U.S. Census Bureau's definition of metropolitan statistical areas (MSAs) includes a central city's neighboring counties to make sure that people in suburban areas who still participate in the economic and social life of the central city are counted among the city's population.

Table 3. Spatial Diagnostics of Standard Regression.

	Value	Probability
LM (lag)	3165.92	.0001
LM (error)	3305.06	.0001
Robust LM (lag)	50.39	.0001
Robust LM (error)	189.53	.0001

can be seen in Table 3, Lagrange Multiplier statistics for spatial lag and spatial error are both highly significant. This indicates that the spatial autocorrelation in the natural increase rates probably violates the homoskadasticity and random sample assumptions[8] for OLS estimation of regression coefficients, and therefore, the OLS estimates can not be considered unbiased. A spatial error model attempts to estimate the effect of heteroskadasticity and a spatial lag model attempts to control for the effect of spatial autocorrelation in the predictor variables (Anselin 1988). A closer inspection of the spatial diagnostics reveals that the value for the Robust LM of the spatial error is more than three times that of the Robust LM of the spatial lag, which may indicate that a spatial error model may fit the data better. The spatial error and spatial lag models have implications for the deconcentration process, which is discussed below.

[8] Homoskadasticity is the assumption that the predictor variables are uncorrelated with the random disturbances (and likewise the regression residuals) in the outcome variable. So for the current study, this would require NIRs to be uncorrelated with all of the unmeasured variables that may affect NMRs. Obviously my above discussion about the possibility that NIRs are indicators of socioeconomic prosperity is a direct refutation of the assumption that NIRs are uncorrelated with the random disturbances in Y. The random sampling assumption within OLS estimation states that for all cases i and j, $\Sigma_{ij}x_i x_j = 0$, or in other words, every NIR is uncorrelated with every other NIR. The Moran's I reported above shows this assumption to be invalid as the NIRs that are located closer geographically are also more similar in their values, which another way of saying the NIRs are spatially autocorrelated.

Spatial Regression Models

Before estimating the spatial lag and spatial error models, it is important to consider the theoretical underpinnings of each model. The spatial lag model takes the form:

$$\hat{y} = \rho \mathbf{W} y + \beta X + \varepsilon$$

where $\mathbf{W}\boldsymbol{y}$ is the multiplication of the neighborhood weight matrix by an n×1 vector of all the other values of y, $\boldsymbol{\beta X}$ is an n×k matrix of explanatory variables multiplied by a k×1 vector of parameters for each X, and ε is a normally distributed disturbance term. In the above equation, ρ (Rho) is a scalar parameter that indicates the effect of the dependent variable in the neighborhood on the dependent variable in the target county; in other words, rho is the amount of spatial lag. So for the present application, the spatial lag model resembles:

$$\begin{aligned}\text{N}\hat{\text{M}}\text{R in target county} = {} & \rho(\text{NMR in neighborhood}) \\ & + \beta_1(\text{Black NIR}) + \beta_2(\text{Hispanic NIR}) + \beta_3(\text{White NIR}) \\ & + \beta_4(\text{X - Coordinate}) + \beta_5(\text{Y - Coordinate}) + \beta_6(\text{X - Coordinate}^2) \\ & + \varepsilon\end{aligned}$$

In general terms, a spatial lag model is often said to take into account the diffusion of a process – in this case migration of individuals – from one unit to other units in its neighborhood. This means that a spatial lag model uses the natural increase rates of the three race/ethnic groups of the target county, the location of the target county within the trend surface, and the net migration of the surrounding neighborhood to predict net migration rates for each county.

The spatial error model takes a form described by the two equations:

$$\hat{y} = \beta X + u,$$
$$u = \lambda \mathbf{Wu} + \varepsilon$$

where ***βX*** is the same as the lag model, ***Wu*** is the neighborhood weight matrix multiplied by an n×1 vector of all the other disturbances of y, and ε is the normally distributed portion of the disturbances in y. In the spatial error model, λ (Lambda) is a scalar parameter for the effect of the disturbances in the neighborhood on the dependent variable in the target county. To summarize the current application:

$$\begin{aligned}\text{N}\hat{\text{M}}\text{R in target county} = {} & \beta_1(\text{Black NIR}) + \beta_2(\text{Hispanic NIR}) + \beta_3(\text{White NIR}) \\ & + \beta_4(\text{X - Coordinate}) + \beta_5(\text{Y - Coordinate}) + \beta_6(\text{X - Coordinate}^2) \\ & + \lambda(\text{disturbances in neighborhood}) + \varepsilon\end{aligned}$$

If the OLS error term is thought of as an accumulation of both random effects and unmeasured predictor variables, the spatial error model disaggregates the two and places the effect of autocorrelated unmeasured variables in Lambda. The spatial error model includes the same predictor variables and trend surface variables as the spatial lag model, but instead of including net migration rates from the surrounding counties, it specifies the amount of autocorrelated error in the neighborhood.

