SPRAWL, MOBILITY GAPS, AND THE TRANSIT SOLUTION

Layla Thomas
Economics Department, Franklin & Marshall College
Honors Thesis

Thesis Advisor:
Tony Maynard, Economics Department

Defense Committee:
Patrick Fleming, Economics Department
Leanne Roncolato, Economics Department
Jerome Hodos, Sociology Department
Defense: May 2, 2017
Abstract

During the Industrial Revolution, American cities boomed as railways and ports anchored firms in bustling central business districts. As transportation technologies evolved, so did urban form. First, mass transportation lured away many high-income workers by providing transit lines to reach inner-city jobs. Later, the invention of the automobile dislodged both high-income labor and firms to spur the ongoing process of omnidirectional dispersion. To understand this process of urban evolution, technological and social change is interpreted in the context of government policy and economic theory. This paper argues for the expansion of current sprawl maturation models to account for the exclusion of low-income labor that, in turn, presents geographic and social barriers for low-income populations’ success in society and the workplace.

Through the aggregation of research, this study concludes with a discussion of public infrastructure as a potential solution to the isolation of low-income urban populations. Using an empirical test, I evaluate the efficacy of public transit infrastructure as a transportation supplement for low-income urban populations in the case of San Diego, California. Using data from the American Community Survey at the Zip Code Tabulation Area (ZCTA)-level alongside an original dataset of bus stop access, this research finds that labor force participation is positively correlated with the quantity of bus stops within 2 kilometers of that ZCTA. Using a simultaneous regression approach, I find that a statistically significant positive relationship between bus stop access and labor force participation remains even after accounting for the endogeneity of the bus stop access variable in the model.

Acknowledgements

First and foremost, I must thank Professor Tony Maynard whose lectures first brought me into the field of economics. He has since shaped both this thesis and my academic development, generally. Providing clear narratives, a process I’ll always remember as “herding cats,” would not have been possible without his insight, criticism, and (perhaps) futile attempts to suppress poetic leanings.

I also sincerely thank Professor Patrick Fleming who has dismissed my math anxiety with his fountain of advice regarding everything from statistics to career paths. I am very grateful for his guidance.

Further, this thesis could not be complete without a committee to critique it. I would like to thank Professor Jerome Hodos and Professor Leanne Roncolato for joining Professor Patrick Fleming as Defense Committee members.

I thank Franklin & Marshall College for the opportunity to work with such inspiring faculty members; for the Hackman Summer Research Scholarship that, with Professor Yeva Nersisyan, first exposed me to research; and for the Marshall Academic Enrichment Fellowship for funding the fieldwork that established the foundation of this report.

Finally, I would like to thank my family, who support each of my “crazy projects” as if the projects weren’t crazy at all; my friends, who provide everything from citations to distractions; and Alex, who never wavers in his generosity, support, and encouragement.
Table of Contents

Abstract ................................................................................................................................. 2
Acknowledgements .............................................................................................................. 2
I. Introduction ...................................................................................................................... 5
II. Economics of Urban Centralization: Considering Labor (1760s-1890s) ......................... 6
   A. Agglomeration Economies: The Mutual Benefits of Clustering ................................. 6
   B. The Strata of Labor: High-Income and Low-Income ............................................... 8
   C. Transportation: Centrifugal or Centripetal? ............................................................... 9
III. Policy for Decentralization, the Market Responds (Early 1900s) ............................... 11
   A. Federal Government and the Fear of Urban Ghettos ................................................. 11
   B. Technology Leads, Suburbanization Follows ............................................................ 12
   C. Decentralize the “Right” Way: Zoning to Keep Suburbs Suburban ........................... 13
   D. American Cities Meet the American Auto ............................................................... 14
IV. Decentralization: Unequal and Unstoppable? (Mid-1950s) ......................................... 15
   A. Firms and the Decision to (De)centralize ................................................................. 15
   B. Sprawl Maturation and Labor Decentralization ....................................................... 17
   C. Low-Income Excluded from Sprawl Maturation ...................................................... 18
V. Government Flip-flop: The Rush to Slow Sprawl (1960s-Present) .............................. 22
   A. Separate and Unequal No More: Riots for Employment? (1960s) ......................... 22
   B. Changes in Zoning Laws: Our Way or No Way (1960s-1970s) ............................. 23
   C. Reforming Welfare, Improving Mobility (1990s) .................................................... 24
VI. Policy for Economic Opportunity .................................................................................. 26
   A. Pro-Highway Advocates: Use What We Have ....................................................... 26
   B. Pro-Transit Advocates: Environment & Social Benefits Require Consideration ..... 27
VIII. An Empirical Analysis of Transit Access and Labor Force Participation .................. 30
   A. Introduction .............................................................................................................. 30
   B. Background ............................................................................................................. 31
   C. Data .......................................................................................................................... 32
   D. Methodology .......................................................................................................... 37
   E. Results ...................................................................................................................... 39
   F. Policy Implications and Limitations ......................................................................... 42
IX. Conclusion ................................................................................................................... 43
References ......................................................................................................................... 46
# Appendices

## Appendix A - History, Policy, and Trends

1. Why Buses? ................................. 52
2. Physical Address .................................. 52
3. Zip Code Tabulation Areas (ZCTAs) .............................. 53
4. Distances between Latitude-Longitude Coordinates .......................... 53

## Appendix B - Creating the Bus Stop Dataset

1. Why Buses? ................................. 52
2. Physical Address .................................. 52
3. Zip Code Tabulation Areas (ZCTAs) .............................. 53
4. Distances between Latitude-Longitude Coordinates .......................... 53

## Appendix C - Incorporated Place of San Diego City

1. Incorporated Places .................................. 54
2. Map of Area .................................... 54
3. Why San Diego? ................................... 55
4. Field Work ........................................ 56

## Appendix D - Incomplete Data

.................................................. 58

## Appendix E - Observed Bus Stop Regression

.................................................. 59

## Appendix F - Bus Stop Prediction Formula

1. Instrumental Variables to Predict Bus Stops .......................... 60
2. Two-Stage Regression .................................. 62
3. Three-Stage Regression .................................. 63
4. Justification of Control Variables .................................. 63

## Appendix G - Consistency of Findings

1. Age Categories .................................... 67
2. 2-Kilometer Radius and the Evaluation of Alternatives .......................... 68
I. Introduction

In the United States, cities have transformed from industrial to technological. In trendier cities, one might see factories and warehouses, once employing masses of blue collar workers, renovated into hip shopping plazas to attract white collar customers (Myers 2016). However, problems arise when not all artifacts of America’s industrial heyday can be so easily refurbished. This paper examines the costs faced by low-income populations as cities developed, grew, and sprawled, leaving a mobility gap in its wake.

Since the 1800s, the introduction of new transportation technologies has pulled growing proportions of, first, high-income laborers and, later, firms, into the suburbs\(^1\) (Glaeser, Kohlhase 2004). Dispersion, for a variety of reasons, rapidly grew into what we now know as urban sprawl (Glaeser, Kohlhase 2004; Glaeser, Kahn 2004). Due to financial limitations, unstable employment, and discriminatory zoning, low-income workers were excluded from the process of sprawl maturation (Moulding 2005; Levine 1990). Instead, they remained isolated in city centers. The result is a core of low-skill workers geographically isolated from the growing proportion of jobs flourishing beyond city limits (Shen 2001).

This paper will argue that the inequalities of sprawl, just like sprawl itself, are not the result of a momentous accident but, instead, the byproduct of shifting transportation technologies, government programs, and social preferences. The few targeted attempts to spur the suburbanization of low-income urban workers were met by a discriminatory “Not in My Backyard” mentality. The troubled populations should disperse, the public declared, just so long

---

\(^1\) Within the context of this paper, the term “decentralization” references movement *within metropolitan areas*. In other words, the decentralization is a regional movement from “inner city” to “suburban” and distinct from the national trend at the national level between “urban areas” and “rural areas.” This national trend has been one of centralization; increasing volumes of the rural population have been relocating to urban areas. As an example, the urban population in the United States has increased from 6% of the population in 1800 to 80.7% in 2010 (U.S. Census Bureau 1993; U.S. Census Bureau 2012). This trend is shown in Figure 1 of Appendix A.
as they aren’t moving into my neighborhood (Hanchett 2000). This protectionism denounced, diluted, and distorted low-income decentralization through discriminatory zoning and lending policies at both the local and federal level (Hanchett 2000).

Following its multistage evolution, sprawl and its associated mobility gap have become the new normal. To improve equity of opportunity, many advocates call for government policy to take a more active role. One prescription is to improve low-income labor’s mobility through the expansion of public transit infrastructure, which, particularly in the highway-centric United States, has been historically neglected (Moulding 2005). Complementing this historical analysis, the final section of this paper evaluates the statistical relationship between public transit access and labor force participation in modern-day San Diego, California. This study finds that public transit infrastructure positively influences labor force participation after accounting for endogeneity of the bus stop variable in the model, though the form policy applications will take remains a process of ongoing design and revision.

II. Economics of Urban Centralization: Considering Labor (1760s-1890s)
A. Agglomeration Economies: The Mutual Benefits of Clustering

From the earliest division of labor to the most extreme forms of modern specialization, trade economies have been inherently reliant on the continued exchange of goods, people, and ideas. In a world with widely dispersed buyers and sellers, transportation costs increase the costs of exchange. To reduce these expenses, firms and labor have historically centralized in cities to create interdependent networks of buyers and sellers all within close proximity to one another. These benefits are what economists call agglomeration economies (Small, Verhoef 2007).
Paul Krugman (1998) describes firms’ clustering in cities through the push-and-pull relationship of centripetal and centrifugal forces. In the (centripetal) formation of cities,² we will examine two primary forces: (1) market-size effects and (2) thick labor markets. Market-size effects are the benefits of firms being near other firms (Krugman 1998; Glaeser Kohlhase 2004). For example, without market-size effects, imagine a country with one baker and one mill. Each firm is highly dependent on the other and competitors aren’t accessible. A scaled market may be a larger city that supports ten bakers and five mills. In this setting, each firm would have alternatives and competitive incentives. Each would be less dependent on any one firm. Due to the economies of scale, increased infrastructure may become viable in this larger city that could boost productivity. For example, each mill may be able to contribute towards the creation of a road, directly connecting the country mills with the urban bakers.

“Thick labor markets,” the second driving centripetal force, describes the pooling of labor. When similar firms cluster, the skills of labor can easily be transferred from one firm to the next (Krugman 1998). The cost of training is only faced by the first employer while the benefits of that training are felt by any employer thereafter. In the early example of a country with one bakery and one mill, the mill worker who fights with his boss would have considerable training to undertake before he could be properly skilled to work in the bakery. With many similar firms, like the scaled ten bakery and five mill variant, this worker could seek employment, without retraining, at any of the other four mills. This benefits both firms and laborers. For the firms, this makes the recruitment of labor faster and simpler, as workers are easier to find and less likely to

² Krugman (1998) uses centripetal and centrifugal forces to describe the location and size of urban areas. Agglomeration would result when centripetal forces are stronger than opposing centrifugal forces (Krugman 1998). Considering the limited scope of this paper, I will not address the centrifugal components of Krugman’s theory as there is considerable overlap with the neoclassical “agglomeration economies” defined previously and what difference exist are non-essential to the argument at hand.
require retraining. While the ease of replacing laborers, many increase turnover, firm clustering increases labor’s ease of finding replacement work and minimize dependence on a single employer. In this way, labor also stands to benefit from following the clustering patterns of firms.

**B. The Strata of Labor: High-Income and Low-Income**

Glaeser and Kohlhase (2004) emphasize that it is the transportation of labor, both in work and in pleasure, that is central to the history of urban formation. With this in mind, the use of homogeneous “labor” in our discussion is an oversimplification. The aggregated “labor” must be segmented to consider the heterogeneity of labor and, by extension, the heterogeneity of the transportation used by labor. For clarity, this paper will concern itself between the broad strata of “low-income” and “high-income” which, in terms of transportation theory, are largely synonymous with “low-skill” and “high-skill.”

The tension between the low-income and high-income labor strata has played a leading role in guiding the form of urban society. Glaeser, Kahn, and Rappaport (2006) document as much in their research of London during the Industrial Revolution. As factory work urbanized, masses of low-income workers flocked to the city. As London became crowded, the high-income kept to themselves, surrounding parks and theatres. While distance from low-income labor was viewed as a luxury, complete separation was too expensive for most. Only the most elite individuals could afford to commute from estates outside the city by carriage (Glaeser et al. 2006).

---

3 The literature on transportation equity is inconsistent in its use of terms and classifications. Studies use groups like “poor” and “non-poor” (Glaeser, Kahn 2004; O’Regan, Quigley 1998); “less-educated” and “more-educated” (Shen 2001); “low-income” and “high-income” (Crane, Chatman 2003); “low socioeconomic levels” and “high socioeconomic levels” (Frias-Martinez et al. 2012); and “welfare-recipient” and “non-welfare recipient” (Sanchez, Schweitzer 2008; Clifton 2004) just to name a few. To aggregate the literature, I have opted to use broader terms.
London’s overcrowding was abated in the 1800s as first attempts at city planning improved sanitation while setting the groundwork for mass transportation (Gifford 2003). Within a few decades, omnibus lines spanned the city (Glaeser et al. 2006). The omnibus, Glaeser, Kahn, and Rappaport (2006) describe, created a lower-cost transportation alternative by utilizing the economies of scale inherent in mass transit. As a lower-cost substitute for the carriages of the elite, it captured untapped demand for city-suburb travel. Now, both the elite and high-income could afford transportation, increasing demand for transportation. Utilizing the new technologies, those who could afford transportation could now live within walking distance to a transit stop rather than within walking distance to their work. From an aerial view, decentralization encircled transit stops like spokes surround an axle. In short order, the high-income workers decentralized along the routes of omnibuses and rail lines to the outer edges of the city. The low-income remained largely stationary in their high-density housing surrounding center-city factories (Glaeser et al. 2006).

