ABSTRACT

TITLE OF DISSERTATION: THE IMPACT OF ENTERPRISE ZONES ON RESIDENT EMPLOYMENT: AN EVALUATION OF THE ENTERPRISE ZONE PROGRAMS OF CALIFORNIA AND FLORIDA

Joel Andrew Elvery, Doctor of Philosophy, 2004

Dissertation directed by: Professor Judith K. Hellerstein
Department of Economics

This dissertation estimates the impact of two geographically targeted economic development programs on the employment of people living in the targeted areas. This question is difficult to address for a number of reasons. Unlike in most program evaluation problems, the process that determines the outcome of interest (resident employment) happens at a lower level of aggregation than the process that determines selection for treatment. Therefore, standard program evaluation techniques have to be modified to address this issue. The programs I study, the enterprise zone programs of California and Florida, were designated at a very detailed level of geography, making it hard to measure the location and the characteristics of the zones.

I develop a methodology to address the unusual selection process of these programs. The first step of the methodology is to create a neighborhood-level measure of the component of residents' employment probabilities explained by the neighborhood that is conditional on the characteristics of area residents. To do this, I estimate the component of employment probability correlated with residential neighborhoods, which I call tracts' conditional employment probabilities. The next step is to estimate the effect of enterprise zones on resident employment by comparing the conditional employment probabilities of neighborhoods containing enterprise zones with those of comparable areas. I do this with tract-level propensity score matching. I also carefully measure the location and attributes of enterprise zones.

I find that a substantial portion of the variation across neighborhoods in employment rates can be explained by controlling for the attributes of residents. This indicates that it is important to control for resident characteristics when making cross-neighborhood comparisons. Using propensity score matching, I find a large pool of non-zone tracts that are observationally similar to tracts containing enterprise zones. I use these non-zone tracts to create an estimate of what the conditional employment probabilities in zone tracts would have been in the absence of the programs. Even though I focus on two very targeted and generous enterprise zone programs, I find no evidence that the programs impacted the employment of zone residents.

THE IMPACT OF ENTERPRISE ZONES ON RESIDENT EMPLOYMENT: AN EVALUATION OF THE ENTERPRISE ZONE PROGRAMS OF CALIFORNIA AND FLORIDA

by

Joel Andrew Elvery

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2004

Advisory committee:

Professor Judith K. Hellerstein, Chair
Professor Jonah B. Gelbach
Professor Marie Howland
Professor Seth Sanders
Professor Jeffrey A. Smith

ACKNOWLEDGEMENTS

So many people have enhanced my work and my life that singling out certain people for acknowledgement seems ludicrous, but I will not let that stop me. I would like to thank the Department of Economics at University of Maryland. It has been a pleasure to be a part of and it has shaped me in many ways. Thanks go to my colleagues at the Center for Economic Studies at the Census Bureau. Not only have they given me great comments and helped me grapple with many issues, they have made me laugh and stay grounded in reality.

My thanks go to Robert Greenbaum for giving my comments and sharing some of his data as I started this research. I appreciate the hard work and valuable comments of all my committee: Jonah Gelbach, Judith Hellerstein, Marie Howland, Seth Sanders, and Jeff Smith. I have been fortunate to have such an engaged and brilliant committee that brought such a broad set of strengths to my aid. My advisor, Judith Hellerstein, deserves special recognition. She provided a magic combination of patience, guidance, confidence, freedom, and pressure that enabled me to complete this dissertation. Working under her as an advisee and a research assistant have been the most rewarding educational and professional experiences I have ever had. I came to graduate school to become a research economist and she has made that possible for me. For this, I am more grateful than I can express.

My kind family has been crucial to my success. I thank my parents, Carol and Dudley Elvery, for instilling in me the curiosity, creativity, and discipline that I needed to complete this dissertation. My grandfathers, the late Frank Elvery and F Grey Butcher, instilled in my family a passion for learning that has helped me throughout graduate school. Whenever I was concerned that I was not able to complete the task at hand, my faith was restored by the unfailing support and generosity of my grandmother, Ruth Elvery, and my mother-in-law, Vivian Phares. Finally, I would like to thank my wife, Lynn Arlington Phares, without whom I would certainly not have finished my studies. I am eternally grateful for the sacrifices she has made for me to pursue my degree. She has been a great help throughout graduate school, reminding me to work when I needed to work and reminding me to play when I needed to play. I love her very much and am lucky to be a part of her life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

Framework for Evaluating Impact of Enterprise Zones on Resident Employment

1.1 Introduction

In the last two decades, policymakers have pursued alternatives to traditional welfare programs. One result has been a sharp increase in the number of enterprise zones. The United States has gone from having no enterprise zones in 1980 to having at least one, and usually many, in 40 states by 2000. Since 1994, there have also been federal Empowerment Zones that are similar to state and local enterprise zones. One of the motivations for creating enterprise zones is to increase the employment of people living in the zones. However, existing research about how these geographically targeted programs impact zone resident employment is inconclusive.

This dissertation estimates the change in zone resident employment probability due to the designation of an enterprise zone conditional on the residents' characteristics. Prior research has ignored the influence of the composition of residents on employment rates. To address this problem, I develop and implement a new estimation strategy to semi-parametrically estimate changes in zone residents' probabilities of being employed. This methodology can be applied to other topics where selection for treatment occurs at a more aggregated level than the process that determines the outcome of interest.

Enterprise zones in California and Florida generally do not have the same boundaries as more common geographic areas, such as Census tracts or Zip Codes. In order to minimize the measurement error that arises from the need to use real world data, I map enterprise zones at a very detailed level of geography and use 1980 Census tracts as my unit of analysis. I show that 1980 Census tracts are a much more accurate measure of zone location than Zip Codes, the unit of analysis used by many studies of enterprise zones. This precision also allows me to estimate the spillover effects of enterprise zones on people living near but not in the zones.

Much of the literature that has examined the impact of enterprise zones on resident employment has found negative effects on employment outcomes. These results are counterintuitive because there is, at best, a weak theoretical basis for zones reducing resident employment. I also find that it is possible to estimate a negative effect of enterprise zones on the employment rate in the zones. However, when I properly condition on the characteristics of zone residents, I find that the estimated effect of the enterprise zone programs of California and Florida on resident employment probability is zero. This clearly demonstrates that ignoring the characteristics of residents when studying the employment effects of these geographically targeted programs can yield misleading results.

The dissertation is organized in three interrelated chapters. This chapter describes enterprise zones, provides detail on the programs of California and Florida, introduces the methodology I use, and describes the data that make this methodology feasible. The next chapter focuses on the estimation of the component of employment probability related to residential neighborhood and shows why this is a superior outcome measure for evaluating geographically determined policies. The last chapter integrates the estimates of the neighborhood-component of employment probability with non-experimental evaluation techniques to estimate the impact of enterprise zones on resident employment.

1.2 Background

Enterprise zones are programs that provide incentives for businesses to operate in certain economically distressed areas selected by the government. The concept of enterprise zones stems from the work of Peter Hall, a British urban planner, who saw them as a way to encourage development in declining industrial cities. The first enterprise zones were enacted in England in the late 1970's. In the early 1980's U.S. policymakers, especially those interested in changing anti-poverty programs, embraced enterprise zones. State and local governments were the first to enact these policies, often as a response to the loss of manufacturing jobs many regions experienced in the 1980's (Van Allen 1995). The types of incentives offered to businesses in the enterprise zones included: reduced property taxes, reduced capital taxes, wage subsidies for hiring zone residents, access to below market rate loans, and less stringent regulation. The types and dollar values of the incentives varied from zone to zone, but all had the potential to reduce businesses' operating costs and make the zones more attractive business locations than they would have been in the absence of the programs.

One of the goals of creating enterprise zones was to increase employment in poor, high-unemployment communities. Policymakers believed that encouraging business activity in disadvantaged areas would increase the employment of the residents of these areas, as suggested by the theory of spatial mismatch (e.g. Kain 1968). Several states used unemployment and/or poverty rates as criteria for designation of a zone. Some enterprise zone programs provided incentives for hiring disadvantaged workers. For example, California offered a wage subsidy to zone businesses of up to 50 percent of wages paid to participants in Job Training Partnership Act (JTPA) or welfare-to-work programs.

Economic theory suggests that enterprise zones may increase resident employment. In zones where subsidies are provided for hiring zone residents, businesses have an incentive to hire zone residents rather than non-zone residents because the net cost of compensation for zone residents would be lower than for non-zone residents. In addition, the subsidies might be used to increase the wages of existing jobs for zone residents, perhaps causing wages to rise above individuals' reservation wages and moving them into employment. Even if enterprise zones do not have specific incentives for hiring zone residents, one would still expect a rise in zone residents' employment probabilities if the policies lead to additional jobs locating in the area. Establishments in an enterprise zone would be able to offer zone residents a lower wage than other workers because the residents would have lower commuting costs, so there would be an incentive to hire zone residents.

Resident employment probability might not rise even if the enterprise zones increase the number of jobs in the area. If the subsidies offered in enterprise zones attract businesses that require workers with skills not possessed by zone residents, then there would be no boost to zone resident employment probability. This might be the case because the value of enterprise zone incentives differs greatly from establishment to establishment (Papke 1994). A labor-intensive establishment will find little value in capital or property based incentives, while a capital-intensive establishment would benefit less from payroll-based incentives. If enterprise zone incentives induce investment into capital that substitutes for the labor of zone residents, the policies could reduce the employment probabilities of zone residents by reducing demand for their labor Therefore, the enterprise zones could either increase or decrease the employment of zone residents depending on how the composition of businesses changes in response to these incentives. I estimate the net effect of these different influences, both positive and negative, of enterprise zones on zone resident employment probability.

1.3 Enterprise Zones in California and Florida

The programs of California and Florida share a number of key features, such as competitive designation of zones and a focus on reducing poverty in the zones. In both states, the zone policies were set primarily at the state level.[1] Both California and

---

[1] In some states, such as Pennsylvania, the state designated enterprise zone programs varied greatly from zone to zone.

Florida had a number of enterprise zones, thirteen and thirty respectively, that were designated in 1986 and implemented at the beginning of 1987. These states also provided more substantial incentives than most states. Lastly, detailed maps of where zones were located during the late 1980's were available for both states. This section summarizes and contrasts the programs of the two states.[2]

1.3.1 California's Enterprise Zone Programs

The enterprise zone program in California was carefully targeted and provided substantial incentives for businesses located in the zones to hire disadvantaged workers. In 1984, California enacted two enterprise zone programs at the same time as a legislative compromise. The Enterprise Zone Act allowed ten zones to be designated in 1986. The Employment and Economic Incentive Areas (EEIA) Act created three zones per year in 1986, 1987, and 1988.[3] The two programs were consolidated in 1994 (California Legislature 1999). Following other research (Dowall 1996, Greenbaum and Engberg 2000, O'Keefe and Dunstan 2001, Bostic and Prohofsky 2002), the two programs will be treated as identical.

The designation process in California was focused on choosing areas that demonstrated need for assistance and potential to benefit from the program. It was a two-tier process implemented by the state's Department of Commerce. Local jurisdictions filed applications with data from the 1980 Census of Population and Housing to establish need and provided descriptions existing and planned business infrastructure to demonstrate development potential. The first tier of the designation

[2] For more detail on the programs, see California Department of Commerce (1987) and Office of the Auditor General (1993).

process was focused on choosing areas that had poor socio-economic conditions, both absolutely and relative to the surrounding counties. The second tier chose the areas with the highest development potential from among the distressed areas. The two programs used similar designation criteria, though the EEIA program was eligible only to areas that combined High Density Unemployment Areas (HDUA) with industrial areas.

Once designated in October of 1986, the incentives of the programs were available starting in January of 1987 and they lasted through 2002. The State of California offered a number of incentives that would reduce the corporate income taxes of businesses located in enterprise zones. When the legislature reviewed the enterprise zone policies in 1998, they concluded that the hiring tax credit was the most substantial incentive (California Legislature 1999). The hiring tax credit allowed businesses to reduce their tax bill by as much as 10 to 50 percent of the wages paid to workers enrolled in specific job training or welfare-to-work programs.[4] The credit started at 50 percent in the first year of employment and declined to 10 percent for the fifth year. While the credit applied to qualified workers regardless of their wage, the wage used to determine the credit was capped at 150 percent of the minimum wage.

The other incentives offered by California were more typical of those offered by other states. Businesses could receive a Corporate Income Tax credit for sales tax paid on purchases of manufacturing machinery to be used in the zone. They were also

[3] Only the three EEIA zones designated in 1986 are included in this analysis.
[4] In order to prevent fraud, qualified workers were given vouchers that their employers were required to submit when claiming the tax credit. Only wages covered under unemployment insurance were eligible so that the state could confirm that the firm was paying the individuals they claimed.

permitted to deduct from corporate income up to $10,000 of purchases of non-real estate, depreciable business property to be used in the enterprise zone. In order to encourage start-up business to locate in the zones, businesses were allowed to carry-over net operating loss for up to 15 years.[5] Net interest income from third party investments in enterprise zone activities could be deducted from corporate or personal income taxes.

The incentives available to businesses in an EEIA were quite similar to those offered to businesses in an enterprise zone. One difference is that location in an EEIA was not sufficient to qualify for the incentives. In order to take advantage of the incentives either: at least 50 percent of the business's employees had to live in a HDUA; at least 30 percent of employees had to be HDUA residents and the business had an approved community service plan; or 30 percent of the business's owners were residents of a HDUA. The types of incentives offered were the same as in the enterprise zone program, but the details of some of the incentives changed. The Credit for Hiring the Unemployed applied to HDUA residents who were unemployed at the time of hire and was less generous than the Hiring Tax Credit in enterprise zones. The business expense deduction was more generous than in enterprise zones, especially in the first four years after designation. The Sales and Use Tax Credit, Net Operating Loss Carryover, and Nontaxable Investments incentives were effectively the same in both programs.

According to the Franchise Tax Board of the State of California, the Sales and Use Tax Credits and the Hiring Tax Credits accounted for $9.9 million in tax credits given to businesses as a result of the enterprise zone/EEIA programs between 1986 and 1990 (California Legislature 1999). Of this amount, \$6.6 million was due to the Hiring Tax Credits. The sum of the tax credits from 1987 to 1990 is equivalent to \$450 per unemployed adult 1990 resident. Unfortunately, the dollars spent are not available on a per zone basis during this period.

In summary, the enterprise zone and EEIA programs in California designated zones in distressed communities that the application reviewers believed could be influenced by the tax incentives that were offered. The incentives offered were a combination of deductions and tax credits for expenses related to expanding or upgrading business and tax credits for hiring targeted workers. In enterprise zones, the targeted workers were those who were receiving job training or search assistance from the state. Unemployed residents of HDUAs were targeted by the hiring credits in EEIAs. Both programs put an emphasis on encouraging program area businesses to hire workers with a history of unemployment.

1.3.2 Florida's Enterprise Zone Program

After the violence in Miami in 1980, Florida became the first state in the U.S. to implement the enterprise zone concept. They developed a program that targeted "slum and blighted areas" for redevelopment through a combinations of tax and regulatory incentives. After frustration with a very broad enterprise zone program that was both expensive and unproductive, the Florida legislature passed a law in 1984 radically changing the states enterprise zone policies (Logan 1991). This law reduced the number of zones to be designated, required a competitive process for designation,

---

[5] Five years for financial institutions.

and set designation criteria that ensured that enterprise zones would go to needy areas. The program created by this law will be described here. Enterprise zone policies in Florida were adjusted in almost every year after 1986, though most of the changes were minor until 1993.

In 1986, the Department of Community Affairs designated six zones in each of five different population categories, for a total of 30 zones.[6] The enterprise zone program was in effect from the first day of 1987 through 1994. Designation was competitive within size category to ensure that the zones would be located in a variety of areas throughout the state. Municipalities and counties applied to the Department of Community Affairs to request designation for zones. The application included the boundaries of the proposed area, information on the condition of businesses in the area, information on the population and housing characteristics of the area, the local development plan, and the local incentives to be offered. In order for an area to be eligible to be designated an enterprise zone, it had to meet three eligibility criteria. The first was that the population of the zone had to be no more than the greater of: 2,500 people, 10 percent of the population of the jurisdiction where it is located, or the percentage of the population of the creating jurisdiction equal to the percentage of families with incomes below the poverty line in the county in which the enterprise zone is located.[7] The boundaries of the proposed enterprise zones had to be

[6] The five size categories were jurisdictions having a population of: fewer than 7500 people; 7,500 to 19,999; 20,000 to 49,999; 50,000 to 124,999; and 125,000 or more.
[7] This eligibility option was created in 1986 so that the North Central Dade zone could be larger.

continuous so that the enterprise zone was a single area.[8] The other eligibility restriction was that at least 40 percent of the land in the zone had to be zoned residential and at least 40 percent had to be zoned commercial or industrial.

The decision of which applications to choose as enterprise zones was based on the level of need and the local participation plan, with weights of 65 and 35 percent respectively. The factors used to determine the level of need were: the age, quality, and vacancy of housing units; unemployment and poverty rates; per capita income and its change from 1970 to 1980; percentage change in per capita taxable value of property in the area from the prior year; and per capita local taxes. The local participation plan was judged based on the incentives planned, with the most weight given to tax abatement and then to improving infrastructure or services. The greatest level of competition for designation was in jurisdictions with 25,000 to 124,999 people. In total, there were 43 applicants for 30 zones.[9]

Most incentives were available only to businesses located in the enterprise zones. These were: a property tax credit, a sales tax exemption for electrical energy used in an enterprise zone, a sales tax exemption for business property used in an enterprise zone, and a sales tax exemption for building materials. Except for the exemption for building materials, at least 20 percent of a business's employees had to be zone residents in order for the business to claim these tax credits. Florida had an unusual set of incentives that applied to all businesses in the state, not just those

[8] In fact, there were exceptions to this rule. Some zones had discontinuous parts, specifically the Haines City, Umatilla, and Ft. Meyers zones. The Ft. Meyers zone bordered three of the four sides of the Franklin Park zone, so they are treated as one zone in this study.

located in the zones. One incentive eligible to all businesses in the state of Florida was a Community Contribution Tax Credit that could be applied to a number of different taxes. A business received a credit of 50 percent of the amount of an approved community contribution to zone development projects, with a maximum credit per business at $200,000.[10] Businesses hiring eligible workers could choose between a corporate income tax credit of 15 percent of wages paid or a monthly sales tax credit of $100 for full-time employees and $50 for part time employees. The credits could be claimed by businesses located in a zone for two years and by other businesses for one year. To be considered an eligible employee, an individual had to be a zone resident, an AFDC recipient, or a JTPA participant for the prior three months and the monthly wages paid could be no more than $1,500. Prior to 1988, the income tax credit was 25 percent of wages paid and JTPA participants were not a category of eligible employees. Also, for non-zone businesses only zone residents were eligible hires. The majority of the hiring tax credits were used for zone residents. In fiscal year (FY) 1990, the employers of 3,793 zone residents and 89 AFDC recipients benefited from the tax credit. No credits were claimed for hiring a JTPA participant (Office of the Auditor General 1993).

In the first three years of the enterprise zone program, 1987-1990, businesses received $26,542,340 in tax credits.[11] The Enterprise Zone Jobs Corporate Income

---

[9] Of the 30 zones that were designated, 25 were in jurisdictions that had enterprise zones in the earlier programs, though it is not possible to tell whether they were the same areas.

[10] Most of the businesses taking advantage of this credit were financial institutions.

[11] Data is missing for the equipment and energy tax credits for 1987. The greatest dollar value for the equipment credit was $47,363 in 1988. The corresponding number for the energy tax credit was $4,846 in 1989. It is likely that the values in 1987 were similarly small.

Tax credit was by far the dominant tax credit. The value of this credit ranged from $9.87 million in FY 1987 to $3.59 million in FY 1989, for a three-year total of $21.3 million. More than 80 percent of these credits were claimed in the first two years and it appears that these credits were claimed for employees hired prior to the start of the zone program. The second largest incentive was the property tax credit, with a three-year total of $4.2 million. The remaining credits accounted for less than $1 million in forgone taxes. The sum of the tax credits from 1987 to 1990 is equivalent to nearly $1,400 per unemployed adult 1990 resident.

1.3.3 Comparison of Zone Programs in California and Florida

The enterprise zone programs of California and Florida shared a number of features. Both states used a competitive designation process in which jurisdictions applied to have one of a fixed number of zones created in their area. The applications were judged on the basis of need, local incentives to be offered, and potential for success. Incentives were offered to employers in both states to hire zone residents or people with a history of unemployment. While the programs of cost each state millions of dollars, the costs per unemployed zone resident were small relative to many training programs.

There are also some distinct differences between the programs of the two states. In Florida, the wages subsidies were offered first for hiring zone residents and only later for hiring other disadvantaged workers. In California, the wage subsidies were offered based on the unemployment and public assistance history of the worker. The Florida program chose areas in a wide number of size classes and in rural areas while all but three California zones were located in larger cities.[12] The California program was more targeted than Florida's. In 1990, 1.5 percent of California's population lived in one of the state's 13 enterprise zones while in Florida 2.9 percent of the population lived in one of the state's 30 zones.

1.4 Enterprise Zone Evaluations

There has been an active literature regarding the effectiveness of state enterprise zone policies in increasing the total number of people employed in the zones, regardless of where they live. Case study research has found a great degree of variation in policy effects.[13] There have been studies of this type for the California and Florida programs (Dowall 1996, Office of the Auditor General 1993). Each of these concluded that the zones had no effect. These studies compared zones to larger geographic areas, such as the counties they are located in, but did not compare zones to similarly distressed areas. Since enterprise zones were created in areas with high levels of poverty and unemployment, it is unlikely that the surrounding county is a relevant comparison to an enterprise zone.

A recent book by Peters and Fisher (2002) provides a thorough summary of the literature regarding the effectiveness of enterprise zones in attracting businesses and presents their analysis of enterprise zones. Their empirical work, which primarily uses Zip Code as the unit of analysis, uses a sample of zones drawn from 13 states and estimates the value of tax incentives offered in zones with representative firm models. They show that enterprise zone incentives are small relative to other business costs.

---

[12] This difference is not likely to be important in reality because very few dollars were spent in the two smallest size classes in Florida (Office of the Auditor General 1993).

Consistent with other studies, they find little growth in the number of jobs and businesses in enterprise zones. Using the Census Transportation Planning Package to study a small number of zones, they show that nearly 10 percent of employed zone residents work in the zone, while roughly 20 percent of zone employees reside in the zone. Unfortunately, they do not provide similar numbers for comparable non-zone areas.

More recent analyses of the California enterprise zone program in the 1990's have found positive employment effects. O'Keefe and Dunstan (2001) use a propensity score matching estimator to test whether tracts containing enterprise zones had higher growth in the number of jobs located in the zone tracts and wages paid by those jobs relative to similar non-zone tracts. Using establishment data and controlling for only the 1990 characteristics of tracts, they find that zones on average had a higher rate of job growth than non-zone areas but lower growth in wages. While their work provides evidence that zones did encourage business development in the 1990's, it does not tell us whether zone residents benefited from the jobs. Another recent paper by Bostic and Prohofsky (2002) uses individual tax returns from 1993 to 1997 to study whether workers employed in zones had faster than expected wage growth. They find that wages increased faster for workers employed in enterprise zones and these workers were more likely to file taxes each year than other similar workers, suggesting more sustained employment. Their results suggest that zone policies did benefit workers, but their analysis cannot distinguish between zone residents and other workers who were hired using the tax credits.

---

[13] See Wilder and Rubin (1996) for a thorough summary of this research.

There have been a few other studies of the impact of enterprise zones on resident employment. In Papke's (1993) work regarding the employment impacts of enterprise zones, she focuses on a set of enterprise zones from Indiana. She finds that the policies had only small effects on resident employment, reducing unemployment by 0.15 percentage points. In other work, Papke (1994) finds that unemployment claims in unemployment offices near enterprise zones fell by 25 percent relative to what they would have been in the absence of the programs. This result runs contrary to her earlier finding of small employment effects, but this difference could be due to some empirical issues. The result is based on 10 enterprise zone unemployment offices and 12 control unemployment offices, so the sample size is small. Furthermore, the offices all served a much larger area than just the enterprise zones, so the change cannot be directly attributed to enterprise zones.

One important strand of the literature on enterprise zones is the work by Bondonio, Engberg, and Greenbaum (1998, 1999, 2000, 2001, 2003). Their research is notable because it addresses the issue of selection for designation of enterprise zones and they collected a large database regarding enterprise zones in several states. Their papers explore the impact of zones on business outcomes by using establishment-side data aggregated to the Zip Code-level. In general, they use data from six states and choose a sample of non-zone areas that are similar to the zone areas to use as controls. These authors have consistently found no significant impact of enterprise zone policies on the number of jobs created and little impact on other business outcomes.

In Greenbaum and Engberg (2000), the authors use Census of Population and Housing data from 1980 and 1990 aggregated to the Zip Code-level to estimate the impact of enterprise zones on the growth rates of employment, income, poverty status, population, housing value, occupancy, and home ownership. The zones evaluated are drawn from California, Florida, New Jersey, New York, Pennsylvania, and Virginia. They find that, relative to similar non-zone areas, enterprise zones had higher rates of growth in poverty and unemployment.