The second and third columns of Table 4 specify the parameters for the spatial lag and spatial error models, respectively. To begin the comparison it is useful to examine the model fit statistics for both spatial models, and it is seen that there are very negligible differences between the two in terms of R^2, AIC, and SC. This was foreshadowed by the diagnostic statistics from the standard OLS model, which indicated that both spatial models would have highly significant contributions to net migration rate predictions. Furthermore, inclusion of a spatial coefficient in both models seems to have the same general effects on the other explanatory variables. In both spatial models black and

Table 4. Regression Coefficients of Standard Regression and Spatial Models.

		Standard OLS Model	Spatial Lag Model	Spatial Error Model	Full Model (SARMA)
	Constant	0.1238***	0.0264***	0.1112***	-0.0090*
Natural Increase Rates	Black NIR	-0.0161*	-0.0125	-0.0122	-0.0104
	Hispanic NIR	0.0160**	0.0120*	0.0113*	0.0095*
	White NIR	0.7168***	0.4452***	0.5033***	0.2239***
Spatial Coefficients	Rho (Lag)		0.7789***		1.0611***
	Lambda (Error)			0.7974***	-0.9416***
Trend Surface	X	-0.0847***	-0.0136	-0.0084	0.0091
	Y	-0.5204***	-0.1309***	-0.4516**	0.0181
	X^2	0.1894***	0.0276	0.1639*	-0.0244**
Model Fit	Adjusted R^2	0.1085	0.3553	0.3556	0.4176
	AIC	-3977.62	-4853.84	-4848.95	-5187.33
	SC	-3935.42	-4805.61	-4806.75	-5139.10
	Moran's I of Residuals	.3205***	-0.0251***	-0.0289***	0.0059

* p<.05, ** p<.01, *** p<.001

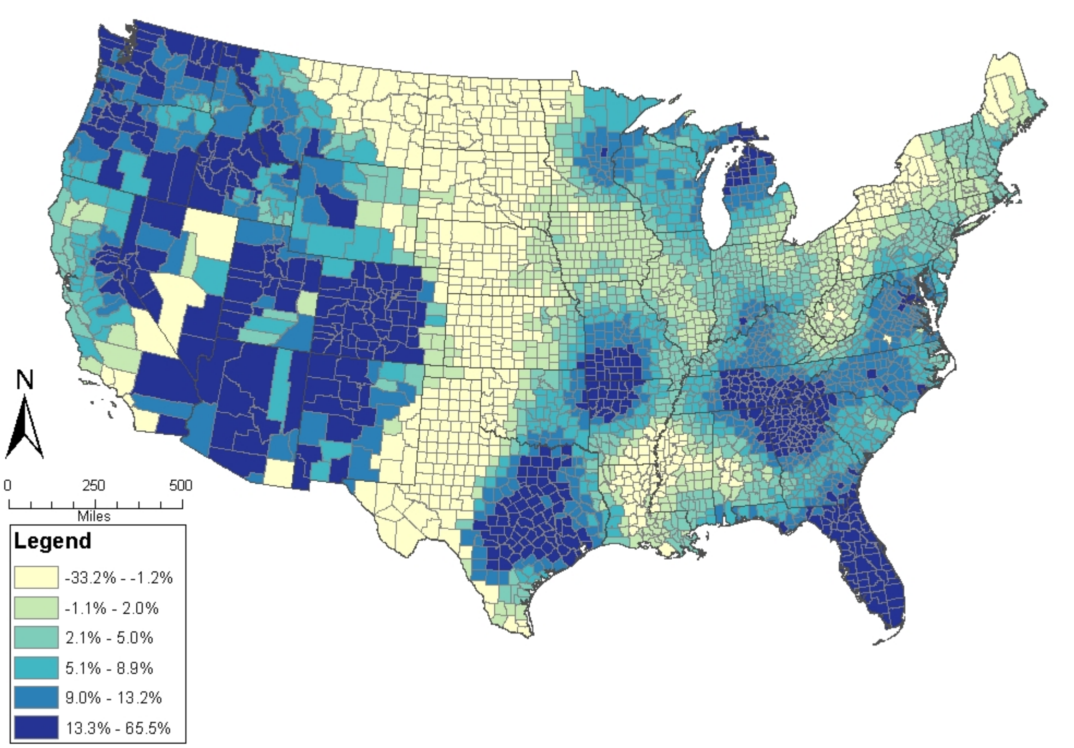
Hispanic natural increase rates lose power in the model, while white natural increase rates remain robust.

