Mapping Poverty and Disability in Austin and San Antonio

Contents

Prepared for: GIS and Water Resources
By John Dinning
December 5, 2014

Introduction

Demographic data is highly useful to policy makers, government agencies, and nonprofits. It can help guide decisions on what types of human service projects are needed and where. And it can be used in the evaluation of social policies and programs.

This paper explores some recent demographic data on poverty and disability rates in Austin and San Antonio, Texas using ArcGIS. It compares several maps showing the distribution of poverty and disability. Projecting available information onto a map provides an information-rich picture of distributions and relative densities of poverty and disability that cannot be communicated through simple, city-wide statistical summaries.

This report will also demonstrate how GIS can be a used as a powerful tool for post-estimation analysis. As an example, we will look consider spatial post-estimation analysis of a linear model predicting full-time work among the disabled using covariates available in the American Community Survey dataset.

Data: The American Community Survey

The analysis prepared for this report was conducted using 5 year data from the American Community Survey (ACS) covering the years 2008-2012. This dataset represents combined observations from five-year rolling samples that permit estimates at the block-group level. This provides a high degree of "resolution" that cannot be obtained in the one-year or three-year ACS datasets.[i]

There are over 15,000 block groups in the state of Texas. The block groups serve as the unit of analysis in this report. It is important to highlight the fact that the ACS is a survey and the variables in the ACS dataset are point estimates for the number of individuals in each block group meeting the specified criteria. These estimates are reported along with margins of error representing 90% confidence intervals.

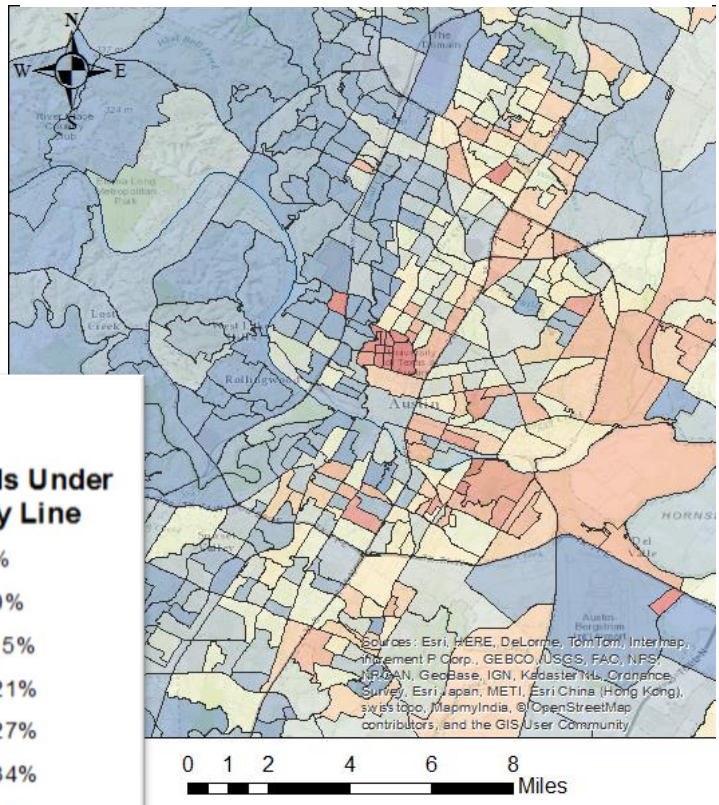
Percentage of Households under the Poverty Line

The following maps illustrate the percentage of households under the federal poverty line for different block groups in Austin and San Antonio. Austin has higher rates of poverty on the East

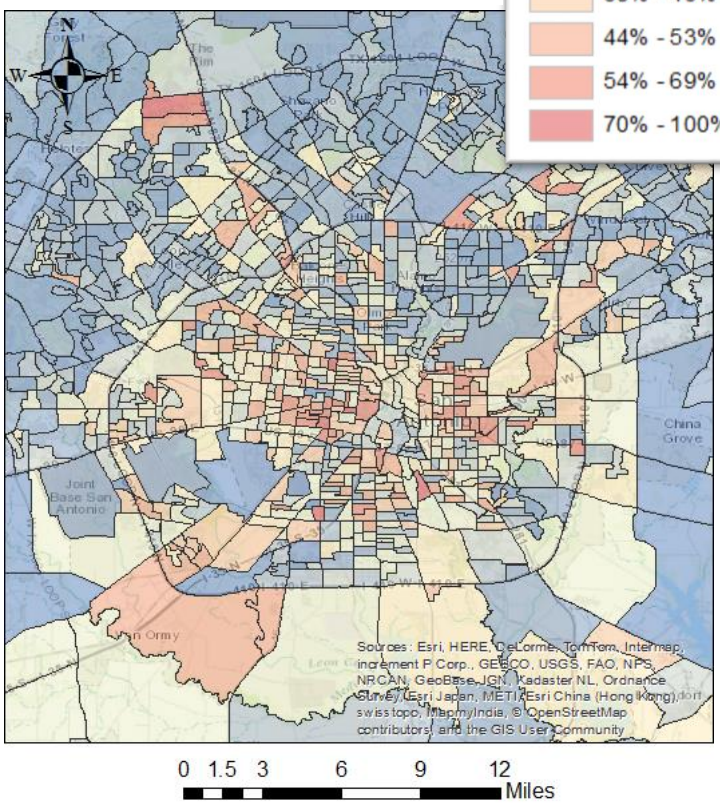
side. There is also a neighborhood with very high rates of household poverty that corresponds to the West Campus area of The University of Texas. These block groups have a very high percentage of university students with little or no income.