While the work of Engberg and Greenbaum is a substantial contribution to the study of the impact of enterprise zones on zone residents, there are a few reasons to question its results. One issue is the method they use to control for Zip Code characteristics prior to designation. They estimate the probability that a zone was designated an enterprise zone as a function of several Zip Code characteristics, including rates of poverty and unemployment in 1980 and changes in the number of jobs in the Zip Code prior to designation. The estimated probability, or propensity score, is the only variable they include in their models to control for Zip Code characteristics. This means that their estimate depends on a linear relationship between growth rates and their propensity score, which is not the usual method for conditioning on propensity score and imposes a severe functional form restriction on the relationship between the observable characteristics and the outcomes of interest. [14] Another issue is that they use Zip Codes as their unit of analysis, which provides an imprecise measure of zone location. Below, I find that Zip Code defined enterprise

[14] See Rosenbaum and Rubin (1983) and Black and Smith (2004) for more typical examples.

zones are often more than five times as large as the true enterprise zones. This is a troubling degree of measurement error in the variable of interest.

One issue that has not been addressed by the literature is the impact of the composition of residents on aggregated measures of resident employment. This is important because zones may have had different demographic trends than other areas. Suppose that after zone designation enterprise zones experienced a sharp exogenous fall in the human capital of residents relative to non-zones. In this case, one would expect a fall in employment rates in the zones due to differences in the characteristics of zone residents. The absence of such a fall would suggest that zones increased the employment probabilities of zone residents. Therefore, looking at employment outcomes without conditioning on the traits of residents could be misleading. In order to estimate whether enterprise zones changed the employment probabilities of residents, rather than the employment rate in the zone, it is necessary to control for the characteristics of zone residents. Another reason it is important to control for the characteristics of zone residents is that individual characteristics are a more important determinant of employment than where one lives. A small difference in the education levels in a neighborhood is likely to have a larger influence on the neighborhood's employment rate than the effect of the neighborhood. In order to identify the neighborhood effect on employment, it is necessary to control for the confounding effect of resident characteristics.

While the literature regarding the evaluation of enterprise zones has grown and improved over time, we still know very little about how zones affected residents. The rationale for such a geographically targeted program is that it should concentrate economic growth in the areas chosen and benefit the members of those communities. This study addresses specifically the question of whether zones increased the employment of people living in tracts that contained a zone. The methodology I develop and implement allows me to condition for the characteristics of zone residents and estimate the effect of enterprise zones on resident employment probability, rather than on the unconditional employment rates used by the rest of the literature. This is possible because I use individual-level data from the 1990 Census rather than aggregate data. Using individual-level data also permits me to use more precise, time-consistent geographic areas as my units of analysis.

1.5 Methodology

This dissertation uses a three-stage estimation strategy to estimate the effects of enterprise zones on resident employment probability. The first stage uses employment probability models in order to calculate the neighborhood effect on employment conditional on the characteristics of the people who live in the neighborhood. The second stage estimates the propensity for an area to be designated an enterprise zone. The third stage estimates the effect of enterprise zone policies on resident employment by matching on the estimated propensity score. This section discusses in detail the estimation strategy and how it is implemented.

The parameter of interest in this thesis is the average effect of containing an enterprise zone on resident employment probability for areas containing a zone, conditional on the traits of residents. This is also called the treatment effect on the treated, where the treatment for a neighborhood is containing an enterprise zone. More formally, the parameter of interest is:

$$\Delta = E[Y_1 - Y_0 \mid T = 1, X = x],$$

where T=1 if the area contains an enterprise zone, Y_0 is the employment rate in an area in the absence of an enterprise zone, Y_1 is the same with an enterprise zone, and X is a vector of the demographic characteristics of the people who live in the area. What makes this non-trivial is that it is not possible to observe Y_1 and Y_0 for the same area and that it is necessary to condition on X. The first stage conditions on X by estimating the probability that an individual is employed as a function of their own characteristics as well as area fixed effects. The model estimated is:

$$y_{ij} = f\left(\beta X_{ij} + \alpha_j + \varepsilon_i\right),$$

where i indexes individuals and j indexes areas, $y_{ij} = 1$ if individual i in j is employed and 0 otherwise, X_{ij} is a set of characteristics of individual i in j, α_j is an area fixed effect, and ε_i is an error term. Because α_j is conditional on X_{ij}, the parameter of interest becomes:

$$\Delta^c = E[g(\alpha_1) - g(\alpha_0) \mid T = 1] = E[g(\alpha_1) \mid T = 1] - E[g(\alpha_0) \mid T = 1],$$

where α_1 is the area effect if T=1 and α_0 is the area effect if T=0 and $g(.)$ is a function that maps the coefficient estimate to a marginal effect.

By estimating the fixed effect, the estimation problem becomes like other program evaluation problems, where the difficulty is in estimating the counterfactual, $E[g(\alpha_0) \mid T = 1]$. I use a propensity score matching estimator for that purpose. The motivation for using propensity score matching rather than regressions to estimate the counterfactual is that enterprise zones were designated in a small number of distressed areas. The vast majority of areas are not similar to enterprise zones, so most non-treated areas provide little information about what would happen to enterprise zone areas in the absence of the programs. Propensity score matching resolves this problem by systematically selecting relevant comparison areas from a large pool of mostly irrelevant areas. Also, enterprise zones would be outliers in most regressions of area traits on employment outcomes because they are so distressed. Therefore, models that fit most areas are likely to fit poorly for enterprise zones. Matching estimates do not suffer from this problem because matching does not impose a specific functional form on the relationship between observable characteristics and the outcome of interest.

To use propensity score matching estimators it is necessary to assume: selection into treatment is a function of observable characteristics *Z*, there exists a set of observations *S* with similar *Z* that contain areas with *T*=1 and areas with *T*=0 (the common support condition), and $g(\alpha_0) \coprod T \mid Z \in S$ (the conditional independence assumption).[15] These conditions are a summary of the conditions in Heckman, Ichimura, Smith, and Todd (1998). It is also assumed there is a function such that $\Pr(T=1 \mid Z) = P(Z)$. Then by propositions one and two in Rosenbaum and Rubin (1983), it follows that:

[15] $\coprod$ symbolizes independence.

$$g(\alpha_0) \coprod T \mid P(Z), Z \in S$$

$$\therefore E[g(\alpha_0) \mid T=1, P(Z), Z \in S] = E[g(\alpha_0) \mid T=0, P(Z), Z \in S].$$

The estimated parameter is now:

$$\Delta^f = E\{E[g(\alpha_1) \mid T=1, P(Z)] - E[g(\alpha_0) \mid T=0, P(Z)]\} \mid Z \in S$$

which differs from Δ^C by requiring that observations be in the common support region. This estimate also ignores general equilibrium effects, so the stable unit treatment value assumption must hold for all areas in the analysis (Lechner 2001). This assumption requires that the zone programs do not induce selective migration. In the final chapter of this thesis, I provide evidence of the plausibility of that assumption. It also must be assumed that individual characteristics do not change as a result of living in a treated area.

Of the assumptions necessary to use this estimator, the one that usually raises the most concern is that selection is strictly on observable characteristics. It is possible that unobservable characteristics influenced which of the areas that met the states' criteria were designated enterprise zones. However, as Greenbaum and Engberg (1999) noted, enterprise zones were designated by state governments in accordance with policies that outline specific levels of poverty, unemployment, or other observable characteristics. In California and Florida, much of the legislated selection process depended on data from the 1980 Census of Population and Housing similar to that which I use to estimate the propensity scores. Therefore, the concern about selection on unobservables is less problematic than in many other contexts.

The effects of enterprise zones are estimated separately for California and Florida. This is because the selection process and incentives offered differed by state, so it would be improper to treat them as equivalent programs. The second chapter of this dissertation details the employment probability models that comprise the first stage and the resulting estimates of the neighborhood component of employment probability. The estimation of $P(Z)$ and the matching estimates of the effect of enterprise zones on resident employment are discussed in detail in the third chapter, which also contains my conclusions.

1.6 Data

The remainder of this chapter discusses the data work that is the foundation for all of the empirical work. Five kinds of data are necessary for this study: data on the location of zones, demographic data from prior the designation of the zones, demographic data from after the designation of the zones, information on employment growth prior to designation, and data that bridge the geographic codes used on each data source. Maps of each of the zones are used to define where enterprise zones were located. Data on the characteristics of the population by tract from 1970, 1980, and 1990 come from the Censuses of Population and Housing. The individual-level data from the 1990 Census long-form sample are also used to estimate the tract effects on employment probability and to create 1980-tract-level statistics from the 1990 Census. In order to control for pre-designation trends in local employment growth, I make special tabulations from the source data for the County Business Patterns data, the

Standard Statistical Establishment List.[16] Finally, to link detailed areas across time, I use geographic information system (GIS) data about how geographic codes changed over time.

This dissertation uses three units of geography that may be unfamiliar to some readers: Census place, tract, and block. The following brief descriptions of these measures are derived from the Census Bureau's Geographic Area Reference Manual (1994), which contains more complete information. Census geography changes somewhat for each decennial census. Census places are typically incorporated areas, such as cities, towns, and villages. They range greatly in size and there is no established minimum size because states have different incorporation criteria.[17] There are also Census designated places, which are unincorporated areas that resemble incorporated areas and, in the case of urban areas, have a population above 2,500.[18] Places vary greatly in population and land area and can cross county lines. For Economic Censuses, data are tabulated for incorporated places with population of at least 2,500, as well as for certain economically important unincorporated places. The Census Bureau also assigns Census tracts as a more detailed unit of geography. Tracts usually have boundaries that follow visible features, population between 2,500 and

---

[16] The individual-level data on people and establishments used in this study are restricted access, confidential data. I have been able to use them because I am employed by the Bureau of the Census's Center for Economic Studies. There is also a process for outside researchers to access similar data. For more information, see http://www.ces.census.gov.

[17] Areas that are close to meeting the criteria for place can encourage the Census Bureau to designate the area as a place.

[18] In 1980, the minimum population for CDPs in larger (smaller) urban areas was 5,000 (1,000). In both 1980 and 1990, the minimum population for a CDP in a rural area was 1,000. There are exceptions in both years because some areas had lower than anticipated populations.

8,000, and do not cross county lines.[19] In my sample, the median tract size is 0.67 square miles and the median 1980 population is 3,501. As a point of reference, an urban Zip Code typically contains four to six tracts, though their boundaries are unrelated. Census blocks are the most detailed unit of geography that the Census Bureau releases data for and, since 1990, all other units of Census geography are aggregations of blocks. In an urban area, a block is often a single city block. Like tracts, block boundaries typically follow visible features. A block group is a group of contiguous blocks and provides a level of detail between tract and block.

1.6.1 Enterprise Zone Location

In California and Florida, enterprise zones do not conform precisely to any geographic areas for which data is available. In order to analyze the zones, it is necessary to convert their ad hoc geography to more common statistical units. After experimenting with alternative definitions, the unit of analysis for this study is 1980 tract-place pair. This is superior to other alternatives because it is possible to build precise, time-consistent definitions of the zones. For example, Greenbaum and Engberg (2000) use Zip Codes as their unit of analysis because they have establishment-level data tabulated at the Zip Code-level and they aggregate 1990 block groups and 1980 tracts to Zip-Code-like areas. Each type of aggregation leads to less precision in the measurement of zone status and area characteristics. This is especially true for Zip Codes, which often have boundaries that are unrelated to those of zones or common statistical units.

[19] Tracts are delineated by committees of local data users or by governments where no such committee is feasible.

I mapped each of the enterprise zones designated in 1986 in California and Florida with GIS software using the Census Bureau's 1992 TIGER/Line files as source data. The 1992 TIGER/Line files provide geographic information at the polygon-level, which is at least as small as a Census block.[20] The 1992 TIGER/Line data have Census geographic codes for 1990 and 1980. By mapping the zones using the TIGER/Line data, it is possible to get a very precise definition of where zones were located and then compare alternative definitions of zone location. The source maps of zone locations were published by the states where the zones were located and I manually mapped the zones to the polygon-level.[21] Except for zones that incorporated rural or water areas, which have very large polygons, the polygon-defined zones are very similar to the source maps.[22]

Tables 1.1 and 1.2 show how different measures of enterprise zone location relate to each other in California and Florida respectively. The first column has the area of the zone calculated from the polygon-level. In California, the zones ranged from 0.46 square miles to 17.38 miles, with an average of 7.11 miles. Of the thirteen zones, eight are between 2.5 and 10 square miles. In Florida, there was more dispersion in the size of zones, ranging from 0.14 to 26.53 square miles with an average of 3.29. The largest zones in Florida were in Jacksonville and North Central Dade. The average number of 1990 residents in a zone was 33,681 in California and

[20] The X-Tools extension for ARCView 3.2 was used to augment the TIGER/Line data with the area of each polygon.

[21] The maps from Florida were very precise and appear to have been generated from similar TIGER/Line files. In California, the boundaries were not as clearly delineated, but the maps showed the shape of the zones and showed major roads in and around the zones, which permitted accurate mapping.

13,910 in Florida. The population size in zones is more heterogeneous in Florida than California, reflecting Florida's decision to place zones in a variety of areas across the state.[23] The Jacksonville and North Central Dade zones were particularly large, containing 63.5 percent of all zone residents. In California, the most populated zones were Watts and San Jose, with 43.8 percent of all zone residents.

The most detailed geographic area for which Census data are available is Census blocks, so I convert the polygon definition of enterprise zones to a block definition. Any block that contains part of a zone is counted as being within an enterprise zone.[24] The block-based definition of zone location is very similar to the polygon definition.[25] In order to use data from the 1980 Census to find control areas, I convert the block-defined zones to geographic codes that were used in the 1980 Census. Publicly available data for areas smaller than Census tract is very limited, so I use 1980 Census tract-place pairs, which I will call tracts throughout this thesis, as the unit of analysis.[26] Tables 1.1 and 1.2 show how tracts can define zones. If one uses all tracts that contain any part of an enterprise zone to define zone location, then the percent of the 1990 population of the tract defined zone that actually lives in the zone ranges from 3 percent to 100 percent, with an average across zones of 47 percent in

---

[22] Because I restrict the analysis to predominantly urban tracts, the poorly mapped zones are not included in the final analysis.
[23] This is also indicative of Florida's decision to let zone size vary with a city's poverty rate.
[24] Identifiable water blocks were excluded because they can contain the population of ships that are not clearly residents of the area.
[25] In California, the polygon-defined area is no less than 73 percent of the block-defined area, with an average of 90.76 percent across zones. In Florida, all but one zone has overlap of 89 percent or higher, with an average of 95.9 percent. The exception was Immoklee, where only 12 percent of the block area was in the polygon area. This zone contained a number of unpopulated areas that led to larger blocks.
[26] I find the combination of Census place and tract more accurate than tract alone.

both states. If instead one counts only tracts where 25 percent or more of residents live in an enterprise zone as being zone tracts, the averages in Florida and California move up to 71 percent and 74 percent respectively. One small zone in each state is excluded from the analysis by using this restriction.[27] Using the 25 percent cutoff to count a tract as within a zone also means that some people who live in a zone are not included in the analysis. When looking across zones in California, 86 percent of people who lived in an enterprise zone lived in a tract where 25 percent or more of the 1990 population lived in a zone. The corresponding number for Florida is 88 percent. If a more restrictive cutoff were used, the percent of people who actually live in zones that live in the tract-defined zones would fall. Therefore, unless mentioned otherwise, tracts with more than zero and less than 25 percent of 1990 population living in a zone are dropped from the sample.

Figures 1.1a to 1.1d illustrate how accurate the definition I use is relative to alternate definitions. Most of the enterprise zone evaluation literature uses Zip Code as the unit of analysis. The darkest gray areas indicate Zip Codes that contain an enterprise zone, the lighter gray shows tracts that contain enterprise zones where no more than 25 percent of the tract's 1990 population lives in the zone, and the white areas are those tracts where more than 25 percent of the 1990 population live in the zone -- the definition of zone location I use. The dark line is the boundary of the zone. Figure 1.1a shows a typical case, the Ft. Lauderdale, FL enterprise zone. Its boundaries are oddly shaped and it is impossible to exactly match the zone boundaries using 1980 tracts. However, using all tracts rather than those where at least a quarter of the population lives in the zone would yield an area nearly double the actual size of the zone. Using Zip Codes to define zone location would yield an area much larger than the actual zone and would include areas more than four miles from the zone boundary as part of the zone. Because these programs are tied directly to geography, such measurement error could be very problematic. Figure 1.1b shows a similar map for the Jacksonville, FL enterprise zone, which is typical of zones in densely populated areas because the boundary of the zone matches closely with a group of 1980 tracts. Figures 1.1c and 1.1d provide examples from California, the zones of San Diego and Fresno.

1.6.2 1980 Tract-level Data from 1970, 1980, and 1990

The data regarding the demographic and housing characteristics of tracts are from the Censuses of Population and Housing from 1970 through 1990. The data from 1990 are my own tabulations using restricted access, individual-level data. Unfortunately, similar individual-level data from 1970 and 1980 are not available at this time.[28] I use publicly available data from the 1970 and 1980 Censuses.

The 1980 data are drawn from the 1980 Summary Tape File 3A (STF3A), which contains population and housing counts for a wide variety of characteristics at various levels of geography. One limitation of these data is that smaller areas have a high rate of suppression for very detailed tables.[29] I deal with the problem of suppression in two ways. First, only tracts with 1980 population of 100 or more are

---

[27] The excluded zones are Porterville, CA and Tarpon Springs, FL.
[28] There is a project working to reclaim such data now, but this work is in progress as of the writing of this thesis.

used in the analysis. This restriction also helps to reduce measurement error for items drawn from the one-in-six sample of households that receive the longer survey form, called the long-form sample. This sample provides data on employment, income, education, and a host of other topics. The second way I cope with suppression is I do not include traits broken out by race or ethnicity in my models using tract-level data from the Censuses prior to 1990. These items, such as the percent of blacks with more than a high school degree, have high suppression rates and including variables derived from these traits could create a problem with missing-data. Some predominantly rural counties were not given Census tracts in the 1980 Census, so I do not include them in the analysis.[30] The 1970 data come from the Fourth Count A files for population and housing, which are very similar in nature to the 1980 STF3A data. These files have data at the Census tract-level. The link between 1970 and 1980 tracts was made with a file created by the Census Bureau that relates 1980 Census tracts to 1970 Census tracts. Of the 9,234 tracts with 1980 population of at least 100, 13 percent were not tracted in 1970.[31] The asterisks in tables 1 and 2 indicate which enterprise zones were not given tracts in 1970.

This study uses data from the 1990 Census of Population and Housing in two ways. In order to get 1990 data at the 1980 tract-level, I make tabulations from the 1990 Sample Edited Detail File (SEDF) and the 1990 Hundred-percent Edited Detail File (HEDF). The SEDF contains the records for the sample of respondents who

---

[29] For example, 2,062 of the 9,935 tracts in California and Florida have suppressed tables for characteristics of blacks even though these tracts have black residents.
[30] There are two enterprise zones in Florida that were not tracted in 1980, Chipley and Starke.

completed a long-form questionnaire while the HEDF has the information from responses to the short-form questionnaire.[32] I aggregate the data to 1980 tract by using 1990 block and the 1992 TIGER/Line geographic data files.[33] In addition, individual-level data from the SEDF was used to estimate the employment probability models. This dataset has the large samples necessary to precisely estimate tract fixed effects.

1.6.3 Business Climate Data

To control for changes in business climate prior to zone designation, I made special tabulations from the Standard Statistical Establishment List (SSEL) from 1982 to 1986. The SSEL is a database of establishments with employees maintained by the Census Bureau and serves as the sampling frame for most establishment surveys.[34] It has limited information, such as total employment and industry codes, for most active establishments. The finest level of geography available in the SSEL in this period is the Census place code used in the 1982 Economic Census.[35] I calculate the number of employees and establishments in each place for each year and link these data to the demographic data by Census place.

---

[31] There were 11 1980 tracts that linked to multiple 1970 tracts, in which case the mean of the 1970 data weighted by 1970 population was used for the 1970 characteristics for the 1980 tract.

[32] As is the case in the STF data, the HEDF was used to calculate 1990 population; percent white, black and Hispanic; vacancy rates; and the distribution of ages. The SEDF was used to calculate the remaining 1990 tract-level variables.

[33] The tract that was the largest portion of the block was used for the 2 percent of blocks that were part of multiple tracts. On average, 78.5 percent of the area of these blocks was in the tract assigned.

[34] Some industries that were out of scope for the Economic Censuses of 1982 had either no data or data at the firm-level. These industries are: government, military, agriculture, FIRE, and colleges and universities.

[35] For incorporated places with 1980 population above 5,000 and unincorporated places with 1980 population above 10,000, the 1982 Economic Census place code is the same as the place code used in the 1980 Census of Population and Housing. Smaller places are not given place codes. Rather than drop places that were not given unique codes in the 1982 Economic Census, I use the employment trends for the non-coded component of each county.

1.7 Conclusion

This chapter lays the groundwork for the rest of the dissertation. One of the many goals of enterprise zone programs is to serve as a substitute for traditional social insurance and job training programs. The literature on enterprise zones has primarily focused on the effectiveness of enterprise zones in attracting business activity. The most robust finding is that enterprise zones had negligible impact on the growth rates of the number of businesses or jobs located in the zones. The three papers that have studied the effect of enterprise zones on resident employment have contradictory findings: one found large positive effects (Papke 1994), one found small positive effects (Papke 1993), and the most recent found mixed but largely negative effects (Greenbaum and Engberg 2000). My work provides definitive evidence by focusing on states where the zone programs were targeted to affect the employment of disadvantaged people, using better measures of where zones were located, and a methodology that is suited to evaluating the effect of a geographically determined policy on an individually determined outcome. A caveat of my work is that I focus on two states and evaluate the impact of the programs from 1987 to 1990. The effects of different enterprise zone programs, or the same programs over a different time period, may be different.

This dissertation adapts program evaluation techniques to evaluate the effects of these programs on the employment of zone residents. The technique can be applied to other situations where selection for treatment occurs at an aggregated level while the outcome of interest is determined at the individual-level. The first step is to generate a tract-level measure of employment outcomes that controls for the characteristics of individuals living in the tract. Chapter 2 details this work. The next step is to find the difference in outcomes between enterprise zone tracts and observationally equivalent non-zone tracts. This gives the estimated effect of the enterprise zones conditional on the characteristics of the population in the zone and non-zone areas. These estimates are presented in Chapter 3.

Figures for Chapter One

Figures 1.1 to 1.1d: Comparison of Different Measures of Enterprise Zone Location

Legend

=Enterprise zone boundary

= Zip Code containing EZ

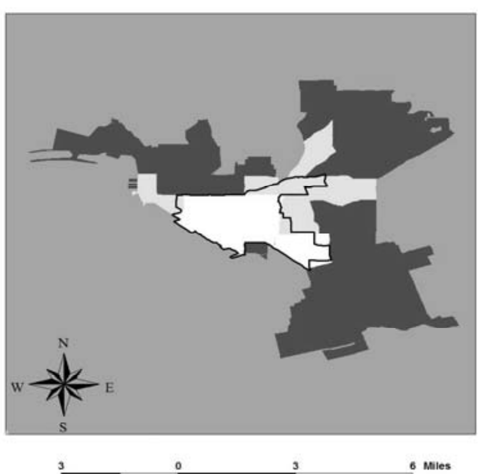
=Tract containing EZ with no more than 25% of population in EZ

=Tract containing EZ with > 25% of population in EZ

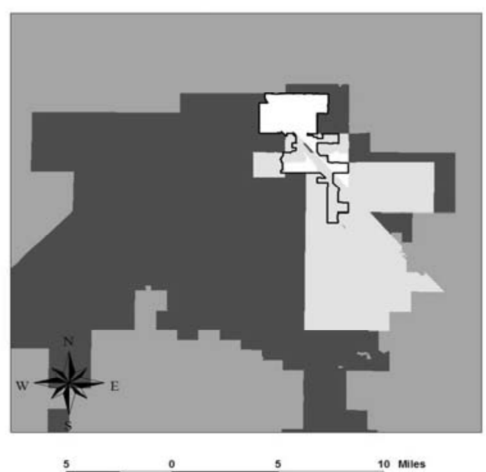
(1.1a) Ft. Lauderdale, FL EZ

(1.1b) Jacksonville, FL EZ

(1.1c) San Diego, CA EZ



(1.1d) Fresno, CA EZ



Tables for Chapter One

Table 1.1: Geographic Information for California Enterprise Zones

				No share restrictions		With at least 25% of population in EZ			With at least 75% of population in EZ		
Enterprise Zone	Square miles	% of block defined area in EZ	Block defined 90 Population	Number of tracts	% of tract population in EZ	Number of tracts	% of tract population in EZ	% of EZ population in tract	Number of tracts	% of tract population in EZ	% of EZ population in tract
California											
Agua Mansa	17.38	86.43	28,717	18	35.62	12	60.61	96.17	6	86.95	35.73
Bakersfield	6.49	85.75	7,350	14	16.36	6	49.37	68.63	2	100.00	1.71
Calexico*	0.90	80.06	2,381	4	26.07	2	71.52	52.83	1	100.00	25.33
Central City	2.88	98.95	45,252	14	73.04	13	78.95	96.72	8	100.00	72.41
Eureka*	5.64	94.54	15,747	6	61.44	4	82.48	92.89	3	99.53	84.81
Fresno	11.44	98.19	38,200	17	68.01	10	99.79	95.07	9	100.00	94.66
North Gate	2.65	73.45	6,702	6	11.70	1	31.42	53.73			
Pacoima	5.69	95.49	46,874	9	71.19	7	87.66	97.94	5	100.00	78.04
Porterville	0.46	99.98	68	1	3.00						
San Diego	4.22	96.04	42,177	13	62.97	8	97.42	90.21	8	97.42	90.21
San Jose	9.61	95.55	88,334	25	70.91	19	90.80	95.59	16	97.94	85.72
Watts	8.23	99.83	103,301	36	62.79	28	89.48	98.05	24	99.06	87.27
Yuba*	16.84	75.68	12,745	9	41.83	6	53.44	94.68	2	98.11	11.83
Average:	**7.11**	**90.76**	**33,681**	**13**	**46.53**	**10**	**74.41**	**86.04**	**8**	**98.09**	**60.70**

Note: * indicates that none of the areas within the zone were assigned tracts for the 1970

Table 1.2: Geographic Information for Florida Enterprise Zones

				No restrictions		With at least 25% of population in EZ			With at least 75% of population in EZ		
Enterprise Zone	Square miles	% of block defined area in EZ	Block defined 90 Population	Number of tracts	% of tract population in EZ	Number of tracts	% of tract population in EZ	% of EZ population in tract	Number of tracts	% of tract population in EZ	% of EZ population in tract
Florida											
Bradenton	1.14	100.00	3,568	4	25.12	2	62.31	57.54	1	87.63	28.59
Clearwater	0.73	99.78	2,869	3	25.65	1	100.00	82.61	1	100.00	82.61
Cocoa	1.07	100.00	4,369	2	88.95	2	88.95	100.00	2	88.95	100.00
Daytona Beach	0.98	89.04	3,050	2	100.00	2	100.00	100.00	2	100.00	100.00
Del Ray Beach	1.03	100.00	3,411	4	17.83	3	32.10	99.79			
Ft Lauderdale	2.48	100.00	12,495	11	27.55	6	58.41	80.68	2	100.00	32.77
Ft Meyers/Franklin Pk*	1.30	100.00	3,981	5	27.78	3	37.22	93.42	1	96.55	1.41
Ft Pierce*	0.56	100.00	2,437	3	18.52	2	33.84	99.79			
Gainesville	0.61	100.00	2,421	1	54.70	1	54.70	100.00			
Haines City*	2.00	99.98	4,047	3	33.34	1	49.97	86.14			
Homestead	1.17	100.00	6,292	4	27.10	2	69.42	98.51	1	97.53	44.52
Immokalee*	6.75	12.06	7,917	4	66.44	4	66.44	100.00	1	100.00	48.95
Jacksonville	22.62	99.10	71,186	24	78.20	20	93.27	99.05	17	100.00	91.20
Leesburg	1.07	91.74	1,886	2	32.52	1	45.48	79.69			
Miami Beach	0.64	100.00	16,395	2	98.91	2	98.91	100.00	2	98.91	100.00
Milton	1.55	99.99	1,753	2	26.76	1	36.83	99.83			
N Central Dade	26.53	99.70	167,417	63	57.96	48	86.81	92.94	43	97.00	85.66
Orlando	2.59	98.93	10,366	11	28.36	3	90.54	71.50	3	90.54	71.50
Palmetto	0.36	99.98	1,880	2	28.38	1	71.99	95.43			
Panama City*	0.70	100.00	1,384	3	22.97	1	94.74	36.42	1	94.74	36.42
South Bay	0.90	100.00	748	2	19.62	1	33.14	23.13			
St Petersburg	4.11	99.20	19,581	10	77.13	9	87.02	99.91	7	99.74	82.63
Tallahassee	2.22	100.00	8,401	5	94.94	5	94.94	100.00	4	99.06	98.27
Tampa	2.29	99.79	10,806	5	100.00	5	100.00	100.00	5	100.00	100.00
Tarpon Springs	0.14	100.00	467	1	5.03						
Umatilla	2.67	99.98	2,350	2	31.24	1	99.07	99.70	1	99.07	99.70
W Palm Beach	0.74	100.00	4,100	4	50.46	3	69.83	96.46	2	83.92	57.17
Average:	**3.29**	**95.90**	**13,910**	**7**	**46.87**	**5**	**71.38**	**88.17**	**5**	**96.31**	**70.08**

Note: * indicates that none of the areas within the zone were assigned tracts for the 1970 Census. Two zones from Florida, Chipley and Starke, were not tracted in 1980 and so are not included in this study. The Ft. Meyers and Franlin Park zones were combined because they are not geographically distinct.

