UNIVERSITY OF CALIFORNIA

Los Angeles

Money for Nothing?

Opportunity Zones and Causal Inference

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Applied Science in Applied Statistics

by

Ryan Alexander Kupyn

2021

ABSTRACT OF THE THESIS

Money for Nothing?

Opportunity Zones and Causal Inference

by

Ryan Alexander Kupyn

Master of Applied Science in Applied Statistics

University of California, Los Angeles, 2021

Professor Chad J Hazlett, Chair

The Tax Cuts and Jobs Act of 2017 permitted US state governments to designate selected low-income census tracts as "Opportunity Zones." This designation permitted investors in projects located in these "Opportunity Zones" (OZs) to avoid or defer capital gains taxes on their investments. This provision was intended to increase the amount of investment in OZs, raising the incomes of households in designated census tracts. The processes of OZ designation was not uniformly transparent, with some indications areas with significant outside investments already in planned were more likely to receive OZ designations. This situation poses a challenge for traditional causal inference techniques, such as difference-in-differences. In this paper, an alternative set of assumptions are used to evaluate the effect of OZ designation on growth in median household income. These results suggest that the Opportunity

Zones program has had a positive effect on income growth in areas that received the Opportunity Zone designation, but highlight the significant uncertainty involved in such an estimate.

The thesis of Ryan Alexander Kupyn is approved.

Frederic R Paik Schoenberg

Hongquan Xu

Chad J Hazlett, Committee Chair

University of California, Los Angeles

2021

To my Katrina, my family, and everyone who has helped me along the way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to acknowledge the many people who have helped me in the development of this thesis, particularly my advisor, Professor Hazlett, and the other members of my committee, Professors Xu and Schoenberg. Their feedback and guidance has been invaluable throughout this process. I would also like to thank Elaine Reardon, Karyn Model, Ali Saad, Shui Tong Wong and Katrina Kaiser for their feedback and support.

1 Introduction

In 2017, the US Congress passed the Tax Cuts and Jobs Act (TCJA), which made broad changes to the US tax code. Among these changes was the creation of the "Opportunity Zones" program, which allowed governors to designate up to 25% of the low-income census tracts in their state as "Opportunity Zones" where investors could avoid paying capital gains taxes on new investments held for at least 5 years. This program was intended to boost incomes in the selected tracts by facilitating new external investments, which would then lead to more jobs and higher wages for workers in these areas.

This program was perceived by some outside commentators to be a "giveaway" to investors that would lead to few public benefits.[9] For the Opportunity Zones to be effective at raising incomes, the tax benefits must have been sufficient to stimulate new investment into designated low-income areas.

For researchers and policymakers, evaluating whether Opportunity Zones increased wages is a useful goal - but the structure and execution of the program makes this challenging. Individual governors had significant discretion over which areas were selected as Opportunity Zones, and there is evidence to suggest that they preferentially selected areas that had significant investments planned.[16] For the governors deciding which areas in their jurisdiction would become Opportunity Zones, this was understandable behavior as the Opportunity Zone designation does not directly affect state tax revenue, and investors with funds already committed to specific areas have an incentive to lobby for an Opportunity Zone designation.

There are many examples of governors using their own discretion to steer the designation of Opportunity Zones towards areas with large investments already planned.

In Maryland, for instance, a relatively wealthy area was designated as an Opportunity Zone after lobbying by a politically connected businessman who has pre-existing plans to redevelop land he owned there.[7] Elsewhere, researchers and journalists have documented more cases where Opportunity Zone designations were steered towards areas where investments were already planned.[12]

Under the putative causal pathway presented in the paper, increases in investment stimulated by a census tract's designation as an Opportunity Zone raise incomes for households in the area. If the areas designated as Opportunity Zones would have received higher levels of investment even without the designation, it would imply that these areas would have had higher income growth even if the Opportunity Zones program had never been implemented.

In addition, the deliberate selection of areas with above-average amounts of future investment planned will directly confound the most common causal inference methodologies that could be used to examine this problem. In an attempted difference-in-differences analysis, for instance, disparities in the amount of investment planned in designated and non-designated areas prior to the designation process would violate the parallel trends assumption, and lead to an overestimate of the impact of the Opportunity Zones program on income growth.