C. Transportation: Centrifugal or Centripetal?

Neoclassical economics and its “agglomeration economies,” suggest that a desire for lower transportation costs that incentivizes firm and labor centralization. In the case of London, decentralization occurred when changing technologies lowered these transportation costs further still (Glaeser, Kohlhase 2004). So, is transportation, to borrow Krugman’s (1998) nomenclature, centripetal or centrifugal? Economists argue it is both, as explained through trade-off decisions. That said, the field remains divided on which factors most influence this trade-off decision.

The popularized Alonso-Muth-Mills theory uses a monocentric city model to argue that decentralization is a result of a trade-off between transportation costs and land preferences (Mieszkowski, Mills 1993). The model relies on the assumptions that (1) land cost is inversely correlated with transportation cost, (2) land is a luxury good, and (3) land is more expensive the
closer one is to the central business district. This model suggests that, as incomes increase, individuals decide either to move further away from the city center or have lower transportation costs (Mieszkowski, Mills 1993). In London, the omnibus’ reduction of transportation cost permitted more people to exchange high land costs (and low transport costs) for lower land costs (and higher transport costs) (Glaeser et al. 2006). At least, that is the Alonso-Muth-Mills interpretation.

Other economists insist that demand for land is not enough to explain this trend. Sanchez (1999) proposes that lowered transportation costs effectively decreased the cost of acting upon high-income individuals’ discrimination against low-income individuals. This can be viewed as an application of Gary Becker’s (1957) taste for discrimination model. Many high-income workers are presumed to discriminate against low-income populations. Being near these low-income population created disutility, and so discriminatory high-income populations will spend more income than non-discriminators to increase their distance from low-income populations.

Whatever the cause, high-income workers’ suburbanization to the outer edges of London increased the density of inner-city low-income populations. For developed economies worldwide, disparities between low-skill workers, who were often minorities as well as low-income, and high-income workers, who were often white and high-income, became the norm both in residential location and lifestyle for developed economies worldwide (Glaeser et al. 2006). In the United States, as urban poverty rates continued to climb, the local and federal governments began to convert concern for urban stability into active policy goals (Glaeser et al. 2006).
III. Policy for Decentralization, the Market Responds (Early 1900s)

A. Federal Government and the Fear of Urban Ghettos

For most of the 20th century, United States policymakers openly encouraged the suburbanization of cities. The 1930s Great Depression only heightened federal concern about the sanitation, safety, and, most importantly, stability, of the nation’s urban ghettos. Beginning with the New Deal of the 1930s, funding for public housing, slum-clearing, and, later, rural investment, became high-visibility national priorities that would persist until the 1960s (Moulding 2005; Hanchett 2000).

While the New Deal attempted to push people out of city centers, Hanchett (2000) credits federal mortgage assistance as the pull that truly suburbanized the urban core. During World War II, Jackson (1985) explains, military necessities were prioritized over consumer goods while marriage and birth rates were on the rise. By the war’s end, there was a dire shortage of housing. To fill the supply gap, the Federal Housing Administration and the Veterans Administration subsidized home ownership through mortgage insurance and down payment assistance programs (Jackson 1985). After the burst of residential construction to follow, homeownership rates rapidly rose from 45 to 65 percent (Hanchett 2000).

Of course, the United States government did not forego an opportunity to influence the types of homes these loans could be used for: explicit preference for new suburban construction was made clear, Hanchett (2000) points out. Additionally, guidelines to ensure homogeneous (and blatantly discriminatory) lending were in place to “retain stability.” Restrictive covenants built into the construction of these properties restricted future uses and ensured that stability was maintained. Overall, these federal guidelines excluded female-led households for many years and
African Americans almost entirely. Direct policies to subsidize suburban housing were supplemented by indirect incentives, such as tax breaks for homeownership and construction, to guarantee that the average, white, nuclear American family could secure the “American Dream” of homeownership (Hanchett 2000).

The Federal Government’s initiatives to encourage non-urban growth did not stop at matters of housing; it also looked to highways. The 1950s saw the birth of the Interstate Highway System that, with an estimated final cost of $128.9 billion in 1991, would become the world’s most expensive public works project (DOT 2017; Hanchett 2000). The Interstate Highway System introduced almost 5,500 miles of freeways cutting directly through the heart of American Cities (DOT 2017; Hanchett 2000). This fiscal policy, true to the Keynesian economics prevalent at the time, provided a surge in investment while cementing roads, literally, into the iconic American landscape.

B. Technology Leads, Suburbanization Follows

In the mid-1900s, market innovations, namely the widespread production of the automobile, unintentionally coincided with federal decentralization initiatives. Before the automobile, one had to walk to a transit stop to reach shopping, school, and work (Glaeser, Kahn 2004). For this reason, suburbanization (both in terms of residential and commercial locations) was limited by one’s proximity to transit lines (Glaeser, Kahn 2004). With the introduction of the auto, transit access mattered less and the definition of “being nearby” changed, too (Glaeser, Kohlhase 2004). If proximity once meant living within a one-mile walk to work (a journey of, say, fifteen minutes) proximity was redefined as a ten-mile drive, also spanning fifteen minutes

---

4 This practice is now widely-known as “redlining,” in reference to the red lines that would denote areas where no investment was to be made (Fair Housing Center of Greater Boston n.d.).

5 $128.9 billion is the estimated total cost reported by the Department of Transportation in 1991 in 1989 price levels. Adjusted for the consumer-price index, this would equal approximately $253.23 billion in 2017 dollars.
(Glaeser, Kohlhase 2004). For this reason, the previous spoke-and-hub decentralization fractured into omnidirectional dispersion, as illustrated in Figure 1. Cars and trucks permitted flexible transportation over larger areas in ways transit expansion could not both for labor and the firms that employ them (Glaeser, Kahn 2004).

**Figure 1. Illustrating an urban dispersion process by the evolution of technology.**

This distinction -- labor decentralization versus labor and firm decentralization -- is key. Unlike the omnibus, Glaeser and Kahn (2004) describe, the automobile was marketable to both firms and labor. Throughout the earlier wave of urban decentralization, spurred by mass transit like the omnibus, labor was the beneficiary. Firms remained reliant on horse-drawn carts, railways, and ports for the receipt of raw materials and shipment of final goods. However, with the engine came the car and with the car came the truck. The truck provided a firm-friendly form of flexibly-routed transportation that required lower upfront investment than, say, railways, while providing superior service than the traditional horse and cart (Glaeser, Kahn 2004).

**C. Decentralize the “Right” Way: Zoning to Keep Suburbs Suburban**

While federal leadership and changing technologies smoothed the way for widespread urban decentralization, local leadership was concerned with the form this decentralization would
take. In the earlier era of small-scale suburbanization, informal self-policing ensured the “right” type of development occurred, describes Fischel (2003). Community pressure and political influence permitted only housing and “clean” industries, like retail, to be along the highly-travelled commuting routes. As suburbanization gained momentum, formalized policies were needed to maintain the separation between factories, garbage facilities, and high-density low-income housing that discriminatory high-income single-family homeowners sought to avoid (Fischel 2003).

In 1916, New York City pioneered the establishment of city-wide zoning laws, Fischel (2003) writes. Initially, zoning’s potential infringement on the constitutional right of private property deterred wide-scale adoption. This concern was dismissed in 1926 when the Euclid v. Ambler Supreme Court ruling upheld zoning’s constitutionality. The protection of residential areas through zoning quickly spread. By 1950, nearly all developed municipalities had zoning regulations. Zoning regulators viewed these measures as a form of protection for the community. Following the precedent of discrimination set by the Federal Housing Administration and Veterans Administration mortgage assistance loans, zoning regulations did not stop their “protection” with matters of noisy factories or ugly warehouses. Zoning often implicitly enforced neighborhood homogeneity in terms of income, class, and race by arguing for “selective growth” policies (Hanchett 2000; Fischel 2003).

D. American Cities Meet the American Auto

Just as labor’s use of the car diminished its reliance on mass transit, firms’ use of the truck diminished their dependence on infrastructure-heavy transportation methods, like railways and ports. Mass transit, railways, and ports are alike in their infrastructure-intensive design. These

---

6 In this case, the judge affirmed the need to protect single family housing from apartments that were, as he puts it, “...a mere parasite, constructed in order to take advantage of the open spaces and attractive surroundings created by the residential character of the district” (Euclid v. Ambler, 1926; Fischel 2003).
systems have high fixed costs of operation and so require large demand to be cost-effective. Automobiles provided lower-cost flexible alternatives that diminished this demand. For many industries, the purchase of trucks allowed access to the economies of agglomeration (supplies, labor, and necessary infrastructure) without the need to be located in the expensive city center (Glaeser, Kohlhase 2004; Glaeser, Kahn 2004). The extent to which the automobile could displace existing infrastructure depended, in large part, on the age of cities and the costs of transitioning to auto-friendly design.

In the case of older cities, the rise of personal vehicles was restrained by logistics. It was costly to widen streets and integrate highways and, as such, travel costs per mile for car owners remained relatively high (Glaeser et al. 2006). Instead, older cities grew taller, introducing skyscrapers to accommodate the rising population density (Glaeser, Kahn 2004). But what about younger American cities? In these places, car culture and urban form developed concurrently. Younger cities were designed with wide lanes for automobiles and with sparse public transportation. They are known today as “edge cities,” or “sprawl cities,” suggestive of the tumorous suburban growth that occurs around the edge of the city, unlike traditional “dense cities” (Glaeser, Kahn 2004).

IV. Decentralization: Unequal and Unstoppable? (Mid-1950s)

A. Firms and the Decision to (De)centralize

Urban decentralization did not affect all firms equally. Crane and Chatman (2003) theorize that the decision to decentralize relies, in large part, on industries’ labor-to-capital production ratios as well as their intended customers. Retail, for example, employs labor and relies on sales to labor. The theory predicts that retail would disperse in accordance with the population to allow for shorter commutes (lower cost of labor) and a larger pool of potential customers. In contrast, capital-intensive industries that sell to other firms, such as manufacturing, are unlikely to benefit
greatly from dispersion. The unique case of government, which does not prioritize cost-
minimization or competition in the same way as the private sector would not decentralize, either7 (Crane, Chatman 2003).

While the automobile and relocation of labor led many firms to depart from center cities, some high-skill industries, including finance, technologies, and research services, became more urbanized (Green 2012). Explanations for these industry-wide trends are imperfect. The popularized clustering theory created by Porter (1998) would suggest that firms will remain in dense regions to maximize entrepreneurial innovation, such as the case of Silicon Valley (cited in Atkinson 1998). That said, recent studies argue that innovation is more heavily influenced by virtual connections, not simply linkages within geographically-nearby networks (Cumbers, MacKinnon 2004). Empirical work, also defies Porter’s clustering theory: research and development-intensive industries are found to be more likely to suburbanize than other firms, established Atkinson (1998).

While the literature did not provide a definitive causal theory, the effect is quite clear: the job composition of the urban core, which is home to a disproportionate number of low-income and low-skill workers, is increasing inappropriate for its residents (Frias-Martinez et al. 2012). As finance and technology industries dominate city centers, there is a surge in high-skill employment for which existing residents are unqualified. In this manner, a “skill mismatch” develops (Green 2012; Shen 2001). As economist Thomas Sanchez puts it, it is “... not simply that central cities are devoid of employment opportunities; rather, the educational background and skills of central city residents are not suited for the jobs that they live near” (Sanchez 1999).

7 Since the time of Crane and Chatman’s work (2003) the United States General Services Administration has revised its guidelines for the location of new federal buildings. These guidelines now prioritize proximity to transit in the decision-making process (Tomer et al. 2011).
Many argue that this separation is a leading contributor to urban poverty, a theory now known as the Spatial Mismatch Hypothesis (Sanchez 1999; Heilmann 2014).

B. Sprawl Maturation and Labor Decentralization

During the initial phase of labor suburbanization, low-income labor was stationary as high-income labor utilized transportation technologies to relocate. The inconvenience to low-income labor was minimal thanks to the location stability of firms. Following the introduction of the automobile, when firms, too, began to decentralize, the low-income workers faced a real problem: How were they to get to their jobs?

Sultana and Weber (2014) use U.S. Census data to conclude that, in the long run, shorter commutes are a common benefit of urban sprawl. For example, as shown in Figure 2, if all of a city’s firms are in one central business district, the ring of labor surrounding this core will be thick, time-intensive to navigate, and expensive to live in due to high demand for a limited supply. The innermost gray ring, where commute times will be shortest, provides residency for a small percentage of the population. If firms divide into, say, three smaller business districts, the total supply of “inner ring” or “short commute” space increases. During the process of sprawl, multiple centers are created and, as time passes, labor relocates to live within these new inner rings, as part of the long-run adjustment process we’ll refer to as sprawl maturation (Sultana, Weber 2014). As labor has easier access to their place of employment, the average commute in the city, post-sprawl, has decreased compared to dense, non-sprawled cities once the new equilibrium is reached.

