“The spatial mismatch of minority groups”

a case study in Phoenix

Floris Jan Sander, s2729652
Ypenmolendrift 1-3
9711CP
f.j.d.sander@student.rug.nl
MSc Economic Geography, Regional Competitiveness & Trade
University of Groningen, Faculty of Spatial Sciences

Supervised by: dr. V.A. Venhorst
Co-assessor: dr. S. Koster

5th of August, 2019

ABSTRACT

This study uses longitudinal U.S. Census data to examine if minorities, living in Phoenix, face a spatial mismatch to employment. We approach the spatial mismatch hypothesis by looking at the distribution of minorities, the distribution of employment centers, and the commuting behavior of minorities. Overall, we find that Hispanics – particularly those living in the CBD – faced a spatial mismatch in 1990, and to a lesser extent in 2000. In more recent years we cannot find evidence that supports the spatial mismatch hypothesis for any minority.

Keywords: spatial mismatch hypothesis, employment accessibility, minorities

Abbreviations

SMH:	Spatial Mismatch Hypothesis
CBD:	Central Business District
(N)SUB:	(Nearest) Subcenter
J/H:	Job/Housing Ratio
TAZ:	Traffic Analysis Zone
ACS:	American Community Survey
CTPP:	Census Transportation Planning Products
TIGER:	Topologically Integrated Geographic Encoding and Referencing

Table of Contents

1. INTRODUCTION

Looking at recent years, the unemployment rate in the United States has steadily decreased (U.S. Bureau of Labor Statistics, 2019). Still, not all residents in the U.S. face the same labor market opportunities. Minorities like Native Americans, African Americans, and Hispanics have a one and a half to two times larger chance to be unemployed as opposed to non-minorities.

Kain (1968) argued that persistent housing segregation of minorities in central cities, combined with increasing suburbanization of metropolitan employment, creates a spatial mismatch to jobs for minority workers. In turn, the worse employment accessibility results in higher unemployment rates, longer commutes, and lower real wages. Kain named this phenomenon the *spatial mismatch hypothesis* (SMH).

In the decades after the publication of Kain, researchers explored to what extent minorities face a spatial mismatch; with ambiguous results. Kasarda (1989; 1995) states that particularly African Americans are unable to gain access to new growth industries in the suburbs of American cities. Taylor and Ong (1995) do not find evidence to support the SMH. They find that African American and Hispanic workers both faced shorter commutes than other groups, while longer commutes for minorities would be evidence to support the SMH.

Since the turn of the century, the SMH gained new attention because of the introduction of more advanced techniques to study the SMH and the rise of other minorities in American cities. Shen (2000) finds that central city minorities face a spatial mismatch to jobs since they face significantly longer commutes as opposed to other central city residents. To explain commuting duration, he uses an employment accessibility measure based on the urban spatial structure of the city he examines. Moreover, Raphael and Stoll (2002) argue that the spatial mismatch to jobs also becomes a problem for thriving minorities like Hispanics and Asians.

Till now, most emphases of the SMH was on old imperial cities in the east and industrial cities in the Midwest – like Boston (Shen, 2000) and Chicago (Wang, 2000) – since the worse labor market opportunities of African Americans spurred research to find causes. However, minorities residing in the Sunbelt do face worse labor market outcomes as well (Bureau of Labor Statistics, 2019). Due to major differences in dominant industries, ethnic composition, and population development between cities in the Sunbelt as opposed to cities in near the east coast (de Pater & Verkoren, 2007); findings and recommendations of previous research cannot automatically be asserted to thriving cities in the Sunbelt without a proper reflection.

Consequently, this study will examine the spatial mismatch hypothesis in a Sunbelt city. In this case: Phoenix, Arizona. The motivation to use Phoenix as a case study will be discussed further on. The research question is as followed.

To what extent is a spatial mismatch present for minority groups residing in Phoenix?

To answer the research question, several sub-questions will be asked. The SMH states that minorities residing in urban areas are spatially clustered in the central city, while employment is located near the outskirts of the city. Sub-questions one and two will thoroughly explore how minorities and employment are distributed across Phoenix. Sub-question three faces the question if minorities face worse employment accessibility as opposed to non-minority groups. Sub-question four is intertwined in all the other sub-questions and faces the question if the spatial mismatch is subject to change over time. Higher unemployment levels for minorities persist over time (U.S. Census, 2018), but is this also the case for the spatial mismatch? The SMH is considered to be a very dynamic phenomenon – e.g. discrimination (and the coinciding segregation) against some minorities tends to decrease over time, while against others it actually increases. These developments, and others, will have an effect on the SMH (Iceland & Sharp, 2013). Accordingly, it is important to approach the SMH for more than one moment in time. By doing this, patterns and trends can be observed as well.

First, *how are minorities distributed in the Phoenix metropolitan area?*
Second, *how is employment distributed in the Phoenix metropolitan area?*
Third, *do minorities face worse employment accessibility as opposed to non-minorities?*
Fourth, *how do these patterns develop over time?*

This study uses U.S. Census data to measure the degree to which minorities – that reside in Phoenix – face a spatial mismatch to employment (Census Bureau, 2018). Census tracts are the unit of observation: i.e. neighborhoods with approximately 4000 inhabitants. The Census Bureau provides data with a wide range of variables, including commuting related variables, on this geographical level.

First of all, descriptives are examined. Maps are constructed to see to what degree minorities are spatially clustered and how this pattern develops over time. Thereafter, urban centers in Phoenix are determined with the method of Giuliano (2007). These urban centers will be represented on a map to examine how they are distributed across Phoenix and how this distribution changes over time.

Thereafter, several regression analyses are performed. Mean commuting duration per census tract – one of the variables – is used as a proxy of a spatial mismatch and functions as the dependent variable. Long commutes indicate the presence of a spatial mismatch while short commutes do not. The share of minorities in a census tract functions as the independent variables of interest. Additionally, a wide range of control variables is added based on data of the U.S. Census, as well as a constructed variable based on distance towards urban centers. This is done for several time periods to see how commuting behavior of minorities changes over time.

The findings for the year 1990 demonstrate that the share of Hispanics in census tracts is positively related to commuting duration. Particularly census tracts in the CBD with large shares of Hispanics face long commutes. The findings for the year 2000 show a similar pattern, yet the effect is smaller. There is no significant difference in commuting duration between minorities in more recent years – i.e. 2010 and 2015. This paper concludes that Hispanics faced a spatial mismatch in the past, yet in more recent years this mismatch cannot be observed anymore. I do not find supportive evidence for a spatial mismatch for minorities other than Hispanics.

The rest of the paper is structured as follows. Section 2 discusses existing literature regarding the Spatial mismatch Hypothesis and the relation of this phenomenon with the suburbanization of jobs. It also discusses the segregation and clustering of minorities, and the case of Phoenix. Section 3 discusses the data sources, the research area, and the research approach to answer the research question. Section 4 presents the results. Section 5 comes up with concluding remarks. Section 6 discusses the results in a bigger context, elaborates on the shortcomings of this study, and proposes opportunities for future research.

2. THEORETICAL FRAMEWORK

2.1 The Spatial Mismatch Hypothesis

A simplified explanation of the SMH is that there are fewer jobs per worker near minority-dominated areas than non-minority dominated areas (Ihlanfeldt & Sjoquist, 1998). The premises of the SMH are the following:

1. Labor demand has shifted away from minority-dominated neighborhoods to mostly suburban areas.
2. Racial discrimination – mainly in the housing and mortgage market – has prevented minorities from moving to job growing regions in the suburbs.
3. Factors like poor information about distant job openings, customer discrimination against minorities, and inadequate public transport linkages between minority-dominated neighborhoods and job-growth areas have restricted minorities to work in job-rich areas.

Kain (1968) emphasized central city minorities and the exodus of employment from the central city to the suburbs. However, this dichotomy between inner-city and suburbs no longer holds. Suburban centers now face similar problems as the central city (Orfield, 1997). Consequences of a spatial mismatch are higher unemployment rates, longer commutes, and lower real wages.

There are several underlying mechanisms that explain why being far away from job opportunities can be harmful and initiate bad labor market outcomes; on the supply side as well as the demand side of labor (Gobillon et al. 2007).

Spatial Mismatch & (un)employment - Supply Side

Firstly, workers might refuse a job opportunity because the commuting to the job involves too many costs in view of the anticipated wage. Coulson et al. (2001) show that adverse labor market outcomes of central city minorities can be explained by the high commuting costs faced by these inner-city residents.

Secondly, workers that live far away from jobs have a lower chance to find a job because they get less information about distant job opportunities. Workers may have little information about suitable job offers, and in they can end up looking for jobs in the wrong locations (Gobillon & Selod, 2014). Especially for low-skilled service jobs, recruiting methods are very local, via advertisements in local newspapers or restaurant managers who use ‘wanted’ signs to reach potential employees. Also, if the general unemployment level is high it becomes even more difficult for individuals to rely on personal connections to lead them to jobs; because many of people in the neighborhood are unemployed themselves (Calvo-Armengol, 2004; Battu et al., 2011).

Thirdly, workers may incur high search costs[1] which limit their spatial search horizon to just the borders of their neighborhood, and this search area is sparse in terms of job opportunities. The consequence is an unsuccessful job search attempt.

Spatial Mismatch & (un)employment - Demand Side

Firstly, suburban employers may assume that inner-city residents have bad work habits, or that they are more criminal and dishonest (Gobillon, 2007). Consequently, suburban employers are less likely to recruit minority workers living in the inner-city.

Secondly, workers with long commutes tend to be less productive and so employers prefer workers who live close to the work location. This especially counts for certain service jobs, like working in restaurants. These jobs involve long breaks during the day and workers who live nearby can go home and relax while workers living further away cannot. Consequently, firms could determine geographical boundaries beyond they will not search for workers.

These mechanisms explain *why* a spatial mismatch is bad and leads to worse labor market outcomes. But then again, how does a spatial mismatch emerge in the first place? This has to do with two developments: (1) employment moving from the inner-city to suburban locations and (2) clustering of minorities due to racial discrimination.

2.2 Employment: The Exodus of Jobs to the Suburbs

In conventional urban models, firms can benefit from agglomeration economies (McCann, 2013). They are willing to pay high rent to locate in the central city, and accordingly employment clusters in the Central Business District (CBD). Workers live around the CBD since they do not benefit from agglomeration economies and have other – often idiosyncratic – preferences.

Workers face costs to travel to work every day. However, this is offset by the less expensive

[1] Search costs are costs involved in looking for a job, e.g. the effort to look for jobs somewhere (Cahuc et al., 2014).

land further away from the center. The implication for commuting is that living further away from the CBD generally coincides with increases commuting distance, time, and costs.

More recently, the monocentric urban structure dissolved into more than just one CBD. Employment suburbanized and firms cluster in suburbs, surrounded by residential areas (Anas et al., 1998). Firms in these new suburban centers value agglomeration economies as well and they are attracted by suburbs because of the cheaper land rents, less congestion and more efficient transport (Liu & Painter, 2012). Gottlieb (1995) argues that there is also an interaction between the location preferences of firms and workers. High-quality amenities, often found in suburbs, enables firms to pay lower wages, which incentivizes firms to move to potential workers in the suburbs.

2.3 Minorities: Segregation & Clustering

Residential Segregation

Early studies about residential segregation focused on the strong black-white divide in U.S. Cities (Burgess, 1928; Myrdal, 1944). This division between Black and White was seen as nearly impenetrable, intensified by discrimination and violence towards African Americans (Clark, 1965). Consequently, the residential segregation of African Americans in U.S. cities is considered to be high in absolute terms, yet it has steadily declined in recent decades (Iceland & Sharp, 2013).

After the 1980s there has been a growing interest in residential patterns of Hispanics, Native Americans, and Asians; which are the most discriminated minorities after African Americans. In contradictory to African Americans, discrimination against Asian and Hispanic minorities have not declined after 1980. In turn, this discrimination – which coincides with spatial segregation – leads to worse job accessibility and fosters a spatial mismatch (Turner, 2008).

Benevolence & Benefits of clustering

Literature discussing the SMH mainly sees the segregation or clustering of minorities as disadvantageous – i.e. minorities are segregated and consequently located far away from job opportunities. However, this is not necessarily the case. Often, minorities *want* to cluster. They establish internal markets and generate employment in their enclaves (Epstein, 2002; Kasarda, 1989).

Epstein (2002) discusses the *herd effects* and *migration networks*, and they mainly apply for minorities with positive net migration, i.e. Asians and Hispanics. The theory of *herd effects* states that the location decision of migrants going to the U.S. is based on imperfect information. Their location decision is often based on previous immigrants' decisions; i.e. they are more likely to locate nearby preceding immigrants. *Migration networks* are the effects of social ties – like kinship and friendship – on the location decision of new immigrants. Former immigrants maintain networks with residents from their homeland. New migrants, therefore, have better information about the labor market in the host country, which can increase the expected wage of this migrant, particularly in the area where former migrants have settled.

Kasarda (1989) states that particularly Hispanics and Asians have been able to establish internal markets and generate employment in their enclaves; employment which is relatively isolated from the national economy. Many family-operated businesses continuously reinvestment their profits. While these firms expand, they favor 'members of their own' when hiring workers. Kotlin (1988) shows that a dollar turns over five times in a Chinese community compared to just once in African American communities. Furthermore, minorities of foreign origin are often overrepresented in entrepreneurial activity (Fischer & Massey, 2000). The *disadvantage theory* depicts entrepreneurship as a survival strategy for minorities that encounter a barrier to local labor markets, like poor English skills, limited educational attainment, and discrimination.

These local multipliers have a positive effect on the labor market accessibility[2] for (some) clustered minorities. Small businesses in these enclaves offer job opportunities for minorities which are likely to live very close by. Consequently, these local multipliers counter the SMH since living segregated can actually increase accessibility to jobs. Figure 1 represents a conceptual model.

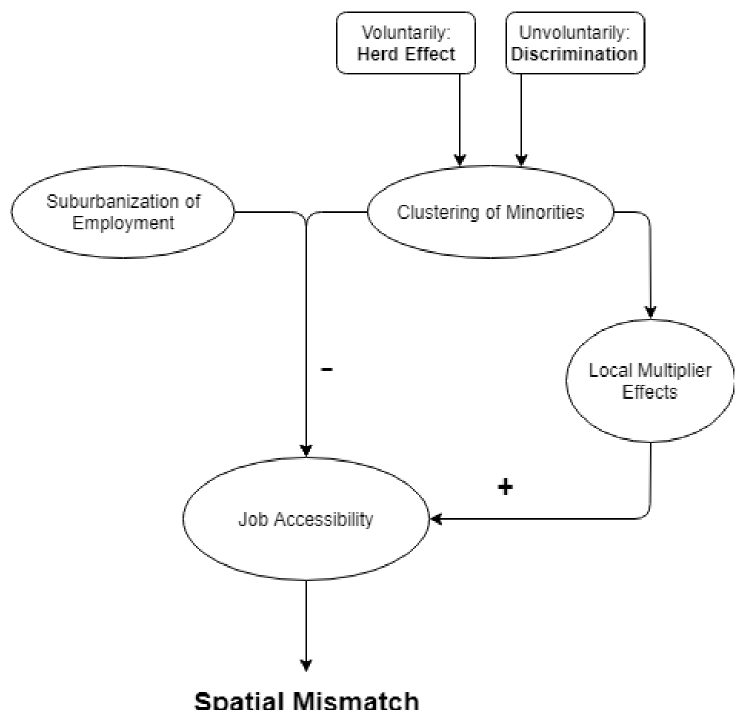


Figure 1

[2] Take note that labor market accessibility in this study is defined as *spatial* labor market accessibility, as in distance to the labor market in a spatial way.

2.4 Local Factors: The case of Phoenix, an extensive car paradise

Like other cities in the Sunbelt – which roughly stretches from Florida to California – Phoenix has grown into a large city during the automotive era in mid-20th century till the 21st century (Census Bureau, 2018). Nowadays Phoenix is one of the most sprawled urban areas in the U.S. with a very low population density and a large share of low-rise buildings (Ross, 2011). Sunbelt cities have different city structures compared to old industrial cities like e.g. New York, Boston, and Philadelphia. They rely much more on road infrastructure and car use. These old industrial cities often have a monocentric city structure with a large and dominant CBD in the center, with suburbs around it. Sunbelt cities are more diversified and often have many subcenters.

Glaeser et al. (2009) argue that public transportation heavily relies on the density of jobs and amenities. So while a city sprawls, public transport becomes less viable. Hence in Phoenix, public transport is less efficient compared to a city like New York – because New York is much denser in terms of jobs, amenities, and people. Consequently, in Phoenix, public transport as a mode of commuting is just 2.2%, while in New York this number is close to 30% (Census Bureau, 2017).

Still, one major development concerning public transportation in Phoenix has to be mentioned. In 2008 the *metro valley rail* is put into operation and it serves three cities in Phoenix, i.e. Phoenix, Mesa, and Tempe. With approximately 16 million passengers per year, and daily 45 thousand passengers the rail is regarded as a big success (Valley Metro, 2017; Liu, 2014; New York Times, 2009). Particularly residents living in the inner city benefit from this metro rail, which positively improves their labor market accessibility. Liu (2014) argues that the light rail mainly improves job accessibility of Hispanic dominated neighborhoods and lower-income groups. Between 1990 and 2015, the use of public transport as mode of commuting increased with 15%, while in most other U.S. cities the use of public transport generally declines. Yet, the share of public transport as travel mode to work is still marginal.

2.5 Hypotheses

As the conceptual model shows, minorities can benefit from clustering while it also can be a disadvantage. Since it can go both ways, and because of the explorative nature of this study, I will not form hypotheses. The formed sub-questions will be the guidelines in this study.

3. RESEARCH DESIGN

3.1 A Quantitative Study in Phoenix, Arizona

A Quantitative Research Approach

This study will mainly use quantitative research methods in combination with secondary data. This decision is based on the following. *First*, quantitative research methods are more suited to study large populations (Clifford et al., 2016). The city that will be studied, Phoenix, is an urban area with approximately 4.5 million inhabitants (Census Bureau, 2018). The Census Bureau provides data that covers the whole of Phoenix, leading to a very representative sample. Qualitative research methods on the other hand face difficulties with representing such large populations.