Previous researchers have used difference-in-differences to analyze the effects of Opportunity Zones on housing prices[15], but have not grappled with the more challenging question of whether the assumptions used are appropriate.

This paper builds on this analysis, describes the ways that traditional methods of causal inference are inappropriate for the purposes of evaluating Opportunity Zones program, and presents an alternative set of assumptions that overcomes some of their shortcomings in this context. Specifically, this paper explores the implications using assorted causal inference techniques in conditions where we have only limited information on treatment assignment, and where we cannot assume that treatment assignment is unrelated to an experimental unit's counterfactual outcome.

The causal inference methodologies examined here are summarized in the table below, which lists the assumptions, advantages, and disadvantages of each one.

Methodology	Assumptions	Advantages	Disadvantages
Naive Comparison of Means	Random assignment	Ease of calculation	Limited utility when analyzing observational data
Regression/Selection on Observables	Random assignment conditional on observables	Ease interpretation, ease of communication of results	Assumption that treatment assignment is random conditional on observed variables is frequently optimistic
Difference-in-Differences	Parallel trends	Provides straightforward estimate of causal effect	Parallel trends assumption can be difficult to justify
Stability Controlled Quasi-Experiment	Counterfactual delta between groups	Less rigid assumption compared to diff-in-diff may be more realistic in practice	Does not provide point estimate of causal effect
Regression Discontinuity	Identical characteristics of experimental units around discontinuity	Easy to implement assuming that placement above/below cutoff is random	Provides useful estimates of causal effect only in the vicinity of discountinuity

Of the five methods that I present here, four have assumptions that cannot be satisfied with the available data. Because of the possibility that areas with higher levels of investment already planned were more likely to be designated as Opportunity Zones, a simple difference-in-differences analysis is unsuitable. OZ designations are not plausibly random conditional on available covariates, which limits the utility of naive comparison of means and selection on observables. In addition, the details of the structure of the Opportunity Zones program limit the utility regression discontinuity, as very few census tracts are designated as Opportunity Zones near the eligibility cutoff.

However, the Stability Controlled Quasi-Experiment offers a way to more pragmatically assess the potential effect of Opportunity Zones on income growth, substituting the point estimates of causal effect used by other techniques for partial identification approach.

2 Analysis

2.1 Data Description

To identify which Census Tracts were designated as Opportunity Zones, as well as tracts that were eligible but not designated, data originally compiled by the Urban Institute[13] for their analysis of the tract-designation process is used. The Urban Institute's analysis covered the process of tract designation, and mirrors many of the summary statistics presented below. For instance, the Urban Institute found that poorer census tracts were more likely to be designated as Opportunity Zones than wealthier ones, and that urban areas were more likely to be designated than rural areas with similar income levels. This data covers 42,176 census tracts eligible for designation as Opportunity Zones, of which 8,762 were actually designated. A regression discontinuity analysis presented later uses the full national population of census tracts (subsequently limited to tracts near the eligibility threshold), for a total of 72,877 census tracts.

This paper also incorporates data drawn from the American Community Survey, which provides tract-by-tract data on demographic and economic variables. This data is drawn from the US Census' 5-year estimates at the census tract level.

Several tract-by-tract covariates covering various indicators of economic deprivation, race/age demographics, location, and education levels were generated as part of the data construction process. Specifically, covariates are drawn for civilian unemployment rate,[5] percent of population below 150% of the federal poverty line,[6] percent of the population below the federal poverty line,[2] median income,[4] income disparity - defined as the log of 100 $\times$ ratio of the number of households with annual income less than \$15,000 to the number of households with annual income greater than \$75,000,[3] percent of the population classified as nonwhite[5] and percent of the population with a bachelor's degree.[1]

2.2 Exploratory Data Analysis and Naive Estimate of Growth Rate

The primary analyses of this paper require only three variables – median household income by tract in 2016 and 2019 and an indicator for whether a census tract is designated as an Opportunity Zone (as opposed to tracts that are eligible for the program but not designated). 2016 was chosen as a baseline for initial income levels because it is the last year to the Opportunity Zone designation process beginning, while 2019 was the most recent year with tract-level income data available. Nominal income values are log-transformed for analyses, and the distribution of untransformed income values is presented in Tables 1 and 2.

