Public Transit and Employment Access The Effect of the Hiawatha Light Rail Transit on Employment in Minneapolis, MN

> Alexander H. Bentz* Labor Economics Macalester College 12/9/13

*A special acknowledgment is owed to Professor Karine Moe, Professor Sarah E. West, Needham B. Hurst, Deveon Kristiansen and Professor Laura Smith for providing help with methodology and data and to Allison Dautrich, Benjamin Lavin, Veronika Jakubovie, Yiwen Hu and Kyle Coombs for their support along the way.

I. Introduction

Minneapolis's Hiawatha Light Rail Transit (LRT) opened in 2004 at a cost of \$715 million and runs from downtown Minneapolis to the Mall of America and Minneapolis-St. Paul International Airport, both employment and economic centers. (Geotz et al., 2010) An investment of this magnitude is expected to generate an economic impact in addition to its transportation role. Other studies have analyzed the Hiawatha LRT's affect on land use, property values and investment (Goetz et al., 2010; Hurst and West). Light rail access affects not only access to the area around LRT stations; it has the potential to increase access to employment opportunities for those living near LRT stations. I estimate this effect at the census tract level using a difference-in-difference technique that controls for citywide employment trends as well as changes in demographics and other employment access factors. Then, I use a similar difference-in-difference-in-difference technique utilizing propensity score matching to control for urban segregation, finding that LRT station proximity was associated with 8%-11% higher changes in employment rates between 2000 and 2010/2011.

Section II reviews the existing literature on employment and public transit and explores the spatial mismatch hypothesis. Section III develops a theoretical model of commuting costs. Section IV describes the data. Section V summarizes the data and describes employment and population trends during the study period for tracts affected by the LRT and a Minneapolis control group. Section VI presents the difference-in-difference estimations and propensity score matching models. Section VII presents the results, Section VIII addresses estimation issues and Section IX concludes.

II. Literature Review

Economic Impact of the Hiawatha LRT

Previous studies have assessed the economic effects of the Hiawatha Light Rail on land use (Hurst and West; Goetz et al., 2010; CTOD, 2013) and property values (Ko and Cao, 2013) along the corridor, finding small effects close to stations. I am aware of no studies to date analyzing local employment effects of Hiawatha Light Rail station proximity. However, the potential to increase employment may be high due to its downtown and Mall of America termini.

Spatial Mismatch Hypothesis

Much of the interest in the relation between transit and employment arises from the Spatial Mismatch Hypothesis (SMH) first posited by Kain (1968); Kain argued that falling employment of poor blacks living in inner-city neighborhoods during post-war suburbanization could be a result of employment suburbanization. He supports the hypothesis by comparing black employment locations with black residential patterns and concludes that residential segregation of urban poor and urban decentralization increases distance between inner-city workers and suburbanizing employers and thus decreases access to employment for inner-city residents. Holzer (1991) summarizes empirical evidence of a spatial mismatch since Kain (1968), finding evidence that earning potential and total jobs for low-skill/minority workers are higher in suburban areas, away from low-income/minority neighborhoods. Recent research has found correlation between commuting times and employment of black youth (Ihlanfeldt and Sjoquist 1989, 1990) but limited support for SMH in the general population.

Increasing attention has been paid to a corollary concept: the "modal mismatch". This theory says that a larger issue is the mismatch between transportation options for many inner-city residents (public transit) and an auto-focused urban structure. Kawabata and Shen (2007) find that employment accessibility differs for car and public transit commuters in the San Francisco Bay Area in time and spatial pattern. They also find that job accessibility for public transit commuters increased between 1990 and 2000 while accessibility for car commuters decreased due to improved public transit and increased congestion. This suggests increases in the level and/or efficiency of public transit or beneficial spatial job reallocation can improve employment accessibility although these specific outcomes are not tested. Taylor (1995) finds that commuting distance and times have varied little over time while minority employment has fallen, suggesting that a 75 percent higher commuting time for public transit commuters (a modal mismatch), who are more likely to be minority, is more responsible for lower minority employment than a spatial mismatch. Minority workers who live in predominantly minority neighborhoods but commute by private vehicle have similar commuting times to whites who live in suburban neighborhoods; it is those that rely on transit that are disadvantaged.

Urban planners have taken the idea of a spatial or modal mismatch as a call for increased inner-city public transit access to connect residents with suburban jobs. Fan (2012) summarizes planning approaches to combat spatial and modal mismatches. Public transportation is considered a solution to problems of limited transportation to employment centers but it is recognized that all transit does not increase access to jobs.

Public Transit and Employment Outcomes

The empirical results looking at the relationship between public transit and employment outcomes are much more mixed. Sanchez (1999) looks at effects of transit access on labor force participation by census block group for Portland and Atlanta in 1990 and finds that, controlling for worker characteristics, access to transit within walking distance is positively correlated with labor force participation. Matas et al. (2010) find that poor public transit accessibility negatively affects employment probability for uneducated workers in Barcelona and Madrid although applicability of this result to U.S. cities may be difficult because of the countries' contrasting urban structures.

Because transit access is likely most important for low-income residents without automobile access and because individual-level micro data are available, many studies look at employment outcomes for welfare recipients. Sanchez et al. (2004) tweak Sanchez (1999) by looking at employment status of Temporary Assistance for Needy Families (federal "welfare-to-work" program) recipients in six US metros and find that access to transit did not significantly affect employment outcomes (as measured by their TANF case status). This result is particularly striking, as they found no significant correlation even when looking at only households without car access. They also suggest including access to commercial centers and daycare in further analysis to account for the complex transit needs of working mothers. Alam (2009) takes the concept a step further by modeling employment access using transportation and traffic-flow modeling software; they find a negative correlation between transit accessibility and length of stay on TANF benefits, supporting the connection between public transit and employment.

By looking at a static relationship between transit access and employment, these studies risk endogeneity between transit access and employment; individuals may be more likely to live near transit because they seek employment, independent of observable characteristics. Sanchez (1999) acknowledges that this is an impediment to establishing a causal relationship. In addition, these studies treat access to all public transit equally, rarely distinguishing between types or accounting for the complex nature of commute decisions.

This paper seeks to add to the literature by analyzing the effects of changing employment accessibility in Minneapolis resulting from the opening of the Hiawatha Light Rail. Previous studies have looked at accessibility varied by individual or geography but not over time. This approach may correct for some of the endogeneity in previous studies, as household location decisions in anticipation of light rail access may be less prevalent soon after opening than decisions made over a longer time period. In addition, the study will attempt to evaluate the impact of a single change in public transit infrastructure, which may reduce some of the complexity surrounding commuting decisions. LRT access will not affect employment accessibility equally for all residents but it would be expected to increase accessibility for at least a portion of residents, allowing the results of that change to be evaluated.

III. Theory

An adapted labor supply model is used to predict the effect of increased transit accessibility on employment outcomes.

Labor Supply

We assume that workers are heterogeneous in their preferences and identical in market wages but that all derive utility from consumption (C) and leisure (L) according to a utility function,

$$U = U(C, L)_i \tag{1}$$

Consumption can be obtained through market work at a constant hourly wage, w or through unearned income, V. However, market hours require a commuting cost, p, per hour worked, so that

$$C = (w - p)M + V \tag{2}$$

where M is the number of market hours worked.¹ Workers maximize utility subject to the constraints imposed by the market wage and finite hours available, T, such that

¹ The commuting cost assumption is required to predict a change in employment status from a change in transportation accessibility.

 $T = M + L \tag{3}$

Formally, workers maximize

$$U = U(C, L) \tag{4}$$

subject to:

$$C = (w - p)M + V \tag{5}$$

$$T = M + L \tag{6}$$

Or substituting Equation 6 into Equation 5:

$$C = wT - wL - (pT - pL) + V \tag{7}$$

Assuming a Cobb-Douglas utility function,

$$U = a\ln(C) + b\ln(L) \tag{8}$$

where a and b are between 0 and 1 and vary across individuals, we can set up a Lagrangian Optimization Problem:

$$\ell = a_i \ln(C) + b_i \ln(L) + \lambda [wT - wL + V - C]$$
(9)

Deriving the first-order conditions we get:

$$\frac{\partial \ell}{\partial C} = \frac{a_i}{C} - \lambda = 0 \tag{10}$$

$$\frac{\partial \ell}{\partial L} = \frac{b_i}{L} - w\lambda = 0 \tag{11}$$

$$\frac{\partial \ell}{\partial \lambda} = wT - wL - (pT - pL) + V - C = 0$$
(12)

Combining Equation 10 and Equation 11 and solving, derives wage as equal to the marginal rate of substitution between leisure and consumption:

$$w = \frac{b_i / L}{a_i / C} = MRS_{L,C}$$
(13)

Workers will devote hours to market work until the ratio of marginal utility of consumption and leisure is equal to the wage. However, the wage is not always high enough to induce market labor. The reservation wage is defined as the wage where a worker is indifferent between working and not working. The reservation wage is equal to the $MRS_{L,C}$ when all hours are devoted to Leisure (T=L). With market hours equal to

zero, Consumption will be equal to unearned income (V). Substituting into the $MRS_{L,C}$ gives the reservation wage,

$$W_{R} = \left(\frac{b}{a}\right)_{i} \frac{V}{T}$$
(14)

Workers will supply market labor if their effective wage (wage minus the commuting cost) is greater than their reservation wage.

$$w - p > (w_{R})_{i}$$

$$w - p > (\frac{b}{a})_{i} \frac{V}{T}$$
(15)

An individual's employment status is a function of market wage, commuting cost, worker preferences and unearned income. It is assumed that the labor market is sufficiently large that changes in labor supply will not affect wages. Because individuals are heterogeneous in their preferences for consumption and leisure, a portion of individuals will be employed while another portion will be unemployed.

An improvement in public transit for some individuals (those nearest to the improvement) will decrease the commuting cost, to p_2 , and increase the effective wage for those individuals. The increased effective market wage will increase the portion of individuals with reservation wages low enough to induce employment. Individuals who are not affected by the improved public transit continue to have a commuting cost of p_1 where p_1 is greater than p_2 . The portion of individuals employed at p_1 is smaller than for those individuals with a commuting cost of p_2 .

$$w - p_1 > (\frac{b}{a})_i \frac{V}{T}$$

$$w - p_2 > (\frac{b}{a})_i \frac{V}{T}$$

$$p_1 > p_2$$
(16)

In conclusion, a public transit improvement that lowers commuting cost is expected to cause an increase in employment for the affected individuals and no effect for those unaffected, holding wages, preferences and unearned income constant.

IV. Data

Data are from 2000 U.S. Decennial Census and the 2010 and 2011 American Community Survey. Data are reported at the census tract level; 2000 Census data are harmonized with 2010 census tract definitions using a methodology provided by Logan and Stults (2012). The 2000 Decennial Census employment data are from Summary Form 3 (SF3) sample data and are not 100% counts. However, more caution needs to be taken in interpreting the 2010/2011 American Community Survey data. These are drawn from 5-year estimates from data collected over the previous 5 years (2010 data are drawn from data collected from 2006-2010 and 2011 data are from 2007-2011). Due to the nature of the estimates, multiple years are included here primarily for redundancy. These estimates cannot be interpreted as giving the effects for that particular year as they are derived from four of the five same years).