CHAPTER TWO

Neighborhood and Employment: Separating Who You Are from Where You Live

2.1 Introduction

Since Kain's seminal paper (1968) introduced the spatial mismatch hypothesis, much has been made of how employment rates vary across neighborhoods.[36] Researchers have also focused on how people's employment is influenced by the characteristics of their neighbors.[37] One reason that addressing these questions is difficult is that neighborhoods vary by the characteristics of residents as well as by any neighborhood effects, such as spatial access, peer interactions, or other neighborhood-determined factors. In this chapter, I estimate the component of employment probability that is directly correlated with residential neighborhood. This provides relative measures of the influence of economic characteristics that are related to place of residence on employment.

Most empirical studies of the spatial mismatch hypothesis have either used neighborhood-level data to study how average employment rates vary by neighborhood (e.g. Ellwood 1986) or individual-level data to study the relationship

---

[36] See Holzer (1991) and Ihlanfeldt and Sjoquist (1998) for literature reviews.

between employment and neighborhood type (e.g. Raphael 1998a). Ihlanfeldt and Sjoquist (1998) note that, "Depending on which type of data is used, a different set of important independent variables gets excluded from estimated equations." Using neighborhood-level data, it is not possible to fully control for the characteristics of residents when studying how employment rates vary by neighborhood. This is especially problematic if there is within-neighborhood variation in characteristics related to individuals' employment probability, such as education, because neighborhood-level data will not capture that variation. Studies that estimate the effect of neighborhood type with individual-level data are also not fully satisfying. These studies typically combine neighborhoods into broad categories, such as central city versus suburban. These categories can contain heterogeneous neighborhoods. For example, the Washington, DC neighborhoods of Trinidad and Friendship Heights are both in a central city but are sharply different.[38]

The spatial mismatch literature focuses on the impact of neighborhood attributes (primarily job access) on the people living in the neighborhoods. Ideally, one would control for other determinants of employment, such as the human capital of people living in the neighborhood, when studying the impact of neighborhood characteristics. Papers that have done this are subject to criticism regarding how they specify neighborhood attributes. For example, Raphael's (1998b) choice of defining job growth rate within a 30 minute commute as a measure of job access has been criticized because it does not capture whether the jobs measured match the

---

[37] See Galsten and Killen (1995) for literature review.

skills of residents (Ihlanfeldt and Sjoquist 1998). Rather than assume a specific relationship between neighborhood attributes and individual employment, it would be good to control for neighborhood in a flexible, detailed manner. One way to make cross-neighborhood comparisons more fruitful is to first estimate the component of employment probability correlated with residential neighborhood when conditioning on the individuals in the neighborhood. This is the approach that I take.

My method for estimating the component of employment that is explained by neighborhood of residence is straightforward: I estimate employment probability models and estimate neighborhood fixed effects, either directly by including neighborhood fixed effects or indirectly by estimating neighborhood-level employment models and using the average residual from these areas. These neighborhood effects provide a measure of the component of resident's employment probability that is explained by the neighborhood. I call this neighborhood component the conditional employment probability because it is estimated conditional on the observable characteristics of residents. I then estimate how the conditional employment probability is correlated with neighborhood characteristics, such as percent of the population that is black or whether the neighborhood is in a central city. I do this for two different types of employment probabilities: the probability that an adult is employed given they have chosen to participate in the labor market and the probability that an adult is employed regardless of labor market

[38] At the time this dissertation is being written, Trinidad is a poor, predominantly black area while Friendship Heights is a wealthy, predominantly white neighborhood.

participation. I look at the employment of men and the employment of all prime age adults. I use a variety of specifications and find the results consistent. Because my ultimate motivation for estimating the conditional employment probabilities of these neighborhoods is to study the effect of enterprise zone policies on the employment probabilities of zone residents, I focus on two states: California and Florida. This also motivates my definition of neighborhood: a 1980 tract-place pair. The reasons for these choices are discussed in the first chapter of the dissertation.

I do not believe that the tract effects I estimate are strictly neighborhood effects that reflect only relatively fixed neighborhood characteristics like job accessibility. Instead, a number of different factors correlated with residential neighborhood are potentially embodied in the tract effects. For example, if people with unobservable propensities to be employed live near people with similar propensities, that would be reflected in the tract effects. Similarly, if there are peer effects or interaction effects between neighbors, that would be picked up in the tract effects. However, I do believe that the tract effects I estimate are closer to the underlying neighborhood effects on employment than the unconditional employment rates are.

This chapter is organized as follows. Section 2.2 discusses how I estimate the conditional employment probabilities and presents the results from the employment probability models. Because these tract effects are not a commonly used measure and are important inputs to the work in Chapter 3, I discuss the characteristics of these effects in detail. I discuss the distributions of the different estimates of conditional employment probability and how they are related in section 2.3. I end with conclusions in section 2.4.

2.2 Employment probability models

I use three different models to estimate the neighborhood effects: individual-level probit, individual-level ordinary least squares (OLS), and tract-level weighted least squares (WLS). In each of the individual-level employment probability models, the dependent variable is an indicator for whether or not an individual is employed. The independent variables included are a number of individual characteristics and a set of neighborhood fixed effects. These fixed effects provide an estimate of the component of employment explained by the tracts when controlling for the characteristics of individuals. In the WLS models, the dependent variable is the neighborhood employment rate and the independent variables are neighborhood level measures of resident characteristics.

As mentioned in the first chapter, the data I use to estimate the employment probability models is the 1990 Sample Edited Detail File (SEDF). The SEDF is the long form sample of households from the 1990 Census of Population and Housing and provides large samples, detailed geographic units, and data regarding individual's housing, labor market activities, education, and a wealth of demographic characteristics. Using the process discussed in Chapter 1, I focus first on California and Florida and convert 1990 Census Blocks to 1980 Census Tract-Place pairs (hereafter, tracts). In order to focus on urban areas similar to enterprise zones, some sample restrictions were placed on all stages. Tracts where less than 95 percent of the 1990 population lived in an urban area are dropped, which eliminated 15 percent of the tracts in the sample. Tracts with 1980 population below 100 are dropped to reduce problems with measurement error and missing data. This cut 662 tracts from the sample. In order to eliminate isolated urban areas that would not be appropriate controls for enterprise zones, the analysis was restricted to tracts located in a Metropolitan Statistical Area in 1990, which eliminated 295 tracts.[39] In addition to the restrictions related to the share of the population of the tract living in an urban area and the 1980 tract population, the sample for the individual-level employment probability model is restricted to adults aged 18 to 55 and not enrolled in school in order focus on prime-age employment.[40] I drop observations that have missing or imputed data for employment status, enrollment status, race, education, or age to reduce measurement error.[41] To have a basis with which to compare California and Florida to the whole United States, I also estimate individual-level employment probability models without fixed effects for men and women living in all urbanized areas in the United States, which I will call the urban US samples. The national samples have the same restrictions as the single-state samples in regard to age, school enrollment, and imputed data.

I look at employment probabilities for twelve different samples. The primary sample, which I call the pooled sample, includes both men and women and is

[39] The zones eliminated from the study by the MSA restriction were Calexico and Eureka in California and Leesburg and Umatilla in Florida.

[40] Note that if enterprise zones induced individual to leave school and search for a job this restriction may produce a bias in my results. Such behavior would be counter to the assumption that zones do not lead people to change their characteristics, which is introduced in Chapter 1.

[41] In the 1990 SEDF, non-response was dealt with by using the response of a similar person through an allocation process.

restricted to those who are in the labor force. The second sample consists of men and women together and includes those people who are out of the labor force; I call this the pooled with non-participants sample. Third, I consider sub-samples of men from the pooled sample (referred to as the men only sample) and, fourth, I consider men for the pooled with non-participants sample (called the men with non-participants sample). Throughout the analysis, I treat California, Florida, and - where applicable - the urban US separately. Three areas times four sample types brings the total number of samples to twelve.

In all of the employment probability models, I include controls for age, education, race, marital status, and immigration status. For the models with the pooled samples, I also include the number of related children in the household. Table 2.1 reports the sample means of all these variables broken out by gender. Table 2.2 contains similar means for the urban US sample. The means show that the state samples that I focus on have similar patterns as the urban US as a whole. On average, a larger fraction of men than women participate in the labor market. Comparing labor market participants and non-participants, participants have higher levels of education in California, Florida, and the urban US. Non-participating men are less likely to be married while non-participating women are more likely to be married. The average levels of employment and education are lower for black men than all men. Hispanic men in the sample are on average younger and less educated but have similar employment rates to all men. This table shows that the samples I use are very large, ranging from 174,884 to 1,239,023 observations. These large samples are critical to my estimation procedure because I need to have a sufficient number of observations per tract to precisely and consistently estimate the tract fixed effects.

The employment probability models estimate the conditional employment probability of each tract in the sample. It has become standard practice to use non-linear, maximum likelihood estimators to model discrete dependent variables, such as an employment indicator variable. I use one of the most common estimators of this class, the probit (Greene 1998). I have also estimated logit models with the same independent and dependent variables. The resulting estimates are extremely similar to the probit estimates; therefore I do not include them in this thesis. There is some fear that including fixed effects in non-linear models can lead to inconsistent estimates of model coefficients (e.g. Maddala 1987). Greene (2002) recently found that as the number of observations per fixed effect approaches 20, estimates from a probit model with 1,000 observations are consistent. The smallest sample I use to estimate individual-level employment probability models has 174,884 observations.[42] The median number of observations per fixed effect ranges across samples from 60 (Florida men in the labor market sample) to 205 (California pooled including non-participants sample). The mean number of observations per fixed effect ranges across samples from 77 to 223.

Because I am interested in the relationship between tract of residence and employment probability, I convert the coefficient estimates from the probit models into marginal effects. The marginal effects can be interpreted as the effect of a one-unit increase in the independent variable conditional on being at the sample average for all characteristics in the models. More formally, the probit models take the form:

$$y = F\left(\beta X + \alpha D + \varepsilon\right)$$

where y is a vector of employment indicators, $F(.)$ is the Normal cumulative density function, X is a matrix of individual-level observable characteristics, D is a matrix of neighborhood indicators, ε is a vector of error terms, and β and α are coefficient vectors. I calculate the marginal effect of continuous variable x_k as:

$$ME_k = \left\{ \frac{dF\left(\beta \overline{X} + \alpha \overline{D}\right)}{d\beta_k x_k} \right\} \beta_k = f\left(\beta \overline{X} + \alpha \overline{D}\right) \beta_k$$

where β_k is the coefficient for x_k, $\overline{X}$ and $\overline{D}$ are the sample means of X and D, and $f(.)$ is the partial density function of $F(.)$. The marginal effect of indicator variable x_l in X is calculated :

$$ME_l = F\left(\beta \overline{X} + \alpha \overline{D} - \beta_l \overline{x}_l + \beta_l\right) - F\left(\beta \overline{X} + \alpha \overline{D} - \beta_l \overline{x}_l\right)$$

where β_l is the coefficient for x_l and $\overline{x}_l$ is the sample mean of x_l. I use this method for calculating marginal effects of the individual-level characteristics for computational ease.

Unlike the marginal effects of the individual-level control variables, the tract fixed effects are central to the rest of this dissertation. Therefore, to calculate the marginal effects for the tract fixed effects I summarize across the distribution of

---

[42] While Greene (2002) finds that the coefficients of non-fixed effect variables can be consistently estimated in probit models with fixed effects, he makes no comment about whether the fixed effects

individual characteristics rather than simply calculating the marginal effect at only one point in the distribution. The marginal effect of tract fixed effect for tract *j* is:

$$ME_j = \frac{1}{N}\sum_{i=1}^{N}\left[F(\beta X_i + \alpha_j) - F(\beta X_i + \alpha_c)\right]$$

where *i* indexes over individuals, *N* is the number of individuals in the sample, α_j is the tract fixed effect for tract *j*, and α_c is a comparison tract fixed effect. I use four alternatives for α_c: the mean α, the median α, the 25th percentile of the distribution of α, and the 75th percentile of the distribution of α.[43] Each of these methods for calculating marginal effects is presented in more detail in Greene (1998).

The other estimator I use to directly estimate tracts' conditional employment probabilities is an ordinary least squares (OLS) model with tract fixed effects. For the OLS models, the coefficients are reported because they have the same scale as the marginal effects. The disadvantage of using the OLS estimator is that there is nothing to restrict the predicted employment probability for a person to be in the range of zero and one. This means that the error term does not fit the assumptions of the OLS estimator. Still, the point estimates from an OLS model should be very close to the marginal effects estimates from a comparable probit model.

---

themselves are consistently estimated.

[43] Greene (1998) shows that the results of these two alternate methods for calculating marginal effects are generally very close. I have calculated the marginal effects of some of the individual-level variables using both methods and the results are very similar. For example, the marginal effect associated with "Black" in the first column of Table2.3 is -0.060 while the marginal effect calculated by summing the effect over individuals is –0.064.

The last way I estimate tracts' conditional employment probabilities is using a weighted least squares (WLS) approach called the minimum chi-square method by Maddala (1997). This approach regresses a grouped discrete outcome, such as neighborhood employment rate, on a set of grouped independent variables, and weights each group by the square root of the inverse of the variance of the outcome for that group. More formally, I estimate the following model:

$$w_t e_t = w_t \beta x_t + u_t ,$$

where e_t is the tract employment rate, x_t is a vector of tract-level independent variables, w_t is the weight for tract *t*, β is the set of coefficients to be estimated, u_t is the error term. With n_t as the number of people in the sample from tract *t,* the formula for the weight is:

$$w_t = \sqrt{\frac{n_t}{e_t(1-e_t)}} .$$[44]

I use the estimated β to estimate the conditional employment probability of tract *t* as the estimated residual for tract *t*:

$$\alpha_{WLS,t} = \hat{\varepsilon}_t = e_t - \hat{\beta} x_t .$$

Maddala notes, and my experimentation with various samples confirms, that this approach only produces meaningful results if n_t is large for all *t*. For this reason, I restrict the WLS models to tracts with at least 25 observations in a specific sample.

[44] This presentation of the model closely follows Maddala (1997).

Also, the weight is not defined when $e_t = 1$ or $e_t = 0$, so I drop tracts where all or no people in the sample are employed from the WLS samples.

The control variables included in the models are age and (age squared)*0.01 and indicators for race, education level, immigrant status, and marital status. For the models with men and women pooled, I add a gender indicator and the number of related children in household. For the probit and OLS models, I interact all controls except the fixed effects with the gender indicator so the coefficients are estimated separately for men and women. To keep the WLS estimates comparable to the probit and OLS estimates, I separately estimate models for men and women. If I were to include measures of men's and women's characteristics in the same tract-level employment regression, the coefficients would have a very different interpretation than those from the probit and OLS models. Consider the example of the coefficients related to marriage for women. In the individual-level OLS models, this coefficient on marriage interacted with an indicator for being female can be interpreted as the decrease in a woman's employment probability that is correlated with being married. If I were to model tract-level employment rate for men and women pooled using the WLS approach, the coefficient for the share of women who are married would reflect the effect of proportion of women who are married on both men's and women's employment. By separately estimating the tract-level employment regressions for women, the coefficient I estimate can be interpreted as the effect of this proportion on the employment rate of women. This is similar to the interpretation of the corresponding coefficient from an individual-level model.

Therefore, separately estimating the tract-level regressions by gender facilitates the comparison of the results from the different types of models. The conditional employment probability for the pooled samples is then:

$$\alpha_{WLS,t} = f_t \alpha_{WLS,t}^{f} + \left(1 - f_t\right) \alpha_{WLS,t}^{m} ,$$

where f_t is the proportion of the sample from tract t that is female and $\alpha_{WLS,t}^{s}$ is the tract residual from the male (s=m) or female (s=f) sample model.[45]

These models are ultimately designed to estimate for California and Florida the component of employment probability that is related to tract of residence after controlling for individual traits, not to give insight into the determinants of employment. That said, it is helpful to know how similar the models I estimate are to what I would get with samples and models that are more typical. To address this, I estimate employment probability models without tract fixed effects for the urban US, California, and Florida samples. The tables with the estimates from the probit and OLS models without tract fixed effects for the pooled samples are Table 2.2 (urban US), Table 2.3 (California), and Table 2.4 (Florida). I provide these results to show that the estimates from the models that include tract fixed effects are not radically different from those from models that do not included tract fixed effects. Because the WLS estimates do not include tract fixed effects, I do not include them in these tables. The comparisons across sample are very similar for each of the estimators. Therefore, I will only discuss the probit results, although the tables also

[45] In earlier work, I also interacted the controls in the men only models with indicators for race. The resulting conditional employment probability estimates were nearly the same as the estimates from the simpler models, so I do not report them here.

include OLS results. I find that the estimates from the employment probability models from the state samples are similar to those from the urban US sample. For example, the marginal effect of being a married man is 0.032 in the Florida participant sample and 0.036 in the California and urban US participant samples. For all the samples including non -participants, the marginal effects for different levels of education are monotonically increasing: more education is associated with higher probability of being in the labor force and employed. The most striking difference is in the estimates in the models including non-participants for the marginal effect of being a Hispanic woman. This estimated marginal effect is 0.018 in California, -0.025 in Florida, and -0.016 in the US. The variation in these estimates is mostly likely due to differences in characteristics related to employment -- such as human capital, family structure, or country of origin -- between Hispanic women in California, Florida, and the US.

The results from the individual-level models including tract fixed effects and the WLS models for the pooled samples from California are in Table 2.5. For the sample of people who are labor market participants, the estimated coefficients on individual characteristics are similar to those without tract fixed effects. For example, in the probit model, the marginal effect associated with being a black male is -0.041. The marginal effect of education is monotonically increasing with two exceptions: men with less than a 9th grade education and women with more than a BA. The probit and OLS estimates are similar and differ by no more than 0.04 and generally by less than 0.01. For example, the marginal effects associated with being a married man are 0.032 and 0.043 for the probit and OLS models respectively. There is more difference between the coefficients from the tract-level WLS models and the individual-level models than between the individual-level models. Generally, the signs are the same but the magnitudes of the WLS coefficients are greater. For example, the OLS coefficient for male participants for having a BA is 0.028 while the WLS coefficient is 0.098. Some WLS coefficients have the opposite sign of the estimates from the individual-level models, such as the number of own children for male participants. This may be partially due to changes in sample because the WLS samples do not include tracts with fewer than 25 people in the sample and tracts where the employment rate is one are dropped. It may also be due to the fact that these models are estimated with tract-level data. Consider again the difference in the interpretation of the coefficients related with being married estimated from individual-level models versus tract-level models. In the individual-level models, the coefficient related with being married is the effect of an individual being married on own employment probability. In the tract-level models, the coefficient on proportion of people who are married reflects the influence of being married on individual employment and any other influence that tract-level marriage rates have on tract-level employment rates. Therefore, it is not surprising that the WLS coefficients are different.

These patterns also hold for the models that include men and women who are not labor force participants. The primary difference is that the magnitudes of most coefficients are higher when the non-participants are included. The magnitude of the

OLS coefficients are typically higher than those of the probit model for women and lower for men. The WLS estimates still typically have the same sign and are of a larger magnitude. Table 2.6 has the corresponding estimates for Florida. The patterns in coefficients for Florida are similar enough to those for California that there is no need for a separate discussion. The hypothesis that the tract fixed effects are equal is rejected for all the individual level models and are not applicable for the tract-level models.

While the estimates from the models with tract fixed effects are similar to the estimates from the models without tract fixed effects, it is interesting to note how they differ. With the exception of the estimates for the Hispanic dummy variable, all of the estimates for dummy variables are closer to zero in the models with fixed effects than those in the models without fixed effects. In absolute terms, the marginal effect of being black is the most sensitive to the inclusion of the tract fixed effects. In the probit model for people in labor force in California, the marginal effect of being a black male is -6.0 percentage points without tract fixed effects and -4.1 percentage points with them. The changes in the samples including non-participants are even starker. The spatial mismatch hypothesis predicts that controlling for access to employment should reduce the difference between the employment probabilities of blacks and whites, so the pattern of my estimates is consistent with this hypothesis. Access to employment is typically defined to include sufficient transportation infrastructure to reach jobs and any neighborhood-related sources of information about jobs, such as informal relationship networks.

The tract fixed effects control not only for job access, but for anything related to employment shared by people living in the same tract, observable or unobservable. Therefore, the changes in the marginal effects of being black from including the tract fixed effects are only suggestive, not evidentiary. The fact that nearly all of the dummy variable marginal effects are attenuated when tract effects are included adds to my caution about interpreting the change in the black marginal effect as supporting the spatial mismatch hypothesis.

I also estimate models for men separately to see whether the tract effect estimates are sensitive to how gender is treated and so I can estimate the impact of enterprise zones on men's employment in chapter 3. The estimates for these models are reported in Tables 2.7 (individual-level models without tract fixed effects for all samples) and 2.8 (models to estimate conditional employment probability for tracts in California and Florida). As in the models I estimate with both men and women, the estimated effects are very similar within sample across the individual-level estimators. The magnitudes of the coefficients are generally highest for the WLS models and some WLS coefficients have different signs than the corresponding OLS and probit estimates. For the samples with only male labor force participants, the estimated marginal effects are very close to the estimates for males from the models with men and women pooled. This is least surprising for the WLS models, where the only difference is that men living in tracts with less than 25 men and at least 25 men and women were included in the pooled model and excluded from the men only models. For the individual-level models, the estimates are less similar between the pooled and men-only samples including non-participants, especially when the models include tract fixed effects. For example, the estimated marginal effect of being a black male in the probit model with tract effects for the pooled California sample is -11.2 percentage points and the corresponding estimate from the male-only sample is -6.6 percentage points. Even though the marginal effects estimates from the male-only non-participant models are of a different magnitude than those from the pooled non-participant models, the patterns discussed above with respect to the pooled samples largely hold. The marginal effects have the expected magnitudes and signs and the marginal effects of the dummy variables are attenuated when the tract effects are added to the models.