Second, one of the benefits of qualitative research methods is that it is able to answer questions like: how do workers perceive a spatial mismatch? Conducting interviews would be useful to answer such questions. However, this study searches for patterns and relations that could indicate the existence of a spatial mismatch. Results of this study could be thought-provoking and lead to follow-up studies which use more qualitative research methods.

Third, the U.S. Census is a very comprehensive dataset with a large variety of socioeconomic variables and many of them can be used as control variables when the spatial mismatch phenomenon is examined.

Fourth, the U.S. Census is held repeatedly, leading to observations over time. In this study, the years 1990, 2000, 2010, and 2015 will be studied (I elaborate on this further on). Because data is available for several years, the SMH can be examined over time, to see if there are trends – e.g. does the spatial mismatch intensifies for Hispanics, while it weakens for African Americans?

A case study in Phoenix

Phoenix will be used as a case study because of three reasons. *First*, I lived in Phoenix for a while, so this background knowledge about Phoenix helps while interpreting the results of this study. *Second*, I got access to data sources of Arizona State University (ASU), data which is very useful for this study. *Third*, most emphasis of the SMH goes to cities in the Midwest and the east coast, cities that often contain large shares of African Americans (the minority Kain’s first study initially focused on). In the more recent years, other minorities (i.e. Hispanics and Asians) boom in cities located in the Sunbelt. Raphael and Stoll (2002) state that the spatial mismatch is also an issue for these minorities. *Fourth*, since cities in the Sunbelt are relatively similar to each other, in terms of spatial setup, prominent industries, and ethnical composition, the recommendations of this study can be useful for other cities in the Sunbelt as well.

3.2 Data Sources & Research Area

U.S. Decennial Census

For the years 1990 and 2000, the U.S. Decennial Census offers data about a variety of socio-economic variables; like commuting characteristics, race, ethnicity, income, and more (Census Bureau, 1990; 2000). The data they offer is aggregated on different geographical units – i.e. national, state, county, census tract, and block group. The smallest *reliable* estimates are available on census tract level. Census tracts are considered *neighborhoods* and they are designed to be relatively homogeneous units with respect to population characteristics. Furthermore, they contain approximately 4000 inhabitants.

Data on an individual level would be preferable since you can make statements and predictions for individuals, but this data is not available. Again, to collect such a dataset myself which is also representative for Phoenix as a whole is not feasible and too time-consuming. Therefore Census data is a good alternative.

American Community Survey (ACS)

Since 2005, the American Community Survey (ACS) has taken over a lot of survey questions of the Decennial Census (Census Bureau, 2018). The ACS contains the same survey questions as the decennial census, but the way the data is collected is different. While the Decennial census collects data within one year, the ACS collects the data over a period of several years. To get reliable estimates for census tracts there are so-called *five-year estimates*, and this has consequences for the data. For example, data in this study for the year 2010 represents data collected between 2008 and 2012. Because the ACS is held continuously, the year 2015 can be added as well. Consequently, the years 1990, 2000, 2010 (*2008 – 2012*) and 2015 (*2013 – 2017*) are chosen to create some consistency in the time periods.

The Census Bureau (2004) warns that the difference in the way data is collected has consequences. Estimates related to work characteristics (e.g. employment, unemployment, commuting time) can be affected. While data of the Decennial Census is collected from March till August, the ACS collects data the whole year-round. So, for example, seasonal workers that are surveyed can be considered unemployed in the Decennial Census, while employed in the ACS. This is a limitation of the data and has to be taken into account[3].

Census Transportation Planning Products (CTPP)

The Census Transportation Planning Products (CTPP) provides data about job and employment locations (Bureau of Transportation Statistics, 2000). With this data, employment density across space can be determined. CTPP offer data for the years 1990, 2000, and 2005 onwards. In 1990 and 2000, data about employment density is only available on an aggregated level, while from 2005 onwards data about exact job locations is available.

Topologically Integrated Geographic Encoding and Referencing (TIGER)

At last, Census and CTPP data can be linked to spatial data. Topologically Integrated Geographic Encoding and Referencing (TIGER) provides uw with spatial data which can be used with GIS software. For example, ACS data can be linked to TIGER data, to show data on a map (Census Bureau, 2018).

Unit of Observation – Census Tracts

The *census tract* is the unit of observation in this study. They are preferable over *block groups* (one geographical unit lower) because the estimates of block groups are often not reliable (because of very high standard errors). Counties, one geographical unit higher, are not preferable as well because Phoenix consists only out of two counties, making the sample very small.

A strength of Census data is that it covers all areas in Phoenix. The data collection process of the Census Bureau is very intensive and they are able to reach a large portion of the population. Accordingly, the sample is quite representative.

Still, there are also some weaknesses. There is variation within census tracts as well, but because each tract is seen as a *case*, data is automatically aggregated. For example, the mean income of a census tract can disguise that there are large differences in income within this tract. Nevertheless, the Census Bureau states that the tracts are designed to be homogeneous units with respect to population, but a clear explanation on how they construct them is absent. Another

[3]Hu (2014) and Hu and Wang (2016) also use employment charactereistics of the Census and perform over time analyses. They argue that it is acceptable to compare the Decennial Census with the ACS Census.

weakness is that not all residents in the tracts are surveyed. So the estimates of the Census data still contain a margin of error.

Research Area

Phoenix is the research area in this study. According to the Whitehouse (2018), the Phoenix metropolitan area consists of the counties Maricopa and Pinal. As a reference, the size of both counties together is close to the size of the Netherlands, in terms of land.

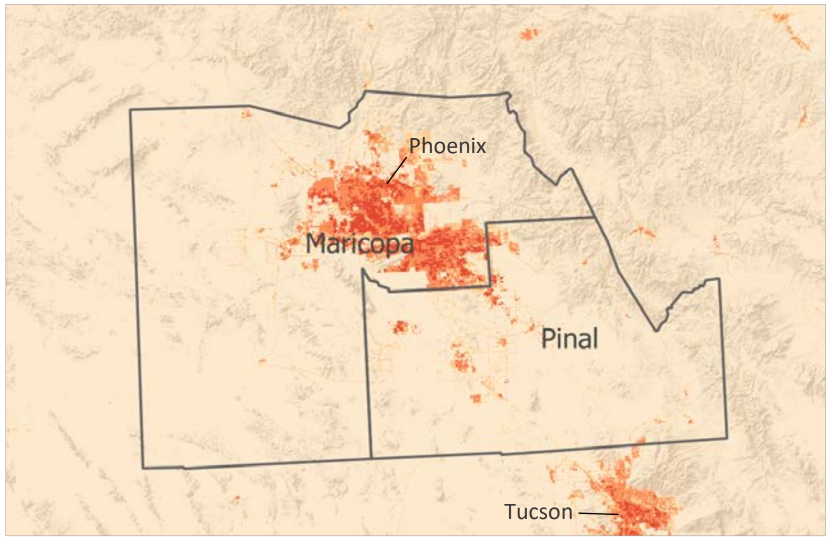


Figure 2 - Urban Area Phoenix (counties Maricopa and Pinal

Figure 2 shows the two counties. It also shows a density measure of human activity (CIESIN, 2016). The lightest color is considered *very rural* and has a population density below 100 residents per square mile. As you can see, most of the area is non-urban. Approximately 90 percent of the land with the lightest color consists out of agricultural land, public land, and open space (MAG, 2019). The other 10 percent consists mainly out of vacant space, vacant state trust, and water. Most of these terms are more related to *rural* than *urban* (Woods, 2010). Therefore the definition of the Phoenix metropolitan area will be altered.

The Census Bureau (2000) defines an area as *urban* when the population density of census tracts is at least 1000 people per square mile (386 per square kilometer). This definition will be used for Phoenix since it better copes with the actual urban area of Phoenix.

Area & Time Dimension

To perform an analysis over time, the data of the different years have to be made consistent. In this study, the urban area in 2000 will be used as a baseline – see figure 3. This approach is similar to the one of Hu (2014). The year 2000 is used since it requires the least changes in census tract conversions.

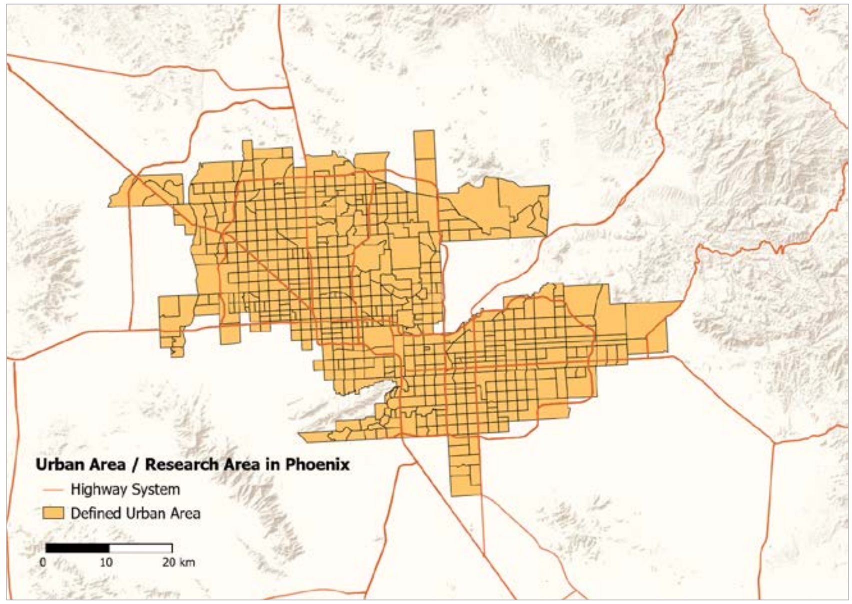


Figure 3 - Defined Urban Area Phoenix

The census tracts of the different time periods have to be made consistent by converting the census tracts of 1990, 2010, and 2015 to the geography of the tracts in 2000 – which is a quite complicated process. Appendix A.1 will further discuss the procedure on how to convert the census tracts to the year 2000. From now on, the data from different time periods can be compared more easily because the sample size stays the same.

It is important to note that choosing a baseline has consequences. Over time Phoenix has grown rapidly in terms of population, but (therefore) also in terms of space. Areas that are urban in 2010 and 2015 are not considered urban in 2000 while areas that are rural in 1990 are urban in 2000. Still, large parts of the urban area in 2000 matches those of 1990, 2010, and 2015.

Ethnicity & Race

The main topic of this study is the spatial mismatch of minorities; therefore the definition of *minorities* needs to be clear. The Census Bureau defines both and *race* and *ethnicity*. As race, persons can define themselves as White, African American, Asian, American Indian, Alaskan Native, Native Hawaiian, Other Pacific Islander, or other race. In this study, the latter three are excluded since their share in the census tracts is often equal to zero. Also, the error margin for these groups is very large, which leads to unreliable estimates.

Furthermore, ethnicity only is made up out of two groups: Hispanics and non-Hispanics. Persons have to answer both a question about race and ethnicity. Hence, a person that defines him or herself as Hispanic, can also be a Hawaiian Native. In most cases, Hispanics report themselves as 'White' on the census form (Humes et al. 2011). But sometimes Hispanics report that they are 'other race' as an alternative for white. In the census of 2010, 53% of the Hispanics identified themselves as white, 37% as other race, and 3% as African American. This problem is recognized by researchers investigating the spatial mismatch hypothesis (see e.g. Shen, 2000). To deal with this, the group Hispanics will only consist of Hispanics that define themselves as White. Consequently, the group other race will also contain a large group of people who see themselves as Hispanic. This is important to take into account while interpreting the results.

All in all, we have discussed why a quantitative research method is chosen, with Phoenix as a case study. We have also elaborated on the used data sources, the research area, and the definition of minorities. *Next*, the approach to answering the research and sub-questions will be discussed.

3.3 The Approach

Commuting Time as Measure of Labor Market Accessibility

A spatial mismatch reflects worse (spatial) accessibility to employment. This accessibility can be measured in different ways. Many studies regarding the SMH use commuting duration as a measure for labor market accessibility (see Shen, 2000; Hu, 2015) and longer commutes for minorities (than non-minorities) would be evidence to support the SMH. On the other hand, Taylor and Ong (1995) use commuting time as well as commuting distance. They found that particularly African Americans and Hispanics travel longer than other groups, but their distance to jobs was the same. Taylor and Ong argue that the longer commutes are due to the use of different transportation modes. African Americans and Hispanics use more public transportation, which on average use more time to cover the same distance compared to another dominant mode of transportation: the car.

We use commuting duration as a measure of employment accessibility because (1) the U.S. Census contains data about the mode of commuting which can be used as control variables, and (2) data about commuting distance – on census tract level – is not available. Consequently, when census tracts face (relatively) long commutes it supports the SMH.

Descriptive Statistics

First of all, some descriptive statistics will be shown. Histograms and maps will be constructed to see how commuting duration is distributed. The emphasis will be on the commuting duration of census tracts in and near the CBD. If these census tracts face long commutes, it could be the first indication of a spatial mismatch.

Distribution of Minorities

The first sub-question is: *how are minorities distributed across Phoenix?* One of the premises of the SMH is that minorities tend to be clustered near the CBD. Therefore we focus on two aspects: (1) to what degree do minority groups tend to cluster together? and (2) where do they tend to concentrate? For this purpose, maps will be constructed.

Distribution of Employment

The second sub-question is: *how is employment distributed across Phoenix?* For this purpose, we will determine employment centers. Employment traditionally have clustered in the CBD, while over time subcenters of employment have emerged. The CBD and subcenters together can be called urban centers (Gregory et al., 2011).

Urban centers can be determined by looking at job densities. An often-used technique to determine urban centers is performed by Giuliano et al. (2007). They use geographical units – often census tracts or Traffic Analysis Zones (TAZs) – which contain information about the number of jobs and the density of jobs. In their view, an urban center is an area, consisting out of one or more adjacent census tracts/TAZs, with a certain density of jobs and a certain total number of jobs. Alternatively, Leslie (2006) and Helbich & Leitner (2010) have access to data about exact job locations, so they use a kernel density approach to determine employment hotspots. The latter approach is more precise since it uses exact job locations instead of aggregated data.

Yet, in this study, the method of Giuliano et al. (2007) will be used. This decision is based on data availability. Employment data based on job points is actually available for Phoenix, but unfortunately not for the equivalent years of the census data. Data about employment locations is only available for the years 2004 and onwards, while data on the aggregated level is available for all matching years. Appendix A.2 exactly explains how centers of employment are determined for Phoenix.

One of the premises of the spatial mismatch hypothesis is that subcenters emerge at the outskirts of the city. But does this actually happen? Moreover, in the regression analyses, two variables will be added based on the distance to the CBD and the distance to subcenters. This will be further explained in the next section.

Regression Analyses

The third sub-question is: *do minorities face worse employment accessibility as opposed to non-minorities?* For this question, an OLS multivariate regression analysis will be performed to find out if minorities in general, or specific minorities, face longer commutes than non-minorities.

Since the census tracts are made consistent over time, it is possible to use the data of different years as panel data, or unite the data and perform a pooled OLS regression. However, we want to observe how the commuting behavior of minorities changes over time – and it is very likely that it does change (e.g. less discrimination against certain minorities, or more decentralization of employment). Consequently, the data will not be merged or used as panel data. Instead, four different models will be constructed for each time period.

Dependent Variable

The dependent variable in the regression analyses is the *mean commute duration*. For every census tract – the unit of observation – a mean commute duration is given. As noted before, long commutes correspond with bad labor market accessibility and could favor the SMH, while short commutes correspond with good labor market accessibility and could oppose the SMH. The commuting duration is given in minutes.

Independent Variables

Variables of Interest

The variables of interest are *African American*, *Asian*, *Hispanic*, and *Native American*. Since the data is aggregated on census tract level, the estimates for minorities are shares in percentages – e.g. 12% of the population in a census tract is considered *Asian*. This setup brings some problems in the regression analyses. One of the conditions of a correct regression does not hold; i.e. when one predictor changes, the others stay constant (ceteris paribus). So, when the share of African Americans in a census tract decreases, this automatically leads to a relative increase of other groups. This has two consequences. *First*, the coefficients for the minorities are slightly biased. *Second*, one group must be left out, to prevent very high or perfect multicollinearity between the minority variables.

Control variables

(1) Socio-economic

The following variables will be used as control variables: *mode of transport, having a car, median household income, share of female workers, educational attainment*, and *share employment in an industry*. This set of control variables will be used because previous research have stressed them as important predictors of commuting duration. In Appendix A.4, this will be further discussed. All variables are listed and summarized there as well.

(2) Spatial Setup

Studies about the SMH before 2000 are often based on the dichotomy between central city and suburbs (see e.g. Taylor & Ong, 1995). This approach is very crude and incorrect because the role of subcenters is completely ignored. After the turn of the century more sophisticated methods were developed, mainly because of advances in GIS software and the rise of spatial data. Two methods have become dominant to control for the spatial setup of cities in examining the phenomenon of the SMH.

Wang (2000) uses the distance between census tracts and urban centers as an explanatory variable for the commuting duration. More specifically, he uses two variables; distance to the CBD and distance to the nearest subcenter. By doing this, he *controls* for distance. So, perhaps minorities face longer commutes, but if they live far away from the CBD, they probably choose to live there (because of idiosyncratic preferences). By using the distance to the CBD and nearest subcenter as predictors of commuting duration, you control for the spatial setup of a city.

Shen (2002) on the other hand constructs an accessibility measure that he implements in his regression models. This accessibility measure is an index based on commuting flows within a city, distance to the biggest job cluster, and mode of transportation. Shen's approach is especially suited for a city with one big employment center – he studies Boston, which is a very monocentric city – while the method of Wang incorporates subcenters much better. The method of Wang is also more convenient since it does not require many calculations, and therefore this method will be used.

Accordingly, two variables will be added to the regression analysis. The first one is the distance[4] from a census tract to the Central Business District (*DCBD*) and the second one is the distance from a census tract to the nearest subcenter (*NSUB*), based on the urban centers determined in the previous section. The constructed variables will serve as control variables. In the Appendix A.3 the method to calculate these distances is explained, plus several limitations and how to deal with it.