Further, margins of error reported for this data are large, particularly for tracks with smaller populations (Figure 1). Standard errors were derived for each variable used; Table 1 shows the average standard errors for Minneapolis census tracts. On average, the standard error of 2010 employment rate within a census tract is 14.2% with an average employment rate of 71.4%. The large sampling error makes it difficult to find small changes in employment as might be expected to find in the following estimations and to obtain robust results.

Employment rate is defined as the percentage of the civilian (non-armed forces) population age, 16-64, that is currently employed. Civilian population is used instead of labor force as the theory outlined here is looking at the decision of whether to seek or accept a job and thus is closer to a question of labor force participation than employment or unemployment. Variables are also included for gender (female working age population), median age (potential experience), education (population with a bachelors degree or higher and population with a high school diploma as their highest education), race, marital status (married individuals) and household vehicle access.

A note needs to be made regarding the construction of the race variables. The U.S. Census reports race in five categories and other categories that include individuals who identify as more than one race. Hispanic origin is reported separately. For the purposes of this estimation, persons of Hispanic origin are selected into distinct category and all other racial categories are constructed using only persons not of Hispanic origin. Figure 15 shows non-Hispanic White Population by Census Tract; Figure 16 shows non-Hispanic African-American population; Figure 17 shows Asian population; Figure 18

shows American Indian population; Figure 19 shows Hispanic Population; Figure 20 shows populations of all other races. Race variables are included to account for the (changing) diversity of each census tract but their imprecise measurement and the nature of racial categorization must also be taken into account when interpreting the results. Another important distinction is that due to data limitations, these race variables represent the entire census tract population and are not limited to the working age population.

Sample Area

This analysis is limited to Minneapolis census tracts. This excludes stations in Bloomington, MN including the Minneapolis-St. Paul airport (MSP) and the Mall of America; Figure 2 gives a reference map of the study area. For the purpose of this research, these stations are not of interest because they lack nearby residential land use. Each tract averages a population of 3,272 to 3,300 over the three years of available data. The smallest tract had a population of 1,163 and the largest tract had a population of 10,346 in 2011. Using Geographic Information Systems (GIS), census tracts are selected to be in the LRT Station Area if any section of the tract is within ½ mile of an LRT station; these tracts are highlighted in Figure 2. This distance has been identified as a walkable distance by previous studies including Hurst and West (2013) and Sanchez (1999).² Control distances are calculated using GIS from the centroid of each census tract to a single point (CBD) or line segment (Lake Street, nearest highway).

V. Summary Statistics

Table 2 presents the summary statistics for each year of data used. Table 3 presents the same means for the LRT Station Area and outside tracts. Finally, Table 4 presents the percent change in each variable by the same Station Area designation and Figures 3-14 show the spatial relationships between these variables. Between 2000 and 2010/2011, Station Area tracts experience higher employment rate increases relative to

² This method is used instead of the traditional centroid method because of the relatively large areas of each census tract. Using only census tracts with a centroid within $\frac{1}{2}$ mile of an LRT station would limit the Station Area sample to only 11 census tracts concentrated primarily in the Central Business District (CBD) instead of 27 evenly distributed census tracts when using the more inclusive method that includes all tracts that intersect a $\frac{1}{2}$ mile radius of LRT stations. This is also problematic as, for some tracts, a majority of the tract is more than $\frac{1}{2}$ mile away from an LRT station and would not be considered within the Station Area if data were available at a finer geographic scale (see Figure 2). Density may also be a confounding factor as the larger census tracts (with a smaller percentage of residents actually within $\frac{1}{2}$ mile) are also less dense so that a very small number of residents are likely within the affected area and further, these tracts are likely to differ from the denser, more centrally located tracts in their urban form and demographics.

the rest of Minneapolis; however, these increases corresponded with increases in education levels and vehicle access, variables correlated with higher employment. Thus, the higher employment outcomes may not have been a result of improved employment access and rather changing neighborhood demographics (gentrification) in part of the LRT Station Area.

Population by census tract falls between 2000 and 2010/2011 although the high end of the range increases for 2010/2011 as one census tract increased in population by 2,961 residents between 2000 and 2011. Population falls between 1.2% and 1.4% outside of the LRT station area while increasing 5.7% to 8.3% within the LRT Station Area. Percent change in employment rate is negative between 2000 and 2010/2011, likely due to the 2009 recession, but is 6.5% to 9.7% higher within the LRT station area. Working age female population increases in Minneapolis and within the LRT Station Area. Median age increases in Minneapolis but less in the LRT Station Area. This is coupled with a larger increase in the working age population in the LRT Station Area and can likely be explained by a decrease in the number of families (using married population as a proxy). The nonwhite population in the LRT Station Area changes only slightly (between -1.2% and +2.3%) but increases outside (+11.0%). Finally, household vehicle access increased more (4.1% to 5.7%) in the LRT Station Area than outside (1.4%). To summarize, employment increased in LRT station areas and fell in the rest of the city over the study period but the area also saw a significant demographic change that confounds the analysis.

VI. Estimation Strategy

Following the stylized Labor Supply model developed in Section III, the portion of individuals employed in a given population is a function of commuting costs, individual preferences, market wage and unearned income. Here, commuting costs are assumed to fall after light rail construction in 2004 for census tracts within the LRT Station Area. Based on this model, I seek to answer the question: did decreased commuting costs caused by LRT increase employment relative to the rest of Minneapolis at a statistically significant level (Model I), when controlling for other employment access factors (Model II)? And, can this effect be explained by demographic changes (Model III)? I assume that census tracts and their employment outcomes are independent of each other and that macroeconomic conditions affect each census tract evenly, subject to controllable factors.³ Employment outcomes are also a function of individual market wages. This can be partially accounted for by differences in education, age and race, but there may be other factors.

Model I – Basic Estimation

First, I evaluate a basic model to determine whether employment outcomes were statistically different for LRT Station Area tracts relative outside tracts. I compare employment in census tracts before (2000) and after the LRT construction (2010/2011). Following Hurst and West (2013), a difference-in-difference technique is used to isolate the differential effect of residing within the LRT Station Area after construction while controlling for the possibility that there are characteristics associated with being within the nonexistent LRT Station Area in both periods that affect employment.

Adapting Hurst and West (2013), Model I estimates the change in employment attributable to the introduction of LRT stations with no other controls. The dependent variable, Emp/cap_{it} , is employment rate by census tract in each period. S_i is a dummy variable equal to one for LRT Station Area tracts in either period. T_t is a dummy variable equal to one in the after period (2010 or 2011) and zero in the before period (2000).

$$Emp / cap_{it} = \alpha + \beta_1 S_i + \beta_2 T_t + \beta_3 S_i T_t + \varepsilon_{it}$$

The coefficient β_3 gives the percent increase in employment for LRT Station Area tracts after LRT opening relative to LRT Station Area tracts before opening, relative to the rest of Minneapolis. This is the employment effect of the LRT.

Model II – Population Characteristics

The theoretical model includes employment determinants other than commuting costs. Thus, Model II adds proxies for population characteristics, individual preferences, wages and unearned income, within each census tract. Adapting Sanchez's (1999) single-period study of employment within census tracts, median age (and median age squared) over 18 (potential experience), working age female population (% of total population), population with a bachelor's degree (% of population over 25), nonwhite population (%

³ This last assumption is difficult to make because of potential differences in investment levels in each neighborhood. Perhaps a large job-training program targeted a neighborhood. In particular, investments may have targeted a neighborhood *because* of the LRT investment.

of total population) and married population (% of population 15 and over) serve as controls for leisure and consumption preferences, market wages and unearned income. A variable for vehicle ownership (% of households with access to one or more vehicles) is added to control for the possibility that light rail access does not affect employment access evenly. These controls are included as a vector, C_{it} .

$$Emp / cap_{it} = \alpha + \beta_1 S_i + \beta_2 T_t + \beta_3 S_i T_t + \beta_4 C_{it} + \varepsilon_{it}$$

Model III – Distance Controls

Model III adds controls for other factors that affect commuting costs. Distance to local employment centers (CBD, Lake Street) and other transportation modes (closest highway) may contribute to light rail's relative employment access effect. D_i is a vector of these distances; an interaction term with the period dummy variable controls for the possibility that these factors may have changed in importance between periods.

 $Emp / cap_{it} = \alpha + \beta_1 S_i + \beta_2 T_t + \beta_3 S_i T_t + \beta_4 C_{it} + \beta_5 D_i + \beta_6 D_i T_t + \varepsilon_{it}$ III

Propensity Score Matching

Models I-III assume that census tracts characteristics are independent of their inclusion in the LRT Station Area sample. However, this may not be the case. Living near LRT stations may be an amenity (or disamenity) independent of employment access or certain groups may live near LRT stations for noneconomic reasons (e.g. if a particular ethnic group is clustered there or if younger, non-married, individuals are more likely to live downtown). Maps 5-20 and Table 3 show that demographic characteristics are not evenly distributed inside and outside the LRT Station Area or within the LRT Station Area itself. There may be reason to believe that census tract selection in the LRT station area group is not a random.

Propensity score matching determines which characteristics are correlated with being in the LRT Station Area sample and assigns a non-LRT Station Area tract as a 'match' for comparison. Minneapolis is a diverse city but also a segregated city, as illustrated by Maps 15-20. There is a large amount of variance in racial makeup along the Hiawatha line but even more in the rest of the city; the Inter Quartile Range for nonwhite population in 2010 is 18.3-55.4% in the LRT Station Area and 15.5%-63.1% outside. This particular statistic; however, masks some outliers. Outside the LRT station area, the 99th percentile is 95.2% nonwhite and 86.8% nonwhite within the LRT station area.

While 2010 Minneapolis employment rates averaged 65.8%, employment rates for the 9 tracts in the 90th percentile or higher of nonwhite population averaged an employment rate of only 43.9%. There is a strong (negative) correlation between nonwhite population and employment rates (-0.73 in 2010).

To build a control group that has the most similar racial makeup to the LRT Station Area, propensity scores are calculated and matched based on the mean percent nonwhite population in each census tract between 2000 and 2010/2011. The LRT Station Area tracts are matched with 21 tracts in 2010 and 23 tracts in 2011. These control tracts are weighted to form a complete control group. Table 6 shows the means and 95% confidence intervals for each variable by LRT station area and the control group. The average nonwhite population in each group is almost identical, as is the intention of the matching. The LRT Station Area has, on average, a smaller African-American and Asian population and a larger American Indian and Hispanic population relative to the control group. The LRT Station Area also has a lower vehicle access rate and is, on average, closer to the CBD and Lake Street and farther from the nearest highway. The summary statistics show that these groups, although similar, are *not* identical. However, if each tract is randomly selected into the LRT Station Area and control groups except for nonwhite population, then employment outcomes are comparable and potentially attributable to LRT access.

