**Distribution Agreement**



Tanner Lewis April 2, 2019

The Effect of Tax Increment Financing Districts on Residential Property Prices: An Analysis of the Atlanta Real Estate Market

by

Tanner Lewis

Zhongjian Lin
Adviser

Department of Economics

Zhongjian Lin

Adviser

Maria Arbatskaya

Committee Member

Seunghwa Rho

Committee Member

2019

The Effect of Tax Increment Financing Districts on Residential Property Prices: An Analysis of the Atlanta Real Estate Market

By

Tanner Lewis

Zhongjian Lin

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Economics

2019

Abstract

The Effect of Tax Increment Financing Districts on Residential Property Prices: An Analysis of the Atlanta Real Estate Market

By Tanner Lewis

This paper uses a difference-in-difference hedonic regression model to measure the effect of Tax Increment Financing (TIF) designation on residential property sales prices. Based on an analysis of tax parcel sales in Atlanta, Georgia, this paper finds that TIF designation has a significant, positive effect on residential property values. This paper also employs several methods to account for selection bias resulting from non-random designation of TIF districts, including a Heckman selection model and propensity score weighting and matching. These methods point toward an even stronger positive effect of TIF. This paper concludes that, at least in Atlanta, TIF has been an effective tool for local economic development.

The Effect of Tax Increment Financing Districts on Residential Property Prices: An Analysis of the Atlanta Real Estate Market

By

Tanner Lewis

Zhongjian Lin
Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Economics

2019

Acknowledgements

I would like to thank my advisor, Zhongjian Lin, for guiding me through my research process. I would also like to thank my other committee members, Maria Arbatskaya and Seunghwa Rho, for evaluating my work. Invest Atlanta, Central Atlanta Progress, and the Fulton County Tax Assessors Office deserve recognition for helping me find data. Josh Jayasundara and Megan Slemons of the Emory Center for Digital Scholarship were invaluable resources for GIS analysis. Finally, I would like to thank my family for their support.

Table of Contents

List of Figures

## I. Introduction

Local government plays a major role in guiding the development of cities. Decisions related to public works projects, business incentives, infrastructure, school funding, affordable housing, and other elements of urban planning can rejuvenate a declining city or stunt the growth of a thriving metropolis. Unfortunately, for decades municipalities have regularly faced difficult trade-offs between development projects imposed by budget realities, and poor investments can have a lasting effect on their constituents.

Thus, it comes as no surprise that local governments have been looking for creative ways to stretch their budgets in the absence of state and federal aid. One such method that has become popular over the last several decades is Tax Increment Financing (TIF). First developed in 1952 in California, TIF rapidly spread nationally in the 1980s (Byrne 2005). By the mid-2000s, TIF had cemented itself as the most popular local development financing mechanism in the United States (Briffault 2010). In cities like Chicago, TIF has practically monopolized funding for economic development (Briffault 2010).

TIF has been mired in controversy in spite of its proliferation. Concerns about program effectiveness, regulatory capture, risks to school funding, and gentrification have fueled substantial research in the social sciences. These misgivings, coupled with post-Great Recession budgetary shortfalls, resulted in the elimination of TIF in its home state of California in 2012 (Swenson 2015). Studies of TIF's effects are vitally important as TIF nears a potential turning point. This paper will attempt to partially quantify TIF's benefit to its target community by examining changes in Atlanta, Georgia's real estate market. The results will contribute to the debate surrounding TIF's effectiveness as a tool for local development in Atlanta and beyond.

Before delving into the specific contributions of this paper, it is important to have a basic understanding of how TIF works and why local governments find it so attractive. TIF was originally envisioned as a targeted method to stimulate development in blighted areas. The local government designates all properties in a geographically-defined blighted area as a TIF district. The government then typically issues bonds that fund public works projects or developer incentives in the district. Property tax revenue is capped at the level of the base year, and any future property tax revenue above the base year value is pledged to paying off the bonds. As the blighted area develops, property values should rise and the incremental tax revenue should increase. By the end of the TIF implementation, the area should be substantially improved, and the local government should reap the rewards of a much larger tax base. Some municipalities maintain the cap until the bonds are fully paid off, while others specify a number of years after which budgeted money will go towards the bonds. In short, TIF funds for a present-day project with its anticipated future revenue.

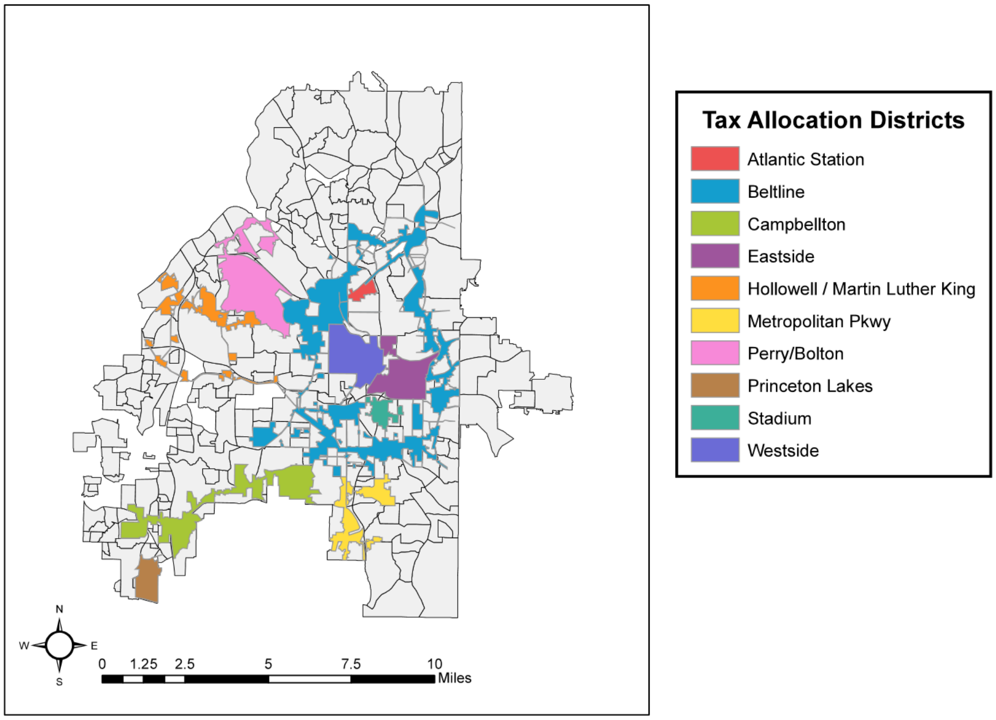
The primary appeal of TIF to local governments boils down to the unique benefit of expanding the tax base while avoiding the political and legal hurdles of tax increases. However, this lack of political accountability may result in inefficiently high usage of TIF. Overuse is concerning, since the costs of TIF are potentially large and difficult to measure. If a local government designates a TIF district in an area what would have grown without treatment, TIF becomes an unnecessary subsidy to developers that poaches from the tax base instead of growing it. Worse, if a local government implements a TIF and the district does not improve, the government will have to cut other programs to pay off the bonds. Given these costs, local governments must be equipped with the data and tools necessary to conduct a thorough cost-benefit analysis.

