ACCESSIBILITY, LOCATION, AND EMPLOYMENT CENTER GROWTH

METRANS Project 11-06

FINAL REPORT

March 2015

Genevieve Giuliano, Principal Investigator
Yuting Hou
Sanggyun Kang
Eun-Jin Shin

Sol Price School of Public Policy
University of Southern California
Los Angeles, CA

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the accuracy of the data and information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, the California Department of Transportation and the METRANS Transportation Center in the interest of information exchange. The U.S. Government, the California Department of Transportation, and the University of Southern California assume no liability for the contents or use thereof. The contents do not necessarily reflect the official views or policies of the State of California, USC, or the Department of Transportation. This report does not constitute a standard, specification, or regulation.

ABSTRACT

The purpose of this research is to examine the relationship between accessibility and the growth of employment centers in order to improve our understanding of how transportation investments influence the spatial organization of metropolitan areas. Although research on the existence of employment centers – concentrations of employment outside the tradition downtown – is extensive, we have little understanding of how these centers emerge and grow, and what role transportation access may play in this process. Studies of employment centers are limited by data availability: there is no publicly available source for reliable, highly detailed and disaggregate employment data. For this research we used the National Establishment Time-Series (NETS) Database, a proprietary database that consists of time series establishment level data for the entire US for the years 1990 through 2009. The NETS data are highly detailed; they provide information on business type, size, location, corporate structure, etc. Despite the richness of the database, however, we discovered that it is not comparable to any other data source, raising serious questions of validity. We devote one chapter of this report to an assessment of the quality and reliability of the NETS data.

We used the NETS data to identify employment centers in California's four largest metro areas (Los Angeles, San Diego, San Francisco and Sacramento) and examine their evolution over the time period in terms of number of firms, employment, and other factors. We find that the four metro areas have quite different urban structures. Los Angeles and San Francisco are highly polycentric, while San Diego and Sacramento are less so, and these characteristics are persistent over the time period of the study. We conduct formal tests of polycentricity, and find that all confirm polycentric urban form. However, the influence of centers outside the CBD is weakest for Sacramento. Consistent with prior studies, the extent of polycentricity is related to metropolitan size. We estimated models of employment center growth as a function of accessibility. We estimate one model for the Los Angeles region, and another model with a pooled sample of Los Angeles and San Francisco. We find that center density is the most consistent factor associated with employment center growth: growth is negatively related to employment density. Results on our access measures – access to airports and freeways, access to labor – are mixed. We attribute our results to the unreliability of the NETS data when used at a highly disaggregate geographic scale, the relatively slow growth that took place over the study period, and the overall maturity of urban structure in Los Angeles and San Francisco. All employment centers have relatively high access to the regional transport system and labor force, hence differences in growth rates are due to more place specific factors.

Our results suggest the need for both more reliable data sources for detailed, spatially disaggregate data and comparable studies of metro areas outside California. The role of transport and labor force accessibility in employment center growth remains uncertain.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF MAPS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Introduction, organization of report, literature review

1.1 Introduction

The purpose of this research is to examine the relationship between accessibility and the growth of employment centers in order to improve our understanding of how transportation investments influence the spatial organization of metropolitan areas. Although research on the existence of employment centers – concentrations of employment outside the tradition downtown – is extensive, we have little understanding of how these centers emerge and grow, and what role transportation access may play in this process. Few studies have examined employment center growth, and only two (Giuliano and Small 1999; Giuliano et al. 2011) have specifically addressed transportation access. These studies generated inconclusive results. Clearly there is a lot more to learn about how transportation systems may influence where firms choose to locate and where large concentrations of economic activity emerge.

Studies of employment centers are limited by data availability: there is no publicly available source for reliable, highly detailed and disaggregate employment data. Giuliano and colleagues used employment data provided by the local metropolitan planning organization. Others (e.g. Lee 2006) have typically used the Census Transportation Planning Package, which generates estimates of jobs by local area from commuting data. These data are limited in several ways: 1) they are available only for census years; 2) industry sector data is limited to aggregate categories; 3) they do not provide establishment information; 4) at the local area level, CTPP has significant variance with other data sources. For this research we used the National Establishment Time-Series (NETS) Database, a proprietary database that consists of time series establishment level data for the entire US for the years 1990 through 2009. The NETS data are highly detailed; they provide information on business type, size, location, corporate structure, etc. The time series data allows for the tracking of firm births, moves, and deaths. Despite the richness of the database, however, we discovered that it is not comparable to any other data source, raising serious questions of validity.

We used the NETS data to identify employment centers in California's four largest metro areas (Los Angeles, San Diego, San Francisco and Sacramento) and examine their evolution over the time period in terms of number of firms, employment, and other factors. We merged the NETS data with transportation network data for two of the four metro areas, and tested relationships between center characteristics, growth, and transport access for Los Angeles and San Francisco.

This project extends our previous research, funded in part by prior METRANS grants (e.g. project 06-16), in the following ways: 1) extends the analysis beyond the Los Angeles region; 2) uses establishment level geography, not census tract aggregates, allowing more precise measurement of centers and access measures; 3) provides detailed data on industry sector, employment, and firm births, deaths and moves; 4) uses shorter time intervals for temporal comparisons.

1.2 Organization of report

This report is organized as follows. The remainder of this chapter presents a brief review of the recent literature on employment centers. Chapter 2 describes the NETS and other data used in this research. It includes an analysis of the comparability of NETS with other employment data sources and explains how the research was restructured as a result of our inability to demonstrate reliability of the data. Chapter 3 describes construction of the final data sets used in our analysis, our method for identifying employment centers, and the resulting centers for the four metropolitan areas across 3 time periods. We also present formal tests of monocentricity and polycentricity. Chapter 4 presents a descriptive analysis of the four metro areas. We compare employment inside and outside centers, changes over time, and the characteristics of centers. Chapter 5 estimates models to explain center growth for Los Angeles and San Francisco, and Chapter 6 presents conclusions, policy implications, and tasks for future research.

1.3 Literature review

There is an extensive literature explaining the evolution of metropolitan spatial structure in economic terms (for example, Mills 1967; Fujita 1989). We have reviewed this literature elsewhere (Giuliano et al. 2007; Giuliano et al. 2011; Agarwal et al. 2012), and hence provide only a brief summary here.

Theories on urban form explain the existence of an employment center, such as the central business district (CBD) based on the economies of scale in production (agglomeration economies) and diseconomies in transportation and congestion. It is argued that firms locate inside employment centers to benefit from external economies of scale associated with locating in spatial proximity to other businesses. For example access to a large skilled labor pool, knowledge spillovers, and input sharing.

The standard urban model assumes a single employment center, and then explains spatial structure as the result of trade-offs between transport costs and housing consumption. However, large metropolitan areas have multiple employment concentrations. Explanations for multiple centers are based on economies and diseconomies of agglomeration: when the existing center(s) grow to a point where the negative externalities of locating inside it outweigh the benefits, at least for some firms, these firms will seek locations outside the existing centers. Agglomeration benefits could lead to the emergence of employment centers at other locations (Anas et al. 1998). Other explanations for multiple employment centers include exogenous changes in transport technology (Chen 1996), entrepreneurial efforts by local governments (Sullivan 1986; Fujita 1989; Zhang and Sasaki 1997, 2000) or by private developers (Henderson and Mitra 1996); or location decisions of large firms (Fujita and Thisse 2002).

Because of the wide availability of population data across metropolitan areas and time intervals, a large number of empirical studies on the evolution of urban spatial structure focus on the population distribution over time. The estimation of the population density gradient is the most common approach. Employment data is far more limited. Studies of employment centers usually utilize the Census Transportation Planning Package (CTPP), which uses commuting data to generate estimates of jobs by local area (transportation analysis zones or census tracts). However, CTPP data are subject to several sources of error, including sampling error and missing data on job locations.

1.3.1 Employment center identification

Despite more than two decades of research, there is no standard, accepted operational definition of a center. The identification of employment centers have developed into many different methods ranging from simple qualitative specification to complex statistical algorithms. The problem is how to identify what constitutes a “center” among the array of large and small employment clusters that characterize metropolitan areas. There are three general types of quantitative methods: (1) methods based on minimum size and density (e.g., Giuliano and Small 1991), (2) estimation of density gradients to identify potential centers (e.g., Craig and Ng 2001), and (3) various two-step methods using locally weighted regression (LWR) to smooth the density surface and then identify centers (e.g., McMillen 2001; Redfearn 2007). There is no agreement on the best method, and the size and number of centers for any metropolitan area varies by different methods. Recent studies have extended these methods in various ways, but none have solved the problem of having to make arbitrary choices to define the boundaries of a center.

Recent studies on employment have extended the minimum cutoff method and the two-stage nonparametric approach. For example, Pan (2003) assigns statistical meaning to the minimum cut-off point method. Assuming a normal distribution of employment density by tracts in a metropolitan area, the minimum density cutoff can be expressed as a point on the distribution, say for example the 90 or 95 percentile of the distribution. However, Pan did not test for whether employment density by tracts is normally distributed within regions.

Garcia-López and Muñiz (2010) extends the two-stage nonparametric approach in identifying population subcenters in Barcelona. In the first stage, a locally weighted regression of a monocentric density function is chosen and the gross employment/population density is regressed over distance to CBD in both north-south direction and east-west direction. Tracts with positive residuals exceeding the critical value are chosen as candidate subcenters. In the second stage, groups of selected tracts defined by “proximity criteria” are chosen as final centers based on the 10,000 employment or population cutoff, or the threshold of “1%” population share of the region. The effects of the two group of centers identified on the overall employment/population distribution are both estimated and the results have no significant differences.

An alternative method of center identification is based on Exploratory Spatial Data Analysis (ESDA), which takes advantage of spatial autocorrelation and spatial heterogeneity of spatial data (Baumont et al. 2004). Guillain et al. (2006) suggest that the use of ESDA would avoid the use of cutoffs and instead relies on statistical test to indicate whether local spatial associations are significant. Specifically, they define employment centers to be related to two types of local spatial association between each observation and its neighbors: the “HH” association indicating an observation with high value (e.g. above mean value) and surrounded by high values and the "HL" association indicating an observations with high values and surrounded by low values (Guillain et al. 2006). In their study of Ile-de-France, the local Moran's I index is applied to examine the spatial pattern of employment-population ratio across the 1280 communes (French municipalities) in the region, and employment centers are identified as a set of neighboring communes of "HH" or "HL" values. By comparing the significance of employment-population ratio of spatial units within the CBD and subcenters to the overall mean of the study area, their study revealed that the importance of CBD is reinforced in 1997 than in 1978, which would not be indicated by the cut-off method.

1.3.2 Intrametropolitan spatial trends

A major issue in the literature is whether employment centers (localized agglomeration economies) are becoming more or less important in the structure of metropolitan areas. Briefly, the argument is whether ICT and lower transport costs reduce the value of spatial proximity, hence reducing demand for high density, high rent locations (see Agarwal et al, 2012 for a recent review). Empirical literature on employment center growth is limited. Giuliano and Small (1999) empirically investigate a series of hypothesis to explain the determinants of growth of employment centers (between 1970 and 1980) in the Los Angeles region. Their study found no significant relationship between center growth and accessibility to labor force or access to the highway system. Our previous study revisits the connection between network accessibility and intrametropolitan employment growth between 1990 and 2000 in the Los Angeles region. The results suggest that employment center growth in the Los Angeles region is a complex process in which traditional forms of accessibility play a limited role. However, we observe indirect effects of highway access through population and employment distribution, and employment center growth is determined by its competitive position with respect to labor force.

Other studies use density gradients to examine trends in population and employment distributions. For example, Garcia-López (2010) examination of population density distribution across 3182 census tracts in the Barcelona Metropolitan Region (BMR) from 1991 to 2005 takes a commonly used linear exponential form,

$$\ln D = \alpha_0 + \beta d_{CBD} + \sum_i \alpha_i d_{sub}^{-1} + \delta d_{INF} \qquad (1)$$

in which the natural log of population density is expressed as function distance to the CBD, the inverse of distance to the nearest subcenter and to the nearest freeway or highway. The population distribution is modeled with regard to population centers instead of employment centers, because the purpose of this study is to "characterize the spatial organization of population as a whole (p. 121)", while population centers can also emerge independently from employment centers due to local amenity and transport access (Garcia-López 2010). The distance to infrastructure variable is added to the density model to test whether transport infrastructure plays a role in the population suburbanization process. Moreover, to test for change over time, the "density model" is extended to a "growth model" taking the following form,

$$\Delta \ln D = \alpha_0 + \beta d_{CBD} + \sum_i \alpha_i d_{sub}^{-1} + \delta d_{INF} \qquad (2)$$

where Δln *D* represents the changes in population density between 1991 and 2005.

The regression results of the density model show that all the distance variables are significant with expected signs, with a negative sign for the CBD and nearest highways and a positive sign for the inverse of distance to the nearest subcenters identified using either 1991 or 2005 population data with either a threshold of 10,000 habitants or 1% of the regional population. On the other hand, a cross-section comparison of density gradient in 1991 and 2005 shows that while a polycentric structure still holds, the influence of the CBD and subcenters on population distribution declines, which is also proved by the significant positive and negative "growth gradient" for the CBD and subcenters, respectively, in the growth models. However, the density gradient of transport infrastructure increased in absolute values comparing the 1991 and 2005 regression results, while its growth gradient is also negatively significant in the growth model. Thus, the study concludes that while the influence of CBD and subcenters are declining, the location of population is more and more influenced by access to transport infrastructure, and evolution of urban form in the BMR is more related to an "accessibility" model instead of a "dispersed model".

In another study of the BMR, Garcia-López and Muñiz (2010) examine the employment decentralization and de-concentration process across 164 municipalities in the region between 1986 and 2005. They adopted the same density function as in Garcia-López's (2010) study, and identified employment subcenters with the cutoff method, using the average employment density of BMR as the density cutoff and the 1% of the BMR employment total as the employment level cutoff. However, adjacent municipalities are considered as historically separate centers rather than part of the aggregate center. The regression results show that employment density always decreases with distance to CBD and transport infrastructure, as well as distance to subcenters identified using either 1986 or 2001 employment data. Different from the results of population distribution, their cross-sectional comparison of the density functions for 1986 and 2001 show that the density gradient of the CBD and transport infrastructure increases over time in terms of absolute values and significance, showing their strengthened role in influencing employment locations outside centers. The density gradient of subcenters decreases in magnitude but increases in significance, showing the maintenance of subcenters' role in structuring employment distribution. They suggest that the polycentric model of employment distribution would have been more "precisely" identified if employment were available at finer geographic scales such as the census tract level.

1.3.3 Specialization of employment centers

Urban economic theory suggests that different economic sectors would value agglomeration economies differently so that the industrial structure within and outside centers would be different (Agarwal et al. 2012). Recent studies provide some new evidence of employment center specialization in different metropolitan areas. For example, Guillain et al.'s (2006) study of the Ile-de-France region applied the location quotient to examine specialization of the CBD and subcenters identified using the local spatial autocorrelation index. The results indicate that different from North America cities, the CBD of the study region in 1997 still plays a dominant role in the provision of diversified high-order services, such as financial and insurance services, legal services, accounting services, and adverting, management consulting. On the other hand, some subcenters that specialized in high-tech sectors such electric and electronics manufacturers, aeronautics, and biotechnology industries, as well as "technical producer service" sectors including IT consultants, engineering and R&D, also emerge in 1997 due to the influence of planning policies. There are also other subcenters functioning as transport junctions and specializing in wholesale trade and "standard services" (e.g. security, cleaning, rental, mailing, and packaging), due to the availability of large space and proximity to highways or airports. The difficulty of these subcenters in attracting business services, however, indicates the non-uniform distribution of the service sector and the complementary roles between the CBD and subcenters in terms of economic bases.

As suggested by Gilli (2009), whether subcenters evolve to be more specialized or diverse suggests different stories of decentralization and agglomeration. The specialization of subcenters is a sign of "vertical disintegration" of industry at the metropolitan scale, while a more diversified economic structure of subcenters is generated more by local growth. Gilli's (2009) study of the Greater Paris Area explores how employment decentralization has led to sprawl or re-agglomeration and how this process is linked to the evolution of sectoral and functional characteristics of subcenters. Using the Ellison and Glaeser's (1994) index for measuring the relative concentration of each sector for the metropolitan area and for the subcenters, respectively, the study find that while the average sectoral concentration remains stable for the Greater Paris region as a whole during the 1975 to 1999 period, the level of specialization of subcenters decreased in terms of both mean value and standard deviation, implying that all subcenters are becoming more diversified. The correlations between specialization and size or rank of the centers are always negative and remained steady, implying that larger centers are always more diversified. The study suggests three explanations for this evolution of subcenter characteristics: (1) the growth of "highly skilled functions" for the whole area affected the subcenters by changing their professional structure, which lead to the "functional diversity" of all subcenters and increased specialization of the core for highly skilled jobs; (2) the transformation from industrial economy to a service-based economy for the whole region also reduces the specialization level of the subcenters; (3) the relocation of jobs from larger and more diverse centers tend to be averse to subcenters specialized in the same sector, which also lead to the "convergence" of subcenters in terms of economic structure. However, the study did not specifically examine how the concentration pattern varies among different sector or how the industrial composition varies among centers of different sizes.

In sum, recent studies on urban form and employment centers extended methods to identify subcenters, examined the influence of subcenters on the overall spatial organization of employment and population within different metropolitan areas, and explored the variation of subcenters in terms of economic structure and specialization or diversity patterns, as well as the differences between the CBD and subcenters in terms of economic functions. However, most of these recent studies focus on metro areas outside the US. Moreover, the availability of employment data at finer geographic scales than the municipality level enables a more precise modeling of polycentric structure of employment distribution within cities. Thus, our study may have the advantage in conducting a more precise modeling of polycentric urban form and identifying the economic structure and function of employment centers, using the NETS data disaggregated at the establishment level.

Chapter 2 Data and comparability investigation

2.1 Data

The primary motivation for this research was the availability of annual, highly detailed establishment level data for California. The National Establishment Time-Series (NETS) database is a proprietary data set developed from the Dun and Bradstreet establishment data (see Walls and Associates, n.d.). For each year starting in 1990, annual data on approximately 34 million firms across the US is generated. This allows for the construction of a firm level time series that gives detailed information on annual activity (e.g. sales, number of employees). Firms can be traced as they move, close down, or expand to multiple locations. Headquarter linkages and corporate hierarchies can be identified. Each firm is geo-located in latitude/longitude. We purchased the data for California, and used the data for the Los Angeles, Sacramento, San Diego, and San Francisco regions for this research. The data series we received includes annual data from 1990 through 2009 for approximately 5.5 million establishments.

This chapter describes the NETS data and the various checks we conducted to test its consistency and reliability. It explains the problems we discovered, and how the research was revised in light of the data limitations. The chapter is organized as follows. We begin with a brief description of the NETS data, including the definition of establishments, description of industrial sectors, employment and sales and the imputation of missing data. Then we conduct a cross check between the 2000 NETS data and the 2000 employment data provided by the Southern California Association of Governments (SCAG) and used in our previous research. The SCAG data compared well with county level employment counts from other sources. Specifically, we test whether identification of employment centers is robust across different data sources. Based on the results of the first data check, we conduct another comparison with the County Business Pattern (CBP) data to assess the accuracy and reliability of NETS at the county level.

2.1.1 The NETS data

The NETS dataset is an annual series database based upon the Dun and Bradstreet (D&B) national establishment data. It includes establishment-level information on industry sector, location, headquarters and performance. Basically, the NETS database identifies establishments of three types of ownership structure: (1) standalone firms, (2) headquarters, (3) branches/divisions. Each establishment is assigned a unique D-U-N-S number that is not allowed to change.[1] The surveyed establishments self-report their Standard Industrial Classification (SIC) code from a codebook developed by D&B with over 18,750 8-digit SICs, or they are given an assigned SIC code, both of which are converted to 6-digit North American Industry Classification System (NAICS) codes by Wall & Associates.[2] The conversion from

[1] Walls and Associates (2008), Understanding the NETS Database
[2] Walls and Associates (2008), Understanding the NETS Database

SIC to NAICS is a reflection of changes in technology and diversification within the service sector. The number of broad sectors in NAICS is twice that in SIC.[3] The NETS database, however, keeps the records of standard SICs since 1989 to maintain the consistency of time-series data in terms of detailed economic performance.

Collecting information from numerous databases and making various efforts to contact establishments to check and update information every year, the establishment counts by Dun and Bradstreet are close to "a complete annual census of American Business" that verifies the existence and operation of establishments every year.[4] Moreover, the birth of new business is detected through multiple sources, including both public records/government registrations and news and media reports, and is distinguished from the relocated ones. A new establishment will not appear in the database until it is actually doing business.[5] Thus, the data quality control in the NETS database is assumed to be more accurate in terms of establishment counts.

There are some significant differences between the NETS database and federal employment data[6]: (1) the NETS database captures both full- and part-time employees and counts all persons, including the owner, as employees in a location; (2) self-employed establishments are reported the NETS dataset; (3) the NETS database is more "sluggish" as new jobs are verified only when it is proved they are "enduring", and changes in employment are reported only when it is considered as "significant" to the establishment manager, which makes the job numbers reported by establishment in the NETS dataset "move in a 'ratchet manner'".

Another important component of the NETS database is the estimation for employment and sales. From 1990 to 2008, about 76% of employment in the entire NETS data is reported and the rest of data are estimated using simple time series regression for existing standalone firms with prior reported employment information, or with the information of median establishment size of the same SIC8 group for new standalone firms, or the median size of branches of the same firm for new branches and divisions.[7] The sales data are estimated in a similar way. Thus, we may assume that the employment and sales data in the NETS database would not necessarily be comparable to employment or sales data collected in a different way.

2.2 Investigating comparability and reliability of NETS data using SCAG data by replicating Giuliano et al. (2007)

2.2.1 Data and methodology

As the first phase of the NETS data investigation, we compared the NETS dataset with the SCAG (Southern California Association of Governments) employment counts. The SCAG dataset is used in Giuliano et al. (2007), and the methodology used in the preliminary statistical

[3] See http://www.bls.gov/ces/cesnaics.htm and http://www.census.gov/eos/www/naics/federal_register_notices/notices/fr09ap97.pdf for a more complete discussion of the conversion between the two industrial classification system.
[4] Wall and Associates (2008), Understanding the NETS Database
[5] Wall and Associates (2008), Understanding the NETS Database
[6] Wall and Associates (2008), Understanding the NETS Database
[7] Wall and Associates (2008), Understanding the NETS Database

analysis is adopted in this chapter. The rationale behind using Giuliano et al. (2007) as a benchmark is to test the compatibility of the NETS dataset with other data sources. While the two data sets are collected in different ways, we expect that the data should be roughly comparable (e.g. have comparable total employment counts at high levels of aggregation, and roughly comparable spatial distributions). As part of the prior research, the SCAG dataset was compared with publicly available county level data, and found to be quite comparable. Thus we reasoned that using the NETS data to replicate the prior study should demonstrate the reliability of the NETS data and help us discover any errors in coding or programming.

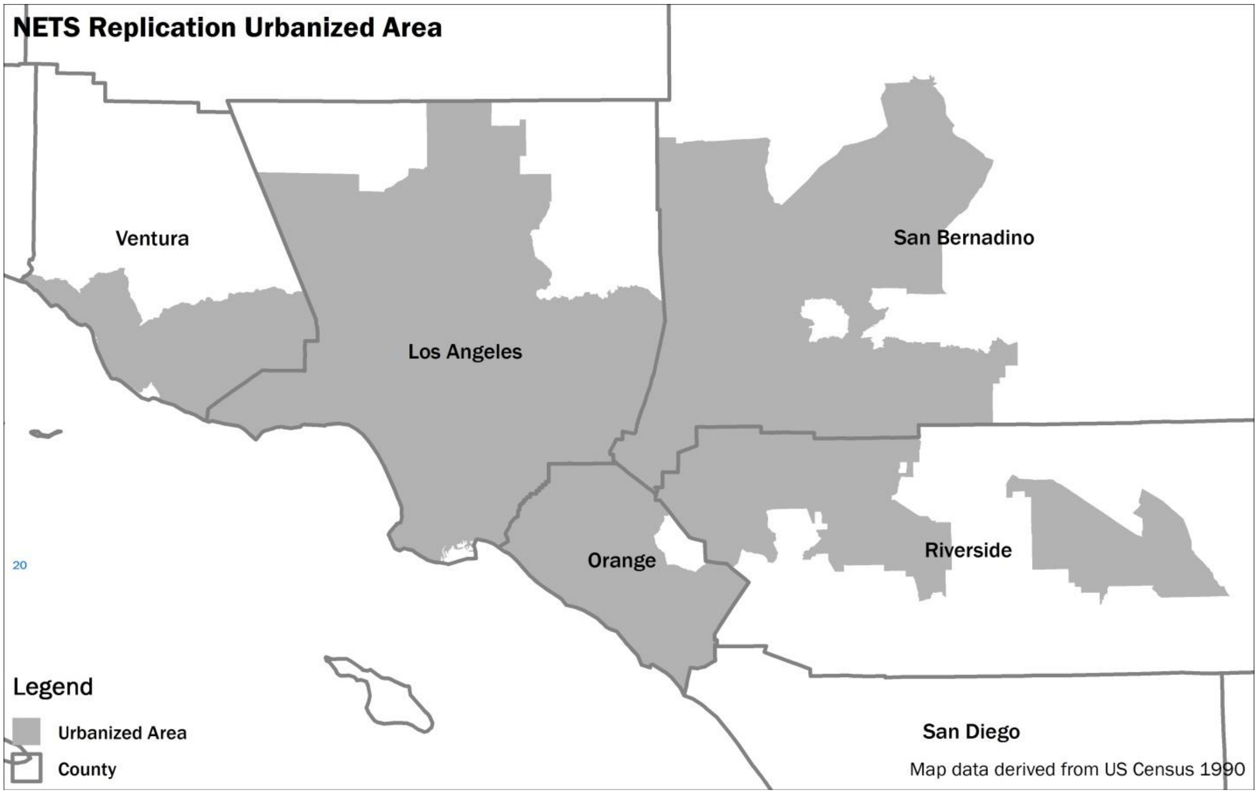
SCAG is the metropolitan planning organization (MPO) in Southern California, which serves the counties of Los Angeles, Orange, Riverside, San Bernadino, Ventura and Imperial. The SCAG dataset is based on the wage and compensation data reported to the State Economic Development Department.[8] Giuliano et al. (2007) use urbanized portion of the Los Angeles CMSA (consolidated metropolitan statistical area, which excludes Imperial County (see Map 2-1).

In order to make this replication comparable to the former study, we aggregated the establishment-level NETS dataset into the exact 1990 census-tract-based urbanized area geography used in Giuliano et al. (2007).. The NETS dataset is transferred to ArcGIS point shapefile format and analyzed within the ArcMap 10.2 application. We chose the year 2000 dataset in NETS as the replication counterpart of the former study. The employment center identification methodology is adopted from Giuliano and Small (1991), which was also utilized in Giuliano et al. (2007). The employment center defined by Giuliano and Small consists of multiple adjoining census tracts which satisfy two cut-off criteria. Firstly, a census tract should satisfy a minimum density cut-off (D). And secondly the total employment count (E) contained in one contiguous cluster should exceed the set threshold. As applied in Giuliano et al. (2007), we adopted the ten-jobs-per-acre density cut-off criterion (D) and the ten thousand total jobs per center cut-off criterion (E). We will denote the cut-off criteria D and E, as ‘D/E.’ For instance, in this replication study, we use 10/10 criteria. The comparison shows that NETS and SCAG datasets are quite different, as will be described in the following sections.

2.2.2 Employment density trends

The comparison begins with comparison of county level employment counts. Table 2-1 shows that county level differences range from 2% in San Bernardino to nearly 20% for Orange County – a difference of 300,000 jobs. In all cases NETS is larger than SCAG; a possible explanation is NETS’ greater emphasis on capturing small and single worker establishments. However, if the difference were simply due to this, the variation across counties should be more systematic, even accounting for differences in industry mix.

[8] Per Giuliano et al (2007), EDD data are based on a random survey of employers. They include all jobs subject to wage (tax) reporting. They exclude self-employment and private household workers. See http://www.calmis.ca.gov/file/resource/indmeth.htm for more information.



Map 2-1 Urbanized area in LA CMSA used in Giuliano et al. (2007) and for the comparison

Table 2-1 Employment (millions) in urbanized area by county in 2000

County	SCAG	NETS	Difference
Los Angeles	4.44	4.96	11.7%
Orange	1.51	1.81	19.8%
Riverside	0.43	0.45	4.4%
San Bernardino	0.55	0.56	2.3%
Ventura	0.31	0.33	6.4%
Total	7.24	8.11	12.0%

Table 2-2 compares the distribution of employment density: the share of all employment in each county that is located in the top 10% of (employment) density tracts. The greater the percentage, the more concentrated is employment. Differences in Table 2-2 are much smaller; for each county SCAG and NETS are within a few percentage points, with the exception of Ventura County.

Table 2-2 Share of jobs contained in the densest 10% of land area

County	SCAG	NETS
Los Angeles	64.4	64.0
Orange	50.6	52.5
Riverside	59.3	55.4
San Bernardino	78.2	80.6
Ventura	63.7	57.4
Total	71.1	71.5

Another way of measuring job density distribution is to generate shares of jobs in various density categories, as in Table 2-3. In this case the share is lower for NETS in the lowest density category, and higher for NETS in the highest density category. Whether these differences are due to the way the data are compiled or some other factor is unknown. It is clear however that even at aggregate levels, there are substantial differences between the two data sets.

Table 2-3 Share of total metropolitan jobs (percentage) in 2000

Tracts	SCAG	NETS
< 10 jobs per acre	58.0	52.8
10 - 20 jobs per acre	22.3	22.8
>= 20 jobs per acre	19.7	24.4

2.2.3 Employment centers

In this section, we compare employment centers using the two data sets. The SCAG centers are exactly those from Giuliano et al (2007). Given that NETS has a larger employment total, we might expect more centers or larger centers, given the fixed land area. The SCAG data generates 48 centers, and the NETS data generates 50. However, the NETS data generates quite a different distribution by rank size, and a substantially different share of total employment in centers. Table 2-4 shows that the differences are greatest in the middle size categories. Although a few larger centers are plausible, there is no apparent explanation for such large differences. Table 2-5 provides a summary of employment inside and outside of centers. It shows that almost all the additional employment in NETS is located in centers. See for example the Orange County numbers. As a result, NETS shows a much larger share of all employment located in centers (43.5% vs 37.5% for SCAG).

Finally we compare the spatial distribution of the two sets of centers. Map 2-2 locates both centers on the same map. SCAG centers are light grey and NETS centers are in dark gray.

It can be seen that there is substantial overlap, but there are also notable differences. Ten centers only exist in NETS; four centers only exist in SCAG. A total of 338 tracts and 272 tracts are included respectively in NETS and SCAG. Of the total 377 tracts identified in either set, 233 tracts (or 62%) are common to both. As might be expected, the smallest centers have the least congruity. We expect this to be due at least partly to the arbitrary cutoff in the center identification method; centers that just make or don't make the cutoff are most vulnerable to differences in the data. We also observe that some of the larger centers are split or combined differently, again possibly due to small differences in the underlying data. While the largest centers exist in both cases, and the general spatial pattern of density is relatively consistent, the difference in results calls into question whether the Giuliano and Small method is robust enough in view of the variation in the data.

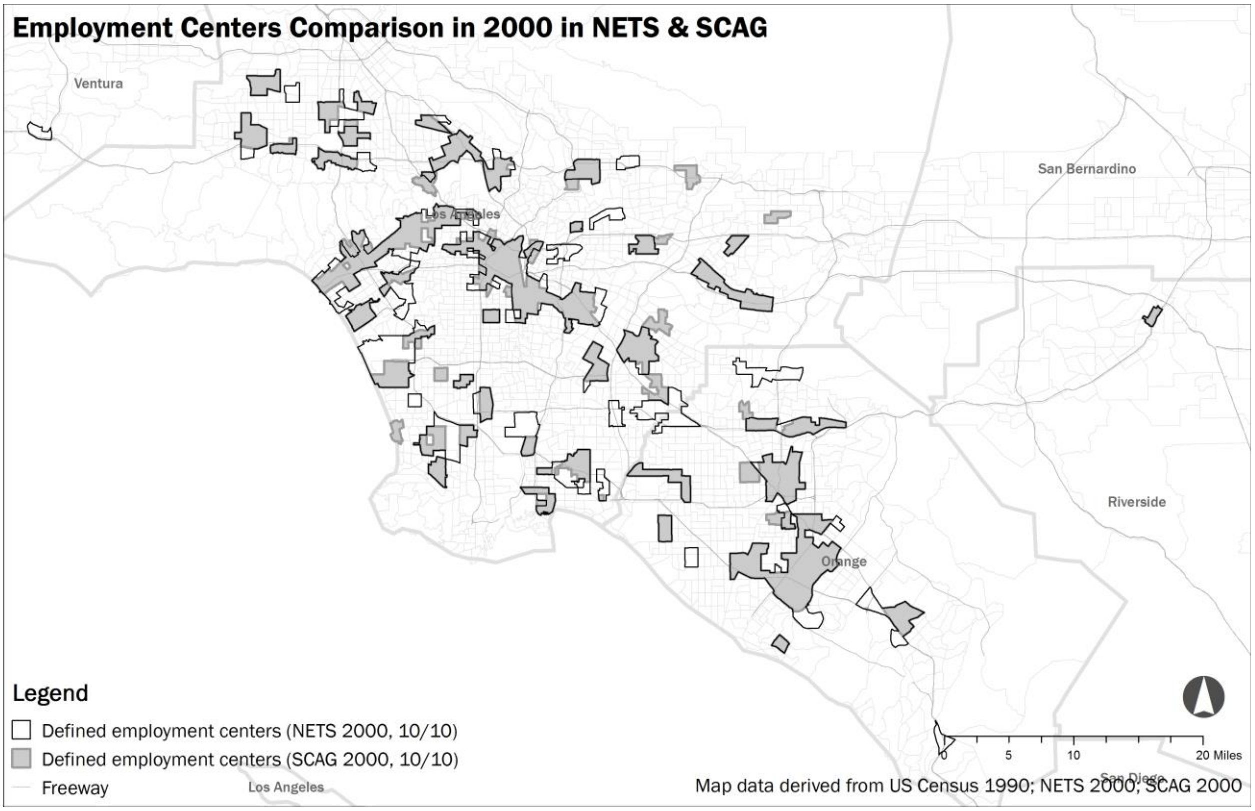
Table 2-4 Rank size of 10-10 centers in 2000

	SCAG	NETS
Number of centers with 10 000 - 20 000 jobs	20	19
Number of centers with 20 000 - 50 000 jobs	18	13
Number of centers with 50 000 - 100 000 jobs	5	11
Number of centers with 100 000 - 500 000 jobs	4	5
Number of centers with > 500 000 jobs	1	2
Total number of centers	48	50

Table 2-5 10-10 employment centers summary table

Area	SCAG			NETS		
	N	emp.	share (%)	N	emp.	share (%)
Within centers						
Los Angeles main	2	985,142	36.3%	2	1,174,707	33.3%
Rest of Los Angeles County	37	1,065,156	39.2%	35	1,422,939	40.3%
Total Los Angeles County	38	2,050,298	75.4%	37	2,597,646	73.5%
Orange County	9	652,593	24.0%	11	905,967	25.7%
Other counties	1	14,674	0.5%	2	28,381	0.8%
Total centers	48	2,717,565		50	3,531,994	
Center share of region total			37.5%			43.5%
Outside centers						
Los Angeles county		2,393,120	52.9%		2,385,820	52.1%
Orange County		861,738	19.0%		881,320	19.2%
Other counties		1,269,099	28.1%		1,316,096	28.7%
Total non-centers		4,523,957			4,583,236	
Non-center share of region total			62.5%			56.5%
Region total		7,241,523			8,115,230	

N, Number of centers; emp., employment.



Map 2-2 Employment Center Comparison in 2000 in NETS & SCAG

2.2.4 Industry sector comparison

Another important dimension for comparability is industry sector data. As with total employment in large units (counties), we should expect consistency or, if data collection methods are related to industry sector, systematic differences. In this section we compare employment count differences between NETS and SCAG datasets using one-digit SIC industry sectors. [9] The census-tract-level one-digit SIC industry sector employment counts were available in the SCAG dataset. Establishment-level NETS data was aggregated into census tracts and into the same SIC codes. Results by county, and by inside/outside centers, are given in Table 2-6. The percentage difference was calculated by dividing the difference between NETS and SCAG count by the SCAG count. Values represent the percentage to which the NETS dataset over- or under-estimates employment relative to the SCAG dataset.

If data set differences were simply a function of more comprehensive counts on the part of NETS, then we would expect about the same 12% difference across all sectors. If small establishments are more concentrated in some industry sectors, we would expect differences in shares across sectors. Unfortunately the SCAG data does not include establishment counts by industry sector, so we cannot test whether observed differences are related to sector level establishment size composition. Table 2-6 shows big differences in total sector share at the regional level (last column in table), and not all differences are intuitively obvious. While we

[9] SIC (Standard Industry Classification)

would expect NETS to do better with FIRE (since there are so many independent consultants in real estate, financial services, etc.), it is not obvious why NETS should greatly over-count mining or under count public administration. Also, for the larger sectors, these differences represent a very large number of jobs. For example, the difference in service employment is nearly 600,000 jobs, or fully 7.3% of total regional NETS employment.

Table 2-6 also breaks out these differences between centers and non-centers, and they are even more dramatic. For example, retail is over-counted by 26% inside centers, but slightly undercounted outside centers. In contrast, the public administration under count is much greater outside centers than inside centers. Clearly the process for assigning industry code must be quite different.