The trend surface variables also show an interesting pattern, such that the effect of western (X) and coastal (X^2) locations decreases but southern (Y) location remains significant. The spatial coefficients in each model are also very similar in their magnitude and level of significance. It is possible that the diffusion of net migration into neighboring counties is also contributing significantly to the autocorrelated portion of the error term. Put another way, it is not only people that migrate into neighboring counties, but also a set of variables attached to the people – income, education, family members, or social ties, for example. Combining the spatial lag and spatial error term in a single model creates a Simultaneous Autoregressive Moving Average (SARMA) model. The

SARMA model is a useful tool for understanding the total effect of the deconcentration process.

The fourth column of Table 4 summarizes the parameters of the full model (SARMA). Inspection of this model's fit statistics reveals that the full model is the only one which yielded residuals free of spatial autocorrelation. In other words, only the SARMA model could accurately predict the spatial distribution of net migration rates. It also exhibited a higher R^2, lower AIC, and lower SC, all of which indicate a much better model specification. The full model's coefficients for the three race/ethnic groups' natural increase rates is consistent with the pattern observed in the first two spatial models; however, the coefficients for the trend surfaces become erratic and difficult to interpret in the full model. The lack of explanatory power left in the trend surface

Figure 5. Full Model Predicted NMR Values



coefficients makes some sense because both Rho and Lambda are picking up spatial effects, which might leave little or no variance across regions in the net migration rates. On the other hand, it would have been convenient to find a model with both significant trend surface coefficients and significant spatial coefficients because such a model would lend evidence to both regionalization and deconcentration processes of population change. The full model seems to indicate strong support to the conception of net migration as a local deconcentration process which also creates unmeasured autocorrelation in neighboring counties. Figure 5 displays the predicted net migration rates from the fully specified spatial model, which can be compared to the actual rates displayed in Figure 1.

CONCLUSIONS

The analyses above offer an interesting, if limited, view into the relationships between the two components of population change: natural increase rates and net migration rates. One potential conclusion would be to discard natural increase rates and regionalization because of their relatively low explanatory power, and proclaim the dramatic effects of the spatial lag and spatial error seen in net migration rates. However, it should be noted that both Hispanic and white natural increase rates continue to be significant predictors of net migration rates even in the fully specified spatial model. Again this brings to mind the debate about what natural increase rates are indicating in the model, as was addressed above, but does suggest that where people are giving birth, where they are dying, and where they are moving is related in important ways. The results of this study also support the argument that race/ethnicity plays an important role in determining the residential sorting, net migration rates, and population growth rates of counties across the nation. Future research could attempt to disaggregate the relationships between natural increase rates and net migration rates by running models that only include one race/ethnic groups' rates for each county. For example, it would be possible to use only Hispanic natural increase rates to predict Hispanic net migration rates. This would create new problems with missing and unspecified populations because the analyses would be based on many counties with no Hispanic population in 1990 or 2000, and spatial analyses may be more sensitive to missing data problems.

In general, these analyses should lend strong support to the argument that the processes of population change – deconcentration, regionalization, and segregation – are spatially dependent. Future research into segregation and deconcentration will need to incorporate space in a more exact way than hierarchical categories along a rural to urban continuum. And although regionalization seemed to show many attractive ideas about population movement in the 1990s, the trend surface coefficients in the fully specified spatial model would suggest that large-scale determinants of migration are trumped by small-scale deconcentration. However, these two processes may be complementary instead of contradictory. For instance, when people move to the southern and coastal areas of the U.S., they may migrate at a higher proportion into suburban counties. Thus, while population distribution processes seem to operate on both the deconcentration and regionalization levels, the small-scale trends pick up some of the effect for the national trends. Building models that contain alternative measures of deconcentration, regionalization, and segregation in addition to the measures of natural increase rates used here, will result in a better understanding of the spatial patterns and trends in population distribution in the United States.

WORKS CITED

Adelman, Robert M., Chris Morett, and Stewart E. Tolnay. 2000. "Homeward Bound: The Return Migration of Southern-Born Black Women, 1940 to 1990." *Sociological Spectrum, 20*, 433-63.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic.

Beale, Calvin. 1975. *The Revival of Population Growth in Non-Metropolitan America*. Report 605. Economic Research Service, U.S. Department of Agriculuture.

Cliff, Andrew D. and J.K. Ord. 1981. *Spatial Processes: Models and Applications*. London: Pion.

Cook, Peggy J. and Karen L. Mizer. 1994. *The Revised ERS County Typology: An Overview*. Rural Development Research Report 89. Economic Research Service, U.S. Department of Agriculture. Available at: http://www.ers.usda.gov/publications/rdrr89/rdrr89.pdf.

