The Rail-ity of Gentrification: Exploring Light Rail Transit and Neighborhood Socioeconomic Status

GIS Project by Nathan Rochester, Jennifer Schofield, and Clement Uduk

GEOG 492/592: GIS II- Advanced GIS Applications

Through GIS and OLS modeling, this project set out to investigate patterns of neighborhood change across the Portland Metropolitan Area (PMA) from 1980 to 2010 in relation to light rail transit (LRT) development that took place throughout that period. The main focus was on a measure of socioeconomic status (SES) and SES change within census tracts (CTs) in the PMA based on census data from the Neighborhood Change Database (NCDB: Geolytics 2013). Using this longitudinal census dataset, SES is operationalized as the sum of three percentages that were identified in the literature as highly reliable SES indicators.

Definitions

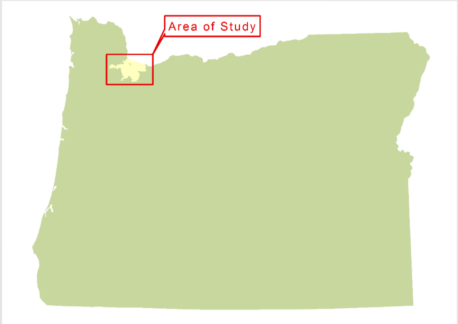
Gentrification at its core is a dense and convoluted concept that has multiple definitions. Within the scope of this particular project, gentrification is defined as neighborhood demographic change associated with the influx of higher SES residents into relatively low SES neighborhoods, the displacement and "pricing out" of the original lower SES residents, and the resulting overall *increase in average SES for the neighborhood*.

SES is a concept that incorporates social and economic factors of class. For the scope of this project, we have operationalized SES as the combination of occupation type, educational attainment and family income.

The LRT lines that were the focus of the project to investigate the relationship between LRT development and demographic change were, in chronological order, Eastside Blue (opened in 1986), Westside Blue (opened in 1998), Red (opened in 2001) and then Yellow (opened in 2004).CT access to LRT is the independent variable that was operationalized as the street network distance from the central housing point in each CT to the nearest station along the each LRT line.

Study Area

305 CTs were selected based on the criteria that the weighted mean center of housing points they contain, as well as at least 20% of their geographic areas, be within the urban growth boundary (UGB), Metro government jurisdiction, and the TriMet service area according to their respective boundaries in 2010. Multnomah County Census Tract 9800 is designated as non-residential and therefore is not included in the sample. 2010 housing points were compiled with records from the Census Bureau's 2010 master address file (MAF) that are located within 2010 taxlots with single-family or multifamily residential landuse designations, combined with records from the 2010 version of the multifamily housing inventory (MFHI). MAF, taxlots, and MFHI datasets were all accessed from Portland State University's RLIS archives.



Demographic Data

SES Index is our dependent variable that was used to operationalize gentrification.

The measure of SES is the sum of three percentage variables pertaining to neighborhood educational attainment, occupation type and family income:

- Educational attainment is defined as the percent of persons 25 and older that have completed a bachelor's degree.
- Occupation is defined as the percent of employed persons 16 and older that have professional, technical or executive occupations.
- Income is defined as the percentage of families in an above average income category calculated by using income ranges above the median family income per census year. For example, in 1980, the median and mean income were $24,386 and $25,488, respectively. The percentage was calculated as the number of families in income categories above $25,000 divided by total number of families.

The Hot Spot Analysis Tool was used to identify statistically significant clustering of CTs with high values ("hot spots") or low values ("cold spots") for the SES index and SES change variables. Figures 1-4 depict hot spot analyses of the SES index for years 1980 through 1990, and Figures 5, 6, and 7 depict hot spot analyses for SES change from 1980-1990, 1990-2000, and 2000-2010, respectively. In interpreting hot spot maps, darker shades of red represent greater statistical significance (i.e., lower p-values) of high clustering, and darker shades of blue indicate greater statistical significance of low clustering.