Table 2-6 Percentage difference in one-digit SIC industry sectors in 2000 between NETS and SCAG

% Difference in one-digit SIC industry sector	Employment Center Census Tracts	Non-Employment Center Census tracts	Total
Agriculture	87.9%	4.5%	13.0%
Mining	104.2%	11.0%	38.7%
Construction	- 2.9%	- 4.2%	- 3.8%
Manufacturing	21.0%	14.4%	17.6%
Transportation	28.7%	6.8%	16.5%
Wholesale	24.9%	14.7%	19.3%
Retail	26.1%	- 2.8%	4.5%
FIRE	54.2%	33.9%	43.2%
Service	26.4%	19.4%	22.1%
Public Administration	- 40.3%	- 62.2%	- 54.1%
Total	21.5%	6.3%	12.0%

Our comparison of SCAG and NETS data show that they have limited comparability. Employment counts, their spatial distribution, and their sectoral distribution are quite different, even at high levels of aggregation. We conclude that we cannot compare analysis results across the two data sets. Nor can we draw any conclusions regarding data accuracy or reliability. Is the SCAG data better than NETS? We cannot say, because we have no way to verify the data. Thus 3 possibilities are equally likely: 1) the SCAG data is more accurate; 2) the NETS data is more accurate; 3) neither is accurate. If accuracy of the data cannot be established, the entire exercise of identifying and analyzing employment centers, by whatever center identification method, is called into question. Given the importance of this question for employment research more generally, we decided to conduct a second set of comparisons, this time with the County Business Pattern data.

2.3 Investigating comparability and reliability of NETS data using County Business Pattern data

Our results from comparing the 2000 NETS and SCAG data sources showed that the two sources have limited comparability. Differences were not systematic, making it difficult to discover the source of these differences. Because we do not know the exact details of how the SCAG data were compiled, we have no way to determine which data are more "correct" and reliable. We therefore conducted a second comparison with County Business Pattern (CBP) data. CBP is produced by the U.S. Census Bureau and is also widely used for studying county-level economic activities or as a benchmark for other economic statistical data. If the cross-checks between the two datasets of the two different sources are comparable in terms of total and sectoral employment and establishment counts, and if we trust in the accuracy and reliability of the CBP dataset, we may conclude the NETS dataset is more "correct" at least at the county level.

This section is organized as follows. We begin with a brief introduction of the CBP dataset and compare it with the NETS database in terms of data sources, collection methods and methods for data imputation. We then examine the comparability between the two datasets in terms of aggregate employment counts, sectoral employment shares and the establishment counts by different size categories. Based on the problems we have examined, this section concludes with a discussion on how the research was revised due to the data limitations.

2.3.1 The CBP data

The source of employer establishments included the CBP dataset is the Business Register, which maintains a record for each known establishment and company that is located in the United States, Puerto Rico and Island Areas with "paid employees".[10] Different from the D&B data, the Business Register database is used for Census Bureau economic data programs. Per federal law, individual establishment level data cannot be released to the public.

Similar to the NETS datasets, an establishment in the CBP datasets is defined as "a single physical location at which business is conducted or services or industrial operations are performed" [11], which could be equivalent to a standalone firm or a headquarter/branch of a multi-unit firm. The CBP dataset includes all the business operating with an Employer Identification Number (EIN) and at least one employee (excluding the owner), while other self-employed business and business in some sectors are excluded (see the following discussion). However, establishments of some small multi-unit companies may be missed because the Census Bureau only creates the "multi-unit" structure for firms with more than 10 employees and the annual Company Organization Survey (COS) only covers multi-unit firms with more than 250 employees), and there is no estimated information for those missed establishments.[12] Another drawback of the dataset is the possible delays in business relocations, particularly for small business, since the dataset geographically assigns the physical location of establishments based

[10] Source: http://www.census.gov/econ/cbp/
[11] Source: http://www.census.gov/econ/cbp/
[12] Source : http://www.census.gov/econ/cbp/methodology.htm

on the administrative records or mailing addresses and would not update the physical location address until it receives the updates.[13]

Historically, the CBP dataset has used SIC codes for industry classification. Starting in 1998, the CBP published county-level summary statistics on a NAICS 2007 basis, and the SIC-based data are no longer available. The industry classification data is derived from the Economic Census or other Census surveys. Most of the NAICS industries are included in the dataset except crop and animal production (NAICS 111,112), rail transportation (NAICS 482), Postal Service (NAICS 491), pension, health, welfare, and vacation funds (NAICS 525110, 525120, 525190), trusts, estates, and agency accounts (NAICS 525920), private households (NAICS 814), and public administration (NAICS 92).

Instead of relying on the self-reported data, the CBP dataset uses the administrative record data to provide annual series of county-level information, including the number of establishments, employment during the week of March 12, first quarter payroll, and annual payroll. The employment data for single establishment firms are data from administrative payroll records by EIN, while employment data for multi-unit companies primarily comes from the Census Bureau's Economic Census conducted every five year or the annual COS.[14] There are, however, still employment data missing from the 12% administrative payroll records and are imputed (1) as average employment of the two adjacent quarters if available; (2) using average wage data for the prior year for the EIN; (3) using the average wage for the industry and geographic area.[15] It is suggested that the usage of administrative records keeps a good coverage of payroll and employment data. The differences between the NETS and CBP dataset is summarized in Table 2-7.

To assess how well the NETS measures employment levels, we compare the NETS dataset with the County Business Pattern (CBP) dataset by aggregating the establishment-level NETS data to the county level. As suggested by previous studies, we dropped the data before 1993 from the NETS dataset due to data accuracy problems and compare the sectoral employment difference between the two datasets. We adjust the NETS data to make it as comparable as possible to CBP by eliminating all the establishments from the NETS dataset with NAICS industries excluded by the CBP database. We also checked the secondary and third SIC codes and eliminated those establishments with some nature of public administration sectors as indicated by the secondary SIC codes belonging to "government/private sector". Finally, we eliminate all self-employed establishments (Employment=1) from the NETS database.

[13] Source : http://www.census.gov/econ/cbp/methodology.htm
[14] Source : http://www.census.gov/econ/cbp/methodology.htm
[15] Source : http://www.census.gov/econ/cbp/methodology.htm

Table 2-7 Comparison of NETS and CBP

		NETS	CBP
Data source		Dun and Bradstreet	Business Register
Primary Purpose		Credit Scoring	Economic Census Programs
Exclusion of establishments		None	Business operating without an EIN or without any employee; self-employed business; business of some sectors; Establishments not responding to the Economic census;
Industrial classification	Data collection	Self-report	Economic Census
	System used	SIC / NAICS 2002	SIC(1986-1997); NAICS2007 (1998 to now) [16]
	Sectors excluded	None	NAICS: 111,112,482,491, 525110, 525120, 525190,525920,814, 92
Employment counts	Standalone firms	Self-report	Administrative payroll records
	Multi-unit firms	Self-report	Economic Census or COS
Dealing with missing employment	Standalone firms	1. Simple time-series regression for existing firms; 2. use median establishment size of the same SIC8s group for new standalone firms	1. Average employment of the two adjacent quarters if available; 2.Using average wage data for the prior year for the EIN; 3. Using the average wage for the industry and geographic area
	Multi-unit firms	1. Simple time-series regression for establishments; 2.Median size of branches of the same firm for new branches and divisions	

[16] The definition of NAICS can be found at http://www.census.gov/eos/www/naics/. According to the definition, there is no difference between the 2002 and 2007 NAICS index at the 2 digit level.

2.3.2 Comparison of aggregate employment and establishments

To examine the employment difference between the two datasets, we first computed the aggregate employment difference for each county in California as the percentage employment difference between NETS data and CBP data. The results for the county-level total employment differences between the two datasets are available in Table A-2-1 in Appendix 2. We observe that 1) the county level differences in employment and establishment numbers are quite large, with the job counts in the NETS dataset approximately 35% to 45% larger than that in the CBP; 2) there is a slight trend of declining difference in total employment over time; 3) there is a pronounced trend of increasing difference in total establishments over time. This is due to the much more rapid rise in NETS counts of establishments compared to CBP.

Because we removed all the establishments and associated employment in NETS that we know do not exist in CBP, we do not know why the counts are so different. One possibility is that NETS does a better job of identifying small firms; hence the greater difference in establishment numbers than employment numbers (see section 2.3.4 below).

We also aggregate the county-level total employment data into the four study regions; summary tables (2a-2 to 2a-5) for each region are available in Appendix 2. The trends of the aggregate employment differences for the four regions as well as the annual mean values of the county-level total employment differences are plotted in Figure 2-1. In general, the employment differences between the two datasets at the CSA level are usually lower than the average county-level differences, as would be expected by further aggregating the data. The percentage employment differences between the two datasets for the four regions also declines throughout the study period. While trends are similar for each of our metro areas, they are not the same; note for example the pattern for San Francisco compared to San Diego.

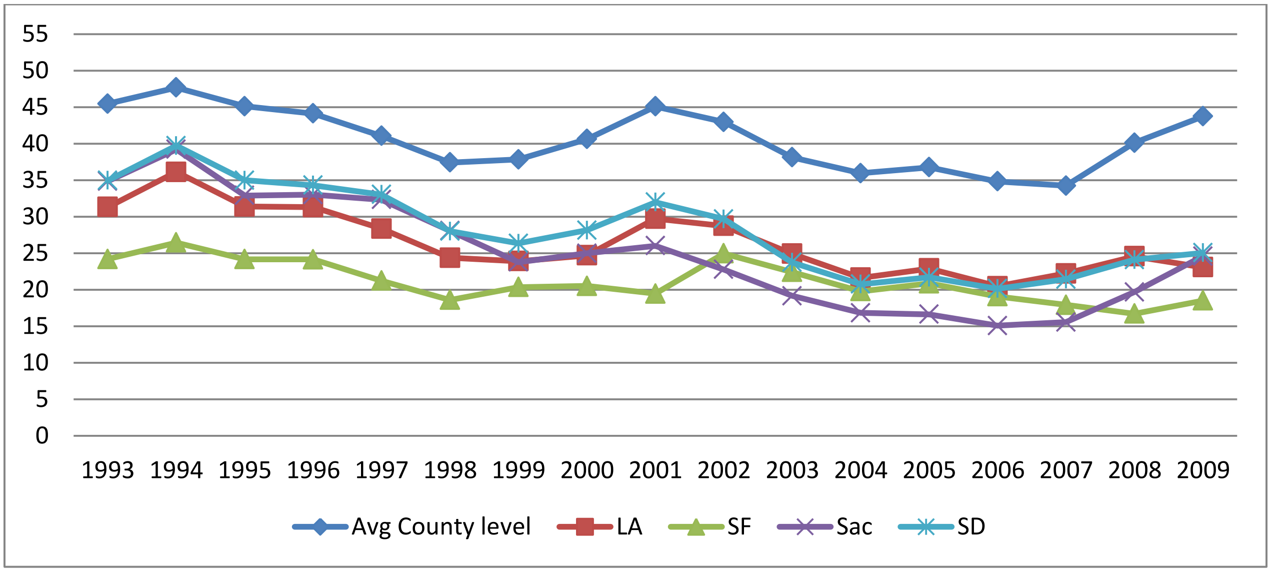


Figure 2-1 Percentage employment difference (%) between NETS data and CBP data

2.3.3 Comparison of sectoral employment

As explained earlier, in CBP industrial classification before 1998 was in the SIC system and after 1998 was replaced by the 2007 NAICS. Following the same format of data reports as the CBP datasets reports, we aggregate the cleaned NETS by county at the 1-digit SIC level from 1993 to 1997 and the 2-digit NAICS from 1998 to 2009. We drop all the county records from both the NETS and the CBP dataset if the county-level employment statistics in the CBP dataset are missing or withheld for any year in a given sector due to confidentiality reasons.[17] In this way, the number of counties used for comparing sectoral employment levels varies among different sectors, but is consistent through the study periods for each individual sector.

To separate the differences in sectoral employment counts from the total employment counts and examine how the employment count differences varies among different sectors, we calculated the ratio of sectoral employment shares between the NETS and CBP datasets:

$$industrial\_dif_i = \frac{NETSemp_i / \sum_i NETSemp}{CBPemp_i / \sum_i CBPemp}$$

where i represents sector i. A value of 1 implies that the employment share of sector i in the two datasets are the same. If the ratio is greater than 1, the share estimate is higher for NETS; if the ratio is less than 1, the share estimate is lower for NETS. We use 10% as a flexible criterion. That is, if the value of industrial ratio falls in the range of 0.9 to 1.1, we consider the employment shares of the given sector relatively comparable between the two datasets. The descriptive statistics of sectoral employment difference between the two datasets at the 1-digit SIC level from 1993-1997, as well as the statistics of sectoral employment difference at the 2-digit NAICS level from 1998-2009 are available in summary Tables 2-1 and 2-2, respectively, in Appendix 2.

Figure 2-2 plots the average value of the industrial differences index across all the CA counties for some industrial divisions from 1993 to 1997. Since the ratio controls for aggregate employment, the index simply shows the differences in sectoral employment as if the total were the same. As indicated by Table A-2-6 and Figure 2-2, the manufacturing, wholesale and retail sectors have indices value lower than 1, with the retail sector showing the lowest average index (approximately 0.8), which indicates that the employment share of the three sectors in the NETS dataset is relatively undercounted compared to the CBP dataset. On the contrary, the transportation, FIRE and service sectors are relatively over-counted in the NETS dataset. The differences in employment share for the manufacturing, transportation, wholesale trade and FIRE sectors are relatively small (around 10%), while the differences in employment shares for the

[17] The sectoral-level statistics for some counties is withheld in the CBP because of confidentiality reasons or because the data does not meet publication standards. Instead of reporting the actual employment number, the CBP uses a “data suppression flag” to denote the employment size class for the withheld data records. Another way to retrieve these missing records is to proxy the employment level by the mean value of each employment size class. For example, the value of 1750 may be used if the employment size belongs to the “1000-2499” group. However, we abandon this method because it does not correctly reflect the employment growth across different years.

retail trade and service sectors are relatively large (more than 20%). The industry ratio in the service sector shows the largest differences, with the sectoral employment share in NETS dataset over-counted by about 25% to 30%. A possible explanation is the large portion of small-sized firms (less than 10 employees) in the service sector in the NETS dataset; more than 81%, ranking first among all the sectors.

Similarly, Figure 2-3 plots the average value of the industrial ratio across all the CA counties for some 2-digit NAICS sectors from 1998 to 2009. As expected, differences in sectoral employment shares are more pronounced when the sectors are classified at a more detailed level. There is no apparent over time: the over-count for FIRE jumps from 2006 – 2008 after years of stability. The over-count in professional services trends downward until 2006, then increases. Retail trade is consistently under-counted, but not by much. From looking at the NETS data, we know that the number of small establishments increases over the period (presumably due to better data gathering). Thus we can expect more over-counting in sectors that have more establishments, meaning upward trends, not varying trends. In any case it is important to note that Figure 2-3 is based on employment, not establishments. Even if many additional small establishments are added to the database over time, this should have a relatively small effect on employment counts over time.

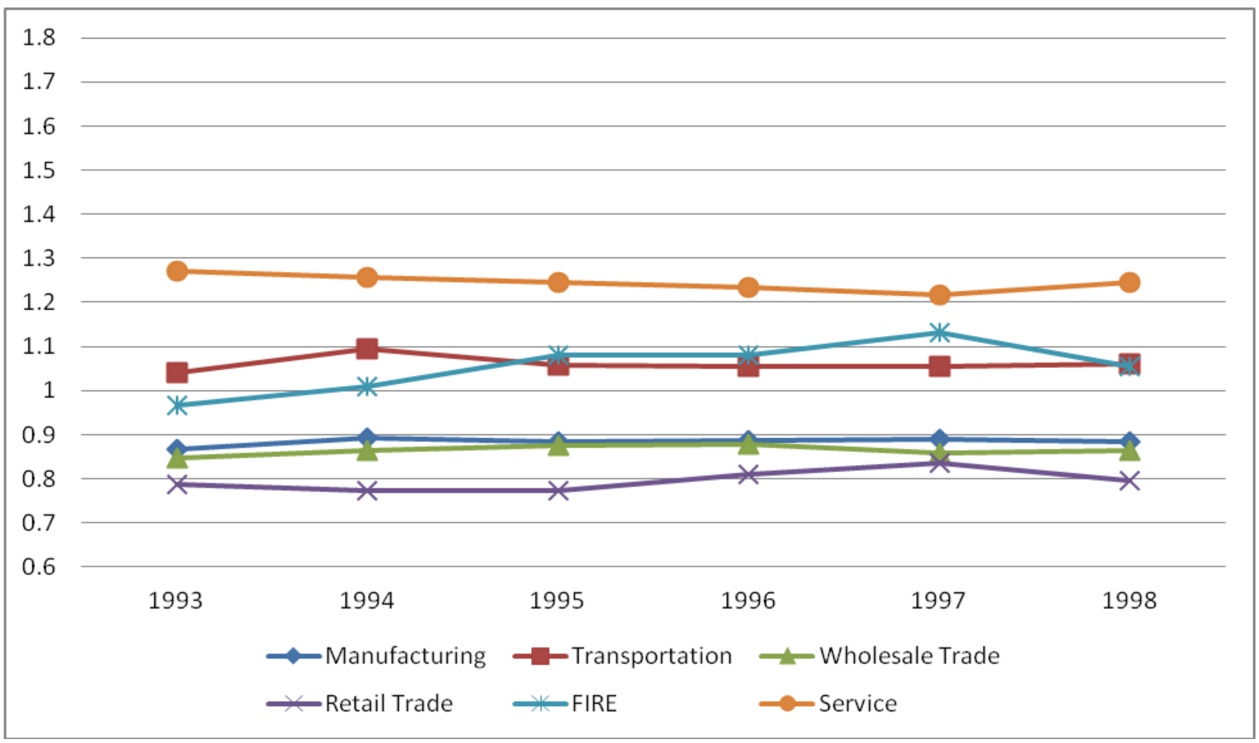


Figure 2-2 Ratio of selected sector employment shares, (NETS emp share / CBP emp share, averaged at the county level, 1993-1997), 1 digit SIC, all California counties

There are systematic differences in sectoral employment data between the two data sources at the county-level. We expected that differences between the two datasets will be reduced when data are aggregated to larger geographic scales. To test for this, we aggregate the county-level data to the CSA level and calculate the industrial difference index for the four regions separately. The data comparison also allows us to determine whether the industrial difference varies among the four study regions. Similar to the county level comparison, we drop some of the counties within the four regions if the sectoral employment statistics for the county is missing or withheld for any year in the CBP dataset.

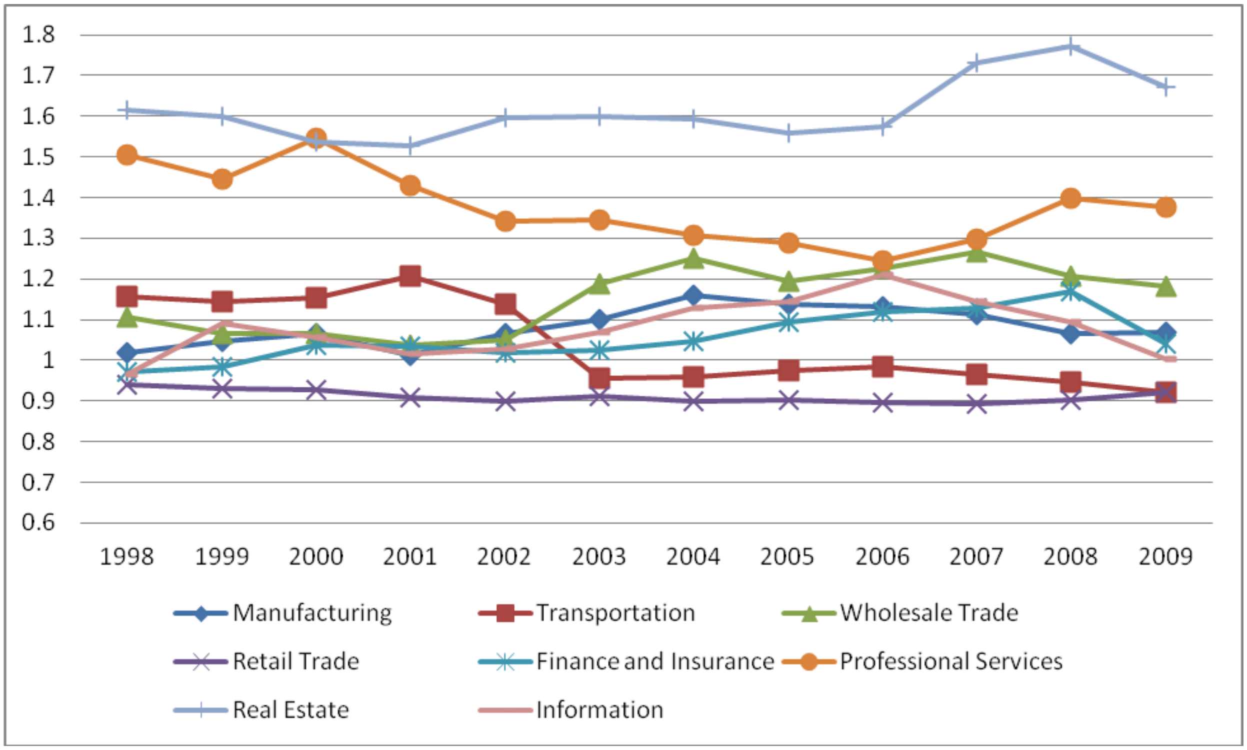


Figure 2-3 Ratio of selected sectoral employment shares (NETS emp share / CBP emp share, averaged at the county level, 1998-2009), 2 digit NAICS, all California counties

2.3.3.1 1993-1997

Table 2-8 gives a summary of average industrial difference index for each region over the 5 year period of 1993-1997. We also show NETS sectoral employment shares and the average county level difference in share for each region. Industrial ratios with differences of greater than 10% are highlighted. Every region has at least two sectors for which the share difference is greater than 10%, and the sectors that have big differences are not consistent across the regions. Some sectors are consistently over- or under-counted (manufacturing, service), while others are not. Some of this may be due to the industry mix of the regions themselves. For example, Sacramento has a relatively small share of manufacturing, but more retail trade and services. Los Angeles and San Francisco have very comparable industry mix, and the same pattern of under- and over-counting across sectors. It is important to note that we are dealing with large aggregations of jobs (millions in most cases) across broad industry categories. Difference in excess of 10% questions the reliability of the data.

Table 2-8 Mean NETS/CBP industry sector ratio and sectoral employment share (NETS), 1993-1997

		Manufacturing	Transport	Wholesale	Retail	FIRE	Services
Los Angeles	Mean ratio	0.91	0.91	0.92	0.88	1.13	1.1
	Emp share (NETS)	16.85	5.12	7.36	16.34	8.13	40.80
San Francisco	Mean ratio	0.98	1.00[a]	0.84	0.87	1.06	1.1
	Emp share (NETS)	16.92	6.60	6.20	16.13	8.37	40.51
Sacramento	Mean ratio	1.04	0.83	0.83	0.79	0.92[b]	1.18
	Emp share (NETS)	10.19	5.25	5.08	19.54	8.09	43.19
San Diego	Mean ratio	1.04	0.91	0.91	0.81	1.15	1.11
	Emp share (NETS)	14.87	4.05	5.39	18.18	8.67	42.71
Average by all CA Counties	Mean ratio	0.88	1.06	0.86	0.80	1.05	1.24
	Emp share (NETS)	15.75	5.51	6.50	17.12	7.93	40.98
	Number of Counties	54	51	50	56	48	57

a. Sonoma County is dropped for calculating the employment share of transportation sector in the region
b. Yuba County is dropped for calculating the employment share of FIRE sector in the region

2.3.3.2 1998-2009

A similar pattern is shown in Table 2-9 with the NAICS sectors, but as might be expected, with more sectors (less aggregation), there are greater differences (although we have aggregated a greater number of years). Los Angeles has the fewest sectors with differences in excess of 10%. Again there is consistency down some columns (retail trade is consistently under-counted, real estate and professional services over-counted), but not others.

Table 2-9 Mean NETS/CBP industry sector ratio and sectoral employment share (NETS), 1998-2009

		Manufacturing	Transport	Wholesale	Retail	Finance & Insurance	Real Estate	Professional service	Information
Los Angeles	Mean ratio	1.09	0.81	1.02	0.98	1.04	1.48	1.01	1.07
	Emp share (NETS)	10.56	2.17	5.16	8.06	3.84	2.48	6.21	3.05
San Francisco	Mean ratio	1.34[a]	0.89[b]	0.88	0.96	1.12	1.42	1.15	0.98
	Emp share (NETS)	9.35	1.76	3.86	7.60	4.32	2.12	8.90	3.64
Sacramento	Mean ratio	1.1	0.85[c]	0.98[d]	0.91	0.86	1.41	1.36	0.91
	Emp share (NETS)	5.64	1.78	3.46	9.59	4.28	2.47	6.25	2.58
San Diego	Mean ratio	1.37	1.05	0.9	0.87	1.03	1.45	1.08	1
	Emp share (NETS)	9.92	1.47	3.62	8.34	3.69	3.01	7.61	2.32
Average by all CA Counties	Mean ratio	1.08	1.04	1.15	0.91	1.06	1.61	1.38	1.08
	Emp share (NETS)	13.08	2.71	6.06	11.36	5.23	3.20	9.15	4.00
	Number of Counties	43	41	47	54	49	38	49	41

a. The Solano county and the Sonoma county are dropped for calculating the employment share of manufacturing sector in the region
b. The San Mateo county is dropped for calculating the employment share of transportation sector in the region
c. The Yuba county is dropped for calculating the employment share of transportation sector in the region
d. The Yuba county is dropped for calculating the employment share of wholesale trade sector in the region

In order to compare year to year variation, we plot the NETS/CBP industrial employment share ratios of selected 2-digit NAICS sectors from 1998 to 2009 for the four regions. See Figures 2-4 through 2-11.We also include the statewide county average for comparison. The level and changes of industrial differences between the two datasets for the four regions are generally consistent with that for the state average and with one another. There is no metro area that is always a relative outlier. Also, temporal trends vary across sector. For the manufacturing sector, while the state average industrial difference index hovers around 1, the index for the SF and SD regions are generally greater, with the employment share in the NETS dataset over-counted by about 30% to 50% (Fig. 2-4). Employment shares in the transportation sector are relatively undercounted in the NETS dataset for the LA and Sacramento regions, while the employment share differences between the two datasets is less obvious for the SF and SD regions for most years (Fig. 2-5). The values of industrial ratios for the sector are less volatile when data are aggregated at the CSA level, except for the SD region.

For the wholesale trade sector, the value of industrial differences index for the four regions also shows less fluctuation than average at the county level for the whole state. The employment shares in the two datasets are also relatively comparable for the four regions for most of the years (Fig. 2-6). However, the sectoral employment share is relatively undercounted in the NETS dataset for all the four regions after 2008, which differs greatly from the average county-level industrial ratio. The values of industrial ratios for the retail sector are less volatile when data are aggregated at the CSA level (Fig. 2-7). With the exception of the SD region, the retail employment share in the two datasets is relatively comparable using the 10% criterion. The real estate sector is the only one that is always over-counted by more than 20% in the NETS data set (Fig. 2-9), which as we have discussed may be attributable to the existence of relatively large shares of small-sized establishments in the sector. Similarly, there is a large gap in the employment share in the professional service sector, with the values of industrial ratio fluctuating over time (Fig. 2-10). Using the 10% criterion, we find that the share between the two datasets is comparable for the LA region for most years, for the SF region after 2003 and for the SD region after 2001.For the information sector, the values of the ratio are volatile for all four regions. LA is over-counted in the early years and under-counted in the later year; the pattern is just the reverse for Sacramento. SF and SD track better with the statewide county average (Fig. 2-11).

In sum, considering the sector ratio for different regions and years, we did not find any sector showing comparability in employment share between the two datasets for all the four regions throughout the time periods. Except for wholesale trade, retail trade and the finance and insurance sector, the values of industry sector ratios are volatile over the period for most of the sectors. We do not, however, find any link between the variations of industrial ratios over the years and share size of the sectors. For example, the professional service sector ranks third in terms of sectoral employment shares among the sectors studied, but has the biggest spread in terms of the value of industrial ratio over the years.

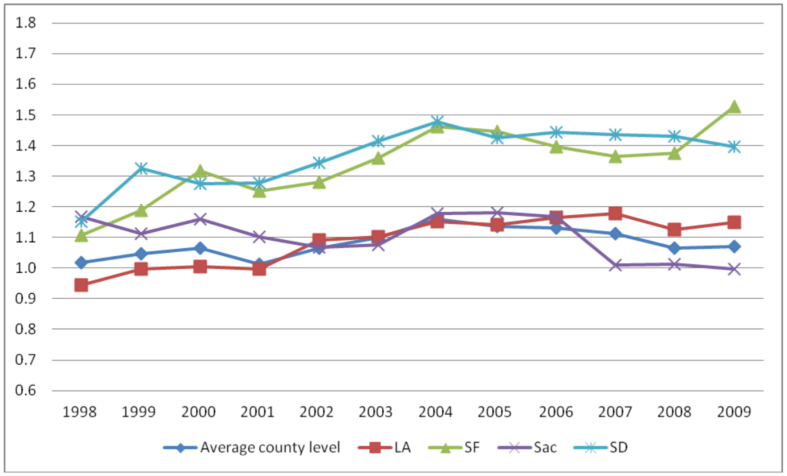


Figure 2-4 Ratio of employment share for the manufacturing sector (1998-2009)

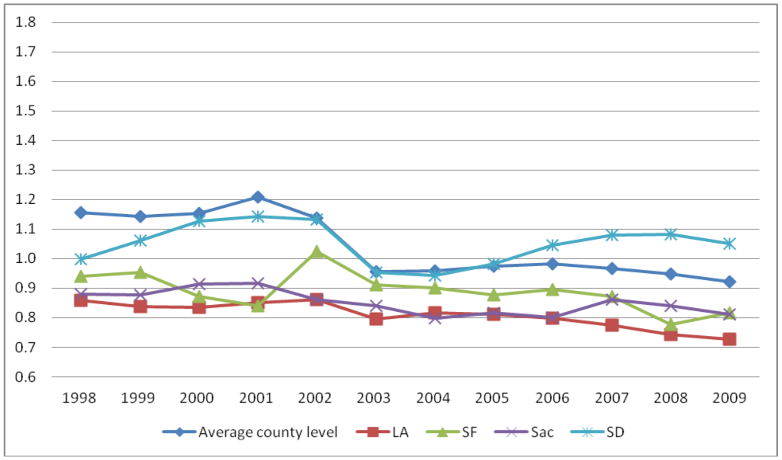


Figure 2-5 Ratio of employment share for the transportation sector (1998-2009)

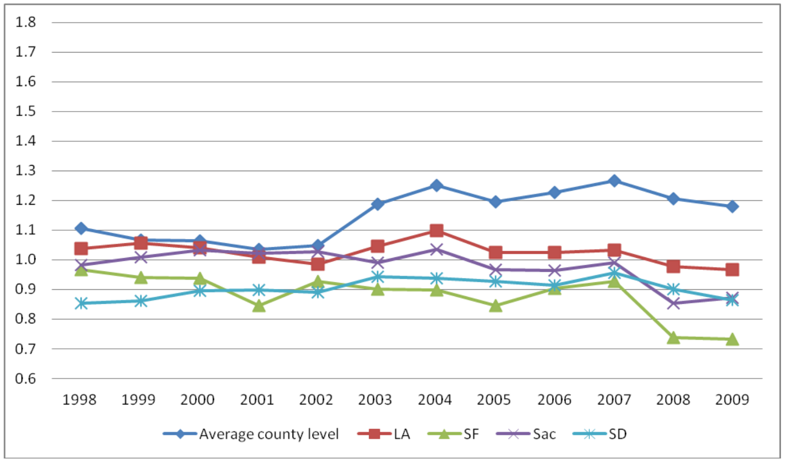


Figure 2-6 Ratio of employment share for the wholesale Trade sector (1998-2009)

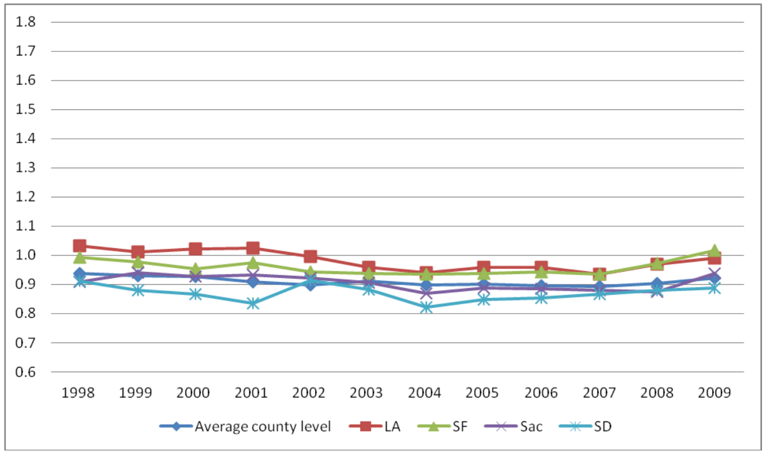


Figure 2-7 Ratio of employment share for the retail trade sector (1998-2009)

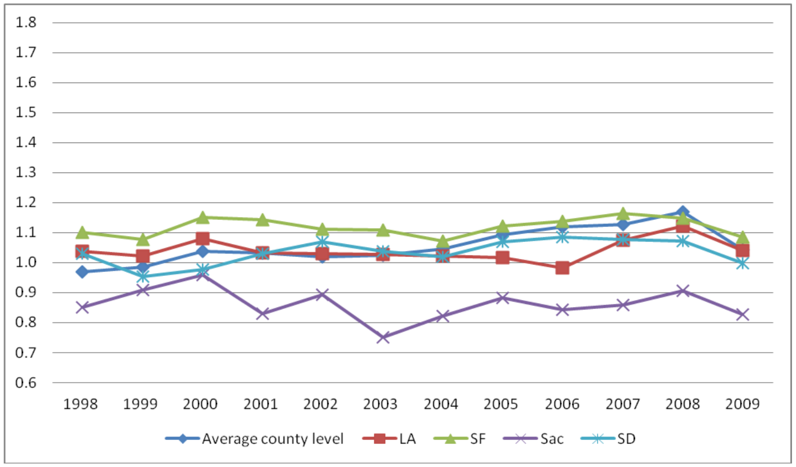


Figure 2-8 Ratio of employment share for the finance and insurance sector (1998-2009)

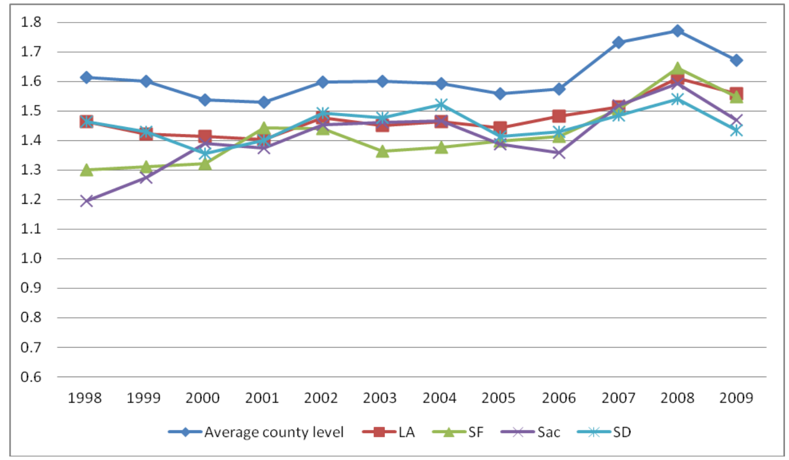


Figure 2-9 Ratio of employment share for the real estate sector (1998-2009)

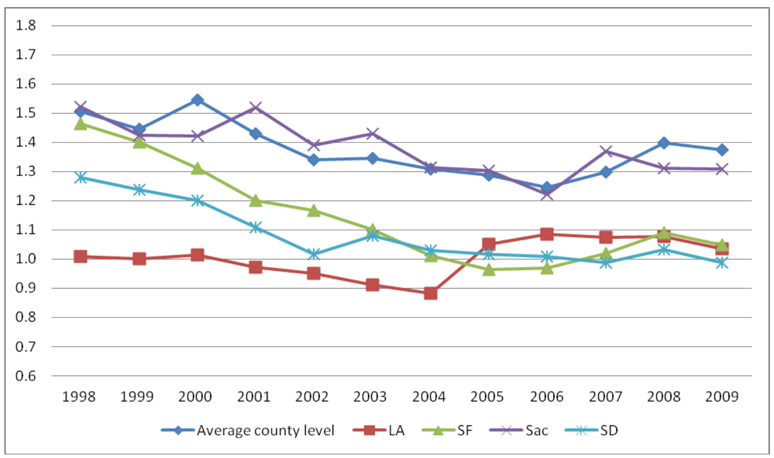


Figure 2-10 Ratio of employment share for the professional service sector (1998-2009)

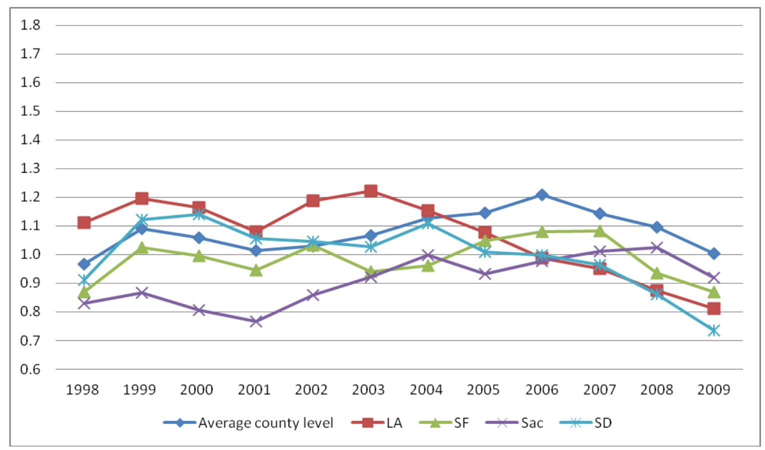


Figure 2-11 Ratio of employment share for the information sector (1998-2009)

2.3.4 Difference by employment size category

To compare the difference in establishment counts between the two datasets by different employment size classes, we also computed the percentage difference between NETS data and CBP data in terms of total establishment counts as well as NETS establishment counts in different size categories. The annual mean value of county level percentage difference in establishment counts of different size category in the two datasets is available in Table A-2-1 in Appendix 2. The percentage difference in establishment counts for the four study regions is available in Tables A-2-2 to A-2-5 in Appendix 2. Figure 2-12 plots the trends of aggregate establishment differences for the four regions as well as the annual mean values of the county-level total employment differences.