These tables show clear differences in income between treated and untreated tracts, with tracts designated as Opportunity Zones having lower incomes than those that were eligible but not selected - as well the presence of some outlier tracts with extremely high and low incomes among both the treated and untreated population of census tracts.

While the presence of tracts with extremely low incomes is not surprising, the

Table 1: Distribution of Household Median Income by Census Tract in 2016

	Minimum	25th Percentile	Median	75th Percentile	Maximum
Designated	2,499	31,631	41,014	51,000	242,292
Not Designated	4,310	41,260	51,563	62,083	250,001
Overall	2,499	39,080	49,788	60,521	250,001

Table 2: Distribution of Household Median Income by Census Tract in 2019

	Minimum	25th Percentile	Median	75th Percentile	Maximum
Designated	2,499	36,784	47,816	59,944	228,804
Not Designated	2,499	47,109	58,889	71,851	250,001
Overall	2,499	44,698	56,772	70,069	250,001

presence of tracts with extremely high incomes is unexpected. This is the result of apparent sampling issues and only affects a small number of tracts (0.3% of eligible tracts have a median income above $150,000 in 2016 and 0.1% of tracts have income above the same threshold in 2019). The distribution of income in the original data is also censored at the $2,499 and $250,001 median income levels. Since the number of tracts affected by censoring is small, the effects of this are not considered here.

Results presented in this paper are in the form of changes in compound annual growth rate (CAGR) - i.e., an effect size of 1.0% implies that the Opportunity Zone designation increases incomes by 1% per year relative to an equivalent non-designated census tract.

Figure 1 provides another view into the pre-treatment income distribution for eligible census tracts. This figure presents a histogram of pre-treatment median income for these census tracts, identifying the selected census tracts and those that were eligible but not selected. As this figure shows, both designated and non-designated census tracts exist at all income levels, but the proportion of tracts receiving the Opportunity Zone designation is greater at the lower end of the distribution of pre-treatment income.