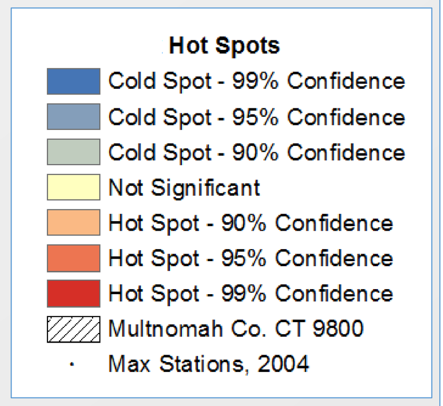


Figure 1: 1980 SES Index Hot Spots

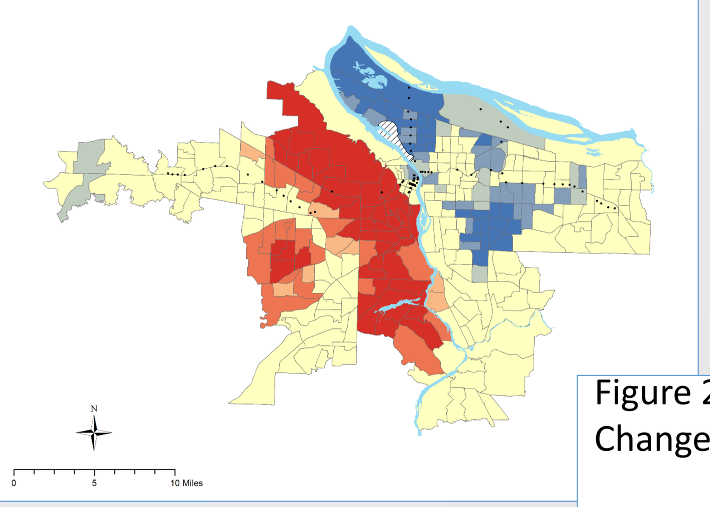


Figure 1 depicts a clear division of high and low SES index clustering at the baseline time point. Statistically significant clustering of CTs with high SES index values are found exclusively in areas northwest, west and southwest of downtown; low SES clustering is predominately located in North and East Portland, with consistent low clustering along the I-205 corridor and extending into the suburb of Gresham. Two low cluster CTs are found downtown west of the Willamette River.

Figure 2: 1980 to 1990 SES Index Change Hot Spots

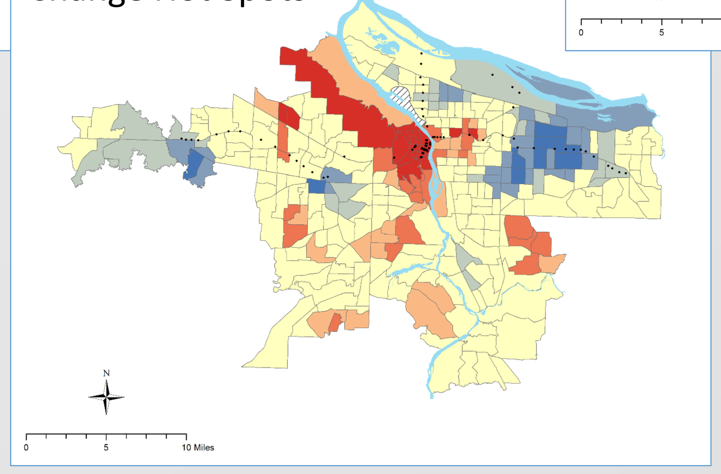


Figure 2 reveals that a large portion of the areas of high SES index clustering in 1980 did not have statistically significant clustering of SES change in any way from 1980 to 1990. High clustering of SES change for this period occurred largely in west Portland and immediately downtown, and the inner east side also saw significant clusters of high SES change, specifically along the newly constructed Blue Line LRT.