Figure 2-12 shows that the trend of percentage employment difference at the CSA level increases over time. NETS is always higher. Again, one possible explanation is NETS capturing more small establishments over time (recall that we have removed self-employment from NETS to make the data sets comparable). Although the pattern is the same for the four metro areas, differences for the LA region are consistently greater -- by about 20 percentage points. In order to check whether these differences are due to the NETS data doing a better job of capturing small firms, we also calculated the percentage difference in establishment counts of different size category between the two datasets, the results of which are shown in Table 2-10 for all California counties. The same data are plotted in Figure 2-13.

There are several observations to be drawn from Table 2-10 and Figure 2-13. First, the percent differences for all establishments follow the general trend of our four regions, but differences are greater in the 1990s and smaller in the 2000s. Second, differences increase in the smallest size category, as would be expected if NETS does an increasingly better job of capturing small firms. Third, the patterns for other size categories are varied: for the next smallest firms (10-49 employees), there is a consistent over-count in the range of 20 - 30%; for the larger firms, there is a downward trend in differences. Perhaps most disturbing is the data for the largest firms (1000 or more employees). We would expect that whatever the method of tracking establishments, numbers for the largest firms would be consistent. This is not the case until the last years of the data set. Even in the last years, some size categories still have large differences (e.g. around 30% for firms with 250 – 499 employees). Given that the numbers in Table 2-10 are averages of 50 or more counties, these differences are indeed large, and would be larger for any given county

Table 2-10 Percentage difference in establishment counts between NETS data and CBP data by employment size class (1993-2009)

Year	All		1-9		10-49		50-99		100-249		250-499		500-999		>= 1000	
	N	Diff (%)	N	Diff (%)	N	Diff (%)	N	Diff (%)	N	Diff (%)	N	Diff (%)	N	Diff (%)	N	Diff (%)
1993	58	52.1	58	60.9	58	20.1	57	66.5	55	42.6	51	35.6	38	52.3	32	75.3
1994	58	51.7	58	58.7	58	26.1	56	66.9	57	51.3	49	62.5	38	62.1	33	52.8
1995	58	52.3	58	59.7	58	26.7	57	70.2	56	47.0	49	65.7	42	57.9	33	38.8
1996	58	51.9	58	59.2	58	26.9	57	64.0	57	35.1	52	41.8	41	55.6	33	39.6
1997	58	49.3	58	56.2	58	26.6	56	53.0	57	37.8	51	42.3	41	64.5	34	48.4
1998	58	48.1	58	54.9	58	26.6	57	47.4	57	41.4	52	35.5	42	62.7	34	52.5
1999	58	45.9	58	52.1	58	27.2	57	50.3	57	39.9	49	45.1	42	52.9	34	40.3
2000	58	50.2	58	57.5	58	28.1	57	54.3	57	43.5	52	50.8	39	57.9	35	62.8
2001	58	71.5	58	86.4	58	26.7	58	47.1	56	59.5	54	63.7	41	47.7	36	46.6
2002	58	65.8	58	79.5	58	24.5	56	55.0	56	51.5	51	31.4	39	51.1	36	19.8
2003	58	67.1	58	81.8	58	23.1	58	60.4	56	38.6	52	26.0	42	36.4	37	15.1
2004	58	69.1	58	85.0	58	22.6	56	45.5	56	33.7	53	18.1	42	28.6	38	12.8
2005	58	72.7	58	88.7	58	25.6	56	55.2	56	24.6	53	15.0	42	30.2	36	20.3
2006	58	75.0	58	93.1	58	23.5	56	59.6	56	23.0	52	29.9	40	19.5	38	9.6
2007	58	82.0	58	102.6	58	21.7	56	57.4	56	18.2	51	26.1	43	11.2	38	2.6
2008	58	104.6	58	132.0	58	27.1	57	51.7	56	28.5	53	29.4	42	0.2	40	-2.0
2009	58	98.7	58	122.5	58	29.4	57	57.9	56	38.0	53	32.9	41	6.9	40	3.6

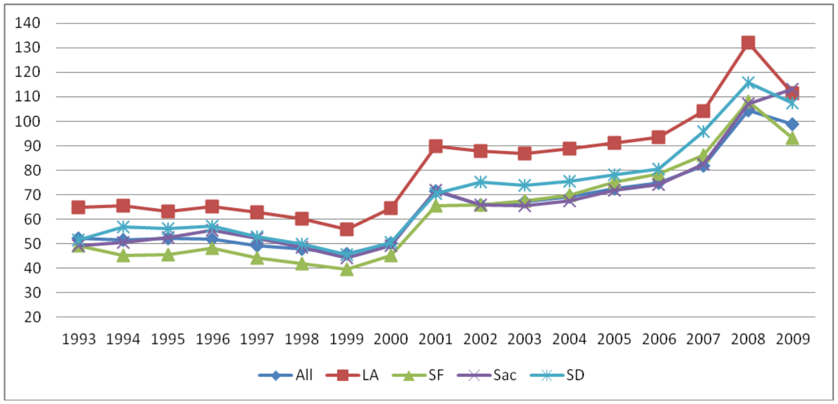


Figure 2-12 Percentage difference in establishment counts

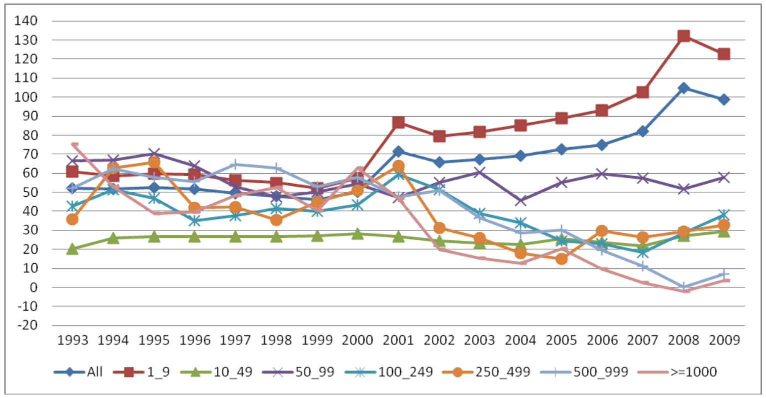


Figure 2-13 Percentage difference in establishment counts by employment size category

2.4 Summary

The results of our analysis are summarized in the following three points:

1) The NETS dataset over-counts about 20% to 35% employment and from 50% to more than 100% establishments relative to the CBP dataset. In general, the over-count of employment declines over time, and the over-count of establishments increases over time. These trends are generally consistent across the four regions and the statewide county average.

2) The industry sector data, both in the early period using SIC and the later period using NAICS, shows large and varying differences across sectors, time periods, and metro areas. The employment shares in the services sector (SIC1) and the real estate and professional service sectors (NAICS2) are least comparable between the two datasets, which might be explained by the large share of small-sized firms in these sectors that might have been missed in the CBP data collection. However, differences are not consistent, and aside from there being more small firms in NETS, have no obvious explanations.

3) Comparing the differences in establishment counts by different employment size classes, we found that the NETS dataset is extremely skewed towards capturing the small-sized firms (less than 10 employees). However, differences exist across all size categories and vary over the years. Thus the differences are not simply due to NETS doing a better job of counting small firms. It is important to note that even the counts of very large firms (over 1,000 employees) are not consistent. It is difficult to understand how the largest firms would not be captured by both data sources.

In sum, the NETS and CBP datasets are not comparable with each other at the county or CSA level, for a given year or across time. Of particular concern is the absence of any systematic differences beyond there being more small firms in NETS. If we could identify such differences, it might be possible to adjust the data to make it more comparable. Unfortunately we have no way of determining which data set is more "correct". We know that data collection techniques continue to change for NETS, but we don't know precisely how these changes should affect establishment counts or employment counts. Differences in industry sector are likely due to the difference in how the data are collected, but we have no way to determine whether self-reporting is more or less reliable than codes assigned by government experts. Clearly the differences between NETS and other sources of data calls into question the reliability of employment data more generally, especially at the level of small spatial units.

Because of these data inconsistencies, we did not conduct any analysis of industry composition, or any analysis of establishments. We proceed with identification and analysis of employment centers, but only for 1995, 2000, and 2005, the range of years that appear to be most stable in the NETS data. A far more comprehensive analysis of employment data sources should be conducted; such an analysis is beyond the scope of this research.

Chapter 3 Identifying employment centers

This chapter presents our method for identifying employment centers, summary results on the employment centers for each case study region, and a series of formal tests of polycentricity to determine whether the identified centers influence the employment density distribution. The chapter begins with a discussion of spatial units and our rationale for translating the establishment data to a uniform hexagonal grid.

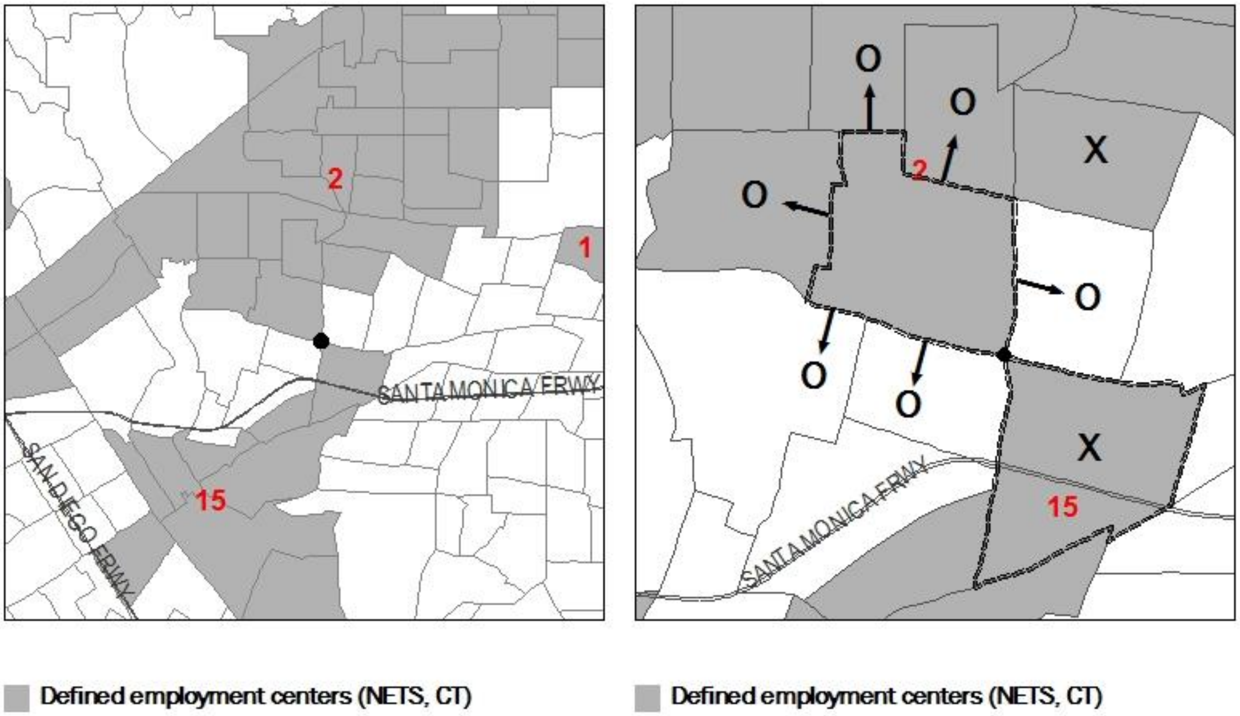
3.1 Developing the geographic and temporal database for center identification

3.1.1 Justification for the use of hexagon grid

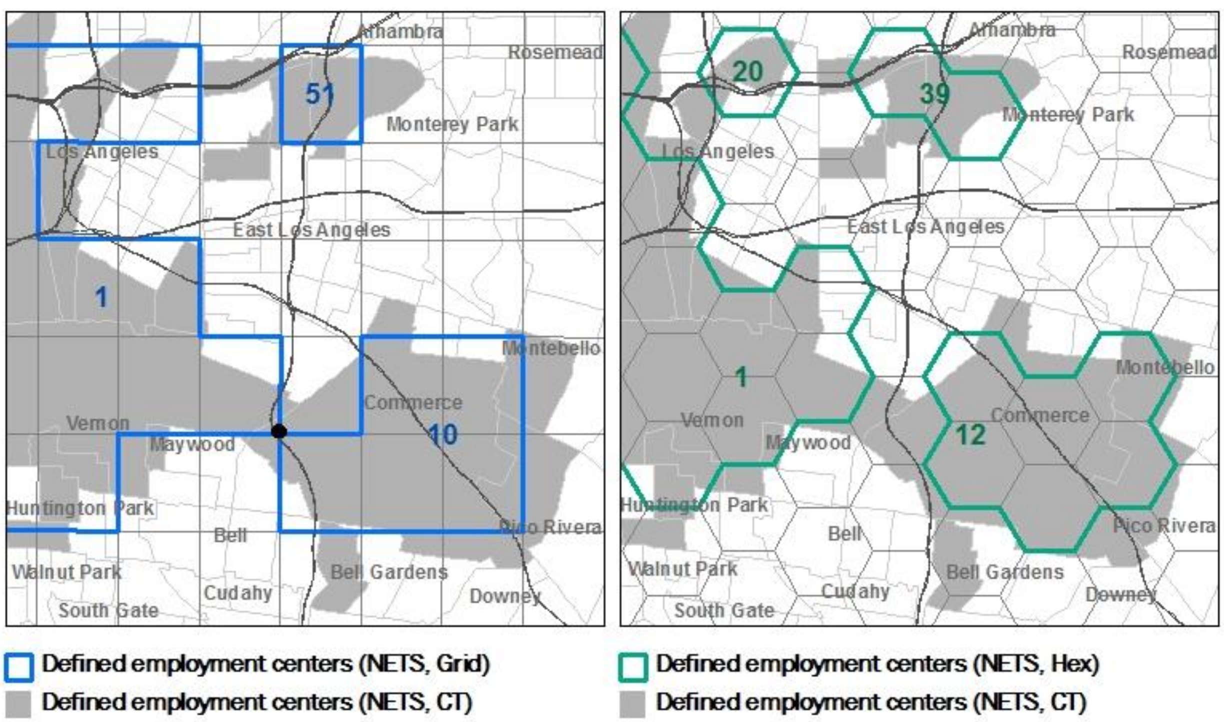
In this section, we justify the use of the hexagon grid system as the spatial unit of analysis. In our former work we used census tracts as the spatial unit of analysis. Because the NETS dataset is based on the address of each establishment, the spatial unit can be defined by the researchers. Census tracts have the disadvantage of being different sizes, and size is correlated with density. The unit of observation is thus related to one of the main variables we want to explain, employment density. We therefore chose a uniform grid of spatial units. Because identifying centers involves aggregating adjacent spatial units based on various criteria, we wanted to avoid point-based aggregation, as would happen with a square or fishnet grid. We therefore chose a hexagon grid, where adjacency requires a common side. The hexagons have an area of one square mile.

We illustrate the benefits of using a hexagon grid with a few simple examples. Map 3-1 (left) shows centers identified by the Giuliano and Small (1991) method, marked as 2 and 15, in the geography of census tracts. The variation in size and configuration of the tracts is evident. In this case, tracts meeting the density criterion share a common point. Map 3-1 (right) shows this common point in greater detail. Six arrows emanating from the center 2 census tract with black boundary show the 'O'-labeled census tracts which share common boundaries. In the meantime, two 'X'-labeled census tracts only share common points with the center 2 census tract. Whether to consider 2 and 15 as one center depends on the arbitrary choice of whether a point constitutes a common boundary.

Map 3-2 (left) shows another group of centers. The grey shading shows the center boundaries as defined by census tracts. The area is overlaid with a mile square fishnet grid, and the blue outlines show the grid units that meet the Giuliano and Small criteria. It can be seen that the same problem of common points appears. Map 3-2 (right) shows the same area, this time with a mile square hexagon grid. Green outlines show the grid units that meet the Giuliano and Small criteria. In this case, each identified center has clear boundaries.



Map 3-1 Ambiguity in defining adjoining census tracts



Map 3-2 Defined employment centers based on fishnet grid (left) and hexagon grid (right)

3.2 Method for identifying centers, application, and results

3.2.1 Methodology: Identifying employment centers, Justification of 95%-10 methods

Theoretically, centers refer to areas with concentration of employment large enough to exert a potential influence on employment and population distributions. Previous studies have made efforts to identify employment centers in many different ways. As noted in Chapter 1, there is no generally agreed upon method for defining or identifying employment centers, and different definitions lead to different numbers and sizes of centers. There are several challenges in center identification. Even within one metropolitan area, there are various kinds of employment centers ranging from central business district (downtown) to suburban office parks. Some areas might be influenced by several surrounding centers, so the theoretical ideal of using influence on the surrounding employment and population is quite difficult to carry out.

As described in Chapter 1, there are three general methods for identifying employment centers: (1) methods based on minimum size and density (e.g., Giuliano and Small 1991), (2) estimation of density gradients to identify potential centers, and (3) various two-step methods using locally weighted regression (LWR) to smooth the density surface and then identify centers. Each of these approaches has advantages and disadvantages. In all cases, they involve some type of arbitrary decision, as for example the criterion of centers having a minimum total number of jobs (method 1), or the criterion of what constitutes a sufficient difference in the smoothed density surface (method 3).

Given that no method has emerged as theoretically more appropriate, it seems the method is best chosen based on the purpose of the research. In this case we want to compare centers across different metropolitan areas, and explain growth patterns over time. In our previous work (Giuliano et al 2010), we found that the centers identified via the Redfearn method, which is based on relative density, were quite unstable over time. Small changes in employment density could generate quite different sets of centers. We therefore chose to use the Giuliano and Small centers for longitudinal comparisons. We take a similar approach here, but adjust the method to allow for differences in the overall size and density across metropolitan areas.

Pan (2003) pointed out that the minimum cut-off of density has statistical meaning under the assumption that employment density by tracts in a region is distributed normally. We tested whether employment density (based on our hexagonal grid) is normally distributed. We find that neither employment density nor natural log of employment density has a normal distribution (test results not shown). Thus we have no basis for using standard deviation units as cutoff values. Instead we use percentiles of the distribution: 95% and 99%. The 95th percentile employment densities are given for each metro area in Table 3-1. As a result, each metro area has different cutoff density as shown in Table 3-1, which reflects its geographic size and magnitude of total employment. Note that the 95 percent cutoff density for the Los Angeles area is very close to the 10 jobs per acre Giuliano and Small criterion. Table 3-1 shows that Los Angeles has the highest 95th percentile density, but LA, SF and SD numbers are quite comparable. The outlier is Sacramento, with a 95th percentile value about 1/3 of the others. Note also that the 95th percentile is relatively stable, with a general upward trend particularly 1995 – 2000, reflecting more job growth during this period.

Table 3-1 95% density criterion for four metro areas

95% employment density	1995	2000	2005
Los Angeles	9.6	10.2	10.5
San Francisco	7.1	9.3	9.8
San Diego	6.8	8.3	8.3
Sacramento	2.4	2.8	2.9

3.2.2 Methodology: Identifying employment centers, Justification of 1995-2000-2005

We want to conduct two types of comparisons in order to better understand the structure and dynamics of employment centers: across metropolitan areas, and across time. Our NETS data is limited to California; therefore we use the four largest metro areas in our research: Los Angeles, San Francisco, San Diego, and Sacramento. With regard to time, we want to capture significant changes over time. The NETS database has annual data from 1993 to 2009. What is the best time interval for capturing change? We expect that annual data is too fine. Absent an economic shock, economic activity tends to be rather stable. We also want to avoid major shocks, such as the Great Recession and financial collapse of 2007. We conducted Kolmogorov-Smirnov tests (K-S tests) on the probability distribution of employment across the hexagons within each metropolitan area for each year, starting at 1993. We tested whether the employment distribution by hexagons in a given year is different from that of the previous year. The K-S test is a nonparametric test, which has no assumption about the distribution of data. Based on K-S results for all metro areas, we decide to use 1995, 2000 and 2005 as our analysis years. Employment distribution by hexagons between year 1993 and year 2008 are likely to fall into three categories (approximately five years), and 1995, 2000 and 2005 are representative years of each category for all metro areas. We chose 2005 to avoid the 2007 bubble and recession. Table 3-2 provides the K-S test results for the San Francisco area to illustrate

We used the 95th and 99th percentile employment densities to identify employment centers in the four metro areas and three time periods. Results are summarized in Table 3-3 below. Table 3-3 shows employment densities at the 95th and 99th percentile in four metropolitan areas in three year categories. The employment density varies widely across metropolitan areas, having rather temporally similar figures within each area. The identified number of employment centers in each area also shows temporally similar figures. The Los Angeles metropolitan area shows the highest employment density and the most employment centers, followed by San Francisco, San Diego, and Sacramento. The number of identified employment center decreases as the cut-off density increases, but it stays similar within an area. The number of centers increases between 1995 and 2000, and then it stays almost consistent between 2000 and 2005. 1995 and 2005 centers are shown in Maps 3-3 through 3-10. A descriptive analysis of the centers is presented in Chapter 4.

Table 3-2 K-S test for San Francisco CSA

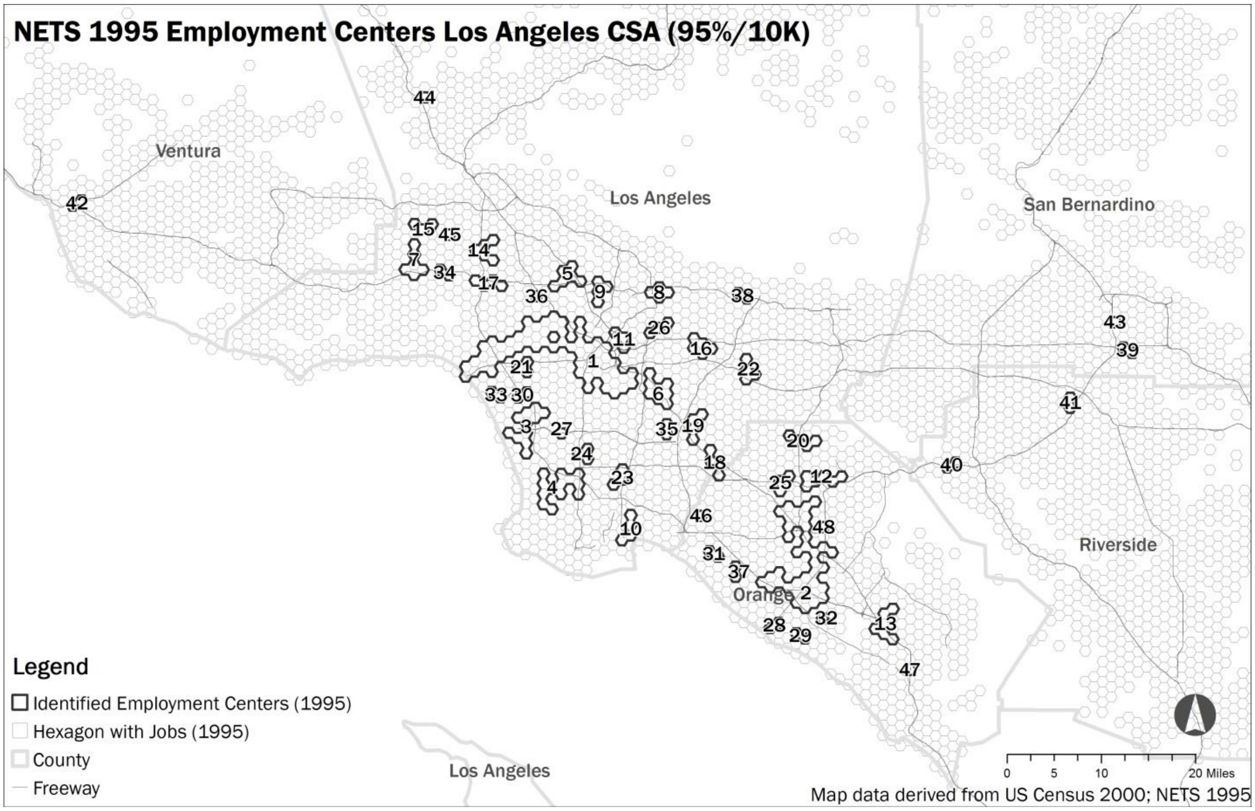
Start year	Compare year	D	P-Value	Corrected P-Value
1993	1994	0.0118	0.919	0.914
	1995	0.0116	0.929	0.926
	1996	0.0100	0.981	0.979
	1997	0.0191	0.400	0.389
	1998	0.0241	0.156	0.149
	1999	***0.0311***	***0.028***	***0.026***

Start year	Compare year	D	P-Value	Corrected P-Value
1999	2000	0.0095	0.988	0.987
	2001	0.0279	0.065	0.061
	2002	0.0095	0.988	0.987
	2003	***0.035***	***0.009***	***0.009***

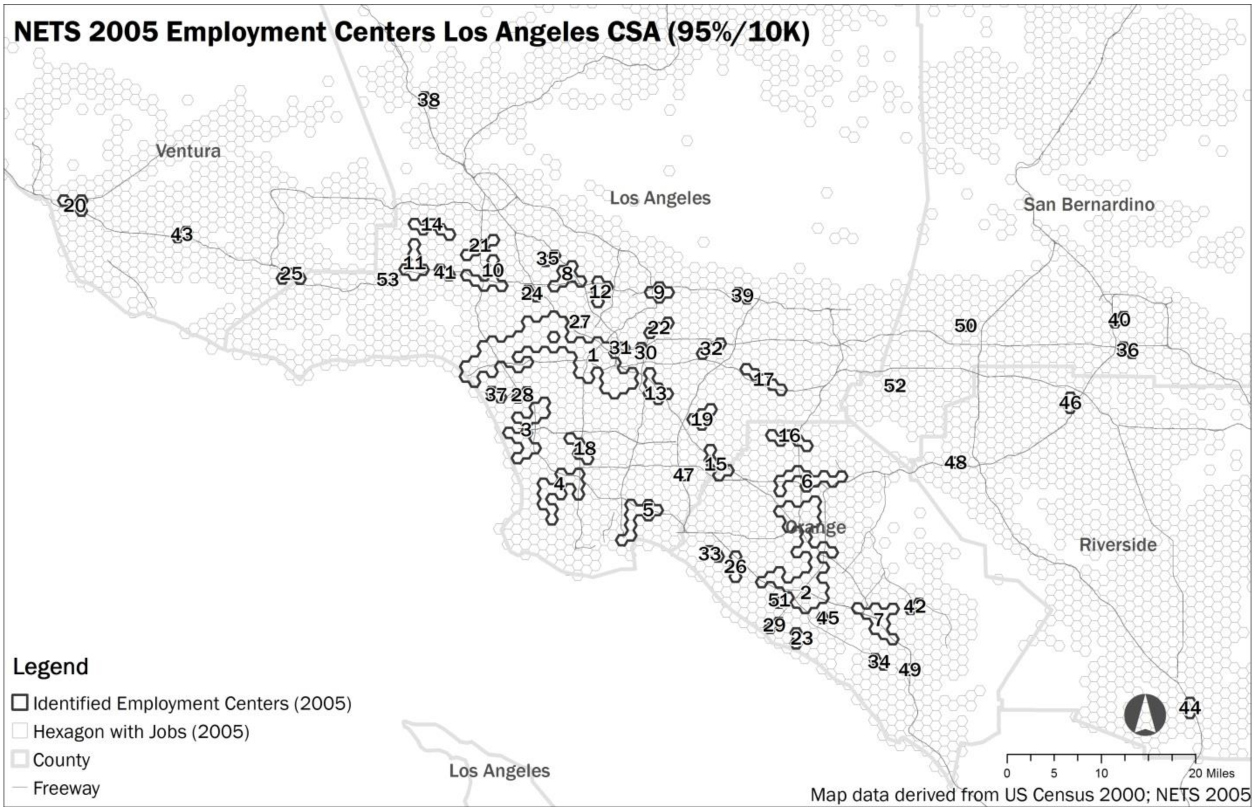
Start year	Compare year	D	P-Value	Corrected P-Value
2003	2004	0.0098	0.985	0.984
	2005	0.0204	0.317	0.307
	2006	0.0234	0.180	0.173
	2007	***0.0311***	***0.028***	***0.026***

Table 3-3 Number of centers by year, metro area

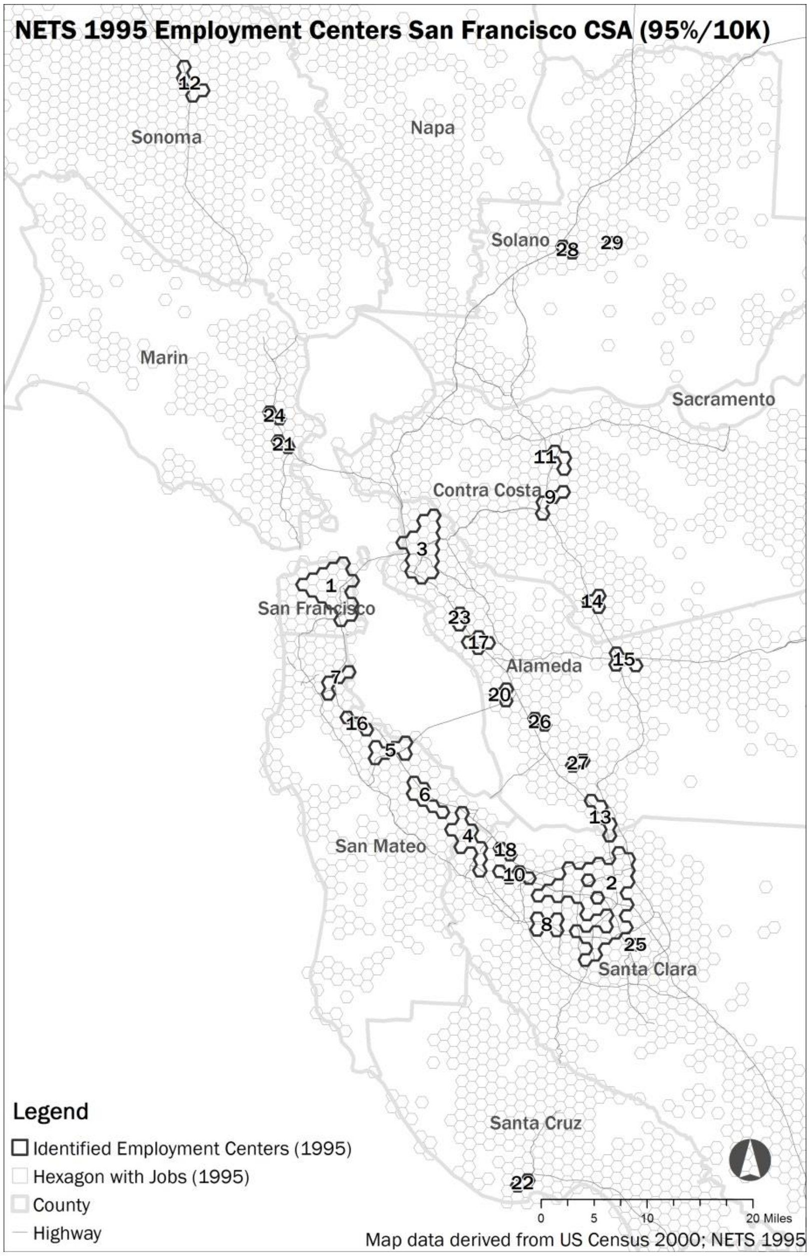
	Los Angeles			San Francisco			San Diego			Sacramento		
	1995	2000	2005	1995	2000	2005	1995	2000	2005	1995	2000	2005
95^{th} Pct	9.6	10.2	10.5	7.1	9.3	9.8	6.8	8.3	8.3	2.4	2.8	2.9
N centers	48	53	53	28	30	30	7	13	12	7	7	8
99^{th} Pct	22.9	23.9	23.0	18.6	20.5	19.4	14.4	18.1	17.2	7.5	8.4	8.2
N centers	15	18	18	6	9	10	3	3	3	2	2	2



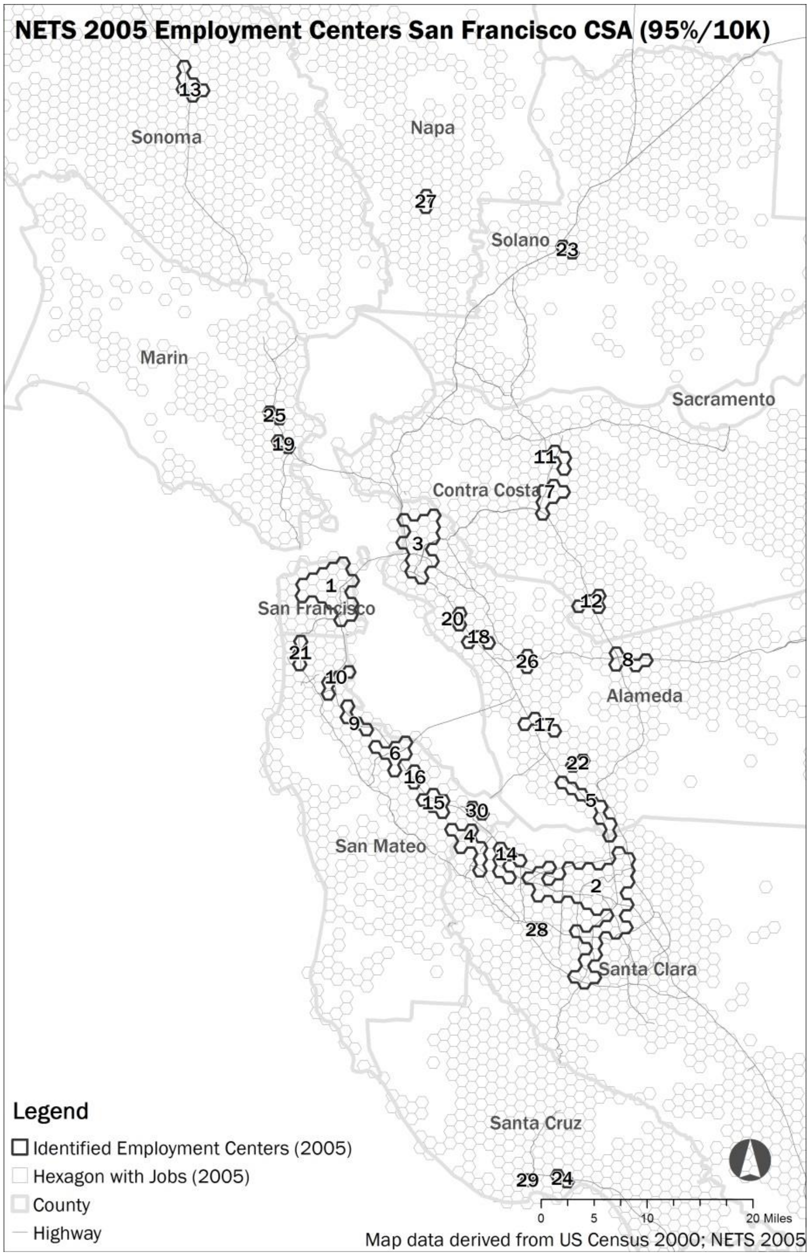
Map 3-3 NETS 1995 employment centers in Los Angeles CSA