I find that my employment probability models are sensible. They are consistent with both the human capital theory and prior empirical literature. One important caveat to note is that I find my models are less effective at explaining employment probability when including people outside of the labor force. For OLS models with tract effects for men and women in California, the R-squared is 0.94 when restricted to people in the labor force and 0.81 when including non-participants.[46] This suggests that there is more unexplained variation in the samples that include non-participants to be incorporated into the tract fixed effects. This may also be due to endogenous selection of neighborhood: people who have opted out of the labor market may choose to live in certain neighborhoods because they place less value on access to employment or they have lower income than others. People who are unemployed but still looking for work may value neighborhood attributes more similarly to the employed, so neighborhood choice may be more similar for people in the labor force than those out of it.

2.3 Tract effects

Each of the models described in the prior section generates a different set of tract effect estimates. How similar are these estimates to one another? How do they compare to the observed employment rates? For simplicity, I will refer to all the estimates of conditional employment probability as tract effects, even though the WLS estimates are actually estimated grouped residuals.

From here forward, I will be discussing tract-level data, estimates, and models. One problem with using the tract-level figures is that tracts have different numbers of observations used to estimate tract-level characteristics. For tracts with very small numbers of observations, it is possible that sampling error leads to bad estimates of tract characteristics. To focus on tracts where I have reasonable measures of tract characteristics, I restrict the sample to those tracts where at least 25 people were included in the employment probability models.[47]

The mean and standard deviation of each of the different sets of tract effect estimates is in Table 2.9. While the scale of the effects is different, the variances are very similar. The standard deviations of the tract effects are at most 82 percent as large as the standard deviations for the corresponding employment rate. For example, the WLS tract effect for the pooled sample in California has a standard

---

[46] For Florida, the corresponding numbers are 0.95 and 0.81.

[47] When comparing estimates from different samples the rule is applied to the smallest sample. For example, if comparing estimates from the men alone sample to the men and women pooled sample, tracts with fewer than 25 men were dropped.

deviation of 0.032, which is 64 percent as large as the standard deviation of the corresponding employment rate, 0.050. This means that much of the variation in employment across tracts is explained by resident characteristics. The tract effects estimated with the samples that include non-participants have higher variance than those estimated with labor for participants, but still have standard deviations 25 percent smaller than those of the corresponding employment rates. The means and standard deviations for the estimated effects for estimates from the pooled samples are similar for the same model and sample restrictions. Excluding women when estimating the tract effect does not markedly change the mean of the tract effects.

Another way to see how similar the tract effects are to one another and how they compare to the employment rate is to look at the degree of correlation between the different sets of estimates. I use four methods to calculate the probit marginal effects that vary by what the effect is relative to: the mean, median, 25th percentile, or the 75th percentile. The four alternate probit marginal effects are perfectly correlated. Therefore, I focus exclusively on the probit tract marginal effects calculated relative to the mean tract effect, but the results are applicable to the other probit marginal effects as well. The correlation coefficients of tract effects across estimators for the pooled labor force participant samples are in Table 2.10. The tract effects are highly correlated across estimators. The least correlated sets of estimates are the probit and WLS tract effects, with a correlation coefficient of 0.912 in California and 0.947 in Florida. Each of the sets of tract estimates is more correlated with the other tract effects than the employment rate, for which the correlation coefficients range from 0.75 for the WLS effects in California to 0.925 for the OLS effects in California. The correlation coefficients have similar patterns for other samples and I do not report them.

The correlations above were for the pooled samples. Table 2.11a has correlation coefficients that show the correlation between effects estimated with the pooled and male participant samples. The most interesting correlations are on the diagonal, which show that the correlation across sample but within estimator is high, ranging 0.766 for the WLS effect in Florida to 0.8702 for the OLS effect in California. The correlation coefficients comparing tract effects estimated with the pooled participant and non-participant samples are in Table 2.11b. The fact that the correlation coefficients are lower than the other ones I have discussed shows that the tract effects are most sensitive to whether or not non participants are included in the model. This is not surprising because including non-participants changes the outcome being modeled in the employment probability models that generate the tract effects. When non-participants are add to the sample, the outcome being modeled changes from the probability that an individual in the labor force is employed to the joint probability that someone chooses to participate in the labor market and is employed.

The remainder of the results that I will discuss in this section is presented in the form of graphs. These graphs are presented in sets of four, with separate graphs for the pooled and men only samples by state. The graphs for California are on the left and the graphs for Florida are on the right. The top rows are for the pooled samples and the bottom rows are for the men only samples. For the graphs and regressions, I restrict the sample to tracts that are not in the first or 99th percentile of the estimated tract effects from any model. This restriction makes the graphs easier to grasp because I drop outliers that would force the range of the graphs to be very large, but it does not have a qualitative effect on the results.[48]

The similarity across estimators can be seen in the kernel density graphs of the tract effect estimates in Figures 2.1a to 2.1d, which has graphs of the densities of the effects estimated when excluding people out of the labor force. For this graph, the tract effects and the employment rate have been demeaned to eliminate the difference in scale. These densities show that the distribution of the tract effects is similar across estimators, especially for Florida. They also show that the distributions of the tract effects is more compressed then that of the employment rate. The densities of the effects estimated when including non-participants are in Figures 2.2a to 2.2d. These graphs show the same patterns: the distributions of the different tract effect estimates are similar and more compressed than the distribution of the employment rate. All of the distributions are single peaked and skewed to the right.

Even though the tract effects are highly correlated within sample, they could be poorly correlated in some parts of the distribution. In Chapter 3, I show that enterprise zone tracts and the matching non-zone tracts have lower than average tract fixed effects. If the tract effects are poorly correlated in the lower tail, it would

[48]As in the restrictions on number of observations per tract, the outlier rules were applied to the smallest sample when comparing effects from different samples.

suggest that the results in Chapter 3 would be sensitive to the tract effect estimator used. This is one motivation for using graphs to compare the various tract effect estimates. The other motivation is that these kinds of effects are not often estimated, so their distribution and robustness to different estimators are not known. Graphing the effects is an intuitive way to look at the properties of the tract effects.

The scatter plots in Figures 2.3a through 2.6d compare the WLS effects to the other tract effects. Figures 2.3a to 2.3d compare the WLS effects to the probit effects for the participant samples. These figures show that the WLS and probit effects are least similar to one another in the upper tail of the distribution. This is also true for the non-participant samples (Figures 2.4a to 2.4d), but the correlation is lower throughout the distribution for these samples than in the participant samples. Figures 2.5a to 2.5d show that the correlation between the WLS and OLS tract effects are equally high throughout the distribution for the participants sample. For the non-participants sample, the relationship between the two sets of estimates is strongest in the upper half of the distribution.

The tract effects I estimate are fairly robust to the estimator used. Within sample, the correlation across sets of estimates is high. This is true even for the WLS estimates, which are the product of tract-level regressions (while the other estimates come from individual-level regressions). The tract effects are most sensitive to whether one is modeling the probability that people in the labor force are employed or the joint probability that people participate in the labor market and are employed.

2.4 Conclusions

This chapter demonstrates the feasibility and importance of controlling for individual characteristics when comparing neighborhood-level employment outcomes. I estimate the component of employment probability explained by neighborhood by estimating individual-level employment probability models with tract fixed effects. I find that the results from those models are sensible and robust. The estimation of the component of employment probability explained by residential tract is robust to which control variables are included and whether the sample includes men and women or only men. These estimates are not sensitive to whether the tract effects are estimated with individual-level OLS models, individual-level probit models, or tract-level WLS models. They are sensitive to whether the probability modeled is the probability that an individual in the labor force is employed or the joint probability that someone is in the labor force and employed.

I find that a quarter of the across-tract variation in employment rates can be explained by the characteristics of tract residents. This shows that it is important to condition on the characteristics of neighborhoods' populations when comparing employment outcomes by neighborhood. In the next chapter, I use this insight and the tract employment effects estimated above to test whether the state designated enterprise zones of California and Florida impacted the employment of zone residents.

Figures for Chapter Two

Figure 2.1a to 2.1d: Densities of Tract Effect Estimates

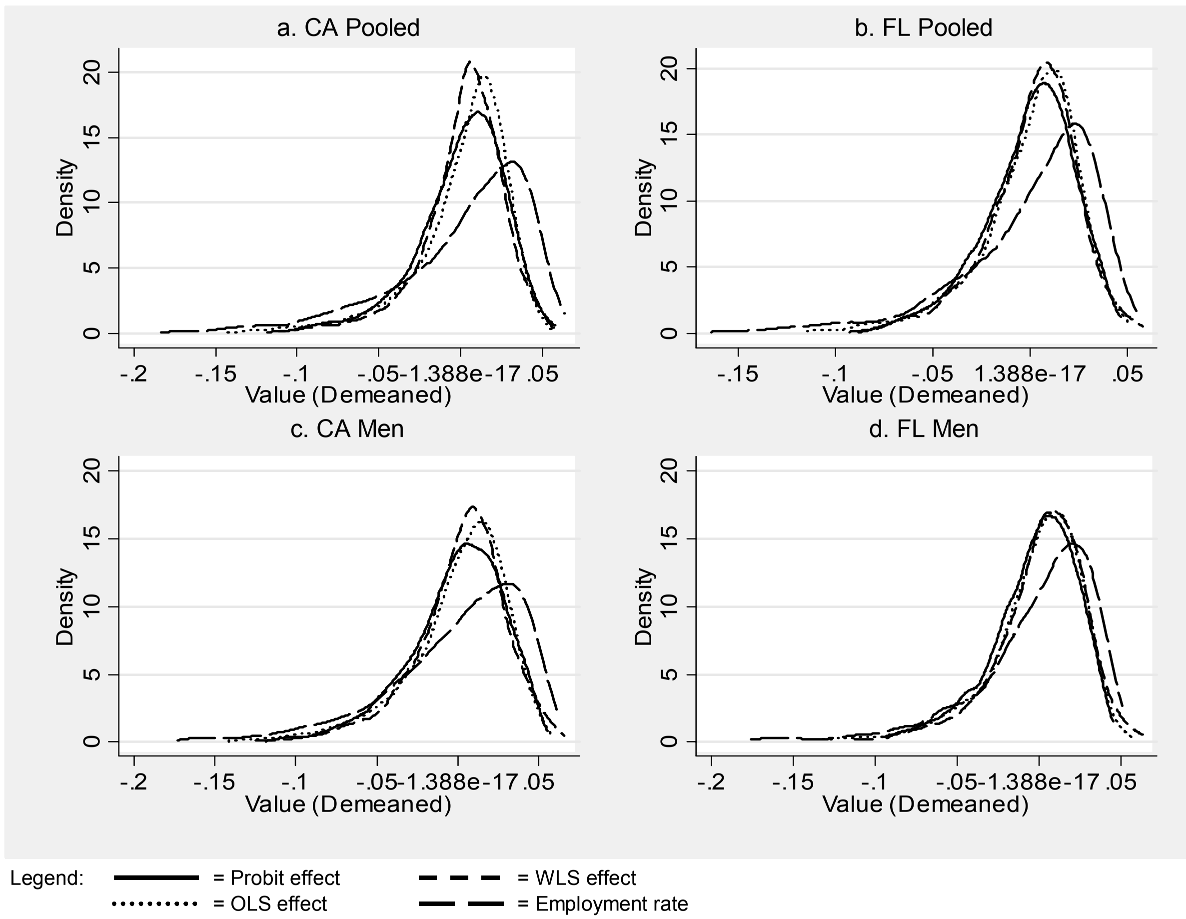


Figure 2.2a to 2.2d: Densities of Tract Effect Estimates for Samples with Non-participants

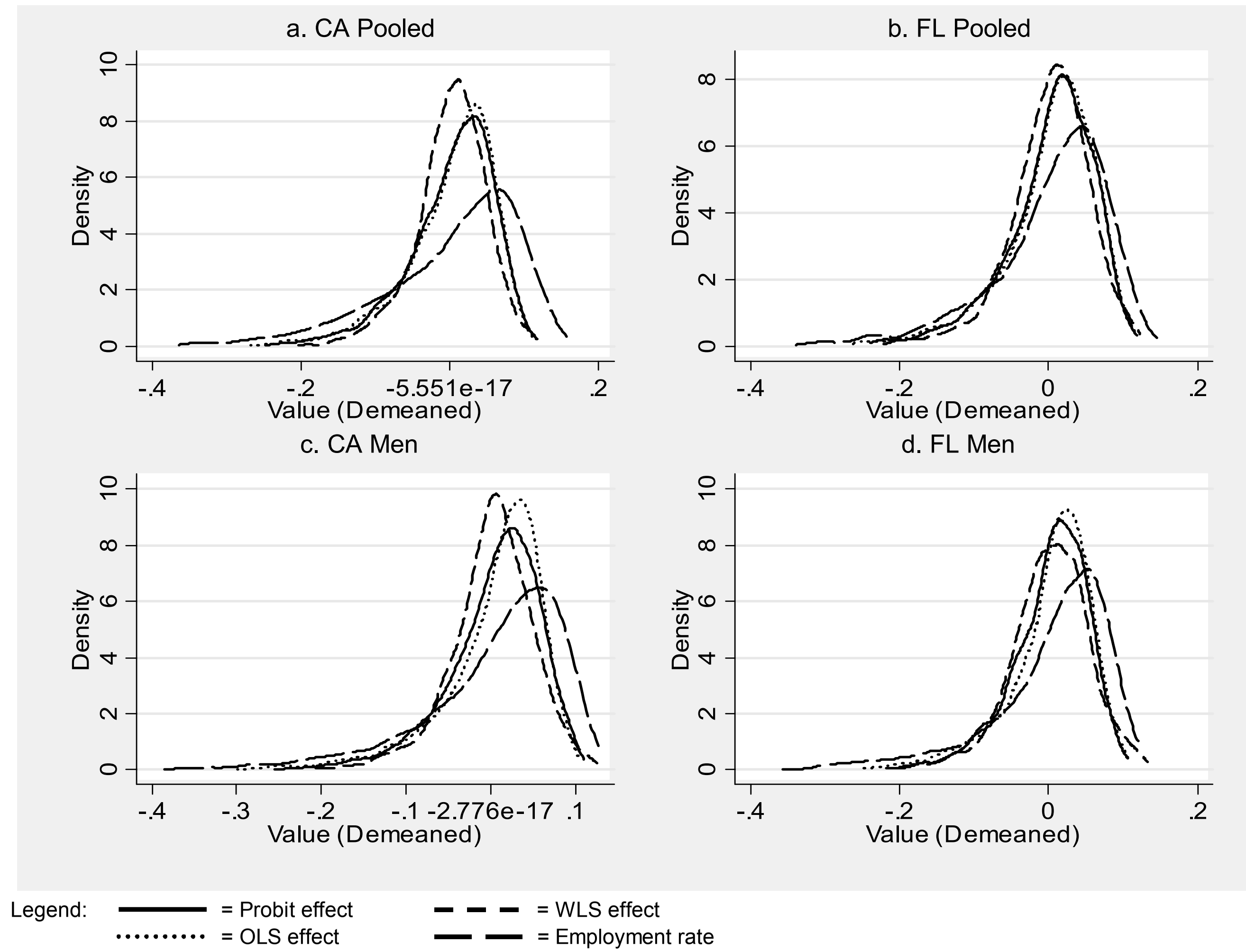
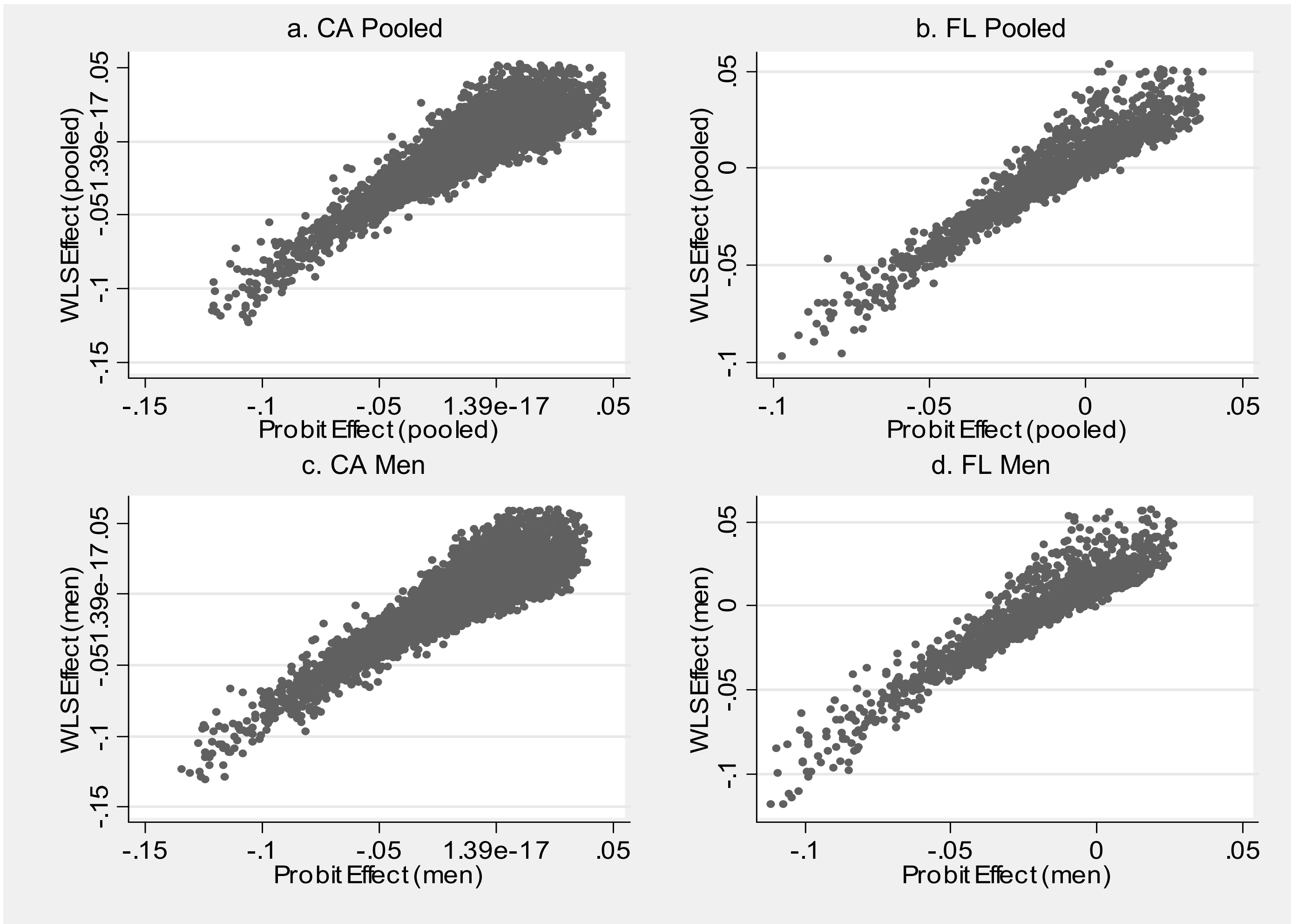
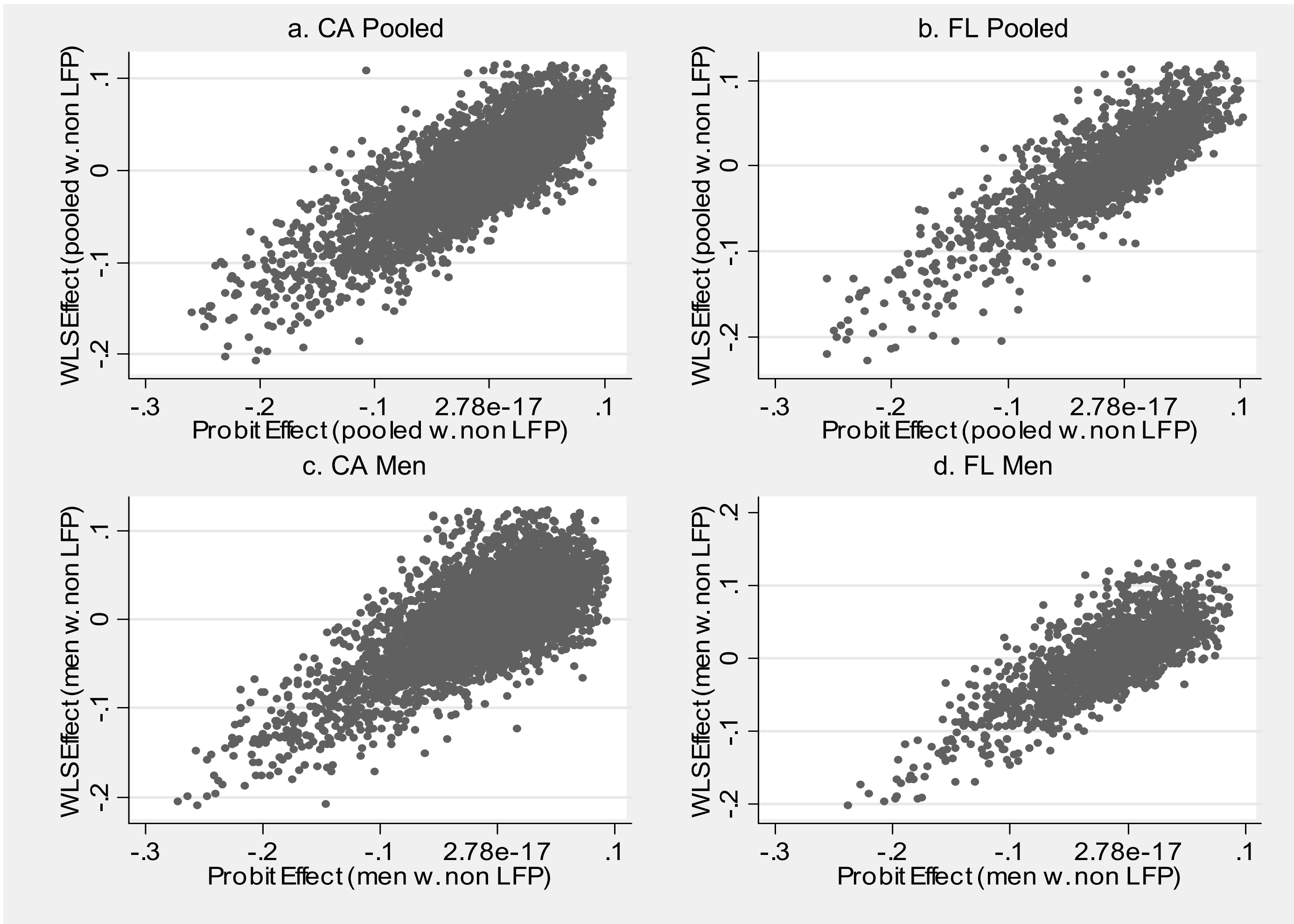


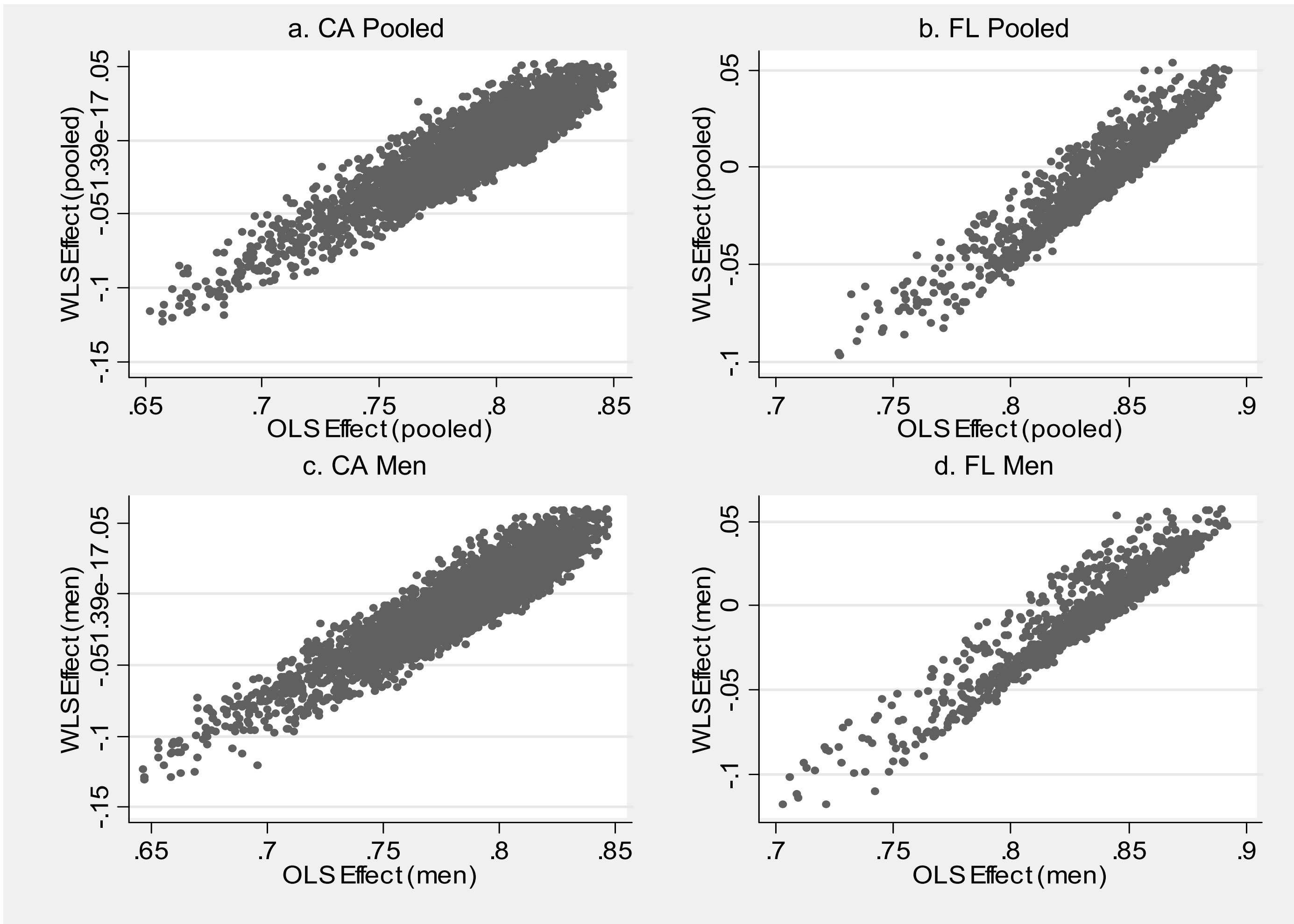
Figure 2.3a to 2.3d: Comparison of WLS and Probit Tract Effect Estimates for In Labor Force Samples