In San Antonio, block groups with higher rates of household poverty are more widely distributed throughout the city with the Northeastern part of the city center being a notable exception as can be seen in the map below.

Percentage of Households under the Poverty Line by Block Group in Austin



Percentage of Households under the Poverty Line by Block Group in San Antonio

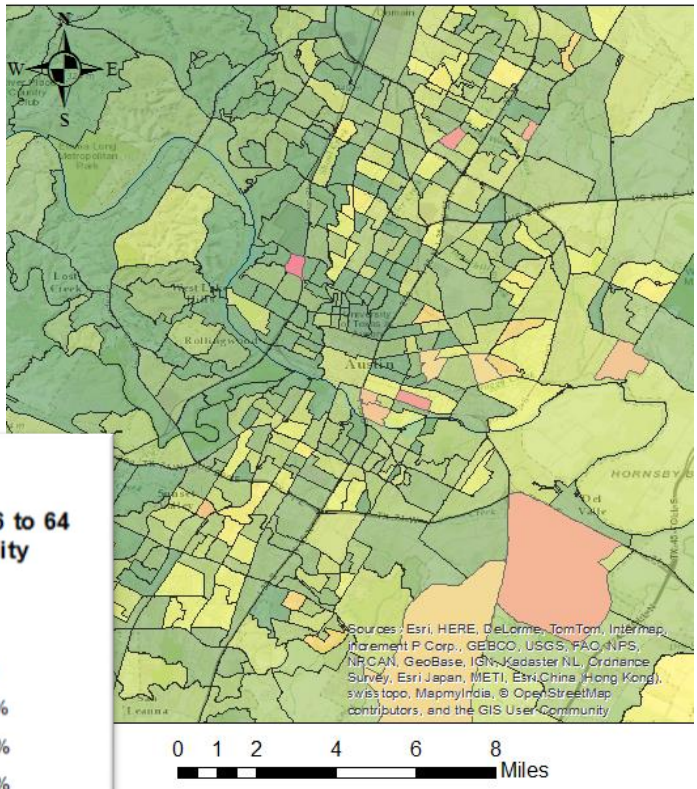


Maps of household poverty can be useful for local organizations trying to understand where poverty is concentrated in their city. Such maps can inform outreach efforts. They can also help useful for planning, monitoring, and evaluating interventions targeted at low income groups.

Percentage of Individuals between the Ages of 16 and 64 with a Disability

The following two maps show striking differences in disability rates in Austin and San Antonio. In San Antonio, nearly 23% of working age individuals reported having a disability according to the 2010 census[ii] compared to only 15% in Austin.[iii] But such statistics provide no information on the spatial distribution or concentrations if disability in these cities. The map of Austin shows that disability rates, while not high in many places, tend to be somewhat higher farther out to the North and to the South from the urban core. Whereas in San Antonio, there are multiple block groups with very high rates of disability clustered around the city center. It is worth noting that in San Antonio there are very low rates of disability in the Northeast corner of the city center which also corresponds to lower rates of household poverty. The same can be said of West Austin.

Percentage of Individuals between the Ages of 16 and 64 with a Disability in Austin