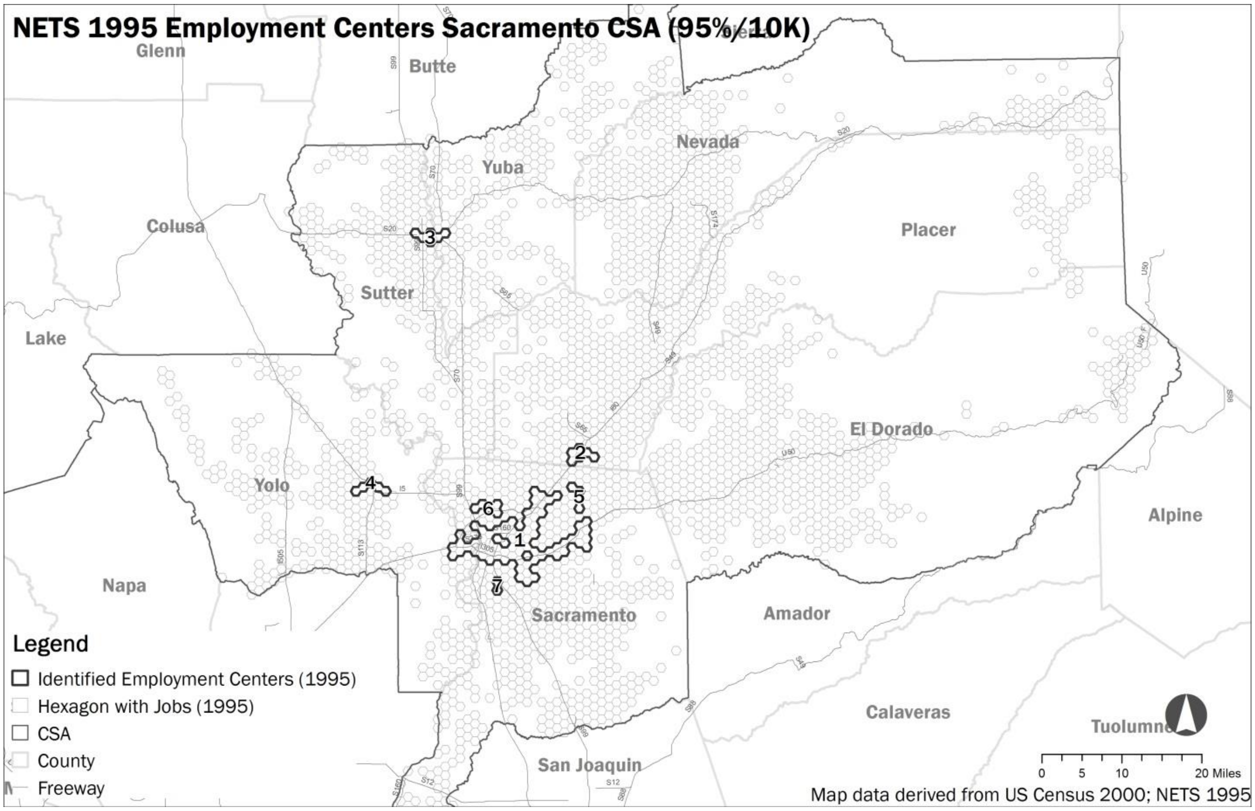
Map 3-4 NETS 2005 employment centers in Los Angeles CSA



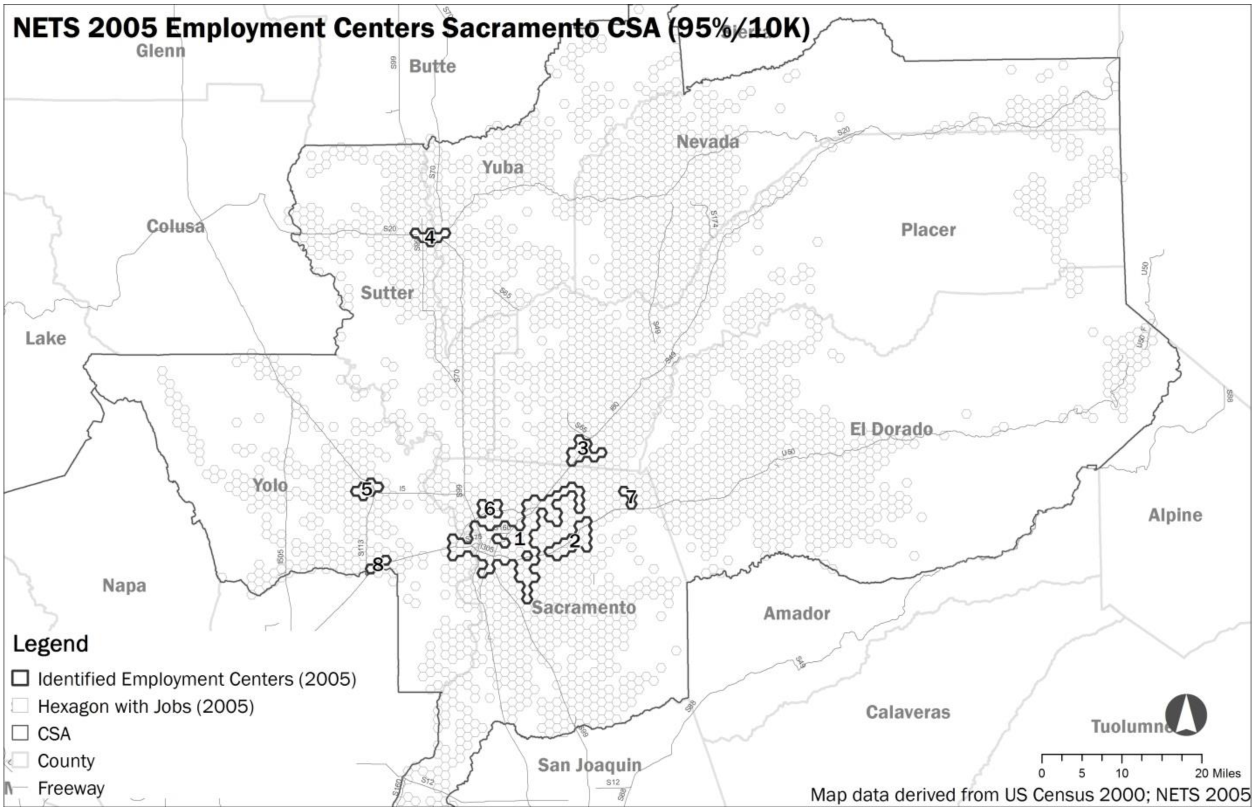
Map 3-5 NETS 1995 employment centers in San Francisco CSA



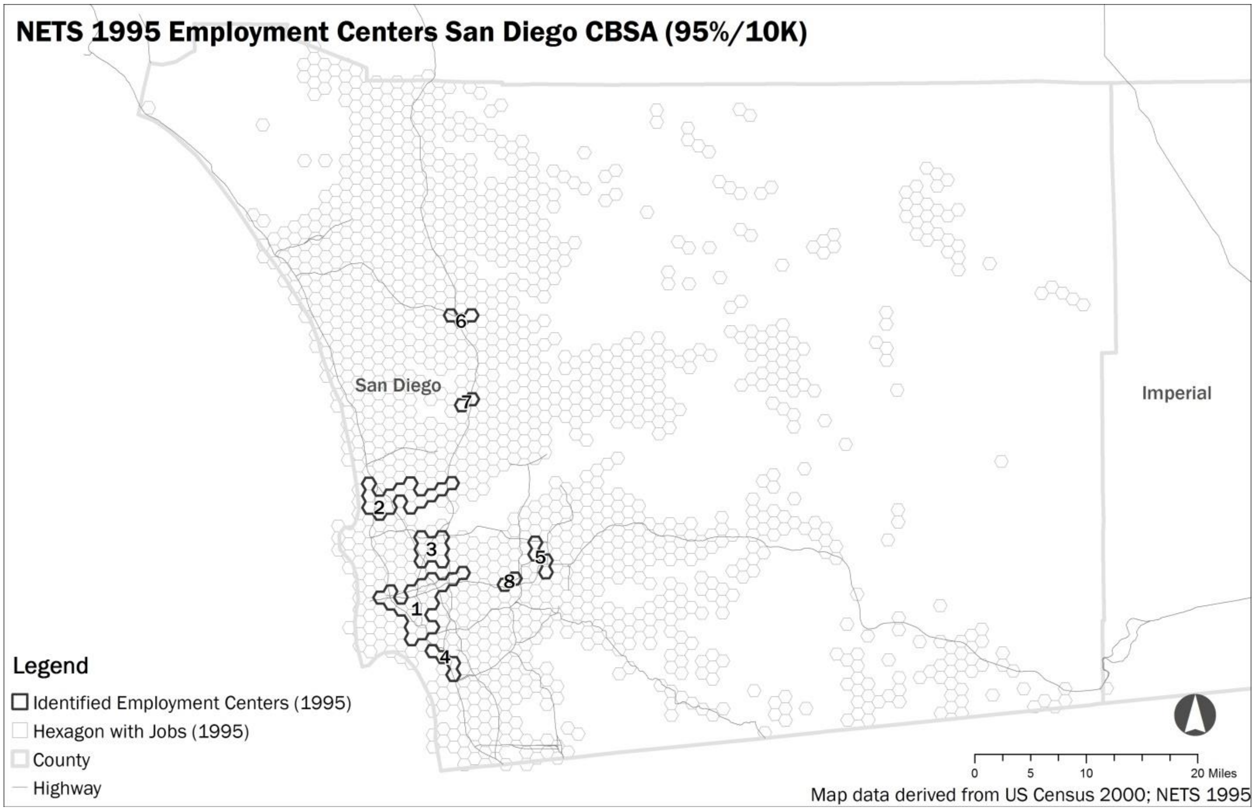
Map 3-6 NETS 2005 employment centers in San Francisco CSA



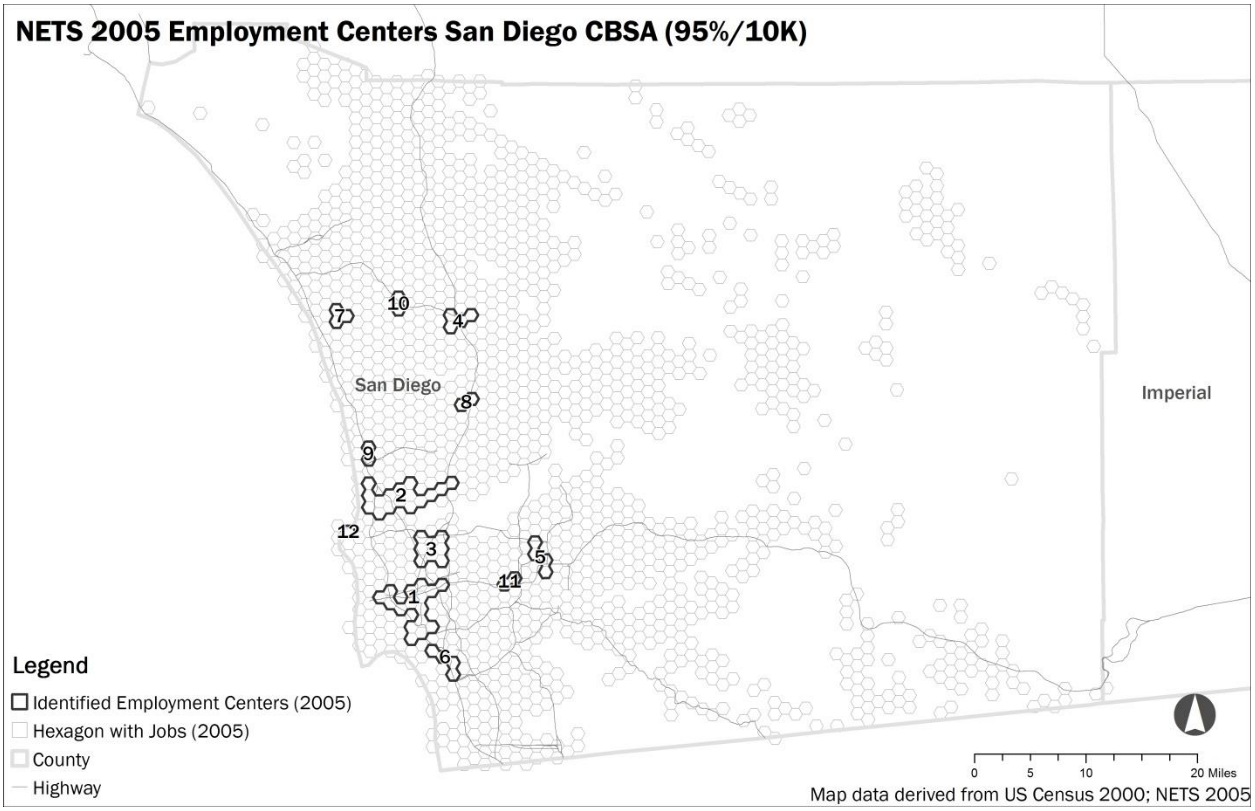
Map 3-7 NETS 1995 employment centers in Sacramento CSA



Map 3-8 NETS 2005 employment centers in Sacramento CSA



Map 3-9 NETS 1995 employment centers in San Diego CSA



Map 3-10 NETS 2005 employment centers in San Diego CSA

3.3 Formal tests for polycentricity

A final question to be explored in this chapter is whether our four metro areas are indeed polycentric. That is, do the centers we have identified contribute to explaining the employment density distribution? We conduct formal tests by estimating polycentric density functions.

3.3.1 The polycentric density function

Derived from the monocentric model, the negative exponential function describes the theoretical distribution of population or employment,

$$D(r) = D_0 e^{-\beta r + u} \qquad (1)$$

where D is the employment density at distance r from the CBD, r is distance to the CBD, D_0 is the employment density at the center, β is the density gradient , and u is the error term. Empirical estimation of the model usually takes the logarithm form of the equation.

The polycentric model is a "natural extension of the monocentric model" that estimates employment/population density as a function of access to both CBD as well as other employment centers (Small and Song 1994). In empirical estimations, the generalization of the monocentric model takes the following form:

$$\ln D = \alpha_0 + \beta d_{CBD} + \sum_i \beta_i d_{sub} + u \qquad (2)$$

where d_{CBD} and d_{sub} measures distance to the center and other subcenters respectively and u is the error term. Based on theory, the coefficients of all the distance measures should be negative. However, correlation between the distances to different centers can generate a multicollinearity problem in the estimation that may render some coefficients insignificant or of the wrong sign (Heikkila, Gordon et al. 1989). McDonald and Prather (1994) deal with this problem by using an inverse distance function for subcenters,

$$\ln D = \alpha_0 + \beta d_{CBD} + \sum_i \alpha_i d_{sub}^{-1} + u \qquad (3)$$

where d^{-1} is the inverse of distance to the nearest subcenter. The basic assumption behind this density function is that while the effects of the CBD extends to the whole metropolitan area, the effects of suburban employment centers are limited to a relatively short distance (McDonald and Prather 1994). Based on this assumption, the density gradient of suburban centers is a declining function of distance,

$$\frac{\partial \ln D}{\partial s} = -\frac{\alpha}{s^2} \qquad (4)$$

where s represents distance to subcenters (McDonald and Prather 1994). The larger value of coefficients for the inverse of distance, the steeper the density gradient.

3.3.2 Estimation results

3.3.2.1 Los Angeles

In this study, we followed McDonald and Prather's (1994) method and take the inverse function to estimate the employment density distribution of the four regions for three cross-sectional years (1995, 2000, 2005). We would confirm a polycentric structure of employment distribution if the coefficient for distance to CBD is significant and negative and the coefficients for some of the inverse of distance to subcenters are significant and positive. We define the employment centers based on the 99% density cutoff and 20,000 total employment thresholds. The number of centers for the LA region is 15 in 1995 and 18 in 2000 and 2005. As indicated by the figures of LA centers, the relative locations of centers are very stable over time.

To check if the inverse of distance eliminates the multicollinearity problem, we first estimate the pairwise correlation matrix for the variables. We find that they are highly correlated, likely because of their spatial proximity. For example, in the year 2000, Centers 2 and 14 are collapsed from the largest center of the LA downtown-Santa Monica corridor defined by the 95%-10,000 cutoff, while centers 6, 8 and 9 are linked by I-210 and centers 3 and 10 by I-5.

Because of the spatial proximity between centers, each center may not exert an independent influence on the employment distribution. We therefore apply stepwise regression by adding the inverse of distance to subcenters one by one based on the rank of their peak hexagon employment density, and drop the influence of a lower rank center if the coefficient of its density gradient is insignificant because of its high correlation with a higher rank center. We use all hexagons with at least one job in the estimation, including those within centers. The different number of hexagon observations is because we have more spatial units with jobs over the study period. The estimation results for the "best" fitted models of the three years are presented in Table 3-4

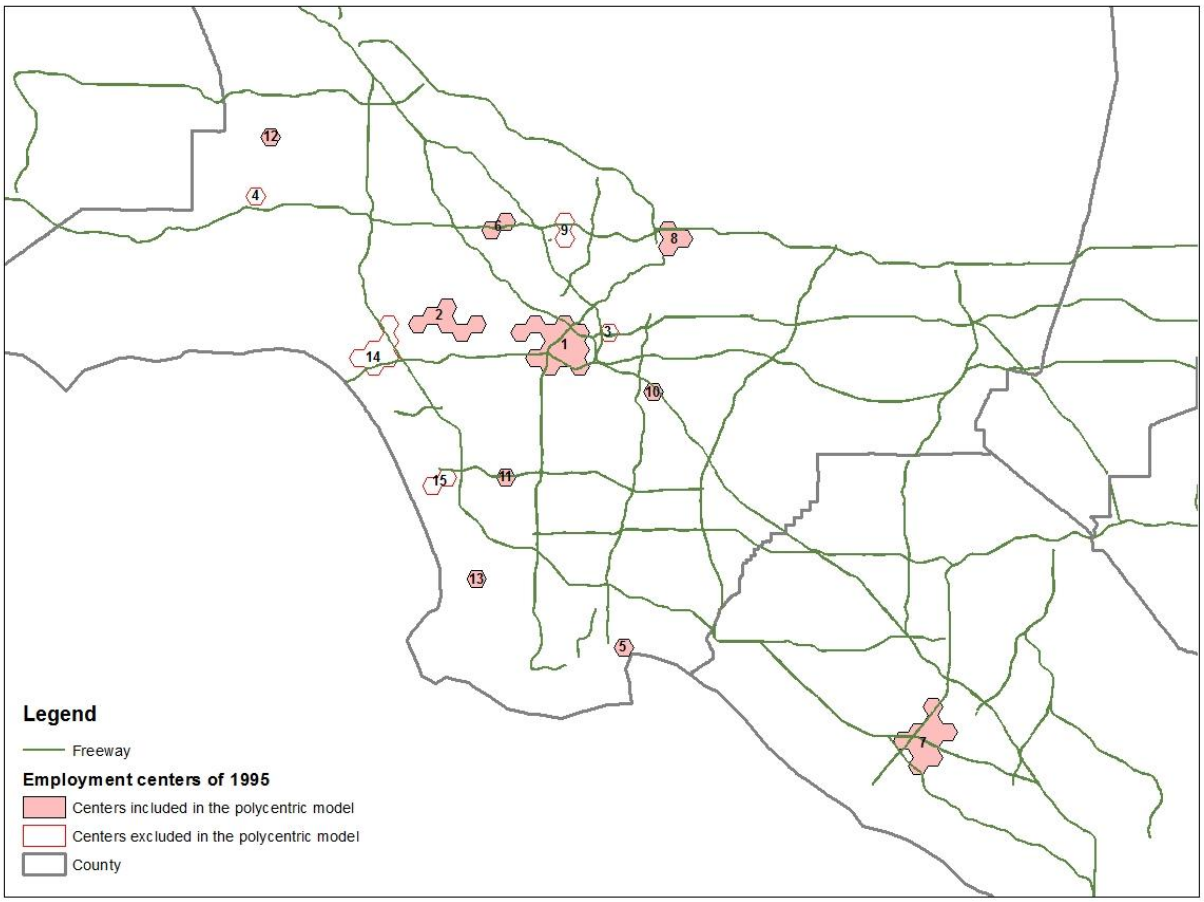
As indicated by the F tests, the polycentric function shows statistically significant improvements over the negative exponential monocentric model for all the three years. All the coefficients have expected signs, with the significant and negative value of the CBD gradient and positive values of subcenter gradients. The coefficients of distance to CBD is -0.04 when distances to subcenters are not added to the function, but drops to -0.01 when all the subcenters are included, implying the coefficients of distance to CBD in the monocentric model is really the "combined effects" of distance to employment centers.

Table 3-4 Regression results for polycentric density function, Los Angeles

	1995		2000		2005	
	Mono	**Poly**	Mono	**Poly**	Mono	**Poly**
DistCBD	-0.04	**-0.01**	-0.04	**-0.01**	-0.04	**-0.01**
	(-43.55)	**(-12.38)**	(-44.02)	**(-10.97)**	(-46.18)	**(-12.21)**
INDCENTER2		**11.46**		**7.32**		**8.48**
		(19.85)		**(10.03)**		**(13.36)**
INDCENTER3		**3.96**		**5.22**		**4.08**
		(5.77)		**(7.88)**		**(5.72)**
INDCENTER4						
INDCENTER5		**2.77**		**4.43**		**3.77**
		(3.32)		**(5.94)**		**(5.58)**
INDCENTER6		**2.40**		**2.08**		**4.09**
		(3.36)		**(2.9)**		**(5.51)**
INDCENTER7		**3.84**		**6.76**		**2.49**
		(5.68)		**(7.81)**		**(3.41)**
INDCENTER8		**6.02**		**3.30**		
		(8.61)		**(4.66)**		
INDCENTER9				**2.48**		**1.64**
				(2.53)		**(2.21)**
INDCENTER10		**3.26**				**8.47**
		(5.41)				**(13.22)**
INDCENTER11		**6.31**		**3.84**		
		(9.46)		**(6.35)**		
INDCENTER12		**3.40**				**5.48**
		(4.28)				**(8.19)**
INDCENTER13		**3.13**				**2.92**
		(4.29)				**(4.39)**
INDCENTER14				**5.04**		
				(7.47)		
INDCENTER15				**3.64**		**5.80**
				(5.09)		**(8.39)**
INDCENTER16				**7.15**		
				(11.06)		
INDCENTER17						**3.99**
						(5.61)
INDCENTER18						
Constant	0.08	**-2.64**	0.21	**-2.77**	0.30	**-2.63**
	(1.55)	**(-27.34)**	(4.04)	**(-27.93)**	(5.76)	**(-27.17)**
Adj R2	0.24	**0.37**	0.24	**0.36**	0.25	**0.38**
F	-	**123.79**	-	**128.03**		**128.33**
Sample Size	5938	**5928**	6139	**6128**	6491	**6480**
SSR	31634.0	**26160.2**	33681.4	**27376.2**	34962.3	28698.0
q	0	**10**	0	**11**	0	**11**
n-q-2	5936	**5916**	6137	**6115**	6489	**6467**

The t values are in parenthesis, the F value is computed as F(q,M - p) = [(SSR` - SSR")/q]/[SSR"/(M - p)], where SSR` and SSR" are restricted (monocentric) and unrestricted (polycentric) sums of squared residuals, M is sample size, p is the number of parameters estimated in the unrestricted model, and q is the number of restrictions in the restricted model.

The results also show that even though the total number of centers changed over the years, the number of subcenters that exert independent influence on the employment distribution is relatively stable, about 10 to 11. For example, the coefficients for center 2 may also combine the effects of distance to center 14 (by the 1995 ranking) because of their spatial proximity, while only two of the three centers (6,8,9 by the 1995 ranking) along the I-210 have significant coefficients in the final estimation model, implying their combined influence. Map 3-11 shows an example the locations of employment centers of 1995 that exert significantly independent influences on the overall employment distribution.



Map 3-11 Employment centers of 1995 (99%-20,000 cutoff, ranked by top employment density)

Our results also show that centers have varying density gradients, and the centers ranked highest based on peak density do not necessarily have the broadest spatial influence. For example, the coefficients for center 2 (West Hollywood) always have the highest value in all the three year estimates. Examples of percentage decline in gross employment density per mile are presented in Table 3-5. The density gradients are calculated based on the coefficients for distance to employment centers of 1995.

Table 3-5 Examples of percentage decline in employment density per mile

	Distance to employment centers of 1995 (miles)				
	2	5	10	15	20
West Hollywood (Center 2)	-2.12	-0.34	-0.08	-0.04	-0.02
Long Beach port (Center 5)	-0.94	-0.15	-0.04	-0.02	-0.01
John Wayne Airport (Center 7)	-1.02	-0.16	-0.04	-0.02	-0.01

3.3.2.2 San Francisco

We conduct the same tests for the San Francisco region. We find the same high correlations among the inverse distances of the centers, and thus use the same stepwise regression approach as described above. The physical geography of the region is such that centers are arrayed in an approximate ring pattern around the bay (See Map 3-5 and 6 in section 3.2.2 above). Results for the best fit models of the three years are presented in Table 3-6.

As indicated by the F test, the polycentric model explains the distribution of employment density better than the monocentric model. All the coefficients have expected signs. The coefficient of distance to CBD is -0.05 in monocentric model, and declines only slightly in the polycentric model. This suggests greater relative influence of the CBD compare to the Los Angeles case. Similar to the Los Angeles case, the higher order centers are consistently significant across the time periods, and the number of centers having significant influence over time is stable. Finally, Table 3-7 shows examples of percentage decline in gross employment density per mile over time. The employment density of a center with higher peak density declines more rapidly than that of a center with lower peak density. Note that the centers have substantially steeper gradients than those of Los Angeles (see Table 3-5).

Table 3-6 Regression results for polycentric density function, San Francisco

	1995		2000		2005	
	Mono	**Poly**	Mono	**Poly**	Mono	**Poly**
DistCBD	-0.05	**-0.04**	-0.05	**-0.03**	-0.05	**-0.03**
	-26.19	**-16.42**	-26.62	**-11.24**	-27.8	**-11.58**
INDCENTER2		**10.91**		**11.7**		**12.01**
		17.46		**18.44**		**20.75**
INDCENTER3		**6.64**		**6**		**6.04**
		9.33		**8.28**		**8.62**
INDCENTER4				**5.32**		
				7.15		
INDCENTER5						
INDCENTER6		**3.51**				
		5.36				
INDCENTER7						**4.31**
						6.63
INDCENTER8				**3.48**		**4.88**
				5.22		**6.58**
INDCENTER9				**2.96**		**4.99**
				4.43		**7.29**
INDCENTER10						
Constant	-0.18	**-1.65**	-0.08	**-2.23**	-0.01	**-2.26**
	1.9	**13.75**	0.89	**-15.24**	0.13	**-15.9**
Adj R2	0.1541	**0.25**	0.15	**0.26**	0.16	**0.27**
F		**172.88**		**123.7**		**135.03**
Sample Size	3759	**3756**	3927	**3922**	4113	**4108**
SSR	20356.7	**17884**	22343.1	**19294.9**	22448.4	19275
q	0	**3**	0	**5**	0	**5**
n-q-2	3757	**3751**	3925	**3915**	4111	**4101**

Table 3-7 Examples of percentage decline in employment density per mile, San Francisco

	Year	Distance to employment centers (miles)				
		2	5	10	15	20
San Jose (Center 2)	1995	-3.01	-0.48	-0.12	-0.05	-0.03
	2000	-3.16	-0.50	-0.13	-0.06	-0.03
	2005	-3.03	-0.48	-0.12	-0.05	-0.03
Oakland-Berkeley (Center 3)	1995	-1.60	-0.26	-0.06	-0.03	-0.02
	2000	-1.60	-0.26	-0.06	-0.03	-0.02
	2005	-1.59	-0.25	-0.06	-0.03	-0.02

3.3.2.3 San Diego

For the San Diego CMSA, we conduct polycentricity tests using 1995, 2000, and 2005 employment centers. Because of the lower overall average employment density of San Diego, we use 95 percentile minimum density and 10,000 total jobs criteria for the identification of the employment centers. The numbers of identified employment subcenters are seven, twelve and eleven in 1995, 2000, and 2005 respectively. We use the same stepwise regression, and results are given in Table 3-8. For all years polycentricity is confirmed. Similar to Los Angeles, adding additional centers substantially reduces the gradient for the CBD. Coefficients of the higher order centers are consistently significant, but in contrast the number of centers for which coefficients are significant increases, perhaps suggesting greater spatial re-organization as the region's employment base grows.

Table 3-8 Regression results for polycentric density function, San Diego

	1995		2000		2005	
	Mono	**Poly**	Mono	**Poly**	Mono	**Poly**
DistCBD	-0.10	**-0.06**	-0.10	**-0.04**	-0.10	**-0.04**
	(-26.63)	**(-9.81)**	(-25.82)	**(-6.15)**	(28.14)	**(-7.88)**
INDCENTER2		**2.79**		**2.44**		**2.73**
		(3.88)		**(3.45)**		**(3.98)**
INDCENTER3		**3.75**		**3.40**		**3.47**
		(4.86)		**(4.48)**		**(4.84)**
INDCENTER4		**4.46**		**2.72**		**2.94**
		(7.31)		**(3.67)**		**(4.85)**
INDCENTER5		**5.18**		**3.52**		**3.00**
		(6.43)		**(4.40)**		**(3.85)**
INDCENTER6		**2.05**		**2.96**		**6.55**
		(3.06)		**(4.73)**		**(10.04)**
INDCENTER7				**6.28**		
				(9.35)		
INDCENTER8						**2.50**
						(3.21)
INDCENTER9						
INDCENTER10				**3.64**		**3.26**
				(5.46)		**(4.73)**
INDCENTER11				**3.03**		**6.35**
				(4.26)		**(8.51)**
INDCENTER12				**6.15**		**2.20**
				(7.84)		**(2.74)**
INDCENTER13						
Constant	-2.37	**-1.83**	0.60	**-3.40**	0.88	**-2.93**
	(5.13)	**(-6.78)**	(5.23)	**(-12.36)**	(7.88)	**(-11.07)**
Adj R2	0.33	**0.37**	0.30	**0.44**	0.33	**0.47**
F	-	**27.13**	-	**48.14**	-	**50.44**
Sample Size	1460	**1455**	1530	**1521**	1597	**1588**
SSR	6029.3	**5513.0**	6739.5	**5236.9**	6665.2	**5175.3**
q	0	**5**	0	**9**	0	**9**
n-q-2	1458	**1448**	1528	**1510**	1595	**1577**

3.3.2.4 Sacramento

Sacramento has by far the lowest average employment density and the smallest number of centers (Table 4-6 and 4-10). We therefore use the centers identified by the 95 percentile minimum density and 10,000 total jobs criteria. The numbers of identified centers are six, six and seven in 1995, 2000, and 2005 respectively excluding the CBD. Results for the best fit models of the three year periods are presented in Table 3-9. Similar to the previous cases, the polycentric model is confirmed for all time periods, and the coefficient on the CBD is lower in the polycentric models. Unlike the other metro areas, the higher order center coefficients are not consistently significant, and the amount of variation explained as measured by pseudo-R square is much lower. Sacramento employment is more dispersed, and hence the centers explain less of the overall density distribution.

Table 3-9 Regression results for polycentric density function, Sacramento

	1995		2000		2005	
	Mono	**Poly**	Mono	**Poly**	Mono	**Poly**
DistCBD	-0.03	**-0.01**	-0.03	**-0.02**	-0.03	**-0.02**
	(-14.34)	**(-2.35)**	(-14.80)	**(-6.99)**	(-16.77)	**(-8.36)**
INDCENTER2				**8.92**		**9.84**
				(14.43)		**(16.24)**
INDCENTER3		**7.19**				
		(12.17)				
INDCENTER4						
INDCENTER5				**3.77**		**3.32**
				(5.02)		**(4.74)**
INDCENTER6		**8.94**		**2.63**		
		(13.80)		**(4.55)**		
INDCENTER7		**2.82**				**2.31**
		(5.06)				**(4.10)**
INDCENTER8						
Constant	-2.37	**-4.17**	-2.20	**-3.44**	-2.00	**-3.28**
	(-29.41)	**(-34.39)**	(-27.40)	**(-28.82)**	(25.75)	**(-28.59)**
Adj R2	0.07	**0.18**	0.07	**0.14**	0.09	**0.16**
F	-	**125.85**	-	**82.65**	-	**99.44**
Sample Size	2665	**2662**	2767	**2764**	2970	**2967**
SSR	11657.5	**10207.1**	12445.8	**11419.6**	13091.5	**11893.6**
q	0	**3**	0	**3**	0	**3**
n-q-2	2663	**2657**	2765	**2759**	2968	**2962**

3.3.2.5 Cross-region comparison

In all metropolitan areas, the F test results show that the polycentric model explains the distribution of employment density better than the monocentric model. All the coefficients have expected signs. The coefficients of distance to CBD are negative in both monocentric and polycentric models across the four metro areas. Even though the magnitude of distance to CBD coefficients became smaller in polycentric model, the coefficients are still statistically significant. This implies that distance to CBD still plays a significant role in explaining the distribution of employment density even after controlling for the influence of other centers. Also, the coefficients of inverse distance to other subcenters are positive, as expected.

In general, the number of centers with significant influence is related to the size of the metro area and total number of centers (note that we use a different cutoff for San Diego and Sacramento, which affects total number of centers). The weakest case for polycentricity is Sacramento. As we will show in the next chapter, the Sacramento CBD is far larger and denser (relative to the region's total employment) than any other center.

Chapter 4 Descriptive analysis

This chapter presents our descriptive analysis of employment centers identified in our four case study metropolitan areas. We begin with an overview of the four metro areas, and then describe the characteristics of centers over the three time periods.

4.1 The metro areas

As discussed in chapter 3, our analysis areas are the four largest metro areas in California: Los Angeles, San Francisco, San Diego, and Sacramento. We use the definition of "Combined Statistical Areas" (CSA) by United States Office of Management and Budget (OMB) for the LA, SF and Sac regions, and the definition of "Core Based Statistical Areas" (CBSA) for the San Diego region because it has only a single urban core and no associated CSA[18]. The component counties and total areas of each metro area are given in Table 4-1.

We choose 1995, 2000 and 2005 as our analysis years and applied the uniform hexagon grid as our basic unit of analysis. The establishment-level NETS employment data across the three analysis years are aggregated to the consistent geography of 1-sq mile hexagons for valid comparison. The census tract level population data are sourced from 1990/2000 U.S. Census data and 2005-2009 American Community Survey (ACS) data. To maintain the consistency of study years, we approximate the 1995 tract-level population with the average of 1990 and 2000 population levels and the 2005 population with the 2005-2009 ACS estimates. The basic trends of employment and population growth of the four regions are summarized in Table 4-2. All regions had positive population growth from 1995 to 2005, but the San Francisco region had much slower growth, while the population growth of Sacramento was double that of Los Angeles and San Diego. Employment growth over the decade was also positive for all regions, but the effect of the "Dot-Com" boom and recession of 2001 is evident for San Francisco. In general, employment growth is about the same for Los Angeles and Sacramento, but employment growth is proportionately greater than population growth for San Francisco and San Diego. It is unclear whether these trends reflect an overall increase in employment levels and labor force participation rate, or increased commuting from outlying areas not captured in the study boundaries.

[18] The United States Office of Management and Budget (OMB) defines a CBSA as one or more adjacent counties or county equivalents that have at least one urban core area of at least 10,000 population, together with adjacent communities having a high degree of economic and social integration with that core. Those CBSAs with an urban core population of 50,000 or more are defined as Metropolitan Statistical Areas (MSAs), while those with a core population 10,000 or more but less than 50,000 are defined as Micropolitan Statistical Areas (µSAs). A Combined Statistical Area as an aggregate of adjacent CBSAs that are linked by commuting ties. Source: http://www.whitehouse.gov/sites/default/files/omb/bulletins/2013/b13-01.pdf

Table 4-1 Definition of the four metropolitan areas [19]

CSA/CBSA	Area (acre)	Component CBSA	county
Los Angeles-Long Beach-Riverside (LA)	22,601,997	Los Angeles-Long Beach-Santa Ana	Los Angeles[20]
			Orange
		Oxnard-Thousand Oaks-Ventura	Ventura[21]
		Riverside-San Bernardino-Ontario	Riverside
			San Bernardino
San Jose-San Francisco-Oakland (SF)	6,521,957	San Francisco-Oakland-Fremont	Alameda
			Contra Costa
			Marin
			San Francisco
			San Mateo
		San Jose-Sunnyvale-Santa Clara	San Benito
			Santa Clara
		Santa Cruz-Watsonville	Santa Cruz
		Santa Rosa-Petaluma	Sonoma
		Vallejo-Fairfield	Solano
		Napa	Napa
Sacramento-Arden-Arcade--Yuba City (Sac)	5,292,878	Sacramento--Arden-Arcade--Roseville	El Dorado
			Placer
			Sacramento
			Yolo
		Truckee-Grass Valley	Nevada
		Yuba City	Sutter
			Yuba
San Diego (SD)	2,896,437	San Diego-Carlsbad-San Marcos	San Diego

19 Original source: http://www.census.gov/population/metro/files/lists/2009/List1.txt

20 Santa Catalina Island and San Clemente U.S. Military Reservation are excluded in our analysis.

21 Anacapa Island and San Nicolas Island are excluded in our analysis.

Table 4-2 Employment and population of 4 regions

	1995		2000				2005					
Region	Emp	Pop *	Emp	Emp Change (%)	Pop	Pop Change (%)	Emp	Emp Change (%)	Emp Change by decade (%)	Pop **	Pop Change (%)	Pop Change by decade (%)
LA	7,495,021	15,452,579	8,296,734	10.69	16,373,645	5.96	8,712,612	5.01	16.25	17,577,378	7.35	13.75
SF	3,683,500	6,672,593	4,380,182	18.91	7,092,595	6.29	4,197,545	- 4.17	13.96	7,257,729	2.33	8.77
Sac	838,391	1,855,492	962,265	14.78	2,028,031	9.30	1,057,669	9.91	26.15	2,335,271	15.15	25.86
SD	1,239,818	2,655,195	1,481,634	19.50	2,813,076	5.95	1,594,595	7.62	28.62	2,986,821	6.18	12.49

* Population in 1995 is imputed average of 1990 and 2000

** Population in 2005 is ACS estimates for 2005-2009

4.2 95%-10k Employment center characteristics

As discussed in Chapter 3, employment centers are identified for each metro area as a contiguous set of hexagons, each with density above top 5 percent cutoff, that include at least 10,000 total employment. This method is devised to reflect the differences in geographies and size across metro areas in the criteria for identifying employment centers. This method allows generating different 95 percent cutoff value of employment density for different metro areas and the cutoff for each area is quite stable over time (Los Angeles CSA: app. 10 jobs/acre, San Francisco CSA: app. 8 jobs/acre, San Diego: approx. 8 jobs/acre; Sacramento CSA: app. 2.5 jobs/acre).