Studying TIF has major implications for the field of economics and society at large. Economists should be interested in TIF since it is at the heart of local economic development. TIF sits at the intersection of a variety of issues in urban and development economics, such as the efficacy of government incentives, the costs and benefits of development, the multiplier effect of public spending, and the importance of blight removal. TIF also presents interesting methodological challenges related to selection bias. It is no surprise that there is already a substantial literature base in economics related to TIF, though the debate is far from settled.

A better understanding of TIF is also important for society. Because it is so popular, additional information on TIF's costs and benefits could affect thousands of projects that reshape our cities and impact local tax revenue. Governments should be able to make evidence-based decisions about TIF and be prepared to fully exploit its opportunities and minimize its harms. Although many municipalities default to TIF whenever they need funds, an economic development opportunity alone does not ensure that TIF is the best tool.

Atlanta's approach to TIF is somewhat unique. In fact, Georgia is the only state to refer to TIF as Tax Allocation District (TAD) Financing. I will refer to the abstract concept as TIF and the Georgia implementation as TAD for clarity. Georgia first legalized TADs in 1985 under the Redevelopment Powers Law and began designating TADs in the 1990s but did not issue bonds until 2001 (Bourdeaux and Matthews 2004). Atlanta's first TAD was the Westside TAD in 1998 and the most recent was Campbellton in 2006. In total, there are ten TADs in Atlanta, seven of which are currently active (Invest Atlanta 2019). TADs have coincided with major revitalization efforts in Atlanta and are major funding sources for Atlantic Station, the Beltline, and other attractions.

Figure 1: Tax Allocation Districts and Neighborhood Statistical Areas in Atlanta




Georgia places several restrictions on TADs to prevent overutilization. First, TADs may not cover more than 10 percent of the taxable value of the jurisdiction (in this case, the City of Atlanta). Second, the jurisdiction must submit a redevelopment plan that includes some form of cost-benefit analysis that finds evidence that the area is blighted and will not develop “but for” the TAD. Finally, overlapping jurisdictions, such as school districts and counties, must consent to the TAD (Bourdeaux and Matthews 2004). However, there is no formal standard for blight or clear selection criteria for TAD designation.

This paper is structured as follows: Section II examines the previous literature on TIF and clarifies the research question and scope. Section III lists the data sources and cleaning methods used in the analysis. Section IV details the estimation strategies considered and used in the study.

Section V presents and interprets the results. Section VI offers concluding remarks and discusses limitations and further research opportunities.

## II. Literature Review

Any approach that examines TIF outcomes must first factor in TIF selection. Based on the criteria laid out in Georgia law, TADs are not randomly selected. On the contrary, the law stipulates that TADs cover areas that are underdeveloped and have a low likelihood of developing in the near future. On the other hand, there is evidence that TIF is susceptible to regulatory capture by developers hoping to receive subsidies for projects they would have built anyways (Weber 2003; Weber and Kohl 2013; Sroka 2017). Therefore, while it is likely that TIF is generally targeted at blighted areas, it unclear whether TIF projects attempt to meet or intentionally violate the “but for” test. Immergluck (2007) sums up the debate succinctly with the phrase “causation or capture.”

Most TIF studies erroneously ignore this endogeneity concern (Lester 2014). Others rely on strong assumptions to justify models that do not take selection bias into account. For example, Immergluck’s (2007) study of the Atlanta Beltline TAD, to my knowledge the only research on Atlanta before until this paper, argues that the mixed-use path follows historical rail lines that ostensibly have little connection to the city’s modern-day economic and demographic makeup. Another paper argues that controlling for community and time fixed effects should handle at least some sources of selection bias (Merriman, Skidmore, and Kashian 2011). Carrol’s (2008) study of Milwaukee uses an instrumental variable approach where downtown location is proposed as a valid instrument since the city is arguably suburban enough that location has little effect on real estate values. In general, instrumental variable approaches are rare since instruments are difficult to find and justify theoretically. This paper will confront endogeneity concerns directly, as theory indicates that TIF suffers from such significant selection bias that failure to take it into account may seriously bias results in either direction. There is certainly debate on this issue in the empirical literature, but many studies find evidence of selection issues (Greenbaum and Landers 2014).

There are two main methods used to counter selection bias in the TIF literature: Heckman selection and propensity scores (Greenbaum and Landers 2014). Heckman selection models attempt to formally model the selection of a treatment. The most common model is used when outcome data is only observed for treated individuals and treatment is not the primary regressor of interest. However, they can be adapted to other cases. Propensity scores are similarly derived from the factors that affect treatment and indicate an observation's "propensity" for treatment. They are frequently used to weight regressions or match treatment and control observations.

TIF research varies greatly in the outcomes of interest. Studies have examined everything from property values to unemployment rates to commercial activity, with property values being the most common. Property values are popular because they directly affect the tax increment through changes in property taxes, so they are the most relevant outcome for local governments. However, a full understanding of the effects of TIF requires examining its other effects on the community as well as nearby communities. Due to the complexity of modelling local economic development, studies tend to focus on a single area and one or two outcomes. This paper will focus on property values for the aforementioned reasons as well as the fact that projects like the Beltline have spurred a major political debate about housing values and gentrification in Atlanta.

Property value estimation strategies typically attempt to predict aggregate assessed property values based on TIF designation, area characteristics, and time dummies. For example, Dye and Merriman (2003) predict the Estimated Appraisal Value (EAV) of municipalities in

Illinois based on TIF status and other measures of TIF, tax rates, community-level variables, county dummies, and distance to Chicago. Studies have recently moved toward more fine-grained analysis of individual residential property values. Since assessments are often unreliable at the parcel level, some researchers prefer examining actual sales records. These studies rely on hedonic regression, which is based on the theory that the effect of TIF can be viewed in the revealed preferences of homebuyers through sales. In hedonic regression, the home is viewed as a bundle of its characteristics (de Haan and Diewert 2013). Sales price, often logged, is estimated as a function of the individual characteristics of the home and time dummy variables. Smith (2006) is representative of this approach and regresses property characteristics such as square foot, age, number of bedrooms, number of stories, and TIF status on the natural log of sales price per square foot. The model also includes community area and time dummies to control for neighborhood and time effects. Because only sales are observed, it is important to note that properties that were sold must be representative of all properties. Since TIF districts tend to encompass areas with more vacant properties, it is possible that hedonic regression selects more valuable properties in TIF districts than randomly selecting which homes are sold (Greenbaum and Landers 2014). This paper will not investigate the issue, but it is a caveat to the results of hedonic modelling.