Figure 1: Distribution of Pre-Treatment Income For Eligible Census Tracts

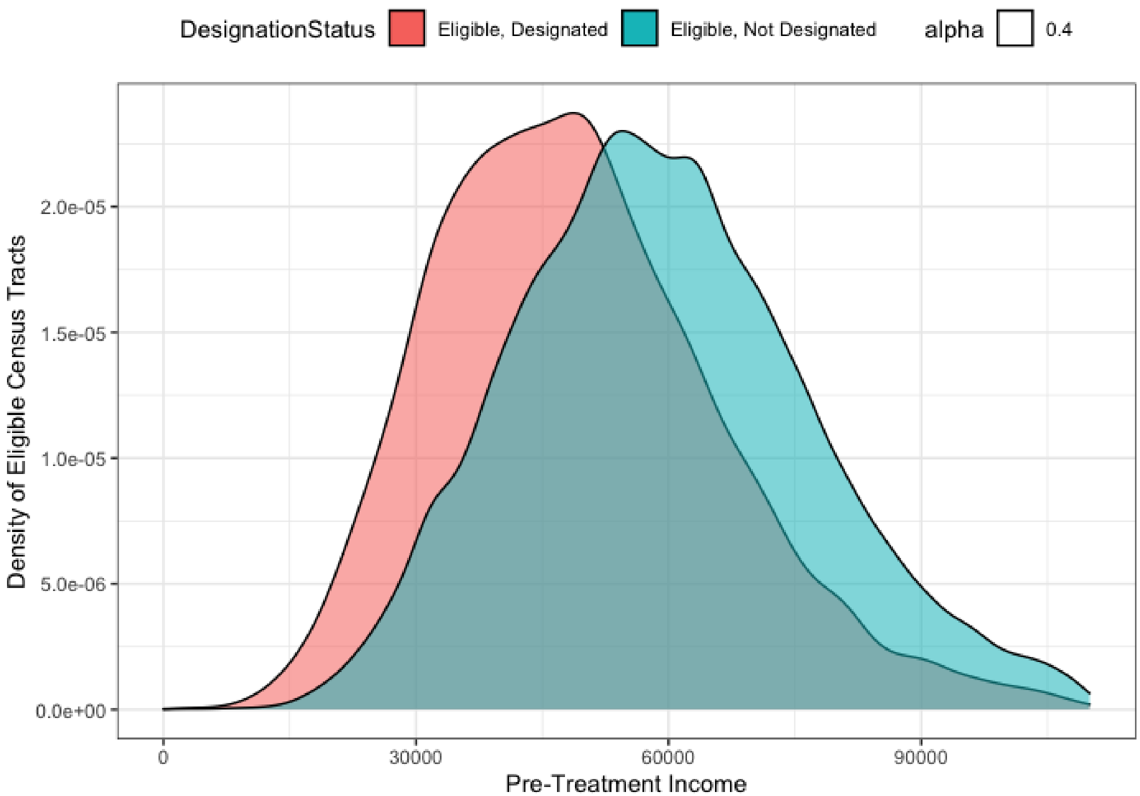
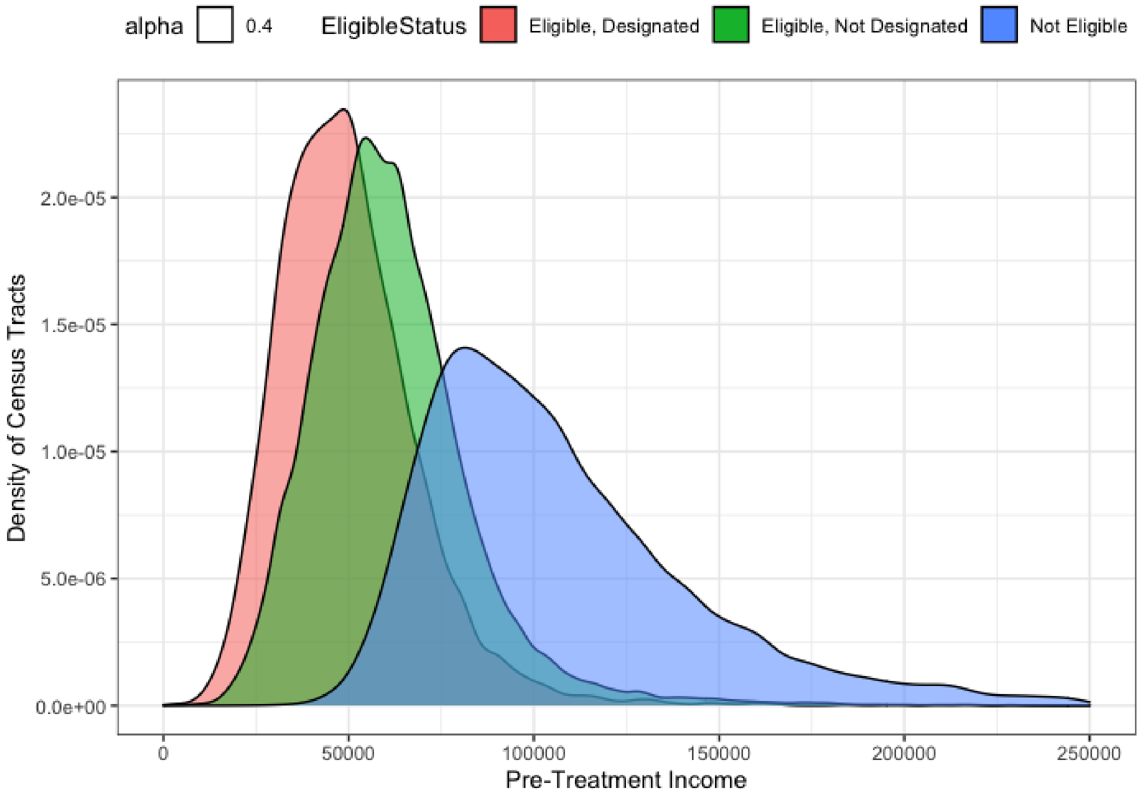


Figure 2 expands the population presented in Figure 1 to include the pre-treatment income distribution for all tracts, including those that were not eligible for designation as Opportunity Zones. Expanding the distribution in Figure 1 to include all census tracts shows the broader range of pre-treatment income. Because the threshold for eligibility varies by state, many income strata have both eligible and ineligible census tracts.