Tables 4-3 through 6 provide the total number of employment centers, total employment within centers, and proportion of employment within centers for each metropolitan area. These Tables show that the number of centers in each metropolitan areas follows the same rank order as population and employment, and that the share of employment in centers is quite large, ranging from about 40 % for Los Angeles to more than 50% for San Francisco and Sacramento. The number of centers in Los Angeles and San Diego areas increased between 1995 and 2000, the number of centers in other areas is quite stable. About 50 centers, 30 centers, 10 centers and 7 centers are identified in Los Angeles CSA, San Francisco CSA, San Diego CSBA and Sacramento CSA respectively. The Tables also show that the number of centers in each metro area is quite stable over time. These results suggest that agglomeration economies within metro areas continue to be a significant force in employment location and spatial organization. As total employment has increased over the period, there is no indication that this new growth is more spatially dispersed.

Table 4-3 Basic information of 95%-10K centers: Los Angeles CSA

	1995	2000	2005
Total number of centers	48	53	53
Total Employment within centers (thousands)	3,126.3	3,528.9	3,593.5
Proportion of employment within centers (%)	41.7	42.5	41.2

Table 4-4 Basic information of 95%-10K centers: San Francisco CSA

	1995	2000	2005
Total number of centers	28	28	30
Total employment within centers (thousands)	1,982.1	2,340.9	2,166.8
Proportion of employment within centers (%)	53.8	53.4	51.6

Table 4-5 Basic Information of 95%-10K centers: San Diego CSBA

	1995	2000	2005
Total number of centers	7	13	12
Total employment within centers (thousands)	494.8	647.7	654.5
Proportion of employment within centers (%)	39.9	43.7	41

Table 4-6 Basic information of 95%-10K centers: Sacramento CSA

	1995	2000	2005
Total number of centers	7	7	8
Total employment within centers (thousands)	472.9	524.1	578.5
Proportion of employment within centers (%)	56.4	54.5	54.7

Tables 4-7 through 10 give some characteristics of the four sets of centers. The first row gives the number of jobs in the largest center. By 2000 two large centers merge in Los Angeles, increasing the number of jobs from about ½ million to about 1 million – nearly 1/3 of all jobs located in centers in the region. In the other metro areas the largest center remains relatively stable. Note that the number of jobs in the largest center is much lower for San Diego and Sacramento.

The second row of the Tables shows that average number of jobs per center is quite consistent across metro areas and time periods. This is in part explained by the lower bound cutoff, but nevertheless shows an interesting regularity across the metro areas. With the exception of Sacramento, there is also remarkable regularity in the average size, employment density, and population density of centers. This is not an artifact of any lower bound, as all centers are aggregated on the basis of density of one square mile. The Sacramento case is different: the region has much lower overall employment density (hence the lower density cutoff). The last row in the Tables gives the employment to population ratio, a measure of population and employment mix. The pattern here is less clear. For Los Angeles, San Francisco, and San Diego, the ratio is in the range of 2 to 3, and is around 1 for Sacramento. Thus on average, there is substantial mixing of population in employment centers.

Table 4-7 Selected characteristics of 95%-10K centers: Los Angeles CSA

	1995	2000	2005
Maximum jobs (thousands)	560	1049.7	1134.7
Average number of jobs per center	65,131	66,583	67,802
Average size (acres)	3,373	3,260	3,454
Average Jobs per acre	16.8	17.8	17.1
Average Population per acre	9.3	9.6	9.6
Average employment to population ratio	1.9	2.3	2.2

Table 4-8 Selected characteristics of 95%-10K centers: San Francisco CSA

	1995	2000	2005
Maximum jobs (thousands)	601.9	678.8	603.2
Average number of jobs per center	70,003	83,219	72,227
Average size (acres)	3,748	3,925	3,920
Average Jobs per acre	14.8	16.8	15.0
Average population per acre	7.9	9.4	9.1
Average employment to population ratio	2.6	2.4	2.1

Table 4-9 Selected characteristics of 95%/10K centers: San Diego CSBA

	1995	2000	2005
Maximum jobs (thousands)	206.6	215.2	209.5
Average number of jobs per center	61,851	49,824	54,540
Average size (acres)	4,720	3,200	3,413
Average jobs per acre	11.32	13.9	14.1
Average population per acre	6.7	6.3	6.3
Average employment to population ratio	2.2	3.6	2.7

Table 4-10 Selected characteristics of 95%-10K centers: Sacramento CSA

	1995	2000	2005
Maximum jobs (thousands)	385.5	423.7	391.5
Average number of jobs per center	67,561	74,874	72,314
Average size (acres)	9,051	9,051	9,040
Average jobs per acre	5.7	6.0	6.6
Average population per acre	5.7	5.6	5.3
Average employment to population ratio	1.1	1.1	1.3

Tables 4-11 through 14 give the number of centers by size category over time for each metro area, from smallest to largest. We observe a general pattern of rank size distribution in every metro area, with only 1 or 2 centers in the largest category and many centers in the smallest category. The pattern is most clear for Los Angeles. The fluctuations across time periods for San Francisco may reflect the unique employment growth and decline in experience compared to the other metro areas.

Table 4-11 Rank size of 95%-10K centers: Los Angeles CSA

	1995	2000	2005
Number of centers with 10 000 - 20 000 jobs	21	26	20
Number of centers with 20 000 - 50 000 jobs	15	13	19
Number of centers with 50 000 - 100 000 jobs	8	9	9
Number of centers with 100 000 - 500 000 jobs	3	3	3
Number of centers with > 500 000 jobs	1	2	2
Total number of centers	48	53	53

Table 4-12 Rank size of 95%-10K centers: San Francisco CSA

	1995	2000	2005
Number of centers with 10,000-20,000 jobs	10	7	12
Number of centers with 20,000-50,000 jobs	14	12	11
Number of centers with 50,000-100,000 jobs	1	5	4
Number of centers with 100,000-500,000 jobs	2	2	1
Number of centers with >500,000 jobs	1	2	2
Total	28	28	30

Table 4-13 Rank size of 95%/10K centers: San Diego CSBA

	1995	2000	2005
Number of centers with 10 000 - 20 000 jobs	3	7	5
Number of centers with 20 000 - 50 000 jobs	2	3	4
Number of centers with 50 000 - 100 000 jobs	1	0	1
Number of centers with 100 000 - 500 000 jobs	2	3	2
Number of centers with > 500 000 jobs	0	0	0
Total number of centers	8	13	12

Table 4-14 Rank size of 95%-10K centers: Sacramento CSA

	1995	2000	2005
Number of centers with 10 000 - 20 000 jobs	6	5	4
Number of centers with 20 000 - 50 000 jobs	0	1	2
Number of centers with 50 000 - 100 000 jobs	0	0	1
Number of centers with 100 000 - 500 000 jobs	1	1	1
Number of centers with > 500 000 jobs	0	0	0
Total number of centers	7	7	8

4.3 99%-20K Employment center characteristics

In this section we discuss centers identified by the more restrictive criteria of 99% and 20K. The 99th percentile densities are 23, 19, 16, and 8 for Los Angeles, San Francisco, San Diego, and Sacramento respectively. Tables 4-15 through 18 give the number of centers and share of employment in centers for each metro area across the 3 time periods. The more restrictive criteria reduce the total number of centers as well as the share of employment in centers. The pattern is similar to that in Tables 4-3 through 6: the number of centers is stable or increasing. The share of employment in centers ranges from 13 – 16% for San Diego and Los Angeles to more than 20% for Sacramento and San Francisco. There is a slight trend of share decline over time, but an increase in total number of centers for all but Sacramento, suggesting that center employment growth is greater in smaller centers (e.g. agglomeration diseconomies may be evident in the largest centers).

Table 4-15 Basic information of 99%-20K centers: Los Angeles CSA

	1995	2000	2005
Total number of centers	15	18	18
Total Employment within centers (thousands)	1,217.0	1,420.4	1,380.4
Proportion of employment within centers (%)	16.2	17.1	15.8

Table 4-16 Basic information of 99%-20K centers: San Francisco CSA

	1995	2000	2005
Total number of centers	6	9	10
Total employment within centers (thousands)	860.1	1,084.1	909.2
Proportion of employment within centers (%)	23.4	24.8	21.7

Table 4-17 Basic information of 99%-20K centers: San Diego CSBA

	1995	2000	2005
Total number of centers	3	3	3
Total employment within centers (thousands)	159.9	242.9	224.3
Proportion of employment within centers (%)	12.9	16.4	14.1

Table 4-18 Basic information of 99%-20K centers: Sacramento CSA

	1995	2000	2005
Total number of centers	2	2	2
Total employment within centers (thousands)	219.3	259.0	278.0
Proportion of employment within centers (%)	26.2	26.9	26.3

Tables 4-19 through 22 provide the same information as Tables 4-15 through 18 above. The patterns indicate different spatial organization across the metro areas. The 99th percentile criterion has the effect of breaking up Los Angeles' main center, generating multiple smaller but more dense centers. Apparently there are multiple local peaks within the 95th percentile main center. The pattern is quite different for San Francisco: just 6 centers remain, but they are very large and very dense. San Diego has three small (in terms of employment) centers, and Sacramento has just two, much larger in geographic size and much less dense, again due to the relatively low employment density of the entire region. Although average population density in centers is close to double that of the 95/10 centers, the employment to population ratio is many times greater due to much higher employment density. This is as expected: higher rent implied by higher densities pushes out population.

Table 4-19 Selected characteristics of 99%-20K centers: Los Angeles CSA

	1995	2000	2005
Maximum jobs (thousands)	367.2	379.2	368.1
Average number of jobs per center	81,130.1	78,911.7	76,689.9
Average size (acres)	2,048	1,956	1,920
Average Jobs per acre	37.7	38.5	36.5
Average Population per acre	14.7	14.7	17.1
Average employment to population ratio	4.5	4.6	3.9

Table 4-20 Selected characteristics of 99%-20K centers: San Francisco CSA

	1995	2000	2005
Maximum jobs (thousands)	528.5	596.5	522.2
Average number of jobs per center	143,351.3	120,270.8	90,924.6
Average size (acres)	2,730	2,246	1,894
Average Jobs per acre	47.6	45.5	38.2
Average Population per acre	14	11.9	11.3
Average employment to population ratio	7.4	6.7	5

Table 4-21 Selected characteristics of 99%-20K centers: San Diego CSBA

	1995	2000	2005
Maximum jobs (thousands)	78,888	92,814	87,953
Average number of jobs per center	53,308.3	80,966.0	74,753.3
Average size (acres)	1,707	2,560	2,347
Average Jobs per acre	34.9	39.8	38.7
Average Population per acre	5.4	6.6	8.4
Average employment to population ratio	8.4	6.3	5.0

Table 4-22 Selected characteristics of 99%/20K centers: Sacramento CSA

	1995	2000	2005
Maximum jobs (thousands)	197,799	228,054	240,914
Average number of jobs per center	109,631.5	129,480.5	139,017.0
Average size (acres)	5,760	6,080	7,360
Average Jobs per acre	18.0	19.2	16.3
Average Population per acre	6.9	6.2	5.5
Average employment to population ratio	2.8	3.4	3.3

Finally, Tables 4-23 and 4-24 show the size distribution of centers for Los Angeles and San Francisco. We do not show results for San Diego and Sacramento due to the small number of centers. All three of San Diego's centers are in the 50 – 100K category, while Sacramento has one center in the 100-500K category, and the other in the 20-50K category. For Los Angeles and San Francisco there is a general pattern of rank size, but it is not as consistent as with the 95/10 centers.

Table 4-23 Rank size of 99%-20K centers: Los Angeles CSA

	1995	2000	2005
Number of centers with 20 000 - 50 000 jobs	10	12	11
Number of centers with 50 000 - 100 000 jobs	2	2	3
Number of centers with 100 000 - 500 000 jobs	3	4	4
Number of centers with > 500 000 jobs	0	0	0
Total	15	18	18

Table 4-24 Rank size of 99%-20K centers: San Francisco CSA

	1995	2000	2005
Number of centers with 20,000-50,000 jobs	3	6	7
Number of centers with 50,000-100,000 jobs	1	1	2
Number of centers with 100,000-500,000 jobs	1	1	0
Number of centers with >500,000 jobs	1	1	1
Total	6	9	10

4.4 Conclusions

There are two main observations to be drawn from this very basic descriptive analysis. First, the presence of centers is consistent across all metropolitan areas and time periods. The number of centers either grows or remains stable as each metro area has grown, and the share of employment in centers is stable, showing that the agglomeration benefits associated with center location continue to be strong. In the case of the 99/20 centers, there is a slight decline in employment share, but no dramatic change. Second, there are similarities and differences across the four metro areas. The total number of centers is associated with metropolitan size, but the relative dominance of the largest center (as measured by center employment) differs across metro area and the two sets of centers. San Francisco is more "centralized" than Los Angeles, meaning that it has fewer but more dense centers. San Francisco also demonstrates the greatest volatility over time, perhaps due to its unique economy. Sacramento is clearly different from the other metro areas. It is the most centralized in the sense of having one large center (downtown area) and a few dispersed small centers. It is possible that Sacramento metropolitan size (the smallest at about 2 million in population), together with the absence of geographical constraints allows low density, dispersed development to proceed. In contrast, San Diego is physically constrained by ocean on the west and hills on the east, perhaps accelerating concentration.

Chapter 5 Explaining employment center growth

In this chapter, we estimate models to explain center growth for Los Angeles and San Francisco. Using the NETS data, we replicate our previous model on the relationship between center growth and accessibility to labor force and the highway system and examine whether our results are robust across different regions for different time periods.

5.1 Research approach, methodology

5.1.1 Conceptual model

Following our previous study (Giuliano et al. 2010), we apply a simple model of employment center growth as a function of accessibility and a set of control variables:

$$\Delta E = f(X, Z) \tag{1}$$

where ***X*** = vector of access measures

Z **=** vector of control measures

As explained in Giuliano et al (2010), we develop appropriate measures of accessibility and control other relevant factors. Our accessibility measures include access to transport networks and facilities, as well as labor force access. Building on our earlier work, we control the following characteristics of centers: size, density, industry composition, and location within the region.

Center size: Large centers might grow faster because of the benefits of agglomeration economies, or might grow slower or decline as a result of congestion, land scarcity, pressures on public facilities, and other diseconomies of agglomeration.

Center density: Less dense centers may grow more quickly because lower density may indicate more land availability and lower land prices.

Industry composition: Centers with larger shares of fast growing industries should fare better than centers with larger shares of slow growing or declining industries. For each center we compute the amount of growth that would have occurred if each industry sector had grown at its rate for the entire region:

$$\Delta E_m^P = \sum_i E_i \overline{g_i} \tag{2}$$

where ΔE^P is the predicted growth of center *m*, E_i is the base period employment in sector *i*, and g_i is the regional growth rate of sector *i*. The predicted growth rate of each center is computed as the ratio of predicted growth and the center's employment in the base year.

$$\Delta g_m^p = \Delta E_p^m / E_{m,t0} \tag{3}$$

where $E_{m,t0}$ is the employment level of center m in the base year (t0).

Proximity to region's center: Proximity to the core may imply both benefits of urbanization economies and overall access to labor force, as well as high costs of doing business such as higher land prices and more congestion. We measure distance to the core as the straight line distance to the peak tract of the CBD.

Access to the coast: The coast is a major amenity in Southern California. The weather is more moderate near the coast, and prevailing westerly winds make air quality much better near the coast. These amenities may be attractive to firms. We therefore include a measure of access to the coastline, measured as straight line distance.

5.1.2 Measuring accessibility

We follow Giuliano et al (2010) and generate the following measures of accessibility:

Network accessibility: The first transport access measure is highway network accessibility. It is measured as,

$$A_i = \sum_j e^{-\beta d_{ij}} \tag{4}$$

where d_{ij} is the travel cost (in distance or time) between nodes *i* and *j*, and β is the impedance function. In our case d_{ij} is the shortest path free-flow travel time (see discussion below). The value of β is estimated as the inverse of the average commute in 2000. Because the highway mode is overwhelmingly dominant, we do not compute a transit network accessibility measure.

Access to Airports: The region has several major airports. We generate simple measures of straight line distance from the four largest airports: LAX, Santa Ana, Burbank, and Ontario.

Because LAX is much larger than the others, it is possible that access to LAX has a much more extensive influence than the other airports.

Labor Force Accessibility: From the perspective of the firm, location considerations include access to potential workers, consumers, and production inputs. We use population distribution as a proxy for both labor force access and consumer market access and construct two different measures of labor force access. The first measures "total" labor force access, a weighted sum of population discounted by distance,

$$A_m = \sum_j L_j^{-\beta d_{jm}} \qquad (5)$$

where L_j is population in tract j, d_{jm} is the distance between j and m, the hexagon of center m with the peak employment density, and β is the impedance parameter. Note that labor force accessibility for each center is calculated as the accessibility of the **peak hexagon**, and takes into account the resident labor force in all tracts within the region.

The second measures "relative" labor force access, and takes into account competition for labor from other employment locations,

$$B_m = \sum_j L_j \left(E_m e^{-\beta d_{jm}} / \sum_k E_k e^{-\beta d_{km}} \right) \qquad (6)$$

where E_m and E_k are total employment in centers m and k respectively. B_m may be viewed as attaching to each member of the labor force a probability, based solely on commuting distance, of choosing to work in the employment center in question. The parameter β is the impedance parameter and set as the same across the studies. We expect all access measures to be positively associated with employment center growth.

5.2 Data (LA region)

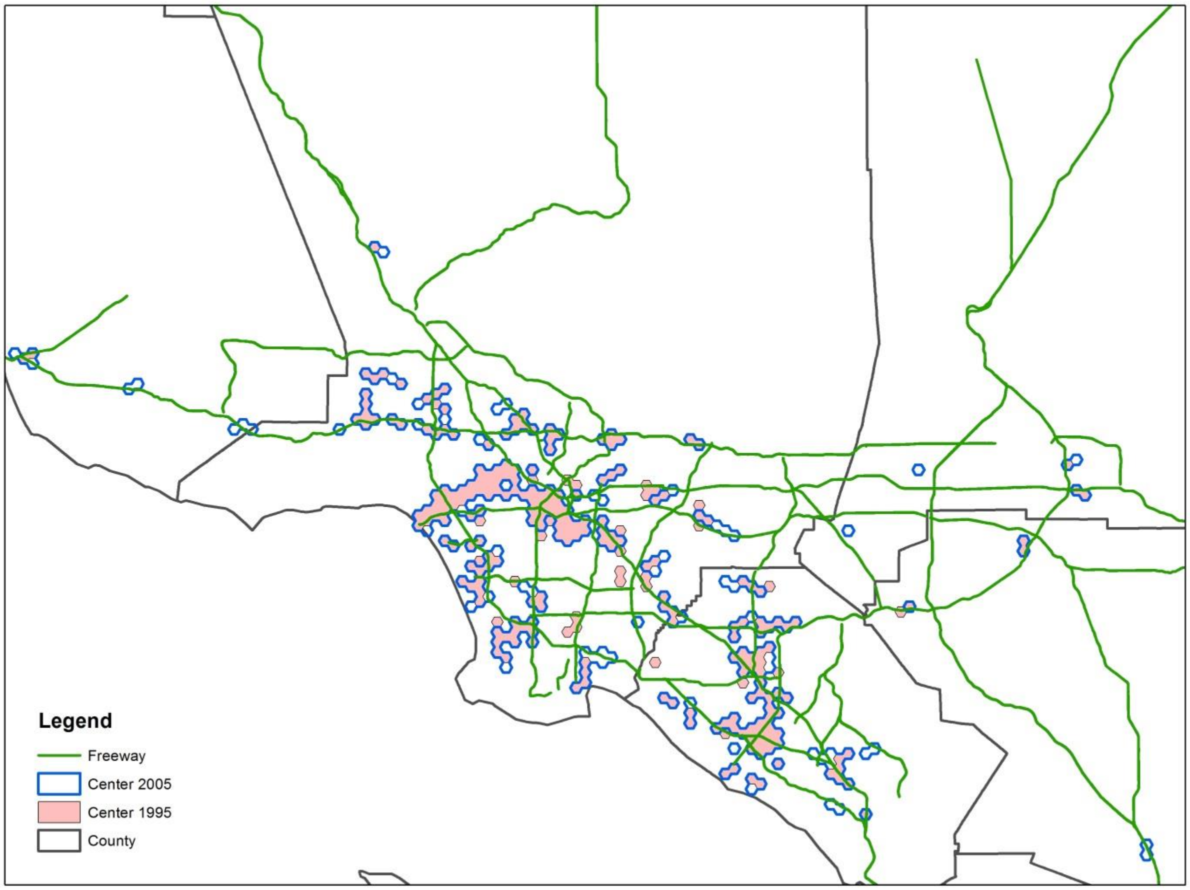
We were able to collect the required employment, population and transport network data for two of our four regions, Los Angeles and San Francisco. In our previous work with Los Angeles, the number of centers was sufficient for quantitative analysis. However, other metro areas, though polycentric, have many fewer centers than Los Angeles. The only way to conduct a quantitative analysis with smaller metro areas is to pool samples, either across metro areas, or across time periods. We present two sets of results here. First, we replicate the Giuliano et al (2010) model for the Los Angeles region, using different sample years. Second we conduct a pooled analysis of Los Angeles and San Francisco.

As discussed in chapter 2, the establishment counts of small-sized firms in the NETS data increases across the time period of our data. We checked the annual distribution of establishments between centers and non-center location for each employment size class and found that the establishment distribution are relatively stable for the 1995-2005 study period (see Figure A-5-1 and Table A-5-1 in Appendix 5). Thus the changing share of small firms does not bias our analysis of employment center growth.

5.2.1 Employment centers characteristics in the LA region

The identification and summary statistics of centers have been discussed in chapter 3 and 4. As employment centers grow or decline, they may expand or shrink in geographic size. What is the appropriate geography for measuring growth? We could use the center as it existed in the base year, the center as it exists in the final year, or the combination of the two (any area that is part of a center at any time period). Ultimately we decided to use the geography of the final year, reasoning that we are explaining current conditions (the final year center) as a function of previous year conditions (e.g. employment, access in the base year). Map 5-1 shows the 2005 centers overlaid on the 1995 centers to provide a visual mapping of where changes in boundaries took place.

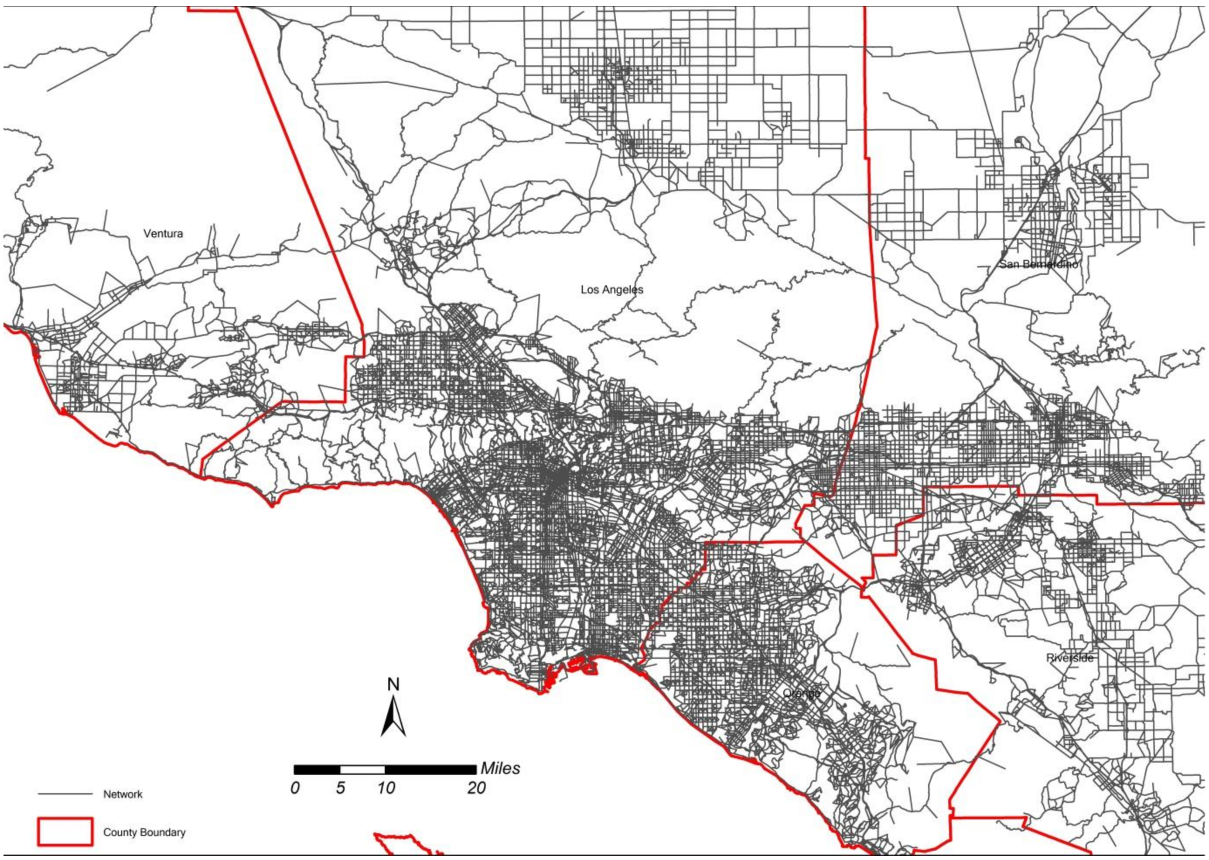
Based on the 2005 boundary, total employment growth for center locations during the decade is 11.73%, much lower than the regional growth rate (16.25%). There is a variation of employment growth and decline across 2005 centers, with 13 centers having declined during the 2000-2005 period and 8 centers having declined for the 1995-2005 decade (see details in Table A-5-2 in Appendix 5). Center 31 (USC Health Center) was the biggest loser in terms of employment for both 1995-2000 and 1995-2005. Center 14 (Northbridge) ranked second and third in terms of job loss for the 1995-2000 and 1995-2005 periods, respectively. Center 13 (Commerce) was also the second biggest employment loser for 2000-2005 and 1995-2005. Centers 3 (El Segundo), 4 (Torrance) and 11 (Hidden Hills) were the first three biggest employment losers for the 2000-2005 period but had positive net employment growth for the 1995-2005 period. On the other hand, Center 2 (Anaheim-Santa Ana- Costa Mesa-Irvine) and center 1 (LA Downtown-West LA-Santa Monica) ranked first and second, respectively, in terms of employment growth for the 1995-2005 and 1995-2000 periods. However, these two centers had much less employment growth for 2000-2005. The mix of centers with job gains and losses enables us to test the hypotheses on the determinants of employment center growth.



Map 5-1 Employment centers identified using 1995 data and 2005 data

5.2.3 Road Network of 2003 in the LA region

The transportation networks used are network files of Los Angeles 2003 provided by SCAG (Southern California Associations of Governments). The files are compiled in the TransCAD software package and contain detailed information for each link and node. Map 5-2 shows the 2003 network.



Map 5-2 2003 Transportation Network (source: SCAG)

To build the travel time OD matrix, the centroid of the peak hexagons of centers and those of census tracts are assigned to the closest network nodes falling within them. For a few census tracts without any nodes falling inside, we manually create additional "centroid connectors" from their centroids to the nearest nodes and assigned a speed of 30 miles per hour to those connectors, whereas travel time of these additional links are calculated by dividing the length by speed[22]. We use the "pre-calculated" minimum travel time for either direction for each link as the travel cost. Using the shortest path algorithm, we generate travel times for every centroid to centroid pair. We use the 2003 data for all panels because the road network does not change between 1995 and 2005 and the free-flow travel time and speed for each link is determined by design and should not change with the land use patterns.

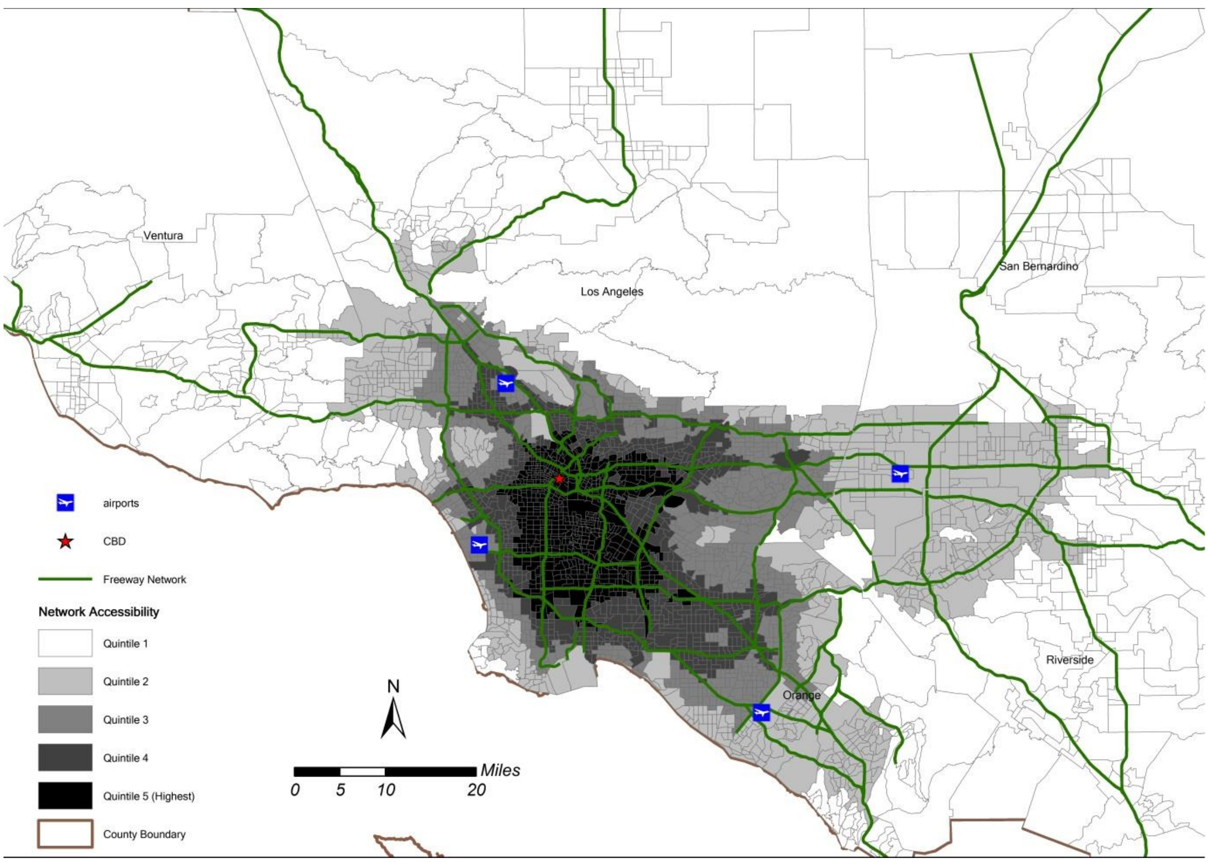
[22]Only three census tracts in the LA region are assigned additional links.

5.3 Results

We discuss descriptive results on our accessibility measures in the first part and regression results on employment center growth in the second part.

5.3.1 Descriptive results on accessibility

Map 5-3 shows 2003 network accessibility, coded in quintiles of normalized accessibility, from lowest (white) to highest (black). We calculate network accessibility for every census tract as the sum of travel times to all other census tracts in the network. Similar to our previous analysis, the highest level of accessibility is in the central core of the network, and covers a rather large area.



Map 5-3 Highway Network Access 2003 (Free Flow)

Table 5-1 gives network access values for the entire region, and for all the centers. To calculate the centers' network accessibility, the component hexagons of each center are located in the 2000 census tract geography. We then calculate the network access of each component hexagon of each center as the area-weighted sum of the access value of all the census tracts intersected with each hexagon. Since hexagons are of regular size, each center's network access reflects the average value of network access of its component hexagons. Table 5-1 shows that network access for center locations are on average much higher than the region as a whole, while the standard deviation is much smaller. The results indicate that centers are generally located in areas of high network access.

Table 5-1 Network access, total region and employment centers

	Mean	Median	SD	Min	Max
Total region	195.26	206.40	112.74	1	448.02
Centers in 2005	229.85	232.95	87.64	13.63	418.46

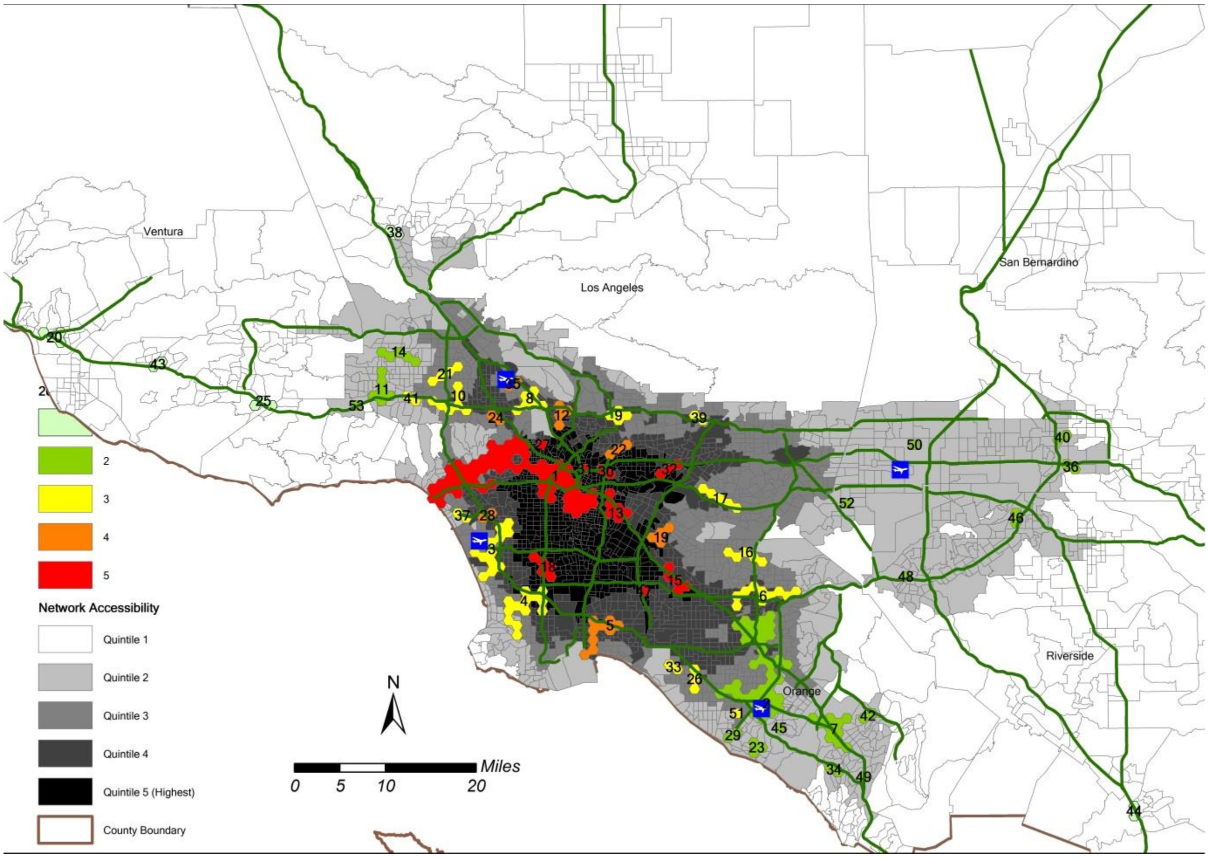
Map 5-5 illustrates the network accessibility of centers identified in 2005 divided by quintiles, using the network accessibility pattern of Map 5-3 as the background. The figure also shows that only a few centers are located in areas of low network accessibility. The largest center, the LA downtown-Santa Monica corridor, has the highest level of accessibility, while the second largest center, Anaheim-Santa Ana- Costa Mesa-Irvine, has a relatively low level of network accessibility due to its location further from the core of the region (and the core of the network). Most of the centers located in the lowest two quintiles of regional accessibility tend to locate near freeway intersections or along the freeway links.