Regression particularities

(1) Before regression:

It is important to see how the data about commuting duration is distributed across census tracts. So first of all, histograms will be constructed. If the distribution is not shaped like the normal distribution, a solution can be to transform the data – e.g. take the natural logarithm. The same goes for the independent variables and corresponding distributions.

Also, the data has to be analyzed thoroughly on errors. It could be that some values for cases exceed a point which makes them invalid. For example, the share of certain minorities could be higher than 100%, which is impossible. These observations must be removed.

(2) After regression:

It is important to test if multicollinearity and heteroskedasticity are present. If independent variables correlate too much with each other, the problem of multicollinearity arises. A result is biased estimates, something we do not want (Wooldridge, 2015). It is important to test for this

[4] Distance is defined as straight-line air distance from the center of a tract to the center of an urban center, this is without taking into account the street pattern. Another way is using the 'Manhattan' distance, which takes into account the raster street pattern in calculating distances. Wang (2000) show that both methods show similar results and straight-line distance even lead to a slightly better fit, thus the straight-line distance will be used.

since there are many independent variables involved. What the exact threshold is between acceptable and problematic is not clear. For convenience, the decision I make is that a VIF score higher than 10 would be seen as problematic (Hair et al., 1995).

To observe if heteroskedasticity is present, the residuals will be plotted and a Breusch-Pagan test will be executed. If necessary, robust standard errors will be used.

(3) Endogeneity:

Endogeneity can arise in many ways, for example when important independent variables are omitted or when variables have a two-way causal relationship (Wooldridge, 2015). Endogeneity is problematic as it leads to biased and inconsistent regression estimates.

We use a regular OLS regression analysis, and therefore we cannot rule out that there is reverse causality, i.e. the mean commuting duration of a census tract affects the share of minorities. As for omitted variables, we try to incorporate as many relevant control variables as possible. Still, some relevant predictors for the commuting duration cannot be measured. For example, cultural aspects in relation to commuting time cannot be added to the regression models. Other regression types, like 2SLS, are able to solve endogeneity issues. However, finding a suitable IV can be hard while it can also lead to other complications. For convenience, a regular OLS regression will be used. Consequently, we have to interpret the results with caution.

4. RESULTS

4.1 Descriptive Statistics of Commuting

Summary Statistics

First of all, we take a look at some descriptive statistics. Figure 4 shows the distribution of commuting time per census tract for the periods 1990, 2000, 2010, and 2015. Table 1 shows the number of cases and the mean commuting duration. As you can see, all time periods have the same number of observations. This is because the data is made consistent.

Year	Obs.	Mean Commuting Duration
1990	602	22,3 minutes
2000	602	26,0 minutes
2010	602	24,0 minutes
2015	602	24,5 minutes

Table 1

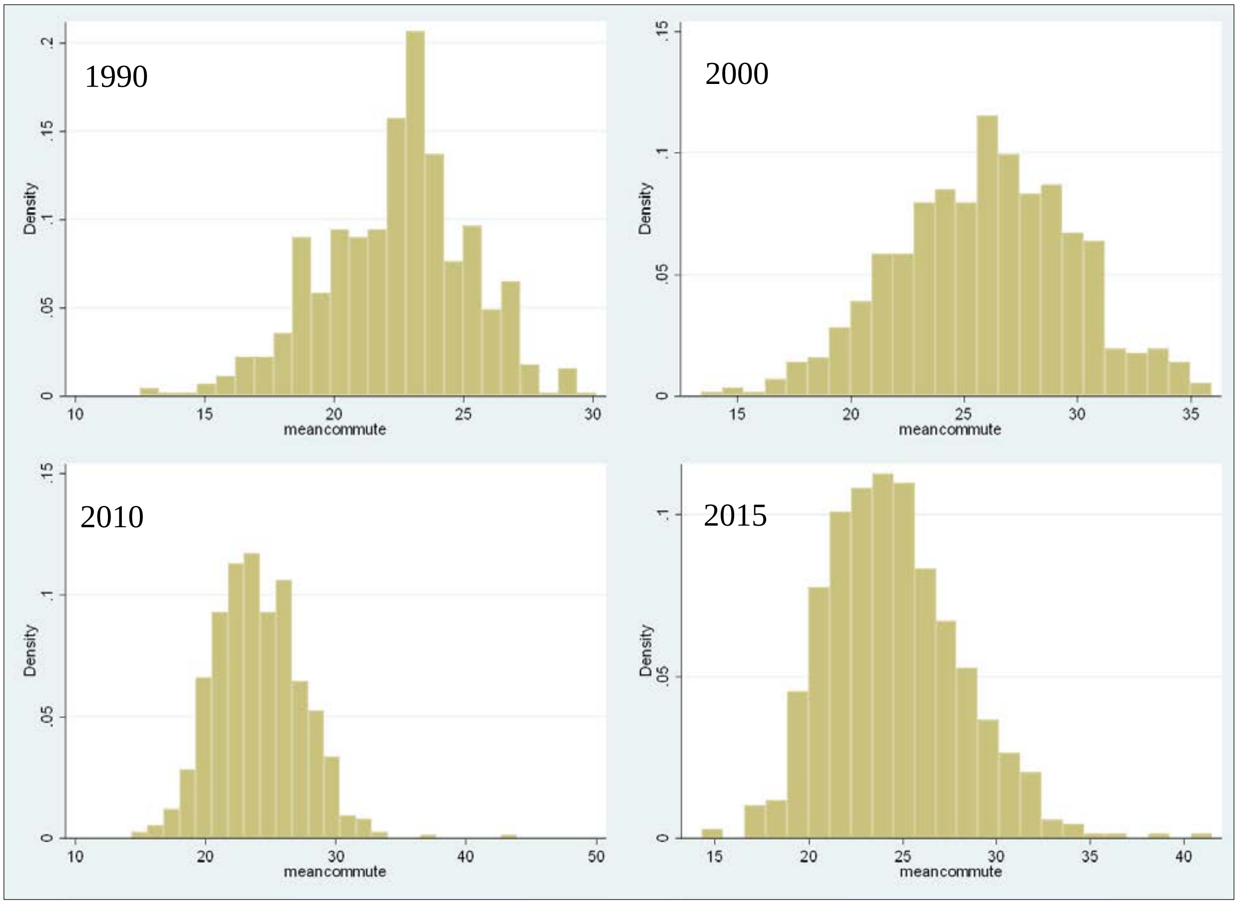


Figure 4 - Distribution of Commuting Duration per Census Tract

Particularly between 1990 and 2000 the mean commute duration (of all tracts together) increased substantially; with almost four minutes. After 2000 this number decreased again with two minutes. Looking at the distributions, particularly the years 2010 and 2015 show outliers of tracts with long commutes.

The range between the tract with the longest average commute and the tract with the shortest average commute is 20 minutes (in 1990). From 2000 onwards, the gap widens. This is important in relation to the spatial mismatch hypothesis because this large variation in commuting duration needs explanation. If there is no variation, there is most likely no spatial mismatch.

Spatial Descriptives

Figure 5 shows a map of Phoenix with city names. This map will help further on when we talk about specific areas in Phoenix. Figure 6 shows the mean commuting duration, per census tract, for all time periods. Starting with the year 1990 and 2000, a clear pattern can be observed as predicted by Alonso (1961) and Brueckner (2011). Close to the center of Phoenix, the commutes are shortest while they gradually increase when moving farther to the outskirts. Nevertheless, areas near the CBD do not face the shortest commutes. Especially areas in southern Scottsdale and Tempe face short commutes. Looking at the years 2010 and 2015, the pattern becomes more ambiguous. There are more census tracts with short commutes in the north and south. This particularly happens north of Phoenix and Scottsdale, and south of Tempe and Chandler.

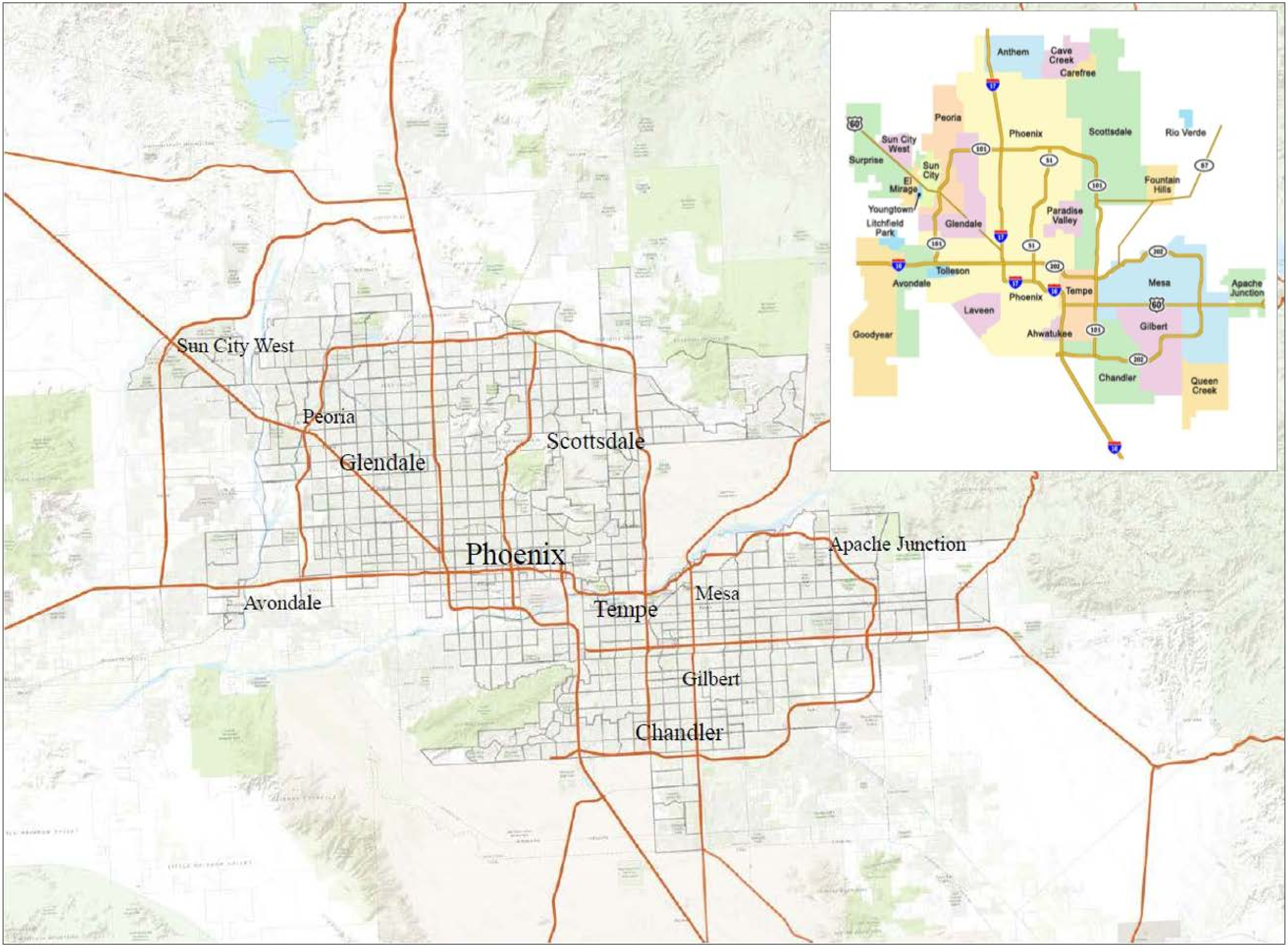


Figure 5 – Map of Phoenix (reference)

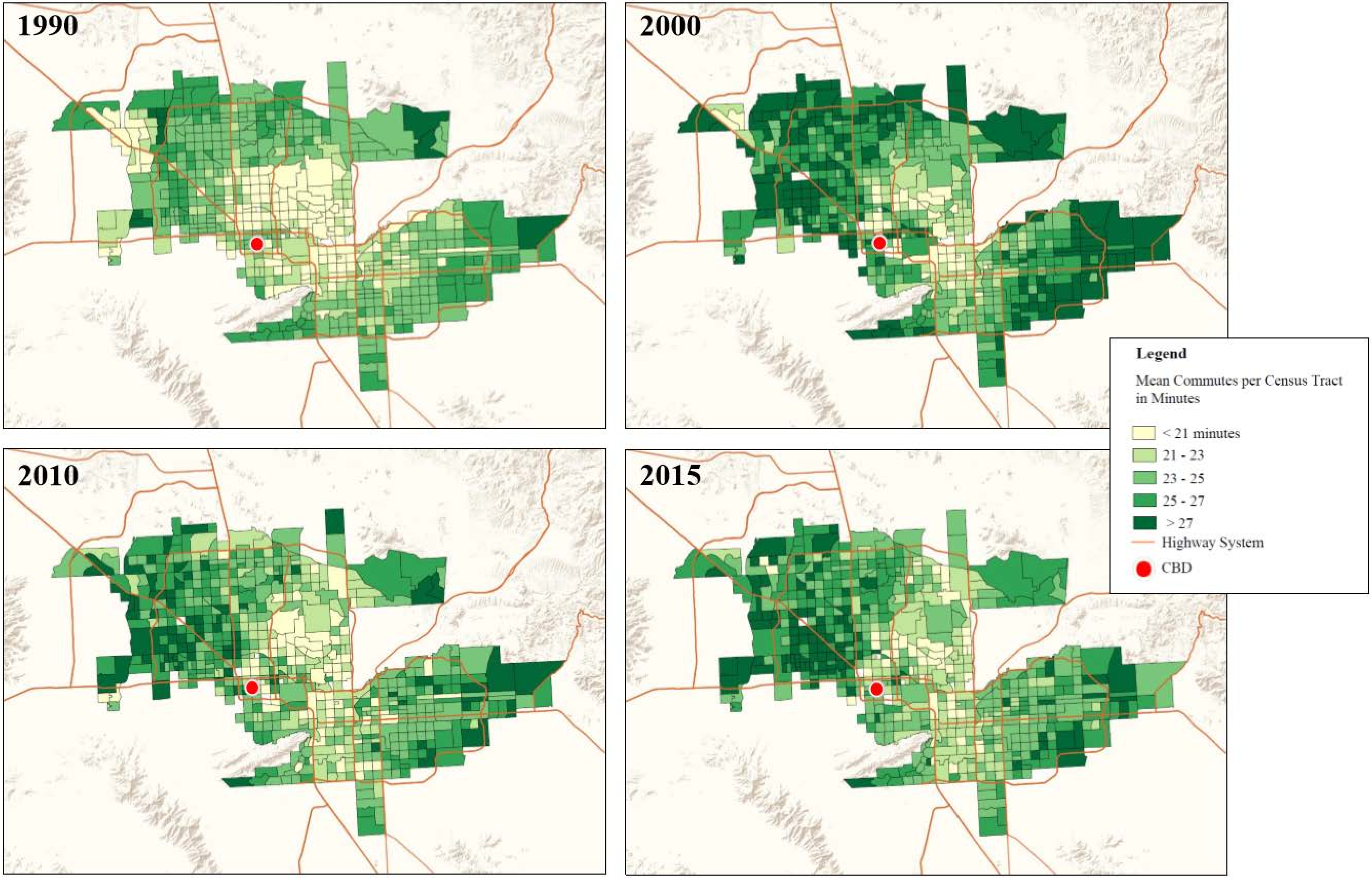


Figure 6 – Mean Commute Duration per Census Tract

4.2 Distribution of Minorities

Kain’s (1968) Spatial Mismatch Hypothesis states that minorities are clustered, mainly around the CBD. Mortgage and housing market discrimination prevents minorities to freely move around Phoenix. Therefore, a pattern of clustering of minorities could be an indication (or prerequisite) of a spatial mismatch.

African Americans

In figure 7 the share of African Americans per census tract is shown[5]. In the years 1990 and 2000 African Americans are very clustered, particularly south of the CBD. In the years 2010 and 2015, African Americans are much more sprawled. Census data (2018) shows that, over time, their share in Phoenix as a whole slightly increases, but census tracts with high shares of Afro-Americans – over 40% – do not appear anymore in 2010 and after.

Asians

In figure 8 the share of Asians per census tract is shown. The first thing to notice is the increase in the share of Asians over time. While in 1990 their share is neglectable, in 2015 their presence is much more significant. The clustering pattern is somewhat vaguer than African Americans. In 1990 and 2000 there is some clustering, but it is just in a handful of tracts. In the years 2010 and 2015, they are spread out in the whole area, although they also tend to cluster in certain places – especially in south Tempe, and not around the CBD.

Hispanics

In figure 9 the share of Hispanics per census tract is shown. Just like the share of Asians, the share of Hispanics has drastically increased over time. Hispanics tend to cluster; especially close to the CBD. Over time the clustering pattern is quite stable; the degree of clustering stays relatively similar. In 2010 and 2015, clusters of Hispanics also emerge in Mesa and Chandler.

[5] The legends of the maps with the share of minorities are all based on equal intervals – i.e. 5-10%, 10-15%, etc.

Native Americans

In figure 10 the share of Native Americans per census tract is shown. Native Americans do not tend to cluster very much. There is some clustering – especially in 1990 and 2000 – but their share in census tracts do never reach high levels – i.e. 99.5% of all tracts have a Native American share below 12%. There is one census tract outlier, with a share of 75%, in 1990, while in 2015 the share Native Americans in this tract decreased again to 39%.

All in all, mainly African American and Hispanic minorities, that reside in Phoenix, tend to spatially cluster. African Americans mainly do so in 1990 and 2000, while Hispanics tend to stay clustered in all time periods. Both minorities most often reside in the area around the CBD. Accordingly, these two minorities need more attention further on in this study.