Figure 3 presents the distribution of average annual income growth for eligible census tracts during the analytical period from 2016-2019. There are no visually apparent differences in the distribution of income growth rates between treated and

Figure 2: Distribution of Pre-Treatment Income For All Census Tracts



untreated areas.

A naive comparison of the mean growth rate for treated and eligible-but-untreated tracts presented in Table 2 provides useful context for the interpretation of later analyses, though it cannot plausibly be used as an estimate of the causal effect itself. Without any controls, areas designated as Opportunity Zones had an annual growth rate six-tenths of a percentage point higher than areas that were eligible to be designated as Opportunity Zones but did not receive the designation. This estimate, though very precise and easy to compute and communicate, is clearly limited in its utility given that the balance statistics presented in this paper clearly indicate that no random assignment occurred in the designation process.

Figure 3: Distribution of Average Annual Income Growth From 2016-2019

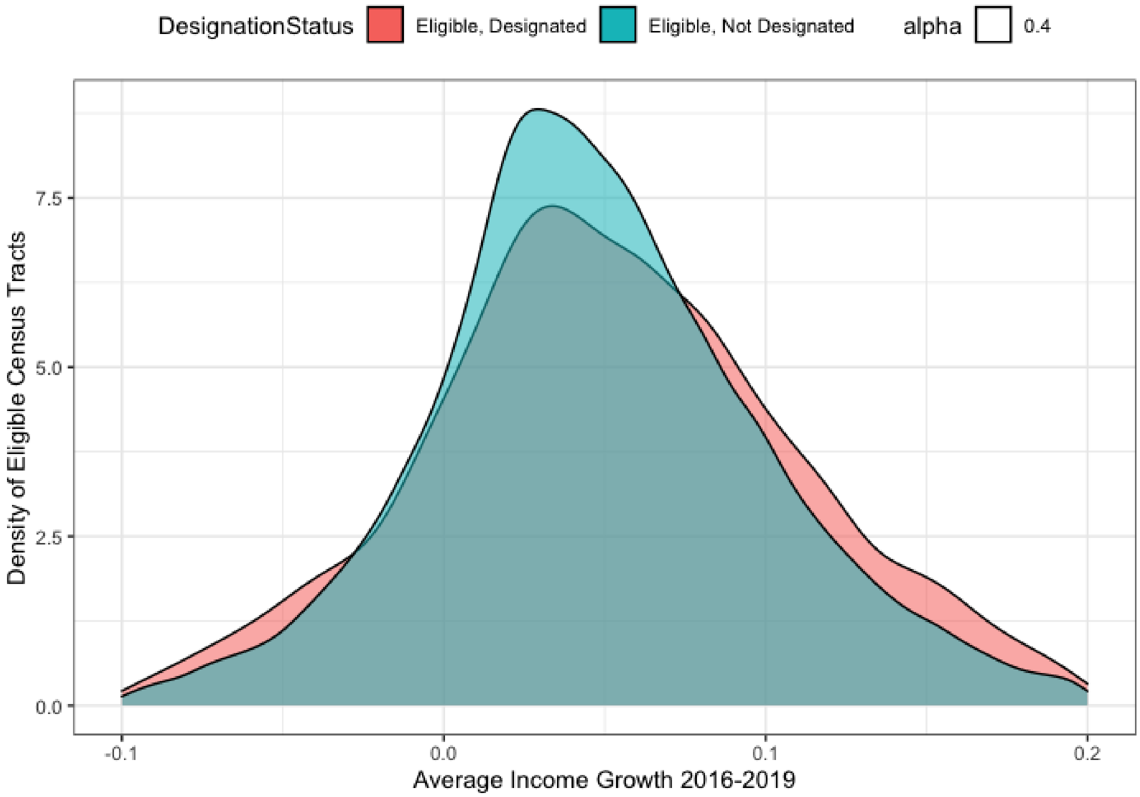


Table 4 presents balance statistics for census tracts census tracts in the regression population, comparing areas designated as Opportunity Zones to tracts that were eligible for the designation but did not receive it.