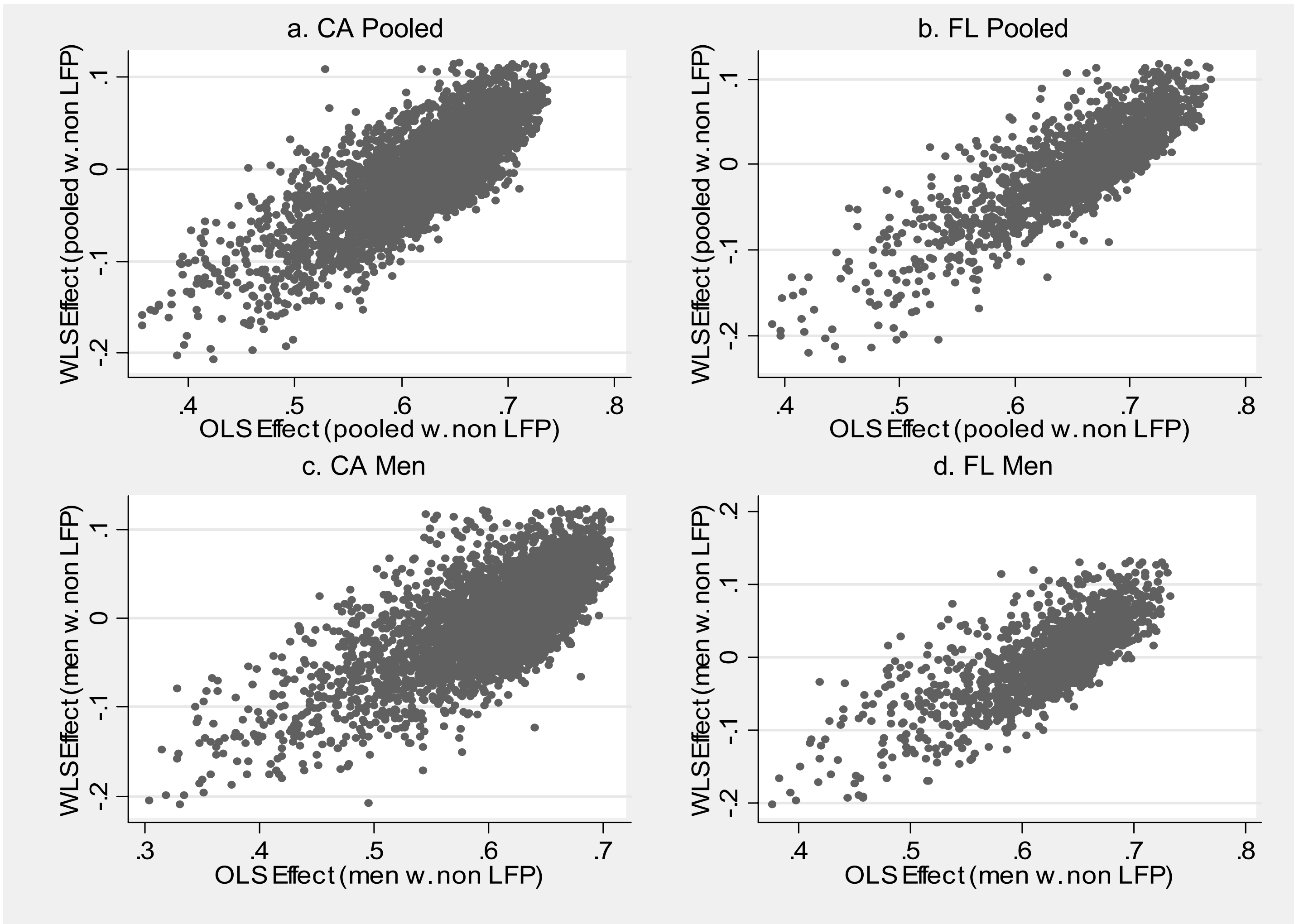
Figures 2.4a to 2.4b: Comparison of WLS and Probit Tract Effect Estimates for Samples with Non-participants



Figures 2.5a to 2.5b: Comparison of WLS and OLS Tract Effect Estimates for In Labor Force Samples



Figures 2.6a to 2.6d: Comparison of WLS and OLS Tract Effect Estimates for Samples with Non-participants



Tables for Chapter Two

Table 2.1: Means for Employment Probability Models

	California, labor force participants		California, include non-participants		Florida, labor force participants		Florida, include non-participants	
	Women	Men	Women	Men	Women	Men	Women	Men
Observations	455556	551865	637596	601427	168363	174884	222770	190321
Employed	0.936	0.935	0.669	0.858	0.942	0.946	0.712	0.869
Black	0.070	0.055	0.070	0.062	0.147	0.120	0.148	0.131
Hispanic	0.218	0.275	0.247	0.276	0.130	0.142	0.140	0.141
Age	35.916	35.400	35.901	35.409	36.122	35.912	36.458	36.020
	(9.497)	(9.577)	(9.688)	(9.703)	(9.653)	(9.629)	(9.867)	(9.783)
Age sq./100	13.801	13.449	13.827	13.479	13.979	13.824	14.266	13.932
	(7.025)	(7.049)	(7.186)	(7.142)	(7.171)	(7.148)	(7.383)	(7.271)
Education level:								
Less than grade 9	0.068	0.104	0.100	0.111	0.033	0.047	0.046	0.055
Grade 9-12	0.096	0.125	0.128	0.137	0.116	0.144	0.141	0.155
H.S. Degree	0.237	0.213	0.245	0.217	0.325	0.291	0.329	0.291
Some college	0.347	0.293	0.315	0.286	0.324	0.286	0.303	0.279
BA	0.180	0.170	0.155	0.161	0.143	0.158	0.131	0.151
More than BA	0.072	0.095	0.058	0.089	0.058	0.074	0.050	0.070
Married	0.559	0.589	0.603	0.571	0.565	0.601	0.606	0.583
Number of own children in HH	0.770	0.801	0.935	0.779	0.740	0.709	0.831	0.678
	(1.110)	(1.243)	(1.252)	(1.259)	(1.038)	(1.105)	(1.115)	(1.103)
Immigrant	0.010	0.010	0.009	0.010	0.009	0.010	0.008	0.010
Percent of pooled sample	45.200	54.800	51.500	48.500	49.100	50.900	53.900	46.100

Where applicable, standard deviations are in parentheses.

Table 2.2: Employment Probability without Tract Effects (United States)

Results from employment probability models without tract fixed effects
All urban US, men and women pooled

	People in the labor force								All people							
	Means		Probit		Logit		OLS		Means		Probit		Logit		OLS	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Gender indicator	--	--	--	-0.042	--	-0.040	--	-0.052	--	--	--	-0.329	--	-0.302	--	-0.471
				(<.0001)		(<.0001)		(<.0001)				(<.0001)		(<.0001)		(<.0001)
Hispanic	0.099	0.078	-0.007	-0.022	-0.006	-0.020	-0.007	-0.026	0.100	0.091	0.003	-0.016	0.005	-0.013	0.014	-0.020
	(0.298)	(0.268)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.301)	(0.288)	(0.000)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Black	0.097	0.125	-0.049	-0.049	-0.044	-0.046	-0.062	-0.056	0.109	0.126	-0.141	-0.032	-0.142	-0.029	-0.121	-0.036
	(0.297)	(0.331)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.312)	(0.332)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Married	0.629	0.576	0.036	0.015	0.033	0.014	0.049	0.019	0.608	0.614	0.126	-0.071	0.125	-0.065	0.108	-0.073
	(0.483)	(0.494)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.488)	(0.487)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Immigrant	0.007	0.007	0.002	0.000	0.002	0.001	0.770	0.003	0.007	0.007	-0.003	0.000	-0.003	0.001	0.659	0.001
	(0.084)	(0.082)	(0.146)	(0.741)	(0.118)	(0.441)	(0.023)	(0.082)	(0.084)	(0.081)	(0.272)	(0.828)	(0.270)	(0.689)	(0.450)	(0.703)
Age	35.846	35.905	0.004	0.006	0.003	0.006	0.007	0.011	35.879	36.105	0.011	0.029	0.011	0.027	0.009	0.035
	(9.564)	(9.578)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(9.696)	(9.743)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Age squared *.01	13.764	13.809	-0.004	-0.007	-0.003	-0.006	-0.008	-0.012	13.813	13.985	-0.018	-0.039	-0.017	-0.036	-0.014	-0.048
	(7.087)	(7.087)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(7.186)	(7.251)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Related children in household	0.818	0.790	0.001	-0.007	0.001	-0.007	0.000	-0.011	0.786	0.921	0.011	-0.054	0.012	-0.050	0.006	-0.068
	(1.132)	(1.071)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.703)	(<.0001)	(1.127)	(1.175)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Education level:																
Less than grade 9	0.046	0.029	-0.038	-0.061	-0.033	-0.054	-0.036	-0.077	0.055	0.047	-0.202	-0.182	-0.208	-0.164	-0.131	-0.257
	(0.209)	(0.168)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.227)	(0.211)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Grade 9-12	0.116	0.093	-0.041	-0.048	-0.035	-0.042	-0.029	-0.067	0.129	0.122	-0.131	-0.129	-0.130	-0.115	-0.064	-0.181
	(0.321)	(0.290)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.335)	(0.327)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Some college	0.207	0.224	0.016	0.014	0.014	0.013	0.007	0.017	0.203	0.210	0.052	0.055	0.053	0.049	0.024	0.066
	(0.446)	(0.463)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.442)	(0.452)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Associates Degree	0.067	0.086	0.025	0.027	0.024	0.026	0.016	0.029	0.064	0.077	0.087	0.094	0.092	0.085	0.036	0.110
	(0.250)	(0.281)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.246)	(0.267)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
BA	0.177	0.170	0.031	0.029	0.030	0.027	0.018	0.034	0.167	0.153	0.113	0.082	0.114	0.072	0.058	0.104
	(0.381)	(0.376)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.373)	(0.360)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
More than BA	0.095	0.076	0.035	0.031	0.034	0.029	0.022	0.031	0.089	0.063	0.171	0.163	0.192	0.153	0.075	0.162
	(0.293)	(0.264)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.285)	(0.244)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Employed	0.938	0.940							0.866	0.699						
	(0.241)	(0.237)							(0.341)	(0.459)						
	52.350	47.650							46.959	53.041						
Observations	4,118,811	3,749,025	7,867,836		7,867,836		7,867,836		4,462,738	5,040,683	9,503,421		9,503,421		9,503,421	

Note: In Means columns, standard deviations are in parentheses, all other columns probability that coefficients are zero in parentheses. The Means columns report means, the OLS columns report all others report marginal effects. Marginal effects calculated using the $\bar{X}\beta$ method, using discrete changes for indicator variables. Hypotheses tested with likelihood

Table 2.3: Employment Probability Models without Tract Effects (California)

	People in the labor force				All people			
	Probit		OLS		Probit		OLS	
	Men	Women	Men	Women	Men	Women	Men	Women
Gender indicator	--	-0.043	--	-0.054	--	-0.385	--	-0.522
		(<.0001)		(<.0001)		(<.0001)		(<.0001)
Hispanic	-0.002	-0.014	0.000	-0.016	0.039	0.018	0.039	0.022
	(0.002)	(<.0001)	(0.871)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Black	-0.060	-0.053	-0.066	-0.053	-0.160	-0.054	-0.132	-0.057
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Married	0.036	0.010	0.044	0.012	0.124	-0.083	0.100	-0.085
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Immigrant	0.000	0.003	0.795	0.005	-0.006	0.013	0.639	0.015
	(0.899)	(0.371)	(0.905)	(0.177)	(0.371)	(0.014)	(0.383)	(0.004)
Age	0.003	0.006	0.006	0.010	0.012	0.031	0.010	0.036
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Age squared *.01	-0.004	-0.006	-0.007	-0.011	-0.020	-0.041	-0.016	-0.047
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Related children in household	0.001	-0.005	0.000	-0.007	0.003	-0.049	0.002	-0.060
	(0.051)	(<.0001)	(0.147)	(<.0001)	(<.0001)	(<.0001)	(0.000)	(<.0001)
Education level:								
Less than grade 9	-0.033	-0.073	-0.036	-0.095	-0.141	-0.157	-0.131	-0.212
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Grade 9-12	-0.041	-0.045	-0.029	-0.061	-0.128	-0.128	-0.064	-0.171
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Some college	0.022	0.019	0.007	0.023	0.072	0.080	0.024	0.094
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Associates Degree	0.029	0.030	0.016	0.031	0.101	0.112	0.036	0.127
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
BA	0.036	0.040	0.018	0.035	0.133	0.110	0.058	0.144
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
More than BA	0.039	0.031	0.022	0.032	0.145	0.138	0.075	0.187
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Observations	1007421		1007421		1239023		1239023	

Note: Probability that coefficients are zero in parentheses. The OLS columns report coefficients, all others are marginal effects. See body of the thesis for description of how marginal effects are calculated.

Table 2.4: Employment Probability Models without Tract Effects (Florida)

	People in the labor force				All people			
	Probit		OLS		Probit		OLS	
	Men	Women	Men	Women	Men	Women	Men	Women
Gender indicator	--	-0.049	--	-0.074	--	-0.291	--	-0.433
		(<.0001)		(<.0001)		(<.0001)		(<.0001)
Hispanic	-0.009	-0.032	-0.007	-0.030	-0.003	-0.025	0.003	-0.029
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.264)	(<.0001)	(0.221)	(<.0001)
Black	-0.047	-0.039	-0.052	-0.043	-0.120	-0.013	-0.099	-0.012
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Married	0.032	0.011	0.039	0.015	0.118	-0.088	0.090	-0.093
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Immigrant	-0.004	0.002	0.850	0.003	0.001	0.006	0.675	0.006
	(0.442)	(0.698)	(0.521)	(0.583)	(0.893)	(0.482)	(0.948)	(0.484)
Age	0.003	0.006	0.004	0.008	0.014	0.028	0.011	0.034
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Age squared *.01	-0.004	-0.006	-0.006	-0.009	-0.023	-0.039	-0.018	-0.047
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Related children in household	0.002	-0.004	0.002	-0.005	0.020	-0.043	0.011	-0.053
	(0.000)	(<.0001)	(0.002)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Education level:								
Less than grade 9	-0.034	-0.056	-0.036	-0.073	-0.156	-0.158	-0.131	-0.222
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Grade 9-12	-0.022	-0.039	-0.029	-0.053	-0.085	-0.125	-0.064	-0.148
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Some college	0.008	0.012	0.007	0.013	0.034	0.048	0.024	0.056
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Associates Degree	0.018	0.024	0.016	0.024	0.058	0.073	0.036	0.082
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
BA	0.021	0.029	0.018	0.026	0.083	0.067	0.058	0.081
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
More than BA	0.026	0.025	0.022	0.025	0.099	0.105	0.075	0.134
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Observations	343,247		343,247		413,091		413,091	

See notes for Table 2.3

Table 2.5: Employment Probability Models to Estimate Tract Effects (California)

California, men and women pooled

	People in the labor force						All people					
	Probit		OLS		WLS		Probit		OLS		WLS	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Gender indicator	--	-0.035	--	-0.051	--	-0.175	--	-0.357	--	-0.499	--	-1.561
		(<.0001)		(<.0001)		(<.0001)		(<.0001)		(<.0001)		(<.0001)
Hispanic	0.004	-0.006	0.008	-0.008	0.043	0.034	0.049	0.028	0.049	0.031	0.204	0.147
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Black	-0.041	-0.033	-0.053	-0.038	-0.070	-0.043	-0.112	-0.016	-0.092	-0.020	-0.147	-0.023
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Married	0.032	0.008	0.043	0.010	0.097	0.052	0.118	-0.090	0.092	-0.090	0.262	0.031
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Immigrant	0.002	0.003	0.002	0.005	-0.020	0.004	-0.004	0.012	-0.003	0.014	-0.175	0.298
	(0.481)	(0.267)	(0.540)	(0.139)	(<.0001)	(0.534)	(0.565)	(0.023)	(0.565)	(0.006)	(<.0001)	(<.0001)
Age	0.003	0.005	0.006	0.009	0.005	0.015	0.013	0.030	0.010	0.035	-0.022	0.072
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Age squared *.01	-0.004	-0.005	-0.007	-0.010	-0.009	-0.019	-0.021	-0.039	-0.016	-0.046	0.018	-0.104
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Related children in household	0.001	-0.004	0.002	-0.006	-0.027	-0.037	0.005	-0.046	0.003	-0.056	-0.027	-0.155
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Education level:												
Less than grade 9	-0.023	-0.056	-0.030	-0.083	-0.062	-0.143	-0.120	-0.139	-0.092	-0.191	-0.268	-0.291
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Grade 9-12	-0.034	-0.037	-0.053	-0.056	-0.091	-0.088	-0.116	-0.117	-0.100	-0.158	-0.378	-0.535
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Some college	0.018	0.015	0.023	0.020	0.083	0.054	0.066	0.075	0.050	0.089	0.174	0.089
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Associates Degree	0.024	0.025	0.027	0.027	0.120	0.098	0.092	0.107	0.063	0.121	0.252	0.145
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
BA	0.028	0.031	0.034	0.029	0.098	0.063	0.124	0.107	0.089	0.138	0.259	0.048
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
More than BA	0.032	0.025	0.035	0.026	0.051	-0.003	0.141	0.137	0.103	0.185	0.105	-0.076
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.052)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Individuals *Tracts*	1,007,421	*5,544*	1,007,421	*5,544*	970,728	*4,890*	1,239,023	*5,561*	1,239,023	*5,561*	1,200,490	*5,086*
Ho: All fixed effects are equal												
Reject at 1% level?	Yes		Yes		N/A		Yes		Yes		N/A	

Note: Probability that coefficients are zero in parentheses. The OLS (Ordinary Least Squares) and WLS (Weighted Least Squares) columns report coefficients while the probit columns report marginal effects. See body of the thesis for description of how marginal effects are calculated.

Table 2.6: Employment Probability Models to Estimate Tract Effects (Florida)

Florida, men and women pooled

	People in the labor force						All people					
	Probit		OLS		WLS		Probit		OLS		WLS	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Gender indicator	--	-0.042	--	-0.069	--	-0.223	--	-0.278	--	-0.418	--	-1.157
		(<.0001)		(<.0001)		(<.0001)		(<.0001)		(<.0001)		(<.0001)
Hispanic	-0.003	-0.023	-0.001	-0.025	-0.012	-0.009	-0.001	-0.024	0.006	-0.028	0.027	0.020
	(<.0001)	(<.0001)	(0.395)	(<.0001)	(<.0001)	(<.0001)	(0.852)	(<.0001)	(0.061)	(<.0001)	(<.0001)	(<.0001)
Black	-0.028	-0.023	-0.036	-0.027	-0.066	-0.030	-0.079	0.017	-0.063	0.023	-0.083	0.020
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Married	0.027	0.009	0.037	0.013	0.060	0.032	0.109	-0.093	0.081	-0.097	0.176	-0.037
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Immigrant	-0.005	0.001	-0.005	0.003	0.049	-0.038	-0.003	0.005	-0.003	0.005	-0.020	-0.098
	(0.302)	(0.798)	(0.664)	(0.620)	(<.0001)	(<.0001)	(0.768)	(0.572)	(0.780)	(0.569)	(0.397)	(0.006)
Age	0.003	0.005	0.004	0.008	0.001	0.013	0.014	0.027	0.011	0.033	-0.004	0.073
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.321)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.013)	(<.0001)
Age squared *.01	-0.004	-0.005	-0.006	-0.009	-0.003	-0.017	-0.023	-0.037	-0.017	-0.046	-0.002	-0.110
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.341)	(<.0001)
Related children in household	0.002	-0.003	0.002	-0.005	-0.003	-0.017	0.018	-0.041	0.010	-0.052	0.032	-0.109
	(<.0001)	(<.0001)	(0.001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Education level:												
Less than grade 9	-0.026	-0.044	-0.036	-0.064	-0.012	-0.195	-0.145	-0.150	-0.131	-0.212	-0.138	-0.415
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Grade 9-12	-0.018	-0.033	-0.029	-0.049	-0.057	-0.118	-0.075	-0.117	-0.064	-0.140	-0.504	-0.539
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Some college	0.006	0.010	0.007	0.012	0.000	0.010	0.034	0.049	0.024	0.056	-0.108	-0.054
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.939)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Associates Degree	0.015	0.020	0.016	0.023	0.007	-0.025	0.057	0.075	0.036	0.083	-0.013	-0.056
	(<.0001)	(<.0001)	(0.000)	(<.0001)	(0.055)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.839)
BA	0.018	0.024	0.018	0.024	0.021	0.002	0.085	0.072	0.058	0.086	-0.014	-0.249
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.218)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.005)	(<.0001)
More than BA	0.022	0.021	0.022	0.023	0.012	-0.006	0.105	0.109	0.075	0.141	-0.131	-0.081
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.033)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Individuals *Tracts*	343,247	*2,276*	343,247	*2,276*	322,747	*1,843*	413,091	*2,282*	413,091	*2,282*	391,122	*1,959*
Ho: All fixed effects are equal												
Reject at 1% level?	Yes		Yes		Yes		Yes		Yes		Yes	

See notes for Table 2.5

Table 2.7: Employment Probability Models without Tract Effects for Men

	Men in labor force		All men	
	Probit	OLS	Probit	OLS
All of Urban US				
Black	-0.048	-0.063	-0.100	-0.121
Hispanic	-0.007	-0.007	0.003	0.015
Age	0.004	0.007	0.009	0.011
Age sq./100	-0.004	-0.008	-0.014	-0.016
Married	0.045	0.049	0.112	0.115
Immigrant	0.002	0.003	-0.002	-0.002
Education level:				
Less than grade 9	-0.029	-0.047	-0.149	-0.177
Grade 9-12	-0.032	-0.060	-0.092	-0.120
Some college	0.019	0.021	0.038	0.041
BA	0.033	0.036	0.076	0.076
More than BA	0.037	0.035	0.082	0.081
Observations	4118811	4118811	4462738	4462738
State=California				
Black	-0.059	-0.066	-0.119	-0.132
Hispanic	-0.002	0.000	0.029	0.040
Age	0.003	0.006	0.009	0.010
Age sq./100	-0.004	-0.007	-0.014	-0.016
Married	0.043	0.045	0.101	0.102
Immigrant	0.000	0.000	-0.004	-0.004
Education level:				
Less than grade 9	-0.032	-0.040	-0.102	-0.110
Grade 9-12	-0.040	-0.057	-0.092	-0.113
Some college	0.025	0.029	0.055	0.062
BA	0.048	0.042	0.091	0.102
More than BA	0.040	0.043	0.096	0.109
Observations	551865	551865	601427	601427
State=Florida				
Black	-0.043	-0.052	-0.083	-0.096
Hispanic	-0.008	-0.007	-0.001	0.005
Age	0.003	0.005	0.012	0.013
Age sq./100	-0.004	-0.006	-0.019	-0.020
Married	0.039	0.041	0.101	0.102
Immigrant	-0.004	-0.003	0.001	0.000
Education level:				
Less than grade 9	-0.032	-0.042	-0.116	-0.141
Grade 9-12	-0.021	-0.032	-0.060	-0.074
Some college	0.010	0.010	0.027	0.027
BA	0.021	0.020	0.057	0.053
More than BA	0.026	0.023	0.067	0.063
Observations	174884	174884	190321	190321

Note: Marginal effects are reported for the probit models and coefficients are reported for the OLS models.