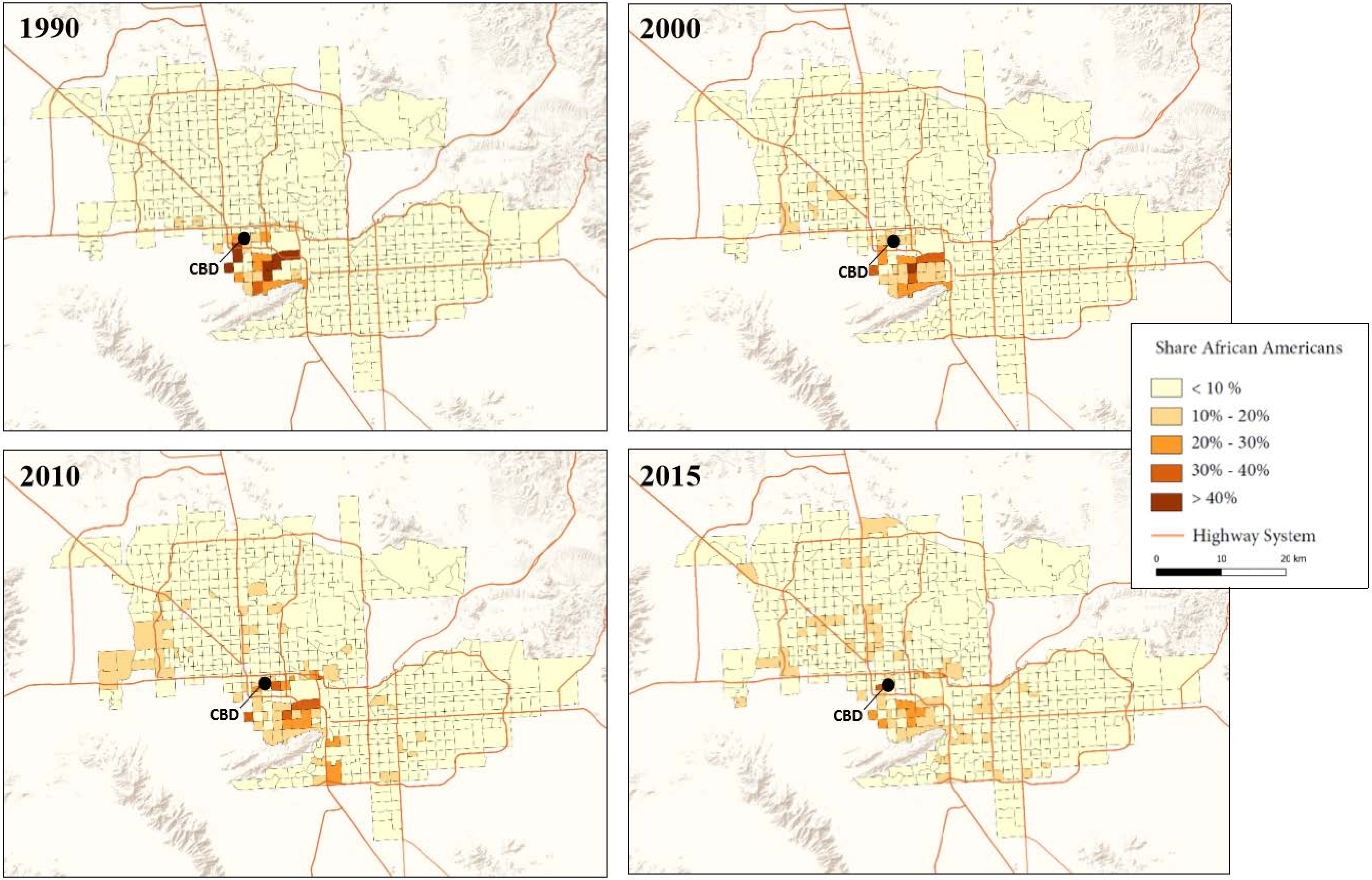


Figure 7 – Share African Americans per Census Tract

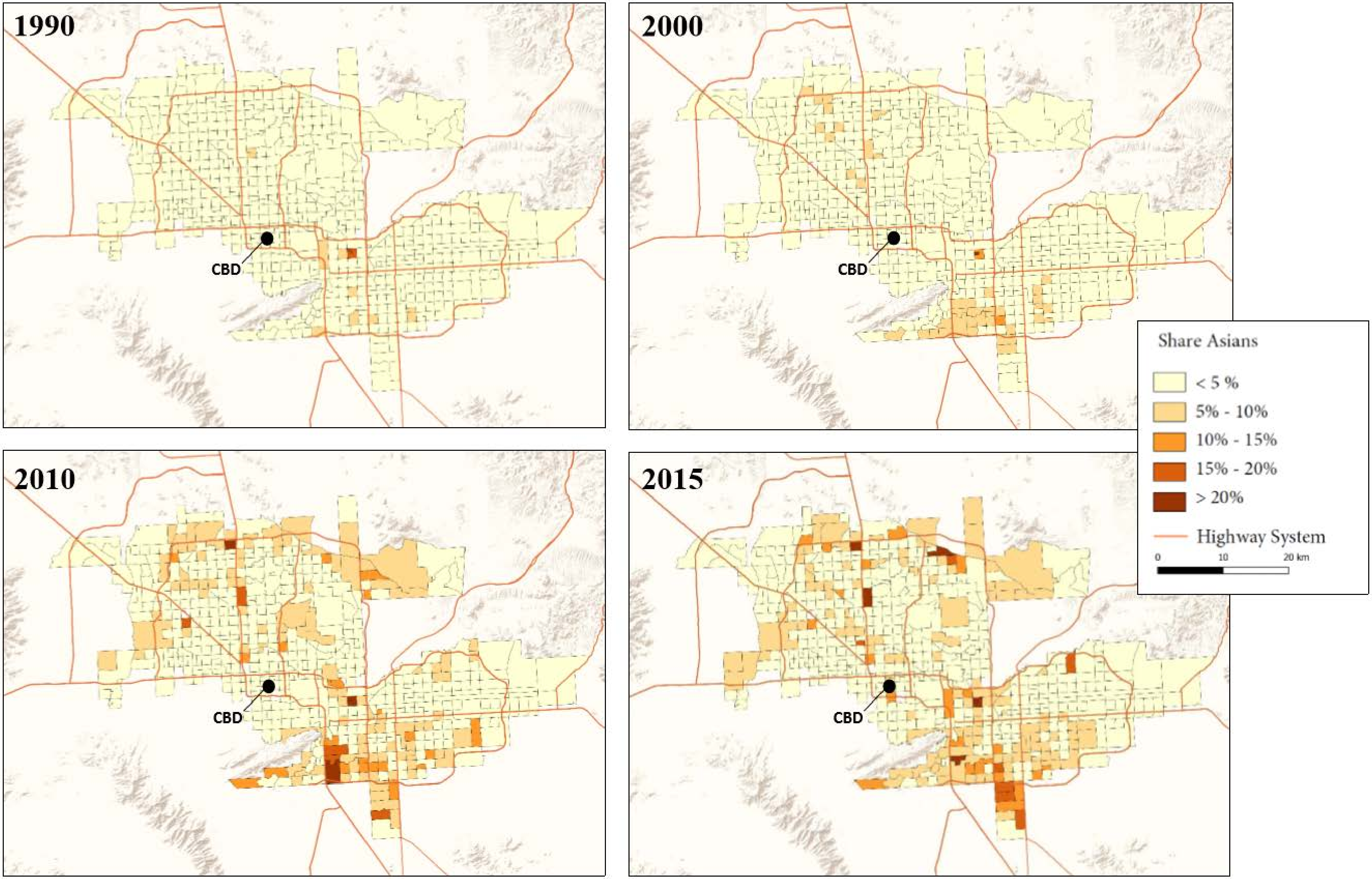


Figure 8 – Share Asians per Census Tract

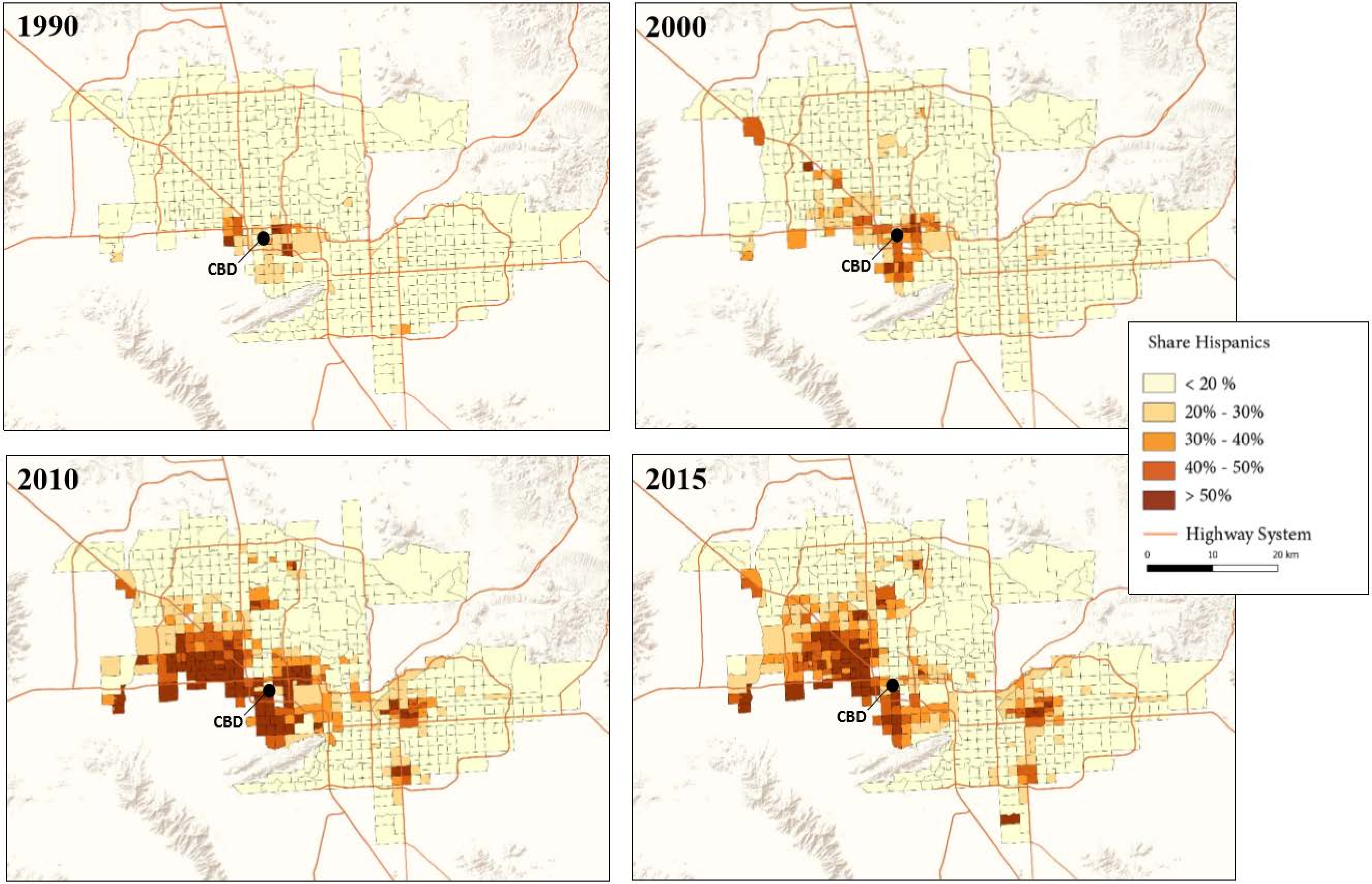


Figure 9 – Share Hispanics per Census Tract

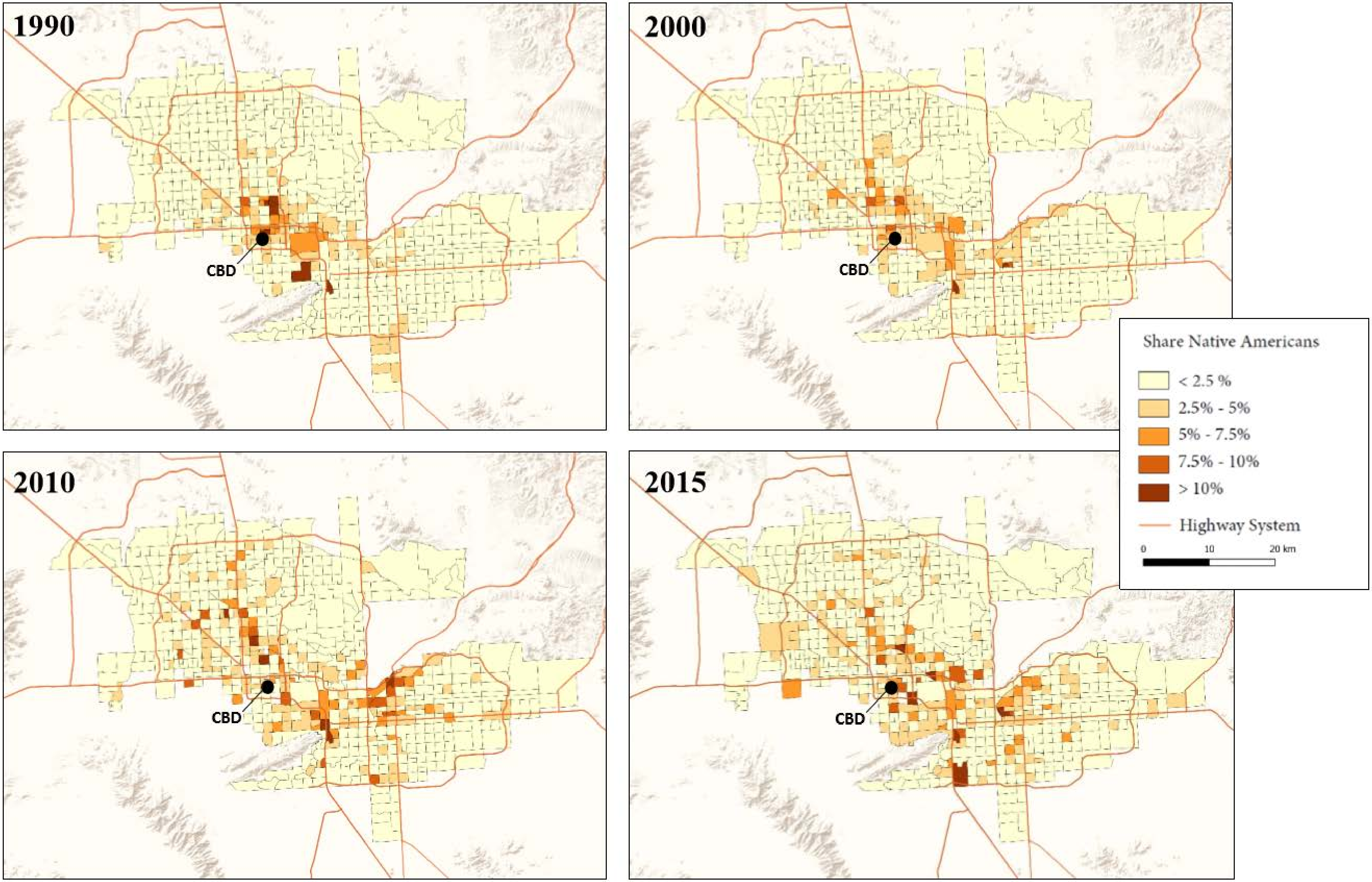


Figure 10 – Share Native Americans per Census Tract

4.3 Distribution of Employment

Another premise of the Spatial Mismatch Hypothesis is that employment is moving towards the suburbs. Therefore it is important to see how employment is distributed throughout Phoenix. In figure 11 the job centers – determined with the method of Giuliano et al. (2007) – in Phoenix are represented.

Urban Centers in Space

The CBD is always labeled with the number ‘1’. In 1990 centers can be found in Phoenix harbor airport area (‘2’), downtown Tempe (‘3’), downtown Scottsdale (‘4’) and the Camelback mountain area (‘5’). In 2000 a new center emerges in Mesa (‘4’) – located in the east. In 2010 the number of centers increases drastically. In the north, around the Scottdale Industrial Airpark (‘5’) a new center emerges, as well as in Metro Center (‘8’) and Deer Valley (‘4’). West of Downtown Phoenix an urban center emerges in the industrial District (‘7’). In the South, in Chandler, two new centers emerge, one near the south mountain range (‘3’) and one in the south part of Chandler (‘9’). In 2015 the urban centers are similar to those in 2010. In Chandler near the Chandler Regional Medical (’10’) center a new urban center arises.

Urban Centers over Time

According to the SMH, employment moves towards the suburbs at the expense of employment in the CBD. The maps in figure 11 show that in 1990, Phoenix had five urban centers, while in 2015 this number increased to seventeen. This pattern seems in line with the SMH, though in 1990 Phoenix was already fairly decentralized.

Yet, the CBD remains an important employment center. Looking at the area the CBD covers, it even grows over time. Also the total number of jobs in the CBD increase. Therefore, the growth of new subcenters is not at the cost of the CBD.

Hu (2016) examines commuting patterns in Los Angelos between 1990 and 2010. He argues that over time enough job opportunities remain in the CBD area, also for low educated minority workers. Accordingly, he states that inner-city minorities do not face a spatial mismatch. The maps in this study initially show a similar pattern, but we further need to control for other important factors – like education, job industry, etc.