Within the set of eligible tracts, tracts selected as Opportunity Zones are significantly poorer and have a smaller proportion of their population with a college degree than those that were eligible but not designated. This alone is cause for concern when

Table 3: Naive Comparison of Growth Rates

Average Growth Rate For Designated Tracts	Average Growth Rate For Eligible But non-designated Tracts
5.6%	5.0%

Table 4: Balance Statistics

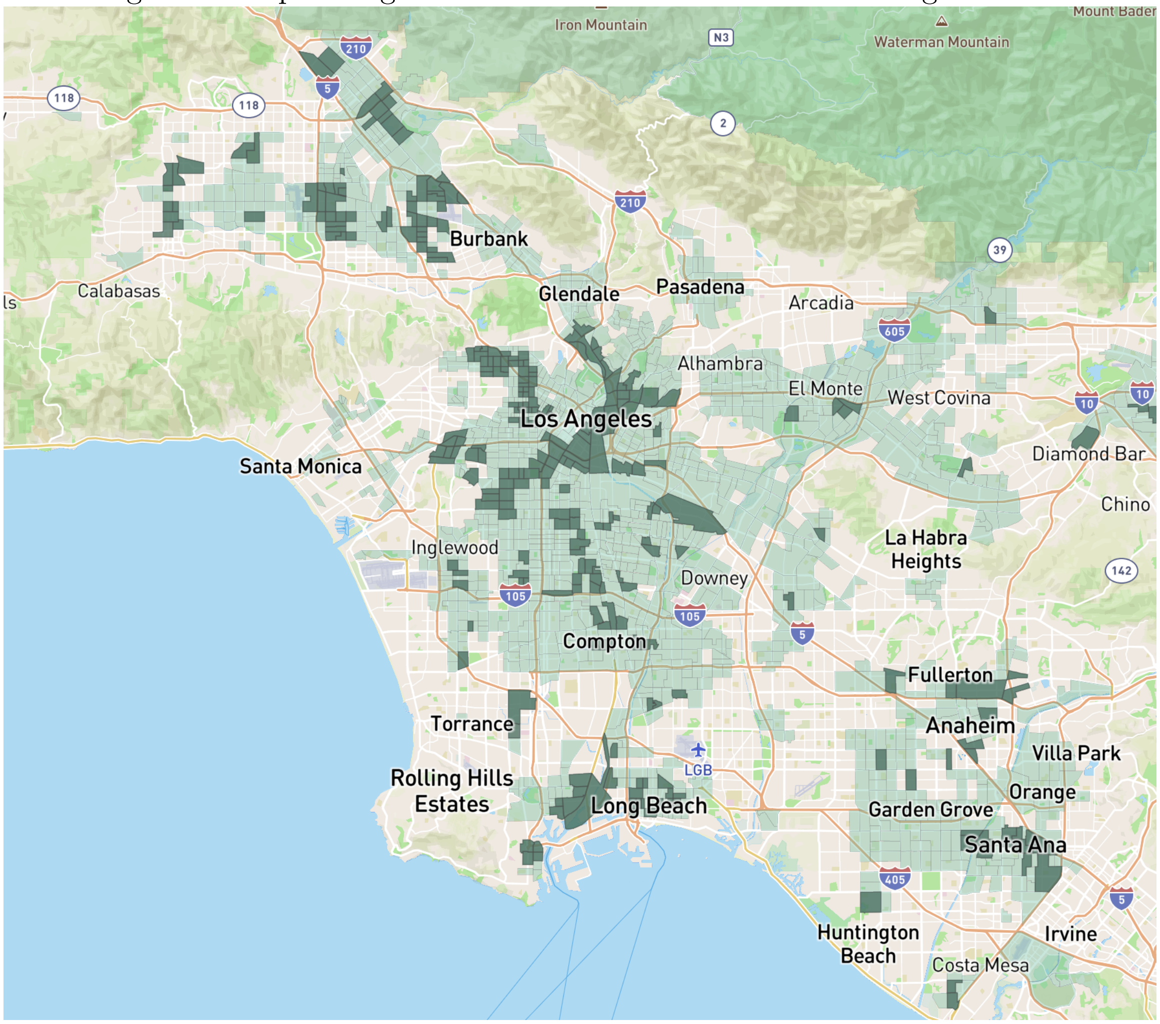
Variable	Mean (Treatment)	Mean (Control)	T-Statistic	P-Value
Percent White	57.5%	68.3%	-28.94	<0.0001
Percent With BA	11.1%	13.6%	-27.58	<0.0001
Unemployment Rate	12.4%	9.3%	36.51	<0.0001
Percent Below Poverty Level	24.8%	16.6%	49.91	<0.0001
Income Disparity	0.799	0.654	62.35	<0.0001