Table 2.8: Employment Probability Models to Estimate Tract Effects for Men

	Men in labor force			All men		
	Probit	OLS	WLS	Probit	OLS	WLS
California:						
Black	-0.035	-0.051	-0.074	-0.066	-0.083	-0.151
Hispanic	0.004	0.009	0.035	0.041	0.053	0.201
Age	0.003	0.006	0.000	0.010	0.012	-0.029
Age sq./100	-0.004	-0.007	-0.002	-0.016	-0.018	0.026
Married	0.035	0.045	0.052	0.092	0.094	0.220
Immigrant	0.002	0.002	-0.047	-0.001	-0.002	-0.203
Education level:						
Less than grade 9	-0.021	-0.031	-0.061	-0.077	-0.089	-0.269
Grade 9-12	-0.030	-0.053	-0.091	-0.077	-0.098	-0.401
Some college	0.018	0.024	0.105	0.047	0.052	0.210
BA	0.033	0.034	0.114	0.077	0.083	0.276
More than BA	0.029	0.034	0.063	0.084	0.091	0.116
Individuals	551865	551865	519660	601427	601427	576436
Tracts	5522	5522	4526	5522	5522	4770
Florida:						
Black	-0.021	-0.034	-0.067	-0.048	-0.057	-0.074
Hispanic	-0.002	0.000	-0.013	0.004	0.010	0.029
Age	0.003	0.005	-0.002	0.012	0.013	0.000
Age sq./100	-0.003	-0.006	0.001	-0.018	-0.021	-0.008
Married	0.028	0.039	0.052	0.087	0.090	0.229
Immigrant	-0.003	-0.004	0.034	-0.002	-0.002	0.107
Education level:						
Less than grade 9	-0.022	-0.039	-0.007	-0.106	-0.134	-0.134
Grade 9-12	-0.014	-0.029	-0.072	-0.049	-0.063	-0.550
Some college	0.007	0.009	0.001	0.024	0.024	-0.112
BA	0.014	0.018	0.019	0.050	0.048	-0.051
More than BA	0.018	0.022	0.012	0.061	0.061	-0.127
Individuals	174884	174884	156437	190321	190321	175897
Tracts	2265	2265	1564	2272	2272	1710

Table 2.9: Means of Tract Employment Measures

Samples with men and women pooled:	California		Florida	
	In labor force	All	In labor force	All
Employment rate	0.933	0.758	0.940	0.773
	(0.050)	(0.110)	(0.043)	(0.095)
OLS tract effect	0.796	0.626	0.842	0.655
	(0.037)	(0.080)	(0.033)	(0.078)
WLS tract effect	-0.007	-0.007	-0.005	-0.007
	(0.032)	(0.066)	(0.029)	(0.067)
Probit tract marginal effect	-0.008	-0.009	-0.008	-0.014
Relative to mean effect	(0.033)	(0.079)	(0.028)	(0.075)
Probit tract marginal effect	-0.006	-0.018	-0.004	-0.019
Relative to median effect	(0.033)	(0.079)	(0.028)	(0.075)
Probit tract marginal effect	0.010	0.022	0.010	0.012
Relative to 25th percentile effect	(0.033)	(0.079)	(0.028)	(0.075)
Probit tract marginal effect	-0.019	-0.046	-0.016	-0.046
Relative to 75th percentile effect	(0.033)	(0.079)	(0.028)	(0.075)
Number of tracts in sample	4995	5087	1900	1959

Men only samples:	California		Florida	
	In labor force	All	In labor force	All
Employment rate	0.934	0.854	0.942	0.857
	(0.050)	(0.106)	(0.046)	(0.099)
OLS tract effect	0.789	0.606	0.836	0.627
	(0.039)	(0.083)	(0.036)	(0.076)
WLS tract effect	-0.008	-0.006	-0.007	-0.005
	(0.035)	(0.063)	(0.033)	(0.067)
Probit tract marginal effect	-0.015	-0.014	-0.018	-0.017
Relative to mean effect	(0.036)	(0.081)	(0.032)	(0.071)
Probit tract marginal effect	-0.005	-0.016	-0.004	-0.015
Relative to median effect	(0.036)	(0.081)	(0.032)	(0.071)
Probit tract marginal effect	0.014	0.020	0.012	0.013
Relative to 25th percentile effect	(0.036)	(0.081)	(0.032)	(0.071)
Probit tract marginal effect	-0.022	-0.044	-0.018	-0.040
Relative to 75th percentile effect	(0.036)	(0.081)	(0.032)	(0.071)
Number of tracts in sample	4725	4798	1672	1725

Table 2.10: Correlation Between Sets of Tract Effect Estimates

Correlation between alternate employment measures

	Probit effect	OLS Effect	WLS effect	Employment rate
California (obs.=4890)				
Probit effect	1.000			
OLS Effect	0.972	1.000		
WLS effect	0.912	0.931	1.000	
Employment rate	0.898	0.925	0.750	1.000
Florida (obs.=1843)				
Probit effect	1.000			
OLS Effect	0.964	1.000		
WLS effect	0.947	0.956	1.000	
Employment rate	0.868	0.922	0.788	1.000

Table 2.11a and 2.11b: Correlations of Tract Effects Across Samples

Table 2.11a: Correlation between measures estimated with pooled and men samples

	Pooled			
	Probit effect	OLS Effect	WLS effect	Employment rate
Men only				
California (obs.=4526)				
Probit effect	0.8519			
OLS Effect	0.845	0.8702		
WLS effect	0.7605	0.7785	0.8318	
Employment rate	0.824	0.845	0.680	0.915
Florida (obs.=1564)				
Probit effect	0.753			
OLS Effect	0.748	0.794		
WLS effect	0.707	0.728	0.766	
Employment rate	0.724	0.790	0.648	0.865

Table 2.11b: Correlation between measures estimated with and without non-participants

	Pooled, in labor force			
	Probit effect	OLS Effect	WLS effect	Employment rate
Pooled, all				
California (obs.=4890)				
Probit effect	0.684			
OLS Effect	0.688	0.715		
WLS effect	0.494	0.504	0.599	
Employment rate	0.693	0.722	0.508	0.827
Florida (obs.=1843)				
Probit effect	0.586			
OLS Effect	0.591	0.634		
WLS effect	0.490	0.486	0.554	
Employment rate	0.610	0.667	0.528	0.749

CHAPTER THREE

The Impact of Enterprise Zones on Resident Employment

3.1 Introduction

Enterprise zones are geographically targeted programs that are usually implemented with the goals of increasing tax revenues, creating jobs, and providing services to underserved areas. The mechanism for achieving those goals is ambiguous. Businesses are offered a combination of subsidies, low-interest loans, or government services and the desired outcomes are supposed to be generated by businesses' responses to the program. Measuring the effects of such a program is difficult. First, the mechanism for the causal relationship is very unclear. Second, it is necessary to find an outcome that can be measured at a level of geographic detail that is similar to that of the programs.

A reduced-form estimate of the effect of enterprise zones on resident employment resolves both of these problems. It imposes very little structure on how enterprise zones influence resident employment, and simply estimates what the effect of the programs are without defining how the programs yield that effect. In the absence of experimental data, resident employment may be the only outcome that can be studied at a sufficiently detailed level of geography. To my knowledge, there is no widely available source of business data with sufficient geographic detail to evaluate business outcomes at the Census tract-level. Most of the evaluations of the impact of enterprise zones on business outcomes that have been done use Zip Code as the unit of analysis (e.g. Engberg and Greenbaum 1999, Greenbaum and Engberg 2000, Bonodonio 2000, Peters and Fisher 2002), which I find to be a very poor measure of where most enterprise zones are located. The Decennial Censuses provide 1-in-6 sample estimates of employment and other characteristics at the tract-level, which allows for good measures of the characteristics of zones before and after designation.

The downside of looking at resident employment is that place of residence is only one of many factors that influence people's employment. In order to sort out the effect of a geographically based program it is necessary to control for confounding factors. Chapter 2 of this dissertation documents the creation of estimates of the component of employment probability associated with each tract. This chapter uses those estimates to estimate the average treatment effect on resident employment probability of containing an enterprise zone for tracts that contain an enterprise zone. To generate these estimated effects, I estimate the counterfactual state using propensity score matching. Essentially, this technique chooses non-treated tracts to function as controls for the zone tracts based on the similarity between the zone and non-zone tracts in observable characteristics correlated with containing an enterprise zone. I find propensity score matching suitable for this application because the states' processes for selecting enterprise zones were based on observable characteristics and because I have a sufficient number of non-zone tracts with propensity scores similar to zone tracts.

My results consistently show that, conditional on the characteristics of zone residents, enterprise zones had no measurable effect on the employment of zone residents who participate in the labor market. The estimated effects range from –1.7 to 2.0 percentage points, are centered near zero, and are rarely more than one standard error away from zero. This is contrary to some of the prior literature, which found substantial (and counterintuitive) negative effects of enterprise zones on resident employment. If I ignore the characteristics of zone residents, the estimated effects are consistently in the neighborhood of –1.0 percentage point. This suggests that what drives the difference between my results and the prior literature is neither the different technique used to estimate the counterfactual nor better measurement of zone location and characteristics. My results are different because I control for the characteristics of zone residents and therefore obtain a better estimate of the component of employment explained by residential neighborhood.

This chapter starts with a description of how I implement propensity score matching. The second section discusses the creation of the propensity scores. The third section presents the estimates of the average treatment effect on the treated for the effect of enterprise zones on resident employment probabilities. The fourth section presents robustness checks and addresses the issue of induced migration. The chapter ends with conclusions for both the chapter and the dissertation.

3.2 Propensity Score Matching Methodology

In the first chapter, I gave a formal description of propensity score matching and the assumptions that must be made to use this methodology. Here I describe the details of propensity score matching. There are two components to propensity score matching: estimating the propensity score and calculating the difference in outcomes between treated and matched observations.

In this thesis, the propensity score serves as an index that summarizes observable characteristics that are correlated with containing an enterprise zone and with the employment rate of tract residents. I estimate the propensity score with a tract-level probit model, where the dependent variable is an indicator for containing an enterprise zone and the independent variables are controls for the demographic and economic characteristics of the tracts. I do this separately for California and Florida because the selection process is different in each state and economic trends may have varied across the two states. As a test of sensitivity to the selection of variables included in the probit models, I estimate a set of propensity scores where the demographic variables included in the model include socio-economic data from the 1970 Census.

This thesis uses two matching estimators, nearest neighbor and Epanechnikov kernel. Both provide a way to estimate what the outcome of a treated observation would be in the absence of the treatment. This discussion of those estimators closely follows Black and Smith (2004). These matching estimators have the shared form for the estimate of the counterfactual outcome for treated observation *i*:

$$\hat{E}(Y_0 \mid \hat{P}(X_i)) = \sum_{j=1}^{J} w(\hat{P}(X_i), \hat{P}(X_j))Y_{0j},$$

where j=1,…,J indexes the observations in the untreated comparison group and $\hat{P}(X_i)$ is the estimated propensity score for the vector of characteristics of observation i, X_i. The matching estimators differ by their weighting functions, w(.). The nearest neighbor estimator uses a weight of 1 for the closest non-treated observation and 0 for all others. Formally, the weighting function for the nearest neighbor estimator with caliper d is:

$$w(\hat{P}(X_i), \hat{P}(X_j)) = \begin{cases} 1 \text{ if } j = \arg\min\limits_{k \in (T=0)} \left\{ | \hat{P}(X_i) - \hat{P}(X_k) | \right\} \text{ and } | \hat{P}(X_i) - \hat{P}(X_k) | \leq d; \\ 0 \text{ otherwise,} \end{cases}$$

with T=1 if the observation is treated and 0 otherwise. Kernel estimates create a composite control observation based on a number of non-treated observations, where the distance from the treated observation using the propensity score as a metric determines the contribution of each non-treated observation. Formally, the weighting functions with kernel function K(.) and bandwidth a_n is:

$$w(\hat{P}(X_i), \hat{P}(X_j)) = \frac{K\left[\dfrac{\hat{P}(X_i) - \hat{P}(X_j)}{a_n}\right]}{\sum\limits_{k \in \{T=0\}} K\left[\dfrac{\hat{P}(X_i) - \hat{P}(X_k)}{a_n}\right]}$$

and the Epanechnikov kernel function is:

$$K(\lambda) = \begin{cases} \dfrac{3}{4}(1-\lambda^2) \text{ if } |\lambda| < 1 \\ 0 \text{ otherwise.} \end{cases}$$

Each of the matching estimators has different strengths and weaknesses. Nearest neighbor matching matches a treated observation to the most similar non-treated observation, so the matches are as similar as possible. The downside is that there may be many close matches, so by choosing a single match one ignores other non-treated observations that are also valid counterfactuals. Because nearest neighbor matching ignores this source of information, nearest neighbor estimates often have higher variance than kernel matching estimates. When there are multiple potential matches, kernel matching uses these potential matches weighted by the closeness of the propensity scores. The downside of kernel matching with finite samples is that there is the potential for non-treated observations that are dissimilar to the treated observation to influence the estimate. The kernel estimator is well suited to my application because for most zone tracts there are a number of non-treated tracts with close propensity scores. Also, I have relatively few treated observations, so the higher variance of the nearest neighbor estimator is especially problematic. I provide estimates using both estimators in the tables.

Another difference between matching estimators is what caliper or bandwidth is used. The caliper used in nearest neighbor matching determines the greatest acceptable difference in propensity score between a zone tract and a non-treated tract. If a zone tract does not have a potential match within the distance of the caliper, the tract is not included in the estimates of the effect. The bandwidth serves a similar purpose for Epanechnikov kernel matching. For a zone tract, all non-zone tracts within the bandwidth are incorporated in the counterfactual estimate while all those outside the bandwidth are ignored. In my work, I found my estimates sensitive to the caliper with nearest neighbor estimators and very robust to bandwidth selection with kernel estimators. For this reason, I provide nearest neighbor estimates with calipers of 0.01 and 0.10 and kernel estimates with bandwidth 0.025.

3.3 Propensity scores

The first step in propensity score matching is to estimate a propensity score that summarizes observable characteristics that are correlated with selection into treatment and the outcome of interest. This section discusses the probit models I use to create the propensity scores and the resulting scores. I find that my models fit the data well. I also find that there is sufficient overlap between the distributions of the scores for the zone and non-zone tracts to use matching estimators. All but the worst-off enterprise zone tracts have potential matches.

In order to focus on areas similar to enterprise zones, some sample restrictions are placed on all stages. Tracts where less than 95 percent of the 1990 population lived in an urban area are dropped. All but six enterprise zone tracts met this restriction, while 15 percent of non-zone tracts are eliminated. Tracts with 1980 population below 100 are dropped to reduce problems with measurement error and missing data. This cut 648 non-zone tracts and 14 zone tracts from the sample. In order to eliminate isolated urban areas that would not be appropriate controls for enterprise zones, the analysis was restricted to tracts located in a Metropolitan Statistical Area in 1990. This eliminated eight enterprise zone tracts and 287 non-zone tracts.[49] The sample is further restricted to those tracts where at least 25 people are in the sample for the employment probability models. Like other statistics, fixed-effects estimated with very small samples have high variance and the point estimate may be far from the true effect. As explained in the previous chapter, because I cannot estimate the weighted least squares (WLS) employment probability models for tracts with employment rates equal to one, those tracts are dropped from the sample. One criticism of enterprise zone policies is that they are likely to pull business development away from areas near but not in the zone because a small change in location would yield a reduction in taxes. If these negative spillovers exist, using tracts near enterprise zones as matches for zone tracts may overstate the effect of zone programs. For this reason, I exclude all non-zone areas that are fewer than five miles from any enterprise zone.[50] In California, the final sample has 4,280 non-zone tracts and 93 zone tracts. In Florida, the corresponding numbers are 1,195 and 102.

The tracts' propensities of containing an enterprise zone are estimated using a probit. The degree to which enterprise zones are distressed can be seen in the difference in means between enterprise zone tracts and non-zone tracts shown in Table 3.1. The precise definitions of the tract-level variables are in the appendix.[51] In both states, the share of households receiving public assistance in 1980 in

---

[49] The zones eliminated from the study by the MSA restriction were Calexico and Eureka in California and Leesburg and Umatilla in Florida.

[50] The distance I use is the distance from the geographic center of the potential control tract to that of the zone.

enterprise zone areas is more than three times that of non-zone areas and the 1980 unemployment rate is nearly double. Enterprise zones were similarly more distressed than non-zone areas in the 1970's. Zone tracts had more than double the poverty and unemployment rates in 1970 of non-zone areas. Zone tracts remained disadvantaged in 1990. They had lower rates of post-secondary education and higher rates of single mother households than non-zone tracts.

The bottom of Table 3.1 provides information for three different 1990 employment measures. The employment rate for men and women together is lower for zone tracts than for all non-zone tracts by 9.8 percentage points in California and 8.9 percentage points in Florida. The unconditional difference in men's employment rates is very similar to that for all people (-9.0 in California and -8.0 in Florida). The difference between zones and non-zones in the conditional employment probability is much smaller than the difference in the unconditional employment rate: -2.9 percentage points in California and –2.4 in Florida. This shows that conditioning on who lives in an area in 1990 does not eliminate the difference between zone tracts and all non-zone tracts in employment rates but does reduce it by more than 70 percent. This is what one would expect if the composition of residents were an important determinant of tract-level employment rates that should be controlled for when estimating the effect of enterprise zones on employment.

---

[51] The means in this Tableare only for those tracts that had data in 1980. All tracts with 1980 data also had 1990 data. Those areas that were not assigned tracts in 1970 are not included in the means for the 1970 characteristics.

The probit estimates for California and Florida, shown in Table 3.2, are largely consistent with the differences in means. The first set of conditioning variables includes only traits from the 1980s and the second set includes variables from 1980 and 1970. The variables included measures of socio-economic distress (unemployment, poverty, education, household structure, building age, and vacancy rate) and the economic climate in the city or town the tract is located in (the job growth rate from 1983 to 1986). The probit models fit very well for non-zone areas; more than 85 percent of non-zone areas in each state have propensity score below 0.05. Figures 3.1a through 3.2c illustrate the degree of support for each of the propensity scores. These graphs show that there are potential matches for all but a few enterprise zone tracts, however there are few potential matches for zones with a propensity score above 0.40. Support is weakest for the zone tracts with the highest propensity scores. Figure 3.1b shows the counts of zone and non-zone areas by range of propensity scores in California.[52] In most ranges of the propensity score distribution, there are non-zone areas that can serve as controls.

3.4 Effects of enterprise zones on resident employment

I estimate the effect of enterprise zones on two outcomes. The first is the employment rate of people aged 18 to 55 and in the labor force, which I refer to as the employment rate. It is the probability that a resident of a tract who is in the labor force is employed and does not condition on the traits of the person. The second outcome is the tract effect estimate from the employment probability models

[52] This figure is typical of the other propensity score distributions.

discussed in Chapter 2. These tract effects are the component of employment probability related to tract of residence after conditioning on the traits of the individuals living in the tract. Therefore, I call the tract effects the conditional employment probabilities. To make the measures comparable, the sample of people used to estimate the conditional employment probability is also used to calculate the employment rate. I also present results for the employment rate and conditional employment probability for people who are out of the labor force.

One common concern when using propensity score matching is whether the score is sufficient for finding appropriate matches. This issue is addressed by figures 3.3a (California) and 3.3b (Florida) and by the "Matched" columns in Table 3.1. The graphs illustrate the average characteristics for all zone tracts, all non-zone tracts, and zone and non-zone tracts included in the nearest neighbor estimates.[53] These graphs demonstrate that the propensity score is sufficient to choose non-zone tracts that match zone tracts on these characteristics. This shows that my propensity scores fulfill their role as balancing scores. Not all treated observations have appropriate matches, which raises the question of whether the treated observations that can be matched are very different from those that cannot. As shown in figures 3.3a and 3.3b, the matched zone tracts are less disadvantaged than the full sample of zone tracts, but the differences are slight.

The average treatment effect of containing an enterprise zone is estimated for two different measures of employment outcomes: the employment rate in the tract

[53] Specifically, the estimates with caliper of 0.01 using the propensity score estimated with 1980 variables.

and the conditional employment probability (the tract effect from the employment probability model). For both outcomes, three different matching estimates are presented for each set of propensity scores. The estimators I use are: nearest neighbor with calipers of 0.01 and 0.10 and Epanechnikov kernel with bandwidth of 0.025. The reason for the variety of estimates is that the standard errors of the estimated effects are high relative to the effects; considering multiple estimates helps to test robustness of the results. While I include estimates using the nearest neighbor with caliper of 0.10 estimates in the tables, I do not discuss them because they suffer from poor support and may be misleading. This is seen in the high rates of repeated use of controls and the low number of control tracts relative to zone tracts, which are reported in the bottom half of each table.

Bootstrapping the employment models, propensity score estimation, and matching 300 times generates the standard errors in tables 3.3 to 3.8. I do this by taking samples of tracts with replacement from the actual sample for each model so that each sample drawn has the same number of tracts as the true sample. I use bootstrapped standard errors because the estimation strategy I use has multiple stages and the relationship between the stages is not straightforward. The bootstrapped standard errors show how sensitive the estimates are to the sample of tracts used to generate the estimates. If a small number of tracts are driving the results, the bootstrapped standard errors should be high. The bootstrapped errors do not adjust for the fact that the propensity scores and the conditional employment probabilities are estimated rather than known, but it does adjust for the sensitivity of those stages to sample composition.[54]

I use the WLS employment models so that it is feasible to bootstrap all stages of the estimation process. The probit and OLS models with tract effects take many hours to estimate, so repeating them 300 times for each of the different samples I consider would be infeasible. For example, each probit model with fixed effects for California takes at least 11 hours to estimate.[55] Using the technology available to me, it would take approximately four months to bootstrap the probit estimation with 300 replications for one sample. I show in Chapter 2 that the tract effects estimated using the WLS models are highly correlated with the effects estimated with the other models. For this reason, I believe that the benefit gained by using the WLS models, being able to estimate standard errors that reflect the estimation of the tract effects, comes at little cost.

The estimates of the effect of enterprise zones on resident employment are in Table 3.3. The first and second columns have the estimated effects of enterprise zones on, respectively, the employment rate and conditional employment probability of zone residents in California. When only conditioning on 1980 variables, the estimated effect of California enterprise zones on the employment rate from nearest neighbor matching is –0.8 percentage points and the kernel estimate is -1.5

[54] Based on Andrews and Buchinsky (2001), the level of certainty that the bootstrapped standard error is within 10 percent of the standard errors estimated with an infinite number of bootstrap replications is at least 95 percent with 300 repetitions for all matching results presented.
[55] The hours of computing time are given for the fastest machine available to me, a workstation with four 20GHz Pentium 4 XEON CPUs, 16 GB of RAM, SCSI Ultra 320 KVE drives, and a Linux operating system. The machine has a speed rating of 3,984.58 BogoMIPS (a standardized measure of

percentage points. The estimated effects of zones on conditional employment probability range from -0.4 to -0.5 percentage points. Controlling for the characteristics of 1990 zone residents, the results indicate that California enterprise zones had no substantive effect on resident employment. Given the semi-parametric estimators I use and the size of the estimated effects, it is not surprising that all of the estimated effects for California are not different from zero at any standard level of significance.

The estimated effects of Florida enterprise zones are in the third and fourth columns of Table 3.3. Conditioning on variables from 1980, the estimates of the impact of enterprise zones on the employment rate of residents range from -1.2 to –0.7 percentage points. The estimated effects of Florida enterprise zones on the conditional employment probability are less negative by about 0.6 percentage points, ranging from –0.4 to -0.1 percentage points. When I control for the characteristics of zone residents, the estimated effect on resident employment is nearly zero. This shows that ignoring the traits of zone residents can yield misleading results.

The employment outcomes discussed above were for men and women together. An alternative measure is the employment outcomes of men. Table 3.4 has the matching estimates of the effect of enterprise zones on the employment rate and conditional employment probability of men. These estimates were generated with the same three stage methodology, but calculating the employment outcomes for only men and restricting the matching sample to those tracts where at least 25 men

---

the relative speed of a Linux machine). The software I use for these models is SAS, version 8.2 for Linux.

were included in the employment probability models.[56] The results for the propensity score matching stage are not reported, but are qualitatively similar to those presented in Table 3.2.

In California, the estimated effects on men's employment rates are more varied than those for all people's employment rates. Looking at the raw employment rate, the estimated effects range from -1.3 to 1.8 percentage points. The effect on the conditional employment probability of men is equally varied, ranging from -0.4 to 2.0 percentage points. Note that the kernel estimates, which are the most precise, are very similar to those for all people. In Florida, the estimated effects of enterprise zones on the male employment rate are more negative than those for the employment rate of all workers in Table 3.3. The nearest neighbor estimate of the effect of Florida zones on the conditional employment probability of men is –0.7 percentage points while the kernel estimate is –1.2 percentage points. Both are closer to zero than the estimated effect on the raw employment rate and all estimates are not statistically different from zero.

The estimates discussed above all use the same characteristics from the 1980 Census and the job growth rate from 1983 to 1986 to estimate the propensity of a tract to contain a zone. However, it might be important to control for long-term trends to find areas that are good matches for enterprise zones, which were typically distressed for many years prior to designation. To address this concern, I estimate the propensity of a tract to contain an enterprise zone adding characteristics from the

[56] The results are qualitatively similar if I use men from the samples in Table3.3 instead.

1970 Census.[57] The marginal effects from these probits are in Table 3.2 and show that some 1970 characteristics are correlated with zone status. Table 3.5 has the estimated effects on employment outcomes when conditioning on these tract characteristics. The estimated effects are qualitatively similar to the estimates using only 1980 characteristics to estimate the propensity score that are presented in Table 3.3. The nearest neighbor estimates of the effect of California zones on the raw and conditional employment measures are –1.6 and –0.7 percentage points. The kernel estimate of the effect on the conditional employment probability is 0.2 percentage points, which is 0.8 percentage points higher than the raw employment rate estimate. In Florida, the estimated effects on the raw employment rate when focusing on matches that were similar in both 1970 and 1980 is –2.7 for both estimators and the conditional employment probability effects range from –1.3 to –1.7 percentage points. These results are suggestive that residents of Florida zones located in major cities – the areas of Florida that were assigned tracts in the 1970 Census – were less likely to be employed due to the zones. However, the estimated effects are not significantly different from zero. In general, these results lead to the same conclusion as the results based on matching only on characteristics from the 1980's: ignoring the characteristics of residents would give an overly pessimistic estimate of the effect of enterprise zones on the employment of zone residents.