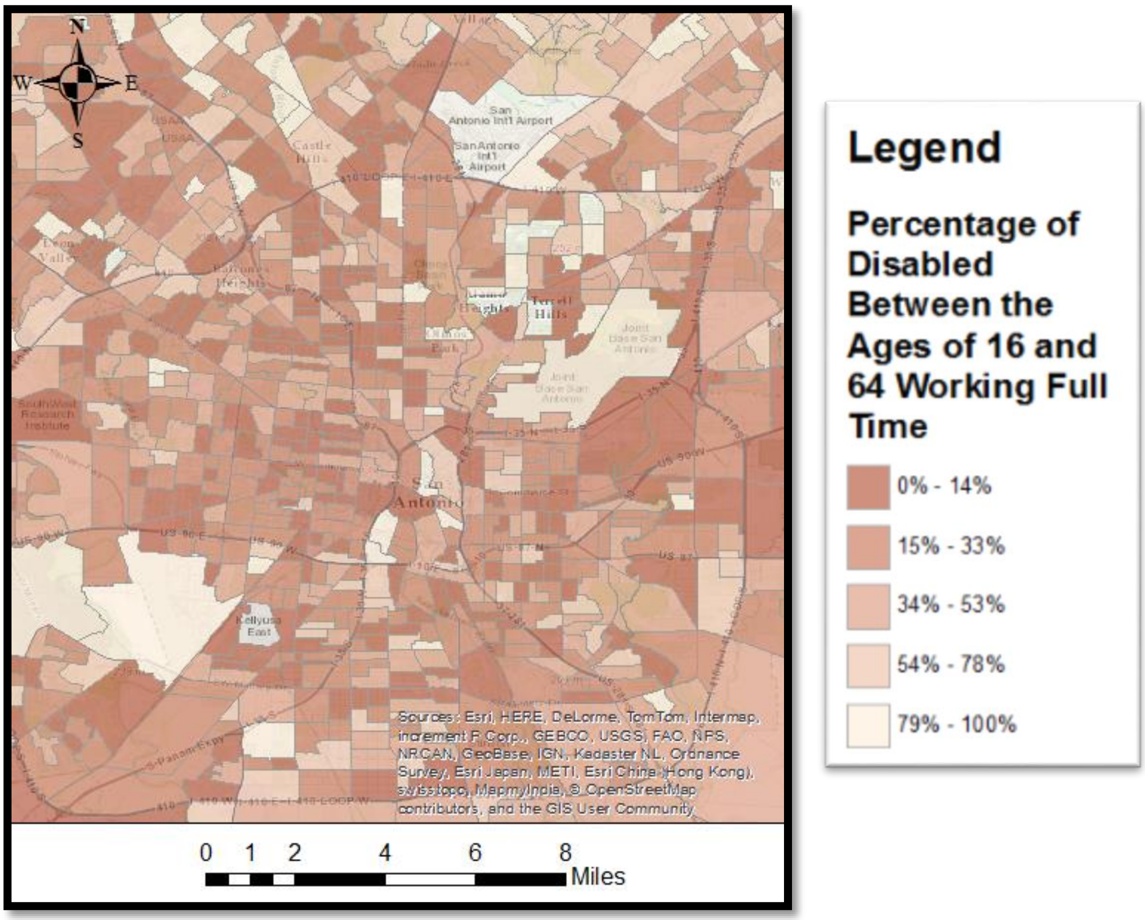
Percentage of Individuals between the Ages of 16 and 64 with a Disability in San Antonio

Maps of disability rates can assist planners at state-level agencies such as the Texas Department of Aging and Disability Services to analyze the distribution and identify clusters of high incidence of disability in Texas cities.

Rates of Full-Time Work among the Disabled in San Antonio

One of the difficulties of working with ACS data occurs when dealing with a very small subset of the population. The following map shows the percentage of full-time workers among the subset of the population that is between 16 and 64 years old and has a disability.