A final discussion in the literature worth mentioning is how to define treatment. Most papers define treatment as inclusion in the TIF district since the district is intended to be the area primarily affected by the project. However, this does not account for differing levels of funding or project types between TIF districts and ignores the possibility of spillover effects. While some studies include measures of distance from TIF, TIF funding, and TIF age, this paper will rely on the simpler approach for ease of interpretability. There is also evidence suggesting that spillover effects are negligible (Weber, Bhatta, and Merriman 2007).

In this paper I will attempt to determine whether TADs in Atlanta have increased the residential property values of treated properties. I will compare naïve hedonic regression with Heckman selection and propensity score-based models to evaluate the importance of selection bias corrections. My results contribute to the existing literature on the effect of TIF on residential property values while producing novel information for Atlanta, which is understudied in the literature. Additionally, my selection corrections are more sophisticated than most TIF research, which rarely grapples with the issue at all. This information is intended to guide policymakers across the country and spur further research into treatment and outcome models of TIF.

## III. Data

Tax-parcel level assessment data was collected from the Fulton County Tax Assessors office from 2000 through 2018, the years for which digital records are available. However, observations before 2003 contained appraisals but not actual parcel characteristics, so the analysis is restricted to 2003 through 2018. Additionally, Fulton County includes cities like Alpharetta, which contain their own real estate markets. The data was therefore limited to parcels within the city of Atlanta. Tax parcels are segments of land that are assessed for property taxes. They generally align with a single structure, such as a home or store, but may contain multiple structures. The data is approximately panel data collected annually, though some parcels are merged or are created in the middle of the time period as city limits changed. To simplify merging data, I eliminated parcels that did not exist in 2018, which totaled roughly 5 percent of 2003 parcels. Parcel characteristics included address; municipal, class, and land use codes; number of living units; calculated acres of land; dummies for location, fronting, streets, topography, and parking measures; tax district; number of stories; type of exterior and style; date built; date remodeled; number of rooms, bedrooms, and bathrooms; dummies for plumbing, basement, heating, and attic; and square footage defined by living area and gross floor area. There were also various appraisal variables irrelevant to the study. TADs have their own tax districts, so they were identified through the tax district variable. There were over 15,000 total observations in 2018 and slightly less each year through 2003, which contained just under 13,000.

I matched the 2018 parcels with GIS data from the City of Atlanta's website. 99 percent of parcels were matched successfully. I then converted the parcel polygons to points and spatially joined them to GIS data on Atlanta Elementary Schools from the ESRI Atlanta office and neighborhoods from the Atlanta Regional Commission's Neighborhood Nexus. Figure 1 (see introduction) also made use of TAD GIS data from the City of Atlanta. These variables were added because neighborhood characteristics and school zone are important determinants of residential property values in Atlanta and are not captured by individual property characteristics.

Next, I joined this data with sales data from the Fulton County Tax Assessor's office, which is also structured by parcel for each year. Sales price is the only relevant added variable. 97 percent of sales were able to be matched to the relevant property's characteristics and the total number of observations fell from approximately 230,000 to 15,000. Although this step is necessary for hedonic regression, the downside is that the data was no longer panel data. Since parcels do not sell each year, no parcels are observed over the entire time period. As a result, the data is closer to pooled cross-sectional data, but the same parcels often show up twice in the dataset. In some cases, parcels were sold twice in the same year. In this situation, one sale was randomly removed. This cut observations by 6.5 percent.

After removing observations without appraisal data (3 percent) or with obviously incorrect sales prices (<1 percent), I downloaded census block data from Social Explorer from the 1990 and 2000 Census for my selection models. These dates were the closest to the actual TIF selection, and the 1990 data allowed me to examine trends. Although neighborhood-level data would be preferable since it most closely aligns with actual neighborhoods, the data only exists aggregated to Neighborhood Statistical Area level and does not go back to before TADs were implemented. Neighborhood statistical areas are similar to block groups but are designated by the City and are likely more accurate. Census block groups are the smallest unit for which the relevant census data is available. They encompass between 600 and 3,000 people and are generally intended to represent neighborhoods. Census block groups are frequently used in the TIF literature in both selection and outcome models. I was able to achieve a 100 percent match.

The Census data included percentage of residents who were white, unemployment rate, median household income, vacant housing units, total housing units, and median home value for 1990 and 2000. I calculated vacancy rate as vacant units divided by total units and calculated changes and percent changes from 1990 to 2000 for several variables. Unfortunately, 10 percent of observations were in block groups without median home value data and had to be dropped.

I then created or transformed more variables for my outcome model. I took the natural log of sales price and labelled the result *lPrice*. I created *TADin*, a dummy variable that indicates whether the parcel is in the treatment group. I also created *TADpost*, a dummy variable that is unity when the parcel is in a TAD in a year after the TAD was designated. It is equivalent to multiplying *TADin* by a dummy set to 1 after treatment. Interestingly, for 2017 parcels the average TAD start year was 2002 and the TAD with the most properties was the Beltline TAD. I turned the parking-, heating-, attic-, and basement-related variables into binary variables.

*Parking* signals whether there is adequate nearby parking. *Heating* indicates whether the home has a heating system. *Atticbin* is unity if there is an attic and *Bsmtbin* is unity if there is a full basement. Finally, I subtracted the year built from the year of the observation to calculate age.

## IV. Estimation Strategy

My estimation strategies are based on hedonic regression, a method that assumes that market prices reflect the combined value of the product's constituent elements. I first estimate an outcome model naïve to selection bias with linear regression. Next, I examine how the model results change as I incorporate three corrections for the selection of TADs. First, I employ a two-step Heckman selection procedure. Next, I calculate propensity scores for inverse probability of treatment weighting in my outcome regression. Finally, I use propensity scores to match treatment and control observations and run my regression on the matched dataset.

The dependent variable of the hedonic regression is *lPrice*, the natural log of the parcel's sales price. I chose this transformation because (1) sales price is strictly positive, (2) semi-log models allow for more useful interpretations of regression coefficients, and (3) inflation adjustments are not necessary. As a result, regression coefficients indicate the percent change in price resulting from a one unit change in the independent variable of interest. Instead of adjusting price for inflation, the coefficients of time dummies will include the effects of inflation on price in addition to changes in the housing market over time (Woolridge 2010, 452).