5.3.2 Descriptive analysis on center growth and location

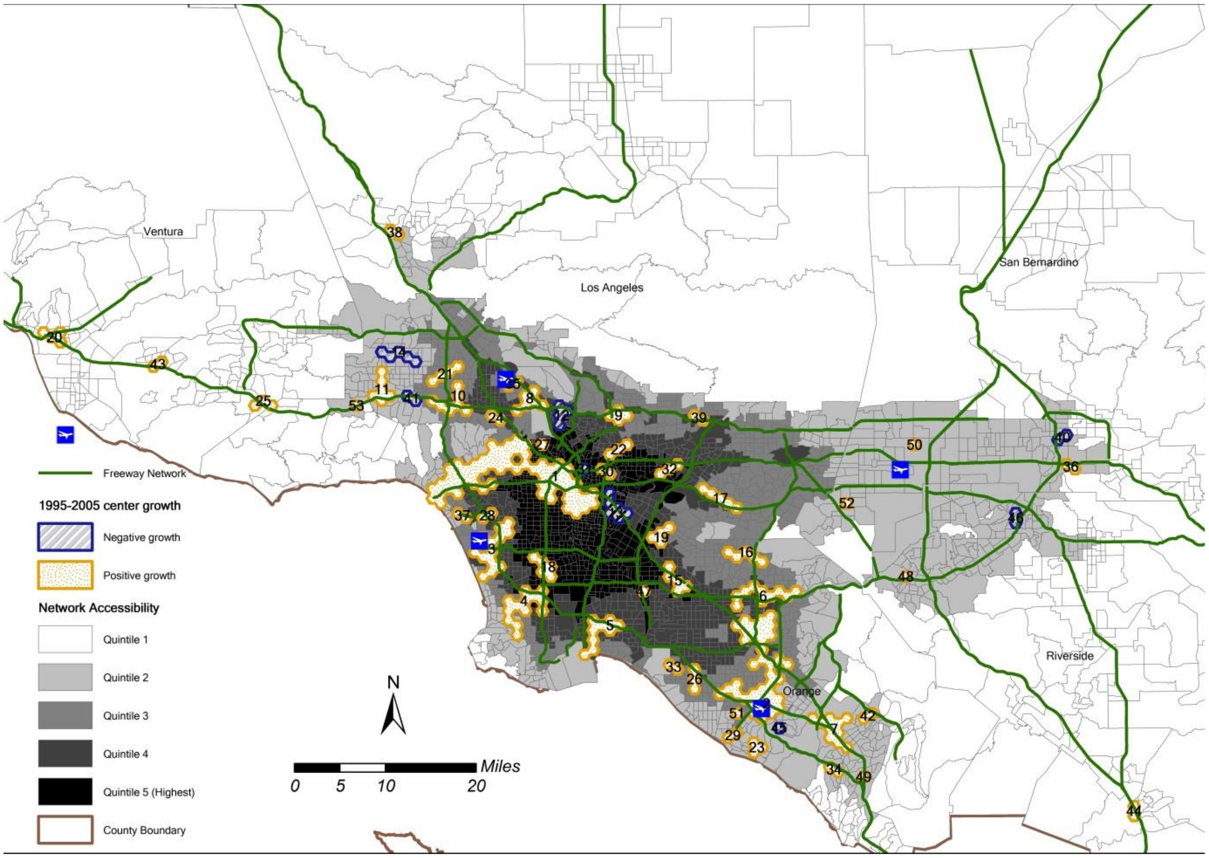
The association of center employment growth with airport location, the CBD, and road network access is shown for 1995-2005 in Map 5-5. Proximity to the airport and the CBD is generally associated with positive employment growth, except for the centers 13 and 31. We cannot observe any clear relationship between network accessibility and employment growth/decline within center from the three figures. Some of centers located in the area with highest quintile in terms of network accessibility had net job loss in 1995-2005, while many centers with the lowest rank of network access also had negative employment growth. For other areas with different ranks of network accessibility, there is also a mixture of center growth and decline for the 1995-2005 period.

The relationship between employment growth within centers and the absolute or relative labor force accessibility is also unclear. As indicated by Tables 5-2 and 5-3, the centers that experienced positive employment growth are rather evenly spread across the access quintiles.

The centers that experienced negative growth are concentrated in one quintile – lower for absolute labor force access and higher for relative labor force access.



Map 5-4 Network Accessibility of employment centers identified in 2005, centers' accessibility calculated based on the peak hexagon identified using the 1995 employment data



Map 5-5 Employment growth within centers identified in 2005 and Network Access (1995-2005)

Table 5-2 Employment Growth of Center identified in 2005 (1995-2005) and Absolute Labor Force Accessibility (1995)

	1 quintile	2 quintile	3 quintile	4 quintile	5 quintile
Negative	0	4	1	1	2
Positive	11	7	9	10	8

Table 5-3 Employment Growth of Center identified in 2005 ((1995-2005) and Relative Labor Force

	1 quintile	2 quintile	3 quintile	4 quintile	5 quintile
Negative	1	1	1	4	1
Positive	10	10	9	7	9

5.3.3 Regression Results and discussions (LA region)

This section presents the regression results of the impacts of accessibility on employment center growth. All centers are included in our analysis. Table 5-4 gives variable names and descriptive statistics. As indicated by the employment size and density of centers in the base year, the size distribution of centers is skewed, with a few very large and very dense centers at the top rank. The average distance to CBD is 25 miles. Centers show some spatial proximity to at least one major airport as well as the coast line, on average about 13 miles. Spatial proximity to inter-state highways are ubiquitous for all centers, with an average distance less than 1 mile.

Table 5-4 Variables and descriptive statistics

Variable	Description	mean	median	sd	min	max
E05_95	Ratio change in employment (05/95)	1.38	1.18	0.91	0.54	6.40
Emp95	1995 employment	60,553	19,854	154,110	2,563	1,042,136
den95	1995 density (jobs per acre)	14.80	12.83	8.72	2.00	61.88
Predgr95_05	Predicted growth rate, 1995-2005 (%)	15.20	15.69	3.88	5.05	22.87
NetwkACC	Network accessibility	229.85	232.95	87.64	13.63	418.46
LFACC95	Absolute labor force accessibility(pop95)	642,656	625,690	305,092	81,221	1,279,470
RLFACC95	Relative labor force accessibility(pop95)	138,466	80,149	164,346	7,540	1,111,603
DCBD_mi	Distance to CBD (miles)	25.07	21.65	16.87	0.06	73.77
Dcoast	Distance to coastline (miles)	14.06	11.55	11.33	0.65	47.74
DLAX	Distance to LAX (miles)	28.56	22.49	17.10	1.59	78.29
DNrstAP	Distance to the nearest airport (mile)	13.07	12.23	9.76	1.20	50.17
DHwy_mi	Distance to nearest Hwy (miles)	0.65	0.48	0.58	0.01	2.65

Table A-5-3 (in Appendix 5) gives pairwise correlations for the variables used in the regression. We use natural log forms for the employment size and density of centers in the base year, as well as for measures of distance to CBD, to the coast, to LAX and to the nearest highway because of their distribution. The correlation matrix indicates that the employment growth ratio is highly correlated with centers' initial density, while the growth ratio of centers also indicates a high correlation with the centers' employment size in the base year (1995). Distance to the CBD is highly correlated with proximity to LAX because of their relative geographic proximity in the region. Network access and absolute labor force access are highly correlated with each other and negatively correlated with distance to CBD, as accessibility tends to decline from the core area to the periphery. Relative labor force access is not correlated with the other two accessibility measures, but is highly correlated with initial employment size.

5.3.3.1 Model results

We estimated models for 1995 – 2000, 2000 – 2005, and 1995 – 2005. We show only the results for 1995 – 2005. We start with a base model that includes only the control variables: employment size and density of the base year (1995), predicted growth rate of 1995-2005, distance to the LA main center, and distance to the coastline. We then added accessibility measures in three groups: airport access, highway access and labor force access. Results are summarized in Table 5-5.

Overall the regression models explain more than 75% to 80% of the variance in centers' growth in the 10-year period. Most of the explanatory power of the model is attributable to the high correlation between initial levels of centers' density and subsequent growth rate. All the regression models show that the coefficients for the initial employment density of centers is always negatively significant, implying that lower density centers are associated with higher growth rate. The coefficients for the initial employment levels of centers are negatively significant in the labor force accessibility model and full model. Contrary to the previous Giuliano et al (2009) results, the shift share variable coefficient is not significant. We suspect this is due to data problems within NETS, as growth rates across industry sectors within NETS are not consistent with growth rates calculated from other employment data sources (e.g. CBP). See Section 3.2.2.

Regarding access variables, coefficients for distance to LAX is positively significant in the second model as well as the full model, implying that proximity to LAX is associated with slower employment growth. In general this is consistent with slower growth within the regional core, where the highest density centers are located. However, as implied by the second model, proximity to other airports is positively correlated with centers' employment growth. The coefficients for distance to CBD variable is not significant in any regression model, nor is the coefficients for distance to the coast. Moreover, neither network access nor absolute labor force has any significant influence on centers' growth, which implies rather ubiquitous access to highways and labor forces across the region. However, there is evidence that relative labor force accessibility plays a significantly positive role in centers' growth.

Table 5-5 Base model and groups of access measures, growth of centers in 2005, 1995-2005

lnE95_05	Base (1)		Add distance to Airport (2)		Add Hwy Acc (3)		Add Labor Acc (4)		Add all Acc measures (5)	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
lnE95	-0.02	-0.45	-0.01	-0.24	-0.02	-0.46	*-0.11*	-2.01	**-0.11**	-2.19
lnD95	**-0.65**	-6.03	**-0.67**	-6.38	**-0.66**	-5.93	**-0.59**	-5.61	**-0.62**	-6.65
predgr9505	0.01	0.84	0.01	0.95	0.008	1.01	0.005	0.66	0.006	0.74
DCBD_mi	-0.002	-1.18	-0.003	-1.31	-0.002	-0.98	-0.004	-1.6	-0.01	-1.58
lnDCoast	-0.007	-0.33	-0.02	-0.57	-0.003	-0.15	-0.02	-0.73	-0.01	-0.38
lnDLAX			<u>0.08</u>	1.83					<u>0.10</u>	2
DNrstAP			*-0.01*	-2.07					-0.005	-1.67
lnDHWY					0.019	0.86			0.02	0.8
NetwkACC					-1.45E-04	-0.21			0.007	1.18
LFACC							-1.49E-07	-1.07	-1.66E-06	-1.33
RLFACC							**6.52E-07**	3.15	**6.61E-07**	3.37
_cons	**2.03**	4.91	**1.86**	5.48	**2.08**	4.61	**2.98**	4.84	**2.81**	5.22
Adj.R2	0.75		0.78		0.76		0.78		0.81	

Bold=P<0.01; Italic=P<0.05 & P>0.01; Underline=P<0.1 & P>0.05

5.3.3.2 A note on predicted growth

The lack of significance of the predicted growth variables is inconsistent with both expectations and prior research. If different industry sectors have different growth rates, the predicted growth variable coefficient should be significant and positive. One explanation for the insignificance of this variable, as discussed in chapter 2, might be related with the quality of NETS datasets at the sector-level, which is not comparable with the federal governments' database such as the CBP (see discussion in chapter 2).

Another explanation, however, would be related to the difference in the industry level employment growth rate between center and non-center locations. Table 5-6 shows that though most of sectors show little difference in terms of employment share between centers and the region as a whole, many sectors show different growth rates between center locations and non-center locations during the 1995-2005 period. For example, service, retail trade and FIRE sectors, which take large employment shares, grow much faster outside centers than within centers, while the manufacturing sector, which takes the third largest employment share, experience faster employment decline within centers than outside centers.

Moreover, Table 5-6 also indicates that the contribution of sector level employment growth to aggregate employment growth for some sectors differs between centers and the whole region. For example, the contribution of retail sector for employment growth within centers is much lower than that for regional employment growth, while the contribution of transportation sector for centers' growth is much higher than for the region as a whole, though this sector takes a relatively small employment share. It is difficult to unravel the many differences in industry mix across centers or inside vs outside centers with 2 digit classifications. However, the unreliability of our data precludes more detailed analysis.

Table 5-6 Sectoral employment share and growth rates within centers and for the LA region (1995-2005)

Industry	Sectoral employment share in 1995		Sectoral employment growth rate, 1995-2005 (%)		Contribution of sectoral employment change to aggregate employment change (%)	
	Within Centers	Region	Within Centers	Region	Within Centers	Region
Agriculture	0.43	1.05	36.54	19.66	1.35	1.27
Mining	0.18	0.19	-25.12	-27.32	-0.38	-0.32
Construction	2.87	4.20	30.74	29.19	7.51	7.55
Manufacturing	17.59	15.78	-2.98	-1.25	-4.47	-1.21
Transportation	5.22	5.46	27.09	17.12	12.04	5.76
Wholesale Trade	7.55	6.89	4.52	11.62	2.91	4.93
Retail Trade	12.08	15.41	15.03	21.37	15.47	20.27
FIRE	9.82	7.83	12.35	21.33	10.33	10.28
Service	40.51	40.35	13.89	17.84	47.96	44.32
Public Administration	3.76	2.84	22.69	40.85	7.28	7.14

5.4 Comparative study

This section presents our results for the pooled sample of Los Angeles and San Francisco. As noted earlier, other metropolitan areas do not have a sufficient number of employment centers to allow quantitative analysis. In this case we pool the centers from Los Angeles and San Francisco.

5.4.1 Employment centers characteristics in the San Francisco region

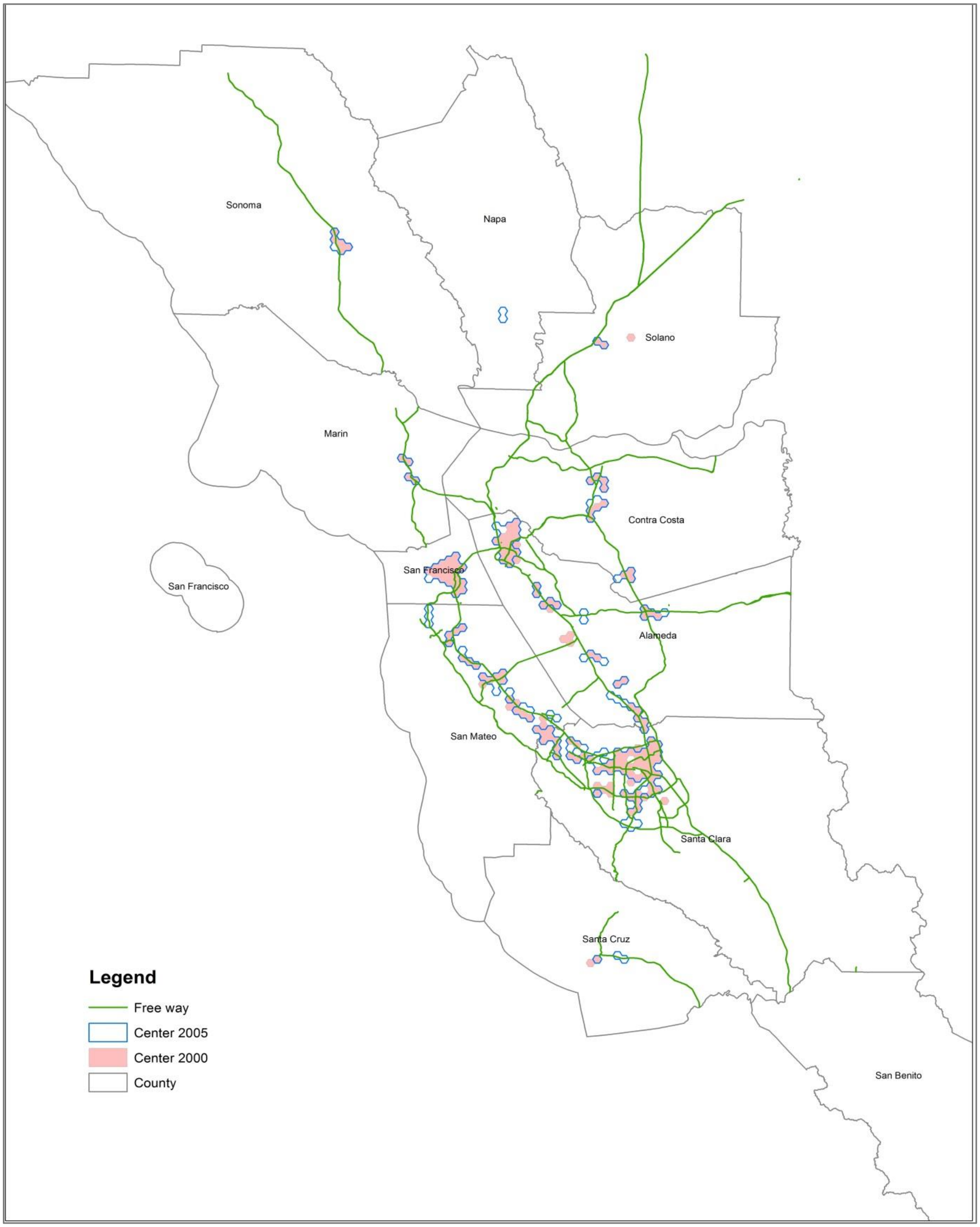
As shown in Chapter 3, San Francisco has a very different physical geography from that of Los Angeles. Development is constrained by the bay in the center and hilly terrain along the west and east perimeters of the bay. The physical geography has forced development along corridors on either side of the bay. Population, employment, and the transportation system are concentrated in these corridors. Thus San Francisco presents a quite different case for exploring the growth of employment centers.

5.4.1.1 The San Francisco employment centers

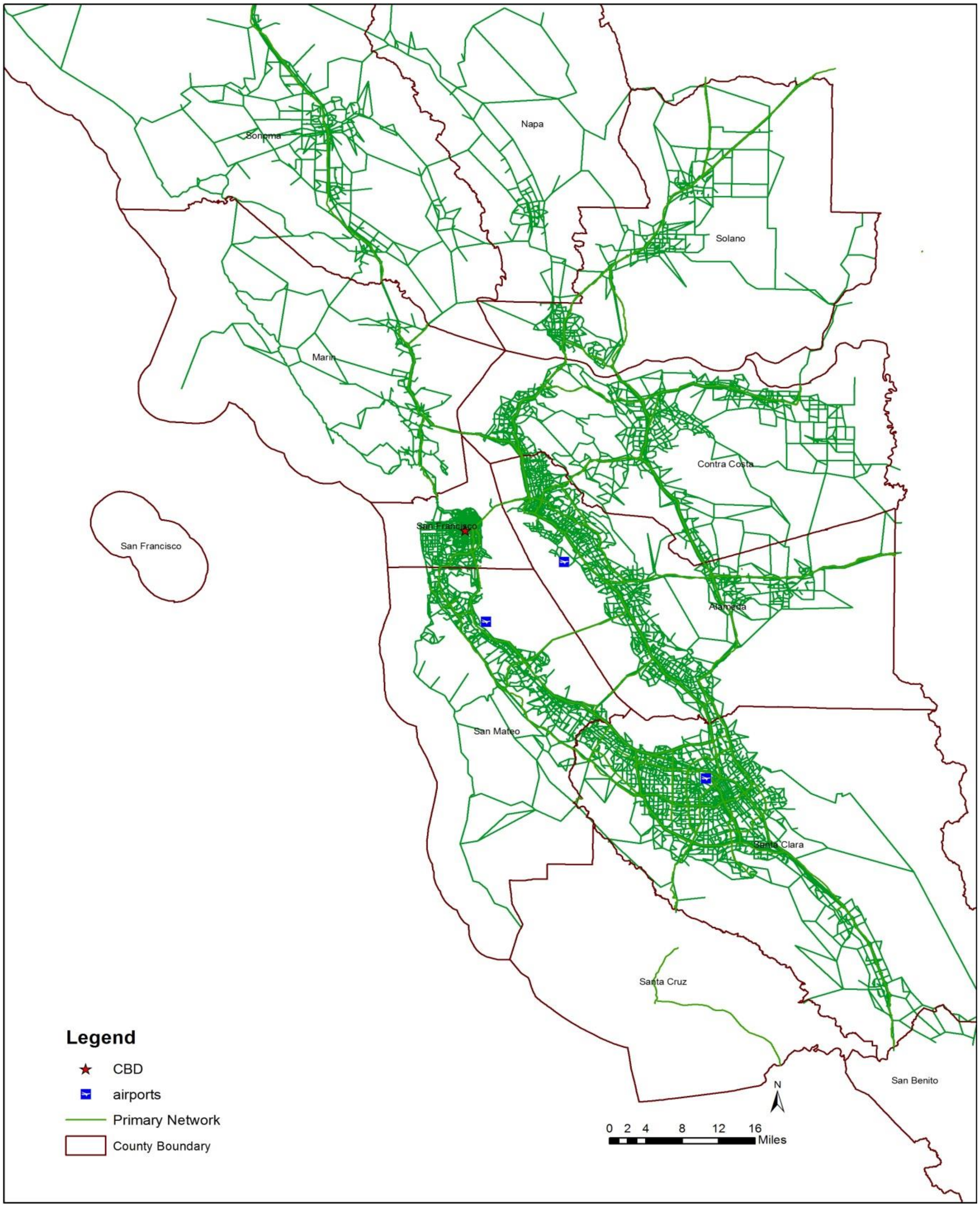
As of 2005 there are 30 employment centers in the region. Like Los Angeles, the San Francisco centers follow a 'rank-size' effect. The largest and second largest centers have a 2005 employment of more than half a million jobs, the next largest centers have more than 200,000 jobs each and subsequent smaller centers have less than 100,000 jobs per center. The smallest centers had just a little over 10,000 jobs. The largest center in terms of area is San Jose downtown, which contains nearly 32000 acres, close to the area of the largest center in the Los Angeles area. On the other hand, the smallest centers identified in 2005 encompass only 640 acres, which means a single hexagon center. San Francisco has fewer single hexagon centers than Los Angeles (2 and 10 respectively). There is also a large variation in employment density; San Francisco downtown has an average employment density of around 50 jobs per acre in 2005 (about twice that of the LA downtown), while the least-dense centers have an employment density of about 8.02 jobs per acre, just above the density cutoff.

As with Los Angeles, we overlay the 2005 employment centers on the 1995 employment centers to show how the centers have changed over time. Figure 5-6 shows how most centers are distributed as "strings of beads" along the corridors surrounding the bay, with the two "strings" meeting at the large San Jose center.

Table A-5-4 (in Appendix 5) shows that all but three centers had positive growth during 1995-2005. Centers that lost employment are Center 1 (San Francisco Downtown), Center 14 (North Mountain view), and Center 25 (San Rafael). The largest employment gainer is Center 16 (Belmont), while the largest loser is Center 25 (San Rafael). It is notable that San Francisco downtown lost employment between 1995 and 2005, while Los Angeles downtown gained employment.



Map 5-6 2005 centers and 1995 centers in San Francisco



Map 5-7 2005 Transportation Network (Source: Association of Bay Area Governments)

5.4.2 Road network of 2005 in the San Francisco region

The transportation networks employed for this study are network files of San Francisco 2005 provided by Association of Bay Area Governments. The original file includes detailed information for each link and node, such as flow speed on each link. To build the travel time OD matrix, the centroid of the peak hexagons of centers and those of census tracts are assigned to the closest network nodes. Since San Francisco network data is not complete for Santa Cruz and San Benito County, there are a few census tracts without any nodes falling inside. Therefore, we manually create additional “centroid connectors” from their centroids to the nearest nodes and assigned an average speed to these connectors. Then, we calculated travel time of these additional links by dividing the length by speed.

Map 5-7 shows the highway network for the region. It too is highly concentrated around the perimeter of the bay, with a secondary corridor further to the east. The three major airports (San Francisco, Oakland, and San Jose) are also located in the main corridors.

5.4.3 Results

This section is two-fold. In the first part, we discuss descriptive results on our accessibility measures. Then, we describe the regression results on employment center growth.

5.4.3.1 Accessibility descriptive analysis

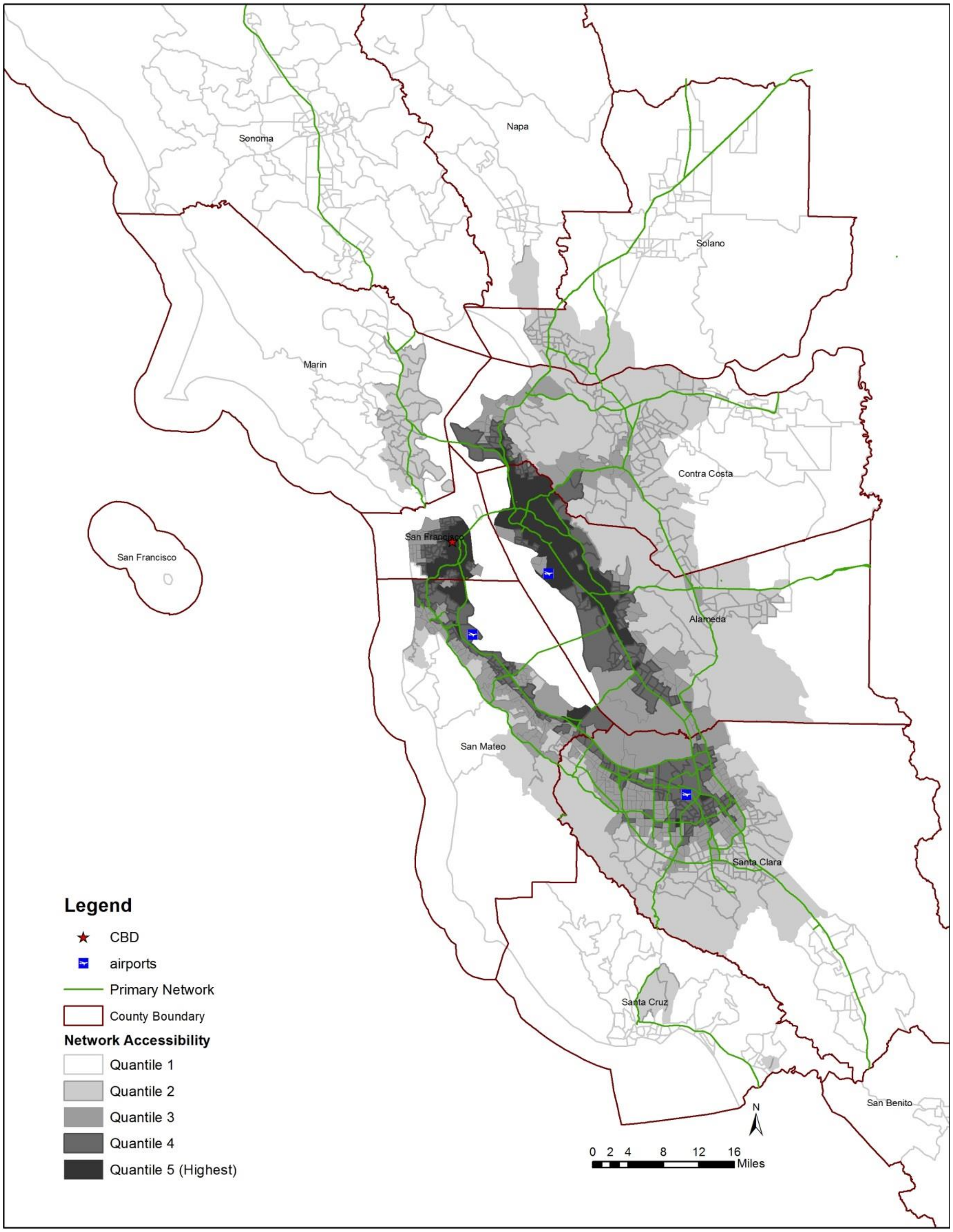
Map 5-8 shows 2005 network accessibility, coded in quintiles of accessibility from lowest (white) to highest (black). Network accessibility for each census tract is calculated for every census tract using travel times to all other census tracts in the network. The highest level of accessibility is in the central core of the network, which is concentrated around the San Francisco Bay Bridge and covers an area surrounded by I-80 and I-280. An area that covers Oakland and San Leandro bordering I-880 on the west and I-580 on the east also has a high level of accessibility. Accessibility concentrates along the bay corridors, with a secondary concentration in San Jose. Visually, the consistency of employment center location and transport access is striking.

Table 5-7 gives network accessibility values for the entire region, and for all centers. It shows that centers have almost 20 percent higher levels of mean network accessibility than the region as a whole. The standard deviation is much smaller for centers, which implies few centers are located in areas with low network accessibility. Note that the average freeway network accessibility of San Francisco is much higher than that of the Los Angeles region for both total region and centers. This is due to the greater spatial concentration of San Francisco. The difference in freeway network access between centers and total region is also greater in the San Francisco region compared to that of the Los Angeles area.

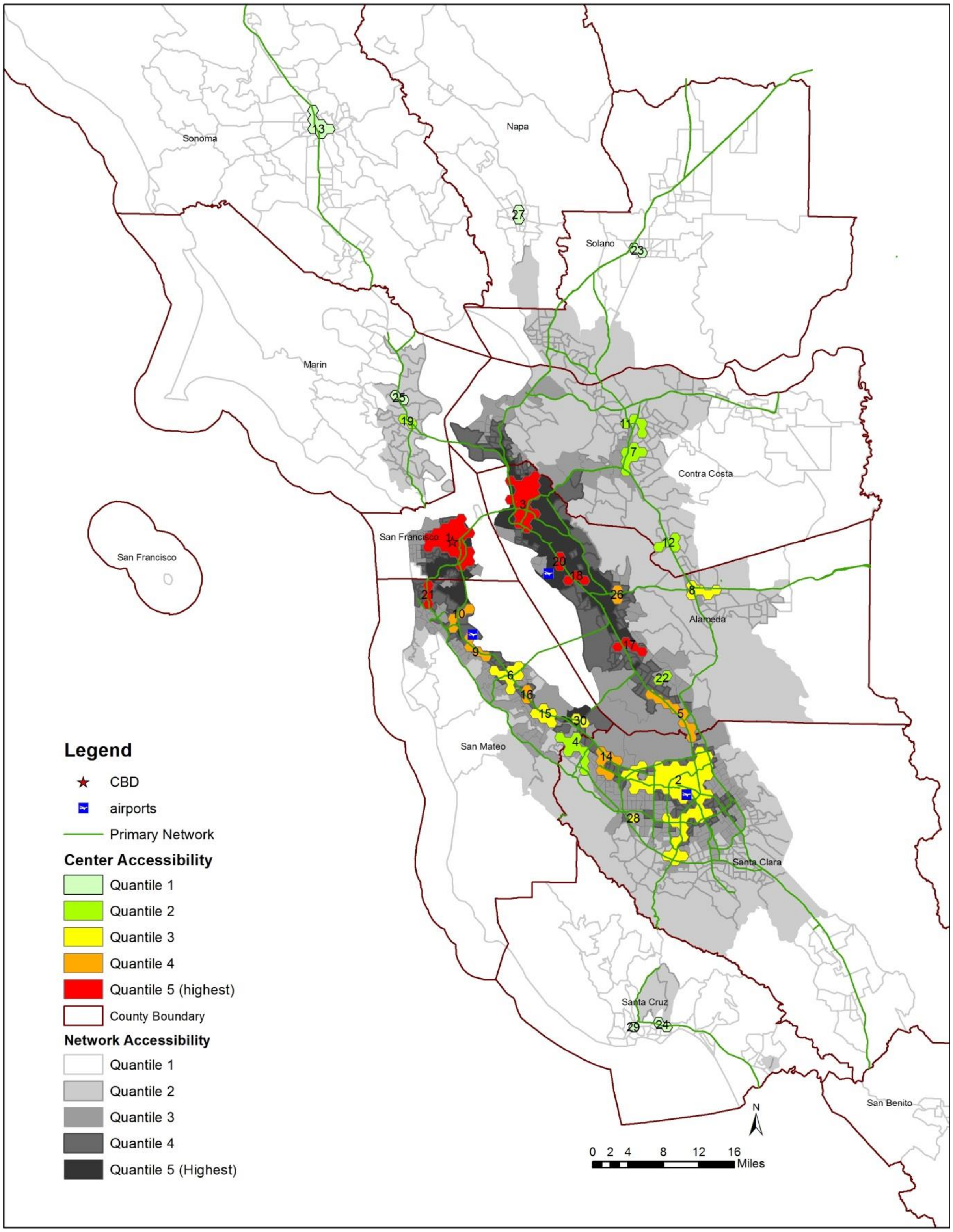
Table 5-7 Freeway network access, total region and 2005 employment centers

	Mean	Median	Std. Dev.	Min	Max
Total region	200.03	222.26	79.64	0.87	343.15
2005 Centers	236.22	253.61	62.71	40.17	315.99

Map 5-9 overlays the centers identified in 2005 on the network accessibility map of Map 5-8. Center accessibility is also divided into quintiles, from lowest (1) to highest (5). The largest center, SF downtown, and the Oakland downtown center have the highest level of accessibility. The second largest center, San Jose downtown, has a relatively low level of accessibility, comparable to the Anaheim/Santa Ana/Costa Mesa/Irvine center in Los Angeles. As in the Los Angeles region, the few centers located in areas with relatively low network accessibility tend to locate near freeway intersections or along the freeway links.



Map 5-8 Highway Network Access 2005 (Free Flow)



Map 5-9 2005 Center Freeway Network Accessibility, 2005 Network

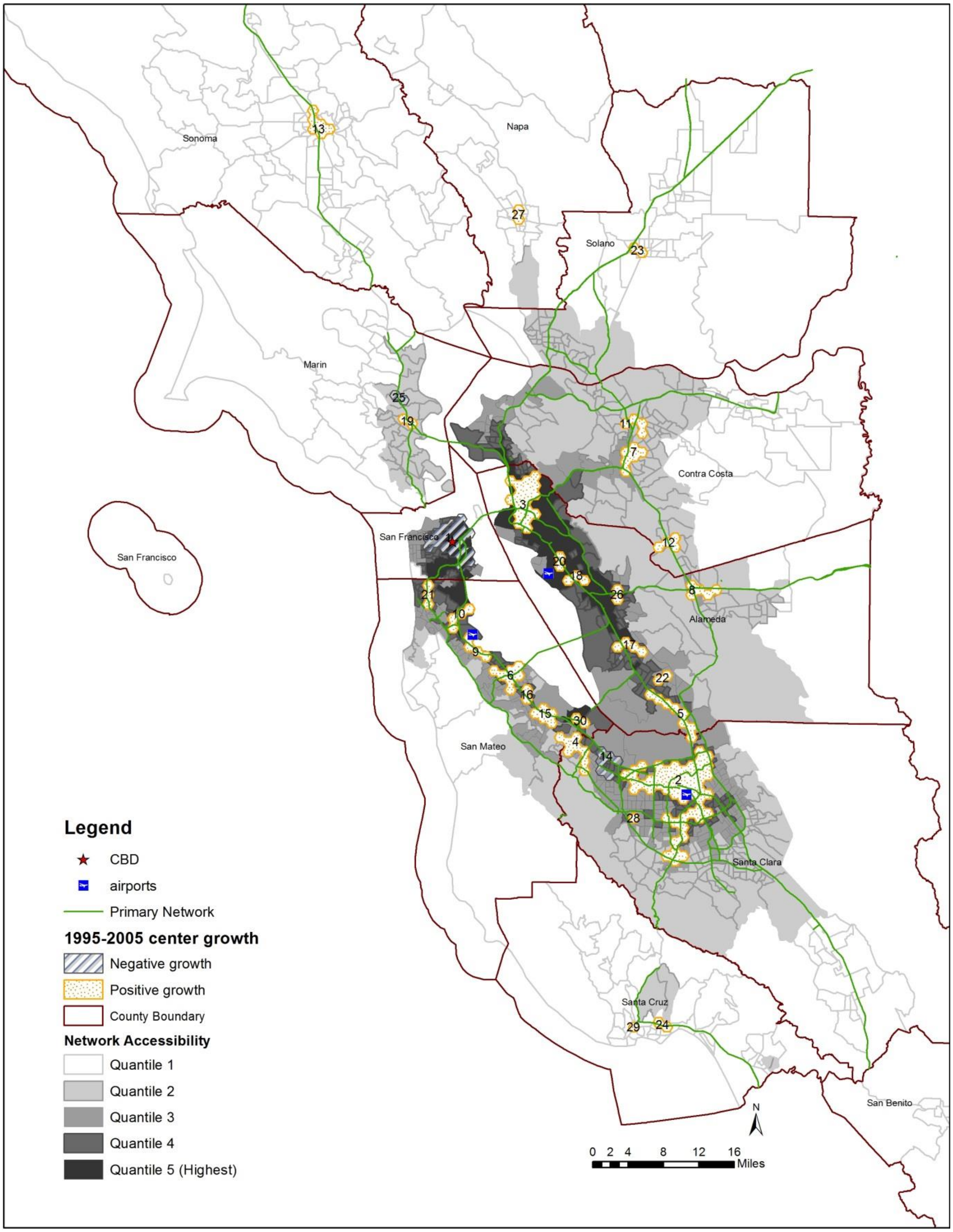
Map 5-10 shows the 2005 centers' growth and decline between 1995 and 2005 and their relative location with respect to airport and network accessibility, which is coded into five categories from lowest (1) to highest (5).Tables 5-8 and 5-9 show the distribution of positive and negative growth centers across quintiles of absolute and relative labor force access respectively. As in the case of San Francisco, positive growth center are distributed evenly across the access categories, suggesting no relationship of growth to labor force access.

Table 5-8 Growth of 2005 Center (1995-2005) and Absolute Labor Force Accessibility (1995)

Absolute labor force	1 quintile	2 quintile	3 quintile	4 quintile	5 quintile
Negative	1	0	0	1	1
Positive	5	6	6	5	5

Table 5-9 Growth of 2005 Center (1995-2005) and Relative Labor Force Accessibility (1995)

Relative labor force	1 quintile	2 quintile	3 quintile	4 quintile	5 quintile
Negative	0	1	0	1	1
Positive	6	5	6	5	5



Map 5-10 Employment Growth of 2005 Centers (1995-2005) and Network Access

5.4.3.2 Regression results

In order to expand our discussion on employment center growth to other metro areas, we conduct the same regression analysis by pooling the San Francisco and Los Angeles centers. The pooled sample totals 80 centers: 50 Los Angeles centers and 30 San Francisco centers. We include a regional dummy variable to control for differences between the two regions[23]

Table 5-10 shows the descriptive statistics of employment centers in the San Francisco region. Table 5-11 presents the summary statistics of centers for both regions. There are some interesting differences in descriptive statistics. As demonstrated by the accessibility figures, San Francisco centers have higher average level of network access than Los Angeles centers.

Table 5-10 Variables and descriptive statistics (San Francisco)

Variable	Description	Mean	Median	SD	Min	Max
E05_95	Ratio change in Emp(05/95)	1.26	1.22	0.41	0.84	3.23
Emp95	2000 Emp	66585.07	24092.5	136912.8	7697	605321
Den95	2000 density (jobs per acre)	12.34	11.14	7.31	6.01	47.29
Predgr05_95	Predicted growth rate, 1995-2005 (%)	13.64	14.42	4.57	1.42	19.98
NetwkACC	Network accessibility	216.99	241.5	76.41	40.17	315.99
LFACC95	Absolute labor force accessibility (pop95)	964840.4	1077411	329529.5	186066.7	1339313
RLFACC95	Relative labor force accessibility(pop95)	175258	171500.5	23304.21	141773.6	240356.5
DCBD_mi	Distance to CBD (miles)	25.4	23.27	14.5	0	61.02
DSFO	Distance to SFO (miles)	23.32	21.55	14.89	1.46	59.16
DNrstAP	Distance to the nearest airport (mile)	15.2	12.02	12.72	1.46	55.99
D_Hwy_mi	Distance to highway (miles)	0.85	0.34	1.97	0.11	10.94

[23] The number of centers in Los Angeles area and San Francisco area are different across the time periods. Please refer to previous chapter for the specific numbers of Los Angeles centers and San Francisco centers.