Full-Time Workers among the Disabled Age 16 to 64 in San Antonio

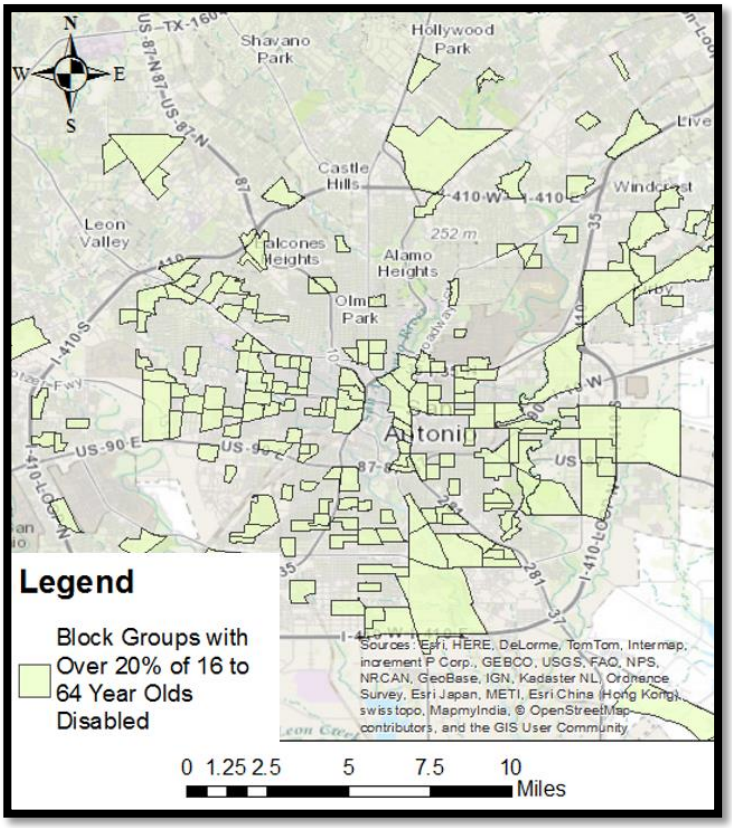


The problem with this map is that the information, as presented, may appear somewhat precise but is, in fact, uncertain for the majority of the block groups. The population of individuals between 16 and 64 with a disability and full-time employment is a very small subset of the entire population. In many of the block groups, the 90% confidence intervals reported in the ACS include zero and the point estimates in the dataset are very unreliable for those block groups where concentrations of individuals with a disability are low. For that reason, the following analyses of full-time employment among the disabled only include data from those block groups with disability rates greater than or equal to 20%. These block groups had smaller confidence intervals for the variable in question and more reliable point estimates.

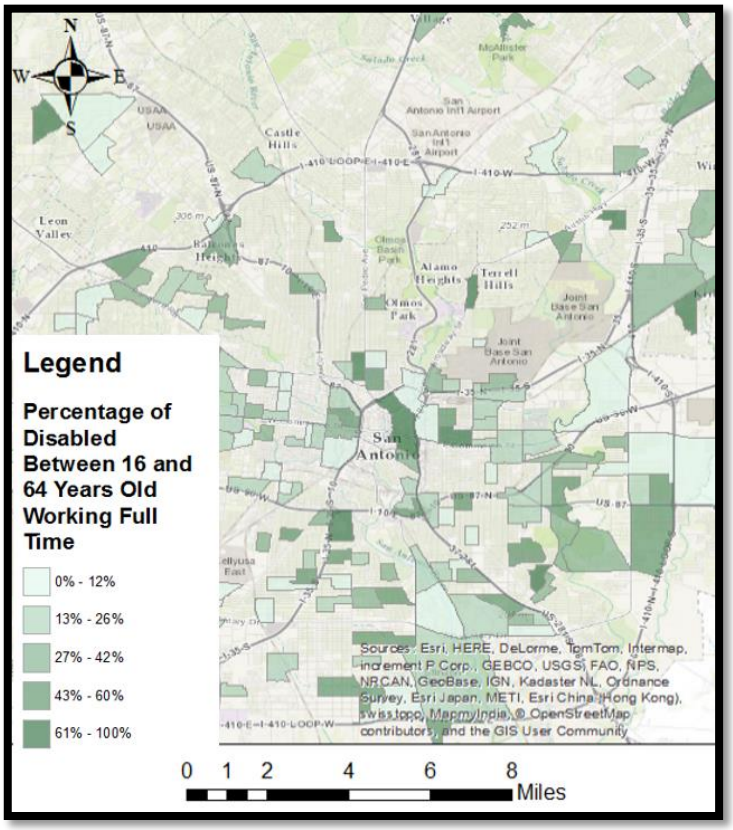
The following section will consider a subset of block groups with more reliable estimates of full-time employment among the disabled. For this analysis, block groups with disability rates over 20% will be categorized as "high-disability" block groups. These high-disability block groups are shown in the map on the right.

High-disability block groups have more observations of employment rates among the disabled and consequently, much more precise estimates. These rates can be observed in the map below.

High Disability Block Groups in San Antonio



Rate of Full-Time Employment among the Disabled Age 16 to 64 in High Disability Block Groups in San Antonio



The ACS data do not permit a reliable analysis of full-time employment among the disabled for all block groups in Texas. But despite this limitation, an analysis of just the high-disability subset can very useful. Drilling down into this subset of the data can provide useful information on the spatial distribution of varying rates of full-time employment among individuals between 16 and 64 with a disability.

Predictors of Disability and Predictors of Employment among the Disabled

The ACS includes a great deal of demographic data at the block group level that cannot be found in the ten-year US Census. Nonprofits, municipal human service departments, and similar agencies and organizations can use this data to analyze where services are needed, and to explore ways to address poverty and unemployment.

The point estimates in the ACS dataset were used to generate proportions at the block group level. The regression models in this report rely on the point estimates which are reported in the ACS along with 90% confidence intervals. These models were fit directly to the ACS point estimate data and the confidence intervals were not accounted for in the models. Due to the large number of block groups observed, this should not have a significant effect on the results. However, it is important to note that due to the nature of the data, the t statistics reported in the following tables are likely to be slightly overstated.