Figures 2 and 3 (below) show descriptive statistics for the relevant variables used in the regression with *Price* instead of *lPrice* for interpretability. As these tables show, roughly 14 percent of observations are in TADs and over 90 percent of these observations occur after the TAD was designated. As expected, treatment group parcels sell for substantially less (33 percent less) than control group parcels. The below statistics generally show that treatment and control parcels are not very comparable, indicating that these variables need to be controlled for to reduce omitted variable bias.

Figure 2: Summary Statistics for Parcels Outside of TADs

| Statistic | N | Mean | St. Dev. | Min | Pctl(25) | Pctl(75) | Max |
|---|---|---|---|---|---|---|---|
| Price | 91,811 | 323,796.600 | 440,305.300 | 5,000 | 113,000 | 387,500 | 47,370,000 |
| Taxyr | 91,811 | 2,010.247 | 4.884 | 2,003 | 2,006 | 2,015 | 2,018 |
| Calcacres | 91,811 | 0.217 | 0.340 | 0.000 | 0.031 | 0.249 | 16.200 |
| Stories | 91,811 | 1.251 | 0.443 | 1 | 1 | 1.5 | 4 |
| Age | 91,811 | 42.938 | 31.502 | 0 | 11 | 69 | 194 |
| Rmtot | 91,811 | 6.104 | 2.345 | 1 | 5 | 7 | 77 |
| Rmbed | 91,811 | 2.632 | 1.107 | 0 | 2 | 3 | 14 |
| Fixbath | 91,811 | 1.850 | 0.943 | 0 | 1 | 2 | 11 |
| Bsmt | 91,811 | 2.158 | 1.120 | 1 | 1 | 3 | 4 |
| Heat | 91,811 | 3.776 | 0.565 | 1 | 4 | 4 | 4 |
| Attic | 91,811 | 1.235 | 0.741 | 1 | 1 | 1 | 5 |
| Sfla | 91,811 | 1,718.927 | 1,089.833 | 0 | 1,056 | 1,996 | 19,740 |
| Parknear | 91,811 | 0.989 | 0.105 | 0 | 1 | 1 | 1 |
| Heating | 91,811 | 0.988 | 0.110 | 0 | 1 | 1 | 1 |
| Atticbin | 91,811 | 0.112 | 0.315 | 0 | 0 | 0 | 1 |
| Bsmtbin | 91,811 | 0.192 | 0.394 | 0 | 0 | 0 | 1 |

Figure 3: Summary Statistics for Parcels Within TADs

| Statistic | N | Mean | St. Dev. | Min | Pctl(25) | Pctl(75) | Max |
|---|---|---|---|---|---|---|---|
| Price | 13,110 | 216,410.000 | 160,040.800 | 5,000 | 116,966.2 | 290,000 | 6,946,760 |
| TADpost | 13,110 | 0.918 | 0.274 | 0 | 1 | 1 | 1 |
| TADpre | 13,110 | 0.082 | 0.274 | 0 | 0 | 0 | 1 |
| Taxyr | 13,110 | 2,010.704 | 4.601 | 2,003 | 2,007 | 2,015 | 2,018 |
| Calcacres | 13,110 | 0.099 | 0.184 | 0.000 | 0.022 | 0.135 | 9.900 |
| Stories | 13,110 | 1.402 | 0.511 | 1 | 1 | 2 | 4 |
| Age | 13,110 | 21.854 | 29.764 | 0 | 2 | 35 | 196 |
| Rmtot | 13,110 | 5.805 | 2.188 | 2 | 4 | 7 | 18 |
| Rmbed | 13,110 | 2.480 | 1.186 | 0 | 2 | 3 | 12 |
| Fixbath | 13,110 | 1.796 | 0.791 | 0 | 1 | 2 | 6 |
| Bsmt | 13,110 | 1.873 | 1.147 | 1 | 1 | 2 | 4 |
| Heat | 13,110 | 3.846 | 0.539 | 1 | 4 | 4 | 4 |
| Attic | 13,110 | 1.091 | 0.482 | 1 | 1 | 1 | 5 |
| Sfla | 13,110 | 1,516.127 | 715.918 | 0 | 1,000 | 1,880 | 7,749 |
| Parknear | 13,110 | 0.971 | 0.168 | 0 | 1 | 1 | 1 |
| Heating | 13,110 | 0.981 | 0.137 | 0 | 1 | 1 | 1 |
| Atticbin | 13,110 | 0.041 | 0.198 | 0 | 0 | 0 | 1 |
| Bsmtbin | 13,110 | 0.183 | 0.386 | 0 | 0 | 0 | 1 |

The regression equation is derived from the basic differences-in-differences (DID) equation. DID is a common method for evaluating the effects of policies that are not randomly selected. Because treatment is not randomly selected, there is no obvious measure of the counterfactual of the outcome of the treated group given they were not treated: $E[Y_0|D_i = 1, T = 2]$, where $Y_0$ indicates the outcome from not receiving treatment, $D_i$ indicates whether the observation was in the treatment group, and $T = 2$ signifies post-treatment period. Instead of measuring this directly, DID estimates the difference in outcomes in the pre- and post-treatment periods ($T = 1$ and $T = 2$, respectively) in the treatment group and subtracts the difference in outcomes in pre- and post-treatment periods from an untreated comparison group. In effect, the comparison group serves as a proxy for the missing control group. Although the treatment and comparison groups may differ in their actual outcomes, differencing should eliminate the differing intercepts of the groups' time trends and compare slopes alone.

DID relies on a key assumption that justifies substituting the comparison group as the control known as the parallel trends assumption. The assumption states that, without treatment, both groups' trend lines would share the same slope. This cannot be tested directly, but it is often justified through theory or showing that the trend lines look similar up until treatment and arguing that no non-treatment shocks that would affect one group more than the other occurred during the testing period. For this study, the parallel trends assumption requires arguing that growth in property values would be the same in TAD and non-TAD areas had TADs not been designated.

The simplified two-period DID model is as follows: $y = \beta_0 + \beta_1 T + \beta_2 D + \mathrm{B}_3(T * D) + \varepsilon$. It is easy to show that $\beta_3$ is the treatment effect defined above. My model simply adapts this equation to multiple treatment periods and adds neighborhood and school zone dummies in addition to the time dummies to account for the spatial elements of the data. I also added property characteristics to assist the parallel trends assumption and improve precision. The equation is $\mathrm{lPrice_{it}} = \beta_0 + \beta_{1_{it}} \boldsymbol{Taxyr} + \beta_{2i} TADin + \beta_{3it}(TADpost) + \beta_{4it}\boldsymbol{X} + \beta_{5i}\boldsymbol{N} + \beta_{6i}\boldsymbol{S} + \varepsilon_{it}$. Bolded terms are vectors of variables. ***Taxyr*** is a set of dichotomous year variables, not the integer used in the summary statistics. ***X*** includes the characteristics listed in Figures 2 and 3 below *Taxyr*. ***X*** includes both *Age* and $Age^2$ to account for diminishing marginal returns. ***N*** and ***S*** refer to neighborhood and school zone dummies, respectively. They do not vary with time along with *TADin*. There are 200 neighborhoods and 45 school codes in the data. Note that *TADpost* is the same as *TADin*****Taxyr***, so $\beta_3$ is the coefficient of interest.