Table 5-11 Variables and descriptive statistics (Los Angeles, San Francisco)

Variable	Description	Mean	Median	SD	Min	Max
E05_95	Ratio change in Emp(05/95)	1.34	1.19	0.77	0.54	6.40
Emp95	2000 Emp	62733	19854	147305	2563	1042136
Den95	2000 density (jobs per acre)	13.91	12.03	8.28	2.00	61.88
Predgr05_95	Predicted growth rate, 1995-2005 (%)	14.63	15.42	4.18	1.42	22.87
NetwkACC	Network accessibility	203.11	216.01	88.04	15.74	390.665
LFACC95	Absolute labor force accessibility(pop95)	759108	750692	348828	81221	1339313
RLFACC95	Relative labour force accessibility(pop95)	151764	156635	132802	7540	1111604
DCBD_mi	Distance to CBD (miles)	25.19	22.66	15.96	0.00	73.77
DLAX/SFO	Distance to LAX/SFO (miles)	26.67	22.47	16.44	1.46	78.29
DNrstAP	Distance to the nearest airport (mile)	13.84	12.23	10.90	1.20	55.99
D_Hwy_mi	Distance to highway (miles)	0.72	0.40	1.26	0.01	10.94

We also generated pairwise correlations among variables (see Table A-5-5 in Appendix 5). As with the Los Angeles sample, the employment growth ratio is highly correlated with the centers' initial density. In addition, distance to the CBD is highly correlated with proximity to the main airport (LAX/SFO). The correlation coefficient between network accessibility and absolute labor force accessibility is consistently very high. Since these two variables can cause a multicollinearity issue, we drop the absolute labor force variable in our regression model.

We estimate a base model that includes only the (1) control variables and add (2) airport access, (3) highway access, (4) labor force access, and (5) all access measures as we did for Los Angeles area. Table 5-12 gives results (beta coefficients and t-values). Results are similar to those for the Los Angeles centers, with only the center base year density variable consistently significant and of the expected sign. Despite the greater number of observations, the explanatory power of the model goes down. As in the Los Angeles estimations, the predicted growth variable coefficient is not significant, nor is access to the CBD or the main airport. Distance to the nearest airport is marginally significant, likely due to the close proximity of most centers to airports in San Francisco. Given the many unmeasured differences between Los Angeles and San Francisco, it is surprising that the dummy variable coefficient is not significant in the final panel of the table (the coefficient is significant, but becomes insignificant when all the access variables are added).

Results on the highway access measures and network accessibility measures are consistently insignificant, indicating that being close to a particular highway and network accessibility does not have any impact on employment center growth. These results are consistent with previous studies. The coefficients of relative labor force become insignificant after adding San Francisco centers. This may be the result of the high accessibility of most San Francisco centers. If most centers are located in high accessibility places, then different growth rates must be explained by some other factor. Based on the final panel of the table, initial employment density and distance to nearest airport are the only significant variables for the model of both regions.

To sum up, the regression results for combined samples of Los Angeles and San Francisco area are different from the regression result base only on Los Angeles area. While the initial employment and distance to the most major airport, as well as relative labor force are statistically significant for the employment growth model based only on Los Angeles area, they become statistically insignificant after adding the samples of San Francisco centers. Based on the expanded model, density is still statistically significant in explaining employment growth/decline of centers, but the only access variable that remains significant is proximity to the nearest airport. Extending our analysis to San Francisco generates more questions than answers about employment center growth, and indicates that our results for Los Angeles are not generalizable to other regions.

Table 5-12 Regression Results (Los Angeles+San Francisco)

lnE95_05	Basic		Add distance to Airport		Add Hwy Acc		Add Labor Acc		Add all Acc measures	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
lnE95	9.1E-05	0	1.5E-03	*0.06*	-1.4E-03	-0.06	-9.0E-04	-0.02	-3.9E-03	-0.1
lnD95	**-5.5E-01**	**-4.64**	**-0.56**	**-4.85**	**-0.55**	**-4.63**	**-5.5E-01**	**-4.59**	**-0.56**	**-4.77**
predgr9505	2.9E-03	0.49	0.01	0.86	3.62E-03	0.58	2.9E-03	0.48	4.9E-03	0.77
DCBD_mi	-8.9E-04	-0.64	-7.9E-05	-0.04	3.07E-05	0.02	-8.9E-04	-0.64	-6.3E-04	-0.23
LA	1.1E-01	1.86	0.09	1.67	0.13	1.76	1.1E-01	1.84	0.08	1.14
lnDLAX/SFO			0.02	0.54					0.02	0.54
DNrstAP			*-0.01*	*-2.62*					*-0.01*	*-2.58*
lnDHWY					4.27E-03	0.23			6.0E-04	0.03
NetwkACC					2.78E-04	0.52			-1.7E-04	-0.32
RLFACC							1.00E-08	0.04	6.4E-08	0.25
_cons	**1.50**	**4.56**	**1.49**	**4.26**	**1.43**	**4.51**	**1.51**	**3.71**		**3.62**
Adj.R2	0.58		0.60		0.58		0.58		0.60	

5.5 Summary

In this chapter, we firstly followed our previous research and examined the impacts of network accessibility on centers' employment growth within the LA region. Our analysis show no evidence that network accessibility and absolute labor force accessibility would impact on centers' employment growth. This could be attributed to the ubiquitous highway access across the LA region. However, we did find that relative labor force accessibility measures, which account for competition for labor force among centers, contributes positively to centers' growth. There is also evidence that centers locating closer to airports other than LAX tend to grow faster. We also found that centers with larger employment size or higher employment density in the beginning year tend to experience slower growth in the study period, while their shift-share of regional growth plays no role. We explained this inconsistency with our previous study as a result of the differences in the sector level growth rates between centers and non-center locations for the LA region during the study period.

We then extend our previous analysis by adding San Francisco centers to examine whether the determinants of centers' growth are consistent across metropolitan areas. The regression results on the combined samples of Los Angeles and San Francisco regions, however, are contrary to our expectation. We found that except for the initial employment density, distance to the nearest airport is the only significant access variable in explaining centers' growth. Relative labor force accessibility, the other significant accessibility variable in the LA model, turn out to be insignificant in the combined model. This result casts some doubts on the consistency of the impacts of accessibility on centers' growth across different metro areas. In the future, we may need further studies on other metro areas to generalize our findings.

Chapter 6 Conclusion

In this project, we examined the relationship between accessibility and growth of centers across different metropolitan areas. We begin with cross-checks on the consistency and reliability of NETS dataset and explained how our research goal was adjusted in light of the problems we have discovered. We then developed methods for employment center identification by adjusting from the classic Giuliano and Small (1993) methods to accommodate for the differences across metropolitan areas. The basic characteristics of employment centers across the four study regions are also compared over different time periods. Finally, we conducted a separate regression analysis on employment center growth for the LA region and for the pooled samples of the LA and SF regions. We then compared the regression results on the determinants of centers' growth in different regions. Our general findings may be summarized as the following.

(1) What we learned about the NETS data

Although the NETS dataset has been widely used in urban and regional economic analysis, few studies have cross-checked the original NETS dataset with other government-issued data sources. In this study, we conducted two cross-checks of the NETS dataset, one with the SCAG data at the micro-geographic level within the LA region for the year of 2000, and the other with the CBP dataset at both the county-level and CSA level for the 1993-2009 periods. We discovered from these two data checks that NETS systematically differs from the CBP dataset at the county or CSA level in terms of either aggregated employment/establishment counts or sectoral employment shares. There are differences across sectors, time periods, and metro areas. We also find that the over-counting of NETS employment is not attributable to the different techniques used by NETS for capturing small firms, because the over-counting of establishments exists in all size categories. Because we could not discover other systematic differences in the way NETS data are collected, we have no way to correct for the differences between NETS and other data sources. Thus, the inconsistency in the NETS dataset limited any analysis related with the industrial composition or size distribution of firms.

(2) What we learned about urban spatial structure

In this study, we extend the classic method for subcenter identification by applying the 95% employment densities as the density cutoff to account for the differences in geographic size and magnitude of total employment of different metro areas. Our formal test for polycentricity shows that employment distribution is better explained by the polycentric model than the monocentric model for all the four study areas. The CBD exerts significant, though weaker, influence when the influence of other centers is introduced. However, the Sacramento region seems to be the weakest case for polycentricity. Our later descriptive analysis of centers for different regions also reveals significant differences across different regions in terms of centers' characteristics such as average geographic size, average employment/population density, employment/population ratios and the total employment share within centers. These differences may be attributable to the general differences across metro areas, such as their natural geographic features, geographic size and total employment sizes. However, we did obverse some similarities in centers' characteristics across different regions. For example, we observe a general pattern of rank size distribution in every metro area, with only 1 or 2 centers in the largest category and many centers in the smallest category. We also found that there are only small differences in the average number of jobs per center for all the four regions.

Another important observation is the persistency of urban spatial structure across all the metro areas. We found that even though the locations of lower-rank centers may change, the number of centers in each metro area remain relatively stable over time. Also, average center characteristics for each metro area, such as average size, employment/population density, change very little in different time periods. This result provides some implication on the path-dependent development of centers and evolution of urban forms.

(3) What we learned about the growth of employment centers

In the last chapter, we examined the determinants of employment growth within centers for the LA region and for the pooled samples of Los Angeles and San Francisco region. Following our previous research, we hypothesized that the growth of centers is positively affected by network accessibility, labor forces accessibility and their shift-share of regional industrial growth, and negatively affected by their distance to the CBD, major airports, highways, as well as their initial size and densities. Our regression results for the LA region indicates that network accessibility and absolute labor force accessibility plays insignificant role in centers' growth, while there is some evidence that relative labor force accessibility measures in which the competition for labors among centers are taken into consideration contributes positively to centers' growth. We also found that centers with larger employment size or higher employment density in the beginning year tend to have slower growth in the following period, while centers with spatial proximity to major airports other than LAX tend to grow faster. In addition to data inconsistency, we also explained the insignificance of shift-share variables as a result of differences in the employment share and employment growth rate of various sectors between center and non-center locations. Sectors with large growth rates within centers are not necessarily fast growing at the regional scale.

However, unlike our expectation, the regression results on the pooled samples of Los Angeles and San Francisco regions differ from the results based only on the Los Angeles region. While initial employment density still plays a significant role in explaining positive employment center growth, other factors such as relative labor force accessibility measure and initial employment size do not play a crucial role anymore. Meanwhile, the results for the combined samples of Los Angeles and San Francisco regions show that the nearer the distance to the nearest airport is, the faster the employment centers grow. Hence, we conclude that initial employment density of centers and access to nearest airports are more universal factors explaining employment center growth than other factors in the model while we need further investigation to generalize the results. Our research results are largely consistent with the literature on urban structure and accessibility. California's largest metro areas are polycentric, and the forces of agglomeration continue to be strong. The results of our analysis of employment center growth can be interpreted in different ways. First, once a metropolitan area matures, the factors that explain variation in growth change. Second, the NETS data may be too unreliable to draw any conclusions. More reliable establishment and employment data are required for fine grain spatial analysis of urban spatial trends and their relationship to accessibility.

REFERENCES

Agarwal, A., G. Giuliano, et al. (2012). "Strangers in our midst: the usefulness of exploring polycentricity." The Annals of Regional Science **48**(2): 433-450.

Anas, A., R. Arnott, et al. (1998). "Urban Spatial Structure." Journal of Economic Literature **36**(3): 1426-1464.

Baumont, C., C. Ertur, et al. (2004). "Spatial analysis of employment and population density: the case of the agglomeration of Dijon 1999." Geographical Analysis **36**(2): 146-176.

Craig, S. G. and P. T. Ng (2001). "Using quantile smoothing splines to identify employment subcenters in a multicentric urban area." Journal of Urban Economics **49**(1): 100-120.

Ellison, G. and E. L. Glaeser (1994). Geographic concentration in US manufacturing industries: a dartboard approach, National Bureau of economic research.

Fujita, M. (1989). Urban economic theory: land use and city size, Cambridge university press.

Fujita, M. and J.-F. Thisse (2002). "Economics of agglomeration: Cities, industrial locations, and regional growth." Cambridge: CambridgeUniversityPress.

Garcia-López, M.-À. (2010). "Population suburbanization in Barcelona, 1991–2005: Is its spatial structure changing?" Journal of Housing Economics **19**(2): 119-132.

Garcia-López, M.-À. and I. Muñiz (2010). "Employment decentralisation: Polycentricity or scatteration? The case of barcelona." Urban Studies **47**(14): 3035-3056.

Garcia-López, M.-À. (2012). "Urban spatial structure, suburbanization and transportation in Barcelona." Journal of Urban Economics 72(2): 176-190.

Gilli, F. (2009). "Sprawl or Reagglomeration? The Dynamics of Employment Deconcentration and Industrial Transformation in Greater Paris." Urban Studies **46**(7): 1385-1420.

Giuliano, G., C. Redfearn, et al. (2011). "Network Accessibility and Employment Centers." Urban Studies.

Giuliano, G., C. Redfearn, et al. (2007). "Employment concentrations in Los Angeles, 1980 - 2000." Environment and Planning A **39**(12): 2935-2957.

Giuliano, G. and K. A. Small (1991). "Subcenters in the Los Angeles region." Regional Science and Urban Economics **21**(2): 163-182.

Giuliano, G. and K. A. Small (1999). "The determinants of growth of employment subcenters." Journal of Transport Geography **7**(3): 189-201.

Guillain, R., J. Le Gallo, et al. (2006). "Changes in Spatial and Sectoral Patterns of Employment in Ile-de-France, 1978-97." Urban Studies **43**(11): 2075-2098.

Henderson, V. and A. Mitra (1996). "The new urban landscape: Developers and edge cities." Regional Science and Urban Economics **26**(6): 613-643.

Lee, B. (2006). Urban spatial structure, commuting, and growth in U.S. metropolitan areas. School of Policy, Planning, and Development. Los Angeles, University of Southern California. **Doctor of Philosophy**.

McMillen, D. P. (2001). "Nonparametric employment subcenter identification." Journal of Urban Economics **50**(3): 448-473.

Mills, E. S. (1967). "An aggregative model of resource allocation in a metropolitan area." The American Economic Review **57**(2): 197-210.

Pan, Q. (2003) The Identification of Employment Subcenters in the Houston Metropolitan Area, Research Report, Texas Southern University.

Redfearn, C. L. (2007). "The topography of metropolitan employment: Identifying centers of employment in a polycentric urban area." Journal of Urban Economics **61**(3): 519-541.

Sullivan, A. M. (1986). "A general equilibrium model with agglomerative economies and decentralized employment." Journal of Urban Economics **20**(1): 55-74.

Walls and Associates (2008), National Establishment Time-Series (NETS) Database: Database Description, Oakland, California, available at https://msbfile03.usc.edu/digitalmeasures/cswenson/intellcont/NETS%20Database%20Description2008-1.pdf

Zhang, Y. and K. Sasaki (1997). "Effects of subcenter formation on urban spatial structure." Regional Science and Urban Economics **27**(3): 297-324.

Zhang, Y. and K. Sasaki (2000). "Spatial structure in an open city with a subcenter." The Annals of Regional Science **34**(1): 37-53.

APPENDIX FOR CHAPTER 2

2.1 Comparison of aggregate employment and establishments

Table A-2-1 Comparison of county-level NETS and CBP (N=58)

Year	NETS Total emp	CBP Total emp	Avg County level Diff.(%)	NETS Total est	CBP Total est	Avg County level Diff.(%)
1993	14,040,305	10,712,274	45.45	1,141,194	730,687	52.10
1994	14,334,002	10,609,632	47.67	1,144,099	729,762	51.74
1995	14,354,342	10,937,576	45.10	1,142,183	732,726	52.33
1996	14,577,538	11,119,083	44.11	1,178,529	746,615	51.86
1997	14,827,851	11,553,666	41.05	1,172,756	757,849	49.31
1998	14,978,252	12,005,269	37.40	1,166,444	765,719	48.09
1999	15,365,300	12,333,478	37.82	1,147,216	771,387	45.91
2000	16,091,248	12,838,178	40.61	1,220,819	785,605	50.16
2001	16,860,250	13,146,208	45.08	1,419,631	796,602	71.47
2002	16,463,408	12,807,012	42.95	1,444,962	816,635	65.78
2003	16,121,230	12,925,512	38.11	1,456,328	823,323	67.10
2004	16,007,196	13,152,029	35.93	1,497,302	836,174	69.14
2005	16,217,827	13,205,638	36.74	1,565,287	857,533	72.70
2006	16,400,181	13,618,236	34.82	1,619,239	874,464	74.99
2007	16,452,881	13,574,339	34.23	1,733,856	890,313	81.97
2008	16,670,180	13,551,619	40.10	1,924,443	877,009	104.60
2009	15,598,646	12,636,839	43.73	1,750,379	855,220	98.68

Table A-2-2 Percentage Employment / Establishment Difference for Los Angeles CSA

Year	NETS Total emp	CBP Total emp	CSA level Diff.(%)	NETS Total est	CBP Total est	CSA level Diff.(%)
1993	7,160,440	5,453,444	31.30	565,271	342,976	64.81
1994	7,275,372	5,344,969	36.12	566,656	342,063	65.66
1995	7,239,139	5,510,634	31.37	561,523	344,281	63.10
1996	7,274,300	5,540,454	31.29	578,296	350,323	65.08
1997	7,364,091	5,737,288	28.35	578,193	355,131	62.81
1998	7,393,345	5,945,796	24.35	574,250	358,731	60.08
1999	7,561,370	6,103,210	23.89	564,453	362,347	55.78
2000	7,891,174	6,328,515	24.69	607,382	369,242	64.49
2001	8,327,460	6,419,657	29.72	713,185	375,639	89.86
2002	8,159,973	6,339,556	28.72	727,237	387,107	87.86
2003	8,023,788	6,423,414	24.91	734,927	393,645	86.70
2004	8,009,876	6,586,393	21.61	758,830	402,186	88.68
2005	8,080,813	6,575,811	22.89	791,926	414,562	91.03
2006	8,169,495	6,782,312	20.45	821,229	424,169	93.61
2007	8,229,545	6,734,012	22.21	884,653	433,294	104.17
2008	8,370,768	6,720,976	24.55	991,059	427,313	131.93
2009	7,733,442	6,283,786	23.07	883,351	417,578	111.54

Table A-2- 1 Percentage Employment / Establishment Difference for San Francisco CSA

Year	NETS Total emp	CBP Total emp	CSA level Diff.(%)	NETS Total est	CBP Total est	CSA level Diff.(%)
1993	3,392,357	2,731,876	24.18	267,739	179,508	49.15
1994	3,434,899	2,716,353	26.45	261,000	179,795	45.17
1995	3,474,595	2,798,159	24.17	263,952	181,286	45.60
1996	3,588,810	2,890,641	24.15	273,898	184,933	48.11
1997	3,681,294	3,036,741	21.23	271,694	188,401	44.21
1998	3,751,869	3,163,860	18.59	271,415	191,052	42.06
1999	3,871,265	3,216,758	20.35	267,248	191,443	39.60
2000	4,094,478	3,397,800	20.50	281,971	194,284	45.13
2001	4,188,007	3,505,873	19.46	322,708	194,772	65.69
2002	4,024,472	3,220,634	24.96	322,594	194,633	65.74
2003	3,861,473	3,154,390	22.42	323,309	193,018	67.50
2004	3,779,603	3,155,632	19.77	329,207	193,833	69.84
2005	3,834,399	3,172,618	20.86	344,146	196,456	75.18
2006	3,857,543	3,239,813	19.07	354,680	198,627	78.57
2007	3,845,320	3,260,908	17.92	374,536	201,009	86.33
2008	3,853,054	3,301,536	16.70	413,288	198,532	108.17
2009	3815196	3218734	18.5	396930	205068	93.6

Table A-2- 2 Percentage Employment / Establishment Difference for Sacramental CSA

Year	NETS Total emp	CBP Total emp	CSA level Diff.(%)	NETS Total est	CBP Total est	CSA level Diff.(%)
1993	691,029	512,463	34.84	61,838	41,471	49.11
1994	720,698	517,723	39.21	62,269	41,325	50.68
1995	737,368	554,929	32.88	63,092	41,323	52.68
1996	745,750	560,702	33.00	65,726	42,232	55.63
1997	770,038	581,926	32.33	65,442	42,973	52.29
1998	777,082	606,753	28.07	64,468	43,365	48.66
1999	790,940	639,033	23.77	63,326	43,848	44.42
2000	830,460	664,480	24.98	67,011	44,945	49.10
2001	880,449	698,696	26.01	78,821	45,843	71.94
2002	868,588	707,509	22.77	79,964	48,205	65.88
2003	868,948	729,293	19.15	80,549	48,689	65.44
2004	866,149	741,532	16.81	83,072	49,593	67.51
2005	884,269	758,201	16.63	88,417	51,418	71.96
2006	901,369	783,351	15.07	91,814	52,679	74.29
2007	898,014	777,215	15.54	97,873	53,644	82.45
2008	905,098	756,354	19.67	109,217	52,734	107.11
2009	872,358	700,234	24.58	108,464	50,883	113.16

Table A-2- 3 Percentage Employment / Establishment Difference for San Diego CSA

Year	NETS Total emp	CBP Total emp	CSA level Diff.(%)	NETS Total est	CBP Total est	CSA level Diff.(%)
1993	1,102,628	817,063	34.95	90,025	59,375	51.62
1994	1,139,044	815,301	39.71	93,403	59,569	56.80
1995	1,138,355	843,338	34.98	93,210	59,620	56.34
1996	1,173,156	873,631	34.29	96,118	61,192	57.08
1997	1,205,891	906,473	33.03	95,779	62,592	53.02
1998	1,227,715	959,264	27.99	95,621	63,720	50.06
1999	1,281,614	1,014,486	26.33	94,370	64,809	45.61
2000	1,345,946	1,050,493	28.13	100,495	66,673	50.73
2001	1,425,974	1,080,592	31.96	116,348	68,243	70.49
2002	1,404,013	1,082,765	29.67	124,410	71,000	75.23
2003	1,387,900	1,121,725	23.73	125,492	72,206	73.80
2004	1,389,136	1,150,390	20.75	130,048	74,081	75.55
2005	1,416,183	1,163,781	21.69	135,918	76,315	78.10
2006	1,448,257	1,205,487	20.14	140,863	78,056	80.46
2007	1,453,694	1,197,231	21.42	154,940	79,080	95.93
2008	1,474,649	1,188,332	24.09	168,300	77,980	115.82
2009	1,399,736	1,119,468	25.04	158,111	76,172	107.57

2.2 Comparison of sectoral employment

Table A-2- 4 Descriptive statistics of industrial differences index (1993-1997)

Division	SIC	Year	N	Employment				Establishment			
				mean	sd	min	max	mean	sd	min	max
Construction	3	1993	45	1.05	0.10	0.68	2.37	0.82	0.25	0.61	0.98
		1994	45	1.01	0.08	0.65	1.33	0.83	0.15	0.63	0.96
		1995	45	1.01	0.08	0.72	1.56	0.82	0.16	0.64	0.93
		1996	45	0.94	0.07	0.69	1.22	0.78	0.12	0.62	0.92
		1997	45	0.91	0.07	0.71	1.39	0.78	0.14	0.63	0.95
Manufacturing	4	1993	54	0.87	0.15	0.52	1.17	1.07	0.15	0.69	1.34
		1994	54	0.89	0.16	0.49	1.31	1.07	0.17	0.47	1.35
		1995	54	0.88	0.16	0.56	1.39	1.06	0.18	0.59	1.43
		1996	54	0.89	0.15	0.47	1.89	1.04	0.21	0.52	1.40
		1997	54	0.89	0.14	0.55	1.22	1.08	0.16	0.74	1.42
Transportation	5	1993	51	1.04	0.13	0.43	1.99	0.92	0.30	0.59	1.38
		1994	51	1.10	0.13	0.52	2.16	0.96	0.35	0.70	1.45
		1995	51	1.06	0.11	0.59	1.97	0.95	0.29	0.72	1.58
		1996	51	1.06	0.12	0.48	2.45	0.95	0.32	0.69	1.54
		1997	51	1.05	0.11	0.55	2.09	0.97	0.31	0.65	1.38
Wholesale Trade	6	1993	50	0.85	0.13	0.45	1.35	0.98	0.18	0.56	1.34
		1994	50	0.86	0.14	0.40	1.56	1.02	0.21	0.60	1.49
		1995	50	0.88	0.14	0.45	1.57	1.03	0.22	0.72	1.37
		1996	50	0.88	0.12	0.43	1.54	0.98	0.20	0.81	1.31
		1997	50	0.86	0.12	0.51	1.51	0.97	0.18	0.70	1.30
Retail Trade	7	1993	56	0.79	0.07	0.57	0.94	1.00	0.09	0.81	1.18
		1994	56	0.77	0.04	0.55	0.98	0.90	0.09	0.81	0.99
		1995	56	0.77	0.05	0.48	1.02	0.89	0.10	0.79	0.99
		1996	56	0.81	0.05	0.51	1.95	0.93	0.19	0.78	1.03
		1997	56	0.84	0.05	0.58	1.89	0.94	0.17	0.81	1.06
FIRE	8	1993	48	0.97	0.09	0.65	1.34	0.90	0.17	0.70	1.15
		1994	48	1.01	0.09	0.44	1.54	1.00	0.21	0.72	1.17
		1995	48	1.08	0.10	0.62	1.57	1.03	0.20	0.68	1.27
		1996	48	1.08	0.11	0.63	1.63	1.00	0.21	0.67	1.31
		1997	48	1.13	0.10	0.67	1.64	0.99	0.22	0.72	1.18
Service	9	1993	57	1.27	0.08	0.82	1.83	1.12	0.19	1.03	1.49
		1994	57	1.26	0.09	0.87	2.08	1.14	0.21	1.04	1.55
		1995	57	1.24	0.08	0.90	1.82	1.15	0.18	1.04	1.52
		1996	57	1.23	0.12	0.72	1.89	1.16	0.19	1.03	1.74
		1997	57	1.22	0.10	0.78	1.93	1.14	0.19	1.04	1.70

Table A-2- 5 Descriptive statistics of industrial differences index (1998-2009)

naics2	Sector	year	N	Employment				Establishment			
				mean	sd	min	max	mean	sd	min	max
23	Construction	1998	50	0.95	0.19	1.63	0.72	0.78	0.07	0.60	0.97
		1999	50	0.86	0.15	1.37	0.66	0.76	0.08	0.61	0.95
		2000	50	0.83	0.11	1.13	0.59	0.73	0.09	0.55	0.96
		2001	50	0.80	0.13	1.10	0.34	0.72	0.11	0.46	0.96
		2002	50	0.84	0.14	1.30	0.42	0.69	0.10	0.45	0.91
		2003	50	0.84	0.12	1.28	0.51	0.70	0.10	0.45	0.94
		2004	50	0.83	0.13	1.39	0.59	0.70	0.09	0.47	0.92
		2005	50	0.83	0.16	1.70	0.63	0.68	0.09	0.45	0.89
		2006	50	0.81	0.19	1.81	0.52	0.66	0.09	0.43	0.85
		2007	50	0.87	0.15	1.48	0.60	0.65	0.09	0.42	0.82
		2008	50	0.99	0.17	1.55	0.66	0.68	0.08	0.48	0.84
		2009	50	1.21	0.20	1.72	0.88	0.77	0.07	0.65	0.94
31	Manufacturing	1998	43	0.97	0.21	1.68	0.65	1.20	0.11	0.99	1.48
		1999	43	1.00	0.21	1.46	0.53	1.23	0.12	0.98	1.52
		2000	43	1.02	0.24	1.83	0.56	1.21	0.11	1.00	1.49
		2001	43	0.97	0.22	1.69	0.58	1.11	0.11	0.85	1.36
		2002	43	1.02	0.25	1.92	0.57	1.16	0.13	0.95	1.49
		2003	43	1.06	0.27	1.89	0.62	1.18	0.15	0.89	1.54
		2004	43	1.11	0.27	1.91	0.68	1.20	0.15	0.91	1.52
		2005	43	1.09	0.28	1.87	0.60	1.21	0.17	0.82	1.62
		2006	43	1.08	0.26	1.88	0.70	1.21	0.17	0.81	1.61
		2007	43	1.07	0.26	1.83	0.60	1.20	0.17	0.88	1.60
		2008	43	1.02	0.27	1.79	0.45	1.13	0.17	0.78	1.54
		2009	43	1.02	0.31	2.04	0.50	1.09	0.15	0.76	1.42
42	Wholesale Trade	1998	47	1.06	0.32	2.49	0.63	1.16	0.20	0.89	1.69
		1999	47	1.03	0.27	1.73	0.21	1.15	0.21	0.87	2.04
		2000	47	1.03	0.29	1.88	0.24	1.17	0.22	0.75	2.15
		2001	47	0.99	0.31	2.32	0.21	1.11	0.22	0.66	1.90
		2002	47	1.01	0.27	1.67	0.20	1.13	0.19	0.77	1.55
		2003	47	1.13	0.47	2.64	0.19	1.20	0.26	0.86	2.05
		2004	47	1.19	0.50	2.83	0.30	1.24	0.28	0.87	2.17
		2005	47	1.13	0.40	2.38	0.33	1.17	0.25	0.84	1.92
		2006	47	1.16	0.46	2.62	0.30	1.16	0.25	0.83	1.95
		2007	47	1.20	0.46	2.74	0.36	1.20	0.28	0.83	2.19
		2008	47	1.14	0.47	2.40	0.33	1.21	0.33	0.78	2.39
		2009	47	1.11	0.42	2.37	0.27	1.15	0.27	0.75	1.86

Table A-2-7 Descriptive statistics of industrial differences index (1998-2009) (continued)

naics2	Sector	year	N	Employment				Establishment			
				mean	sd	min	max	mean	sd	min	max
44	Retail Trade	1998	54	0.90	0.27	2.63	0.54	1.06	0.08	0.86	1.27
		1999	54	0.90	0.31	2.96	0.54	1.05	0.07	0.88	1.28
		2000	54	0.90	0.35	3.32	0.57	1.05	0.08	0.83	1.32
		2001	54	0.85	0.14	1.31	0.55	1.01	0.08	0.79	1.21
		2002	54	0.84	0.11	1.10	0.56	1.02	0.08	0.82	1.21
		2003	54	0.85	0.11	1.20	0.52	1.02	0.07	0.86	1.23
		2004	54	0.84	0.10	1.12	0.62	1.02	0.08	0.85	1.24
		2005	54	0.84	0.11	1.10	0.52	0.98	0.08	0.79	1.19
		2006	54	0.83	0.11	1.11	0.58	0.97	0.07	0.80	1.18
		2007	54	0.84	0.12	1.21	0.54	0.96	0.08	0.78	1.17
		2008	54	0.85	0.11	1.08	0.54	0.92	0.08	0.68	1.14
		2009	54	0.86	0.14	1.29	0.53	0.85	0.08	0.61	1.05
48	Transportation and Warehousing	1998	41	1.11	0.42	2.63	0.61	0.95	0.15	0.53	1.39
		1999	41	1.10	0.40	2.32	0.61	0.96	0.16	0.53	1.37
		2000	41	1.11	0.37	2.22	0.63	0.95	0.18	0.61	1.67
		2001	41	1.16	0.42	2.45	0.62	0.95	0.18	0.60	1.73
		2002	41	1.09	0.25	1.87	0.66	0.93	0.16	0.51	1.33
		2003	41	0.92	0.19	1.49	0.55	0.91	0.16	0.60	1.34
		2004	41	0.92	0.24	1.78	0.52	0.91	0.15	0.55	1.21
		2005	41	0.93	0.26	1.75	0.47	0.93	0.14	0.59	1.28
		2006	41	0.94	0.25	1.72	0.46	0.94	0.15	0.65	1.24
		2007	41	0.93	0.26	1.77	0.47	0.97	0.15	0.69	1.33
		2008	41	0.91	0.28	2.03	0.48	0.97	0.14	0.69	1.31
		2009	41	0.88	0.26	1.76	0.44	0.95	0.15	0.65	1.37
51	Information	1998	41	0.94	0.30	2.06	0.34	1.22	0.24	0.74	1.97
		1999	41	1.06	0.34	2.03	0.38	1.25	0.23	0.70	1.79
		2000	41	1.03	0.39	2.34	0.34	1.25	0.24	0.77	1.77
		2001	41	0.98	0.36	2.58	0.34	1.19	0.27	0.64	2.08
		2002	41	0.99	0.31	2.23	0.30	1.27	0.24	0.67	1.99
		2003	41	1.02	0.34	2.12	0.40	1.29	0.20	0.88	1.73
		2004	41	1.08	0.47	3.16	0.29	1.32	0.23	0.89	1.98
		2005	41	1.10	0.49	3.08	0.34	1.27	0.20	0.93	1.96
		2006	41	1.16	0.51	3.08	0.34	1.31	0.22	1.01	2.03
		2007	41	1.10	0.47	2.71	0.49	1.29	0.23	0.99	1.89
		2008	41	1.05	0.48	3.13	0.55	1.22	0.20	0.80	1.69
		2009	41	0.96	0.39	2.50	0.45	1.15	0.23	0.71	1.76

Table A-2-7 Descriptive statistics of industrial differences index (1998-2009) (continued)

naics2	Sector	year	N	Employment				Establishment			
				mean	sd	min	max	mean	sd	min	max
52	Finance and Insurance	1998	49	0.93	0.23	1.51	0.46	0.83	0.10	0.56	0.99
		1999	49	0.94	0.24	1.44	0.40	0.84	0.11	0.48	1.01
		2000	49	0.99	0.27	1.58	0.42	0.87	0.10	0.51	1.03
		2001	49	0.98	0.27	1.53	0.20	0.87	0.11	0.50	1.03
		2002	49	0.96	0.23	1.56	0.31	0.95	0.14	0.60	1.50
		2003	49	0.96	0.25	1.48	0.28	0.96	0.14	0.58	1.49
		2004	49	0.98	0.23	1.62	0.35	1.00	0.14	0.70	1.49
		2005	49	1.01	0.24	1.70	0.39	1.01	0.15	0.64	1.45
		2006	49	1.04	0.26	1.72	0.42	1.01	0.14	0.67	1.39
		2007	49	1.06	0.26	1.67	0.32	1.02	0.16	0.58	1.37
		2008	49	1.09	0.27	1.84	0.45	1.02	0.18	0.58	1.42
		2009	49	0.97	0.22	1.63	0.44	0.90	0.16	0.43	1.25
53	Real Estate	1998	38	1.56	0.35	2.39	0.91	1.17	0.16	0.91	1.65
		1999	38	1.55	0.45	3.30	0.79	1.16	0.15	0.92	1.55
		2000	38	1.50	0.29	2.30	0.90	1.17	0.16	0.90	1.56
		2001	38	1.48	0.25	2.16	0.88	1.17	0.13	0.94	1.45
		2002	38	1.55	0.30	2.35	0.92	1.19	0.14	0.97	1.57
		2003	38	1.54	0.34	2.40	1.03	1.17	0.14	0.94	1.62
		2004	38	1.53	0.35	2.31	0.86	1.15	0.14	0.94	1.61
		2005	38	1.50	0.37	2.57	0.75	1.12	0.12	0.89	1.46
		2006	38	1.51	0.33	2.36	0.81	1.12	0.11	0.85	1.41
		2007	38	1.66	0.39	2.75	1.19	1.18	0.13	0.95	1.54
		2008	38	1.71	0.29	2.54	1.29	1.17	0.11	0.95	1.36
		2009	38	1.62	0.24	2.21	1.25	1.09	0.09	0.92	1.34
54	Professional Service	1998	49	1.43	0.27	2.17	0.85	1.11	0.12	0.75	1.36
		1999	49	1.38	0.37	3.09	0.70	1.08	0.12	0.72	1.42
		2000	49	1.47	0.64	4.44	0.85	1.07	0.11	0.80	1.54
		2001	49	1.35	0.28	2.19	0.81	1.06	0.10	0.83	1.41
		2002	49	1.26	0.29	2.23	0.75	0.98	0.08	0.84	1.19
		2003	49	1.27	0.29	2.15	0.75	0.96	0.09	0.82	1.30
		2004	49	1.23	0.30	2.51	0.73	0.91	0.08	0.72	1.20
		2005	49	1.21	0.28	2.29	0.42	0.90	0.09	0.67	1.32
		2006	49	1.17	0.26	2.09	0.48	0.88	0.10	0.60	1.27
		2007	49	1.22	0.28	2.32	0.57	0.90	0.10	0.67	1.13
		2008	49	1.32	0.37	2.91	0.73	0.90	0.12	0.61	1.39
		2009	49	1.29	0.38	2.70	0.72	0.90	0.12	0.66	1.41