Predictors of Disability

There are various demographic characteristics that tend to covary with disability. Examples of covariates available in the ACS data are race, age, veteran status and poverty status. Using block group data from the state of Texas, we can regress the disability rates on the available covariates. The table on the following page shows the results of a log-linear model with the log-odds of the proportion of disability regressed on the percentages of individuals in different race, age, poverty and veteran status groups. These observed characteristics, available in the ACS, can predict some of the variance in the percentage of individuals with disabilities in each block group. Note that the dependent variable, proportion of disabled, has been transformed from a variable bounded between zero and one to a continuous variable by converting it to log-odds (lnOddsPDis).[1]

The reference categories for the following models are percentages of whites, non-veterans, minors, and households above the poverty line.[2] The model captured about a quarter of the variance in disability rates. As would be expected, higher percentages of older adults predicted higher rates of disability. As the percentage of individuals in a block group identifying as Asian

[1] Log odds transformation of the proportion (p) requires converting it to odds and taking the natural logarithm as follows: **ln[p/(1-p)]**

[2] The race categories used in this analysis are neither comprehensive nor orthogonal. In the ACS, many individuals identify as two or more races. For example, many individuals identifying as white also identify as Hispanic.

increases, the expected rates of disability decrease. The percentage of Hispanics, African Americans, and veterans is positively correlated with disability rates. The most significant predictor in the model was poverty. Poverty has been shown in numerous studies to be positively correlated with disability.[iv] The observations in this dataset are no exception.

	lnOddsPDis
Percent of households with income below the poverty line	0.016 (31.57)**
Percent of veterans in the pop. of 18 to 64 years olds	0.018 (12.93)**
Percent Hispanic	0.340 (12.54)**
Percent African American	0.661 (18.78)**
Percent Asian	-2.549 (23.46)**
Percent Adults Age 18 to 44	-0.000 (0.25)
Percent Adults Age 45 to 64	0.017 (16.33)**
Percent Adults Age 65 and over	0.020 (20.22)**
Constant	-3.501 (53.29)**
R^2	0.25
N	15,265

* $p<0.05$; ** $p<0.01$; t statistic in parenthesis

Predictors of Full Time Work Among the Disabled Age 16 to 64 in High Disability Block Groups

Similarly, organizations and agencies working to promote employment among the disabled would be interested in knowing predictors or risk factors that affect the rates at which the disabled tend to join the workforce.

The following table represents the log-odds of full time work among the disabled age 16 to 64 in high disability block groups in the State of Texas (lnOddsDisFTwork) regressed on percentages of age, race, veteran and poverty status.

	lnOddsDisFTwork
Percent of households with income below the poverty line	-0.017 (10.25)**
Percent of veterans in the pop. of 18 to 64 years olds	0.006 (1.47)
Percent Hispanic	-0.039 (0.47)
Percent African American	-0.435 (4.15)**
Percent Asian	2.600 (3.13)**
Percent Adults Age 18 to 44	-0.007 (1.74)
Percent Adults Age 45 to 64	-0.015 (3.98)**
Percent Adults Age 65 and over	-0.008 (2.26)*
Constant	0.271
	(1.08)
R^2	0.10
N	1,840

* $p<0.05$; ** $p<0.01$; t statistic in parenthesis

The results of the model show that percentages of certain racial or age groups can serve as weakly significant predictors of full time employment rates among the disabled age 16 to 64. However, for high-disability block groups in Texas, neither percentages of Hispanics nor of veterans are significant predictors of full time employment among the working-age disabled. As in the previous model, percentage of households in poverty was the most significant predictor of rates of full-time employment. Household poverty rates are negatively correlated meaning that, as poverty rates rise, rates of full time work go down.[3] However, the magnitude of the correlation is relatively small.

It is important to note that only 10% of the variance in full-time work rates among the disabled age 16 to 64 in high-disability block groups in Texas was captured by the model. As a result, the above model has only limited predictive power. A spatial evaluation of the residuals can provide important clues as to how the model can be improved.