The Heckman selection model utilizes the DID outcome model as well, but includes a first stage equation that models TAD selection. The first equation, a probit model, will produce an Inverse Mills Ratio (IMR) that will be added to the second stage. Though there is theoretical support for Heckman selection, it is often criticized in practice because it is highly sensitive to the functional form of the first stage equation and often requires an instrumental variable to deal with collinearity between the equations, and does not naturally produce (Puhani 2002).

The first stage equation is $TADin_i = Probit(PctWhite2000 + UnemploymentRate2000 + MedianHouseHoldIncome2000 + VacancyRate2000 + ChangeVacancyRate1990\_2000 + MedianHomeValue2000 + PctChangeHomeVal1990\_2000)$. The variables are self-explanatory. Because Atlanta lacks clear criteria to evaluate blight, I defined it as high unemployment, low median income, high and increasing vacancy rate, and low and decreasing median home value. The differenced variables are important because negative trends are evidence that the area meets the "but for" test. I also included a variable for race in case the City takes racial composition into account. From the results of the model, I simply estimated the IMR and added it as an additional regressor interacted with *TADin*. The typical Heckman model's variables of interest are ***X*** and the "treatment" is a decision that truncates observations, so treatment is not included in the final model. Since I wish to examine the treatment effect, I adjusted the model following Tucker (2007).

Propensity scores estimate the likelihood of receiving treatment based on characteristics that influence selection. The propensity scores were calculated the same way but with a logit model. Heckman selection requires a probit model, but logit models are equally valid for propensity scores (Caliendo and Kopeinig 2005). I used the same variables as the Heckman selection model to obtain the scores.

One use of the scores is as weights in the outcome model. The Inverse Probability of Treatment Weighting (IPTW) method constructs weights with the formula $w_i = \frac{1}{ps_i}$ for the treatment group and $w_i = \frac{1}{1-ps_i}$ for the control group (Austin 2011). The idea is that cases that are extremely likely to be treatment or control are weighted down, so observations that are closer to approximating random selection have more influence on the model. There is debate about the effectiveness about IPTW, with some researchers suggesting it can be worse than making no corrections in some cases (Freedman and Berk 2008). However, it seems to be a relatively well-accepted and common tool in economics.

In propensity score matching, treatment observations are matched with untreated observations with similar propensity scores for analysis. The matched data can be compared directly for a basic understanding of the treatment effect or imputed into an outcome model in place of the original data. I used nearest neighbor matching and set a caliper of .25 to eliminate distant matches that would make poor comparisons. A downside of matching with this mixed pooled cross section/panel data is the same parcel will be matched for multiple years, increasing the weight of those parcels in the final match. In fact, this concern may affect the other selection methods as well. Fortunately, many parcels sell once and parcels rarely sell more than twice, so this issue should be somewhat limited. Matching on parcels would be preferable to matching on observations, but that would necessitate drastically altering the data structure.

In total, 10,552 control observations were matched with treatment observations. The results of the propensity scores are represented in Figures 4 and 5. Figure 4 is composed of histograms showing the differing distributions of propensity scores between treatment and control observations before matching as well as the much more similar distributions after matching. The jitter plot in Figure 5 shows the results of the match more concisely. Overall, these figures indicate that, as long as the propensity score model is valid, there was significant selection bias before the correction that was improved significantly by matching. The Appendix contains the results of matching for each variable in the logit model to assess balance. The data indicates that there were major differences before matching and that every variable's differences fell substantially after matching.

Figure 4: Histogram of the Propensity Score Distribution of Treatment and Control Groups Before and After Matching

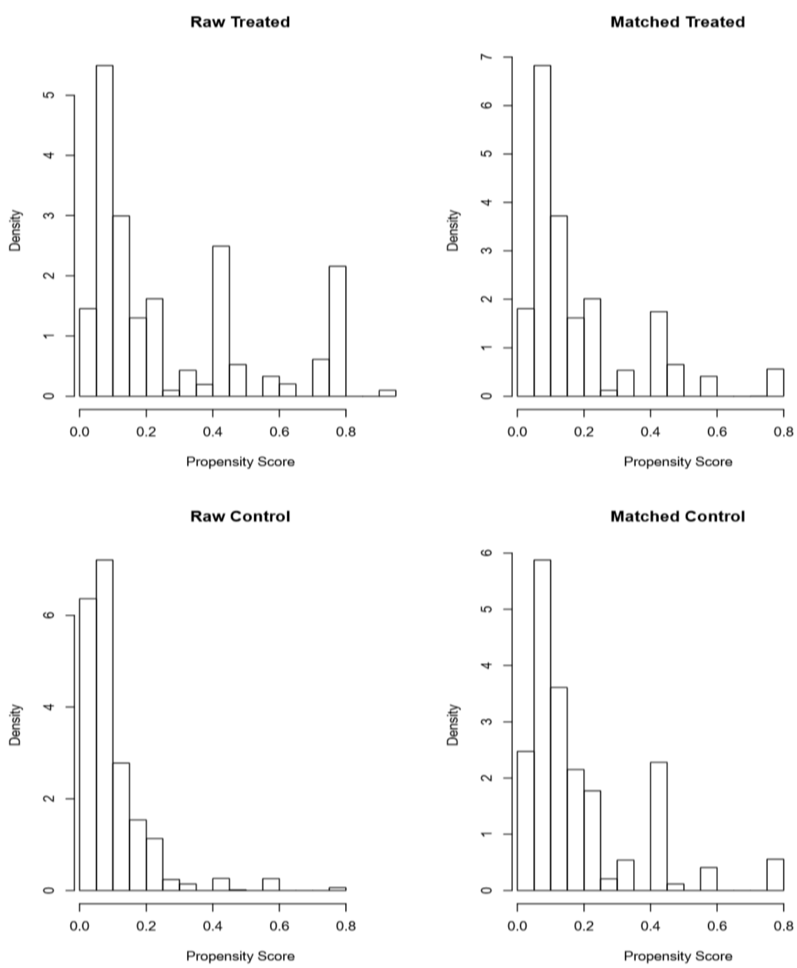
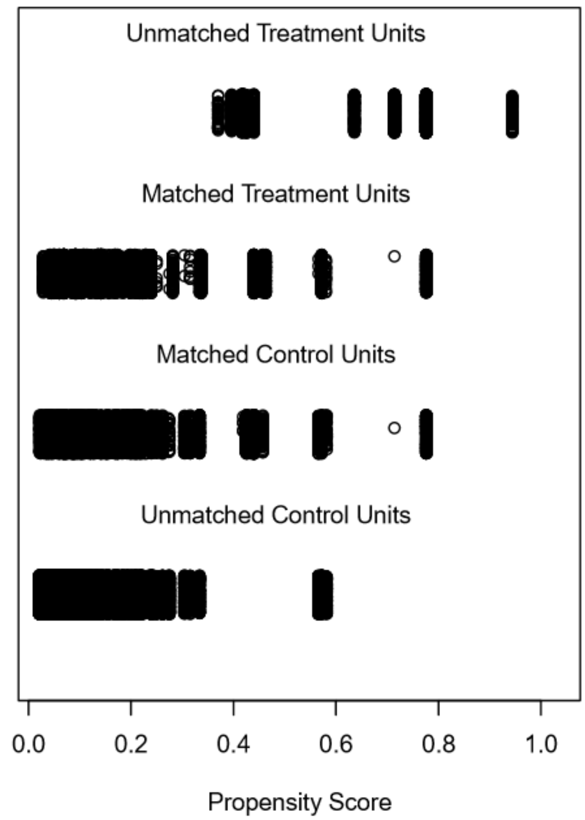