VII. Results

If LRT access/proximity affects employment outcomes, we would expect, holding everything else constant, employment to be higher in LRT Station Area tracts in the after period relative to outside tracts. Specifically, the coefficient on the interaction term between the after period dummy and the LRT Station Area dummy would be positive and statistically significant. Results for each specification are shown in Table 5. Depending on the specification and year, this LRT employment effect is between 1.7% and 3.8%; but not statistically significant. The LRT Station Area and after period dummy variables are negative in each specification, as would be expected from the trends shown in Section V. The after period dummy is statistically significant (99% confidence) in Model II and the station area dummy is statistically significant (99% confidence) in Model III.⁴

Overall, Models II and III explain much of the employment variation with R^2 values of 0.68/0.71 and 0.73/0.77. However, the results do not significantly support the theoretical predictions of improved employment access.

Propensity Score Matching

Table 6 shows the means and 95% confidence intervals for the LRT Station Area and the matched and weighted control group. The variable of interest is percent change in employment rate (between 2000 and 2010/2011) to answer the question: did LRT Station Area tracts have greater changes in employment relative to a control area? And, was this difference statistically significant? Between 2000 and 2010, LRT Station Area tracts had an average percent change in employment rate 11.0% higher than the control area; between 2000 and 2011, that difference was 8.0%. This result is statistically significant at the 95% level as the confidence intervals do not overlap for either period. If nonwhite population is the only non-randomly distributed independent variable between the LRT Station Area and other Minneapolis tracts, then LRT access improved employment.

VIII. Estimation Issues

The first issue affecting estimation is the reliability of American Community Survey data at the census tract scale. This is *not* taken into account when calculating statistical significance for Models I-III or propensity score matching. Secondly, there is likely omitted variable bias because of limitations in data and the difficulty of predicting employment at an aggregated (not individual) level. Further, none of the LRT or distances variables directly measure commuting costs or employment access; they are simply proxies.

The controls' behavior is important to validate the estimation even if the estimated result is inconsistent with employment access theory. The age and vehicle access coefficients are significant (99% confidence) in Models II and III. The age variable is negative, which may be explained by changing consumption/leisure preferences for older individuals or multicollinearity. The vehicle access coefficient is

⁴ The models are also estimated male and female employment rates. For these, the controls are considered neighborhood-effects. For men and women separately, the LRT effects are positive and not statistically significant.

positive (22.3%-27.4% increase in employment for a 100% increase in household vehicle access) and statistically significant for both Models II and III. Unexpectedly, the nonwhite coefficient is positive and statistically significant (99% confidence) in Models II and III. This may be a product of the imprecise race definition or due to an omitted variable (or multicollinearity). ⁵ As defined, nonwhite population is (negatively) correlated with median potential experience.⁶

The distance controls relative to major employment centers (CBD and Lake Street) are both statistically significant (99% confidence). CBD proximity is associated with lower employment rates while Lake Street proximity is associated with higher employment rates. The negative correlation between CBD proximity and employment may be explained by suburbanization and the inner-city 'donut' effect although it would be expected that demographic characteristics would control for that effect.⁷ Distance to nearest highway is not statistically significant in either period, which may be expected because of Minneapolis's high overall highway density (the largest distance is only 1.56 miles).

Multicollinearity and Endogeneity

Potential multicollinearity (correlation between independent variables) is a major concern when including variables that affect employment. Individuals (and tracts) with characteristics that increase their employment can also be expected to be wealthier and thus have greater vehicle access. Other multicollinearity results from Minneapolis's urban layout. Highways are concentrated close to downtown. Thus, the distance to downtown for a given census tract will be (negatively) correlated with the distance to the nearest highway. Vehicle access may also be correlated with transportation factors, including light rail access; vehicle access is lower closer to downtown area as seen in Figure 5 (correlation of 0.67 with distance to downtown in 2010) and within the LRT Station Area (correlation of -0.14 if within LRT station area). We expect vehicle access to explain some of variation in employment rates; however, vehicle access may not affect employment similarly for tracts close to downtown or within the LRT Station Area.

⁵ Estimating the same models with the six racial categories established in Section IV does not change the qualitative results and maintains joint statistical significance (99% confidence).

⁶ Estimating the model without median potential experience results in negative and statistically significant (99% confidence) coefficients for nonwhite population pointing to multicollinearity as a likely culprit.

⁷ These results are consistent when using all six race categories.

A further confusion for this estimation arises because of the potential for LRT access to affect vehicle access. Overall, between 2000 and 2010/2011, the average vehicle access for LRT Station Area tracts increases 4.1% to 5.7% (compared to 1.4% for the rest of the city). Much of this may be explained by demographic changes. There may also be a problem of endogeneity with regard to vehicle access. Vehicles aid in access to employment but also may result from increased wealth derived from employment.

Variance Inflation Factors (VIF) are calculated for each variable post estimation. Multicollinearity is high among some of the controls as might be expected; however, this does not introduce bias into the interaction coefficient that is of greatest interest. Multicollinearity may explain some of the confusing results in Section VII.

Heteroskedasticity

The Bruesch-Pagan/Cook Weisberg Heteroskedasticity Test reveals significant heteroskedasticity. Robust standard errors are used here to correct for heteroskedasticity. Using robust standard errors makes it more difficult to find statistically significant results for small sample sizes; however, statistical significance was not found on the variable of interest without robust standard errors.

IX. Conclusion

I estimate the effect of a large public transit investment on employment outcomes for residents living near the new LRT stations. During the study period, employment changes were higher within the LRT station area; however, the demographics of the area also changed dramatically, attracting younger, more educated workers. Conventional difference-in-difference estimations with controls for population demographics and other employment access factors find statistically insignificant results. Using a difference-indifference propensity score matching estimation to account for nonrandom racial segregation, I find that LRT station area tracts had employment rate changes 8%-11% higher than a Minneapolis control group that are statistically significant at a 95% confidence level. This result is unreliable because of American Community Survey sampling errors. The results play a role in determination of the societal cost and benefits of public transit investment. This analysis provides weak support for improved employment outcomes resulting from public transit investment. However, future studies may seek to also look at changing demographics as a goal of public transit investment.

	20	010	20	011
Variable	Mean	Standard Error	Mean	Standard Error
Employment Rate (%)	71.4	14.2	71.2	12.0
Population	3,272	222	3,291	212
Working Age Population (16-64) (%)	72.6	12.5	72.8	10.6
Female Working Age Population (%)	49.6	13.8	49.4	11.2
Median Age	32.8	2.0	33.3	1.9
Population with High School Diploma (%)	21.0	5.6	19.8	4.8
Population with Bachelor Degree (%)	40.0	9.0	41.0	7.6
Married (%)	36.1	4.6	35.5	4.3
Access to Vehicle (%)	80.8	10.5	80.6	10.1
Nonwhite (%)	39.9	10.0	40.0	9.3
Non-Hispanic White (%)	60.1	6.6	60.0	6.3
Non-Hispanic African-American (%)	19.3	4.7	19.1	4.5
Non-Hispanic Asian (%)	6.0	2.6	5.7	2.5
Non-Hispanic American Indian (%)	1.7	1.8	1.7	1.4
Hispanic (%)	9.6	3.8	9.9	3.8
Other Races (%)	6.4	3.9	6.9	3.4
Standard errors are derived from A Source: US Census				

 Table 1: ACS Average Derived Standard Errors

Table 2: Summary Statistics By Census Tract							
	US Census	American Community Survey					
Variable	2000	2010	2011				
Employment Rate (%)	74.0 (11.3)	71.4 (12.3)	71.2 (11.6)				
	44.9-91.2	37.1-89.1	37.5-87.6				
Population	3,300 (1,186)	3,272 (1,303)	3,291 (1,326)				
	1,381-7,551	1,466-10,216	1,163-10,346				
Working Age Population	68.4 (10.7)	72.6 (9.3)	72.8 (8.6)				
(16-64) (%)	48.6-97.0	50.9-97.1	54.9-97.1				
Female Working Age	48.7 (4.3)	49.6 (6.0)	49.4 (5.6)				
Population (%)	30.0-55.7	29.0-65.3	31.6-63.3				
Median Age	31.7 (5.8)	32.8 (6.6)	33.3 (6.2)				
	19.6-42.5	19.8-48.3	20.0-47.5				
Population with High	21.9 (8.3)	21.0 (9.3)	19.8 (9.2)				
School Diploma (%)	5.6-38.5	2.2-42.9	4.0-41.7				
Population with Bachelor	34.7 (19.7)	40.0 (20.7)	41.0 (20.9)				
Degree or Higher (%)	3.7-77.7	5.9-91.8	5.4-87.3				
Married (%)	38.3 (10.5)	36.1 (12.9)	35.5 (13.1)				
	10.2-64.8	4.1-67.8	4.2-68.7				
Access to Vehicle (%)	79.5 (13.8)	80.8 (13.4)	80.6 (13.7)				
	36.8-98.6	4.4-100	42.3-100				
Nonwhite (%)	39.1 (26.7)	39.9 (26.4)	40.0 (26.0)				
	5.6-96.7	5.6-95.2	6.7-93.9				
Non-Hispanic White (%)	60.9 (26.7)	60.1 (26.4)	60.0 (26.0)				
	8.3–95.1	4.8-94.4	8.3-93.9				
Non-Hispanic African-	19.0 (17.9)	19.3 (18.0)	19.1 (18.0)				
American (%)	7.1–66.4	0.0-70.9	0.0-67.9				
Non-Hispanic Asian (%)	6.1 (5.2)	6.0 (70.2)	5.7 (6.6)				
	1.0-2.6	0.0-41.1	0.0-32.9				
Non-Hispanic American	2.0 (2.6)	1.7 (2.9)	1.7 (3.0)				
Indian (%)	0.04-21.2	0.0-21.3	0.0-17.1				
Hispanic (%)	7.9 (7.6)	9.6 (10.6)	9.9 (10.9)				
	1.0-33.1	0.0-58.2	0.0-58.4				
Non-Hispanic	4.1 (1.8)	6.4 94.8)	6.9 (5.2)				
Other Race(s) (%)	1.0-9.7	0.0-30.5	0.0-39.9				
Distance (miles)							
To Closest LRT Station		1.83 (1.16)					
		0.19-4.88					
To CBD		2.86 (1.39)					
		0.10-6.13					
To Lake Street		2.27 (1.64)					
		0.048-6.63					
To Nearest Highway		0.49 (0.38)					
		0.017-1.56					

Table 2: Summary Statistics By Census Tract

Values are averages of sample 2010 census tracts. 2000 Census data are harmonized with 2010 Census Tracts. Standard Deviations are shown in parentheses. Ranges are shown in italics

Employment Rate is number of individuals employed as a percentage of civilian population age 16-64. Control Distances are calculated from centroid of census tract.