Even if enterprise zones did not increase the employment probabilities of zone residents, they might benefit people in the area surrounding the zones by

[57] Adding these traits reduces the sample because some 1980 tracts were not given tracts in the 1970 Census.

encouraging development in the zones. Alternatively, the zone programs could also reduce the employment probabilities of people living near but not in the zones by encouraging businesses to locate in the zone and hire zone residents. Given the fact that the estimated effect of zones on resident employment is close to zero, it would be surprising to find that the zone programs reduced employment prospects for people living near but not in a zone. Nonetheless, in order to see if there are any measurable spillover effects of this sort, I estimate the effect of enterprise zones on residents of tracts not in the enterprise zones but within five miles of the center of the zones. The methodology is very similar to that used to estimate the effect of zones on zone residents. The key difference is that the propensity score estimated is the propensity of a tract to be located within five miles of zone and not contain a zone. The estimates of the average treatment effect of enterprise zones on the employment of residents of tracts surrounding a zone are in Table 3.6. The raw employment rate effect estimates range from -0.6 to –0.3 percentage points in California and are both 0.2 percentage points in Florida. Controlling for the characteristics of people living around the zone brings the estimated effects up by no more than 0.8 percentage points. While the precision of the estimates of the effect of enterprise zones on nearby tracts is much higher than the precision of the estimates of the effect of zone tracts, all but one of these estimates is statistically insignificant at the ten percent level. My results do not support the hypothesis that zones reduced the employment of people living in the areas surrounding enterprise zone.

The results to this point have only looked at the employment of people who are labor market participants, which means they are either employed or looking for employment. During the 1980's, many communities experienced rapid growth in the number of men out of the labor force. Therefore, it is interesting to know whether enterprise zones impacted the proportion of all working age people who are employed. If enterprise zones created new jobs, the zones may have encouraged marginal workers to enter or remain in the labor market. I estimated the effect of enterprise zones on the proportion of all men who are employed using the samples of men including non-participants. The individual-level samples and the conditional employment probability estimates for these samples are discussed in Chapter 2. I focus on only men because it is unambiguous whether a change in the proportion of prime-aged men employed is good or bad. Because most men aged 18-55 participate in the labor market, it is likely that men out of the labor force are either not able to work or discouraged workers, meaning they have dropped out of the labor market because of inability to find an acceptable job. I repeat the propensity score estimation process for the sample of tracts with at least 25 men in the individual-level employment models when non-participants are included. The propensity score stage results are very similar to those in Table 3.2 and are not reported.

The matching estimates of the effect of enterprise zones on the proportion of all working-age men who are employed are in Table 3.7. In California, the estimated effect on the raw employment rate of all men ranges from –3.3 to –1.8 percentage points. The estimated effects on the conditional employment probability range from

–2.6 to –1.4 percentage points. In Florida, the estimated effects on the employment rate of all men range from –3.4 to –1.0 percentage points. The effects on the conditional employment probability range from -0.9 to 0.1 percentage points. These point estimates suggest that enterprise zone programs reduced the probability that a man living in a zone is participating in the labor market and employed. In California, the estimated effects on the conditional employment probability are less negative than the effects on the employment rate. While the point estimates are consistently negative, they are not statistically significant. The results for Florida are more typical of the other matching results: when conditioning on resident characteristics, the estimated effect of enterprise zones on resident employment is close to zero.

3.5 Sensitivity analysis

To test the robustness of the results discussed above, I repeated the estimation with alternative sample specifications, conditioning variables, and matching schemes. For example, I estimated the individual-level employment probability models with interactions between indicators for race and all the regressors except the fixed effects; the average treatment effects from the final stage were very similar. I estimated the propensity score with a reduced set of controls and the effects on conditional employment probability were not markedly different. This section discusses in more detail two sensitivity checks: trimming outliers and using the conditional employment probability estimates from individual-level probit models.

This section ends with an investigation of whether or not enterprise zone tracts had differential migration patterns than matching non-zone tracts.

Propensity score matching can be very sensitive to outliers, especially when using finite samples. This is because an outlying observation that is not in the treatment group can be matched to multiple treated observations and have a large influence on the estimated effects. To see whether outliers drive my results, I repeated the three stages for samples that excluded observations in the lower tail of the employment rate and conditional employment probability distributions. I do not drop observations from the upper tail because I drop tracts with employment rates of one in order to use the WLS models. In addition, there are no striking outliers in the upper end of the distribution, only the lower end. Prior to estimating the WLS employment probability models, I drop tracts in the first percentile of the employment rate distribution from the sample. After estimating the WLS models, I drop observations in the first percentile of the tract effect distribution. Aside from these restrictions, I use the same methodology for these results as those discussed above.

The results when dropping outliers are in the first four columns of Table 3.8. For the California pooled sample, the trimmed estimates are not more than 0.3 percentage points from the corresponding estimates from the fuller sample. The kernel estimate for the Californian men sample is not sensitive to outlying observations, but the nearest neighbor estimate is. The nearest neighbor estimate of the effect of enterprise zones on men’s conditional employment probability is –0.5

percentage points in the trimmed sample while the estimate from the fuller sample is 2.0 percentage points. This unusually high estimated effect is not robust to trimming outliers. In Florida, the estimates are more sensitive to outliers than in California. The kernel estimate of the effect of zones on the pooled conditional employment probability is –0.1 in the full sample and is –1.2 in the trimmed sample. For the men samples, the corresponding figures are –1.2 and 0.1. The gap between the employment rate and the conditional employment probability effects are very similar in the trimmed and full samples. Even with the sensitivity of the point estimates to outliers, the results show that ignoring the characteristics of tract residents gives a more pessimistic estimate of the effect of zones on resident employment.

The last column of Table 3.8 has the estimated effects for Florida men using the probit estimates of the tract effect rather than the WLS estimates. When using the probit estimates, the marginal effect of a tract on an individual depends on the probability that individual is employed, as reflected in the z-score of the individual. To get a treatment effect estimate that is comparable to those already reported, it is necessary to convert the coefficients for the tract effects into the employment probability scale. To do that, I use propensity score matching to estimate the counterfactual tract coefficient and then convert that to a marginal effect on employment. Formally, the average treatment effect estimate becomes:

$$\Delta^{p} = \frac{1}{Z}\sum_{z=1}^{Z}\frac{1}{N_z}\sum_{\substack{i=1\\ i\in z}}^{N_z}\{F(\hat{\beta}X_i + \hat{\alpha}_z) - F(\hat{\beta}X_i + \hat{\alpha}_z^{M})\},$$

where Z is the number of zone tracts in the sample, N_z is the number of people living in the sample living in tract z (which contains a zone), $F(.)$ is the Normal cumulative density function, X_i is a vector of observable characteristics for individual i, $\hat{\beta}$ is the vector of coefficient estimates from the probit model, $\hat{\alpha}_z$ is the tract fixed effect estimate from the probit model for tract j, and $\hat{\alpha}_z^M$ is the matching estimate of the counterfactual tract fixed effect.[58] This is the formula I use for the estimates in the last column of Table 3.8. The probit-based estimates are similar to the WLS-based estimates. These estimates are within 0.5 percentage points of the estimates using the WLS tract effects reported in the last column of Table 3.4. The probit-based kernel estimate is more negative than the WLS-based kernel estimate while the probit-based nearest neighbor estimate is less negative than the corresponding WLS-based estimate. These results suggest that my results are robust to the method chosen to estimate the tracts' conditional employment probabilities.

One potential critique of my results is that conditioning on the post-designation characteristics of zone residents is improper because changes in these characteristics could be an effect of the programs. If the programs make the zone tracts more attractive to different kinds of individuals, the composition of residents might change as a result of the policy and controlling for those changes ignores an important effect. I disagree with this critique. Given the small effects of enterprise zones on the outcomes they were intended to affect, it would be surprising if the zone programs had a significant impact on the composition of residents.

In order to see if there are substantial differences between zones and observationally similar non-zone tracts in who moved to the area, I look at the attributes of adults who moved to their 1990 residence after 1984 (movers) relative to the attributes of adults who moved to their 1990 residence no later than 1984 (stayers).[59] I look at the ratio of the traits of new residents to the traits of prior residents, ${}^{X_{move}}\!/\!{}_{X_{stay}}$. X_{move} are the characteristics of movers while X_{stay} are the characteristics of stayers. I use ratios to make it easier to interpret the resulting estimates. I also look at the difference between zone tracts and matching tracts in the percent of all adults who are movers. I use the same propensity score matching approach as above to compare zone tracts to tracts that were observationally similar in 1980 using the Epanechnikov estimator with bandwidth of 0.025.

The comparisons of movers to stayers are in Table 3.9. On average, 57.9 percent of adults in the zone tracts in California and 60.8 percent in Florida are movers, compared to 55.2 percent and 59.9 percent respectively for matching non-zone tracts. In both states, zones did not experience an atypical amount of moving. Comparing ${}^{X_{move}}\!/\!{}_{X_{stay}}$ for different characteristics shows that generally the relationship between the traits of movers and stayers are not very different in zones than in matching areas. If ${}^{X_{move}}\!/\!{}_{X_{stay}}=1$, then the average X for movers is the same

[58] I thank Judith Hellerstein and Jonah Gelbach for helping me determine the best way to create this estimate.
[59] Unfortunately, the 1990 Census has limited migration data. It is not possible to identify who arrived after 1986, when the zones were designated, or to identify who left the zone. It could also be

as for stayers. If ${X_{move}}/{X_{stay}}$ =0.5 (2), then the average X is half (double) that of stayers. For all but the ratio for the share of the population with more than a high school degree, in both states the difference in the ratios between zones and similar non-zone areas are less than 0.10. In California, the ratio for the share with more than high school was 1.14 in zones and 1.29 for matches, suggesting that zones were not as attractive to higher education people than non-zones. In Florida, the opposite was the case: the corresponding ratio for zone tracts was 0.15 higher than for non-zone tracts. Both of these differences are less than half of a standard deviation. The similarity between zone and matching tracts in the relationship between movers and stayers in most characteristics suggests that selective migration is not a major issue when evaluating the effects of zones.

The results discussed in this section suggest that the main average treatment effect results reported in the beginning of the chapter capture the impact of enterprise zones on resident employment. The estimates are fairly robust to whether outliers are included in the sample. At least for one sample, the estimated effects when using probit tract fixed effects are similar to the WLS-based estimates. Finally, there is no strong evidence that enterprise zone policies induced selective migration.

3.6 Conclusions

My work provides a comprehensive analysis of the impact of the enterprise zone programs of California and Florida on the employment of zone residents. In

---

that some of the people who moved into their residence after 1984 lived in the tract before in a different residence.

order to estimate the effects of enterprise zones on resident employment, I develop and implement a methodology that recognizes the fact that the selection into treatment occurs at the geographic level and the determination of employment occurs at the individual-level. In addition, I am careful to minimize error in the measurement of zone location and the characteristics of zones both before and after designation.

Most of the literature that has looked at the effects of enterprise zones on the employment rate of zone residents has found negative effects. When looking at the unconditional employment rate of residents, I also find that zone residents were less likely to be employed than residents of observationally similar areas, though the standard errors are too high to make a definitive conclusion. These results are misleading because they do not control for the characteristics of zone residents. Once I control for the characteristics of zone residents, the estimated effect of enterprise zones on resident employment probability is indistinguishable from zero in both California and Florida. The estimated effects that are large in magnitude appear to be sensitive to the inclusion of outliers in the sample. My estimates also do not support the belief that zones had negative spillovers for residents of nearby areas. One caveat that might explain the lack of measurable effects is that the programs studied had been in effect for only three years prior to 1990.

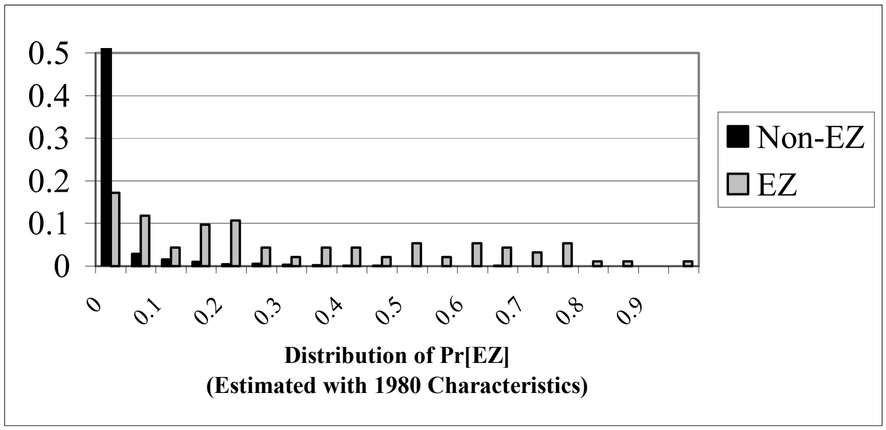
The enterprise zone programs of California and Florida were atypical in that they were carefully targeted and provided relatively large incentives for hiring zone residents and people with a history of unemployment. The majority of the zone program expenditures in these states were spent on hiring tax credits. If one were to expect a positive impact of enterprise zones on resident employment, it would be in these two states. I carefully measure zone location, control for the characteristics of people who lived in the zones in 1990, and systematically choose observationally similar non-zone tracts to use as comparison samples. In the end, I find that the enterprise zones of California and Florida had no measurable impact on the employment of residents. This provides further evidence that, at least at the historical level of expenditures, enterprise zones are not an effective way of increasing the probability that people in distressed communities are employed.

Figures for Chapter Three

Note: Throughout figures and tables, enterprise zone is abbreviated as EZ.

Figures 3.1a to 3.1c: Distribution of Propensity Score (California)

Fig. 3.1a



Note: 92.8% of non-EZ tracts had Pr[EZ] less than .05

Fig. 3.1b

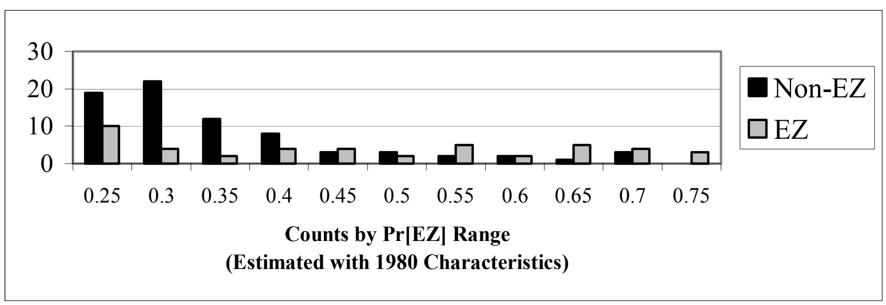
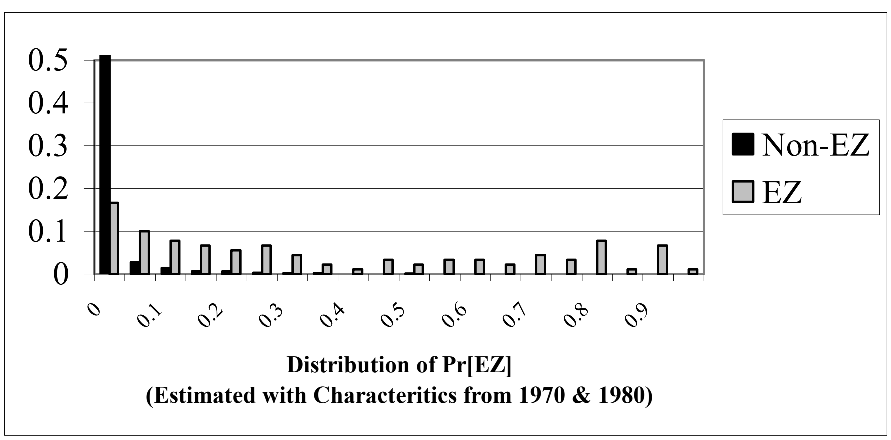


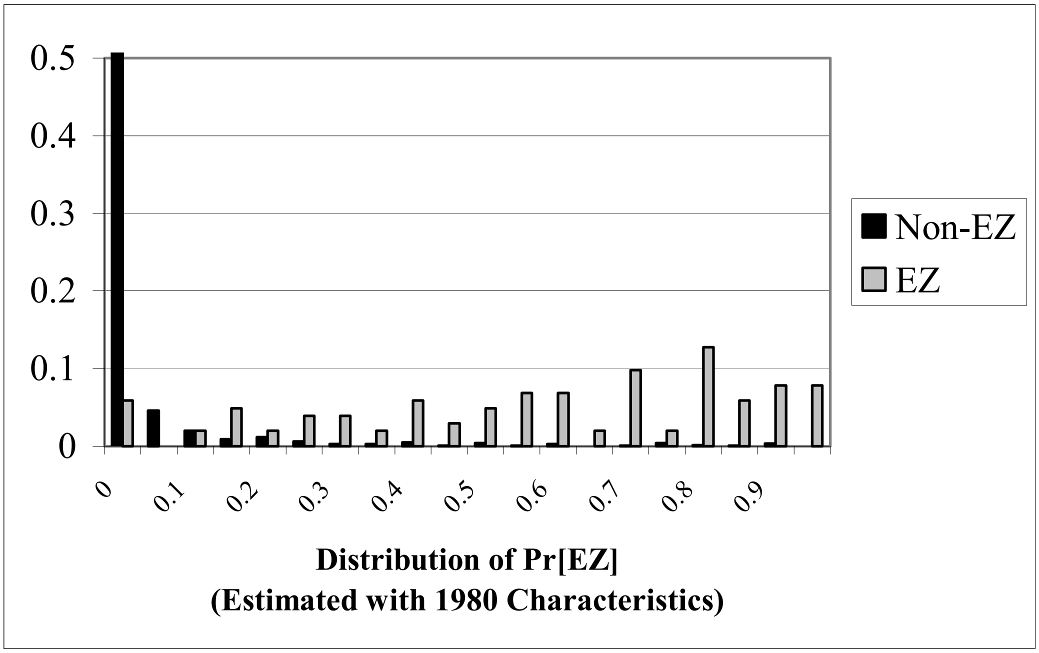
Fig. 3.1c



Note: 93.4% of non-EZ tracts have Pr[EZ] less than .05

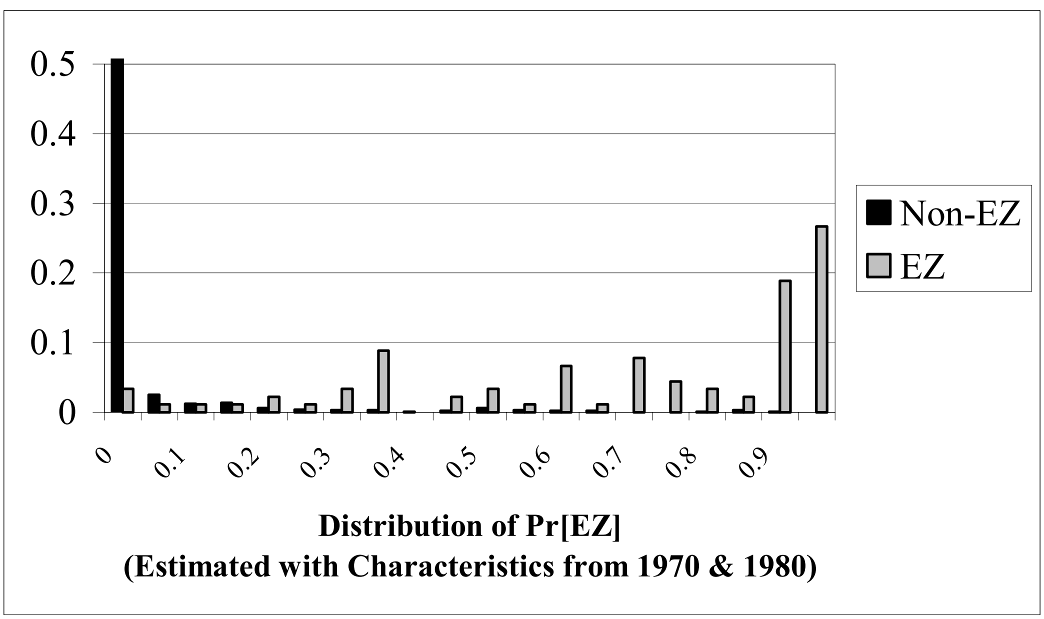
Figures 3.2a to 3.2b: Distribution of Propensity Score (Florida)

Fig. 3.2a



Note: 87.8% of non-EZ tracts have Pr[EZ] less than .05

Fig. 3.2b



Note: 91.0% of non-EZ tracts have Pr[EZ] less than .05.

Figures 3.3a and 3.3b: Match Quality

EZ compared to all areas and areas matched
Based on Nearest Neighbor matches with a caliper of .01
Using only 1980 conditioning variables to estimate propensity score

Fig. 3.3a: California

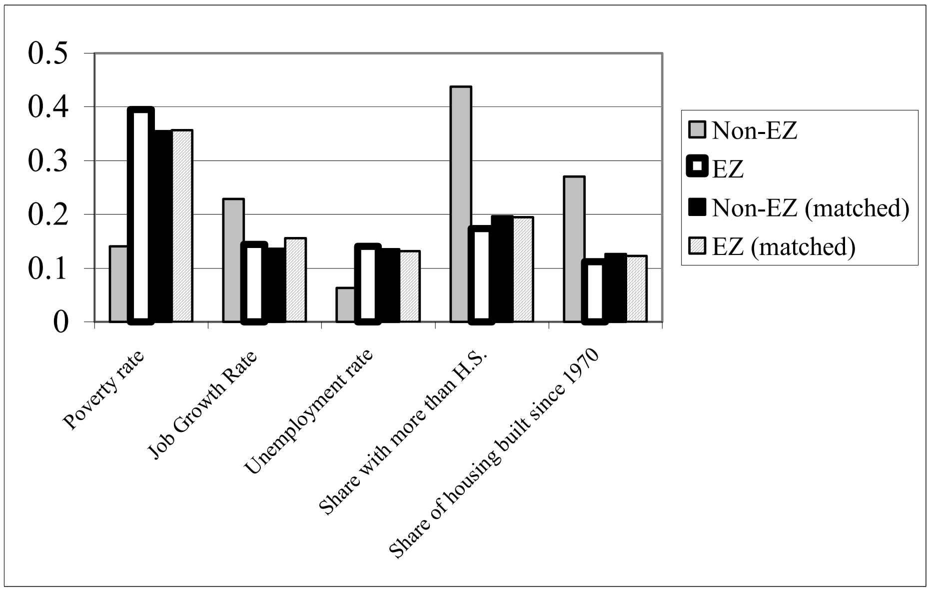
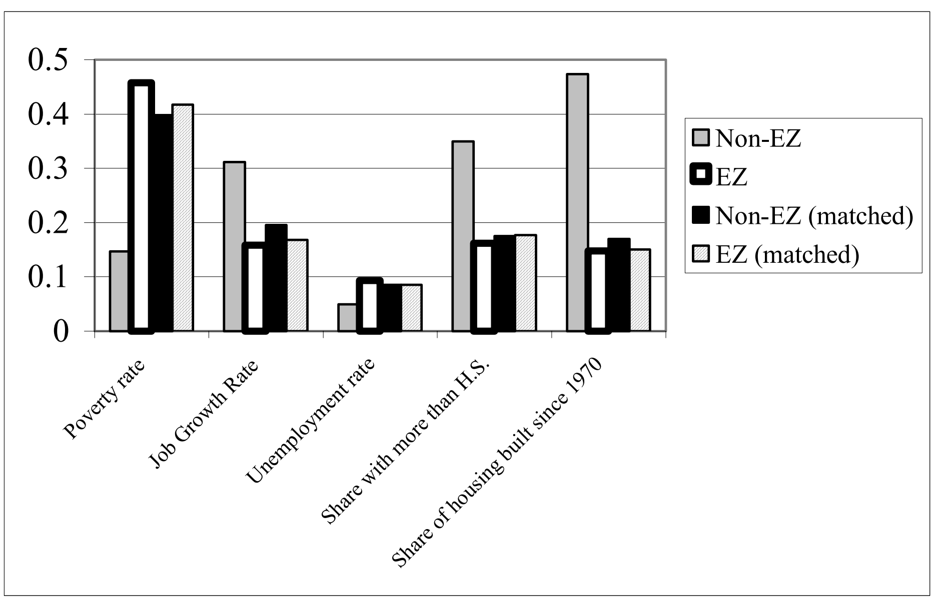