Elliott, James R. 1997. "Cycles within the System: Metropolitanisation and Internal Migration in the US, 1965-90." *Urban Studies, 34*, 21-41.

Frey, William H. and Kenneth M. Johnson. 1998. "Concentrated Immigration, Restructuring and the 'Selective' Deconcentration of the United States Population." In Paul Boyle and Keith Halfacree (eds.), *Migration into Rural Areas: Theories and Issues*. New York: John Wiley & Sons, Inc.

Johnson Jr., James H. and Curtis C. Roseman. 1990. "Increasing Black Outmigration from Los Angeles: The Role of Household Dynamics and Kinship Systems." *Annals of the Association of American Geographers, 80*, 205-22.

Johnson, Kenneth M. 1993. "Demographic Changes in Nonmetropolitan America, 1980 to 1990." *Rural Sociology, 58*, 347-65.

Johnson, Kenneth M. and Glenn V. Fuguitt. 2000. "Continuity and Change in Rural Migration Patterns, 1950-1995." *Rural Sociology, 65*, 27-49.

Johnson, Kenneth M., Alfred Nucci, and Larry Long. 2005. "Population Trends in Metropolitan and Nonmetropolitan America: Selective Deconcentration and the Rural Rebound." *Population Research and Policy Review, 24*, 527-542.

Johnson, Kenneth M., Paul R. Voss, Roger B. Hammer, Glenn V. Fuguitt and Scott McNiven. 2005. "Temporal and Spatial Variation in Age-Specific Net Migration in the United States." *Demography, 42*, 791-812.

Logan, John R., Brian J. Stults, and Reynolds Farley. 2004. "Segregation of Minorities in the Metropolis: Two Decades of Change." *Demography, 41*, 1-22.

Manson, Gary A. and Richard E. Groop. 2000. "U.S. Intercounty Migration in the 1990s: People and Income Move Down the Urban Hierarchy." *Professional Geographer, 52*, 493-504.

Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, J. Edward Taylor. 1993. "Theories of International Migration: A Review and Appraisal." *Population and Development Review, 19*, 431-66.

Massey, Douglas S. and Nancy A. Denton. 1985. "Spatial Assimilation as a Socioeconomic Outcome." *American Sociological Review, 50*, 94-106.

Massey, Douglas S. and Nancy A. Denton. 1987. "Trends in the Residential Segregation of Blacks, Hispanics, and Asians: 1970-1980." *American Sociological Review, 52*, 802-825.

Massey, Douglas S. and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

McGranahan, David. 1999. *Natural Amenities Drive Rural Population Change*. Agriculture Economic Report Number 781, Economic Research Service, U.S. Department of Agriculture. Available at: http://www.ers.usda.gov/publications/aer781/aer781.pdf.

Newbold, K. Bruce. 1997. "Race and Primary, Return, and Onward Interstate Migration." *Profession Geographer, 49*, 1-14.

Noyelle, Thierry J. and Thomas M. Stanback, Jr. 1984. *The Economic Transformation of American Cities*. Totowa, NJ: Rowman & Allanheld.

Otterstrom, Samuel. 2001. "Trends in national and regional population concentration in the United States from 1790 to 1990: From the Frontier to the Urban Transformation." *The Social Science Journal, 38*, 393-407.

Plane, D. A., C. J. Henrie, and M. J. Perry. 2005. "Migration up and down the Urban Hierarchy and Across the Life Course." *Proceedings of the National Academy of Sciences, 12*, 15313-15318.

Tobler, Waldo. 1970. "A computer movie simulating urban growth in the Detroit Region." *Economic Geography, 46*, 234-240.

Voss, Paul R., Scott McNiven, Roger B. Hammer, Kenneth M. Johnson, & Glenn V. Fuguitt. 2004. *County-specific net migration by five-year age groups, Hispanic origin, race and sex 1990-2000*. Working Paper No. 2004-24, Center for Demography and Ecology, University of Wisconsin-Madison. Available at: http://www.ssc.wisc.edu/cde/cdewp/2004-24.pdf.

Voss, Paul R., Scott McNiven, Roger B. Hammer, Kenneth M. Johnson, and Glenn V. Fuguitt. COUNTY-SPECIFIC NET MIGRATION BY FIVE-YEAR AGE GROUPS, HISPANIC ORIGIN, RACE, AND SEX, 1990-2000: [UNITED STATES] [Computer file]. ICPSR4171-v1. Madison, WI: University of Wisconsin-Madison, Department of Rural Sociology [producer], 2003. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2005-05-23.