Source: US Census

By Census Tract and LRT Station Proximity						
Variable	2000 2010		2011			
Station Proximity	Station Area	Outside	Station Area	Outside	Station Area	Outside
Employment Rate (%)	71.1	74.9	71.3	71.4	70.1	71.6
Population	3,484	3,244	3,638	3,160	3,715	3,161
Working Age	71.3	67.5	76.2	71.5	75.7	71.9
Population						
(16-64) (%)						
Female Working Age	46.6	49.3	46.8	50.4	47.5	50.0
Population (%)						
Median Age	32.6	31.5	33.5	32.6	33.0	33.5
Median Age						
Population with High	23.1	21.5	20.9	21.0	19.5	19.9
School Diploma (%)						
Population with	31.0	35.8	40.9	39.7	41.6	40.8
Bachelor Degree (%)						
Married (%)	36.6	38.8	32.2	37.3	31.9	36.6
Access to Vehicle (%)	74.5	81.1	77.4	81.9	76.3	82.0
Nonwhite (%)	38.7	39.2	38.5	40.3	39.6	40.1
Non-Hispanic	61.3	60.8	61.5	59.7	60.4	59.9
White (%)						
Non-Hispanic	16.5	19.8	16.6	20.2	17.2	19.7
African-American (%)						
Non-Hispanic	4.6	6.6	4.3	6.5	4.5	6.1
Asian (%)						
Non-Hispanic	4.0	1.4	3.5	1.1	3.7	1.1
American Indian (%)						
Hispanic (%)	8.9	7.6	10.2	9.4	10.2	9.9
Non-Hispanic	4.7	3.9	7.2	6.1	7.5	6.8
Other Races (%)						

Table 3: Employment and Demographic CharacteristicsBy Census Tract and LRT Station Proximity

Values are averages of sample 2010 census tracts. 2000 Census data are harmonized with 2010 Census Tracts. 27 (of 115) Census Tracts are included in the less than ½ mile LRT Station Area.

Employment is number of individuals employed as a percentage of civilian population age 16-64. Source: US Census

By Census Tract and LRT Station Proximity						
20	10	2011				
Station Area	Outside	Station Area	Outside			
+1.7	-4.8	-0.16	-4.3			
+5.7	-1.2	+8.3	-1.4			
+7.2	+6.9	+6.5	+7.5			
+0.97	+2.4	+2.6	+1.5			
+2.5	+4.1	+1.4	+6.8			
-8.6	-1.6	-15.8	-7.2			
+40.2	+22.9	+42.2	+24.7			
-14.4	-4.7	-15.0	-7.0			
+5.7	+1.4	+4.1	+1.4			
-1.2	+11.0	+2.3	+11.0			
-0.6	+0.2	-2.0	+2.5			
-1.1	+20.7	+2.1	+16.0			
-3.4	-0.01	-4.2	-3.2			
-15.4	-0.6	-12.6	-19.1			
+5.9	+33.6	+10.1	+42.8			
+72.1	+87.9	+84.0	+108.4			
	$\begin{array}{r} 20\\ \hline \text{Station Area}\\ +1.7\\ +5.7\\ +7.2\\ +0.97\\ +2.5\\ \hline -8.6\\ +40.2\\ -14.4\\ +5.7\\ \hline -1.2\\ -0.6\\ -1.1\\ -3.4\\ -15.4\\ +5.9\\ \end{array}$	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	$\begin{array}{c c c c c c c c c c c c c c c c c c c $			

Table 4: Percent Change from 2000Employment and Demographic CharacteristicsBy Census Tract and LRT Station Proximity

Values are averages of sample 2010 census tracts. 2000 Census data are harmonized with 2010 Census Tracts. 27 (of 115) Census Tracts are included in the less than ½ mile LRT Station Area.

Employment is number of individuals employed as a percentage of civilian population age 16-64. Source: US Census

Table 5: Affect of LRT Proximity on Employment: Models I-III						
	(1)	(2)	(3)	(4)	(5)	(6)
Variable	I - 2010	I - 2011	II - 2010	II - 2011	III - 2010	III - 2011
Station Area Proximity (within ¹ / ₂ mile)	-0.0381	-0.0381	-0.0233*	-0.0242*	-0.0375***	-0.0436***
• •	(0.0268)	(0.0268)	(0.0140)	(0.0137)	(0.0138)	(0.0139)
After Period (2010/2011)	-0.0348*	-0.0327*	-0.0375***	-0.0403***	-0.0225	-0.0116
	(0.0177)	(0.0170)	(0.00931)	(0.00814)	(0.0269)	(0.0197)
After Period*Station Area	0.0376	0.0227	0.0255	0.0206	0.0216	0.0166
	(0.0372)	(0.0369)	(0.0210)	(0.0206)	(0.0224)	(0.0211)
Percent Nonwhite Population			0.0181***	0.0194***	0.0152***	0.0149***
			(0.00366)	(0.00390)	(0.00377)	(0.00376)
Potential Experience (Median Age above 18)			-0.000499***	-0.000561***	-0.000432***	-0.000459***
			(0.000108)	(0.000116)	(0.000115)	(0.000113)
Potential Experience ²			-0.155***	-0.181***	-0.182***	-0.237***
			(0.0380)	(0.0355)	(0.0363)	(0.0316)
High School Diploma-Highest Level (%)			-0.208*	-0.204**	-0.0755	-0.0393
			(0.121)	(0.100)	(0.116)	(0.0930)
Bachelors Degree or Higher (% of Population)			0.0194	-0.0146	0.0290	-0.0320
			(0.0643)	(0.0547)	(0.0724)	(0.0512)
Percent Female			-0.0930	-0.127	-0.129	-0.153
			(0.130)	(0.106)	(0.145)	(0.104)
Percent Married			-0.0778	-0.0721	-0.101	-0.0422
			(0.0597)	(0.0536)	(0.0621)	(0.0555)
Vehicle Access (Percent)			0.274***	0.288***	0.223***	0.253***
			(0.0575)	(0.0489)	(0.0618)	(0.0463)
Distance to CBD (Miles)					0.0173***	0.0131***
					(0.00479)	(0.00457)
After Period (2010/2011)*Distance to CBD					-0.00143	-0.00280
					(0.00583)	(0.00528)
Distance to Lake Street (Miles)					-0.0119***	-0.0156***
					(0.00334)	(0.00331)
After Period*Distance to Lake Street					-0.00424	-0.00335
					(0.00507)	(0.00437)
Distance to Nearest Highway (Miles)					-0.0148	-0.0149
					(0.0117)	(0.0110)
After Period*Distance to Nearest Highway					0.00308	-0.00775
					(0.0209)	(0.0180)
Constant	0.749***	0.749***	0.564***	0.585***	0.622***	0.655***
	(0.0115)	(0.0115)	(0.0831)	(0.0694)	(0.0949)	(0.0712)
Observations (per period)	115	115	115	115	115	115
Adjusted R ²	0.008	0.013	0.725	0.757	0.758	0.798

.

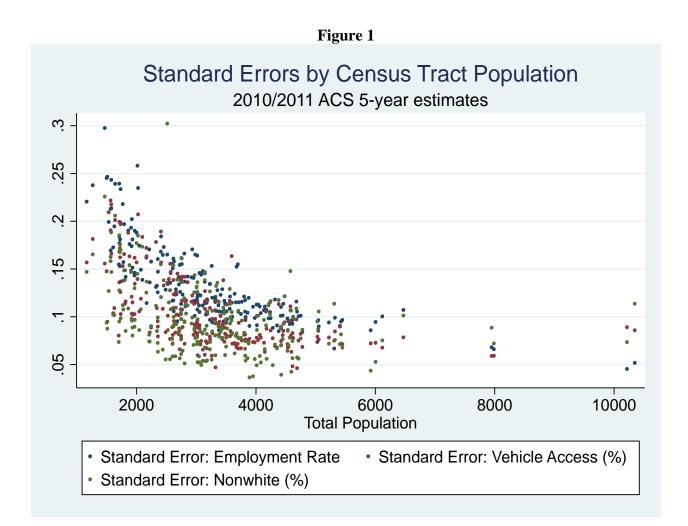
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