Figure 5: Jitter Plot of the Propensity Score Distribution of Treatment and Control Groups Before and After Matching




**V. Results**

**1. Outcome model**

The results of the model are below in Figure 6. Neighborhood, school district, exterior, style, and year variables were omitted from the table to make it a manageable length, but the variables were generally significant at the p<.01 level. Unsurprisingly, the coefficients on the year dummies became negative soon after the 2008 recession.

Figure 6: DID Regression Results Without Selection Corrections

| | Dependent variable: |
|---|---|
| | lPrice |
| TADin | -0.090*** (0.018) |
| TADpost | 0.142*** (0.019) |
| Calcacres | 0.168*** (0.008) |
| Sfla | 0.0002*** (0.00000) |
| Age | -0.012*** (0.0003) |
| I(Age2) | 0.0001*** (0.00000) |
| Rmtot | 0.005*** (0.001) |
| Rmbed | 0.024*** (0.003) |
| Fixbath | 0.115*** (0.004) |
| Bsmtbin | 0.038*** (0.005) |
| Atticbin | 0.054*** (0.006) |
| Parknear | -0.058*** (0.016) |
| Heating | 0.163*** (0.016) |
| Constant | 11.757*** (0.033) |
| Observations | 104,921 |
| R2 | 0.746 |
| Adjusted R2 | 0.746 |
| Residual Std. Error | 0.564 (df = 104643) |
| F Statistic | 1,111.658*** (df = 277; 104643) |
| Note: | *p<0.1; **p<0.05; ***p<0.01 |

The coefficient on *TADpost*, 0.142, indicates that a parcel becoming a member of a TAD increases its property value by 14 percent. This is consistent with the belief that TADs have a substantial positive effect on residential property values. Other variables generally behave as expected as well. The small, negative coefficient on *TADin* indicates that properties in TAD districts are disadvantaged relative to non-TAD properties. Increasing acres and square footage raise property value, and the small coefficient on square footage is a result of measuring the percent change in price from adding a single square foot. *Age* harms property values, but its effect lessens for each additional year. Rooms and amenities also increase property values. The coefficient on *Parknear* is puzzling, but it is not directly relevant to the analysis of the effects of TADs.

**2. Heckman Selection Model**

The Heckman Selection Model's results (Figure 7) are extremely similar to the DID model without the IMR. *TADin* is more negative, but *TADpost*'s estimated parameter, the Average Treatment Effect, has only changed by .001. The *IMR* term is small and not significant, offering evidence against significant selection bias. Note however that the standard errors of all of these models are likely biased since additional corrections are necessary when using these methods. Taking all of this into account, this model does not seem to change results enough to justify the additional assumptions of Heckman models or the different standard errors.

Figure 7: DID Regression Results with IMR

| | Dependent variable: |
| --- | --- |
| | lPrice |
| TADin | -0.118*** (0.025) |
| TADpost | 0.143*** (0.019) |
| Calcacres | 0.168*** (0.008) |
| Sfla | 0.0002*** (0.00000) |
| Age | -0.012*** (0.0003) |
| I(Age2) | 0.0001*** (0.00000) |
| Rmtot | 0.005*** (0.001) |
| Rmbed | 0.024*** (0.003) |
| Fixbath | 0.115*** (0.004) |
| Bsmtbin | 0.037*** (0.005) |
| Atticbin | 0.054*** (0.006) |
| Parknear | -0.059*** (0.016) |
| Heating | 0.163*** (0.016) |
| TADin:IMR1 | 0.018 (0.011) |
| Constant | 11.756*** (0.033) |
| Observations | 104,921 |
| R2 | 0.746 |
| Adjusted R2 | 0.746 |
| Residual Std. Error | 0.564 (df = 104642) |
| F Statistic | 1,107.686*** (df = 278; 104642) |
| Note: | *p<0.1; **p<0.05; ***p<0.01 |

### 3. IPTW Model

The results of the weighted model (Figure 8) are similar to the basic outcome model for many variables, but there is a large difference in TADin and TADpost. The signs are the same, but the IPTW model indicates that the treatment group faces significantly lower prices before treatment and that treatment causes a nearly 25 percent change in prices. These results indicate that the parcels that the propensity score model finds less extreme (and therefore more comparable) benefit far more from treatment than parcels in general.

A less relevant but interesting observation is that *Parknear* is no longer significant, so it is possible that the counterintuitive negative effect came from parcels that were extremely likely or unlikely to be selected into a TAD based on local demographic and economic characteristics. However, once again, the standard errors have a high risk of bias.