Fig. 3.3b: Florida



Tables for Chapter Three

Table 3.1: Means of Tract Characteristics

Non-zone areas restricted to those at least 5 miles from a zone

	California				Florida			
	All		Matched		All		Matched	
	EZ	Non-EZ	EZ	Non-EZ	EZ	Non-EZ	EZ	Non-EZ
Characteristics from 1980's								
Log of median household income	0.026	0.666	0.094	0.078	-0.244	0.497	-0.141	-0.086
	(0.316)	(0.389)	(0.312)	(0.339)	(0.331)	(0.339)	(0.309)	(0.287)
Unemployment rate	0.140	0.064	0.131	0.136	0.093	0.049	0.085	0.085
	(0.069)	(0.045)	(0.058)	(0.075)	(0.038)	(0.029)	(0.038)	(0.045)
Poverty rate	0.395	0.140	0.357	0.355	0.458	0.147	0.417	0.398
	(0.142)	(0.102)	(0.128)	(0.123)	(0.142)	(0.104)	(0.133)	(0.144)
Public assistance rate	0.282	0.086	0.251	0.243	0.189	0.050	0.166	0.153
	(0.139)	(0.072)	(0.124)	(0.108)	(0.090)	(0.047)	(0.080)	(0.090)
Share of adults with more than	0.174	0.437	0.195	0.197	0.162	0.350	0.177	0.175
HS	(0.106)	(0.175)	(0.111)	(0.087)	(0.100)	(0.151)	(0.110)	(0.081)
Share non-white	0.643	0.199	0.585	0.548	0.647	0.102	0.594	0.544
	(0.240)	(0.194)	(0.242)	(0.280)	(0.337)	(0.204)	(0.358)	(0.368)
Share single mother households	0.145	0.065	0.131	0.122	0.143	0.054	0.132	0.119
	(0.108)	(0.042)	(0.098)	(0.065)	(0.080)	(0.045)	(0.081)	(0.069)
Vacancy rate	0.060	0.050	0.061	0.062	0.100	0.114	0.095	0.096
	(0.038)	(0.055)	(0.037)	(0.046)	(0.060)	(0.107)	(0.055)	(0.056)
Share of units built in prior ten	0.112	0.270	0.123	0.126	0.148	0.474	0.151	0.170
years	(0.146)	(0.258)	(0.155)	(0.164)	(0.113)	(0.278)	(0.115)	(0.120)
Share of workers in	0.310	0.200	0.279	0.269	0.131	0.123	0.135	0.134
manufacturing	(0.135)	(0.101)	(0.121)	(0.120)	(0.078)	(0.061)	(0.081)	(0.071)
Job growth rate	0.144	0.229	0.156	0.137	0.159	0.311	0.168	0.196
	(0.088)	(0.210)	(0.093)	(0.104)	(0.126)	(0.256)	(0.125)	(0.111)
In remainder place	0.226	0.194	0.197	0.194	0.108	0.510	0.141	0.205
	(0.420)	(0.395)	(0.401)	(0.398)	(0.312)	(0.500)	(0.350)	(0.408)
Characteristics from 1970								
Unemployment rate	0.116	0.061	0.109	0.107	0.051	0.037	0.049	0.062
	(0.048)	(0.034)	(0.042)	(0.054)	(0.018)	(0.018)	(0.017)	(0.024)
Poverty rate	0.327	0.103	0.291	0.257	0.348	0.136	0.314	0.330
	(0.143)	(0.080)	(0.137)	(0.113)	(0.139)	(0.088)	(0.133)	(0.134)
Public assistance rate	0.254	0.068	0.225	0.193	0.131	0.030	0.117	0.120
	(0.123)	(0.061)	(0.108)	(0.121)	(0.069)	(0.035)	(0.063)	(0.069)
Share of adults with more than	0.123	0.319	0.137	0.139	0.112	0.246	0.122	0.129
HS	(0.073)	(0.157)	(0.076)	(0.063)	(0.076)	(0.128)	(0.087)	(0.074)
Share non-white	0.464	0.076	0.370	0.319	0.539	0.077	0.484	0.545
	(0.364)	(0.148)	(0.347)	(0.337)	(0.396)	(0.187)	(0.402)	(0.378)
Share single mother households	0.179	0.069	0.158	0.134	0.142	0.052	0.126	0.125
	(0.105)	(0.044)	(0.099)	(0.075)	(0.083)	(0.036)	(0.079)	(0.065)
Vacancy rate	0.066	0.048	0.060	0.057	0.074	0.078	0.069	0.074
	(0.029)	(0.054)	(0.025)	(0.026)	(0.053)	(0.073)	(0.047)	(0.040)
Share of workers in	0.279	0.217	0.263	0.234	0.138	0.148	0.144	0.112
manufacturing	(0.108)	(0.102)	(0.109)	(0.111)	(0.085)	(0.070)	(0.091)	(0.054)
Measures of employment in 1990								
Employment rate	0.840	0.938	0.852	0.865	0.856	0.945	0.863	0.885
	(0.083)	(0.043)	(0.077)	(0.076)	(0.069)	(0.032)	(0.055)	(0.055)
Mens employment rate	0.849	0.939	0.852	0.871	0.867	0.947	0.871	0.901
	(0.079)	(0.047)	(0.081)	(0.082)	(0.076)	(0.039)	(0.069)	(0.066)
Fixed effect from employment	-0.034	-0.006	-0.031	-0.024	-0.028	-0.004	-0.026	-0.021
probability model	(0.068)	(0.029)	(0.062)	(0.064)	(0.060)	(0.024)	(0.043)	(0.047)
Number of tracts	93	4,188	71	62	102	1,158	71	44
Tracts without 1970 data	3	254	2	1	12	235	9	14

Table 3.2: Probit Models to Estimate Tracts Propensity of Containing an Enterprise Zone

Non-zone areas restricted to those at least 5 miles from a zone

	Only 1980 Traits			1970 and 1980 Traits		
	Marginal Effect (thousandths)	Standard Error (thousandths)	Prob. that coefficient equals 0	Marginal Effect (thousandths)	Standard Error (thousandths)	Prob. that coefficient equals 0
California						
1980 Characteristics						
Log of median household income	0.097	(1.782)	0.96	0.399	(2.590)	0.88
Unemployment rate	-6.104	(6.028)	0.29	-13.882	(10.420)	0.14
Poverty rate	15.184	(7.975)	0.00	15.355	(9.057)	0.03
Public assistance rate	14.545	(8.392)	0.01	4.590	(8.421)	0.58
Share of adults with more than HS	-4.015	(2.884)	0.20	-5.712	(6.834)	0.39
Share non-white	5.809	(2.572)	0.00	4.846	(3.506)	0.15
Share single mother households	-11.375	(7.312)	0.04	-8.425	(10.229)	0.38
Vacancy rate	0.766	(6.353)	0.90	-4.681	(10.328)	0.64
Share of units built in prior ten year	-0.136	(1.841)	0.94	-0.666	(2.607)	0.80
Share of workers in manufacturing	14.082	(6.080)	0.00	23.767	(9.659)	0.00
Job growth rate	-0.495	(2.122)	0.82	1.749	(2.740)	0.52
In remainder place	1.236	(1.193)	0.17	1.051	(1.389)	0.37
1970 Characteristics:						
Unemployment rate				-6.728	(14.177)	0.63
Poverty rate				15.648	(9.670)	0.03
Public assistance rate				19.965	(12.154)	0.03
Share of adults with more than HS				5.728	(8.564)	0.47
Share non-white				3.340	(3.084)	0.22
Share single mother households				-13.881	(11.616)	0.19
Vacancy rate				7.027	(11.419)	0.53
Share of workers in manufacturing				-3.216	(6.457)	0.62
Observations	4281			4019		
Psuedo R sq.	0.464			0.489		
Florida						
1980 Characteristics						
Log of median household income	-0.005	(0.004)	0.08	-6.902	(5.748)	0.00
Unemployment rate	-0.003	(0.013)	0.80	2.201	(9.964)	0.82
Poverty rate	0.016	(0.013)	0.04	9.241	(10.138)	0.15
Public assistance rate	-0.001	(0.008)	0.90	1.115	(7.229)	0.87
Share of adults with more than HS	-0.004	(0.005)	0.41	12.591	(11.449)	0.02
Share non-white	0.005	(0.004)	0.02	2.645	(3.633)	0.36
Share single mother households	-0.016	(0.015)	0.14	-11.796	(13.592)	0.19
Vacancy rate	0.008	(0.008)	0.22	9.787	(9.500)	0.07
Share of units built in prior ten year	-0.009	(0.005)	0.00	-7.998	(6.063)	0.00
Share of workers in manufacturing	0.017	(0.012)	0.01	8.147	(8.785)	0.20
Job growth rate	-0.011	(0.007)	0.00	-3.911	(3.873)	0.09
In remainder place	-0.004	(0.003)	0.00	-2.947	(2.472)	0.00
1970 Characteristics:						
Unemployment rate				-35.917	(33.771)	0.05
Poverty rate				-14.290	(12.964)	0.02
Public assistance rate				2.610	(9.894)	0.79
Share of adults with more than HS				-19.873	(16.758)	0.00
Share non-white				1.931	(3.221)	0.49
Share single mother households				14.506	(15.478)	0.15
Vacancy rate				-5.749	(8.596)	0.45
Share of workers in manufacturing				13.903	(12.553)	0.02
Observations	1260			1012		
Psuedo R sq.	0.634			0.684		

Table 3.3: Matching Estimates of Effect of Enterprise Zones on Resident Employment

Outcome: employment measures for men and women in the labor force

Estimator	California Outcome: Employment rate	California Outcome: Conditional Employment Probability	Florida Outcome: Employment rate	Florida Outcome: Conditional Employment Probability
Epan. Kernel	-0.015	-0.005	-0.007	-0.001
Bandwidth=0.025	(0.012)	(0.010)	(0.013)	(0.011)
Nearest Neighbor	-0.008	-0.004	-0.012	-0.004
Caliper=0.01	(0.015)	(0.013)	(0.015)	(0.014)
Nearest Neighbor	-0.011	-0.005	-0.007	0.002
Caliper=0.10	(0.017)	(0.015)	(0.018)	(0.015)

Supplemental information:

Tracts contributing to estimate:	California Zone tracts	California Non-zone tracts	California Max. times repeated	Florida Zone tracts	Florida Non-zone tracts	Florida Max. times repeated
Epan. Kernel	79			92		
Nearest Neighbor, Caliper=0.01	71	62	3	71	44	6
Nearest Neighbor, Caliper=0.10	90	65	10	102	45	11

Number of people per tract (Nearest Neighbor, Caliper=0.01)	California Average	California Minimum	Florida Average	Florida Minimum
Zone tracts	166	29	116	25
Non-zone tracts	162	25	101	27

Notes: Bootstrapped standard errors with 300 repetitions in parentheses. The tract-level conditional employment probabilities are estimated with a WLS employment probability model. Propensity scores estimated with tract characteristics from the 1980's.

Table 3.4: Matching Estimates of Effect of Enterprise Zones on Male Resident Employment

Outcome: employment measures for men in the labor force

Estimator	California Outcome: Employment rate	California Outcome: Conditional Employment Probability	Florida Outcome: Employment rate	Florida Outcome: Conditional Employment Probability
Epan. Kernel	-0.013	-0.004	-0.014	-0.012
Bandwidth=0.025	(0.010)	(0.010)	(0.018)	(0.017)
Nearest Neighbor	0.018	0.020	-0.016	-0.007
Caliper=0.01	(0.015)	(0.014)	(0.022)	(0.021)
Nearest Neighbor	0.019	0.025	-0.015	-0.014
Caliper=0.10	(0.017)	(0.017)	(0.021)	(0.020)

Supplemental information: Tracts contributing to estimate:	California: Zone tracts	California: Non-zone tracts	California: Max. times repeated	Florida: Zone tracts	Florida: Non-zone tracts	Florida: Max. times repeated
Epan. Kernel	70			63		
Nearest Neighbor, Caliper=0.01	65	55	4	45	32	4
Nearest Neighbor, Caliper=0.10	76	55	8	83	30	10

Number of people per tract (Nearest Neighbor, Caliper=0.01)	California: Average	California: Minimum	Florida: Average	Florida: Minimum
Zone tracts	110	34	71	25
Non-zone tracts	111	25	57	26

Notes: see Table 3.3.

Table 3.5: Estimates of Effect - Propensity Score Estimated with

1970 and 1980 Characteristics

Outcome: employment measures for men and women in the labor force

Estimator	California Outcome Employment rate	California Outcome Conditional Employment Probability	Florida Outcome Employment rate	Florida Outcome Conditional Employment Probability
Epan. Kernel	-0.006	0.002	-0.027	-0.013
Bandwidth=0.025	(0.011)	(0.010)	(0.014)	(0.012)
Nearest Neighbor	-0.016	-0.007	-0.027	-0.017
Caliper=0.01	(0.014)	(0.012)	(0.019)	(0.016)
Nearest Neighbor	-0.016	-0.004	-0.037	-0.024
Caliper=0.10	(0.016)	(0.014)	(0.018)	(0.016)

Supplemental information: Tracts contributing to estimate:	California Zone tracts	California Non-zone tracts	California Max. times repeated	Florida Zone tracts	Florida Non-zone tracts	Florida Max. times repeated
Epan. Kernel	67			67		
Nearest Neighbor, Caliper=0.01	63	54	2	57	29	7
Nearest Neighbor, Caliper=0.10	86	56	11	90	30	17

Number of people per tract (Nearest Neighbor, Caliper=0.01)	California Average	California Minimum	Florida Average	Florida Minimum
Zone tracts	168	29	109	28
Non-zone tracts	175	27	114	27

Notes: Bootstrapped standard errors with 300 repetitions in parentheses. The tract-level conditional employment probabilities are estimated with a WLS employment probability model. Propensity scores estimated with tract characteristics from the 1980's and the 1970 Census.

Table 3.6: Estimates of Effect on People Who Live Near Zones

Effect of living within 5 miles of a zone but not in a zone
Outcome: employment measures for men and women in the labor force

Estimator	California Outcome: Employment rate	California Outcome: Conditional Employment Probability	Florida Outcome: Employment rate	Florida Outcome: Conditional Employment Probability
Epan. Kernel	-0.006	0.002	0.002	0.005
Bandwidth=0.025	(0.003)	(0.003)	(0.002)	(0.002)
Nearest Neighbor	-0.003	0.005	0.002	0.004
Caliper=0.01	(0.004)	(0.004)	(0.003)	(0.003)
Nearest Neighbor	-0.003	0.005	0.003	0.004
Caliper=0.10	(0.004)	(0.004)	(0.003)	(0.003)

Supplemental information: Tracts contributing to estimate:	California: Near-zone tracts	California: Distant tracts	California: Max. times repeated	Florida: Near-zone tracts	Florida: Distant tracts	Florida: Max. times repeated
Epan. Kernel	614			599		
Nearest Neighbor, Caliper=0.01	605	469	8	595	375	13
Nearest Neighbor, Caliper=0.10	619	470	9	600	375	13

Number of people per tract (Nearest Neighbor, Caliper=0.01)	California: Average	California: Minimum	Florida: Average	Florida: Minimum
Near-zone tracts	172	25	146	25
Distant tracts	180	27	168	25

Notes: The probability that a tract is located with-in five miles of an enterprise zone is modeled in the propensity score stage. Zone tracts are not included in the sample. Bootstrapped standard errors with 300 repetitions in parentheses. The tract-level conditional employment probabilities are estimated with a WLS employment probability model. Propensity scores estimated with tract characteristics from the 1980's.

Table 3.7: Matching Estimates of Effect of Enterprise Zones on Resident Men's Employment When Including Men Out of the Labor Force

Outcome: employment measures for men, including those out of the labor force

Estimator	California Outcome: Employment rate	California Outcome: Conditional Employment Probability	Florida Outcome: Employment rate	Florida Outcome: Conditional Employment Probability
Epan. Kernel	-0.018	-0.014	-0.010	0.001
Bandwidth=0.025	(0.022)	(0.017)	(0.027)	(0.023)
Nearest Neighbor	-0.033	-0.026	-0.034	-0.009
Caliper=0.01	(0.027)	(0.022)	(0.035)	(0.028)
Nearest Neighbor	0.013	0.005	0.010	0.019
Caliper=0.10	(0.034)	(0.024)	(0.034)	(0.024)

Supplemental information:	California Zone tracts	California Non-zone tracts	California Max. times repeated	Florida Zone tracts	Florida Non-zone tracts	Florida Max. times repeated
Tracts contributing to estimate:						
Epan. Kernel	81			71		
Nearest Neighbor, Caliper=0.01	73	62	3	52	38	4
Nearest Neighbor, Caliper=0.10	88	63	7	95	39	15

Number of people per tract (Nearest Neighbor, Caliper=0.01)	California Average	California Minimum	Florida Average	Florida Minimum
Zone tracts	120	26	73	27
Non-zone tracts	132	26	56	26

Notes: See Table 3.3.

Table 3.8: Sensitivity tests

	Dropping observations in the first percentile of the employment rate or conditional employment rate distributions				Alternate tract effect estimate
Equivalent to table:	3.3		3.4		3.4
Conditioning set:	1980's		1980's		1980's
Sample:	Women and men pooled		Men		Men in Florida
Outcome:	Employment rate	Conditional employment probability	Employment rate	Conditional employment probability	Probit estimate of conditional employment prob.
California					
Epan. Kernel	-0.016	-0.005	-0.015	-0.007	
Bandwidth=.025	(0.006)	(0.005)	(0.007)	(0.006)	
Nearest Neighbor	-0.007	-0.001	-0.007	-0.005	
Caliper=.01	(0.009)	(0.008)	(0.010)	(0.009)	
Florida					
Epan. Kernel	-0.020	-0.012	-0.006	0.001	-0.017
Bandwidth=.025	(0.009)	(0.009)	(0.016)	(0.014)	
Nearest Neighbor	-0.021	-0.010	-0.011	-0.003	-0.005
Caliper=.01	(0.012)	(0.011)	(0.018)	(0.015)	

Notes: Bootstrapped standard errors from 300 repetitions in parentheses.

Table 3.9: Comparison of Differences Between Movers and Stayers Between EZ and Matching Tracts

Matching with propensity score estimated with 1980 characteristics and Epanechnikov kernel estimator with bandwidth of .025.
Mover if moved to 1990 housing unit after 1984, stayer if moved no later than 1984.

	California			Florida		
	EZ's	Matches	Difference	EZ's	Matches	Difference
Share of adults living in different house than five years ago	0.579	0.552	0.027	0.608	0.599	0.010
	(0.132)	(0.053)		(0.125)	(0.058)	
Ratios of characteristics: (X move) / (X stay)						
Share non-white	1.064	1.143	-0.080	1.109	1.111	-0.002
	(0.230)	(0.212)		(0.570)	(0.378)	
Share with more than HS degree	1.137	1.289	-0.152	1.338	1.185	0.153
	(0.354)	(0.387)		(0.820)	(0.334)	
Share of HH headed by single mother	1.686	1.654	0.032	2.052	1.969	0.082
	(0.964)	(0.326)		(1.043)	(0.627)	
Employment rate	0.978	0.996	-0.018	0.966	0.960	0.006
(only people in labor force)	(0.094)	(0.051)		(0.078)	(0.047)	
Employment rate	1.190	1.227	-0.037	1.274	1.211	0.063
(including people not in labor force)	(0.192)	(0.125)		(0.369)	(0.303)	
Share very poor	1.526	1.522	0.003	1.325	1.326	-0.001
	(0.477)	(0.336)		(0.397)	(0.239)	

APPENDIX

Tract-level Variable Definitions

Data source listed in parentheses. The abbreviations used are: 1980 Census of Population and Housing: Summary Tape File 3A (STF80), 1970 Census of Population and Housing: Fourth Count (4C70), and my own tabulations from the Standard Statistical Establishment Lists of 1983-1986 (SSEL).

Variable definitions:

Employment rate: Share of people in the tract-place pair aged 18 to 55 and not enrolled in school who are in the non-military labor force and employed. (STF80 and 4C70).

In remainder place: Place not included in the 1982 Economic Census geography so that the 1982 Economic Census place code equals “9990”. (SSEL).

Job growth rate: (place-level employment in 1986 – place-level employment in 1983)/(place-level employment in 1983). (SSEL).

Poverty rate: The fraction of people in a tract-place pair who live in a family with income below 125 percent of the poverty line. (STF80 and 4C70).

Share of adults with more than HS: The fraction of adults in a tract-place pair aged 25 or above with more than a high school degree. (STF80 and 4C70).

Share of workers in manufacturing: The fraction of employed workers in the tract-place pair who worked in a manufacturing industry. (STF80 and 4C70).

Share non-white: 1 minus the share of tract-place pair population that is white. (STF80 and 4C70).

Share single mother households: Share of the tract-place pair households headed by an unmarried female with own children. (STF80 and 4C70).

Unemployment rate: Share of all people in the tract-place pair aged 15 or above who are in the non-military labor force and not employed. (STF80 and 4C70).

REFERENCES

Andrews, Donald W. K. and Moshe Buchinsky. "A Three-Step Method for Choosing the Number of Bootstrap Repetitions". Econometrica. 68(1), 23-51. 2000.

Black, Dan and Jeffrey Smith. "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching". Journal of Econometrics. forthcoming 2004.

Bondonio, Daniele. "Evaluating decentralized policies: How to compare the performance of state and local economic development programs across different regions". Paper presented in the Fourth EES Conference, Lausanne. September 8, 2000.

Bondonio, Daniele and John Engberg. "Enterprise Zones and Local Employment: Evidence from the States' Programs". Heinz School of Public Policy Working Paper 1999-13. 1999.

Bostic, Raphael W. and Allen C. Prohofsky. "Enterprise Zones and Individual Welfare: A Case Study of California". Mimeo. May 2002.

California Department of Commerce. <u>Employment and Economic Incentive Program: Report to the Legislature</u>. State of California: Sacramento, CA. 1987.

California Department of Commerce. <u>Enterprise Zone Act: Report to the Legislature</u>. State of California: Sacramento, CA. 1987.

California Legislature, Committee on Housing and Community Development. "SB 511 (Alarcon) Enterprise Zones: Where Do We Go From Here? : A Summary Report from the Special Hearing of the Senate Committee on Housing and Community Development". Senate Publications, State of California: Sacramento, CA. 1999.

Dowall, David E.. "An Evaluation of California' Enterprise Zone Programs". Economic Development Quarterly. Vol. 10 Issue 4. 352-69. November 1996.

Elvery, Joel. "The Impact of Enterprise Zones on Residents' Employment: An Evaluation of the Enterprise Zone Programs of California and Florida". Mimeo, University of Maryland. Sept. 2003.

Ellwood, David T. "The Spatial Mismatch Hypothesis: Are There Teenage Jobs Missing in the Ghetto?" in The Black Youth Employment Crisis, edited by Richard B. Freeman and Harry J. Holzer. 147-87. Chicago, IL: University of Chicago Press, 1986.

Engberg, John and Robert Greenbaum. "The Impact of State Enterprise Zones on Urban Housing Markets". Heinz School of Public Policy Working Paper 1999-16. 1999.

Greenbaum, Robert and John Engberg. "The Impact of State Urban Enterprise Zones on Business Outcomes". Center for Economic Studies Working Paper 98-20. 1998.

Greenbaum, Robert and John Engberg. "An Evaluation of State Enterprise Zone Policies". Policy Studies Review. 30-45. Summer/Autumn 2000.

Greene, William H. Econometric Analysis. 3rd Edition. Upper Saddle River, New Jersey: Prentice Hall, 1997.

Greene, William H. "The Bias of the Fixed Effects Estimator in Nonlinear Models." Mimeo. October 2002.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. "Characterizing Selection Bias Using Experimental Data". Econometrica. 66(5), 1017-1098. 1998.

Holzer, Harry J. "The Spatial Mismatch Hypothesis: What Has the Evidence Shown?". Urban Studies. 28(1), 105-22. 1998.

Ihlanfeldt, Keith R. and David L. Sjoquist. "The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform". Housing Policy Debate. 9(4), 849-892. 1998.

Kain, John. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." Quarterly Journal of Economics. 175-97. May 1968.

Lechner, Michael. "Identification and estimation of causal effects of multiple treatments under the conditional independence assumption". in Econometric Evaluation of Active Labour Market Policies. Edited by M. Lechner and F. Pfieffer. 43-58. Heidelberg: Physica, 2001.

Logan, Willie Jr. and Lee Ann Barron. "Florida's Enterprise Zone Program: The Program After Sunset". in Enterprise Zones : New Directions in Economic Development. Edited by Roy E. Green. 91-104. Newbury Park, Calif. : Sage Publications, 1991.

Office of the Auditor General, State of Florida. "Review and Evaluation of the Enterprise Zone Program". Office of the Auditor General Report No. 12003. February 1, 1993.

O'Keefe, Suzanne and Roger Dunstan. "Evaluation of California's Enterprise Zones". California Research Bureau Working Paper. CRB 01-012. August 2001.

Papke , Leslie. "Tax Policy and Urban Development: Evidence from the Indiana Enterprise Zone Program". Journal of Public Economics. 54(2), 37-49. 1994.

Papke, Leslie. "What Do We Know About Enterprise Zones?". NBER Working Paper 4251. Jan. 1993.

Peters, Alan H. and Peter S. Fisher. State Enterprise Zone Programs: Have They Worked?. Kalamazoo, Mich.: Upjohn Institute, 2002.

Raphael, Steven. “Inter- and Intra-Ethnic Comparisons of the Central City-Suburban Youth Employment Differential: Evidence from the Oakland Metropolitan Area”. Industrial and Labor Relations Review. 51(3), 505-24. 1998a.

Raphael, Steven. "The Spatial Mismatch Hypothesis and Black Youth Joblessness: Evidence from the San Francisco Bay Area". Journal of Urban Economics. 43. 79-111. 1998b.

Rosenbaum, Paul R. and Donald R. Rubin. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." Biometrika. 70. 1983. 41-55.

U.S. Dept. of Commerce, Bureau of the Census. CENSUS OF POPULATION AND HOUSING, 1970 [UNITED STATES]: SUMMARY STATISTIC FILE 4A - HOUSING [FOURTH COUNT] [Computer file]. ICPSR version. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census [producer], 1972. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001.

U.S. Dept. of Commerce, Bureau of the Census. CENSUS OF POPULATION AND HOUSING, 1970 [UNITED STATES]: SUMMARY STATISTIC FILE 4A - POPULATION [FOURTH COUNT] [Computer file]. 2nd ICPSR version. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census [producer], 1972. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001.

U.S. Dept. of Commerce, Bureau of the Census. CENSUS OF POPULATION AND HOUSING 1980 [UNITED STATES]: 1970-PRE 1980 TRACT RELATIONSHIPS [Computer file]. ICPSR version. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census [producer], 1981. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001.

U.S. Dept. of Commerce, Bureau of the Census. CENSUS OF POPULATION AND HOUSING, 1980 [UNITED STATES]: SUMMARY TAPE FILE 3A [Computer file]. ICPSR version. Washington, DC: U.S. Dept. of Commerce, Bureau of the

Census [producer], 1982. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1999.

U.S. Dept. of Commerce, Bureau of the Census. Geographic Areas Reference Manual. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census, 1994.

Van Allen, Terry. The Impact of Enterprise Zones on Employment. Austin & Winfield Publishers: Bethesda, MD. 1996.

Wilder, Margaret G. and Barry M. Rubin. "Rhetoric versus Reality: A Review of Studies on State Enterprise Zone Programs". Journal of the American Planning Association. V. 62 N. 4 (Autumn 1996). 473-91.