---

8 Sultana and Weber (2014) reference this maturation process through an analysis of how commuting times change over time, or a “commuting transition” to react to sprawl. I expand this commuting-centric conclusion to the broader process of sprawl and so adjust the terminology to be more a more general, “sprawl maturation.”

9 This trend has been strongest for larger metropolitan areas, though the authors believe this may be due to the limitations of Census data (Sultana, Weber 2014). Generally, this transition is also stronger for East and West coast cities (Sultana, Weber 2014). The closing section of this study will analyze the reasonably mature sprawl (no longer experiencing unusually rapid growth that requires adjustment) in the West coast city of San Diego, California.
C. Low-Income Excluded from Sprawl Maturation

By reverting to aggregate terms like “labor” and “average,” this theory is neat but not representative. The theory of Sultana and Weber (2014) relies on sprawl maturation for the transition to the new equilibrium. Maturation occurs when labor (1) utilize private alternatives to mass transit and (2) relocate to be nearer to their place of employment (Sultana, Weber 2014). I seek to expand Sultana and Weber’s (2014) theory to incorporate what happens if low-income labor is excluded from sprawl maturation. First, I will justify why low-income labor may not (1) utilize private alternatives to mass transit and/or (2) relocate to be nearer to their place of employment. Second, I will show the theoretical implications of this exclusion.

Sprawl, owing to its dispersive nature, is unable to sustain the economies of scale necessary for public transit infrastructure. Regardless of its economic feasibility for low-income workers, the ownership of a car is practically essential for the search and maintenance of employment in edge cities (O’Regan, Quigley 1998; Moulding 2005). In a nation where household vehicle expenditures are second only to that spent on housing itself, a circular paradox arises: Low-
income workers need a car to find and maintain employment but need employment to afford a car (Glaeser, Kahn 2004).  

To meet the requirements of life in a sprawled society, many low-income households will make extreme sacrifices to afford car ownership. This begins a reinforcing cycle of a high-cost and low-savings living. Struggling to afford autos alongside basic needs, low-income households with vehicles are highly susceptible to changes in gas prices or other maintenance costs (Moulding 2005). Financial barriers prevent (or at least deter) many low-income workers from private vehicle ownership. Figure 3 shows that, as of 2001, car ownership for those earning less than $25,000 annually had a ten times lower car ownership rate than those earning more than $25,000\(^{10}\) (BTS, n.d.).

![Figure 3](image.png)

**Figure 3.** The Bureau of Transportation Statistics found the preceding data using their 2001 National Household Travel Survey. From here, we can see that approximately 20% of households earning less than $25,000 are without vehicles versus approximately 2.5% for higher income households.

Relocation of labor to be closer to the newly-dispersed firms is the second component of sprawl maturation (Sultana, Weber 2014). Low-income workers are rationally unwilling to
relocate near their place of employment due to uncertainty of employment (Crane, Chatman 2003). Low-income workers often hold high-turnover positions (Holzer, Martinson 2006). As in any high-turnover occupation, it would be irrational to relocate to be closer to a job that one may lose at any time, Crane and Chatman (2003) argue. For these reasons, workers in high-turnover positions, especially those who are also low-income, must hedge their risk. They bear expensive and time-intensive commutes from a centrally-located residence to reach many dispersed jobs, rather than relocate to be closer to any one job. This strategy is observable in both low- and high-skill occupations, from fast food workers to visiting professors. In this manner, if one job is lost, a replacement employment is available without further relocation (Crane, Chatman 2003).

Even if low-income workers were willing to relocate, financial and societal restrictions ensure this is not an option for most (Crane, Chatman 2003). As mentioned previously, zoning regulators’ “selective growth” argument to prevent high-density (and low-cost) housing in suburbs was powerful (Fischel 2003). In fact, while maturing sprawl changes almost everything about suburban structure -- from transportation access to demographic composition-- the percentage of single-family dwellings is remarkably constant (Sultana, Weber 2014). This means, while land is cheaper outside the city, housing options can be prohibitively limited if one cannot afford a single-family home

So, lacking in ability, due to financial constraints, lacking in willingness, due to job insecurity, and lacking opportunity, due to discrimination, low-income populations forego opportunity during the process of sprawl maturation (Sultana, Weber 2014). Instead, they participate in “reverse-commuting,” Sanchez (2008a) writes. Reverse-commuting consists of an

---

11 That is not to say that no urban poor are relocating to the suburbs. In actuality, suburban poverty rates are growing faster than urban poverty rates, but urban poverty rates are still substantially higher than the suburban poverty rates (Allard, Paisner 2016). For further discussion, see Appendix A Figure 3.
urban resident commuting away from the central business district, often into the suburbs. These routes are usually only traveled by low-income labor, earning the nickname “poverty transportation” within the United States (Sanchez 2008; Cox et al. 2014). Resulting from this low ridership and, some argue, the limited means of the riders, these reverse-commuting routes have infrequent service and limited hours making the search for and commute to suburban employment more difficult still (Sanchez 2008; Shen 2000).

With this logical foundation for the exemption of low-income populations during sprawl maturation, I will now return to the theory of Sultana and Weber (2014). If a stratum of labor does not use private vehicles and/or move to the new locations of firms, they cannot, by definition, participate in maturation nor receive the benefits the maturation process. Instead, low-income populations remain centralized in the previously-thriving central business district, as shown in Figure 4. Sultana and Weber (2014) imply that the post-maturation multicentric equilibrium makes society better off by reducing commutes. I argue this multicentric form can create a less-than-ideal equilibrium. Maturation often excludes low-income populations, for whom commutes become longer, because of their exclusion from the process of sprawl. This theory is consistent with empirical findings of Levine (1990) that concluded that low-income populations tended to have longer commutes post-sprawl relative to high-income populations.
V. Government Flip-flop: The Rush to Slow Sprawl (1960s-Present)

A. Separate and Unequal No More: Riots for Employment? (1960s)

While government originally dreamed of urban decentralization to remedy overpopulation and associated problems in cities, federal opinion reversed in the 1960s following a series of civil riots. The most notable of these riots occurred in the Watts ghetto of Los Angeles in 1965 when the arrest of a drunk black driver was escalated by onlookers into an altercation involving the officer, the driver, and the driver’s mother, Fogelson (1967) recounts. This tense exchange spurred thousands of Watts residents to violently protest the structural and racial inequities of the ghetto. When the Watts Riots ended three day later, with 34 dead and almost 4,000 arrested, the governor of California sent a commission to investigate. The mostly-white commission pointed fingers at “irresponsible agitation by black leaders,” and an unrepresentative, unemployed, and uneducated black subgroup for the disturbance.\(^\text{12}\) The

\(^{12}\) Later analysts correct this claim. The riots were an action of the ghetto majority. These statements by the commission, along with its racial and political stances, were widely criticized following the release of the commission’s report (Fogelson 1967).
commission’s final report, among other things, strongly recommended creating policies to increase employment opportunities for all ghetto residents, regardless of race.

Two years later, in the summer of 1967, a series of over 164 uncoordinated riots broke out across the United States (Kerner Commission 1968). Much like the Watts Riots, the commission charged with investigation returned with calls for employment stimuli, such as transportation investment, primarily for urban blacks who did not have adequate job access (Fogelson 1967; Sanchez 2008). We can view this lack of employment as a side effect of the decentralization of firms and high-income labor, described above. Those left behind in the urban core did not have adequate access to appropriate employment.

The riots of 1965 and 1967 filled newspaper headlines and, at last, mustered enough public support to demand change (Sanchez 2008). In 1967, the Johnson administration established the cabinet-level Department of Transportation, the largest governmental restructuring since the National Security Act of 1947 (DOT 2017). The Department of Transportation was given the responsibility of managing the $150 million grant budget designated for reverse-commuting issues nationwide, a fund first established by the Urban Mass Transportation Act of 1964. (Adjusted for the Consumer Price Index, this would equal approximately $1.2 billion in 2017 dollars.) This transition marked the redefinition of transit from a failing private business venture to a necessary public work (Sanchez 2008). The government was beginning to recognize the costs of the suburbanization it once encouraged.

B. Changes in Zoning Laws: Our Way or No Way (1960s-1970s)

In light of the civil rights riots and several decisive court cases in the 1960s, Fischel (2003) writes, zoning regulation began to change, though reluctantly. Due to the pressure of the judicial branch, race- and income-based zoning discrimination, dubbed “selective growth,” were forbidden. (Fischel 2003) With selective growth struck down, zoning bodies grew desperate to
maintain protection for high-income (and mostly white) residential neighborhoods (Fischel 2003). Zoning bodies dedicated to maintaining their discriminatory practices required a new, legal, justification for their exclusionary practices. This justification came in the 1970s as environmentalists actively campaigned for “reduced growth” policies across the country to preserve natural ecosystems, Fischel (2003) writes. It is unlikely coincidental that zoning regulators began citing environmental concerns as justifications for denying certain construction around this same time (Fischel 2003). With more laws come more loopholes, it seems.

At the same time, United States’ post-World War II housing boom dramatically increased homeownership rates, particularly for the iconic, white, nuclear family. This group of homeowners were politically unified through their reverence for single-family homes. In short order, the single-family home was no longer just a desirable asset; it was a “national virtue” (Fischel 2003). (I would note how this specific terminology, “virtue” sits well within the context of American religiosity.) This mentality sowed seeds of empathy for those interested in maintaining residential discrimination. In the courts, judges, who were largely single-family homeowners themselves, empowered citizens to litigate against new development in their neighborhoods (Fischel 2003). If the zoning regulators couldn’t control the types of people new construction would attract, it seemed, zoning regulators would attempt to restrict all new construction.

C. Reforming Welfare, Improving Mobility (1990s)

After the 1960s, the equity of opportunity for low-income urban populations went largely unaddressed until the aftermath of the 1990s Savings & Loans Crisis. With stories of poverty flooding newsstands and the public once again unsettled, Washington began a comprehensive

---

13 The Savings & Loan Crisis was a minor financial crisis stemming from the banking sector, generally, and saving and loans associations, specifically. Its roots were in the 1970s and the crisis spanned the 1980s.
reform of the United States welfare system and, by extension, transportation policy (Sanchez 2008; WiredPen n.d.). The major emphasis of the lawmakers was to (1) decentralize welfare spending through the use of block grants and (2) decrease long-term welfare dependency (Sanchez 2008).

The process began, Sanchez (2008) continues, with the Intermodal Surface Transportation Efficiency Act of 1991 which promised to improve the mobility gap through, “intermodal connections between people and jobs, goods and markets, and neighborhoods.” This was followed by programs such as Welfare to Work and Temporary Assistance for Needy Families, which provided generous federal block grants to states with few restrictions. Many of these funds were spent on roadways and transit to improve job accessibility. In 1998, the Transportation Equity Act of the 21st Century provided a comprehensive transit-focused policy reform, which targeted job access and reverse-commuting programs (Sanchez 2008). This followed a General Accountability Office report in the same year that identified transportation barriers as a leading cause of unemployment (Sanchez 2008).

Throughout the first half of the paper, I have established that the changing transportation technologies stratified the mobility of labor based on income level. I determined that standard explanations of sprawl (as a long-term equilibrium following a period of maturation) fail to consider the costs to low-income labor excluded from sprawl maturation. Using historical evidence, I have illustrated how government policies and social preferences have intentionally, and unintentionally, reinforced the spatial segregation of metropolitan areas through lending and zoning practices. Finally, I’ve established that, though there has been federal policy aimed at resolving this issue, the mobility gap and associated opportunity gap remain.
VI. Policy for Economic Opportunity

A. Pro-Highway Advocates: Use What We Have

I have argued that labor strata’s reaction to changing transportation technologies influenced urban societies transition from centralized to sprawled. It’s impossible to undo technological innovation and, similarly, impossible to disconnect the automobile from the present American society. However, car-enabled sprawl leaves the involuntarily car-less geographically isolated. If ensuring equality of opportunity is a national priority, this mobility gap presents an opening for meaningful government intervention. That said, even those who would prefer intervention cannot reach a consensus on how to intervene.

Historically, the United States has invested more heavily in highway infrastructure than any other form of infrastructure (Aschauer, Campbell 1991; Moulding 2005). Mass transit advocates argue that if intensive investment, like that of the Interstate System, had been made into mass transit rather than highway infrastructure, transportation systems would better serve populations who cannot own or operate private vehicles (Moulding 2005). Yet what is done is done. Many of those concerned with urban immobility now believe policy should prioritize relocating jobs or utilizing the existing, flexible, car-centric infrastructure before discussion moves to expanding infrastructure, whether it be to transit or elsewhere (Sanchez, Schweitzer 2008; Moulding 2005; O’Regan, Quigley 1998). The mobility benefits of transit, Clifton (2004) proposes, are only felt in the long run and do not meet the immediate needs of isolated populations. Instead, advocates recommend subsidizing car purchases, subsidizing commuting costs, and permitting higher value of vehicle assets for welfare recipients (O’Regan, Quigley 1998; Martin 2001; Ong, Blumenberg 1998).

Further, some argue the fundamental structure of transit makes transit-users uncompetitive with car-owning populations. Studies have found that transit service limitations may largely
diminish any potential benefits of geographic transit expansion. Shen (2000) analyzed the determinants for job access in the city versus the suburbs. He concluded that transit provided increased access to job-rich areas but provided far less access than an automobile (Shen 2000). A study in Detroit, confirmed the superiority of car ownership for job-seekers as well as job-holders (O’Regan, Quigley 1998). The limited hours and routes of transit are often unusable for workers with non-peak-hour work shifts or parents who require flexibility to meet childcare needs (GAO Report 1998). If we really want to give the poor equal opportunity, advocates argue, we must give them equal tools for travel: automobiles, not inflexible mass transit (O’Regan, Quigley 1998).