Figure 8: IPTW DID Regression Results

| | Dependent variable: |
| --- | --- |
| | lPrice |
| TADin | -0.259*** (0.009) |
| TADpost | 0.246*** (0.009) |
| Calcacres | 0.214*** (0.008) |
| Sfla | 0.0003*** (0.00000) |
| Age | -0.011*** (0.0003) |
| I(Age2) | 0.0001*** (0.00000) |
| Rmtot | 0.004** (0.001) |
| Rmbed | 0.021*** (0.003) |
| Fixbath | 0.125*** (0.004) |
| Bsmtbin | 0.047*** (0.005) |
| Atticbin | 0.040*** (0.007) |
| Parknear | -0.012 (0.014) |
| Heating | 0.173*** (0.015) |
| Constant | 11.787*** (0.027) |
| Observations | 104,921 |
| R2 | 0.733 |
| Adjusted R2 | 0.732 |
| Residual Std. Error | 0.738 (df = 104643) |
| F Statistic | 1,034.753*** (df = 277; 104643) |
| Note: | *p<0.1; **p<0.05; ***p<0.01 |

## 4. Propensity Matching Model

Figure 9 (below) lists the results of the outcome regression on the propensity score-matched data. The results are closest to the IPTW model, though the estimates for *TADin* and *TADpost* are slightly lower. The 23 percent estimated increase in sales price from TAD designation is still much larger than the basic model. All three of these selection bias corrections point toward the true treatment effect being at least as large as the estimated effect in the original DID model. This result supports the “causation” hypothesis over the “capture” hypothesis since TADs seem to improve residential property values and TAD selection seems more likely to favor blighted areas with increasing blight instead of areas that are developing quickly. In other words, there is stronger evidence that TADs truly improve property values than that they simply subsidize developers at the expense of taxpayers.

Figure 9: Propensity Score-Matched DID Regression Results

| | Dependent variable: |
|---|---|
| | lPrice |
| TADin | -0.207*** (0.023) |
| TADpost | 0.232*** (0.022) |
| Calcacres | 0.211*** (0.023) |
| Sfla | 0.0003*** (0.00001) |
| Age | -0.008*** (0.001) |
| I(Age2) | 0.0001*** (0.00001) |
| Rmtot | -0.002 (0.003) |
| Rmbed | 0.008 (0.007) |
| Fixbath | 0.144*** (0.009) |
| Bsmtbin | 0.058*** (0.012) |
| Atticbin | 0.038** (0.016) |
| Parknear | 0.022 (0.030) |
| Heating | 0.200*** (0.030) |
| Constant | 11.690*** (0.059) |
| Observations | 21,104 |
| R2 | 0.708 |
| Adjusted R2 | 0.704 |
| Residual Std. Error | 0.564 (df = 20838) |
| F Statistic | 190.347*** (df = 265; 20838) |
| Note: | *p<0.1; **p<0.05; ***p<0.01 |

## VI. Conclusion

TIF is an extremely popular local development tool, yet its effects are difficult to measure. As a result, many citizens are concerned that local governments advocate TIF’s use without a full understanding of how it will impact the target area. Academia has reached minimal consensus on major questions surrounding TIF, and many studies ignore selection bias. Clearly, there is a critical need for new research on TIF.

This paper contributes to this broader TIF discussion while supplying Atlanta with much-needed data analysis. As a city that has recently funded its signature projects through TADs, Atlanta requires high-quality research more than most municipal areas. Fortunately, this paper's outcome model, coupled with multiple models of selection, points toward a positive view of TIF's effect on residential property values. This paper estimates that TAD implementation increases treated property values by well over 10 percent regardless of the selection model.

While these results are based on relatively rigorous analysis, there are still clear limitations and opportunities for further research. First, TAD status was defined by official inclusion in the district and all TADs were treated equally. However, TAD funding levels, age, project makeup, and geographic range may significantly affect TAD effectiveness. Coupled with a sample of only ten TADs (compared to over a hundred in Chicago), these concerns may limit the general applicability of the analysis. For example, it is possible that the Beltline is an especially good infrastructure project and that most future TADs will be far less successful. At the very least, this paper suggests that Atlanta has implemented TADs well so far. Future studies could include these complex treatment variables in their analysis.

Another concern is that the selection models are overly simplistic or insufficiently tested for sensitivity. Data limitations and the lack of formal treatment criteria in Georgia were barriers in this analysis, but more variables can certainly be added to the analysis. However, although my selection models may not match the complexity of a study like Dye and Merriman (2003), it is still far ahead of most TIF studies and it incorporates the main components of blight.

This paper offers sensible results indicating that TIF is a successful tool even with these concerns. At the same time, the study demonstrates the challenges of satisfactorily modelling selection and operationalizing the treatment variable. Fortunately, the fact that every model points in the same direction instills a reasonable amount of confidence in the results. I still recommend caution and thorough cost-benefit analysis before TIF implementation, but my results unmistakably indicate that TIF expands the local tax base and benefits landowners in the district.

## VII. References

Austin, Peter C. “An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.” *Multivariate Behavioral Research* 46, no. 3 (May 2011): 399–424. https://doi.org/10.1080/00273171.2011.568786.

Bourdeaux, Carolyn, and John W Matthews. “Georgia’s Redevelopment Powers Law: A Policy Guide to the Evaluation and Use of Tax Allocation Districts,” n.d., 82.

Briffault, Richard. “The Most Popular Tool: Tax Increment Financing and the Political Economy of Local Government Symposium: Reassessing the State and Local Government Toolkit.” *University of Chicago Law Review* 77 (2010): 65–96.

Byrne, Paul F. “Strategic Interaction and the Adoption of Tax Increment Financing.” *Regional Science and Urban Economics* 35, no. 3 (May 1, 2005): 279–303. https://doi.org/10.1016/j.regsciurbeco.2004.04.002.

Caliendo, Marco, and Sabine Kopeinig. “Some Practical Guidance for the Implementation of Propensity Score Matching.” *Journal of Economic Surveys* 22, no. 1 (February 2008): 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x.

Carroll, Deborah A. “Tax Increment Financing and Property Value: An Examination of Business Property Using Panel Data.” *Urban Affairs Review* 43, no. 4 (March 2008): 520–52. https://doi.org/10.1177/1078087407308516.

Dye, Richard F., and David F. Merriman. “The Effects of Tax Increment Financing on Economic Development.” *Journal of Urban Economics* 47, no. 2 (March 2000): 306–28. https://doi.org/10.1006/juec.1999.2149.

Freedman, David A., and Richard A. Berk. “Weighting Regressions by Propensity Scores.” *Evaluation Review* 32, no. 4 (August 2008): 392–409. https://doi.org/10.1177/0193841X08317586.

Greenbaum, Robert T., and Jim Landers. “The TIFF over TIF: A Review of the Literature Examining the Effectiveness of the Tax Increment Financing.” *National Tax Journal* 67, no. 3 (September 2014): 655–74. https://doi.org/10.17310/ntj.2014.3.06.

Immergluck, Dan. “Large Redevelopment Initiatives, Housing Values and Gentrification: The Case of the Atlanta Beltline.” *Urban Studies* 46, no. 8 (July 2009): 1723–45. https://doi.org/10.1177/0042098009105500.

Invest Atlanta. “Tax Allocation District Financing.” Accessed April 2, 2019. https://www.investatlanta.com/developers/opportunities-incentives/tax-allocation-district-financing.