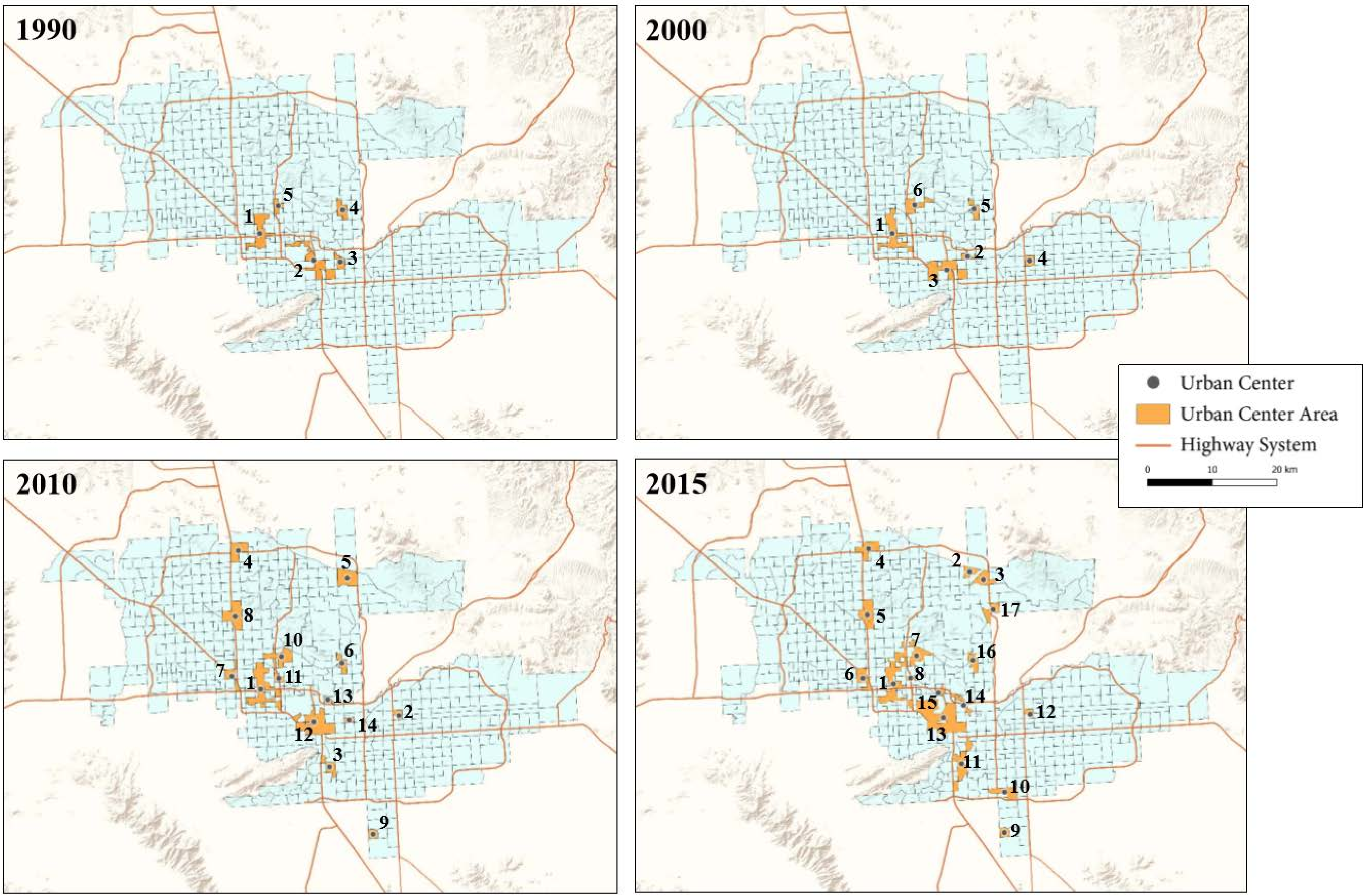


Figure 11 – Urban Centers in Phoenix

4.4 Explaining Variation in Commuting

The maps in the previous two sections show us important information. Particularly Hispanic and African American minorities that reside in Phoenix tend to cluster near the CBD. Yet, African Americans tend to sprawl over time. The maps about urban centers demonstrate that Phoenix has several urban centers, and over time this number increases heavily. However, while subcenters emerge and grow, the CBD area grows and expands as well. In this section, several regression analyses will be performed to examine if minorities – particularly African Americans and Hispanics – also travel longer.

Variables, Summary Statistics, and Correlation Table

In Appendix A.4 all variables that will be used in regression analysis are listed. Summary statistics and correlation tables can be found as well.

An interesting development, looking at the summary statistics, is that the (mean) share of Hispanics in census tracts grows drastically; from 7.5% in 1990 to 22% in 2015. Also, the (mean) share of public transportation as commuting mode increases from 2% to 2.8%, while this number in many U.S. cities drops (The Economist, 2018).

Model Results

Table 2 shows the results of the regression analyses. For each time period, three models are constructed. Model 1 tries to explain variation in commuting duration only by the share of minorities per census tract. Model 2 adds a range of socio-economic controls – i.e. income, mode of commuting, etc. – and industry controls to the regression analysis. The *socio-economic* and *industry* control variables are listed in Appendix A.4. Model 3 adds *spatial* variables, as in the distance of a census tract to the CBD and its nearest subcenter (NSUB).

The regression models do not have multicollinearity problems: in none of the models the VIF value of any variable is higher than ten – i.e. the threshold of problematic multicollinearity (Hair et al., 1995). Moreover, the Breusch-Pagan test shows that there are no problems with heteroskedasticity. Thus, robust standard errors are not necessary.

General observations

Overall, model 1 in all years is a very limited model and the R^2 stays quite low. Model 2, with socio-economic and industry controls, is a big improvement. The R^2 increases with 30 to 40 percent in all time periods and the adjusted R^2 increases likewise. In model 3 the spatial variables are added. These variables are in most models very significant – particularly NSUB – and the R^2 again increases significantly. The coefficients of the spatial variables are both positive, meaning that when tracts are located further away from an urban center, the mean commuting duration in these tracts are generally longer – ceterius paribus. Note that the coefficients are very low because the distance is measured in meters.

Particularly the coefficients of the minorities are heavily affected by the addition of the spatial variables in model 3. In 1990 the coefficient for *hispanic* switches from negative and significant to positive and significant. Meaning that according to model 2; a higher share of Hispanics within a census tract coincides with shorter commutes (ceterius paribus), while in model 3 a higher share of Hispanics coincides with longer commutes. This finding is striking and further on the robustness check will give some explanation for this finding which is very relevant for the spatial mismatch hypothesis.

Coefficients for minorities

As noted, in 1990 – model 3 – the coefficient for Hispanics is positive and very significant. To interpret this; when the share of Hispanics in a census tract increases with 10%, the commuting duration increases with 34 seconds. The coefficient for Asians is positive and significant as well. Every 10% increase in the share of Asians coincides with an increase in commuting duration of 1 minutes and 17 seconds. In the year 2000 – again model 3 – the coefficient for Hispanics is also positive and significant. The coefficient in 2000 compared to 1990 increases slightly – from .0562 to .0574 – but stays quite constant. In 2010 – model 3 – we do not find a disparity in commuting duration for the distinguished minorities. In 2015 – model 3 – the coefficient for Asians is positive and significant with a value is .0675.

Interpretation

According to the maps presented in section 4.2, particularly African Americans and Hispanics that live in Phoenix tend to cluster together. Both minorities are over-represented in the area around the CBD. Furthermore, the regression analyses show that in 1990 and 2000, higher shares of Hispanics are correlated with longer commutes, which could be an indication of a spatial mismatch. Also the coefficients for Asians in 1990 and 2015 are positive and significant, but this minority tends to sprawl much more as opposed to Hispanics and African Americans. Perhaps the longer commutes for this group are associated with something else than a spatial mismatch to jobs – e.g. idiosyncratic preferences.

	1990			2000			2010			2015		
	model 1	model 2	model 3	model 1	model 2	model 3	model 1	model 2	model 3	model 1	model 2	model 3
VARIABLES	meancommute	meancommute	meancommute	meancommute	meancommute	meancommute	meancommute	meancommute	meancommute	meancommute	meancommute	meancommute
afro_american	-0.0258	-0.0311*	-0.0113	-0.101**	-0.0484	0.0245	-0.0233	-0.0307	-0.0323	-0.0187	-0.0659**	-0.0511*
	(0.0185)	(0.0173)	(0.0152)	(0.0414)	(0.0364)	(0.0326)	(0.0281)	(0.0267)	(0.0257)	(0.0299)	(0.0291)	(0.0282)
native	-12.14***	-0.0781***	-0.0416*	-0.297***	-0.161**	-0.000329	-0.144***	-0.135***	-0.0589	-0.205***	-0.227***	-0.136***
	(0.0306)	(0.0254)	(0.0224)	(0.0834)	(0.0705)	(0.0632)	(0.0518)	(0.0461)	(0.0444)	(0.0547)	(0.0495)	(0.0478)
asian	-0.160**	0.0069	0.1284**	-0.210**	0.153*	0.0979	-0.140***	-0.00635	0.00829	-0.0616*	0.0521	0.0772**
	(0.0739)	(0.0654)	(0.0576)	(0.0904)	(0.0825)	(0.0741)	(0.0397)	(0.0380)	(0.0360)	(0.0351)	(0.0329)	(0.0313)
hispanic	-0.0324**	-0.0414**	0.0562***	0.0860***	-0.0112	0.0574**	0.0209***	-0.0378***	-0.0119	0.0728***	-0.00650	0.0205
	(0.0153)	(0.0177)	(0.0170)	(0.0185)	(0.0255)	(0.0235)	(0.00736)	(0.0132)	(0.0134)	(0.00853)	(0.0140)	(0.0136)
CBD			3.29e-05			9.47e-05***			2.17e-05			2.48e-05
			(2.00e-05)			(2.34e-05)			(1.79e-05)			(1.71e-05)
NSUB			0.000173***			0.000239***			0.000178***			0.000176***
			(2.39e-05)			(2.59e-05)			(2.78e-05)			(2.71e-05)
controls:												
socio-economic	NO	YES	YES	NO	YES	YES	NO	YES	YES	NO	YES	YES
industry	NO	YES	YES	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	23.17***	12.04***	1.990	26.36***	26.49***	15.70**	24.37***	22.45***	16.87***	23.63***	21.56***	14.30***
	(0.207)	(3.602)	(3.269)	(0.378)	(8.425)	(7.455)	(0.292)	(4.638)	(4.426)	(0.307)	(5.296)	(5.091)
Observations	602	602	602	602	602	602	602	602	602	602	602	602
R-squared	0.052	0.434	0.572	0.066	0.475	0.596	0.056	0.358	0.430	0.149	0.411	0.479
Adj. R^2	0.046	0.409	0.552	0.059	0.453	0.577	0.050	0.330	0.404	0.143	0.386	0.456

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2

Robustness Check

The findings in the previous section generally show that Hispanics in 1990 and 2000 face longer commutes. Yet, the regression models are also very sensitive to the addition of *spatial* variables: i.e. without the addition of *spatial* variables (distance to CBD/nearest subcenter) the coefficients for Hispanics are negative and significant, while when these variables are included the coefficients become positive and significant. To find out why the coefficients flip around, the setup of the regression models will be altered.

The maps in figure 9 show that high concentrations of Hispanics are located in and around the CBD area, particularly in the years 1990 and 2000. Consequently, instead of using a continuous spatial distance measure – i.e. distance to CBD and nearest subcenter – a dichotomy will be made between *inner-city* and *suburbs* to see if the results hold, particularly for Hispanics.

This will be done by distinguishing census tracts by (1) CBD tracts and (2) non-CBD tracts, or 'suburb' tracts. The CBD area is demarcated by using the interstate highways as boundaries. This method is also used by Hu & Schneider (2015) in the case of Chicago. However, the north border will not be a highway border since the interstate 10 crosses through the CBD[6]. Instead of using this highway as border, the *tail* of the CBD area – defined in section 4.3 (see figure 11) – will be used. The tracts adjacent to the CBD will function as transition zones and are also considered CBD. Figure 12 shows a map of this dichotomy between CBD and non-CBD in Phoenix.

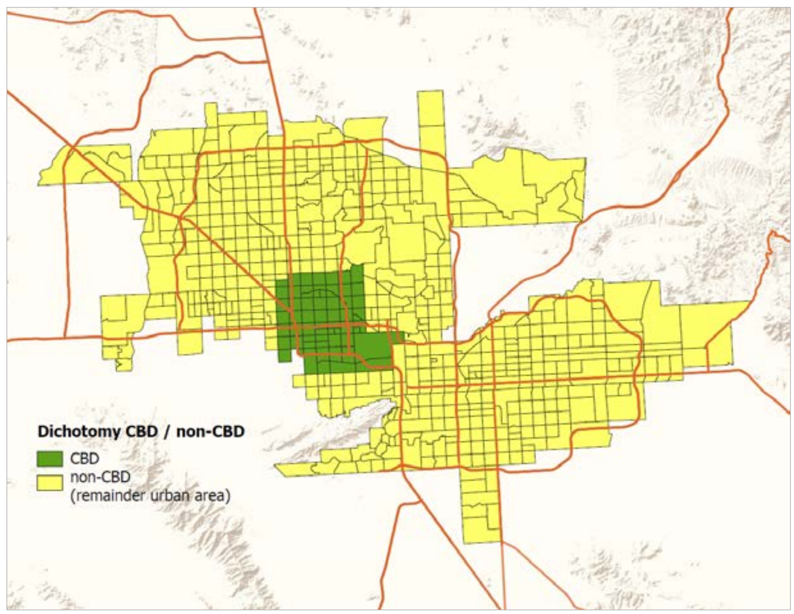


Figure 12

Note that the dichotomy between CBD and non-CBD is very crude. Orfield (1997) argues that also subcenters face issues with job accessibility for minorities. Yet, the maps about the distribution of Hispanics in Phoenix shows us that – especially in 1990 and 2000 – this group mainly clusters around the CBD area (see figure 9).

[6] As defined by figure 11 and Leslie (2006).

Robustness Check Regression Model

Table 3 shows the results of four regression models – i.e. for each time period one. Instead of using a continuous distance measure – *CBD* and *NSUB* – a dummy variable is incorporated:. *CBD_yn*. This means that a census tract can have 2 values: or it is part of the CBD (1), or not (0). Additionally, the variable *hispCBD_yn* is introduced, which is an interaction variable based on the share of Hispanics in a census tract (*hispanic*) and *CBD_yn*.

Particularly the results for the year 1990 are interesting. The coefficient *hispanic* is negative and significant, meaning that an increase in the share of Hispanics in a census tract coincides with a *shorter* commuting time (consistent with model 2 in table 2). The variable *CBD_yn* is negative and significant, meaning that tracts in the CBD area face, on average, shorter commutes. Nevertheless, the interaction variable *hispCBD_yn* is positive and strongly significant. This means that if the share of Hispanics increases in the CBD area, there is an additional positive effect on commuting duration.

The results suggest that an increase in the share of Hispanics outside the CBD coincides with *shorter* commutes. Yet, the coefficient of *hispCBD_yn* is greater[7] than *hispanic* (if you add them together, the slope becomes positive). Therefore, an increase in the share of Hispanics in the CBD area leads to *longer* commutes. This is evidence that Hispanics in 1990 – particularly those living in the inner-city – face a spatial mismatch.

This finding is can give be an explanation for the flip of the variable *hispanic* in table 2, i.e. there is an interaction between location of tracts in space and the ethnic composition. Apperently do census tracts with large shares of Hispanics in the inner-city of Phoenix face longer commutes, while outside the inner-city they face shorter ones. The SMH assumes that particularly minorities living in the central city face longer commutes, therefore these results are in line with the original spatial mismatch hypothesis.

[7] The coefficient for *hispanic* is -0,10, while for *hispCBD_yn* it is 0,16. If they are added together, the net outcome is (+)0,06. This implicates that the slope for *hispanic* in the CBD is positive while for *hispanic* outside the CBD it is negative.

In other years, the regressions results are more ambiguous. In 2000 the interaction variable is positive and significant as well. However, the variable *hispanic* is not significant and thus we cannot interpret the interaction term properly. Surprisingly in 2010, the results are similar as in 1990. However, the coefficient of *hispCBD_yn* is smaller than the coefficient of *hispanic* – i.e. if you add them together, the coefficient stays negative – and therefore the slope remains negative. Consequently, in 2010, the increase in the share of Hispanics coincides with shorter commutes, even in the CBD, yet the effect there is smaller. At last, in 2015 none of the coefficients of interest are significant.

These results are mostly in line with the previous regression analyses. Particularly the results for 1990 show that Hispanics residing in the CBD area face longer commutes.

	1990	**2000**	**2010**	**2015**
	model 1	model 2	model 3	model 4
VARIABLES	meancommute	meancommute	meancommute	meancommute
afro_american	-0.0207	-0.0408	-0.0281	-0.0679**
	(0.0168)	(0.0367)	(0.0266)	(0.0290)
native	-0.0605**	-0.128*	-0.129***	-0.225***
	(0.0243)	(0.0708)	(0.0460)	(0.0490)
asian	0.0270	0.138*	-0.0110	0.0479
	(0.0630)	(0.0824)	(0.0379)	(0.0326)
hispanic	**-0.1023*** **	**-0.0273**	**-0.0414*** **	**-0.00400**
	(0.0207)	**(0.0283)**	**(0.0134)**	**(0.0142)**
CBD_yn	**-0.03932*** **	**-3.268*** **	**-2.021** **	**-0.719**
	(0.005)	**(1.036)**	**(0.856)**	**(0.800)**
hispCBD_yn	**0.1607*** **	**0.0719* **	**0.0350** **	**-0.0295**
	(0.026)	**(0.0368)**	**(0.0173)**	**(0.0210)**
controls:				
socio-economic	YES	YES	YES	YES
industry	YES	YES	YES	YES
Constant	12.20***	29.92***	22.39***	26.19***
	(3.442)	(8.426)	(4.642)	(5.380)
Observations	602	602	602	602
R-squared	0.485	0.485	0.364	0.426
Adj. R^2	0.461	0.462	0.335	0.400

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3

5. CONCLUSION

To restate the research question: “to what extent is a spatial mismatch present for minorities in Phoenix, Arizona? Kain observed that (especially) African American minorities face a spatial mismatch to employment. This is because (1) employment is moving towards the suburbs, (2) racial discrimination in the housing market prevented African Americans to move to areas where job growth exists, and (3) poor information about distant job openings and inadequate public transport linkages between minority-dominated neighborhoods and jobs-growing areas.

To start off, minorities living in Phoenix do tend to cluster. Particularly Hispanics and African Americans are spatially clustered. Yet, African Americans sprawl over time. *Second*, new subcenters of employment emerge over time. However, this is not at the expense of employment in the CBD: employment in and around the CBD grows as well. *Third*, the regression analyses show that in 1990 and 2000, workers in census tracts – particularly in the CBD – dominated by Hispanics travel significantly longer than workers in census tracts with low shares of Hispanics. In 2010 and 2015 we do not find a significant difference in commuting duration between minorities. These findings support the existence of a SMH in the past (for Hispanic workers), yet the findings from recent years do not indicate the existence of a spatial mismatch for any minority.

6. DISCUSSION

Kain (1968) and Kasarda (1989; 1995) found that particularly African Americans face a spatial mismatch in U.S. cities. This study does not find supportive evidence for a spatial mismatch for African Americans, but instead finds evidence for spatial mismatch for Hispanics during and before the turn of the century. Yet, Kain and Kasarda both focused on old-industrial cities which have a different population and industry structure as opposed to Phoenix and Sunbelt cities in general.