attempting to make assertions about causal inference using the simplest techniques - it is clear, at least, that the Opportunity Zone designations were not handed out at random conditional on the available covariates.

Figure 4 shows a map of census tracts eligible to be designated as Opportunity Zones in Los Angeles, with tracts that were eligible but not selected in light green and selected tracts in darker green. Areas that were not eligible for the program are left unshaded. This map shows the small areas covered by individual census tracts, which may influence the dispersion of benefits, with some designated census tracts very near to tracts that were eligible but not designated or ineligible to be designated as Opportunity Zones.

The relative proximity of selected and unselected areas is potentially significant when considering the practical implications of these results. Within a relatively small area such as Los Angles metro, with fairly easy access to Opportunity Zones from neighboring non-designated areas, it is possible that geographically constrained tax incentives do not increase aggregate investment so much as they draw investment from one area to another. This may make it harder to discern a causal effect, as investments in eligible census tracts lead to the employment of individuals outside the census tracts as well.

Figure 4: Map of Eligible and Selected Tracts in the Los Angeles Area



2.3 Causal Inference Methodologies

For this analysis, I review a series of causal inference tools and their applicability to this problem, and the describe ways that details of the implementation of the Opportunity Zone program limit their ability to accurately determine the causal effect of the program on income growth in designated areas. As discussed earlier, the selection process for Opportunity Zones is decidedly nonrandom, and may have frequently led to patterns in the designation of Opportunity Zones that are correlated with counterfactual income growth. Despite this, alternative causal inference methodologies still provide a useful baseline against which to compare results.

2.3.1 Regression

A simple regression analysis would be sufficient to estimate the causal impact of the Opportunity Zones program if one were able to assume that Opportunity Zones were selected at random from the pool of eligible census tracts, or if random assignment could be assumed conditional on the available covariates. If that were the case, a straightforward regression incorporating all relevant non-treatment covariates would be sufficient to infer a causal effect.

However, the nature of the program precludes making these assumptions. Because Opportunity Zones are designated at the discretion of each state's governor, we cannot assume that the selection process is random even after controlling for the additional factors available - and thus cannot draw useful inferences on potential causal effects of Opportunity Zone designation on income growth. However, these analyses still provide a meaningful comparison with other methods.

In the regression analysis, three regressions are conducted in order to quickly investigate the overall structure and relationships present in the data. These regressions vary only in the covariates they include - the first one is designed for purely exploratory purposes and includes no covariates, while the second includes covariates that might plausibly affect income growth, such as pre-treatment demographics (race, age, and education levels), location (the state and county the census tract is located in, as well as whether the census tract is in a metropolitan or micropolitan area) and measures of economic deprivation prior to treatment (the percentage of the population that was unemployed or rent burdened in 2016).

A detailed investigation of the relationship between the response variable and the covariates reveals three variables (poverty rate, unemployment rate, and percentage of the population with a BA or higher) that have a nonlinear relationship with income growth. The third regression accounts for this, by log-transforming these three variables. This third model can be compared the second, which is specified identically but without the log-transformation. This comparison reveals that the model which includes the log-transformation does better job explaining variation in income growth between census tracts, but leads to only a negligible change in the estimate of the causal effect. An investigation of further potentially nonlinear relationships between income growth and the available covariates reveal none that notably increase the explanatory power of the model.

Table 5 shows the estimated effect size for the initial linear regressions, both including and excluding covariates and with the covariates with an apparently nonlinear relationship with income growth transformed.