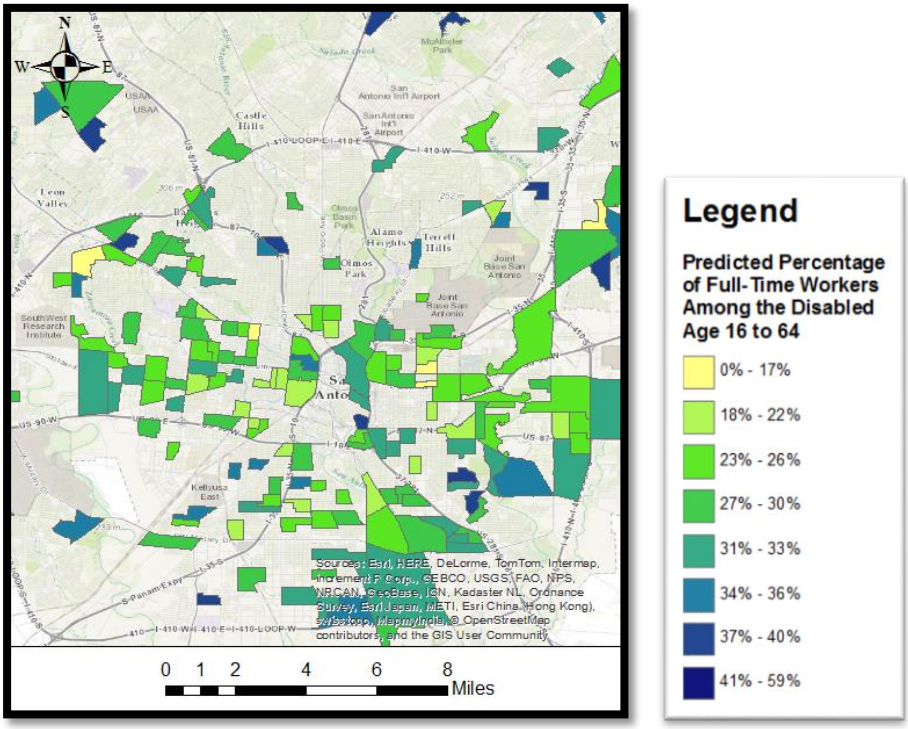
[3] This analysis provides no indication of causal direction between rates of poverty rates and full-time employment among the disabled.

Spatial Evaluation of Statistical Models

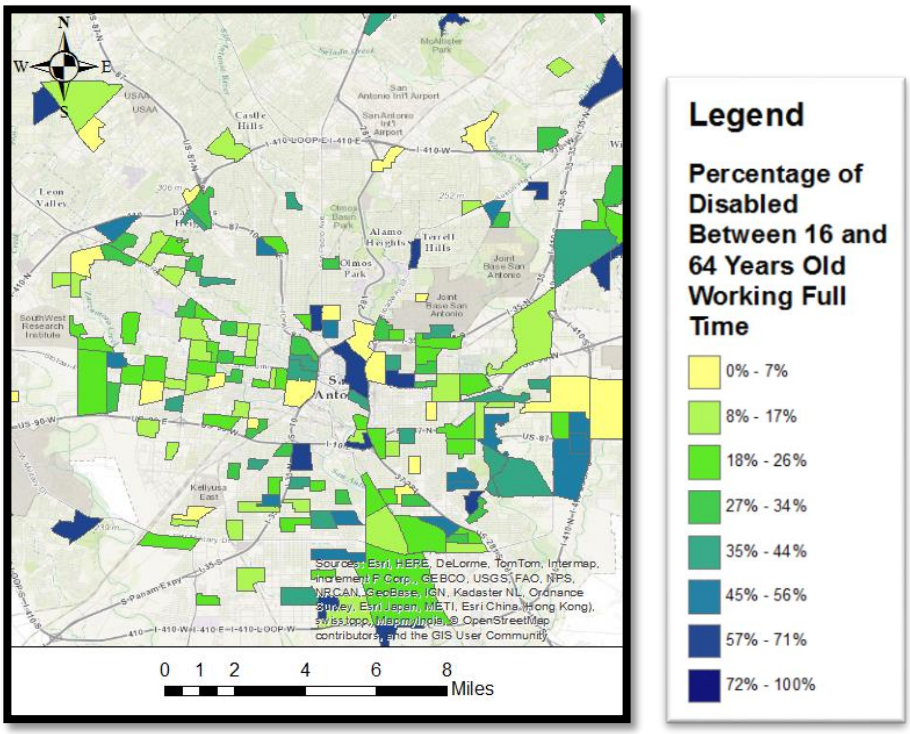
The models described above are rudimentary and limited by the available data. There are certainly a great many omitted variables that could better predict rates of full-time work among the disabled. These omitted variables are correlated with the residual which likely includes systematic variance in full-time employment rates among the disabled in Texas that was not explained by the model.

The residual is the difference between the predicted and the observed rates of the dependent variable. That is, it is the difference between to two maps on the right.