Hu (2015) examines Los Angeles – a city more similar to Phoenix – but does not observe a spatial mismatch for any minority group, not in the past and not in the present. Hu argues that this is because enough employment remains in and around the CBD, the area minorities often are clustered. Looking at Phoenix, the supply of employment in the CBD also grows over time. Perhaps this job growth improved the accessibility to employment for Hispanics in more recent years.

All things considered, in recent years there are no signs for a spatial mismatch for minorities that reside in Phoenix. Accordingly, policymakers should focus on other factors which have to do with the worse labor market outcomes for minorities – like differences in educational attainment, work experience, and discrimination (which all are not necessarily spatial). Currently, Hispanics and other minorities still have worse labor market outcomes (i.e. higher shares of unemployment) as opposed to non-minorities. Since Phoenix as a city is very similar to other cities in the Sunbelt, this policy recommendation can be asserted to other cities as well, to a certain extent.

It is important to mention that this study did not discuss the *why* question – as in why did Hispanics faced a spatial mismatch in the past but not anymore. Still, I will discuss some possible explanations. Liu (2014) performed a study about employment accessibility in Phoenix and argues that the light rail – a public transportation connection, opened in 2008 – improved the accessibility of especially Hispanic dominated neighborhoods. While in 2000 Hispanics face longer commutes, in 2010 and 2015 they do not anymore. Therefore, the introduction of the light-rail could have had a positive impact on the accessibility of Hispanics in Phoenix. Kimball et al. (2013) also argue that the introduction of the light rail in Phoenix had a significant impact on the development of employment in the CBD. And as said before, this growth of jobs in the CBD area again increases accessibility to jobs for minorities living close to the CBD. Another explanation could again be related to the internal markets. Between 1990 and 2015 the share of Hispanics living in Phoenix have grown drastically. Perhaps it takes time for internal markets to develop; i.e. small businesses need time to grow, hire new personnel, and the new influx of Hispanics fosters this process. Small businesses have a larger customer base, sell more, grow, and again need more personnel. However, the preceding arguments are purely based on correlations, I cannot argue that one or both of them have any to do with better job accessibility in 2010 and 2015, relative to 2000. Furthermore, it would be very interesting to investigate how internal markets actually work. How goes the process of growth, hiring new workers, and so forth.

There are several shortcomings in this study. *First*, the regression model is a simple OLS regression. Therefore, I cannot rule out endogeneity is present – i.e. because important omitted variables were not included, or due to reverse causality – and thus the coefficients of the minorities can be biased to a certain extent. Future studies could come up with alternative methods, like 2SLS, that deal with endogeneity issues. Unfortunately, I was not able to find a proper IV in the given time for this study. Also, one of the conditions of the regression analysis – i.e. while one predictor increases, the other predictors stay constant – does not fully hold, since the shares of minorities are always relative to each other. Again, this leads to some bias in the coefficients. *Second*, variation *within* census tracts is ignored. The unit of observation in this study are census tracts and they are neighborhoods with approximately 4000 inhabitants. The dependent variable is the mean commuting time within a tract. Therefore, even if the findings – particularly in 2010 and 2015 – do not support the spatial mismatch for minorities, it cannot be ruled out that there is a spatial mismatch for certain minorities because within census tract variation in commuting is ignored. *Third*, the data for Phoenix is not fully reliable. The urban area of Phoenix in the year 2000 is chosen as baseline, while the city have grown drastically over time. Areas urban in 2000 were rural in 1990 while areas rural in 2000 are urban in 2010. Perhaps new urban areas in 2010 face difficulties with job accessibility, but this is ignored.

There are more than a few fruitful areas for future research. I found evidence that Hispanics living in Phoenix faced a spatial mismatch to jobs in the past, but not anymore. However, I did not present the story behind this development: *why* did they face a spatial mismatch in the past, but not anymore? A future study can examine the causes for the positive development of job accessibility of Hispanics residing in Phoenix. Furthermore, findings of recent years cannot support the spatial mismatch anymore, but this does not have to continue that way. The maps with the distribution of minorities and employment centers in Phoenix showed us something interesting. While the area where minorities cluster tends to be very static – especially for Hispanics – the development of urban centers is much more dynamic. In ten years it could be that the relocation of employment centers leads to a situation where certain minorities again face a spatial mismatch to jobs. A good development, on the other hand, is the growth of employment in the CBD in Phoenix. Many minorities live close to that area and they are able to benefit from this[8].

[8] Certainly, the jobs that come into existence in the CBD area do not necessarily have to be the right jobs. But even the growth of jobs in sectors like finance and insurrance – often overrepresented in the CBD area of U.S. cities (de Pater & Verkoren, 2007) – leads to the growth in low-skilled jobs as well (Moretti, 2012).

SOURCES

Alonso, W. (1964). Location and land use.

Anas, A., Arnott, R., Small, K.A., 1998. Urban spatial structure. *J. Econ. Lit.* 36 (3), 1426–1464.

Audretsch, D. B., & Keilbach, M. (2010). Entrepreneurship and growth. In *Knowledge Intensive Entrepreneurship and Innovation Systems* (pp. 309-320). Routledge.

Battu, H., Seaman, P., & Zenou, Y. (2011). Job contact networks and the ethnic minorities. *Labour Economics*, *18*(1), 48-56.

Blumenberg, E. (2004). En-gendering effective planning: spatial mismatch, low-income women, and transportation policy. *Journal of the American Planning Association*, *70*(3), 269-281.

Boeri, T., & Van Ours, J. (2013). *The economics of imperfect labor markets*. Princeton University Press.

Brueckner, J. K. (2011). *Lectures on urban economics*. MIT Press.

Bureau of Labor Statistics (2018). Retrieved from https://www.bls.gov/

Burgess, E. W. (1928). Residential segregation in American cities. *The Annals of the American Academy of Political and Social Science, 140,* 105–115.

Cahuc, P., Carcillo, S., & Zylberberg, A. (2014). *Labor economics*. MIT press.

Calvó-Armengol, A. (2004). Job contact networks. *Journal of economic Theory*, *115*(1), 191-206.

Center for International Earth Science Information Network (2000) Gridded population of the world: http://sedac.ciesin.columbia.edu/plue/gpw via the Internet. Accessed June 29, 2019.

Census Bureau (2018). Decennial Census Data. Retreived from: https://www.census.gov/programs-surveys/decennial-census/data.html

Clark, K. B. (1965). *Dark ghetto: Dilemmas of social power*. New York: Harper and Row.

Clark, W. A., Huang, Y., & Withers, S. (2003). Does commuting distance matter?: Commuting tolerance and residential change. *Regional Science and Urban Economics*, *33*(2), 199-221.

Coulson, N. E., Laing, D., & Wang, P. (2001). Spatial mismatch in search equilibrium. *Journal of Labor Economics*, *19*(4), 949-972.

Crane, R. (2007). Is there a quiet revolution in women's travel? Revisiting the gender gap in commuting. *Journal of the American planning association*, *73*(3), 298-316.

de Pater, B.C. & Verkoren, O. (2007). Noord-Amerika. Een geografie van de Verenigde Staten en Canada.

Epstein, G. S. (2002). Informational cascades and decision to migrate.

Fischer, M. J., & Massey, D. S. (2000). Residential segregation and ethnic enterprise in US metropolitan areas. *Social Problems*, *47*(3), 408-424.

Giuliano, G. (1998). Information technology, work patterns and intra-metropolitan location: A case study. *Urban Studies*, *35*(7), 1077-1095.

Giuliano, G., Redfearn, C., Agarwal, A., Li, C., & Zhuang, D. (2007). Employment concentrations in Los Angeles, 1980–2000. *Environment and planning A*, *39*(12), 2935-2957.

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of urban Economics*, *63*(1), 1-24.

Gobillon, L., & Selod, H. (2014). Spatial mismatch, poverty, and vulnerable populations. In *Handbook of regional science* (pp. 93-107). Springer Berlin Heidelberg.

Gobillon, L., Selod, H., & Zenou, Y. (2007). The mechanisms of spatial mismatch. *Urban studies*, *44*(12), 2401-2427.

Gordon, P., Kumar, A., & Richardson, H. W. (1989). The influence of metropolitan spatial structure on commuting time. *Journal of urban economics*, *26*(2), 138-151.

Gordon, P., Richardson, H., & Jun, M.-J. (1991). The commuting paradox: Evidence from the top twenty. *Journal of the American Planning Association*, 57, 416–420.

Gregory, D., Johnston, R., Pratt, G., Watts, M., & Whatmore, S. (Eds.). (2011). *The dictionary of human geography*. John Wiley & Sons.

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. (1995). Multivariate Data Analysis (3rd ed). New York: Macmillan.

Helbich, M., & Leitner, M. (2010). Postsuburban spatial evolution of Vienna's urban fringe: evidence from point process modeling. *Urban Geography*, *31*(8), 1100-1117.

Hu, L. (2015). Job accessibility of the poor in Los Angeles: Has suburbanization affected spatial mismatch?. *Journal of the American Planning Association*, *81*(1), 30-45.

Hu, Y., & Wang, F. (2016). Temporal trends of intraurban commuting in Baton Rouge, 1990–2010. *Annals of the American Association of Geographers*, *106*(2), 470-479.

Humes, K. R., Jones, N.A, Ramirez, R.R. (2011). *Overview of race and Hispanic origin: 2010*. Census 2010 Brief # C2010BR-02. Washington: U.S. Government Printing Office.

Iceland, J., & Sharp, G. (2013). White residential segregation in US metropolitan areas: Conceptual issues, patterns, and trends from the US census, 1980 to 2010. *Population research and policy review*, *32*(5), 663-686.

Ihlanfeldt, K. R., & Sjoquist, D. L. (1998). The spatial mismatch hypothesis: a review of recent studies and their implications for welfare reform. *Housing policy debate*, *9*(4), 849-892.

Kain, J. F. (1968). Housing segregation, Negro employment, and metropolitan decentralization. *The Quarterly Journal of Economics*, 82, 175–197.

Kasarda, J. (1989). Urban industrial transition and the under- class. *The Annals of the American Academy of Political and Social Science*, 501, 26–47.

Kasarda, J. (1995). Industrial restructuring and the changing location of jobs. *State of the Union: America in the 1990s*, *1*, 215-267.

Kawabata, M., & Shen, Q. (2007). Commuting inequality between cars and public transit: The case of the San Francisco Bay Area, 1990-2000. *Urban Studies*, *44*(9), 1759-1780.

Kneebone, E. (2009) Job sprawl revisited: the changing geography of metropolitan employment. *Metropolitan Policy Program,* The Brookings Institution, Washington, DC.

Lee, B. S., & McDonald, J. F. (2003). Determinants of commuting time and distance for Seoul residents: The impact of family status on the commuting of women. *Urban Studies*, *40*(7), 1283-1302.

Leslie, T. F., & Huallacháin, B. Ó. (2006). Polycentric phoenix. *Economic Geography*, *82*(2), 167-192.

Liu, C. Y., & Painter, G. (2012). Immigrant settlement and employment suburbanisation in the US: Is there a spatial mismatch?. *Urban Studies*, *49*(5), 979-1002.

Liu, L. (2014). *Impacts of light rail in job accessibility in Phoenix*. Arizona State University.

Maricopa Association of Governments. (2018) Maricopa Land-use Explorer: https://ims.azmag.gov/ via the Internet. Accessed June 29, 2019.

McCann, P. (2013). *Modern urban and regional economics*. Oxford University Press.

McQuaid, R. W., & Chen, T. (2012). Commuting times–The role of gender, children and part-time work. *Research in transportation economics*, *34*(1), 66-73.

Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.

Myrdal, G. (1944). *An American Dilemma: The Negro problem and modern democracy*. Reprint, New Brunswick: Transaction Publishers, 1996.

Orfield, Myron. 1997. Metro Politics: A Regional Agenda for Community and Stability. Washington, DC: Brookings Institution.

O'sullivan, A. (2007). *Urban economics*. McGraw-Hill/Irwin.

Renne, J. L. (2009). From transit-adjacent to transit-oriented development. *Local Environment*, *14*(1), 1-15.

Robertson, K. A. (1995). Downtown redevelopment strategies in the United States: An end-of-the-century assessment. *Journal of the American Planning Association*, 61(4), 429-437.

Ross, A. (2011). *Bird on fire: Lessons from the world's least sustainable city*. OUP USA.

Schularick, M., & Taylor, A. M. (2012). Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review*, *102*(2), 1029-61.

Shen, Q. (2000). Spatial and social dimensions of commuting. *Journal of the American Planning Association*, *66*(1), 68-82.

Taylor, B., & Ong, P. (1995). Spatial mismatch or automobile mismatch? An examination of race, residence and com- muting in U.S. metropolitan areas. *Urban Studies*, 32, 1453–1473.

The Economist. (2018) Retrieved from https://www.economist.com/international/2018/06/21/public-transport-is-in-decline-in-many-wealthy-cities

Turcotte, M. (2011). Commuting to work: Results of the 2010 general social survey. *Canadian Social Trends*, *92*(August), 25-36.

Turner, M. A. (2008). Limits on Housing and Neighborhoods Choice: Discrimination and Segregation in US Housing Markets. *Ind. L. Rev.*, *41*, 797.

U.S. Census Bureau (1990; 2000; 2010; 2015; 2016; 2017; 2018). Retrieved from https://www.census.gov/

Von Thunen, J.H. (1826). The Isolated State, Translated by Carla M. Pergamon Press, Wartenburg. New York.

Wang, F. (2000). Modeling commuting patterns in Chicago in a GIS environment: A job accessibility perspective. *The Professional Geographer, 52*(1), 120-133.

Wang, F. (2003). Job proximity and accessibility for workers of various wage groups. *Urban Geography*, *24*(3), 253-271.

Whitehouse. (2018) Retrieved from https://www.whitehouse.gov/

Woods, M. (2010). *Rural*. Routledge.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.

APPENDIX

A.1 Making the Geography of Census Tracts consistent over time

To make data of census tracts consistent over time, GIS software is needed (e.g. ArcGIS or QGIS). Over time, the number of census tracts in Phoenix increases considerably. This has consequences for the census tracts. For example, a census tract in 2000 can be made up out of two census tracts in 1990 (see figure 13) . Data estimates from 1990 therefore need to be split into two tracts in 2000. For 2010/2015 to 2000 it is the other way around. A census tract from 2000 can be split up to two or more census tracts in 2010 – figure 14 shows this. What happens is that an average value is calculated to make the data of 2010/2015 consistent with the year 2000. This can be done with GIS software by (1) join attributes by location, and (2) calculate averages. I prefer to calculate the average with Excel and create pivot tables. But this can be done in many ways.

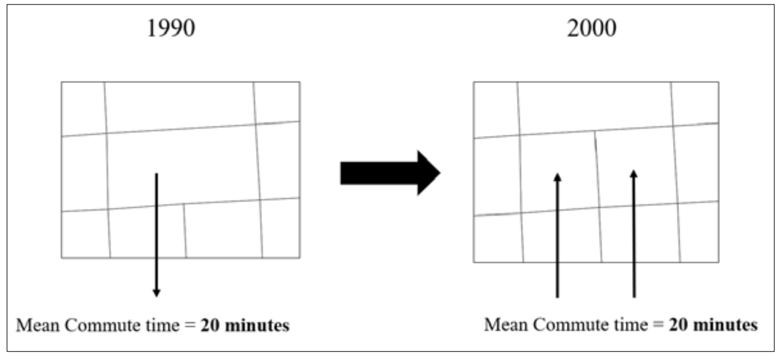


Figure 13

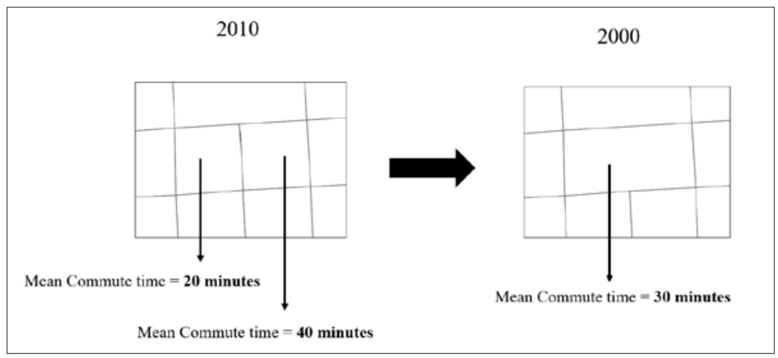


Figure 14

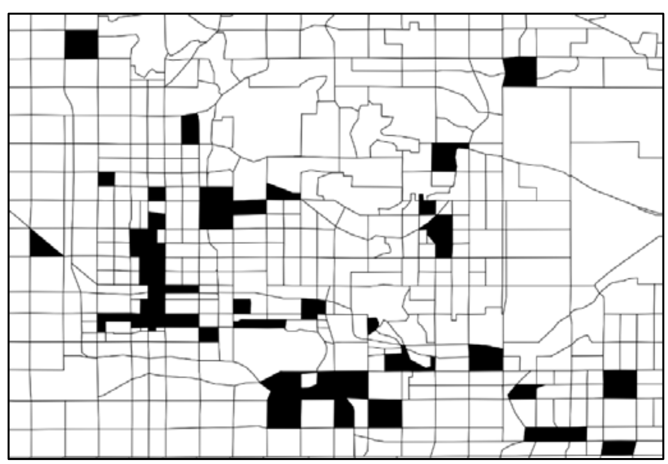
A.2 Determining Urban Centers

In this section, the method of Giuliano (2007) to determine urban centers will be explained. Giuliano defines urban centers as a group of contiguous TAZ's (Transportation Analysis Zone[9]) which have at least 10 employees per acre and in total consists out of 10.000 employees. In total means the contiguous tracts together must at least contain 10.000 jobs. To be adjacent or contiguous, tracts have to share a common border and even touching a corner is enough to be considered adjacent. Giuliano used this method in Los Angeles, a similar city compared to Phoenix, so the same cut-off values will be used to determine centers. For executing this for Phoenix, QGIS software is used again. The next steps will explain step by step the process to determine urban centers and to attain the variables (distance towards the CBD/nearest subcenter) used in the regression.