B. Pro-Transit Advocates: Environment & Social Benefits Require Consideration

Another camp of pro-infrastructure reformers turn to transit, not highway transportation methods, as the solution. They argue that current infrastructure investment fails to adequately consider environmental costs and social benefits. For that reason, current investment disproportionately favors highway infrastructure (Sanchez 2008). Transit advocates believe greater funding should go towards transit once the returns on investment incorporate environmental and social goals.

The continuation of car dependence, with all the pavement, parking, fossil fuels use, and emissions that accompany it, is an environmentally, logistically, and economically unsustainable goal, argue public transit advocates (Gifford 2003; Sanchez, Schweitzer 2008; Moulding 2005). Highways seem to present a real-world micro-application of Say’s Law: regardless of supply, demand will rise to maintain a constant transportation shortage14 (Gifford 2003; CBO Report 2016). In this manner, a car-centric future is one bearing large environmental costs (from, say,

14 The additional demand triggered by increase supply is known as “induced demand.” The Congressional Budget Office report finds that the benefits of highway expansion, specifically, is generally non-existent within ten years of completion due to this “induced demand” (CBO Report 2016).
construction damage and emission pollution) with little long-standing benefits like diminishing congestion (Gifford 2003). These externalized costs of highway investment should be weighed more heavily into cost-benefit analyses when investment decisions are made, environmentally-minded advocates insist (Gifford 2003).

The benefits to human populations is also under-considered when investment decisions are made, transit advocates argue (Shen 2001; Sanchez, Schweitzer 2008). Mass transit, unlike highways, boosts aggregate demand for travel because transit has no barriers (other than a small fare) for use. There is no minimum level of physical ability, financial investment, or license necessary to use public transit. Transit provides transportation for those who would otherwise not have access such as the elderly, the young, and the disabled (Cox et al. 2014; Shen 2000; Shen 2001). Rather than validating transit investment as a means to an end (such as higher employment) transit should be viewed as an end itself as it provides valuable social assistance to otherwise immobile populations (Sanchez, Schweitzer 2008).

Finally, one must consider the economic side effects of this transit form decision. While both infrastructure investments will spend public funds to boost the nation’s aggregate demand—a policy choice demand-side economists following the Keynesian tradition might applaud—it is not to say that both types of spending are created equally. Research by Aschauer and Campbell (1991) find that, on a macroeconomic scale, transportation spending stimulates long-run growth almost twice as much as highway spending.15 While both have hefty construction costs (and so hefty employment during construction), I theorize this is in part due to public transit infrastructure’s provision of ongoing employment for day-to-day maintenance and operations. Apart from the occasional toll booth, the same cannot be said for roadways. For these reasons,

15 The work of Aschauer and Campbell (1991) was made in association with the American Public Transit Association.
transportation advocates demand more funding, research, and analysis be dedicated to the environmental, social, and macroeconomic benefits of mass transportation (Sanchez 2008).

**C. Empirical Evidence for Transit Boosting Economic Success: Inconclusive**

Limited work has examined the impact of public transit on economic outcomes. While the existing relevant literature provides excellent conceptual frameworks for discussion by use of logic (Mieszkowski, Mills 1993), intuition (Krugman 1998), or mathematical models (Chatman, Noland 2014), these a priori studies are difficult to apply to tangible policy strategies. While theories possess an elegant simplicity that is clear and illustrative, it is this same simplicity that limits application to a reality that is neither simple nor elegant (Uhlig 2012). For questions of policy, I posit that empirical work plays a leading role. Before discussing my own empirical contribution, I will provide a brief summary of the existing literature that precedes it.

The work of Glaeser, Kahn, and Rappaport (2006) uses a hybrid research structure mixing theoretical with empirical. They conclude that, rather than reducing poverty, transit may attract poverty by offering an ease of commutes for (often poor) non-car owners. After analyzing the location of the poor in 16 United States cities before and after rail expansions, they found that increasing rail access in regions led to increase in poverty rates at the census tract-level (Glaeser et al. 2006). Glaeser, Kahn, and Rappaport’s time-series analysis dismisses concerns of simultaneity and helps, in part, to justify public transportation’s colloquial name of “poverty transportation” in the United States (Cox et al. 2014; Sanchez 2008). Should this be the case, expanding transit could increase the costs to local governments; not only must governments fund transit, but they must then provide additional social services and support for areas with increased poverty rates (Allard, Paisner 2016).
In 1999, Thomas Sanchez analyzed the effect of private vehicle ownership and straight-line distance from transit stops on employment levels in Portland, Oregon and Atlanta, Georgia. Both rail and bus stations were evaluated. The endogeneity of private vehicle ownership and employment were accounted for using a two-stage least squares statistical regression model with instrumental variables. Using cross-sectional data, he found that employment levels decreased as transit stop distance increased and as vehicle ownership decreased (Sanchez 1999). This work suggests that transit access may help labor maintain employment. It does not, however, weigh in on populations who are currently unemployed and who, through increased transit access, may be more inclined to seek work. (To evaluate both employed labor and those seeking employment would require use of the labor force participation rate, as opposed to employment rate).

VIII. An Empirical Analysis of Transit Access and Labor Force Participation

A. Introduction

To add to the existing empirical literature on the effectiveness of public transit on employment metrics, the final portion of this research will analyze transit access and labor force participation in the city of San Diego, California using a unique dataset of bus stop access created for the purpose of this thesis. The resulting regressions find a statistically significant relationship between bus stop access (the number of bus stops within 2 kilometers from a Zip Code Tabulation Area (ZCTA) centroid) and labor force participation, even after accounting for potential endogeneity of the bus stop access variable. Findings using a two-stage least squares regression indicate that an increase in 1 accessible bus stop would be expected to increase labor force participation by 0.13% at a 5% significant level. Due to the relatively small size of this coefficient estimate, further work would be necessary to make a persuasive case for the cost effectiveness public transit service expansions as a labor force participation stimulant.
Before turning to a discussion of the data and empirical methodology, I will provide a brief background on the economic, social, and transit history of San Diego, which differs in certain ways from the general case of sprawl maturation described above.16

**B. Background**

As San Diego’s transportation technology evolved-- expanding from horse-drawn buses to automobiles-- the mass transit ownership passed through a series of private owners (San Diego MTS “History” 2017). The evolution of San Diego’s transit was slow and careful; a benefit of the monopoly-like control of transit in the early part of the 1900s was the assurance that transit was not over-supplied (San Diego MTS “History” 2017). The transit network entered public hands in 1967 when the City of San Diego purchased the transit system. This is in line with the national trend of transit privatization spanning from the 1950s to the 1970s17 (Pashigian 1976). Following a series of coverage expansions, ownership was transferred for the final time to the San Diego Metropolitan Development Board which presently maintains a 15-member Board, to manage the system and represent the interests of the county18 (San Diego MTS “History” 2017). This county-level integration should, theoretically, imply more connectivity between the urban and suburban area that are scattered across the county.

In terms of San Diego’s economy, they boast favorable unemployment of 4.2%, lower than both the state (5%) and national (4.5%) rates (Brady 2017). Job growth is at 1.3%, slightly

---

16 This cross-sectional analysis could be repeated in just about any city that has public transportation. San Diego was selected due to its manageable size, availability of data, and its limited modes of transit. For a further discussion of why San Diego was selected, see Appendix C.
17 It is unclear what spurred the wave of transit privatization in cities, those the most prevalent hypothesis point to regulation and low profits making private sector ownership unsustainable (Pashigian 1976).
18 The 15-member committee consists of: four representatives of the San Diego City Council, one representative from Chula Vista, Coronado, El Cajon, Imperial Beach, La Mesa, Lemon Grove, National City, Poway and Santee city councils, one representative from the San Diego County Board of Supervisors, and one San Diego County resident elected by other Board members to serve as Chairman.
higher than the national average (0.76%) (Brady 2017). In terms of San Diego’s job composition, Employment growth divided by sector shows that “Financial Activities,” “Leisure and Hospitality,” and “Education and Health Services” have seen the greatest year-over-year growth in 2016. This is shown below in Figure 5. Truck drivers, registered nurses, salespeople, and application developers were the highest growth occupations, according to Brady (2017).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Private</td>
<td>1,194.4</td>
<td>1,184.4</td>
<td>1,171.8</td>
<td>10.0</td>
<td>22.6</td>
<td>0.8%</td>
<td>1.9%</td>
</tr>
<tr>
<td>Professional and Business Services</td>
<td>2395</td>
<td>239.4</td>
<td>234.0</td>
<td>0.1</td>
<td>5.5</td>
<td>0.0%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Prof., Scientific, and Tech. Services*</td>
<td>131.2</td>
<td>129.3</td>
<td>127.4</td>
<td>1.9</td>
<td>3.8</td>
<td>1.5%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Mgmt. of Companies and Enterprises</td>
<td>229</td>
<td>22.7</td>
<td>22.4</td>
<td>0.2</td>
<td>0.3</td>
<td>0.9%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Administrative</td>
<td>854</td>
<td>87.4</td>
<td>542</td>
<td>(2.0)</td>
<td>1.2</td>
<td>-2.3%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Trade, Transportation, and Utilities</td>
<td>231.3</td>
<td>200.3</td>
<td>226.8</td>
<td>11.0</td>
<td>2.5</td>
<td>0.0%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Wholesale Trade</td>
<td>46.1</td>
<td>45.4</td>
<td>44.0</td>
<td>0.7</td>
<td>2.1</td>
<td>1.5%</td>
<td>4.5%</td>
</tr>
<tr>
<td>Retail Trade</td>
<td>155.2</td>
<td>146.1</td>
<td>155.3</td>
<td>9.2</td>
<td>-</td>
<td>6.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Transportation and Utilities</td>
<td>299</td>
<td>28.8</td>
<td>255.5</td>
<td>1.1</td>
<td>0.4</td>
<td>3.6%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Education and Health Services</td>
<td>294.9</td>
<td>201.4</td>
<td>195.2</td>
<td>5.5</td>
<td>5.7</td>
<td>1.7%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Education Services</td>
<td>31.2</td>
<td>29.6</td>
<td>30.4</td>
<td>1.6</td>
<td>0.8</td>
<td>5.4%</td>
<td>2.6%</td>
</tr>
<tr>
<td>Healthcare and Social Assistance</td>
<td>173.7</td>
<td>171.8</td>
<td>168.8</td>
<td>1.9</td>
<td>4.9</td>
<td>1.1%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Leisure and Hospitality*</td>
<td>187.8</td>
<td>194.3</td>
<td>182.4</td>
<td>(6.5)</td>
<td>5.4</td>
<td>-3.3%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Manufacturing*</td>
<td>195.1</td>
<td>195.7</td>
<td>196.7</td>
<td>(0.6)</td>
<td>(1.6)</td>
<td>-0.6%</td>
<td>-1.5%</td>
</tr>
<tr>
<td>Financial Activities</td>
<td>75.1</td>
<td>73.3</td>
<td>72.9</td>
<td>1.8</td>
<td>2.2</td>
<td>2.5%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Finance and Insurance</td>
<td>45.9</td>
<td>44.9</td>
<td>45.0</td>
<td>0.6</td>
<td>0.3</td>
<td>1.3%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Real Estate and Rental and Leasing</td>
<td>296</td>
<td>28.4</td>
<td>27.9</td>
<td>1.2</td>
<td>1.7</td>
<td>4.2%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Construction</td>
<td>71.8</td>
<td>71.9</td>
<td>71.4</td>
<td>(0.1)</td>
<td>0.4</td>
<td>-0.1%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Other Services</td>
<td>54.3</td>
<td>53.7</td>
<td>52.2</td>
<td>0.6</td>
<td>2.1</td>
<td>1.1%</td>
<td>4.0%</td>
</tr>
<tr>
<td>Information*</td>
<td>242</td>
<td>240.0</td>
<td>238.8</td>
<td>0.2</td>
<td>0.4</td>
<td>0.8%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Mining and Logging</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>-</td>
<td>-</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Total Government</td>
<td>246.9</td>
<td>238.4</td>
<td>240.6</td>
<td>8.5</td>
<td>5.3</td>
<td>3.6%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Federal Government</td>
<td>47.2</td>
<td>46.7</td>
<td>46.7</td>
<td>0.5</td>
<td>0.3</td>
<td>1.1%</td>
<td>1.1%</td>
</tr>
<tr>
<td>State Government</td>
<td>48.7</td>
<td>45.3</td>
<td>47.0</td>
<td>3.4</td>
<td>1.7</td>
<td>7.5%</td>
<td>3.6%</td>
</tr>
</tbody>
</table>

Source: Bureau of Labor Statistics
Notes: Italics denote sectors, asterisks (*) denotes sectors strongly associated with San Diego's traded economies. Quarter-end monthly data.

Figure 5. San Diego sees steady job growth, especially in service sectors (Brady 2017).