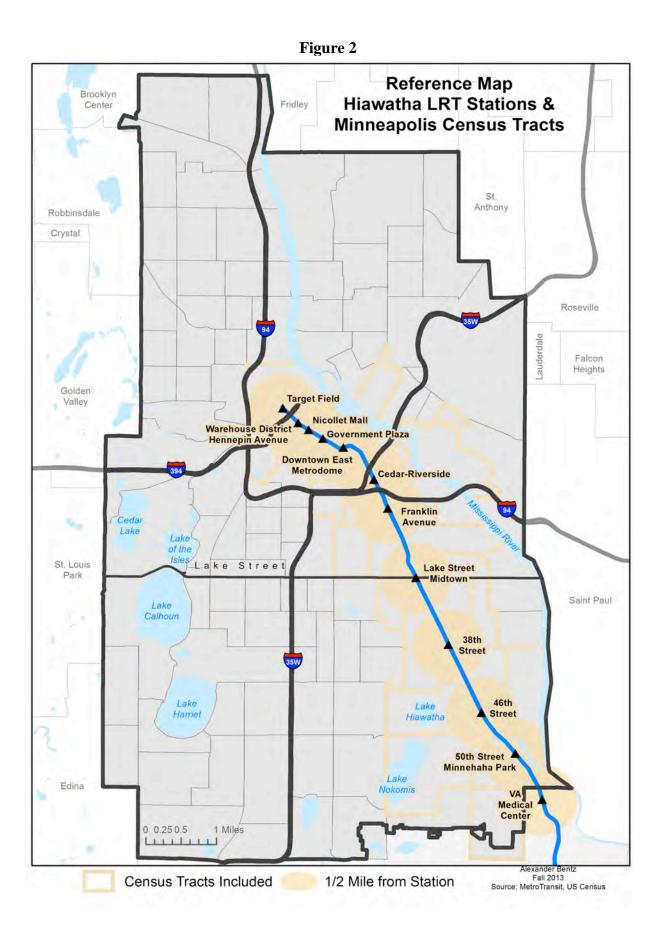
Station Area and Control Group–Means and 95% Confidence Intervals						
	2010 2011					
Variable	LRT Station	Control	LRT Station	Control		
Variable	Area	Group	Area	Group		
Percent Change in	1.7	-9.3	-0.16	-8.2		
Employment Rate	(-4.20–7.58)	(-13.74.9)	(-6.03–5.72)	(-12.53.9)		
Employment Rate (%)	71.3	68.2	70.1	69.1		
Employment Rate (70)	(66.8–75.9)	(62.4–74.1)	(65.5–74.6)	(64.7–73.6)		
Nonwhite (%)	38.5	40.9	39.6	39.3		
	(29.3–47.8) 61.5	(30.1–51.6) 59.1	(30.4–48.8) 60.4	(29.4–49.2) 60.7		
Non-Hispanic White (%)	(52.2–70.7)	(48.4–69.9)	(51.2–69.6)	(50.8–70.6)		
Non-Hispanic	16.6	20.7	17.2	19.5		
African-American (%)	(11.3–22.0)	(12.4–28.9)	(11.6–22.9)	(13.0–26.0)		
	4.3	8.8	4.5	6.9		
Non-Hispanic Asian (%)	(2.7–5.9)	(5.0–12.6)	(2.7–6.3)	(3.6–10.3)		
Non-Hispanic	3.5	0.81	3.7	1.6		
American Indian (%)	(1.7–5.4)	(0.18–1.4)	(1.8–5.6)	(0.67–2.5)		
	10.2	6.7	10.2	7.7		
Hispanic (%)	(5.7–14.7)	(4.2–9.3)	(5.8–14.6)	(2.9-12.5)		
Non-Hispanic Other (%)	7.2	7.6	7.5	7.1		
Non-Inspanie Other (%)	(7.2–8.5)	(4.6–10.6)	(6.4–8.6)	(3.7–10.5)		
Population	3,639	3,160	3,715	3,155		
-	(2,926–4,351)	(2,744–3,575)	(3,005–4,424)	(2,619–3,691)		
Female Working Age	46.8	49.9	47.5	50.5		
Population (%)	(44.7–49.0)	(47.8–52.0)	(45.5–49.6)	(48.1–52.9)		
Working Age Population	76.2	73.3	75.7	73.4		
(16-64) (%)	(72.5–79.8)	(67.8–78.7)	(72.4–78.9)	(68.9–77.8)		
Median Age	33.5	30.2	33.0	31.5		
internation in the	(30.7–36.3)	(27.7–32.8)	(30.4–35.6)	(29.1–34.0)		
Bachelor Degree (%)	40.9	37.6	41.6	38.8		
High School Diploma	(33.7–48.1)	(26.0–49.2)	(34.8–48.5)	(29.6–47.9)		
High School Diploma	20.9 (19.0–23.7)	22.8 (17.5–28.1)	19.5 (16.5–22.6)	22.3 (18.2–26.4)		
(Highest) (%)	· · · · · · · · · · · · · · · · · · ·		31.9			
Married (%)	32.2 (27.6–36.9)	35.5 (29.6–41.5)	(27.3–36.4)	34.7 (29.5–40.4)		
	77.4	80.6	76.3	85.0		
Access to Vehicle (%)	(71.9–83.0)	(74.8–86.4)	(70.5-82.0)	(81.3-88.7)		
Distance (miles)						
	2.49	2.79	2.49	3.14		
To CBD	(1.83-3.15)	(2.19–3.38)	(1.83–3.15)	(2.59–3.70)		
To Lake Street	1.41	3.50	1.41	3.28		
10 Lune Street	(1.06-1.76)	(2.75–4.26)	(1.06–1.76)	(2.52–4.05)		
To Nearest Highway	0.31	0.52	0.31	0.63		
10 Treatest Highway (0.22-0.39) (0.31-0.72) (0.22-0.39) (0.44-0.83) 27 LRT Station Area Tracts are matched with 21 (23) tracts in the rest of Minneapolis for 2010 (2011)						

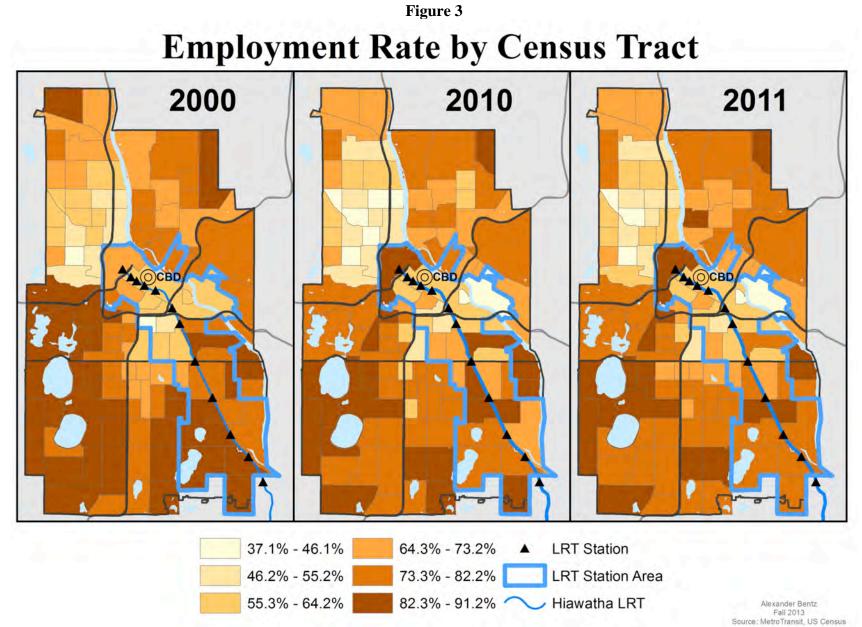
 Table 6: Propensity Score Matching

 Station Area and Control Cooper Magne and 05%

27 LRT Station Area Tracts are matched with 21 (23) tracts in the rest of Minneapolis for 2010 (2011) based on nonwhite population to form a weighted control group for each year. 95% Confidence Intervals for the means are shown in parenthesis.

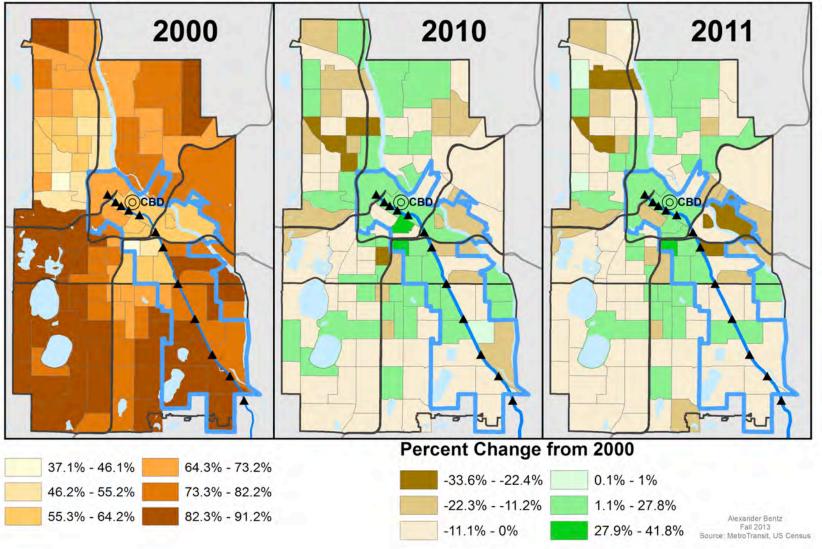


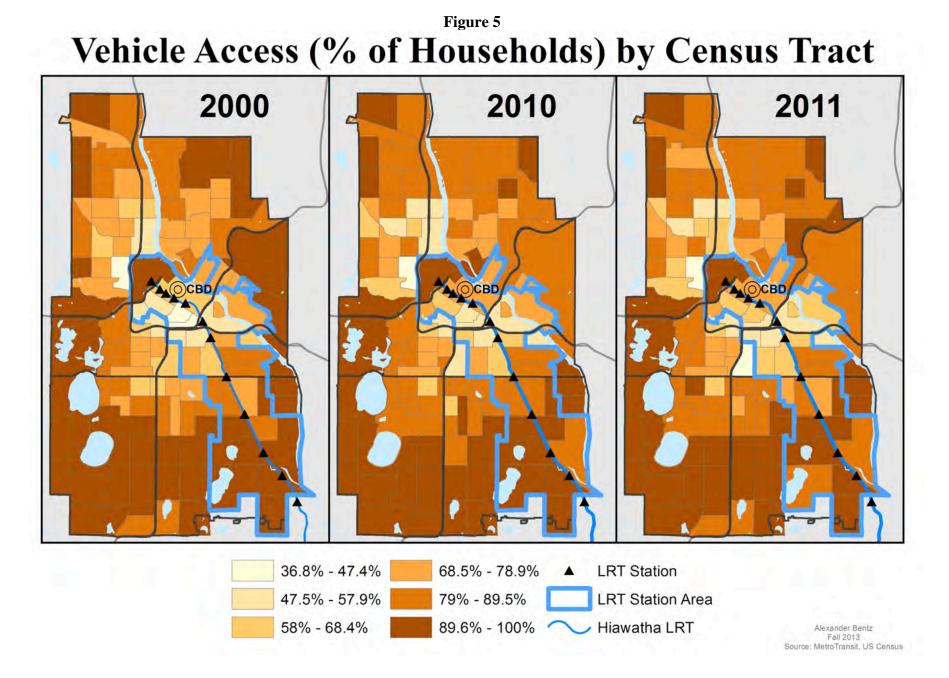


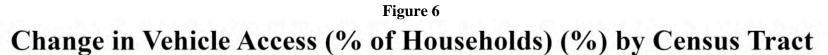




Change in Employment Rate (%) by Census Tract







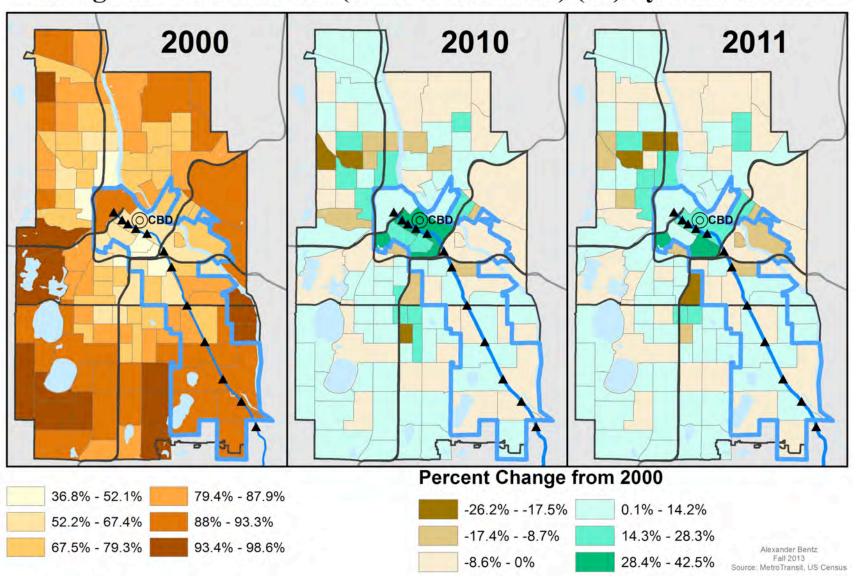
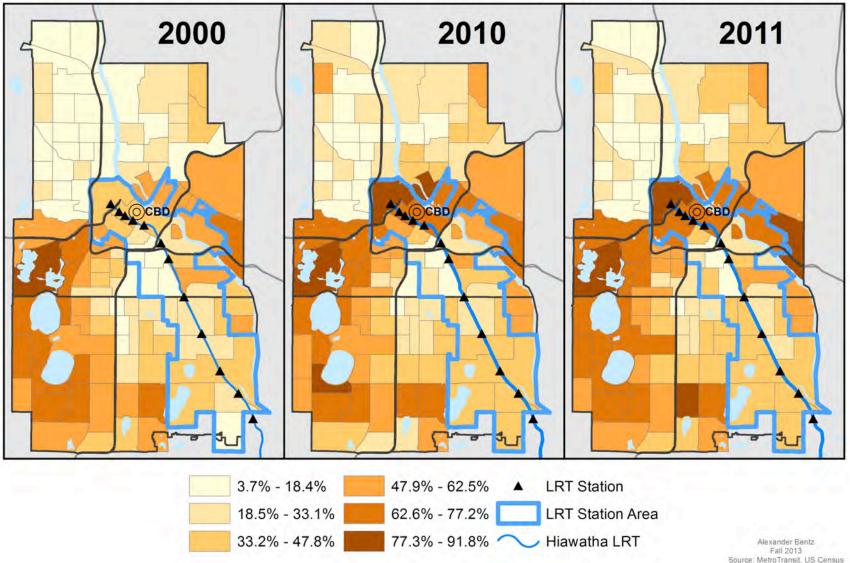