Predicted Percentages of Full-Time Workers among the Disabled Age 16 to 64 for High-Disability Block Groups in San Antonio



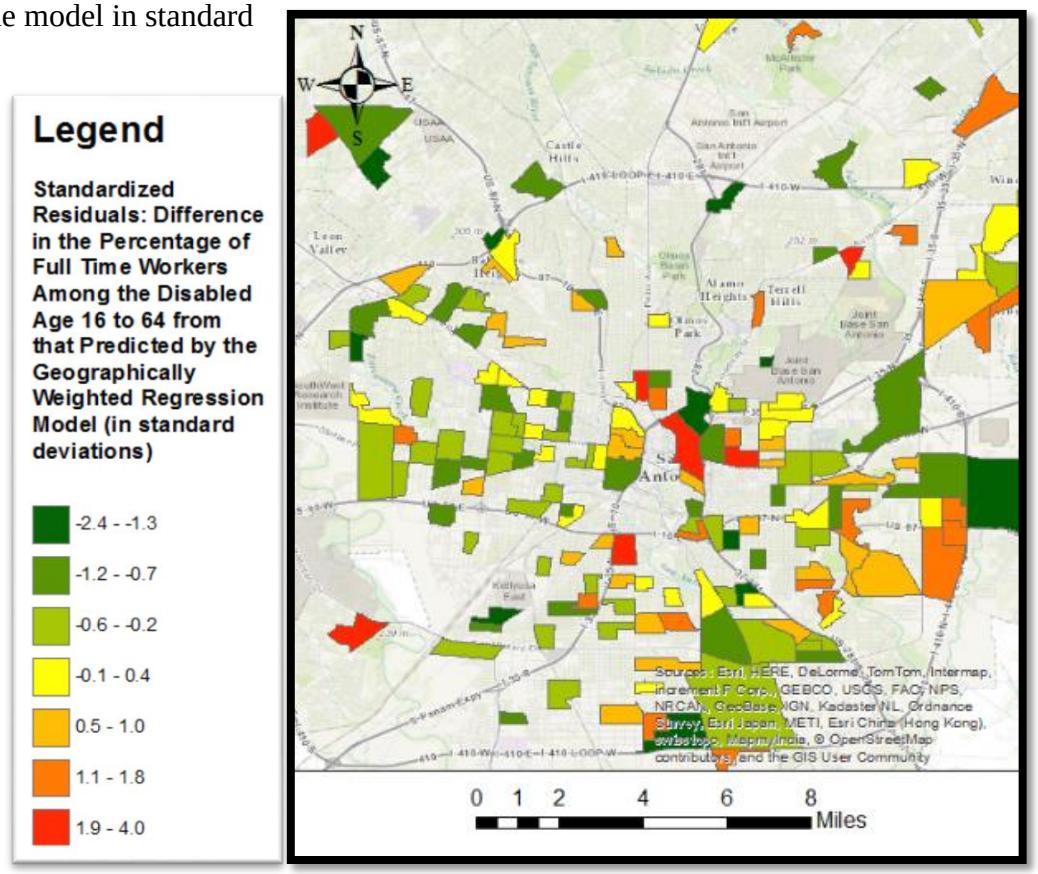
Actual Percentages Reported in the ACS of Full-Time Workers among the Disabled Age 16 to 64 for High-Disability Block Groups in San Antonio



Spatial Post-Estimation Analysis and Exploration of Omitted Variable Bias

Block group data from the ACS lends itself well to spatial post-estimation analysis. This map shows the residuals from the model in standard deviations. Areas in yellow were predicted with relative accuracy. Areas in green had lower rates of full-time employment than expected and areas in orange and red had higher rates than expected. By looking at the variation in full-time employment among the disabled *controlling for the observed variables*, it is possible to develop and test theories for what is contributing to the unexplained variation and to identify the omitted variables that should be included in the model. The following are some examples of omitted variables with important spatial considerations:

Differences in the Rates of Full-Time Employment among the Disabled Age 16 to 64 in High Disability Block Groups in San Antonio Between that Predicted by the Model and that Reported by the ACS