Table A-2-7 Descriptive statistics of industrial differences index (1998-2009) (continued)

naics2	Sector	year	N	Employment				Establishment			
				mean	sd	min	max	mean	sd	min	max
55	Management of Companies	1998	28	0.03	0.02	0.08	0.00	0.11	0.05	0.03	0.26
		1999	28	0.03	0.03	0.13	0.00	0.11	0.06	0.02	0.29
		2000	28	0.03	0.03	0.13	0.00	0.12	0.07	0.02	0.41
		2001	28	0.06	0.13	0.69	0.00	0.14	0.07	0.04	0.35
		2002	28	0.07	0.14	0.75	0.00	0.30	0.11	0.13	0.49
		2003	28	0.08	0.12	0.64	0.01	0.32	0.13	0.13	0.59
		2004	28	0.09	0.12	0.64	0.01	0.38	0.16	0.14	0.71
		2005	28	0.11	0.14	0.66	0.01	0.45	0.20	0.16	0.84
		2006	28	0.15	0.21	0.89	0.01	0.48	0.22	0.17	0.99
		2007	28	0.18	0.27	1.19	0.01	0.50	0.21	0.16	0.92
		2008	28	0.19	0.28	1.39	0.02	0.54	0.23	0.15	0.93
		2009	28	0.17	0.27	1.33	0.01	0.46	0.20	0.17	0.83
56	Administrative Management	1998	41	0.79	0.29	1.69	0.32	1.02	0.10	0.82	1.40
		1999	41	0.79	0.33	2.15	0.26	0.99	0.10	0.66	1.34
		2000	41	0.78	0.40	2.45	0.29	1.02	0.10	0.77	1.20
		2001	41	0.83	0.29	1.66	0.35	1.22	0.16	0.80	1.65
		2002	41	0.93	0.30	1.95	0.63	1.25	0.22	0.89	2.11
		2003	41	0.90	0.26	1.62	0.55	1.25	0.18	0.85	1.69
		2004	41	0.96	0.29	1.78	0.61	1.41	0.22	0.90	1.94
		2005	41	1.05	0.33	2.08	0.65	1.97	0.34	0.77	2.76
		2006	41	1.10	0.38	2.39	0.62	2.12	0.33	1.08	2.85
		2007	41	1.19	0.39	2.34	0.74	2.35	0.37	1.05	3.12
		2008	41	1.31	0.41	2.46	0.82	2.82	0.41	1.24	3.59
		2009	41	1.46	0.38	2.65	0.92	3.29	0.51	1.54	4.57
61	Educational Services	1998	27	7.29	5.87	29.13	1.54	2.16	0.80	1.11	4.22
		1999	27	6.92	5.31	25.49	1.48	2.13	0.71	1.15	4.35
		2000	27	6.61	5.05	23.45	1.35	2.07	0.70	1.09	4.02
		2001	27	6.99	6.49	32.57	1.28	1.90	0.70	0.99	3.84
		2002	27	6.98	7.13	35.19	1.14	1.91	0.83	0.97	4.29
		2003	27	6.90	7.51	38.81	1.16	1.77	0.73	0.92	4.17
		2004	27	6.74	8.03	42.39	1.16	1.72	0.74	0.93	4.07
		2005	27	6.87	9.09	49.1	1.11	1.74	0.98	0.98	5.59
		2006	27	6.41	9.25	50.31	1.16	1.66	0.87	1.02	5.06
		2007	27	6.12	8.48	46.51	1.23	1.58	1.01	0.88	5.83
		2008	27	5.57	7.08	38.63	1.04	1.44	0.81	0.78	4.59
		2009	27	5.69	8.50	45.96	1.00	1.49	0.84	0.75	4.73

Table A-2-7 Descriptive statistics of industrial differences index (1998-2009) (continued)

naics2	Sector	year	N	Employment				Establishment			
				mean	sd	min	max	mean	sd	min	max
62	Health Care and Social Assistance	1998	54	0.83	0.17	1.29	0.38	0.87	0.11	0.58	1.23
		1999	54	0.84	0.16	1.34	0.54	0.87	0.08	0.70	1.08
		2000	54	0.84	0.16	1.23	0.50	0.88	0.08	0.70	1.06
		2001	54	0.88	0.18	1.47	0.54	0.93	0.10	0.69	1.18
		2002	54	0.83	0.14	1.08	0.49	0.94	0.10	0.60	1.26
		2003	54	0.83	0.14	1.25	0.52	0.93	0.10	0.64	1.23
		2004	54	0.83	0.13	1.17	0.54	0.91	0.10	0.63	1.13
		2005	54	0.83	0.13	1.12	0.35	0.89	0.09	0.69	1.13
		2006	54	0.83	0.14	1.18	0.39	0.87	0.10	0.69	1.23
		2007	54	0.85	0.15	1.32	0.57	0.88	0.11	0.70	1.28
		2008	54	0.82	0.14	1.27	0.55	0.83	0.10	0.67	1.20
		2009	54	0.74	0.11	1.07	0.51	0.78	0.10	0.59	1.11
71	Arts and Entertainment	1998	33	1.10	0.38	2.45	0.41	1.39	0.27	0.65	2.02
		1999	33	1.14	0.39	2.28	0.33	1.41	0.29	0.63	1.99
		2000	33	1.15	0.41	2.17	0.33	1.45	0.29	0.63	1.90
		2001	33	1.13	0.34	1.86	0.42	1.42	0.30	0.60	1.99
		2002	33	1.08	0.36	2.15	0.36	1.37	0.27	0.55	1.83
		2003	33	1.03	0.30	1.72	0.48	1.36	0.27	0.52	1.85
		2004	33	1.00	0.32	1.60	0.35	1.30	0.26	0.50	1.78
		2005	33	0.98	0.32	1.73	0.35	1.27	0.24	0.48	1.71
		2006	33	0.97	0.30	1.78	0.42	1.29	0.25	0.48	1.74
		2007	33	0.92	0.31	1.75	0.38	1.21	0.24	0.44	1.70
		2008	33	0.87	0.27	1.74	0.32	1.11	0.22	0.40	1.52
		2009	33	0.84	0.29	1.54	0.27	1.07	0.22	0.38	1.53
72	Accommodation	1998	49	0.79	0.13	1.09	0.48	0.71	0.08	0.56	0.90
		1999	49	0.80	0.14	1.17	0.52	0.75	0.08	0.65	0.94
		2000	49	0.81	0.15	1.20	0.52	0.74	0.08	0.64	0.97
		2001	49	0.77	0.13	1.10	0.54	0.66	0.08	0.55	0.85
		2002	49	0.78	0.13	1.05	0.46	0.68	0.07	0.56	0.84
		2003	49	0.77	0.12	1.01	0.47	0.67	0.07	0.56	0.89
		2004	49	0.77	0.13	1.18	0.46	0.65	0.08	0.53	0.91
		2005	49	0.78	0.15	1.24	0.44	0.65	0.08	0.52	0.87
		2006	49	0.79	0.14	1.17	0.46	0.69	0.08	0.54	0.88
		2007	49	0.73	0.10	0.97	0.43	0.62	0.10	0.48	0.86
		2008	49	0.68	0.12	1.06	0.41	0.53	0.11	0.38	0.80
		2009	49	0.67	0.11	0.98	0.36	0.51	0.10	0.37	0.74

Table A-2-7 Descriptive statistics of industrial differences index (1998-2009) (continued)

naics2	Sector	year	N	Employment				Establishment			
				mean	sd	min	max	mean	sd	min	max
81	Other Services	1998	52	1.41	0.35	2.51	0.84	1.21	0.11	0.92	1.63
		1999	52	1.40	0.37	2.73	0.86	1.22	0.10	0.90	1.51
		2000	52	1.44	0.35	2.58	0.90	1.22	0.10	0.86	1.45
		2001	52	1.55	0.34	2.69	0.91	1.39	0.13	1.20	1.97
		2002	52	1.49	0.40	2.79	0.92	1.35	0.16	1.12	1.90
		2003	52	1.47	0.34	2.42	0.95	1.36	0.18	1.12	2.13
		2004	52	1.43	0.33	2.44	0.81	1.31	0.15	1.06	1.78
		2005	52	1.36	0.28	2.30	0.90	1.23	0.11	1.04	1.65
		2006	52	1.40	0.29	2.28	0.91	1.22	0.13	1.02	1.75
		2007	52	1.40	0.31	2.33	0.79	1.19	0.13	1.00	1.67
		2008	52	1.43	0.27	2.18	1.02	1.19	0.14	1.01	1.61
		2009	52	1.38	0.29	2.41	1.09	1.16	0.12	0.92	1.59

APPENDIX FOR CHAPTER 3

1. Los Angeles[24]

Table A-3- 1 Correlation matrix for distance to CBD and inverse of distance to subcenters (1995), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7	INCN8	INCN9	INCN10	INCN11	INCN12	INCN13	INCN14	INCN15
DCN1	1														
INCN2	-0.42	1													
INCN3	-0.46	0.37	1												
INCN4	-0.33	0.34	0.15	1											
INCN5	-0.35	0.16	0.23	0.05	1										
INCN6	-0.44	0.50	0.45	0.29	0.13	1									
INCN7	-0.22	0.00	0.05	-0.05	0.19	-0.01	1								
INCN8	-0.42	0.25	0.52	0.12	0.15	0.39	0.04	1							
INCN9	-0.44	0.42	0.57	0.21	0.15	0.72	0.01	0.52	1						
INCN10	-0.45	0.29	0.67	0.10	0.29	0.32	0.10	0.44	0.39	1					
INCN11	-0.42	0.39	0.40	0.15	0.35	0.29	0.07	0.24	0.30	0.41	1				
INCN12	-0.33	0.30	0.14	0.72	0.03	0.29	-0.06	0.12	0.22	0.09	0.13	1			
INCN13	-0.36	0.26	0.26	0.11	0.42	0.19	0.08	0.16	0.19	0.28	0.55	0.09	1		
INCN14	-0.41	0.89	0.33	0.38	0.15	0.46	-0.01	0.22	0.37	0.25	0.36	0.33	0.25	1	
INCN15	-0.40	0.42	0.32	0.19	0.29	0.28	0.05	0.20	0.27	0.31	0.67	0.16	0.54	0.41	1

[24] Cells are highlighted in Tables A-3-1 to A-3-3 if the correlation coefficients they contain are larger than 0.5.

Table A-3- 2 Correlation matrix for distance to CBD and inverse of distance to subcenters (2000), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7	INCN8	INCN9	INCN10	INCN11	INCN12	INCN13	INCN14	INCN15	INCN16	INCN17	INCN18
DCN1	1																	
INCN2	-0.42	1																
INCN3	-0.34	0.34	1															
INCN4	-0.23	0.00	-0.05	1														
INCN5	-0.46	0.37	0.15	0.06	1													
INCN6	-0.44	0.40	0.22	0.01	0.52	1												
INCN7	-0.40	0.42	0.19	0.05	0.32	0.25	1											
INCN8	-0.44	0.50	0.29	-0.01	0.45	0.74	0.28	1										
INCN9	-0.35	0.16	0.05	0.19	0.23	0.14	0.29	0.13	1									
INCN10	-0.42	0.27	0.13	0.04	0.53	0.56	0.20	0.42	0.15	1								
INCN11	-0.36	0.27	0.12	0.09	0.26	0.18	0.55	0.19	0.42	0.16	1							
INCN12	-0.41	0.89	0.38	-0.01	0.33	0.36	0.41	0.46	0.15	0.24	0.25	1						
INCN13	-0.45	0.29	0.11	0.10	0.67	0.37	0.31	0.32	0.29	0.42	0.29	0.26	1					
INCN14	-0.42	0.64	0.37	0.00	0.30	0.30	0.48	0.37	0.18	0.21	0.30	0.71	0.25	1				
INCN15	-0.46	0.52	0.23	0.02	0.63	0.67	0.34	0.67	0.18	0.46	0.24	0.46	0.44	0.39	1			
INCN16	-0.29	0.02	-0.04	0.49	0.11	0.05	0.08	0.03	0.23	0.09	0.11	0.02	0.17	0.02	0.06	1		
INCN17	-0.26	0.02	-0.04	0.72	0.09	0.03	0.08	0.02	0.25	0.06	0.12	0.01	0.13	0.03	0.05	0.52	1	
INCN18	-0.27	0.02	-0.05	0.65	0.09	0.03	0.07	0.01	0.23	0.07	0.11	0.01	0.14	0.02	0.04	0.77	0.67	1

Table A-3- 3 Correlation matrix for distance to CBD and inverse of distance to subcenters (2005), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7	INCN8	INCN9	INCN10	INCN11	INCN12	INCN13	INCN14	INCN15	INCN16	INCN17	INCN18
DCN1	1																	
INCN2	-0.43	1																
INCN3	-0.23	0.00	1															
INCN4	-0.44	0.51	0.00	1														
INCN5	-0.36	0.17	0.20	0.14	1													
INCN6	-0.34	0.35	-0.05	0.30	0.05	1												
INCN7	-0.42	0.26	0.05	0.39	0.16	0.12	1											
INCN8	-0.41	0.89	0.00	0.46	0.15	0.38	0.23	1										
INCN9	-0.46	0.53	0.03	0.67	0.19	0.24	0.44	0.46	1									
INCN10	-0.44	0.42	0.02	0.73	0.15	0.22	0.52	0.38	0.74	1								
INCN11	-0.39	0.53	-0.03	0.48	0.09	0.56	0.20	0.56	0.38	0.35	1							
INCN12	-0.46	0.38	0.06	0.46	0.24	0.16	0.53	0.33	0.64	0.57	0.26	1						
INCN13	-0.37	0.18	0.18	0.15	0.87	0.06	0.17	0.17	0.21	0.17	0.10	0.26	1					
INCN14	-0.40	0.43	0.05	0.28	0.30	0.20	0.20	0.41	0.35	0.27	0.26	0.33	0.32	1				
INCN15	-0.37	0.27	0.09	0.19	0.43	0.12	0.16	0.26	0.25	0.20	0.17	0.27	0.45	0.55	1			
INCN16	-0.42	0.64	0.01	0.38	0.18	0.38	0.20	0.71	0.39	0.32	0.47	0.30	0.20	0.48	0.31	1		
INCN17	-0.29	0.03	0.50	0.03	0.24	-0.04	0.10	0.02	0.07	0.06	-0.01	0.12	0.23	0.08	0.12	0.03	1	
INCN18	-0.45	0.61	0.02	0.70	0.18	0.27	0.38	0.53	0.89	0.65	0.43	0.55	0.20	0.36	0.25	0.43	0.05	1

2. San Francisco[25]

Table A-3- 4 Correlation matrix for distance to CBD and inverse of distance to subcenters (1995), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6
DCN1	1					
INCN2	-0.09	1				
INCN3	-0.50	-0.02	1.00			
INCN4	0.00	0.54	-0.05	1.00		
INCN5	-0.14	0.79	0.01	0.43	1.00	
INCN6	-0.22	0.34	0.04	0.20	0.43	1.00

Table A-3- 5 Correlation matrix for distance to CBD and inverse of distance to subcenters (2000), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7	INCN8	INCN9
DCN1	1								
INCN2	-0.10	1							
INCN3	-0.50	-0.01	1						
INCN4	-0.43	0.04	0.26	1					
INCN5	-0.01	0.55	-0.04	0.00	1				
INCN6	-0.14	0.79	0.01	0.07	0.43	1			
INCN7	-0.14	0.69	0.01	0.07	0.39	0.85	1		
INCN8	-0.22	0.35	0.05	0.15	0.21	0.43	0.49	1	
INCN9	-0.39	-0.04	0.29	0.08	-0.06	-0.02	-0.03	-0.01	1

Table A-3- 6 Correlation matrix for distance to CBD and inverse of distance to subcenters (2005), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7	INCN8	INCN9	INCN10
DCN1	1									
INCN2	-0.04	1								
INCN3	-0.50	-0.03	1							
INCN4	-0.10	0.68	-0.01	1						
INCN5	-0.15	0.44	0.01	0.64	1					
INCN6	-0.01	0.88	-0.04	0.57	0.39	1				
INCN7	-0.35	-0.07	0.21	-0.06	-0.06	-0.08	1			
INCN8	-0.43	0.01	0.26	0.04	0.08	0.00	0.05	1		
INCN9	-0.33	0.12	0.14	0.17	0.25	0.10	0.01	0.30	1	
INCN10	-0.33	0.04	0.20	0.06	0.06	0.02	0.23	0.09	0.10	1

3. Sacramento[26]

[25] Cells are highlighted in Tables A-3-4 to A-3-6 if the correlation coefficients they contain are larger than 0.5.

Table A-3- 7 Correlation matrix for distance to CBD and inverse of distance to subcenters (1995), centers ranked by top employment densit

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7
DCN1	1.00						
INCN2	-0.50	1.00					
INCN3	-0.34	0.07	1.00				
INCN4	-0.41	0.16	0.58	1.00			
INCN5	-0.31	0.09	-0.01	0.01	1.00		
INCN6	-0.51	0.33	0.19	0.28	0.15	1.00	
INCN7	0.04	-0.11	-0.07	-0.09	-0.02	-0.07	1.00

Table A-3- 8 Correlation matrix for distance to CBD and inverse of distance to subcenters (2000), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7
DCN1	1.00						
INCN2	-0.35	1.00					
INCN3	-0.41	0.58	1.00				
INCN4	-0.51	0.19	0.28	1.00			
INCN5	-0.37	0.00	0.03	0.20	1.00		
INCN6	0.03	-0.07	-0.09	-0.07	-0.06	1.00	
INCN7	-0.32	0.00	0.02	0.17	0.34	-0.01	1.00

Table A-3- 9 Correlation matrix for distance to CBD and inverse of distance to subcenters (2005), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7	INCN8
DCN1	1.00							
INCN2	-0.35	1.00						
INCN3	-0.44	0.32	1.00					
INCN4	-0.33	0.41	0.43	1.00				
INCN5	-0.36	0.00	0.06	-0.02	1.00			
INCN6	-0.51	0.20	0.29	0.14	0.19	1.00		
INCN7	0.03	-0.07	-0.11	-0.11	-0.06	-0.06	1.00	
INCN8	-0.31	-0.02	0.01	-0.05	0.36	0.14	-0.02	1.00

[26] Cells are highlighted in Tables A-3-7 to A-3-9 if the correlation coefficients they contain are larger than 0.5.

4. San Diego[27]

Table A-3- 10 Correlation matrix for distance to CBD and inverse of distance to subcenters (1995), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7	INCN8
DCN1	1.00							
INCN2	-0.42	1.00						
INCN3	-0.55	0.46	1.00					
INCN4	0.04	0.01	-0.08	1.00				
INCN5	-0.58	0.09	0.27	-0.16	1.00			
INCN6	-0.44	0.08	0.22	-0.12	0.19	1.00		
INCN7	-0.19	0.21	0.09	0.33	-0.07	0.01	1.00	
INCN8	-0.52	0.12	0.32	-0.13	0.31	0.69	0.00	1.00

Table A-3- 11 Correlation matrix for distance to CBD and inverse of distance to subcenters (2000), centers ranked by top employment density

	DCN 1	INCN 2	INCN 3	INCN 4	INCN 5	INCN 6	INCN 7	INCN 8	INCN 9	INCN1 0	INCN1 1	INCN1 2	INCN1 3
DCN1	1.00												
INCN2	-0.41	1.00											
INCN3	-0.54	0.47	1.00										
INCN4	-0.59	0.19	0.48	1.00									
INCN5	-0.43	0.48	0.41	0.21	1.00								
INCN6	0.05	0.02	-0.07	-0.13	-0.04	1.00							
INCN7	0.05	0.04	-0.06	-0.13	0.01	0.21	1.00						
INCN8	-0.21	0.26	0.12	0.01	0.11	0.31	0.13	1.00					
INCN9	-0.28	0.55	0.23	0.06	0.34	0.08	0.17	0.28	1.00				
INCN1 0	-0.45	0.08	0.22	0.39	0.06	-0.12	-0.15	0.02	0.00	1.00			
INCN1 1	0.09	0.01	-0.09	-0.16	-0.03	0.41	0.55	0.18	0.11	-0.16	1.00		
INCN1 2	-0.60	0.12	0.30	0.47	0.20	-0.15	-0.13	-0.05	0.02	0.18	-0.16	1.00	
INCN1 3	-0.53	0.12	0.33	0.63	0.12	-0.13	-0.15	0.01	0.02	0.69	-0.16	0.30	1.00

[27] Cells are highlighted in Tables A-3-10 to A-3-12 if the correlation coefficients they contain are larger than 0.5.

Table A-3- 12 Correlation matrix for distance to CBD and inverse of distance to subcenters (2005), centers ranked by top employment density

	DCN1	INCN2	INCN3	INCN4	INCN5	INCN6	INCN7	INCN8	INCN9	INCN10	INCN11	INCN12
DCN1	1.00											
INCN2	-0.43	1.00										
INCN3	-0.55	0.47	1.00									
INCN4	0.02	0.03	-0.06	1.00								
INCN5	-0.45	0.49	0.42	-0.03	1.00							
INCN6	0.02	0.05	-0.06	0.22	0.02	1.00						
INCN7	-0.23	0.27	0.13	0.31	0.11	0.14	1.00					
INCN8	-0.45	0.09	0.23	-0.12	0.07	-0.14	0.02	1.00				
INCN9	-0.31	0.55	0.23	0.09	0.35	0.17	0.29	0.00	1.00			
INCN10	0.06	0.02	-0.08	0.41	-0.02	0.55	0.18	-0.16	0.12	1.00		
INCN11	-0.59	0.12	0.31	-0.15	0.20	-0.13	-0.05	0.18	0.03	-0.16	1.00	
INCN12	-0.52	0.13	0.33	-0.13	0.12	-0.14	0.01	0.69	0.03	-0.16	0.30	1.00

Appendix for Chapter 5

1. The data problem

If NETS changed its techniques for capturing data over the years and has been more successful in capturing certain types of firms or sizes of firms, are those changes correlated with centers?

To answer this question, we use the same division of employment size class as in chapter 2 and calculated the ratio of establishment counts of each size class between centers and non-center locations (Table 1). The trend of the ratio is plotted in the following figure, which indicates that centers (identified in 2005) and non-center locations are different in terms of composition of firms with different size class. Compared to non-center locations, centers capture far less share of smaller-sized firms (Employment <250) but far more share of extra-large firms (Employment >=1000) employment. The distribution of firms with employment between 250 and 1000 are relatively even between centers and non-center locations. However, except for the extra-large firms, the distribution of firms of different size class between centers and non-center locations are relatively stable over time. This implies that if NETS become more successful in capturing small-sized firms, the growth rate of firms in this size group between centers and non-centers would be almost the same, and the changes in data-collection would NOT be biased towards either center or non-center locations.

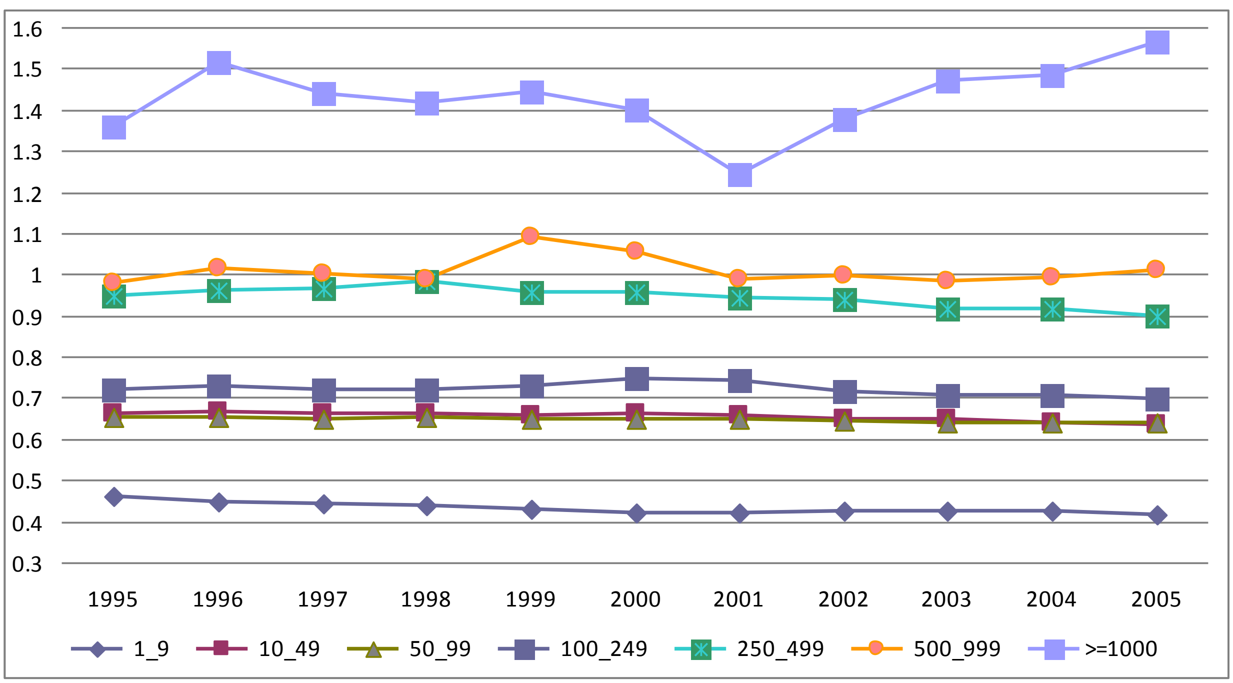


Figure A-5- 1 Ratio of establishment shares between center and non-center locations for each size category

Table A-5- 1 Ratio of establishment counts between centers and non-center locations

	1_9	10_49	50_99	100_249	250_499	500_999	>=1000
1995	0.46	0.66	0.65	0.72	0.95	0.98	1.36
1996	0.45	0.67	0.65	0.73	0.96	1.01	1.51
1997	0.44	0.66	0.65	0.72	0.96	1.00	1.44
1998	0.44	0.66	0.65	0.72	0.99	0.99	1.42
1999	0.43	0.66	0.65	0.73	0.96	1.09	1.44
2000	0.42	0.66	0.65	0.75	0.95	1.06	1.40
2001	0.42	0.66	0.65	0.74	0.94	0.99	1.24
2002	0.42	0.65	0.64	0.72	0.94	0.99	1.38
2003	0.43	0.65	0.64	0.71	0.92	0.98	1.47
2004	0.42	0.64	0.64	0.70	0.92	0.99	1.49
2005	0.42	0.63	0.64	0.70	0.90	1.01	1.57

Table A-5- 2 Selected characteristics of employment centers identified in 2005[28]

ID	Location	Area (acres)	Emp 2005	Emp 2000	Emp 1995	Emp Density 2005	Emp Growth 00-05	Emp Growth 95-05
1	LA Downtown-West LA-Santa Monica	39,680	1,088,700	1,083,836	1,042,136	27.44	4,864	46,086
2	Anaheim-Santa Ana- Costa Mesa-Irvine	24,960	561,975	561,925	493,166	22.52	50	68,833
3	El Segundo	8,320	142,859	166,245	134,181	17.17	-23,386	8,678
4	Torrance	8,960	129,948	145,898	116,682	14.50	-15,950	13,538
5	Long Beach	5,760	110,274	104,890	89,628	19.14	5,384	20,646
6	Fullerton-City of Orange	6,400	99,470	98,146	80,671	15.54	1,324	18,799
7	Irvine Spectrum	6,400	93,187	96,698	65,690	14.56	-3,511	27,497
8	Burbank	3,840	89,789	84,282	83,128	23.38	5,507	6,661
9	Pasadena	2,560	80,924	75,175	72,548	31.61	5,749	8,376
10	Van Nuys Airport-Sherman Oaks	4,480	79,015	67,141	57,629	17.64	11,874	21,386
11	Hidden Hills	3,840	76,594	87,174	74,949	19.95	-10,580	1,645
12	Glendale	2,560	62,952	71,932	64,388	24.59	-8,980	-1,436
13	Commerce	3,840	55,072	65,445	62,584	14.34	-10,373	-7,512
14	Northbridge	3,200	54,835	54,619	60,322	17.14	216	-5,487
15	Santa Fe Springs	3,200	47,476	43,386	40,443	14.84	4,090	7,033
16	La Habra	3,200	43,869	39,282	36,999	13.71	4,587	6,870
17	City of Industry	3,200	41,172	37,996	33,436	12.87	3,176	7,736

[28] Cells are highlighted in Tables A-5-2 if they contain negative values.

Table A-5-2 Selected characteristics of employment centers identified in 2005 (continued)

ID	Location	Area (acres)	Emp 2005	Emp 2000	Emp 1995	Emp Density 2005	Emp Growth 00-05	Emp Growth 95-05
18	Gardena-Campton	3,200	36,779	36,175	36,335	11.49	604	444
19	Whitter	2,560	36,337	33,014	29,116	14.19	3,323	7,221
20	Camino-Ventura	2,560	35,792	31,959	23,962	13.98	3,833	11,830
21	Van Nuys West-Van Nuys City	2,560	33,010	33,662	32,842	12.89	-652	168
22	Alhambra	1,920	29,736	25,995	22,455	15.49	3,741	7,281
23	New Port Center	1,920	28,656	25,593	23,555	14.93	3,063	5,101
24	Glendale	1,280	24,892	24,314	21,557	19.45	578	3,335
25	Thousand Oaks	1,920	24,513	20,870	17,192	12.77	3,643	7,321
26	Huntington Beach, shopping center	1,920	23,496	22,288	19,854	12.24	1,208	3,642
27	Sunset Blvd-Fountain Ave	640	23,391	20,442	19,117	36.55	2,949	4,274
28	Fox Hills Plaza-Los Angeles	1,280	22,416	19,807	18,322	17.51	2,609	4,094
29	Newport Beach (North)	1,280	21,967	19,441	19,833	17.16	2,526	2,134
30	Monterey Park	1,280	21,468	16,422	12,976	16.77	5,046	8,492
31	USC Health Center	640	21,389	32,516	39,604	33.42	-11,127	-18,215
32	El Monte-South El Monte	1,920	20,907	21,265	20,272	10.89	-358	635
33	Huntington Beach	1,280	20,099	20,972	17,622	15.70	-873	2,477

Table A-5-2 Selected characteristics of employment centers identified in 2005 (continued)

ID	Location	Area (acres)	Emp 2005	Emp 2000	Emp 1995	Emp Density 2005	Emp Growth 00-05	Emp Growth 95-05
34	Aliso Viejo	1,280	19,771	15,881	4,059	15.45	3,890	15,712
35	Burbank	1,280	19,720	17,098	14,981	15.41	2,622	4,739
36	South Pointe-San Bernardino	1,280	19,565	18,509	15,332	15.29	1,056	4,233
37	Los Angeles-Culver City	1,280	19,491	18,487	17,093	15.23	1,004	2,398
38	Santa Clarita	1,280	18,748	17,754	17,764	14.65	994	984
39	Monrovia	1,280	18,072	17,274	15,353	14.12	798	2,719
40	Carosel-San Bernardino	1,280	16,940	14,992	17,321	13.23	1,948	-381
41	South Valley	1,280	16,456	17,240	16,900	12.86	-784	-444
42	Lake Forest	1,280	16,399	7,900	2,563	12.81	8,499	13,836
43	Camarillo	1,280	16,089	13,861	9,102	12.57	2,228	6,987
44	San Juan Capistrano	1,280	16,047	13,373	11,458	12.54	2,674	4,589
45	University of California Irvine	640	15,353	15,109	17,144	23.99	244	-1,791
46	Riverside downtown	1,280	15,105	17,280	15,109	11.80	-2,175	-4
47	Cerritos	640	11,419	9,471	9,782	17.84	1,948	1,637
48	Corona	640	11,147	9,258	7,291	17.42	1,889	3,856
49	Camino Capistrano	640	10,837	11,655	10,254	16.93	-818	583

Table A-5-2 Selected characteristics of employment centers identified in 2005 (continued)

50	Rancho Cucamonga	640	10,832	6,103	5,283	16.93	4,729	5,549
51	Orange Coast College, Coasta Mesa	640	10,563	7,700	6,303	16.50	2,863	4,260
52	Chino	640	10,196	5,920	4,471	15.93	4,276	5,725
53	The Commons at Calabasas	640	10,152	9,286	8,550	15.86	866	1,602

Using either 2000 or 2005 employment data, the largest center (the LA Downtown-West LA-Santa Monica corridor) identified also has the largest area, spreading over 36,480 acres. It is also the densest center and has an employment density of approximately 29 jobs per acre in 2000 and 28 jobs per acre in 2005. The smallest centers identified in either 2000 or 2005 are all single-hexagon centers with employment density of approximately 15 to 16 jobs per acre. However, these centers are not the least dense ones.

Table A-5- 3 Pairwise correlations (Employment centers identified in 2005 for Los Angeles region, base year=1995)[29]

	ln (E95_05)	ln(Emp95)	ln(den95)	predgr	DCBD	ln(Dcoast)	ln(DLAX)	DNrstAP	ln(Dhwy)	NWAcc	LFAcc	RLFAcc
ln(E95_05)	1											
ln(Emp95)	-0.54	1										
ln(den95)	-0.86	0.6	1									
predgr	-0.21	0.08	0.3	1								
DCBD	0.37	-0.42	-0.48	0.05	1							
ln(Dcoast)	0	-0.15	-0.02	-0.04	0.06	1						
ln(DLAX)	0.32	-0.4	-0.36	0.03	0.77	0.37	1					
DNrstAP	0.07	-0.16	-0.23	0.04	0.43	0.2	0.4	1				
ln(Dhwy)	-0.04	0.13	0.09	-0.24	-0.06	-0.13	-0.08	-0.2	1			
NWAcc	-0.35	0.36	0.44	-0.02	-0.84	0.04	-0.56	-0.31	0	1		
LFAcc	-0.34	0.34	0.42	-0.03	-0.82	0.07	-0.53	-0.31	0.01	1	1	
RLFAcc	-0.23	0.81	0.32	0.06	-0.23	-0.01	-0.22	-0.05	0.08	0.29	0.28	1

[29] Cells are highlighted in Tables A-5-3 if the correlation coefficients they contain are larger than 0.5.

Table A-5- 4 Selected characteristics of 2005 employment centers[30]

Center ID	Location	Area (acres)	Employment 2005	Employment 1995	Emp. density 2005	Employment growth 1995-2005	Percentage growth 1995-2005
1	San Francisco downtown	12800	603201	605321	47.12	-2120	-0.35
2	San Jose downtown	32000	507386	494461	15.86	12925	2.61
3	West Oakland	10880	207300	197161	19.05	10139	5.14
4	Menlo Park	5120	84088	83528	16.42	560	0.67
5	Newark-Fremont	5120	65937	44675	12.88	21262	47.59
6	San Mateo-Foster City	4480	62435	50195	13.94	12240	24.38
7	Walnut Creek	3840	57211	47277	14.90	9934	21.01
8	Dublin	3200	47963	32966	14.99	14997	45.49
9	Millbrae	2560	44099	28370	17.23	15729	55.44
10	South San Francisco	2560	43535	42441	17.01	1094	2.58
11	Concord	2560	43247	33710	16.89	9537	28.29
12	San Ramon	2560	41982	32548	16.40	9434	28.98
13	Santa Rosa	3840	40493	37146	10.55	3347	9.01
14	North Mountain View	3840	37574	44610	9.78	-7036	-15.77
15	Redwood City	3200	37375	36093	11.68	1282	3.55
16	Belmont	1280	34713	10740	27.12	23973	223.21
17	North Union City	2560	23480	19420	9.17	4060	20.91
18	San Leandro	1920	20538	19815	10.70	723	3.65
19	San Rafael	1280	19955	17763	15.59	2192	12.34

[30] Cells are highlighted in Tables A-5-4 if they contain negative values.

Table A-5-4 Selected characteristics of 2005 employment centers (continued)

Center ID	Location	Area (acres)	Employment 2005	Employment 1995	Emp. density 2005	Employment growth 1995-2005	Percentage growth 1995-2005
20	North East Oakland Airport	1280	16782	15397	13.11	1385	9.00
21	Daly City	1920	15657	12016	8.15	3641	30.30
22	North Fremont	1280	15611	12292	12.20	3319	27.00
23	Fairfield	1280	13132	10873	10.26	2259	20.78
24	North East Santa Cruz	1280	13095	10078	10.23	3017	29.94
25	North Santa Venetia	1280	12902	13841	10.08	-939	-6.78
26	Castro Valley	1280	12330	9742	9.63	2588	26.57
27	Napa	1280	11745	9328	9.18	2417	25.91
28	Cupertino	640	11726	9531	18.32	2195	23.03
29	North Santa Cruz	640	11074	8517	17.30	2557	30.02
30	East Palo Alto	1280	10267	7697	8.02	2570	33.39

Table A-5- 5 Pairwise correlations (Employment centers identified in 2005 for Los Angeles/San Francisco region, base year=1995)[31]

	ln(E00_05)	ln(Emp00)	ln(den00)	pred_emp	DCBD	ln(DLAX/SFO)	DNrstAP	ln(Dhwy)	NWAcc	LFAcc	RLFAcc	LA
ln(E00_05)	1											
ln(Emp00)	-0.48	1										
ln(den00)	-0.74	0.62	1									
pred_emp	-0.13	-0.01	0.24	1								
DCBD	0.3	-0.37	-0.43	0.06	1							
ln(DLAX/SFO)	0.18	-0.3	-0.25	0.12	0.73	1						
DNrstAP	0.02	-0.2	-0.23	0.17	0.46	0.5	1					
ln(Dhwy)	-0.07	0.09	0.1	-0.1	-0.01	0.02	-0.03	1				
NWAcc	-0.24	0.35	0.26	-0.23	-0.71	-0.59	-0.44	-0.06	1			
LFAcc	-0.24	0.34	0.25	-0.24	-0.7	-0.57	-0.44	-0.05	1	1		
RLFAcc	-0.23	0.67	0.28	0.03	-0.19	-0.16	-0.02	0.06	0.25	0.25	1	
LA	0.05	-0.04	0.15	0.18	-0.01	0.15	-0.09	-0.01	-0.47	-0.45	-0.13	1

[31] Cells are highlighted in Tables A-5-5 if the correlation coefficients they contain are larger than 0.5.