C. Data
The effect of public transit on employment opportunities in San Diego is analyzed using the number of accessible bus stops alongside a series of independent sociodemographic variables from the U.S. Census Bureau’s 2013 and 2014 American Community Survey 5-Year Estimates. These two datasets were linked at the level of Zip Code Tabulation Area (ZCTA). ZCTAs are 5-digit federal identifiers similar, but not identical, to United States Postal Service’s ZIP codes. This study limits itself to the 50 ZCTAs that are registered in the State of California under the
Incorporated Place of San Diego City that also have Census-provided demographic data available.\textsuperscript{19}

The primary explanatory variable in this study is access to public bus stops. This variable represents the \textit{quantity} of bus stops within a 2-kilometer radius from the center of each ZCTA.\textsuperscript{20} This variable was created through mapping the straight-line distance between the latitude and longitude of each of the 2,443 unique bus stops listed on the San Diego Transit Authority’s website and the centroid of each of the ZCTAs. Distances were then summarized by counting the number of bus stops falling within a 2-kilometer threshold of each ZCTA centroid. In order to test for robustness this methodology was repeated at several radiuses, as shown in Appendix G.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
Radius & Minimum & 1st Q & Mean & Median & 3rd Q & Max \\
\hline
2 km & 0.00 & 4.00 & 27.84 & 27.00 & 42.25 & 104.00 \\
1.5 km & 0.00 & 3.00 & 17.20 & 14.50 & 25.75 & 72.00 \\
1 km & 0.00 & 0.00 & 7.92 & 5.00 & 14.25 & 34.00 \\
0.75 km & 0.00 & 0.00 & 4.56 & 3.00 & 6.00 & 25.00 \\
0.25 km & 0.00 & 6.25 & 39.88 & 40.00 & 65.00 & 122.00 \\
\hline
\end{tabular}
\caption{Descriptive Statistics for Bus Stop Variable}
\end{table}

\textbf{Figure 6.} This table outlines the descriptive statistics of the bus stop variables using different radiuses. For example, on average, 7.92 bus stops are located within 1 kilometer of the ZCTAs evaluated.

The response variable, “labor force participation rate,” is reported by the \textit{American Community Survey} as a percentage of civilians over the age of 16 that are in the labor force (U.S. Census 2014). The U.S. Census measure of “in the labor force” includes employed and unemployed civilians as well as all active members of the United States Armed Forces.\textsuperscript{21} It excludes students, homemakers, prisoners, or institutionalized persons. It also excludes any person not actively looking for work, such as discouraged workers (U.S. Census 2014).

\textsuperscript{19} This excludes the five ZCTAs with incomplete data. For more information, see Appendix D.

\textsuperscript{21} Labor force is defined by the U.S. Census as the number employed and unemployed civilians \textit{as well as} members of the military (U.S. Census 2013; U.S. Census 2014). This is distinct from the Bureau of Labor Statistics, where in the labor force is derived from only the number of employed and unemployed \textit{civilians} (Bureau of Labor Statistics, “Glossary,” n.d.). The sociodemographic data for control variables was provided by the \textit{American Community Survey} created by the Census Bureau and for this reason the Census Bureau’s definition and dataset for labor force participation was used, rather than the Bureau of Labor Statistics’ version.
The regression analysis also includes a number of control variables from the U.S. Census American Community Survey (U.S. Census 2013; U.S. Census 2014). These variables are defined as follows (further discussion of variables and the justifications for their inclusion can be found in Appendix G). Table of descriptive statistics is shown below in Figure 7 and Figure 8.

<table>
<thead>
<tr>
<th>Explanatory</th>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus Stops within 2 Kilometers of Centroid</td>
<td>0.00</td>
<td>4.00</td>
<td>27.00</td>
<td>27.84</td>
<td>42.25</td>
<td>104.00</td>
<td>26.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response</th>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor Force Participation Rate</td>
<td>42.10</td>
<td>62.85</td>
<td>65.50</td>
<td>66.19</td>
<td>70.08</td>
<td>79.20</td>
<td>5.94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control</th>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Earnings</td>
<td>20037.00</td>
<td>30464.25</td>
<td>38524.00</td>
<td>41024.76</td>
<td>46252.50</td>
<td>104645.00</td>
<td>16556.75</td>
</tr>
<tr>
<td>Total Housing Occupancy Rate</td>
<td>78.10</td>
<td>91.50</td>
<td>93.15</td>
<td>92.11</td>
<td>95.70</td>
<td>97.10</td>
<td>4.54</td>
</tr>
<tr>
<td>Owner Vacancy Rate</td>
<td>0.00</td>
<td>1.10</td>
<td>1.40</td>
<td>1.83</td>
<td>2.40</td>
<td>8.80</td>
<td>1.49</td>
</tr>
<tr>
<td>Rental Vacancy Rate</td>
<td>0.00</td>
<td>2.35</td>
<td>3.75</td>
<td>3.87</td>
<td>5.25</td>
<td>8.10</td>
<td>1.86</td>
</tr>
<tr>
<td>Median Age</td>
<td>25.80</td>
<td>31.65</td>
<td>35.50</td>
<td>35.86</td>
<td>38.65</td>
<td>50.30</td>
<td>5.37</td>
</tr>
<tr>
<td>Work-at-Home Workers</td>
<td>12.00</td>
<td>69.75</td>
<td>97.50</td>
<td>108.76</td>
<td>141.75</td>
<td>291.00</td>
<td>61.91</td>
</tr>
<tr>
<td>Gross Rent</td>
<td>102.00</td>
<td>3199.00</td>
<td>5732.50</td>
<td>6553.32</td>
<td>9017.50</td>
<td>15201.00</td>
<td>4128.26</td>
</tr>
<tr>
<td>Households with 0 Vehicles</td>
<td>8.00</td>
<td>338.25</td>
<td>733.50</td>
<td>947.96</td>
<td>1196.75</td>
<td>4454.00</td>
<td>874.21</td>
</tr>
<tr>
<td>Proportion White</td>
<td>0.23</td>
<td>0.59</td>
<td>0.73</td>
<td>0.70</td>
<td>0.83</td>
<td>0.92</td>
<td>0.16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instrumental</th>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees Working in ZCTA</td>
<td>549.00</td>
<td>5316.25</td>
<td>9688.00</td>
<td>17177.72</td>
<td>20816.00</td>
<td>93155.00</td>
<td>19477.77</td>
</tr>
<tr>
<td>Population Density</td>
<td>214.40</td>
<td>2504.58</td>
<td>3808.25</td>
<td>4897.84</td>
<td>7672.00</td>
<td>12533.50</td>
<td>3293.55</td>
</tr>
</tbody>
</table>

*Figure 7. Descriptive statistics grouped by variable.*
<table>
<thead>
<tr>
<th>Case</th>
<th>Explanatory</th>
<th>Response</th>
<th>Control</th>
<th>Instrumental</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91215.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91214.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91225.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91224.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91213.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91212.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91211.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 8. Summary of data grouped by ZCTA.
**Natural Log of Median Earnings** is the median of annual income received by individuals over 16 before deductions, grouped by households, within each ZCTA. The data is provided in dollars.

**Total Housing Occupancy Rates** the proportion of all available housing (owner- and tenant-occupying) that is occupied. That is, not “for sale,” “for rent,” or abandoned. The data is provided as a proportion.

**Owner Vacancy Rates (Owner Lives in)** the proportion of homeowner housing units that are listed “for sale” divided by total homeowner housing units. The data is provided as a proportion.

**Rental Vacancy Rates (Tenant Lives in)** the proportion of rental housing units that are listed “for rent” over total rental housing units. The data is provided as a percentage.

**Median Age Squared** the median age of each person, surveyed by household. The data is provided as number of years.

**Work-At-Home Workers** the number of workers who primarily worked from home. This data is provided as a number of workers.

**Gross Rent** the monthly housing costs for renters, grouped by housing unit, including utilities and fuels. The data is provided in dollar values.

**Quantity of Households without Vehicles** the number of households with access to zero personal vehicles. The data is provided as number of households.

**Proportion of Population White** the proportion of people who identified as “white,” Irish, German, Italian, Lebanese, Arab, Moroccan, or Caucasian. This data is provided as a proportion of the population over 16.

For clarity, these factors were grouped into the variable “Control,” for the initial equation.
**D. Methodology**

To estimate the effect of access to public transit (as proxied by the quantity of bus stops) on labor force participation, I use single- and multi-stage regressions. The initial equation used is defined as follows and estimated a simple ordinary least squares regression model:

**Equation 1.** \( \text{LaborForceParticipation}_i = \beta_0 + \beta_1 \text{BusStops}_{1i} + \beta_2 \text{Control}_{2i} + \text{Error}_i \)

The results of this initial equation are detailed in Appendix E. Note that the coefficient \( \beta_1 \) represents the estimated effect of one additional bus stop within 2 kilometers of a ZCTA on the labor force participation rate within that ZCTA. However, with this initial formula, there is a concern of endogeneity of the bus stop variable. While the formula attempts to use bus stops \( (x) \) to predict labor force participation \( (y) \), without resolving this endogeneity, it could be the case that labor force participation itself determines the quantity of bus stops established near a ZCTA. For example, areas with low labor force participation may not need transportation in the eyes of city planners, as they have less a need to travel to work. With this initial formula, it is impossible to determine whether labor force participation influences bus stops or whether bus stops influence labor force participation, which makes this first model more or less useless from a policy perspective.

Following Cameron and Trivedi (2005), I use a two-stage and three-stage least square to unravel this loop of circularity. Two-stage least squares (2SLS) uses instrumental variables to estimate a predicted bus stop variable in a first-stage equation that holds the same controls as Equation 1 while adding in a set of instrumental variables. The instrumental variables are expected influence the number of bus stops but would not be expected to influence labor force participation (except through the bus stop variable). More information on instrument variable estimation is contained in Appendix F. Note that this is the same empirical approach taken by Sanchez (1999).
The instrumental variables used in this study are, like the control variables, from the American Community Survey and are as follows:

*Number of Employees Working* the number of workers who work in-- not necessarily live--the ZCTA. Data is provided as the number of workers.

*Population density:* the population per square mile of land area. Data is provided as the number of people per square mile.

Logically, I believe that the number of employees would likely affect bus stop access because an increasing quantity of potential riders would increase the demand for and economies of scale of transit. This may then determine where new transportation was installed. It would not impact labor force participation because changing quantities of people should not impact a rate which our labor force measure is. Like the number of employees, I expected population density to impact bus stops because it indicates greater potential demand for occupational and recreational travel. It should not impact labor force participation rate because, again, rates should remain largely independent of any changes in quantities.

Empirically, an ordinary least squares (OLS) regressions confirmed that the instrumental variables were both correlated with bus stop access at a 1% level while being independent of labor force participation, except through the bus stop variable. The full justifications for these instrumental variables are shown in Appendix F.

Accordingly, labor force participation in ZCTA \( i \) is estimated in a 2SLS model as follows, with the aforementioned instrumental variable grouped together into the variable “Instrumental”:

**Equation 2.** \( \text{PredictedBusStops}_i = \beta_0 + \beta_2 \text{Control}_i + \beta_3 \text{Instrumental}_i + \text{Error}_i \)

**Equation 3.** \( \text{LaborForceParticipation}_i = \beta_0 + \beta_1 \text{PredictedBusStops}_i + \beta_2 \text{Control}_i + \text{Error}_i \)

Once weakness of the 2SLS model is that is does not take advantage of all the relevant information contained in Equation 2 and Equation 3 above. For example, there may be potential
omitted variable bias, or unobserved factors that influence both bus stops and labor force participation that are not contained in the controls. For example, more detailed information on the racial composition within a ZCTA – beyond that explained in the control variable “Proportion White”—may influence both variables of interest. These common unobservables will inevitably end up in the error terms of both equations, reflecting imprecision in the estimates. The three-stage least squares model (3SLS) improves upon the 2SLS estimates by taking advantage of correlation in the error terms of both equations to make inferences about the influence of these unobserved by relevant factors. The full set of results for this model are show in Figure 7, to follow. Further explanation is available in Appendix F.

In short, by using both instrumental variables and correlation in the error terms of the two estimating equations, we can reasonable conclude that $\beta_1$ represents the effect of bus stop access on labor force participation, not the reverse.

**E. Results**

Figure 5 shows the results for the ordinary least squares (OLS), two-stage least squares (2SLS), three-stage least squares (3SLS) models. The results indicate that the effect of the bus stop access variable was robust, even after correcting for endogeneity. This is important from a policy perspective, as it provides evidence that the quantity of bus stops within a 2-kilometer radius of a ZCTA increases labor force participation by 0.12 to 0.13%. This is relatively consistent across models.\(^{22}\)

---

\(^{22}\) As mentioned, full regression results for each model are shown in Appendix F.
Figure 7. Using multi-stage regressions to correct for potential endogeneity by replacing observed bus with predicted bus stops to predict labor force participation rates did not result in materially different results. Therefore, we can dismiss the endogeneity concerns of the bus stop variable.

As seen in Figure 7, the 2SLS model had a slightly lower significant level than the OLS, at 5% rather than a 1% level. Interestingly, it also had a slightly increased coefficient estimate. Using the 3SLS model, an increase in 1 bus stop within 2 kilometers of a ZCTA centroid was associated with a 0.13% increase in the labor force participation rate within that ZCTA, whereas the simple OLS model resulted in a similar 0.12% increase. These estimates are not statistically different from each other. Therefore, through use of the more sophisticated 2SLS and 3SLS models, I conclude that there is no strong statistical evidence for endogeneity.