Lester, T. William. “Does Chicago’s Tax Increment Financing (TIF) Programme Pass the ‘But-for’ Test? Job Creation and Economic Development Impacts Using Time-Series Data.” *Urban Studies* 51, no. 4 (March 2014): 655–74. https://doi.org/10.1177/0042098013492228.

Merriman, David F., Mark L. Skidmore, and Russ D. Kashian. “Do Tax Increment Finance Districts Stimulate Growth in Real Estate Values?” *Real Estate Economics* 39, no. 2 (2011): 221–50. https://doi.org/10.1111/j.1540-6229.2010.00294.x.

OECD. *Handbook on Residential Property Price Indices*. Accessed April 2, 2019. https://read.oecd-ilibrary.org/economics/handbook-on-residential-property-price-indices_9789264197183-en.

Puhani, Patrick. "The Heckman Correction for Sample Selection and Its Critique." *Journal of Economic Surveys* 14, no. 1 (2000): 53–68. https://doi.org/10.1111/1467-6419.00104.

Smith, Brent C. "The Impact of Tax Increment Finance Districts on Localized Real Estate: Evidence from Chicago's Multifamily Markets." *Journal of Housing Economics* 15, no. 1 (March 1, 2006): 21–37. https://doi.org/10.1016/j.jhe.2006.02.003.

Sroka, Robert. "Pyrrhic Victory: Tax Increment Financing, 'But For,' and Developer Capture in the Dallas Arena District." Marquette Sports Law Review 28, no. 1 (January 1, 2017): 201.

Swenson, Charles W. "The Death of California Redevelopment Areas: Did the State Get It Right?" *Economic Development Quarterly* 29, no. 3 (August 1, 2015): 211–28. https://doi.org/10.1177/0891242414567944.

Tucker, Jennifer W. "Is Openness Penalized? Stock Returns around Earnings Warnings." *The Accounting Review* 82, no. 4 (2007): 1055–87.

Weber, Rachel. "Can Tax Increment Financing (TIF) Reverse Urban Decline?," 2003. https://www.jstor.org/stable/resrep18512.

Weber, Rachel, Saurav Dev Bhatta, and David Merriman. "Spillovers from Tax Increment Financing Districts: Implications for Housing Price Appreciation." *Regional Science and Urban Economics* 37, no. 2 (March 1, 2007): 259–81. https://doi.org/10.1016/j.regsciurbeco.2006.11.003.

Weber, Rachel, and Sara O'Neill-Kohl. "The Historical Roots of Tax Increment Financing, or How Real Estate Consultants Kept Urban Renewal Alive." *Economic Development Quarterly* 27, no. 3 (August 1, 2013): 193–207. https://doi.org/10.1177/0891242413487018.

Woolridge, Jeffrey. *Introductory Econometrics: A Modern Approach*. 5th ed. Accessed April 2, 2019. /c/introductory-econometrics-a-modern-approach-6e-wooldridge/9781111531041.

## VIII. Appendix

Figure 10: Balance Statistics for the Propensity Score Matching Algorithm

Summary of balance for matched data:

| | Means Treated | Means Control |
|---|---|---|
| distance | 0.1900 | 0.1843 |
| PctWhite2000 | 38.2144 | 40.6512 |
| UnemploymentRate2000 | 12.8347 | 13.4489 |
| MedianHouseHoldIncome2000 | 58878.9609 | 62452.6932 |
| VacancyRate2000 | 12.3719 | 12.7388 |
| ChangeVacancyRate1990_2000 | -2.7900 | -1.7107 |
| MedianHomeValue2000 | 227538.8193 | 243744.4632 |
| PctChangeHomeVal1990_2000 | 0.7430 | 0.5747 |

| | SD Control | Mean Diff |
|---|---|---|
| distance | 0.1673 | 0.0057 |
| PctWhite2000 | 37.3522 | -2.4368 |
| UnemploymentRate2000 | 13.2774 | -0.6142 |
| MedianHouseHoldIncome2000 | 43118.2034 | -3573.7324 |
| VacancyRate2000 | 6.4228 | -0.3668 |
| ChangeVacancyRate1990_2000 | 8.5919 | -1.0793 |
| MedianHomeValue2000 | 197266.3016 | -16205.6440 |
| PctChangeHomeVal1990_2000 | 0.9822 | 0.1683 |

| | eQQ Med | eQQ Mean |
|---|---|---|
| distance | 0.0070 | 0.0077 |
| PctWhite2000 | 2.2900 | 4.2896 |
| UnemploymentRate2000 | 1.1400 | 2.0586 |
| MedianHouseHoldIncome2000 | 3849.0000 | 6859.0233 |
| VacancyRate2000 | 0.8913 | 1.3089 |
| ChangeVacancyRate1990_2000 | 1.8748 | 2.2663 |
| MedianHomeValue2000 | 7994.0000 | 22095.0738 |
| PctChangeHomeVal1990_2000 | 0.0551 | 0.1940 |

| | eQQ Max |
|---|---|
| distance | 0.0322 |
| PctWhite2000 | 20.5900 |
| UnemploymentRate2000 | 15.5500 |
| MedianHouseHoldIncome2000 | 147662.0000 |
| VacancyRate2000 | 12.8408 |
| ChangeVacancyRate1990_2000 | 20.9247 |
| MedianHomeValue2000 | 568835.0000 |
| PctChangeHomeVal1990_2000 | 2.3394 |

Percent Balance Improvement:

| | Mean Diff. | eQQ Med | eQQ Mean |
|---|---|---|---|
| distance | 96.7448 | 91.1810 | 95.5989 |
| PctWhite2000 | 84.2526 | 78.4977 | 72.5064 |
| UnemploymentRate2000 | 94.1670 | 83.4543 | 80.4479 |
| MedianHouseHoldIncome2000 | 85.5921 | 81.6540 | 72.3780 |
| VacancyRate2000 | 89.9521 | 65.7832 | 65.3425 |
| ChangeVacancyRate1990_2000 | -75.7757 | -48.2585 | 0.7453 |
| MedianHomeValue2000 | 85.5392 | 92.7081 | 80.2869 |

```
PctChangeHomeVal1990_2000     64.5773  64.6679  59.6654

                         eQQ Max
distance                   94.4783
PctWhite2000                 53.0338
UnemploymentRate2000       67.5298
MedianHouseHoldIncome2000   0.0000
VacancyRate2000             59.3815
ChangeVacancyRate1990_2000 -2.5234
MedianHomeValue2000         0.0000
PctChangeHomeVal1990_2000  17.4580

Sample sizes:
          Control Treated
All         91811   13110
Matched     10552   10552
Unmatched   81259    2558
Discarded       0       0
```