1. Obtain spatial data about the tracts in all the time periods (2000, 2010 & 2015). This data can be found by accessing the U.S. Census website and access the TIGER[10] data (Census Bureau, 2017). The TIGER products contain shapefiles for the TAZ's in Arizona. For the years 1990 and 2000 data belonging to the Arizona State University will be used. By selecting the two counties – Maricopa and Pinal – the right TAZ's are selected. Shapefile format is very common file format which is compatible with most GIS software packages.

2. The second step is to get the data about the number of jobs per TAZ. For the year 1990 and 2000, Census Transportation Planning Package (abbreviation: CTPP) data will be used (Bureau of Transportation Statistics, 2000). This data will be linked to the spatial TAZ data. This can be done by using the 'join' tool in ArcGIS or QGIS. For the years 2010 and 2015 another data source is used. OnTheMap.com provides data for job points and they can be downloaded easily as a shapefile. After downloading the shapefile, it can be imported into QGIS. In QGIS a tool is available to count how many points are present within a certain geographical unit. So using both files – TAZ polygons & job points – the number of jobs per tract can be calculated with the command 'Count Points in polygon'. Nonetheless, job points can consist out of more jobs – e.g. a job point can contain 100 jobs – the points have to be weighed by the number of jobs per point.

3. Now the number of jobs per TAZ is clear for all time periods, the next step is to calculate the density of jobs. In QGIS there is an option to calculate the acreage of a TAZ. By doing this, for every polygon the acreage will be calculated. By dividing the number of jobs by the area acreage, the jobs per acre can be calculated. Now the number of jobs and the job density per TAZ is calculated.



4. Now a map can be drawn like in the figure 15. It shows TAZs for the year 2000. The dark color means that the density in that tract is higher than 10 jobs per acre. The next step is to see which tracts are contiguous and determine if the total employment of contiguous tracts exceeds 10.000 jobs.

5. After step four, a number of urban clusters is determined. The next step is to determine the center of gravity. To do this, the first step is to calculate the centroid of all the units – this can be done by the tool 'polygon centroid' in QGIS. This will determine the

[9] TAZ's are geographical units like census tracts, both are aggregates of census blocks. However, the boundaries of TAZ's are defined by functional characteristics. It also does not have a fixed population like census tracts have.
[10] TIGER stands for **T**opologically **I**ntegrated **G**eographic **E**ncoding and **R**eferencing

centroid of a unit based on its shape.

6. The centroid points still contain the number of jobs that specific tract has. The next step is to calculate a spatial weighted average of the contiguous units, weighted by number of jobs. This can be done by the tool 'coordinate average' and using the option weighted by: employment/ number of jobs. This is not necessary per se, but it makes the 'center' of the urban center more precise. All in all the difference between a weighted and a non-weighted center is not big. *Figure 15*

7. By now, spatial points are constructed which can be seen as urban centers. Additionally, the clusters of TAZ areas (which exceed the 10.000 jobs threshold) can be seen as urban centers.

A.3 Calculating the Distance from Census Tracts to Urban Centers

For explaining variation in commuting time, the distance towards urban centers is used. To determine the distance towards urban centers, some additional actions have to be performed. However, most of the work is already done. The following steps can be used to calculate the distances.

1. Calculate the centroid of the census tract polygons (using ‘polygon centroid’).
2. Add the layer with the centroids of the urban centers (calculated following the steps of ‘determining urban centers’) to the layer with the centroids of census tracts.
3. Distinct the CBD from the other centers (sub centers) by putting it in another layer.
4. Perform a distance matrix analysis for both the distance towards the CBD and the distance towards the nearest subcenter. The nearest subcenter can be calculated by using ‘nearest point’.
5. Two outputs are generated (e.g. in .xls format). One is the distance from tracts towards the CBD and one from the tracts towards the nearest sub center. The measure unit chosen is meter. As noted in the methodology, the distance is straight line from point to point.

The method of Giuliano is very useful to determine urban centers as well as to determine distances towards urban centers. However, there are stern imperfections by using this method that have to be elaborated:

- Data about employment is aggregated on TAZ level. So the way TAZs are constructed affects the total number of jobs and density of jobs within a certain tract. To circumvent this problem, data about exact job locations would be more precise, but this data is not used because of reasons mentioned earlier.
- The approach uses a kind of cut-off approach. For tracts to be ‘contiguous’ they only need to touch each other by some sort of common border. The approach of Giuliano is used in different settings by different researchers and their interpretation of what can be seen as ‘contiguous’, differs. McMillan (2003) uses a similar approach but he sees tracts as contiguous if they are within a reach of 1.5 mile. Their reasoning is that without this ‘looser’ interpretation of contiguous, too few urban centers would occur by using the data. The fact that these decisions – to choose between a 1.5 mile radius or a zero mile radius – have implications for the number of urban centers in Phoenix as well. Luckily, in the case of Phoenix, the difference is not significant. The motive to use the zero mile radius is because without this radius of 1.5 mile, a big part of the central Phoenix area would be considered as one big urban area instead of several distinct urban centers. According to Leslie (2006) it is surely not the case that in the area around the CBD can be considered as one urban area.
- The method used to determine the CBD and subcenters is questionable. A center is put together in just one point, while a urban center often covers a greater region. An assumption made in the regression is that the distance towards that exact point would determine the commuting time, while the job is not necessarily located there.
- The center of gravity of the census tracts – so not the TAZ’s – is used to calculate the distance to the urban centers. Again, not all workers live on that exact location, making this variable less accurate. Especially tracts at the outskirts of Phoenix are very big. For these tracts the center of gravity is far from accurate and quite unrealistic.
- Using the variable ‘nearest subcenter’ is questionable. Workers do not necessarily work at the closest subcenter. It could be the case that a worker lives very close to an urban center, but works at the other side of the city. Yet, because the cases are tracts – so a group of workers – it is likely that being close to a urban center will have an negative effect on commuting time of the group in total.

All in all, these are some downsides for using this variable. However, the distance towards urban centers it is a control variable. Consequently the goal is that this variable controls for the spatial pattern of the city and is of additional value to explain variation in commuting time. These inconsistencies are therefore not seen a too problematic.

A.4 Variables, Correlation Table, and Summary Statistics

Now several important determinants of commuting time will be discussed. For these factors will be controlled in the regression analyses.

First, mode of transportation. Using public transport increases the commuting duration – a measure of labor market accessibility – of workers considerably (Gordon et al. 1989; Kawabata & Shen, 2007), because the majority of time spent commuting is not taken up by riding, but by waiting. Workers who walk to their work location often tend to have the shortest commutes since their distance towards their work location is so short, that they are able to walk instead of using a other modes.

Second, income. The role of income on labor market accessibility tends to be ambiguous. Giuliano (1998) argues that higher income workers tend to travel longer while Blumenberg (2004) and Hu & Schneider (2015) argue that low income workers do as well. The relationship if more or less parabolic. Still, there is a big difference between both groups. High-income workers choose to live farther away from their jobs, while low-income workers often do not have a choice.

Third, level of education. Lee & McDonald (2003) argue that education is an important factor in explaining variation in travel time. They argue that workers with a higher education travel longer than workers with a lower education. Shen (2000) argues that particularly workers with a bachelor degree or higher, travel longer compared to workers with lower education. Jobs associated with higher degrees are often more specific, and to find the right job 'fit', workers often have to search in a larger area.

Fourth, female workers. Female workers in general tend to have lower commute times (Rouwendal, 1999; Clark et al., 2003) The American Commuting Survey (2009) shows that the average time difference between male workers and female workers in the U.S. is three minutes. Factors like child-care are often the responsibility of female workers, and it affects their commuting patterns of females.

Fifth, job classification. Executives and managers tend to long commutes compared to other classifications (Shen, 2000). Also, Gottlieb & Joseph (2006) argue that graduates with different educational attainments have different residential preferences. Employees working in higher education institutions put more emphasis on the characteristics of amenities. This could implicate that they are more choosy on their residential location, which in turn can lead to longer commutes.

- On the next page a summary of the variables can be found -

SUMMARY VARIABLES

	Variables	Meaning
Dependent Variable		
Job Accessibility	*meancommute*	The average commuting time within a cencus tract
Independent Variables		
minorities	*afro_american*	Share (%) of African Americans
	native	Share (%) of Native Americans
	asian	Share (%) of Asians
	hispanic	Share (%) of Hispanics
socio-economic	*car*	Share residents which use the car as mode of transport to work
	carpool	Share residents which use carpooling as mode of transport to work
	pt	Share residents which use public transport as mode of transport to work
	vehicle0	Share of residents without a vehicle
	medhincom	Medium household income
	edumid	Share residents with college degree, or associate degree
	eduhigh	Share residents with bachelor degree, or graduate degree
	femalework	Share of the workers per tract which is female
	INDagri	Share workers in Agriculture, forestry, fishing and hunting, and mining
industry	*INDconst*	Share workers in Construction
	INDmanu	Share workers in Manufacturing (durable and non-durable goods)
	INDwholesale	Share workers in Wholesale trade
	INDretail	Share workers in Retail trade
	INDeduc	Share workers in educational services
	INDtrans	Share workers in Transportation and warehousing, and utilities
	INDcommu	Share workers in Communication Services
	INDfinan	Share workers in Finance, insurance, real estate, and rental and leasing
	INDbusi	Share workers in Professional, scientific, management, administrative, and waste management services
	INDpers	Share workers in Arts, entertainment, recreation, accommodation and food services
	INDent	Share workers in Other services (except public administration)
	INDpubl	Share workers in Public administration
spatial	*CBD*	Euclidean distance from census tract to CBD
	NSUB	Euclidean distance from census tract to nearest subcenter
robustness check	*CBD_yn*	Dummy variable, or a census tract is located in the CBD area, or not
	hispCBD_yn	Interaction variable, based on *CBD_yn* and *hispanic*

SUMMARY STATISTICS

1990

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
meancommute	606	22.39955	2.845373	12.5	30.1
black	606	3.48935	6.56226	0	66.8
native	606	1.48724	3.76209	0	75.4
asian	606	1.62539	1.54128	0	16.5
hispwhite	606	7.47734	8.10038	0	59.3
car	606	73.79546	10.56335	15.84699	95.9
carpool	606	14.48142	6.5311	0	50.9
pt	606	2.04797	3.09329	0	33.0
vehicle0	606	6.98495	9.11168	0	100
medhincom	606	3460952	1512103	516200	13748900
edumid	606	33.81188	7.7064	0	48.7
eduhigh	606	22.29059	13.93646	0	100
femalework	606	0.444774	0.045209	0.10929	0.62
INDagri	602	3.30175	5.01475	0	34.56
INDconst	602	6.43093	3.54698	0	50
INDmanu	602	14.83303	6.14574	0	42.64
INDwholesale	602	4.1563	2.01513	0	11.50
INDretail	602	17.40266	4.93043	0	50
INDeduc	602	14.73074	4.82326	0	34.85
INDtrans	602	4.55549	2.75411	0	31.43
INDcommu	602	2.91881	1.8868	0	13.29
INDfinan	602	8.15531	4.37063	0	28.49
INDbusi	602	6.00497	2.62719	0	31.25
INDpers	602	4.31873	2.75164	0	27.32
INDpubl	602	4.62664	3.56921	0	45.96
CBD	606	19255.59	9797.559	371.5151	53405.32
NSUB	606	13097.73	7885.177	116.0753	39547.08

2000

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
meancommute	602	26.00947	4.6006	13.4	85.6
black	603	3.860199	4.691061	0	56.3
native	603	1.599502	2.281665	0	44.2
asian	603	2.209121	2.075206	0	31.3
hispwhite	603	11.40032	10.97798	0	61.3314
car	602	73.47591	11.86499	15.6	100
carpool	602	15.9608	8.570529	0	58.7
pt	602	2.385548	3.248981	0	35
vehicle0	602	7.371595	7.833337	0	58.6
medhincom	603	48535.36	20843.09	0	174840
edumid	602	33.01645	8.653845	4.6	53.4
eduhigh	602	25.00532	15.91482	0	70.3
femalework	602	0.44239	0.052917	0.151351	0.625749
INDagri	602	0.495017	0.836148	0	7
INDconst	602	9.257475	5.513673	0	65.1
INDmanu	602	11.70498	4.888171	0	34
INDwholesale	602	3.693023	1.66152	0	12.8
INDretail	602	12.04834	3.463233	0	34.9
INDtrans	602	4.937043	2.08445	0	12.5
INDinfo	602	3.070598	1.624055	0	14.2
INDfinance	602	9.207807	3.619625	0	23.8
INDprof	602	11.5098	4.005593	0	31.8
INDeduc	602	15.9103	4.587844	0	38.6
INDarts	602	9.113787	4.017225	0	32.7
INDpubl	602	4.363621	2.337616	0	21.2
CBD	603	19010.41	9407.732	395.1653	51045.08
NSUB	603	11285.89	7297.727	26.73335	39210.28

2010

Variable	Obs	Mean	Std. Dev.	Min	Max
meancommute	602	24.01632	3.415628	14.3	43.9
black	603	4.95257	5.113297	0	38.5
native	603	1.828883	2.693474	0	38
asian	603	3.293444	3.655805	0	27.3
hispwhite	603	23.31232	20.4707	0	81.25198
car	602	74.91054	8.565838	25.3	89.5
carpool	602	12.08602	6.299506	0	36.4
pt	602	3.037984	3.602374	0	30.7
vehicle0	602	7.984815	8.276294	0	52.1
medhincom	602	55566.29	25700.57	9579	179306
edumid	603	32.12506	8.570728	6.3	54.1
eduhigh	603	27.61626	17.05002	0.7	76.5
femalework	603	0.454257	0.060916	0	0.649485
INDagri	602	0.547968	1.047088	0	7.75
INDconst	602	8.001556	5.515117	0	46.7
INDmanu	602	8.166094	4.376891	0	36.3
INDwholesale	602	2.721935	1.970967	0	15.8
INDretail	602	12.42023	4.745384	1.4	38
INDtrans	602	4.765028	2.616024	0	16.7
INDinfo	602	1.960357	1.838028	0	23.2
INDfinance	602	9.054169	4.460312	0	22.7
INDprof	602	12.81418	4.93988	0	33.7
INDeduc	602	20.0257	5.806738	3.2	39.9
INDarts	602	10.3394	4.827528	0	29.7
INDpubl	602	3.989322	2.507685	0	25.8
CBD	603	22841.79	11228.6	254.0054	61040.31
NSUB	603	8278.956	6120.35	254.0054	36544.6

2015

Variable	Obs	Mean	Std. Dev.	Min	Max
meancommute	603	24.53312	3.608537	14.3	41.5
black	604	4.985184	4.879075	0	38.7
native	604	1.874393	2.578462	0	39.4
asian	604	3.733011	4.070175	0	31.4
hispwhite	604	22.04015	17.3292	0.314582	82.26545
car	603	75.28524	8.544988	36.4	89.9
carpool	603	11.06984	6.049593	0	35
pt	603	2.778095	3.218161	0	20.6
vehicle0	603	7.452055	7.528776	0	45.5
medhincom	603	59547.84	27784.34	14489	194729
edumid	604	31.58965	7.496472	10.2	49.9
eduhigh	604	29.82332	17.7364	0.7	82.9
femalework	603	0.456784	0.05201	0.208873	0.623639
INDagri	603	0.621606	2.247314	0	50
INDconst	603	7.450388	4.740799	0	27.3
INDmanu	603	7.332563	3.575168	1	31.3
INDwholesale	603	2.645936	1.733198	0	11
INDretail	603	12.20942	3.914762	1.8	25.7
INDtrans	603	5.159065	2.670756	0	18.7
INDinfo	603	1.882892	1.414373	0	9.3
INDfinance	603	9.39433	4.15756	0	24.85
INDprof	603	13.63034	4.694228	0	30.3
INDeduc	603	20.36951	5.669964	5.3	38.3
INDarts	603	10.67018	4.370654	1.9	29.8
INDpubl	603	3.384496	1.908779	0	10.7
CBD	604	22670	11190.62	761.2083	60811.26
NSUB	604	7970.973	6096.113	302.44	36327.36