Figure 3: 1990 SES Index Hot Spots

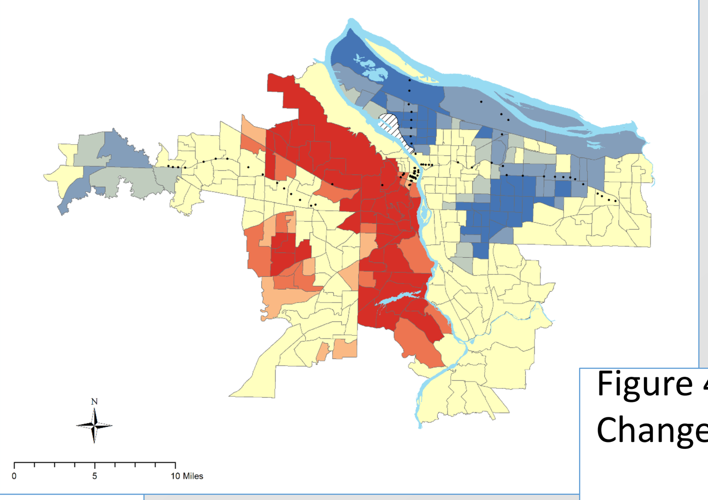


Figure 3 depicts similar clustering as to what was seen in 1980, but with some important changes. There appears to be a reduction in the clustering of high SES CTs on the west side, and an increase in the clustering of low SES CTs on the eastside, particularly in the easternmost suburb of Gresham. Additionally, low SES clustering develops in the westernmost suburb of Forest Grove. Note the absence of the low SES cluster downtown west of the Willamette River that was apparent at baseline..

Figure 4: 1990 to 2000 SES Index Change Hot Spots

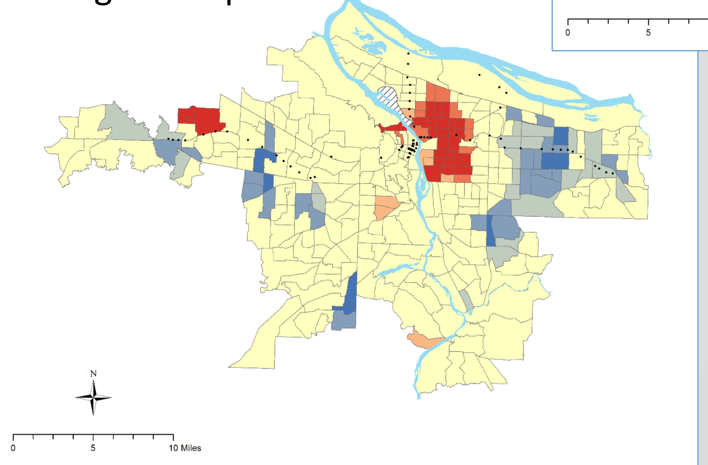


Figure 4 shows that the inner eastside of Portland saw significant high clustering of SES change from 1990 to 2000 , specifically along the newly constructed Blue Line LRT. This is similar to the change from 1980 to 1990. However, significant increase in SES change has grown into northeast and southeast. Low SES change clustering continued in east suburbs.

Figure 5: 2000 SES Index Hot Spots

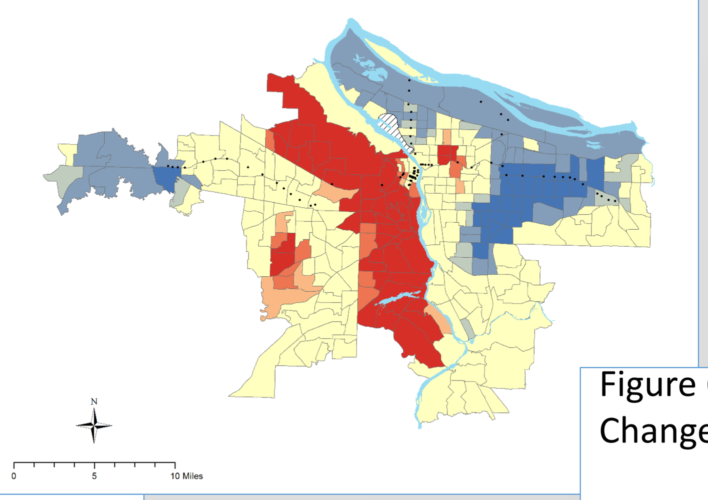


Figure 5 depicts continued decrease of west side high SES clustering as those patterns shifted closer to downtown. Low SES clustering occurred in the west and east suburbs. Additionally North Portland low SES begins to break up and lose significance in many CTs, and the emergence of high clustering begins in the central eastside.