Figure 7

Population with Bachelor's Degree (%) by Census Tract



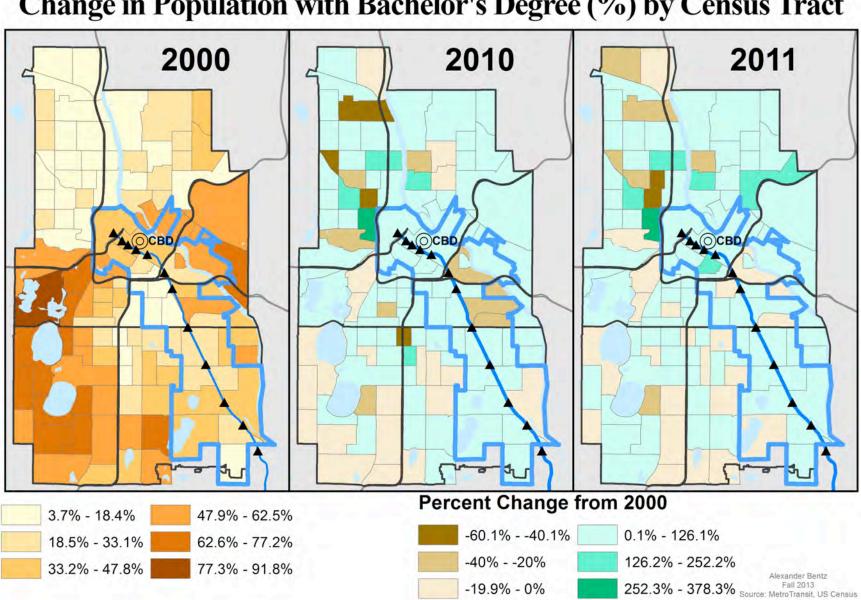


Figure 8 Change in Population with Bachelor's Degree (%) by Census Tract

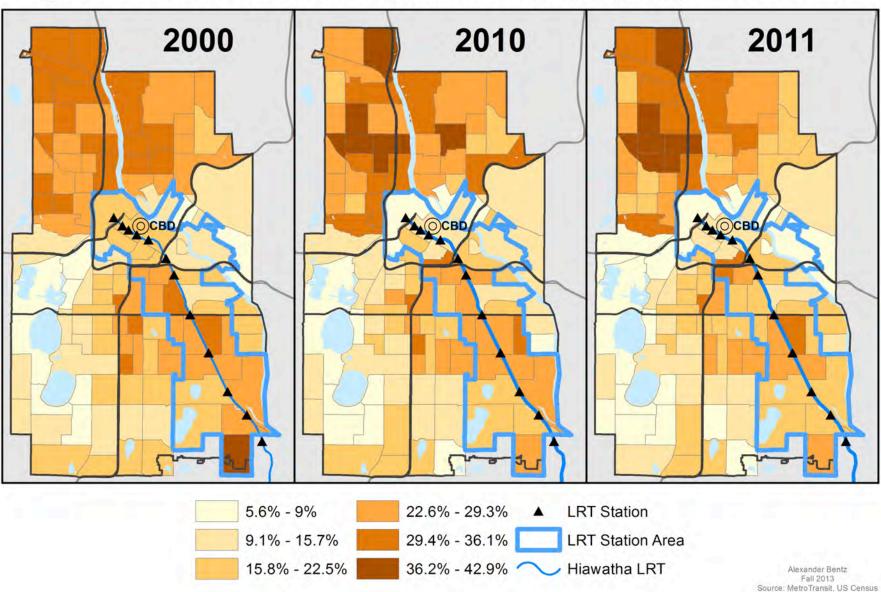
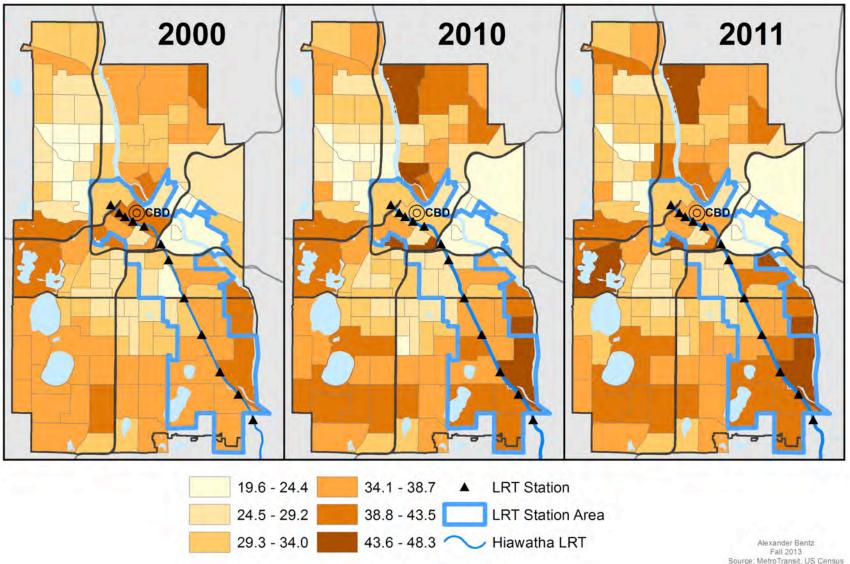
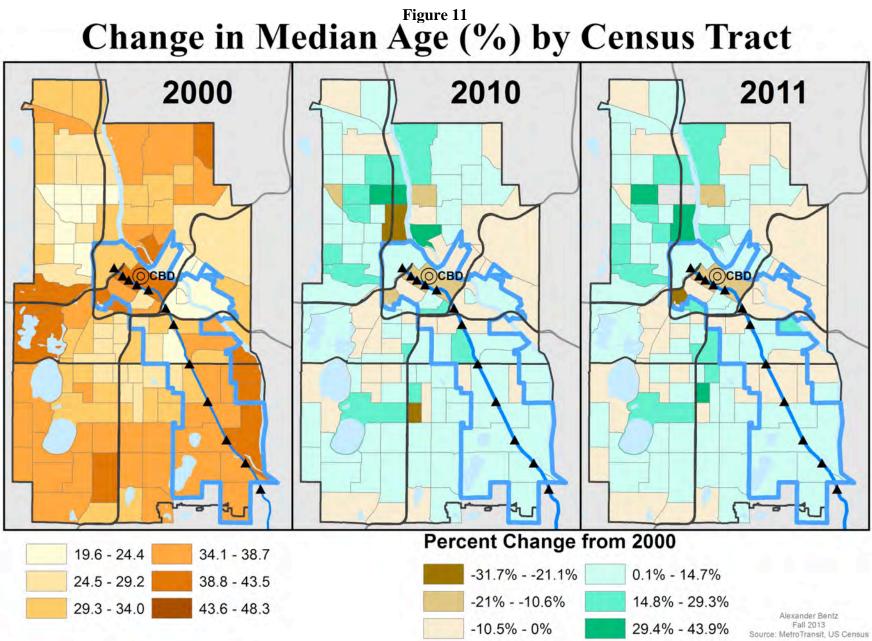
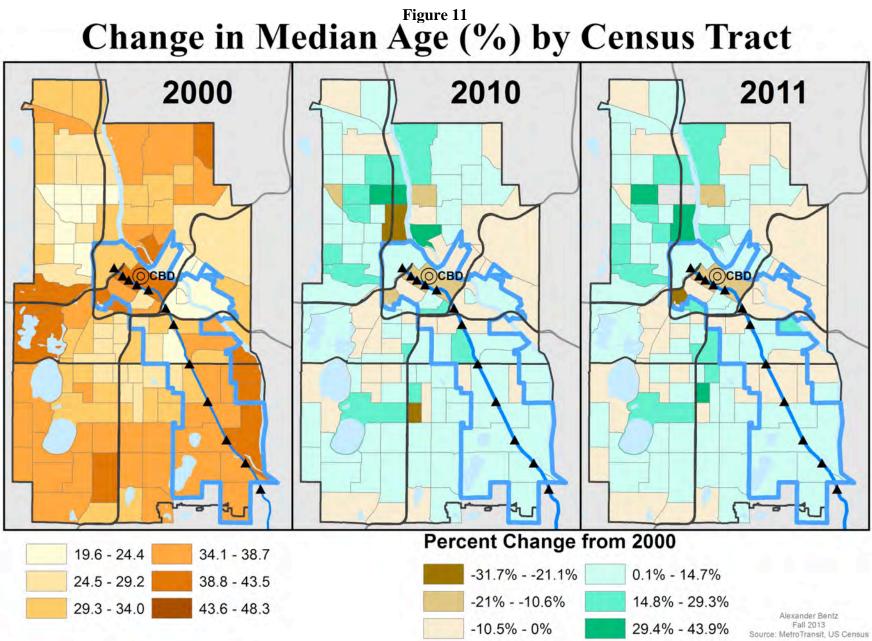