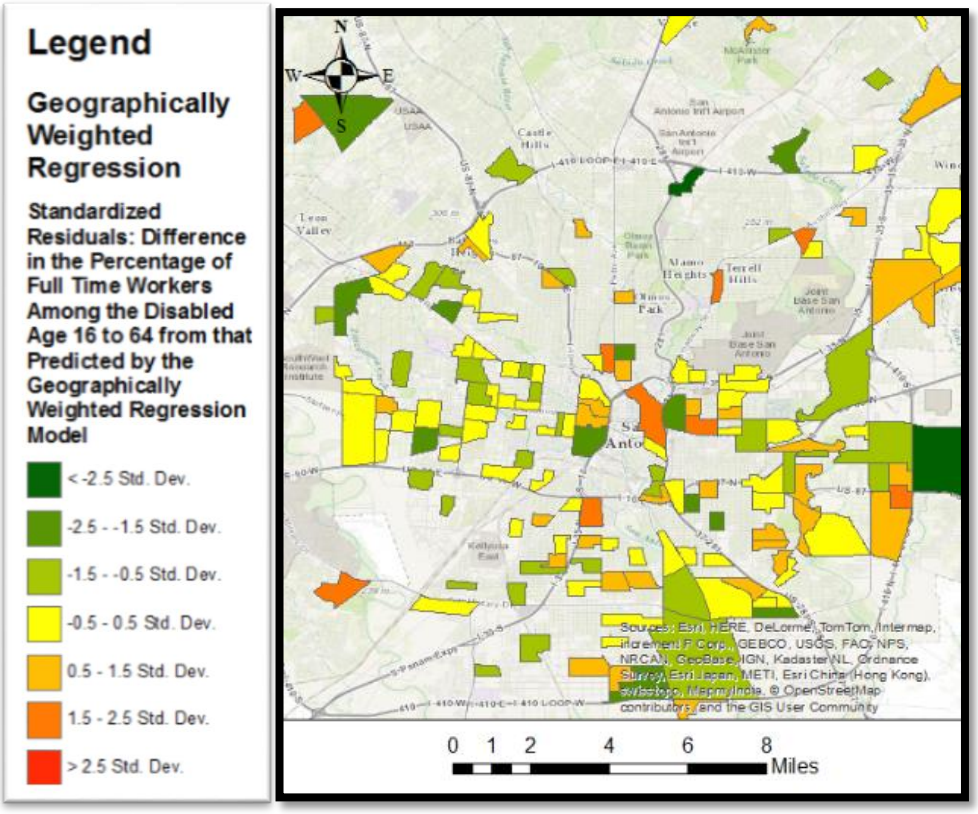
- Proximity to centers offering specialized support services to the disabled.
- Accessibility of public transportation.
- Areas where nonprofits or government agencies have focused outreached efforts or programs designed to help the disabled find work.
- Proximity to sources of employment.

Comparison with the Geographically Weighted Regression Feature in ArcGIS

The data for the residuals mapped on the previous page were derived from the log-linear model using the formula: **Residual = (Observed Values) – (Predicted Values)**

The ArcGIS software has a Geographically Weighted Regression feature which permits a similar analysis but takes into account the proximity of other block groups.

Standardized Residuals from the Model of Full-Time Employment among the Disabled Age 16 to 64 in High Disability Block Groups in San Antonio Using Geographically Weighted Regression in ArcGIS



The results of the geographically weighted regression are quite similar to that from the basic linear model. The R Squared statistic shown in the table below indicates that 10% of the overall variance in rates of full-time work among the disabled in high disability block groups is explained by the geographically weighted model. This is the same amount of variance that was captured in the non-geographically weighted linear model. This indicates that there may no additional benefit from using geographic weighting in this case. The lack of improvement on the model may be related to the way in which the high-disability block groups evaluated in this model are distributed throughout the state of Texas.

VARNAME	VARIABLE	DEFINITION
Bandwidth	2235259.195237	
ResidualSquares	1707.241249	
EffectiveNumber	21.534385	
Sigma	0.968936	
AICc	5123.62867	
R2	0.101504	
R2Adjusted	0.091358	
Dependent Field	0	lnOddsDisFTwork
Explanatory Field	1	percHHPov
Explanatory Field	2	percVets
Explanatory Field	3	percHisp
Explanatory Field	4	percAfroAm
Explanatory Field	5	percAsian
Explanatory Field	6	percAdult45to64
Explanatory Field	7	percAdult18to44
Explanatory Field	8	percAdult65plus

Final Considerations

The models considered in this report certainly suffer from omitted variable bias. For example, an important omitted variable is the nature of the local economy. Different cities in Texas have very different economies which could reasonably be expected to play a role in employment.

This could be evaluated using a multilevel model that accounts for the characteristics of specific economic ecosystem in which each block group exists. In this case, which the level-one units would be block groups and the level-two units would be metropolitan statistical areas. This could help to capture some of the variance in work rates among the disabled that changes systematically from city to city depending on factors such as the local economy, availability of city-wide support services, accessible public transportation, etc.

When evaluating models for demographic data, omitted variables could potentially be revealed through an investigation of the spatial component of the residuals. Spatial post-estimation can be very useful to social science researchers interested in exploring omitted variables and improving on existing models. And on a more practical level, they could be used by service providers and policy makers seeking to provide better services and to monitor, evaluate, and improve existing programs.

Notes

[i] U.S. Census Bureau, "A Compass for Understanding and Using American Community Survey Data," October 2008.
[ii] "San Antonio, Texas, Social Statistics," accessed December 5, 2014, http://www.infoplease.com/us/census/data/texas/san-antonio/social.html.
[iii] "Austin, Texas, Social Statistics," accessed December 5, 2014, http://www.infoplease.com/us/census/data/texas/austin/social.html.
[iv] Meredith Minkler, Esme Fuller-Thomson, and Jack M. Guralnik, "Gradient of Disability across the Socioeconomic Spectrum in the United States," *New England Journal of Medicine* 355, no. 7 (August 17, 2006): 695–703, doi:10.1056/NEJMsa044316.