The null hypothesis in this study is that bus access has no relationship to labor force participation rates. The alternative hypothesis is that bus access does have a relationship with labor force participation. Within the Incorporated Place of San Diego City, I can conclude at a 5-percent level that increasing the quantity of bus stops within 2 kilometers of a ZCTA by 1 stop correlates to an increase in labor force participation of approximately 0.13%.

Concerning the non-primary variables, work-at-home workers held the highest significance level, 0.1%, finding that an increase in one work-at-home worker is associated with
a 0.05% increase in labor force participation. This may reflect the fact that telecommuting is often viewed as “providing less disutility” or “easier” than commuting jobs and, therefore, make work-at-home workers more likely to work.

Of the remaining control variables, the results for log median earnings and total housing occupancy rate were most interesting. Results indicate a 1% increase in median earnings would result in a 0.08% increase in labor force participation, at a 1% significance level. This is a modest coefficient, but may signify that high-income workers who are better compensated or receive investment income are more likely to be in the labor force. However, this raises the question of whether high-income households elect to work because they are better-compensated (with a higher opportunity cost for leisure) or due to necessity, such as a high-debt lifestyle.

Total housing occupancy also has a positive impact, whereby an increase in 1% is expected to result in a 0.48% increase in labor force participation rate, at a 5% significance level. A strong real estate market may be viewed as an indicator of a thriving local economy which, oftentimes, is linked to a strong job market in that region.
F. Policy Implications and Limitations

While results indicated that there is a statistically significant, positive, relationship between bus stop access and labor force participation in the Incorporated Place of San Diego, the size of this relationship makes policy prescriptions less clear. The bus stop coefficient estimate appears small relative to the labor force participation distribution, which has a standard deviation of 5.94% (Figure 7). However, of the ZCTAs reviewed, 0.13% of the population would represent about 40 individuals23 entering the labor force, per ZCTA. Further work examining the costs of expansion relative to these benefits would need to be made to establish whether the policy prescription of improved transit access is warranted.

This work would be enhanced by (1) replications in other cities and time periods (2) an improved measure of transit access and (3) use of large samples with smaller margins of errors. Firstly, cross-sectional analysis provides a “snapshot” of the time and place reviewed. Future work might expand this research to evaluate whether the relationship holds in other cities and during other time periods. It is possible that there is an unanticipated uniqueness to the case of San Diego in the 2010s, whereby other cities and time periods might have stronger or weaker results.

Secondly, the use of straight-line distance may not be the strongest proxy for bus stop access. This distance may be unrepresentative of true accessibility. For example, actual distance to walk to that bus stop may be hindered by obstacles or restricted to sidewalks, which would likely be a longer distance than the straight-line method utilized. Further, geographic location may not be truly representative of access as it excludes factors such as service frequency, fare costs, and the aesthetic quality of the travel mode. Mamun and Lownes (2011) have developed

---

23 This value is derived from a simple average of the population over sixteen years of age in each of the 50 ZCTAs reviewed. This mean, 31,386, was then multiplied by the coefficient estimate for bus stop access, 0.13%, equals approximately 40 individuals.
such an accessibility access, and the substitution of their method would likely provide a more holistic measure of “access.”

Finally, the data utilized for analysis was compiled from the American Community Survey 5-Year Estimates wherein some variables had large margin of errors. This is not uncommon for survey sampling datasets but, for this reason, this data may not be truly representative of the population.

IX. Conclusion

The evolution of United States cities from industrial hubs to borderless masses is a byproduct of shifting (1) transportation technologies, (2) government programs, and (3) social preferences. Today, discussion must turn to the impact urban decentralizing is having on the present. Throughout this paper, I’ve established the advantages sprawl has provided for firms and high-income workers, including lower land costs and shorter commutes (Glaeser, Kahn 2004; Sultana, Weber 2014). At the same time, sprawl creates costs for low-income urban workers.

To adapt to a sprawled geographic equilibrium, many high-income workers relocate and utilize private transportation to receive the benefits of shorter commutes (Sultana, Weber 2014; Crane, Chatman 2003). This process is a component of sprawl maturation (Sultana, Weber 2014). Many low-income workers are unable to participate in this maturation due to the financial limitations and job insecurity that typically accompanies low-income employment (Moulding 2005; Crane, Chatman 2005; Holzer, Martinson 2006). Instead, many low-income populations hedge their risk by performing time-intensive and money-intensive reverse-commutes from centrally located residences (Crane, Chatman 2005).
Advocates call for government policies to bridge the growing opportunity gap between those who benefit from sprawl (the high-income) and those who do not (the low-income). This paper reviewed two outspoken parties who believe public infrastructure is the proper solution: the pro-highway and pro-mass transit camps. To evaluate at least one camp’s policy option, the final portion of this paper examined the relationship between public bus stop access and labor force participation in San Diego, California.

Using a straight-line distance-based transit accessibility variable alongside a string of control and instrumental variables, a two-stage least squares regression indicated that an increase in access to one bus stop would be expected to increase the labor force participation rate by 0.12-0.13%. This finding held, even after accounting for endogeneity. Based on the averages for the ZCTAs evaluated, this would be equivalent to approximately 40 individuals entering the labor market.

Like most research, this paper raises more questions than it answers. Further work might expand this study to include additional cities to confirm whether these findings are unique to San Diego. Secondly, the integration of a holistic measure of transit access that includes transit quality and service frequency, such as that developed by Mamun and Lownes (2011) would likely increase the accuracy of these results (Dill et al. 2013). Moreover, interesting findings may result from the incorporation of time-series data, as opposed to cross-sectional data, to evaluate how transit-related benefits grow or diminish over time following transit investment.

This paper follows the historical evolution of sprawl and explores a narrow sampling of potentially-equalizing policy solutions. For scenarios where policy was part of the problem, there’s considerable pressure to maintain conservative prescriptions to ensure new policy doesn’t have any unwanted side effect. Apprehension and naysaying upon a history of failure makes it
easy to avoid taking meaningful steps and, instead, to rely on the safety of abstract theory
difficult to apply to the real world. Unfortunately, the costs of sprawl are a matter of reality, not theory.
References


Crane, Randall, and Daniel G. Chatman. "As jobs sprawl, whither the commute?" *ACCESS Magazine* 1, no. 23 (2003).


Appendices

Appendix A - History, Policy, and Trends

Urban and Rural Populations of the United States (1800-2010)

Figure 1. By graphing data from the U.S. Census Bureau, the overall trend of United States urbanization is clear. As the paper refers to “suburbanization,” “sprawl,” and “decentralization,” it must be noted that these are trends within the metropolitan area. (U.S. Census Bureau 1993; U.S. Census Bureau 2012)

Figure 2. This describes the poverty rate in relationship to distance from city center. As seen in this Census-derived graphic by Luke Juday (n.d.) poverty rates have been decentralizing from 1990 to 2015. In visual terms, the slope of the line has been smoothing out (Juday n.d.).
<table>
<thead>
<tr>
<th></th>
<th>Total Poverty Population (in 1000s)</th>
<th>Percentage Change (1990-2014)</th>
<th>Mean Poverty Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Urban Tracts</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9,544</td>
<td>12,741</td>
<td>33.5%</td>
<td>14.4%</td>
</tr>
<tr>
<td><strong>Suburban Tracts</strong></td>
<td>8,616</td>
<td>16,909</td>
<td>55.5%</td>
</tr>
</tbody>
</table>

Source: US Census 1990; American Community Survey, 2010-2014

**Figure 3.** This trend is confirmed by Allard and Paisner (2016) who find that, though poverty growth is higher in suburbs than in urban tracts, the mean poverty rate is still higher, by far, in urban areas.
Appendix B - Creating the Bus Stop Dataset

1. Why Buses?
Bus lines were used as a proxy for public transit in general. This is because bus lines present the lowest-cost and highest-flexibility public transit option in most cities. Buses are also the most common form of transportation for the low-income workforce, which is the population most often priced-out of private transportation and is the demographic this research targets (Moulding 2005). If more time were available, expanding this research to incorporate access to any type of transit, including the small and more-expensive trolley system that exists in San Diego, would improve the accuracy of these findings.

2. Physical Address
The first step in the construction of the bus stop variable was the acquisition of public bus stop locations. Bus routes are more flexible than other modes, so finding up-to-date comprehensive information on these routes proved more difficult than anticipated and, as discussed in Appendix C, was a key determinant in the choice of city. The data for San Diego was accessed through the San Diego Metropolitan Transit Authority’s website (San Diego MTS “Bus Routes” 2017). On this site, as seen in Figure 4, bus routes were listed by number.

![Figure 4; Figure 5. List of bus line routes and numerical identifiers (San Diego MTS “Bus Routes” 2017). The second image shows the interactive mapping tool for bus routes (San Diego MTS “Bus Routes” 2017).](image)

Using the interactive bus route tool, each of these bus lines were then mapped, with dots representing each stop, as seen in Figure 4. One by one, these stops -- reported as intersections-- were recorded in a spreadsheet for each of the 100 bus lines. These physical addresses were then converted into latitude and longitude coordinates using an online application for batch geocoding (Zwiefelhofer, n.d.). Using Microsoft Excel, these values were then screened for replications.
Several bus lines overlapped at high-volume stations and, following these reductions, 2,443 unique bus stop coordinates remained.

This data is available for viewing at: https://drive.google.com/drive/folders/0BztBnS-iOrgxOXJMEJqXzBMZVE?usp=sharing

3. Zip Code Tabulation Areas (ZCTAs)
The cases for analysis in this report are Zip Code Tabulation Areas, which are a unit of demographic information reported by the American Community Survey. They are similar, but not identical to, United States Postal Service Zip Codes. I feared aggregation would suppress findings, which is why such a small unit was selected. In the creation of the bus stop variable, I evaluated bus stop access in terms of the quantity of bus stops within a 2-kilometer radius from the centroid of each ZCTA. The centroids of each of the 55 ZCTAs were found using MapTechnica.com, where it was reported as latitude and longitude measure.

This data is available for viewing at: https://drive.google.com/drive/folders/0BztBnS-iOrgxOXJMEJqXzBMZVE?usp=sharing

4. Distances between Latitude-Longitude Coordinates
The straight-line distance (in kilometers) between each of the 55 ZCTAs and 2,443 bus stop locations were calculated using the following function in Microsoft Excel, as designed BlueMM.blogspot.com (n.d.):

\[
\text{ACOS} \left( \text{COS} \left( \text{RADIANS} \left( 90 - \text{Lat1} \right) \right) \cdot \text{COS} \left( \text{RADIANS} \left( 90 - \text{Lat2} \right) \right) + \text{SIN} \left( \text{RADIANS} \left( 90 - \text{Lat1} \right) \right) \cdot \text{SIN} \left( \text{RADIANS} \left( 90 - \text{Lat2} \right) \right) \cdot \text{COS} \left( \text{RADIANS} \left( \text{Long1} - \text{Long2} \right) \right) \right) \cdot 6371
\]

The resulting 134,365 data points (the result of 55 ZCTA x 2,443 bus stops), were then evaluated and converted into a singular quantity: the quantity of bus stops for which the distance to the ZCTA centroid was less than 2 kilometers. The Excel function used was:

\[
\text{=COUNTIF} \left( \text{Bus Stop's Distance from Centroid},"<=2" \right)
\]

This data is available for viewing at: https://drive.google.com/drive/folders/0BztBnS-iOrgxOXJMEJqXzBMZVE?usp=sharing

The resulting dataset had a minimum of 0, a maximum of 103, first quartile value of 4, mean of 27.84, third quartile value of 42.25, and a median was 27.
Appendix C - Incorporated Place of San Diego City

1. Incorporated Places

A place is defined by the Bureau of the Census as, “a concentration of population; a place may or may not have legally prescribed limits, powers, or functions. This concentration of population must have a name, be locally recognized, and not be part of any other place” (U.S. Census Bureau “GARM” 2013). The Census Bureau (2013) publication goes on to explain that places can be census-designated, whereby the census decides where its boundaries are, or incorporated. An Incorporated Place is a legally incorporated area that meets the conditions of its respective state. In California, an incorporated place is required to have a minimum of 500 registered voters to incorporate as a city or town (U.S. Census Bureau “GARM” 2013). This research utilized the Incorporated Place of San Diego City. San Diego is identified by the U.S. Census with the Place number of 66000.

2. Map of Area

![Map of Incorporated Place of San Diego City](image)

**Figure 6.** Provided by the U.S. Census-sponsored TigerMap (n.d.), this image shows the boundaries of the Incorporated Place of San Diego City.

It is important to note that the Incorporated Place of San Diego City doesn’t follow ZCTA boundaries precisely--there are some ZCTAs that are only partially within the bounds of the Place that, for the purposes of this research, are treated equally to those ZCTAs which are entirely within the Place.
3. Why San Diego?