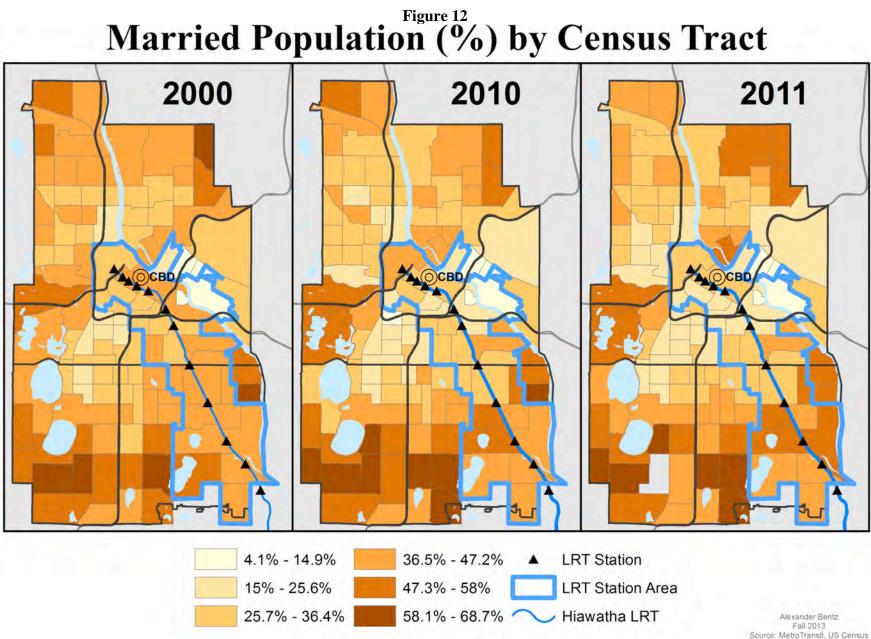
Figure 9 Population with High School Diploma (Highest) (%) by Census Tract

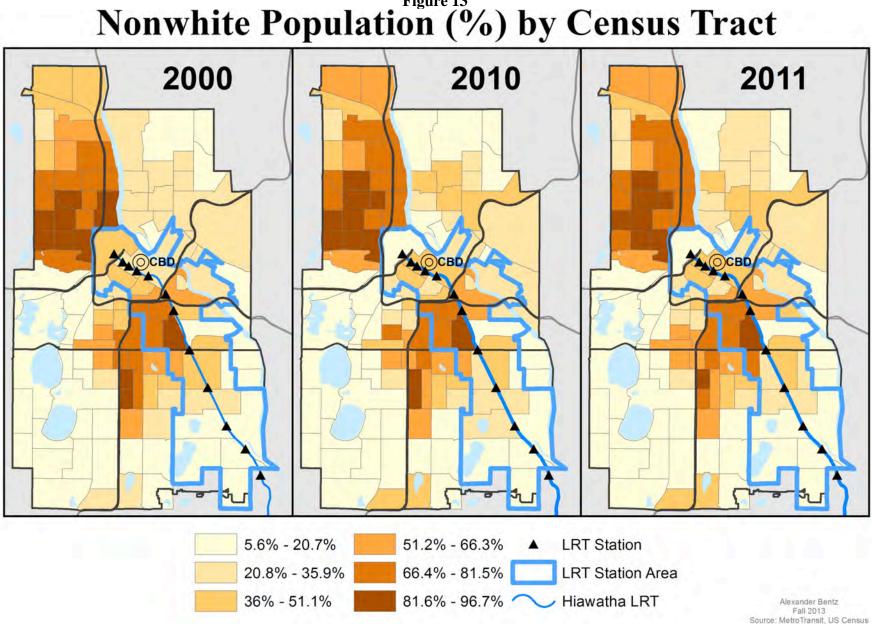












Nonwhite Population (%) by Census Tract

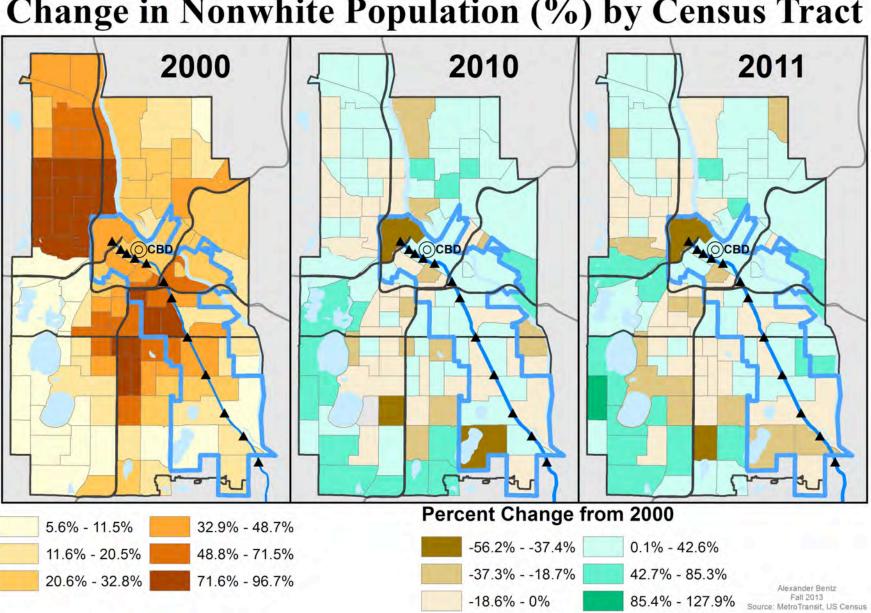
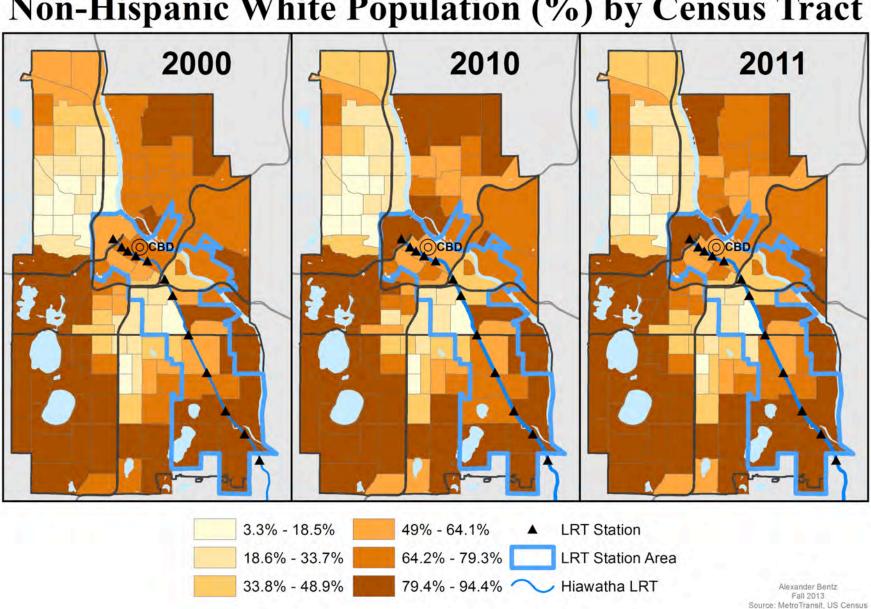


Figure 14 **Change in Nonwhite Population (%) by Census Tract**



Non-Hispanic White Population (%) by Census Tract

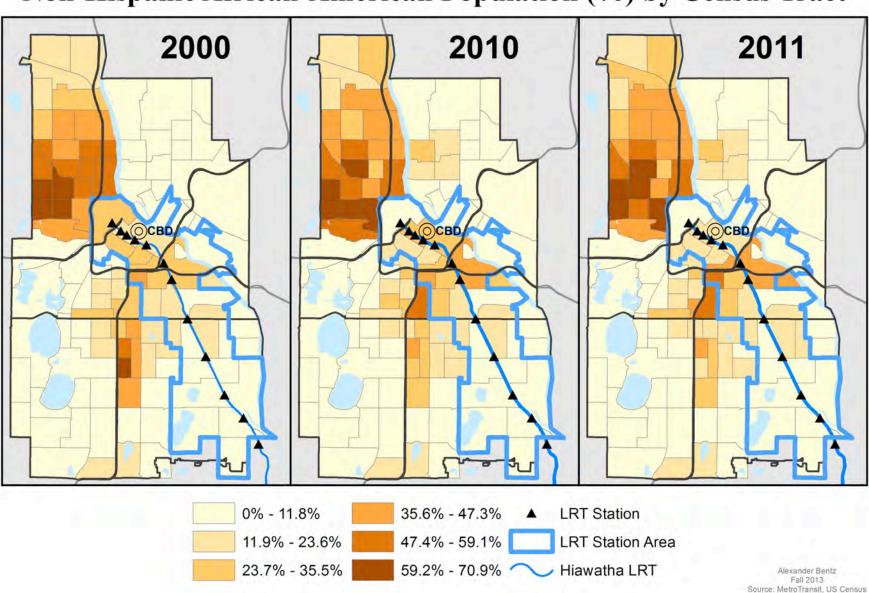
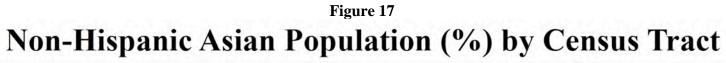
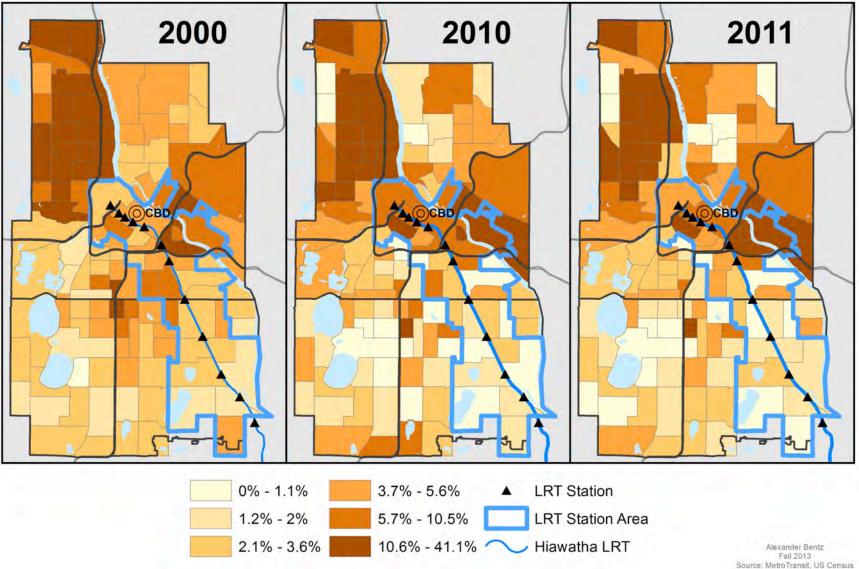
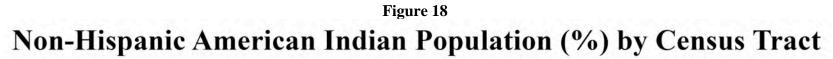
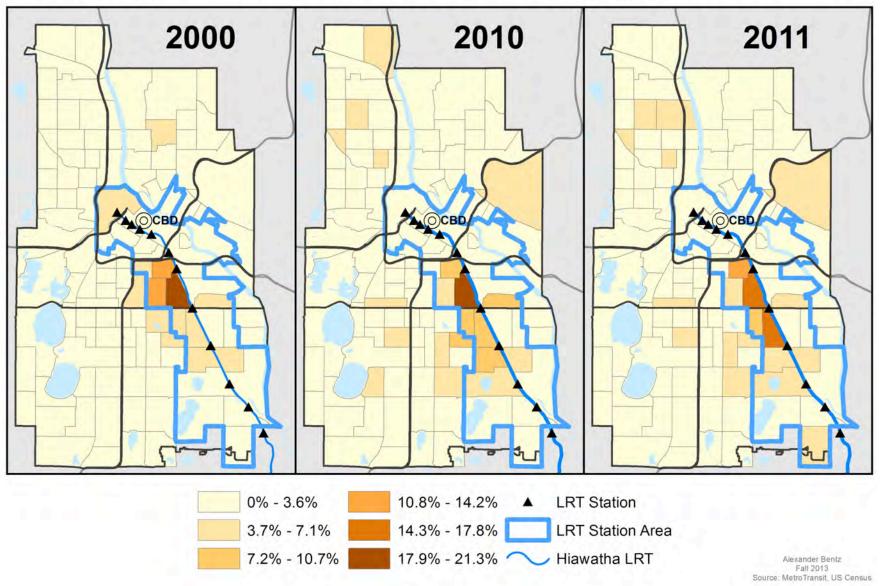