Figure 6: 2000 to 2010 SES Index Change Hot Spots

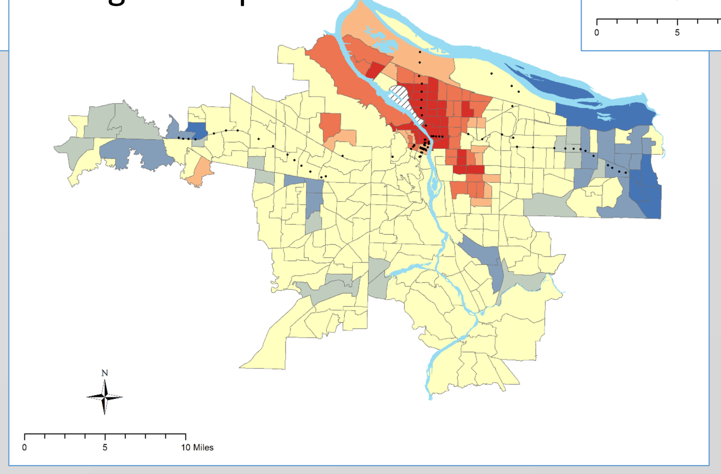


Figure 6 depicts significant high clustering of SES change in north and inner-northeast Portland, specifically along the Yellow Line LRT (which was constructed in 2004) between 2000-2010. Additional significant high SES clusters occur northwest of downtown Portland.

Figure 7: 2006-10 SES Index Hot Spots

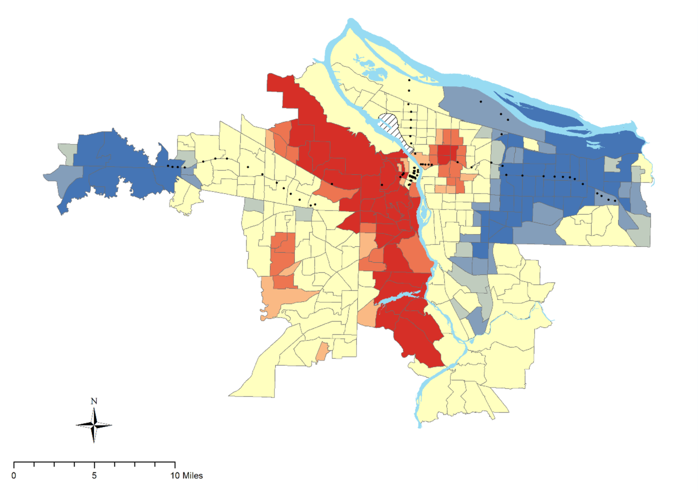


Figure 7 indicates that clustering of low SES index values has disappeared entirely from North Portland along the Interstate corridor and the recently opened Yellow Line, while becoming more intense and wider spread in the farthest east and west suburbs. Additionally, two CTs on the west side, one with close proximity to LRT stations, gained statistical significance for low clustering. Meanwhile, high clustering on the west side shifted further into Portland, reaching downtown CTs and spreading further into the central east side of the region.

OLS Linear Regression

The OLS regression detailed below predicts 2000 to 2006-10 SES change using 7 explanatory variables. Three variables measure street network distance in miles from CT central housing features to the nearest LRT stations along different lines, and the remaining four variables, based on data from the 2000 census, control for various CT demographic factors.

All explanatory variables are statistically significant with 95% or greater confidence. Based on the listed coefficients, this model predicts that for every additional mile of network distance between a CT's central housing feature and the nearest Yellow Line station, there is an estimated decrease of .016 in its SES index value between 2000 and 2010, holding all controls constant. In other words, all else being equal, a CT that is 7 miles from a Yellow Line station (sample mean = 7.7) is predicted to decrease in SES index by .112 points during that period (sample mean = .09).

Change in SES Index from 2000 to 2010

Variable	Coefficient
Intercept	0.018868
Distance to Nearest MAX Station in 1998	0.008426*
Distance to Nearest New Red Line MAX Station (2001 and later)	0.006764**
Distance to Nearest New Yellow Line MAX Station (2004 and later)	-0.015999***
2000 Percent Non-Hispanic Black	0.003846**
2000 Percent of Housing Units that are Single Family Detached	0.001938***
2000 Percent of Workers Taking Transit to Work	0.009969***
2000 Median Rent ($100 units)	-0.014673*
Koenker Statistic:	26.441345***
Joint Wald Statistic:	146.526255***
Jarque-Bera Statistic:	25.119913***
AICc:	-219.372446
Adj. R-Squared:	0.303556