As stated in the body of this report, these regressions utilized San Diego as an example. These regressions may have been easily done in other, similar, cities. San Diego was selected for its (a) size, (b) data availability, (c) modes of transportation, (d) familiarity, and (e) reputation as a city with above-average public transit. I will first expand each of these points before highlighting the many other cities that may have been equally as appropriate for study.

a. Size - San Diego is a mid-sized city with a simple transportation system. As noted in Appendix B, much of the data collection for this report was done manually. To minimize the time costs of data collection, I targeted small- and mid-sized cities who did not have overly-expansive transit infrastructure that would be cumbersome to collect data on as opposed to large, complex, cities like New York City.

b. Data Availability - As a central and time-intensive portion of this research was the creation of a bus stop variable, accessible and clear data on where transit stops were located was essential. After reviewing several cities, I was impressed by the clarity with which San Diego’s transit information was presented. Comparable data in Los Angeles, for example, was incredibly difficult to access due to route images-- not a string of physical addresses-- being the most commonly provided information about where stops were located.

c. Modes of Transportation - San Diego has two forms of public transit: bus and trolley. The trolley’s coverage was limited and expensive. As such, bus routes make up the vast majority of transit stops. As I was limiting the research to bus stops specifically, San Diego’s reliance on bus stops almost exclusively was a positive factor.

d. Familiarity - The use of San Diego provided qualitative benefits for the sake of relaying this research as well. Simply put, people know San Diego, where it is, and, often, what it is like. With this popularity, San Diego (and Southern California more generally) was easier to find in literature and to explain to readers than otherwise comparable sites.

e. Reputation - Not unrelated to my previous point, San Diego also has a reputation in media as a site for strong public transit. Wanting to provide a best-case scenario for research, the repeating sources that cited California, generally, and San Diego, specifically as a place with outstanding public transit informed my early decision to narrow down to the San Diego region (Tomer et al. 2011).
4. Field Work

**Figure 7.** To evaluate San Diego’s transit personally, I first secured a multi-use pass. The pass itself cost $2, and I could load it for one day of unlimited ride of the transit and bus for an additional $5. If I had wanted the same level of access, which is the lowest offered, for a month-long period, it would cost $72.

**Figure 8; Figure 9.** Using the bus in San Diego was an odd experience. On routes in the more touristy areas, I was often one of only one or two riders. During comparable trips that ventured into the low-income portions of the city, the bus was heavily crowded with many strollers filling aisles and several languages being spoken at once. On those trips, the clear majority of riders were non-white. The buses were well-maintained and stops seemed to be thoroughly distributed throughout the city. Some bus stops had seating for waiting riders, others did not.
The trolley of San Diego was a different story. On the trolley, cars were full and fewer routes were offered. Many more riders who appeared to be tourists were on these routes, but a greater proportion of locals—who appeared to be of all income levels—were present. In terms of race, the trolley seemed to be more representative of the city itself with even proportions of people who appeared to be white, black, and Latino. Waiting for the trolley was luxurious compared to buses—large platforms with seating, trash cans, and, often, palm trees were the norm. Due to their sprawling size, these stations were more inconvenient to access.
Appendix D - Incomplete Data

Of the 55 ZCTAs assessed, five ZCTAs had incomplete data as provided in the American Community Survey. To avoid biasing regression by including blanks (which were read as zeros) that, in actuality, are non-zero, these cases were excluded from the data. These ZCTAs were: 92135, 92140, 92134, 92147, 92132. These ZCTAs are all geographically small ZCTAs, as seen below.

Figure 12. These excluded ZCTAs may have had missing data because of their small size, but other similarly-sized ZCTAs, like 92155, have had not this issue. Whatever the reason for the missing data, this is a shortcoming of the dataset and regression.
Appendix E - Observed Bus Stop Regression

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Labor Force Participation Rate</th>
</tr>
</thead>
</table>
| Bus Stops within 2 km of Centroid | 0.128***  
|                    | (0.038) |
| Log Median Earnings | 7.403***  
|                    | (2.399) |
| Total Housing Occupancy Rate | 0.476**  
|                    | (0.191) |
| Owner Vacancy Rate | 0.673  
|                    | (0.616) |
| Rental Vacancy Rate | -0.610  
|                    | (0.380) |
| Median Age Squared | -0.006**  
|                    | (0.002) |
| Work-at-Home Workers | 0.058***  
|                    | (0.015) |
| Gross Rent | -0.040  
| **Households with 0 Vehicles | -0.375**  
|                    | (0.188) |
| Proportion White | 0.032  
|                    | (0.049) |
| Constant | -54.698  
|          | (35.767) |

Observations: 50
R²: 0.663
Adjusted R²: 0.576
Residual Std. Error: 3.69 (df = 39)
F Statistic: 7.661*** (df = 10; 39)

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 13. Multivariate regression results prior to adjusting for the endogeneity of bus stops and excluding incomplete cases find that bus stops are significant at a 1% level. This regression tells us that, if the number of bus stops within two kilometers of a ZCTA centroid increased by 1 stop, the labor force participation in that ZCTA would be expected to increase by 0.12%. Concern regarding the validity of these findings, stemming from the potential endogeneity of the bus stop variable, led to further regressions using a predicted bus stop variable in place of the observed values.
Appendix F - Bus Stop Prediction Formula

1. Instrumental Variables to Predict Bus Stops

An array of variables were considered when selecting instrumental variables to predict bus stops in each ZCTA, including: (a) Number of businesses (b) Number of employees working (not living) in a ZCTA (c) Total annual payroll firms paid (d) Population density (e) Housing Unit density (f) Land area (distinct from “Area” as it excludes bodies of water).

Number of businesses and number of employees were strongly related and, when combined, their correlations to bus stops were reduced. When entered separately, the number of employees had a stronger correlation (significant at the 1% level) than number of businesses (significant at the 10% level) with bus stops. The number of employees were then paired with another significant variable, the population density, and the control variables discussed in Appendix G in a multivariate regression the results are shown below in Figure 14.

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Bus Stops within 3km of Centroid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees Working in ZCTA</td>
<td>0.047*** (0.012)</td>
</tr>
<tr>
<td>Population Density</td>
<td>3.403*** (0.557)</td>
</tr>
<tr>
<td>Log Median Earnings</td>
<td>-2.126 (5.611)</td>
</tr>
<tr>
<td>House Language Not English (Over 5)</td>
<td>-0.119 (0.144)</td>
</tr>
<tr>
<td>Total Housing Occupancy Rate</td>
<td>-0.082 (0.610)</td>
</tr>
<tr>
<td>Owner Vacancy Rate</td>
<td>0.289 (1.994)</td>
</tr>
<tr>
<td>Rental Vacancy Rate</td>
<td>0.101 (1.707)</td>
</tr>
<tr>
<td>Median Age Squared</td>
<td>-0.0003 (0.008)</td>
</tr>
<tr>
<td>Work at Home Workers</td>
<td>-0.052 (0.049)</td>
</tr>
<tr>
<td>Gross Rent</td>
<td>-0.0003 (0.004)</td>
</tr>
<tr>
<td>Households with 0 Vehicles</td>
<td>0.010* (0.006)</td>
</tr>
<tr>
<td>Proportion White</td>
<td>-0.804 (17.001)</td>
</tr>
<tr>
<td>Constant</td>
<td>21.819 (135.244)</td>
</tr>
</tbody>
</table>

Observations: 50
R^2: 0.847
Adjusted R^2: 0.797
Residual Std. Error: 11.728 (df = 37)
F Statistic: 17.008*** (df = 12. 37)

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 14. Number of employees working in a ZCTA and population density are strongly correlated with bus stops, which helped make the case for their inclusion as instrumental variables.

In this case, instrumental variables are only viable if their independence from labor force participation can be maintained while keeping correlation with bus stops high. To select variables, a logical argument for independence was crucial. The instrumental variables and their
logical reasons for correlation with bus stops (but not with labor force participation) are outlined below:

**Number of Employees**

*Impacts Bus Stops* because an increasing quantity of potential riders (who earned incomes with which to pay for rides) would increase the demand for and economies of scale of transit. This may then determine where new transportation was installed.

*Doesn’t Impact Labor Force Participation Rate* because changing quantities of people should not impact a rate which our labor force measure is.

**Population Density**

*Impacts Bus Stops* because, like the number of employees, it indicates greater potential demand for occupational and recreational travel.

*Doesn’t Impact Labor Force Participation Rate* because, again, rates should remain largely independent of any changes in quantities.

| Number of Employees Working in ZCTA | 0.003 | (0.004) |
| Population Density | 0.712* | (0.309) |
| Log Median Earnings | 5.485* | (2.797) |
| House Language Not English (Over 5) | -0.123** | (0.047) |
| Total Housing Occupancy Rate | 0.427** | (0.158) |
| Owner Vacancy Rate | 0.803 | (0.648) |
| Rental Vacancy Rate | -0.775* | (0.361) |
| Median Age Squared | -0.028*** | (0.003) |
| Work-at-Home Workers | 0.045*** | (0.016) |
| Gross Rent | -0.0002 | (0.0004) |
| Households with 0 Vehicles | -0.0002 | (0.0025) |
| Proportion White | -0.470 | (5.509) |
| Constant | -19.916 | (43.823) |

| Bus Stops within 2 km of Centroid | 0.112*** | (0.055) |
| Number of Employees Working in ZCTA | -0.001 | (0.005) |
| Population Density | 0.206 | (0.428) |
| Log Median Earnings | 8.325*** | (2.635) |
| Total Housing Occupancy Rate | 0.510*** | (0.201) |
| Owner Vacancy Rate | 0.838 | (0.667) |
| Rental Vacancy Rate | -0.584 | (0.396) |
| Median Age Squared | -0.007*** | (0.002) |
| Work-at-Home Workers | 0.058*** | (0.017) |
| Gross Rent | -0.074 | (0.410) |
| Households with 0 Vehicles | -0.381* | (0.192) |
| Proportion White | 0.042 | (0.053) |
| Constant | -69.164*** | (40.410) |

**Figure 15; Figure 16.** Figure 15 shows the result of an ordinary least squares regression that correlates the instrumental variables, alongside control variables, with labor force participation. This regression finds that population density is correlated with labor force participation, but the
number of employees working in a ZCTA is not. However, it is possible that this relationship may be through the bus stop variable. To evaluate this possibility in the hopes of discovering these instrumental variables are truly empirically independent of labor force participation, bus stops are added into the regression in Figure 16. Once incorporating the bus stop variable into the ordinary least squares regression, it appears that any significance of the instrumental variables has diminished; in other words, the interaction between our instrumental variables and labor force participation appears to be through our bus stop variable.

2. Two-Stage Regression

Utilizing the R program “systemfit” (Henningsen, Hamann 2007), I was able to insert the predicted bus stop function into the labor force participation function. This process, known as a two-stage least squares regression, first, evaluated the ordinary least squares regression and, second, reevaluated the regression using the predicted bus stop variable. The output of this second stage is summarized below.

**Model Formula:** 
\[ LF16 + bus\_stops + log\_Median\_Earnings + Occ\_Housing + Own\_Vac + Rent\_Vac + Med\_Age\_Sq + At\_Home + Gross\_Rent + V0 + Prop\_White \]

**Instruments:** 
\[-Employees + Pop\_Density + log\_Median\_Earnings + Occ\_Housing + Own\_Vac + Rent\_Vac + Med\_Age\_Sq + At\_Home + Gross\_Rent + V0 + Prop\_White \]

| (Intercept) | Estimate | Std. Error | t value | Pr(>|t|) |
|-------------|----------|------------|---------|----------|
| -56.72 | 36.40 | -1.55 | 0.12724455 * |
| Bus Stops within 2 km of Centroid | 0.1349 | 0.0529 | 2.54 | 0.01481974 * |
| Log Median Earnings | 7.53 | 2.4096 | 3.12 | 0.00332153 ** |
| Total Housing Occupancy Rate | 0.4797 | 0.1912 | 2.50 | 0.01638439 * |
| Owner Vacancy Rate | 0.6459 | 0.6227 | 1.03 | 0.30594389 |
| Rental Vacancy Rate | -0.61 | 1.22 | -0.50 | 0.611106 |
| Median Age Squared | -0.06 | 0.01 | -0.01 | 0.991072 |
| Work-at-Home Workers | 0.05 | 0.01 | 3.74 | 0.00058609 *** |
| Gross Rent | -0.02 | 0.40 | -0.02 | 0.9122119 |
| Households with 0 Vehicles | -0.39 | 0.16 | -2.47 | 0.0574268 |
| Proportion White | 0.00 | 0.05 | 0.71 | 0.47957773 |

Significance codes: 0 ‘****’ 0.001 ‘***’ 0.01 ‘**’ 0.05 ‘*’ 0.1 ‘ ’ 1

Residual standard error: 3.8733 on 39 degrees of freedom
Number of observations: 50 Degrees of Freedom: 39
SSR: 58.5100426 MSE: 15.002375 Root MSE: 3.873316
Multiple R-Squared: 0.661848 Adjusted R-Squared: 0.575142

**Figure 17.** Once incorporating the predicted bus stop variable, which accounted for potential endogeneity of the bus stop variable, the significance of the bus stop variable decreased to a 5% level, as opposed to the 1% level previously. However, the coefficient has increased. This model finds that an increase in 1 bus stop will see a 0.13% increase in labor force participation, versus as 0.12% increase in Figure 14 using an ordinary least squares regression. This consistency indicates that these findings are robust, or consistent despite alterations and modifications to the regression model.
3. Three-Stage Regression

\textbf{3LS estimates for eq2 (equation 2)}

Model Formula: \( LF16P = \text{bus stops} + \log \text{Med Earnings} + \text{Occ Housing} + \text{Own Vac} + \text{Rent Vac} + \text{Med Age Sq} + \text{At Home} + \text{Gross Rent} + V0 + \text{Prop White} \)

Instruments: - \text{Employees} - \text{Pop Density} - \log \text{Med Earnings} - \text{Occ Housing} - \text{Own Vac} + \text{Rent Vac} + \text{Med Age Sq} + \text{At Home} + \text{Gross Rent} + V0 + \text{Prop White} \)

| (Intercept) | Estimate | Std. Error | t value | Pr(>|t|) |
|-------------|----------|------------|---------|----------|
| -56.72680429 | 36.40335452 | -1.55829 | 0.12724455 |

| Bus Stops within 2 km of Centroid | 0.1492822 | 0.03291647 | 2.54985 | 0.01481974 * |
| Log Median Earnings | 7.53782157 | 2.40695513 | 3.12218 | 0.00332153 ** |
| Total Housing Occupancy Rate | 0.47972394 | 0.19122439 | 2.50870 | 0.01638459 * |
| Owner Vacancy Rate | 0.64598601 | 0.62270532 | 1.03739 | 0.30594389 |
| Rental Vacancy Rate | -0.61555264 | 0.38942615 | -1.58063 | 0.12204106 |
| Median Age Squared | -0.00624553 | 0.00240995 | -2.39156 | 0.01337488 * |
| Work-at-Home Workers | 0.03538573 | 0.01582422 | 3.74269 | 0.00038609 *** |
| Gross Rent | -0.03440426 | 0.040155562 | -0.08558 | 0.93216119 |
| Households with 0 Vehicles | -0.39390104 | 0.19972642 | -1.97191 | 0.05574269 |
| Proportion White | 0.03651713 | 0.05115603 | 0.71384 | 0.47957773 |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.873316 on 39 degrees of freedom.
Number of observations: 50
Degrees of Freedom: 39
SSR: 585.100426
MSE: 15.002575
Root MSE: 3.873316
Multiple R-Squared: 0.661848
Adjusted R-Squared: 0.575342

\textbf{Figure 18.} To confirm the validity of our two-stage least square regression, I performed a three-stage least squares regression to compare findings. A three-stage least squares regression performs a similar function as a two-stage least squares regression, except that it performs stages simultaneously and considers the interactions between each equation rather than exploring each equation independently. The results were identical.