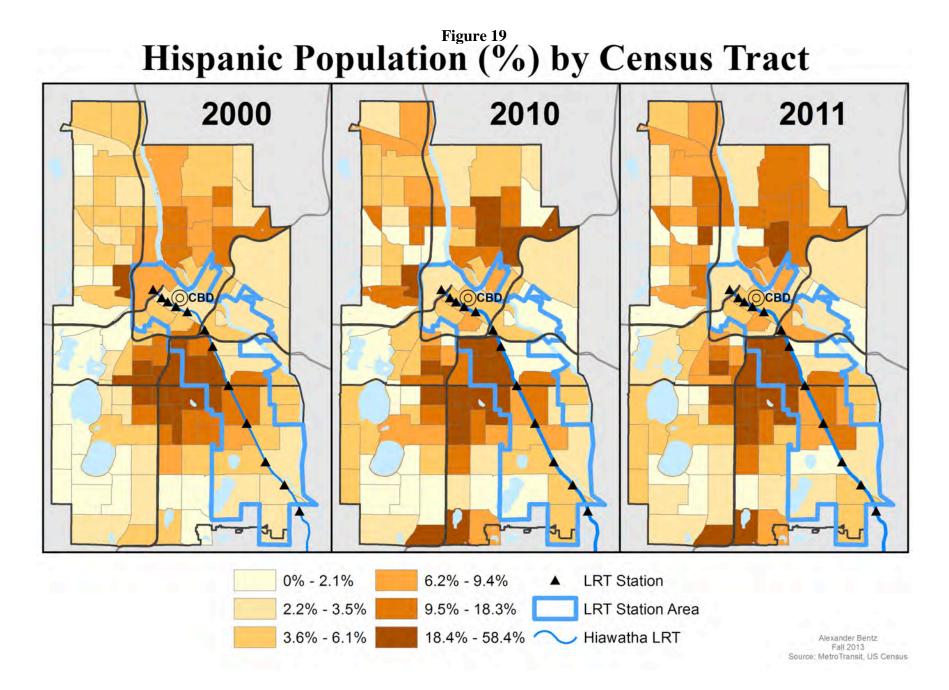
Figure 16 Non-Hispanic African-American Population (%) by Census Tract

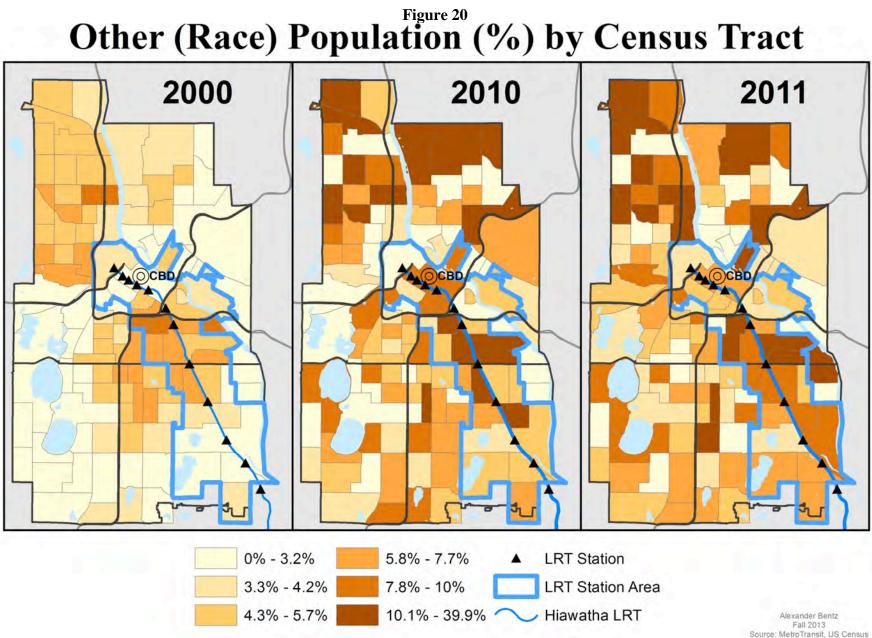








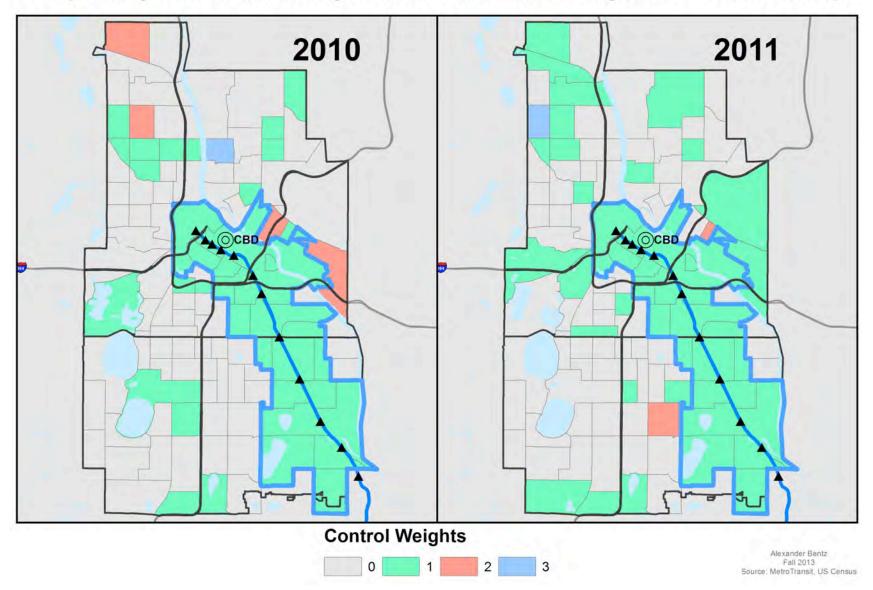




letro transit, US Census



Propensity Score Matching: Station Area and Weighted Control Group



References

- Alam, Bhuiyan M. (2009) "Transit Accessibility to Jobs and Employment Prospects of Welfare Recipients Without Cars: a Study of Broward County, Florida, Using Geographic Information Systems and an Econometric Model" *Transportation Research Record*. 40.2110: 78.
- Center for Transit Oriented Development (CTOD). (2011) "Rails to Real Estate: Development Patterns along Three New Transit Lines"
- Fan, Y. (2012) "The Planners' War against Spatial Mismatch: Lessons Learned and Ways Forward" *Journal of Planning Literature*. 27.2: 153-169.
- Goetz, Edward G., et al. (2010) "The Hiawatha Line: Impacts on Land Use and Residential Housing Value" Center for Transportation Studies–Humphrey Institute of Public Affairs. Minneapolis, MN.
- Holzer, Harry J. (1991) "The Spatial Mismatch Hypothesis: What Has the Evidence Shown?" *Urban Studies*. 28.1: 105-122.
- Hurst, Needham B. and West, Sarah E. "Public Transit and Urban Development: The Effect of Light Rail Transit on Land Use in Minneapolis, Minnesota" Macalester College. Draft obtained from author 25 September 2013.
- Ihlanfeldt, Keith R, and David L. Sjoquist. (1989) "The Impact of Job Decentralization on the Economic Welfare of Central City Blacks" *Journal of Urban Economics*. 26.1.
- Ihlanfeldt, K. and Sjoquist, D. (1990) "Job accessibility and racial differences in youth employment rates". *American Economic Review*, 80, pp. 267-276.
- Kain, J. F. (1968). "Housing segregation, Negro employment, and metropolitan decentralization" *Quarterly Journal of Economics*, 82, 175–197.
- Kawabata, Mizuki and Shen, Qing. (2007) "Commuting Inequality between Cars and Public Transit: the Case of the San Francisco Bay Area, 1990-2000. (Case Study)" *Urban Studies*. 44.9.
- Ko, K, and X.J Cao. (2013) "The Impact of Hiawatha Light Rail on Commercial and Industrial Property Values in Minneapolis." *Journal of Public Transportation*. 16.1: 47-66.
- Logan, John R., Zengwang Xu, and Brian Stults. (2012) "Interpolating US Decennial Census Tract Data from as Early as 1970 to 2010: A Longitudinal Tract Database" Professional Geographer, forthcoming.
- Matas, Anna, Josep-Lluis Raymond, and Josep-Lluis Roig. (2010) "Job Accessibility and Female Employment Probability: the Cases of Barcelona and Madrid." *Urban Studies*. 47.4: 769-787.

- Sanchez, Thomas W.T., Shen QIng, and Peng Zhong-Ren. (2004) "Transit Mobility, Jobs Access and Low-Income Labor Participation in Us Metropolitan Areas." *Urban Studies*. 41.7.
- Sanchez, Thomas W. (1999) "The Connection between Public Transit and Employment: the Cases of Portland and Atlanta" *Journal of the American Planning Association*. 65.3: 284.
- U.S. Census Bureau. 2000 Decennial Census. http://factfinder2.census.gov. 1 November, 2013.
- U.S. Census Bureau. American Community Survey, 2010 American Community Survey 5-Year Estimates. http://factfinder2.census.gov. 1 November, 2013.
- U.S. Census Bureau. American Community Survey, 2011 American Community Survey 5-Year Estimates. http://factfinder2.census.gov. 1 November, 2013.
- U.S. Census Bureau. 2013 TIGER/Line Shapefiles. http://factfinder2.census.gov. 1 November, 2013.
- Taylor, Brian D. (1995) "Spatial Mismatch or Automobile Mismatch? An Examination of Race, Residence and Commuting in US Metropolitan Areas" *Urban Studies*. 32.9: 1453-1473.