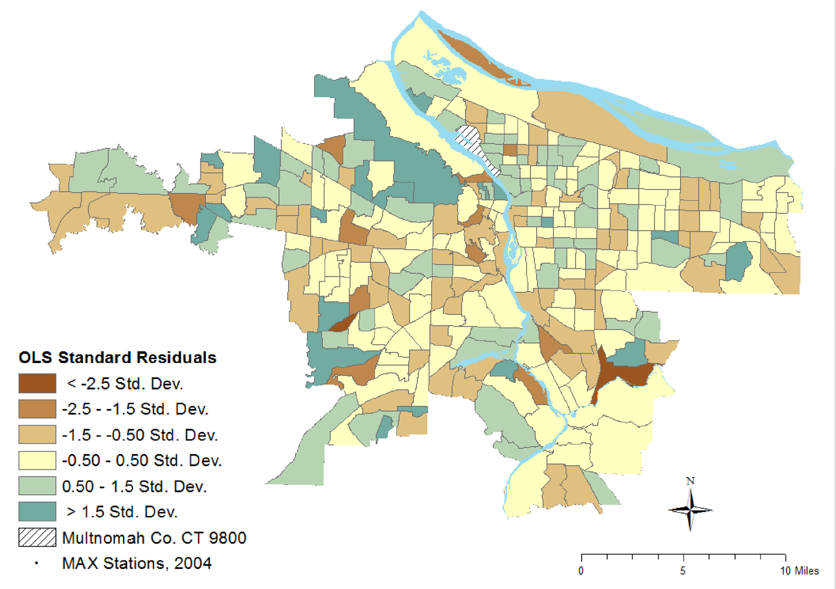
*p < .05
**p < .01
***p < .001
Note: Due to significant Koenker statistic, coefficient p-values reflect robust probabilities.

The Joint Wald statistic indicates overall model significance at p < .001, corresponding to 99.9% confidence that it reflects real relationships and is not simply a product of chance; and according to the adjusted r-squared, this model accounts for about 33.4% of variation in SES change in this period.

However, Koenker and Jarque-Bera statistic p-values (both < .001) indicate heteroscedastic or nonstationary residuals, and non-normally distributed residuals, respectively.

Global Moran's I was run on the residuals and confirmed that no spatial autocorrelation exists in the distribution of residuals (z-score = 0.961022).

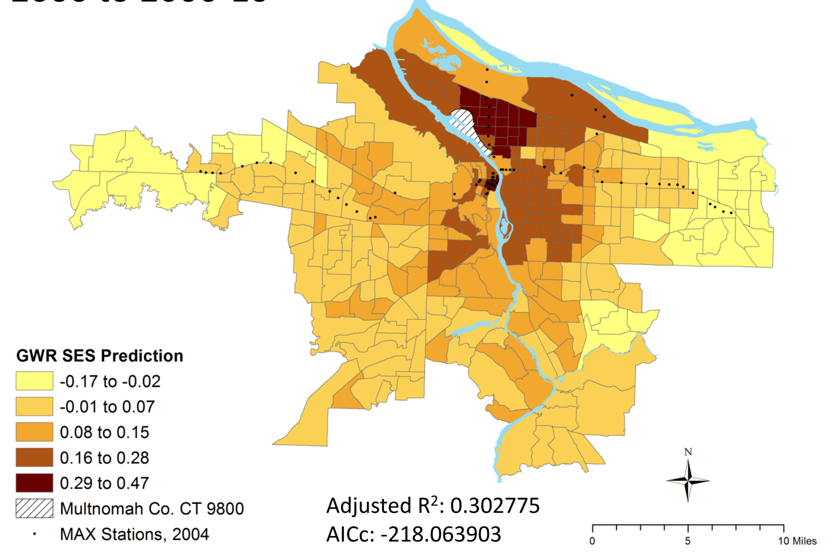
Figure 8: Spatial Distribution of OLS Residuals



The same variables used in the global OLS model were input to the Geographically Weighted Regression (GWR) Tool to produce localized coefficients and prediction values.

The highest predicted values of SES change are in North Portland and one downtown CT corresponding approximately to the Pearl District. Predicted low values of SES change similarly reflect those observed in the hot spot analyses above.