4. Justification of Control Variables

The following list outlines the control (independent) variables utilized in the statistical regressions as well as the justifications for using each:

**Natural Log of Median Earnings**

\textit{Definition}: is the median of annual income received by individuals over 16 before deductions, grouped by households, within each ZCTA. The data is provided in dollar values.

\textit{Manipulation}: Median was selected because it would be more skew resistant than mean. A natural log is applied to normalize the right-skewed data, which is common with income-related variables.

\textit{Logical Reason for Use}: It has been observed that increasing incomes are strongly associated with longer commute times in the United States and this is largely related to the increased use of private vehicles (to improve commute speed) and the increased disposable income (to cover the costs of transportation) (Crane, Chatman 2003). Transit, likely for these reasons, is often used most heavily by low-income populations (Cox et al. 2014). For these reasons, it seemed important to incorporate a measure of earnings when determining the impact that transit would have on employment.

\textit{Descriptive Statistics Without Log}: Minimum: 20037; 1st Quartile: 30464.3; Mean: 41024.8; Median: 38524; 3rd Quartile: 46252.5; Maximum: 104645
Total Housing Occupancy Rates

*Definition:* the proportion of all available housing (owner- and tenant-occupying) that is occupied. That is, not “for sale,” “for rent,” or abandoned. The data is provided as a proportion.

*Logical Reason for Use:* The state of the housing market may be an indicator of the economic success of an area which, oftentimes, is correlated with the job prospects.

*Descriptive Statistics:* Minimum: 78.1; 1st Quartile: 91.5; Mean: 92.1; Median: 93.2; 3rd Quartile: 95.7; Maximum: 97.1

Owner Vacancy Rates (Owner Lives in)

*Definition:* the proportion of homeowner housing units that are listed “for sale” divided by total homeowner housing units. Note this is not just the inverse of Total Housing Occupancy because it excludes abandoned properties. The data is provided as a proportion.

*Logical Reason for Use:* The state of the housing market may be an indicator of the economic success of an area which, oftentimes, is correlated with the job prospects. High owner-occupancy, specifically, could be related to the wealth and stability of the area as it relies on personal mortgages rather than lease agreements.

*Descriptive Statistics:* Minimum: 0; 1st Quartile: 1.1; Mean: 1.8; Median: 1.4; 3rd Quartile: 2.4; Maximum: 8.8

Rental Vacancy Rates (Tenant Lives in)

*Definition:* the proportion of rental housing units that are listed “for rent” over total rental housing units. Note this is not just the inverse of Total Housing Occupancy because it excludes abandoned properties. The data is provided as a percentage.

*Logical Reason for Use:* The state of the housing market may be an indicator of the economic success of an area which, oftentimes, is correlated with the job prospects. Tenant occupancy is more volatile than homeowner occupancy, so it would provide a faster indication of the area’s economic health.

*Descriptive Statistics:* Minimum: 0; 1st Quartile: 2.4; Mean: 3.9; Median: 3.8; 3rd Quartile: 5.3; Maximum: 8.1

Median Age Squared

*Definition:* the median age of each person, surveyed by household. The data is provided as number of years.

*Manipulation:* The squaring was applied to better display a non-linear relationship between median age and labor force participation. This follows logically, as the very young and the very old tend to be less active in the labor force and may, as the raw data’s scatterplot between labor force and median age showed, followed a more bell-like shape (Figure 19).
**Reason for Use:** Age undoubtedly affects labor force participation; a 16-year-old and an 86-year-old would be unlikely to work the same amount as a 26-year-old.

**Descriptive Statistics Prior to Squaring:** Minimum: 25.8; 1st Quartile: 31.7; Mean: 35.9; Median: 35.5; 3rd Quartile: 38.7; Maximum: 50.3

![Scatterplot of Labor Force Participation and Median Age for ZCTAs](image)

**Figure 19.** The relationship between median age and labor force participation appears to be non-linear.

**Work-At-Home Workers**

*Definition:* the quantity of workers who primarily worked from home. This data is provided as a quantity of workers.

*Reason for Use:* if a worker does not need to travel to find or maintain employment, variables relating to transportation should not impact that worker’s labor force participation.

*Descriptive Statistics:* Minimum: 12; 1st Quartile: 69.8; Mean: 108.8; Median: 97.5; 3rd Quartile: 141.8; Maximum: 291

**Gross Rent**

*Definition:* the monthly housing costs for renters, grouped by housing unit, including utilities and fuels. The data is provided in dollar values.

*Reason for Use:* The amount a household spends on rent is funds that cannot be used to spend on transportation and may also indicate the degree of wealth (not income) the household has. The wealthiness of a household and amount of funds available for transport spending are assumed to positively impact labor force participation.

*Descriptive Statistics:* Minimum: 102; 1st Quartile: 3199; Mean: 6553.3; Median: 5732.5; 3rd Quartile: 9017.5; Maximum: 15201
Quantity of Households without Vehicles

*Definition:* the number of households with zero personal vehicles. The data is provided as quantity of households.

*Reason for Use:* Vehicle use is the primary alternative to transit for commuters. The availability of a vehicle would likely indicate a household is unlikely to need or use transit, even if it is accessible.

*Descriptive Statistics:* Minimum: 8; 1st Quartile: 338.3; Mean: 948; Median: 733.5; 3rd Quartile: 1196.8; Maximum: 4454

Proportion of Population White

*Definition:* the proportion of people who identified as “white,” Irish, German, Italian, Lebanese, Arab, Moroccan, or Caucasian. This data is provided as a proportion of the population over 16.

*Reason for Use:* Labor force participation could be affected by discriminatory factors; some firms may not want to hire non-white individuals. Race-based discrimination would then impact the quantity of jobs available and, by extension, labor force participation.

*Descriptive Statistics:* Minimum: 0.23; 1st Quartile: 0.59; Mean: 0.70; Median: 0.73; 3rd Quartile: 0.83; Maximum: 0.92
Appendix G - Consistency of Findings

1. Age Categories

Consistency of Ordinary Least Squares Regression to Explain Different Labor Force Participation Rate Age Categories

<table>
<thead>
<tr>
<th>Age Category:</th>
<th>Overall (16+)</th>
<th>16-19</th>
<th>20-24</th>
<th>25-44</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-54.6988</td>
<td>0.1593</td>
<td>26.4806</td>
<td>48.6731</td>
</tr>
<tr>
<td>Bus Stops within 2 km of Centroid</td>
<td>0.1324</td>
<td>0.0022**</td>
<td>0.0624</td>
<td>0.3130</td>
</tr>
<tr>
<td>Log Median Earnings</td>
<td>7.4055</td>
<td>0.0033**</td>
<td>-0.0001</td>
<td>0.1320</td>
</tr>
<tr>
<td>Total Housing Occupancy Rate</td>
<td>0.4760</td>
<td>0.0169*</td>
<td>0.0412</td>
<td>0.8980</td>
</tr>
<tr>
<td>Owner Vacancy Rate</td>
<td>0.6732</td>
<td>0.2810</td>
<td>0.0436</td>
<td>0.9650</td>
</tr>
<tr>
<td>Rental Vacancy Rate</td>
<td>-0.6100</td>
<td>0.1245</td>
<td>0.0295</td>
<td>0.9620</td>
</tr>
<tr>
<td>Median Age Squared</td>
<td>-0.0062</td>
<td>0.0136*</td>
<td>0.0008</td>
<td>0.8310</td>
</tr>
<tr>
<td>Work-at-Home Workers</td>
<td>0.0581</td>
<td>0.0004***</td>
<td>0.0094</td>
<td>0.6990</td>
</tr>
<tr>
<td>Gross Rent</td>
<td>-0.0400</td>
<td>0.9210</td>
<td>0.0000</td>
<td>0.9670</td>
</tr>
<tr>
<td>Households with 0 Vehicles</td>
<td>-0.3732</td>
<td>0.0542</td>
<td>-0.0023</td>
<td>0.4490</td>
</tr>
<tr>
<td>Proportion White</td>
<td>0.0321</td>
<td>0.3169</td>
<td>3.0010</td>
<td>0.7060</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age Category:</th>
<th>45-54</th>
<th>55-64</th>
<th>65-74</th>
<th>75+</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>74.6600</td>
<td>0.0209*</td>
<td>-84.9900</td>
<td>0.0603</td>
</tr>
<tr>
<td>Bus Stops within 2 km of Centroid</td>
<td>0.0449</td>
<td>0.4075</td>
<td>-0.0003</td>
<td>0.9973</td>
</tr>
<tr>
<td>Log Median Earnings</td>
<td>0.0000</td>
<td>0.7033</td>
<td>0.0001</td>
<td>0.6169</td>
</tr>
<tr>
<td>Total Housing Occupancy Rate</td>
<td>0.1355</td>
<td>0.6317</td>
<td>1.3230</td>
<td>0.0019**</td>
</tr>
<tr>
<td>Owner Vacancy Rate</td>
<td>-0.4677</td>
<td>0.5925</td>
<td>2.1910</td>
<td>0.0821</td>
</tr>
<tr>
<td>Rental Vacancy Rate</td>
<td>-0.5922</td>
<td>0.3122</td>
<td>0.8828</td>
<td>0.2608</td>
</tr>
<tr>
<td>Median Age Squared</td>
<td>1.0000</td>
<td>0.7196</td>
<td>0.0092</td>
<td>0.0555</td>
</tr>
<tr>
<td>Work-at-Home Workers</td>
<td>0.0484</td>
<td>0.0271*</td>
<td>0.0345</td>
<td>0.2548</td>
</tr>
<tr>
<td>Gross Rent</td>
<td>0.0001</td>
<td>0.8664</td>
<td>0.0010</td>
<td>0.1995</td>
</tr>
<tr>
<td>Households with 0 Vehicles</td>
<td>-0.0051</td>
<td>0.0604</td>
<td>-0.0068</td>
<td>0.0387*</td>
</tr>
<tr>
<td>Proportion White</td>
<td>-8.1550</td>
<td>0.2462</td>
<td>5.9980</td>
<td>0.5716</td>
</tr>
</tbody>
</table>

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Figure 20. Labor force participation is an aggregate term. As described in Appendix F, it is a measure of labor force participation among all populations over the age of 16. This aggregation may hide some of the nuances involving labor force participation at various age categories. Fortunately, the American Community Survey also supplies labor force participation rates within age subsectors. Using the same independent variables used previously, Figure 20 outlines how outcomes differ by replacing the dependent variable of overall labor force participation with labor force participation for certain age categories.

Bus stops hold a significant relationship with labor force participation for 20-24 and 25-44 year olds, which is what we’d expect as those age categories embody core working years. Working at home seems most influential for 25-44 and year olds and less significant for 20-24 and 45-54. Interestingly, households without vehicles becomes and remains significant from all age categories from the 25-44 group up. Clearly, private transportation plays a prominent role for labor force participation.
2. 2-Kilometer Radius and the Evaluation of Alternatives

Calculations for quantity of bus stops accessible were based on a 2-kilometer radius from ZCTA centroids. This distance was selected based on the distance covered on a twenty-minute moderately-based walk (Bumgardner, 2017). Of course, the distance I deem reasonable to consider a bus stop accessible may not be reasonable to others. To address this concern, I also evaluated other radiuses.

Figure 22. The use of these alternative radiiuses is outlined above, including metrics where average and median distance to bus stops were calculated for each ZCTA. In the instance of mean and median, the bus stop variable is no longer a “Quantity of Bus Stops” within a range but, instead, a distance measured in kilometers. I find that my results hold for radiiuses of both 2.5 and 1.5 kilometers, though at 1.5 the significance level decreases to 5% as opposed to that at radiiuses of 2 and 2.5 kilometers where significance is at a 1% level.