CORRELATION TABLE

1990	meancommute	afro_american	native	asian	hispanic	car	carpool	pt	vehicle0	medhincc	edumid	eduhigh	femalework	INDagri	INDconst	INDmanu	INDwholes	INDretail	INDeduc	INDtrans	INDcommu	INDfinan	INDpers	INDent	INDpubadı	CBD	NSUB
meancommute	1																										
black	-0.0971	1																									
native	-0.178	0.0654	1																								
asian	-0.0624	-0.0767	-0.0386	1																							
hispwhite	-0.135	0.366	0.189	-0.122	1																						
car	0.265	-0.412	-0.298	0.0986	-0.673	1																					
carpool	0.0710	0.367	0.204	-0.149	0.609	-0.677	1																				
pt	-0.205	0.457	0.241	-0.0482	0.518	-0.700	0.292	1																			
vehicle0	-0.299	0.449	0.282	-0.126	0.532	-0.817	0.377	0.727	1																		
medhincom	0.231	-0.310	-0.259	0.160	-0.482	0.645	-0.499	-0.452	-0.608	1																	
edumid	0.247	-0.273	-0.221	0.236	-0.569	0.612	-0.518	-0.426	-0.520	0.371	1																
eduhigh	-0.117	-0.316	-0.204	0.296	-0.517	0.524	-0.579	-0.347	-0.431	0.760	0.280	1															
femalework	0.0528	-0.0351	-0.0770	-0.00804	-0.400	0.427	-0.134	-0.341	-0.391	-0.00155	0.337	0.0776	1														
INDagri	0.0671	-0.0183	0.0249	-0.0450	0.0458	-0.0139	0.0852	-0.0512	-0.0147	-0.0257	-0.0356	-0.0507	-0.0416	1													
INDconst	-0.0631	-0.00890	0.0531	0.0559	0.0879	-0.0925	0.0850	0.0464	0.0461	-0.0768	-0.0103	-0.0900	-0.0548	0.124	1												
INDmanu	-0.206	0.0727	0.0384	-0.0661	0.186	-0.159	0.0628	0.148	0.133	-0.115	-0.0751	-0.0580	-0.116	0.0552	0.157	1											
INDwholesale	-0.0946	0.0248	0.00660	-0.0282	-0.0495	0.0155	-0.0434	0.00353	-0.0212	-0.0226	0.0278	-0.0284	0.0683	-0.216	-0.0653	0.101	1										
INDretail	0.0740	-0.101	-0.0725	0.0338	-0.114	0.172	-0.0994	-0.122	-0.191	0.146	0.0847	0.125	0.0458	-0.213	0.170	-0.159	-0.116	1									
INDeduc	0.149	-0.134	-0.0429	-0.0188	-0.150	0.169	-0.145	-0.160	-0.166	0.0882	0.109	0.0736	0.0630	-0.190	-0.328	-0.402	-0.0429	-0.151	1								
INDtrans	-0.0782	0.0790	0.00369	-0.116	0.0726	-0.133	0.0658	0.0938	0.191	-0.0685	-0.137	-0.0416	-0.0648	-0.120	-0.183	0.0190	0.159	-0.302	-0.135	1							
INDcommu	0.0507	-0.0427	0.0195	-0.121	-0.0646	0.0103	-0.0446	0.0178	0.0364	0.00597	0.0118	-0.0172	0.0578	-0.165	-0.131	0.0434	0.196	-0.159	0.0192	0.287	1						
INDfinan	0.0353	0.0987	0.0225	0.0322	-0.0969	0.0812	-0.0335	-0.0827	-0.0966	0.0361	0.0732	-0.0127	0.110	-0.402	-0.305	-0.371	0.0958	-0.0216	0.192	-0.121	0.00834	1					
INDpers	-0.0878	0.0478	0.0155	0.0930	0.118	-0.0903	0.0424	0.107	0.0687	-0.0666	-0.0303	-0.0175	-0.0240	0.0673	0.0137	-0.0683	-0.230	0.0515	-0.241	-0.0994	-0.212	-0.222	1				
INDent	0.117	0.00332	-0.0388	0.000066	-0.0729	0.0788	-0.112	-0.0405	-0.0741	0.110	0.0354	0.0747	-0.0178	-0.0607	-0.208	-0.227	-0.00277	-0.0431	0.121	0.0496	-0.124	0.278	-0.0391	1			
INDpubl	0.0838	-0.0503	-0.00892	0.0349	-0.0805	0.0303	-0.00629	-0.00655	0.00536	0.0700	-0.0252	0.0590	-0.00936	0.0458	-0.0533	-0.166	-0.126	-0.160	0.0724	-0.0575	0.0801	-0.175	-0.0996	-0.115	1		
CBD	0.435	-0.368	-0.160	-0.147	-0.453	0.420	-0.206	-0.590	-0.470	0.255	0.245	0.101	0.0772	0.142	-0.0404	-0.0120	-0.0732	0.0992	0.201	-0.104	-0.0294	-0.0515	-0.226	0.120	0.0288	1	
NSUB	0.585	-0.276	-0.172	-0.218	-0.349	0.319	-0.0410	-0.459	-0.381	0.171	0.139	-0.112	0.0682	0.126	-0.0519	-0.210	0.0171	0.0601	0.145	-0.0571	0.104	0.0537	-0.179	0.160	0.0920	0.817	1

2000	meancommute	afro_american	native	asian	hispanic	car	carpool	pt	vehicle0	medhincc	edumid	eduhigh	femalework	INDagri	INDconst	INDmanu	INDwholes	INDretail	INDtrans	INDinfo	INDfinanc	INDprof	INDeduc	INDarts	INDpubadı	CBD	NSUB
meancommute	1																										
black	-0.0585	1																									
native	-0.107	0.203	1																								
asian	-0.128	-0.0709	-0.0627	1																							
hispwhite	0.154	0.332	0.269	-0.243	1																						
car	0.0290	-0.385	-0.418	0.185	-0.703	1																					
carpool	0.117	0.345	0.372	-0.249	0.727	-0.838	1																				
pt	-0.0137	0.424	0.377	-0.169	0.565	-0.741	0.503	1																			
vehicle0	-0.0997	0.373	0.411	-0.160	0.580	-0.748	0.512	0.701	1																		
medhincom	0.0386	-0.334	-0.364	0.295	-0.536	0.624	-0.571	-0.522	-0.653	1																	
edumid	0.107	-0.285	-0.228	0.170	-0.673	0.667	-0.571	-0.504	-0.574	0.348	1																
eduhigh	-0.279	-0.337	-0.293	0.452	-0.641	0.562	-0.690	-0.429	-0.465	0.774	0.294	1															
femalework	-0.160	-0.152	-0.223	0.101	-0.554	0.622	-0.523	-0.467	-0.461	0.187	0.533	0.263	1														
INDagri	0.0698	0.0243	0.0552	-0.149	0.293	-0.216	0.284	0.0168	0.154	-0.129	-0.252	-0.233	-0.262	1													
INDconst	0.442	0.173	0.221	-0.342	0.704	-0.580	0.675	0.401	0.390	-0.451	-0.453	-0.666	-0.581	0.226	1												
INDmanu	0.101	0.126	-0.0282	0.179	0.105	0.0256	0.141	-0.120	-0.142	0.134	-0.0383	-0.0641	-0.0787	0.0948	0.0281	1											
INDwholesale	-0.0248	0.0590	-0.0105	0.0250	0.0170	0.0457	0.0142	-0.0667	-0.124	0.133	-0.0276	0.0191	-0.0173	0.0500	-0.0812	0.120	1										
INDretail	0.208	-0.260	-0.150	-0.0604	-0.254	0.331	-0.301	-0.280	-0.225	-0.00976	0.377	-0.0324	0.234	-0.104	-0.105	-0.202	-0.0650	1									
INDtrans	0.0577	0.164	-0.0345	0.0756	-0.0658	0.167	-0.0338	-0.165	-0.183	0.0776	0.241	-0.0688	0.0994	0.0346	-0.145	0.194	0.162	-0.0801	1								
INDinfo	-0.0761	-0.113	-0.0707	0.174	-0.339	0.340	-0.338	-0.214	-0.200	0.198	0.335	0.287	0.244	-0.157	-0.386	-0.0595	0.00693	0.0299	0.121	1							
INDfinance	-0.114	-0.248	-0.268	0.131	-0.603	0.542	-0.584	-0.338	-0.390	0.427	0.408	0.548	0.459	-0.241	-0.549	-0.353	0.00113	0.158	-0.112	0.223	1						
INDprof	-0.308	0.00681	0.0794	0.0340	-0.00143	-0.153	-0.0695	0.268	0.213	0.0799	-0.249	0.310	-0.146	-0.0922	-0.133	-0.316	-0.144	-0.316	-0.259	-0.0107	0.135	1					
INDeduc	-0.262	-0.248	-0.171	0.318	-0.554	0.474	-0.555	-0.414	-0.353	0.428	0.381	0.553	0.425	-0.201	-0.569	-0.202	-0.0578	0.0373	0.00492	0.172	0.316	-0.0467	1				
INDarts	-0.263	0.153	0.246	-0.157	0.321	-0.517	0.337	0.465	0.495	-0.478	-0.382	-0.243	-0.238	-0.0131	0.156	-0.276	-0.198	-0.203	-0.285	-0.170	-0.229	0.187	-0.316	1			
INDpubadm	0.0634	0.112	-0.0404	0.0164	-0.107	0.138	-0.00796	-0.130	-0.167	0.0796	0.201	-0.0143	0.228	0.0259	-0.143	0.00876	-0.0161	-0.0965	0.197	0.00222	-0.0203	-0.252	0.0738	-0.202	1		
CBD	0.275	-0.395	-0.339	0.0320	-0.527	0.534	-0.422	-0.586	-0.518	0.354	0.519	0.214	0.344	0.0310	-0.274	0.209	-0.0191	0.366	0.0774	0.124	0.222	-0.279	0.264	-0.380	-0.0493	1	
NSUB	0.445	-0.232	-0.318	-0.0281	-0.283	0.362	-0.223	-0.402	-0.382	0.215	0.305	-0.0127	0.302	0.116	-0.118	-0.0415	0.00915	0.274	0.115	0.0186	0.263	-0.329	0.131	-0.329	0.268	0.559	1

2010	meancommute	afro_american	native	asian	hispanic	car	carpool	pt	vehicle0	medhincc	edumid	eduhigh	femalework	INDagri	INDconst	INDmanu	INDwholes	INDretail	INDtrans	INDinfo	INDfinanc	INDprof	INDeduc	INDarts	INDpubad	CBD	NSUB
meancommute	1																										
black	-0.0268	1																									
native	-0.0840	0.165	1																								
asian	-0.183	0.0287	-0.0775	1																							
hispwhite	0.146	0.244	0.186	-0.333	1																						
car	-0.0200	-0.172	-0.200	0.0329	-0.438	1																					
carpool	0.254	0.126	0.0893	-0.166	0.596	-0.567	1																				
pt	0.0154	0.362	0.230	0.0499	0.326	-0.616	0.151	1																			
vehicle0	-0.00947	0.355	0.310	-0.150	0.468	-0.603	0.248	0.579	1																		
medhincom	-0.0619	-0.295	-0.310	0.334	-0.592	0.444	-0.408	-0.404	-0.613	1																	
edumid	0.110	-0.118	-0.0647	0.106	-0.613	0.393	-0.281	-0.269	-0.389	0.223	1																
eduhigh	-0.379	-0.218	-0.241	0.452	-0.713	0.352	-0.588	-0.264	-0.428	0.782	0.141	1															
femalework	-0.0107	-0.134	-0.120	0.0244	-0.462	0.450	-0.354	-0.307	-0.300	0.153	0.449	0.252	1														
INDagri	0.00961	0.179	0.0665	-0.0560	0.232	-0.145	0.163	0.0931	0.200	-0.145	-0.206	-0.191	-0.186	1													
INDconst	0.232	0.0250	0.112	-0.313	0.609	-0.333	0.558	0.137	0.241	-0.420	-0.378	-0.606	-0.461	0.163	1												
INDmanu	0.0718	0.0339	-0.0272	0.160	0.149	0.0146	0.133	-0.0433	-0.135	0.190	-0.118	0.0219	-0.161	0.0210	-0.0189	1											
INDwholesale	0.0194	-0.00585	-0.0611	-0.0180	0.0293	0.107	-0.0352	-0.0940	-0.153	0.125	-0.0587	0.0671	0.0167	-0.0244	-0.0231	0.136	1										
INDretail	0.0933	-0.0298	-0.0184	0.0296	-0.114	0.0352	0.00780	0.0186	-0.0487	-0.152	0.193	-0.170	0.159	0.00310	-0.0866	-0.122	-0.116	1									
INDtrans	0.112	0.0993	-0.00604	-0.0762	0.0712	0.122	-0.0131	-0.0410	-0.0301	-0.0359	0.0534	-0.141	0.0261	-0.0290	0.000583	-0.00597	-0.0425	-0.0177	1								
INDinfo	-0.159	0.0708	0.0596	0.0533	-0.177	0.0785	-0.131	-0.0882	-0.0462	0.109	0.172	0.189	0.107	-0.0246	-0.255	-0.0619	-0.0669	-0.0722	-0.142	1							
INDfinance	-0.131	-0.126	-0.150	0.220	-0.525	0.353	-0.420	-0.218	-0.283	0.473	0.253	0.550	0.279	-0.179	-0.477	-0.126	-0.0404	-0.0807	-0.0751	0.115	1						
INDprof	-0.121	0.00153	-0.0245	-0.00487	0.0841	-0.178	0.0521	0.117	0.160	0.0118	-0.262	0.122	-0.191	-0.0361	-0.0267	-0.102	-0.113	-0.321	-0.205	0.0874	-0.0380	1					
INDeduc	-0.102	-0.0898	-0.131	0.165	-0.541	0.314	-0.380	-0.252	-0.266	0.375	0.397	0.453	0.451	-0.105	-0.448	-0.206	-0.0419	-0.141	-0.0124	0.0395	0.209	-0.209	1				
INDarts	-0.185	0.0563	0.200	-0.0407	0.228	-0.291	0.0489	0.266	0.339	-0.379	-0.140	-0.169	-0.100	-0.0121	0.00918	-0.263	-0.123	-0.0746	-0.143	-0.0589	-0.224	-0.0227	-0.203	1			
INDpubadm	0.101	0.0480	0.00312	-0.0169	-0.130	0.117	-0.109	-0.0114	-0.0835	0.107	0.166	0.0730	0.0456	-0.0424	-0.176	-0.0278	0.0845	-0.0952	0.0640	0.0410	0.0335	-0.136	0.0936	-0.211	1		
CBD	0.228	-0.351	-0.241	0.0939	-0.530	0.372	-0.193	-0.483	-0.507	0.356	0.493	0.198	0.354	-0.0891	-0.229	0.0779	-0.0109	0.289	-0.0103	0.00227	0.177	-0.237	0.263	-0.249	0.00192	1	
NSUB	0.401	-0.132	-0.241	-0.110	-0.211	0.258	-0.0816	-0.358	-0.297	0.137	0.249	-0.0568	0.277	0.0400	-0.108	0.0484	0.0417	0.262	0.0973	-0.0964	0.0100	-0.288	0.167	-0.243	0.189	0.654	1

2015	meancommute	afro_american	native	asian	hispanic	car	carpool	pt	vehicle0	medhincc	edumid	eduhigh	femalework	INDagri	INDconst	INDmanu	INDwholes	INDretail	INDtrans	INDinfo	INDfinanc	INDprof	INDeduc	INDarts	INDpubad	CBD	NSUB
meancommute	1																										
black	0.0326	1																									
native	-0.114	0.233	1																								
asian	-0.163	-0.00180	-0.0640	1																							
hispwhite	0.350	0.262	0.0952	-0.293	1																						
car	-0.0656	-0.288	-0.159	0.0549	-0.389	1																					
carpool	0.361	0.297	0.111	-0.186	0.653	-0.514	1																				
pt	0.133	0.433	0.245	-0.108	0.379	-0.608	0.271	1																			
vehicle0	0.0542	0.438	0.285	-0.126	0.375	-0.607	0.278	0.676	1																		
medhincom	-0.148	-0.346	-0.254	0.301	-0.570	0.432	-0.427	-0.473	-0.608	1																	
edumid	0.00881	-0.0708	-0.0399	0.0393	-0.459	0.409	-0.340	-0.241	-0.269	0.0875	1																
eduhigh	-0.461	-0.294	-0.175	0.423	-0.741	0.323	-0.607	-0.365	-0.413	0.774	0.0520	1															
femalework	-0.0151	-0.141	-0.117	0.0419	-0.343	0.339	-0.250	-0.238	-0.268	0.118	0.309	0.182	1														
INDagri	-0.00846	-0.0160	-0.0168	-0.0357	0.134	-0.0468	0.00882	0.0293	0.0112	0.0280	-0.126	-0.0480	-0.228	1													
INDconst	0.402	0.0844	0.0457	-0.355	0.680	-0.281	0.570	0.214	0.264	-0.463	-0.315	-0.681	-0.296	-0.0120	1												
INDmanu	0.104	0.0185	-0.0345	0.156	0.106	0.110	0.0928	-0.128	-0.165	0.183	-0.0580	0.00108	-0.0841	-0.0152	0.0135	1											
INDwholesale	0.0850	-0.0740	-0.0841	-0.0487	-0.0278	0.107	-0.0392	-0.0923	-0.126	0.119	-0.0452	0.0441	0.0555	-0.0472	-0.0549	0.0727	1										
INDretail	0.127	0.0605	0.0142	-0.0838	0.0500	0.0508	0.0878	-0.00751	0.0344	-0.289	0.278	-0.322	0.169	-0.112	0.0470	-0.103	-0.0320	1									
INDtrans	0.180	0.160	0.0530	-0.104	0.143	-0.0231	0.126	0.0572	0.103	-0.125	0.0429	-0.240	0.0133	-0.0259	0.117	0.0230	0.0255	0.00536	1								
INDinfo	-0.199	-0.00966	-0.0448	0.123	-0.218	0.127	-0.175	-0.133	-0.109	0.166	0.0842	0.288	0.0964	-0.00725	-0.234	-0.0691	-0.0254	-0.0724	-0.113	1							
INDfinance	-0.269	-0.212	-0.215	0.250	-0.581	0.383	-0.496	-0.320	-0.371	0.497	0.250	0.598	0.208	-0.0841	-0.497	-0.0876	0.00753	-0.196	-0.170	0.148	1						
INDprof	-0.203	0.00616	0.0139	0.0224	0.0242	-0.151	0.0540	0.0927	0.0792	0.0955	-0.261	0.198	-0.197	-0.0331	-0.130	-0.143	-0.0777	-0.324	-0.240	-0.00789	0.0641	1					
INDeduc	-0.227	-0.152	-0.0108	0.222	-0.521	0.177	-0.406	-0.203	-0.198	0.390	0.206	0.523	0.301	-0.108	-0.470	-0.221	-0.0461	-0.189	-0.201	0.125	0.190	-0.133	1				
INDarts	-0.0666	0.157	0.149	-0.0813	0.192	-0.369	0.127	0.359	0.333	-0.422	-0.110	-0.245	-0.133	-0.0333	0.0864	-0.234	-0.210	-0.00915	-0.0769	-0.108	-0.294	-0.0847	-0.194	1			
INDpubadm	0.0707	-0.0197	0.113	0.0734	-0.164	0.231	-0.175	-0.0915	-0.121	0.169	0.194	0.145	0.106	-0.0654	-0.167	0.0149	0.0450	-0.0683	0.0499	0.0386	0.0514	-0.177	0.166	-0.232	1		
CBD	0.128	-0.376	-0.227	0.0847	-0.395	0.373	-0.251	-0.507	-0.461	0.305	0.473	0.164	0.282	0.0461	-0.206	0.174	0.0165	0.192	-0.0292	-0.0255	0.186	-0.215	0.187	-0.230	-0.0161	1	
NSUB	0.363	-0.161	-0.219	-0.148	-0.141	0.177	-0.0482	-0.324	-0.230	0.0612	0.267	-0.126	0.238	-0.0193	-0.00311	0.0442	0.0425	0.220	0.169	-0.107	-0.0533	-0.299	0.0828	-0.137	0.0753	0.617	1