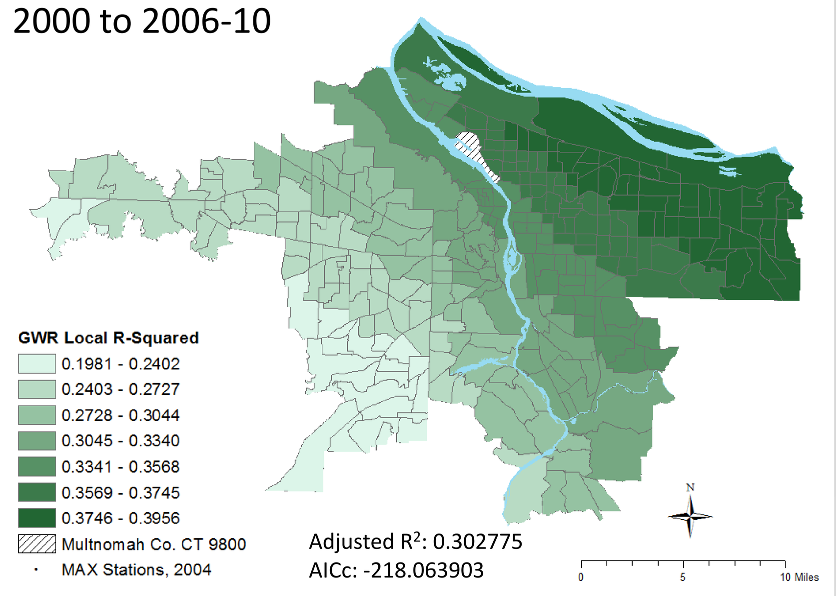
Figure 9: GWR Localized Predictions of SES Change: 2000 to 2006-10



Although the GWR adjusted r-squared does not improve on the *global* predictive power of the OLS model, it does provide *local* r-squared values that reveal wide variations in the predictive power of the model. The model explains 35.7% to 37.5% of variation in North Portland around the Yellow Line, and up to 39.6% of variation in areas further east and in Gresham.

Predictive power declines rapidly in areas west of Portland, with a minimum local r-squared of 0.20 to 0.24 in the easternmost reaches of Forest Grove and the southwest Metro suburbs. This suggests the model might be improved upon with the addition of data pertaining to those areas, or by removing them from the analysis.

Figure 10: GWR Localized R-Squares of SES Change: 2000 to 2006-10



Tools Used

Mean Center Tool was utilized to approximate the center point of all of the housing in each CT.
- Weighted by number of housing units.
- Used in sample selection.

Central Feature identified the real address closest to the center of all housing in the CT.
- Weighted by number of housing units.
- Used as the incident point in Network Analyst Tool.

Closest Facility in the Network Analyst Toolbox was used to calculate the street network distance, which is in miles, from the central housing feature to the nearest LRT station along each individual LRT line.

High/Low Clustering (Getis-Ord General G) was used to assess the distribution of the data.
- Reports confirmed high clustering prior to performing Hot Spot Analysis was performed.

Hot Spot Analysis (Getis-Ord Gi*) was used to identify statistically significant clustering of high and low values for SES index and SES change.
- Contiguity (edges and corners) was the conceptualization of distance that was decided on because the CTs have a topology that connects them to a neighbor.

Ordinary Least Squares Tools (OLS) was used to run initial linear regressions.
- Used to properly specify the regression model for Geographical Weighted Regression (GWR).

Spatial Autocorrelation (Moran's I) was used to check residuals for autocorrelation.
- No spatial autocorrelation was found; distribution of data was determined to be random.

GWR was run to account for non-stationarity of the OLS regression.
- Adaptive kernel type with bandwidth method based on AICc

Sources

GeoLytics, Inc, Urban Institute, and U.S. Census Bureau. 2013. *Neighborhood change database (NCDB) 2010 tract data for 1970-80-90-00-10*. E. Brunswick, NJ: GeoLytics.

Zuk, Miriam, Ariel H. Bierbaum, Karen Chapple, Karolina Gorska, Anastasia Loukaitou-Sideris, Paul Ong, and Trevor Thomas. 2015. "Gentrification, Displacement and the Role of Public Investment: A Literature Review." *Federal Reserve Bank of San Francisco*, No. 2015-55.

Oregon shapefile for "Study Area," Portland Metro RLIS

Acknowledgements

Nathan Rochester's thesis work on the relationship between light rail transit and gentrification was the catalyst for this project. His definitions of gentrifying factors and the operationalization of gentrifying factors were invaluable in this project.