Table 5: OLS Regression Results

Model	Effect Size	T-Statistic	P-value	R^2
1. Naive Regression, No Covariates	0.7%	8.73	<0.001	0.0019
2. Expanded Regression	-0.01%	-0.10	0.920	0.08
3. Expanded Regression With Transformation of Selected Variables	0.00%	0.04	0.99	0.11

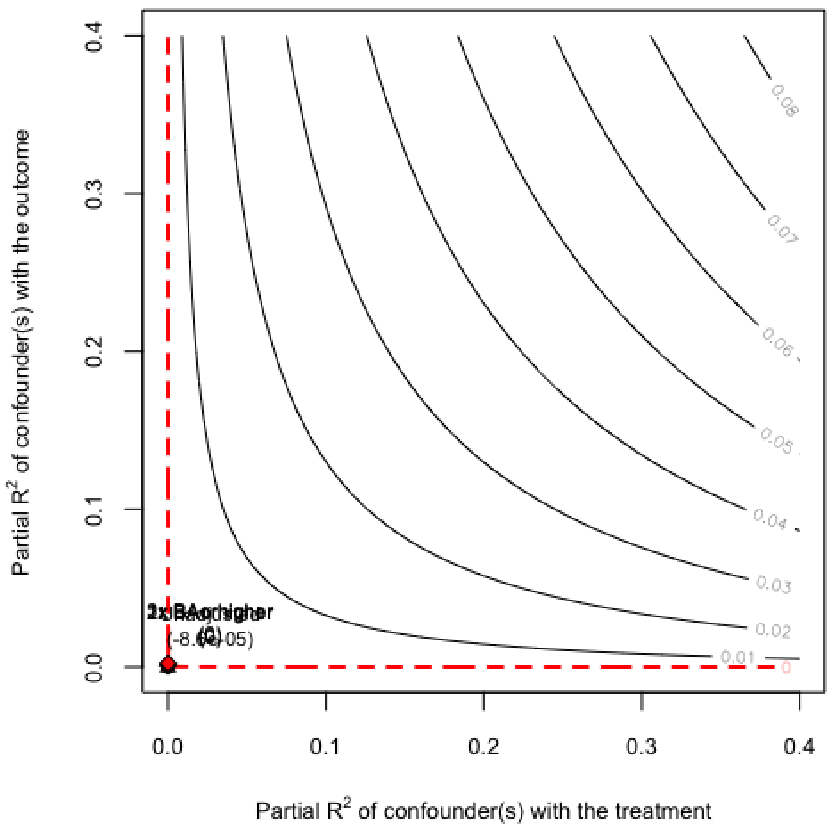
Taken at face value, these results suggest that designation as an Opportunity Zone has no statistically significant effect on the median income growth rate. The first regression - a naive one with no covariates - indicates that designation as an Opportunity Zone boosted the growth rate of median household income by 0.7 percentage points, while the incorporation of relevant covariates suggests no noticeable effect of Opportunity Zone designation at all.

For purposes of determining the practical significance of the observed effect sizes, it is possible to convert these changes in income growth into changes in annual dollar income for the median household in the median census tract. In this case, the 0.7% increase in income suggests a fairly substantial practical effect. For the median household in the median designated census tract, an increase in income growth rate of this size is the equivalent of an income that is $835 higher than it otherwise would have been in 2019 had their tract not been designated as an Opportunity Zone.

These regressions explain only a small portion of the variation in income growth between different census tracts. This is concerning - it suggests that the small effect size observed could easily be driven by unobserved confounding variables. I conduct a sensitivity analysis to evaluate this possibility. In this sensitivity analysis, I determine the sensitivity of these results to a hypothetical confounding effect across a range of partial R^2 values between the treatment and outcome variables. In Figure 5, the red line indicates the range of combinations of partial R^2 between the treatment and outcome of a hypothetical confounding variable that would be necessary to completely eliminate the observed treatment effect.

This sensitivity analysis shows that these regression results are highly sensitive to potential confounding effects. As Figure 5 (corresponding with the expanded Regression 2) indicates, a confounding effect with a partial R^2 of less than 1% with the outcome and treatment would completely switch the sign of the supposed effect. In addition, even the analysis with the added covariates has an R^2 of only .08, which suggests that these models do a generally poor job of explaining variation in growth rates between census tracts over the relevant analytical period. This poor showing

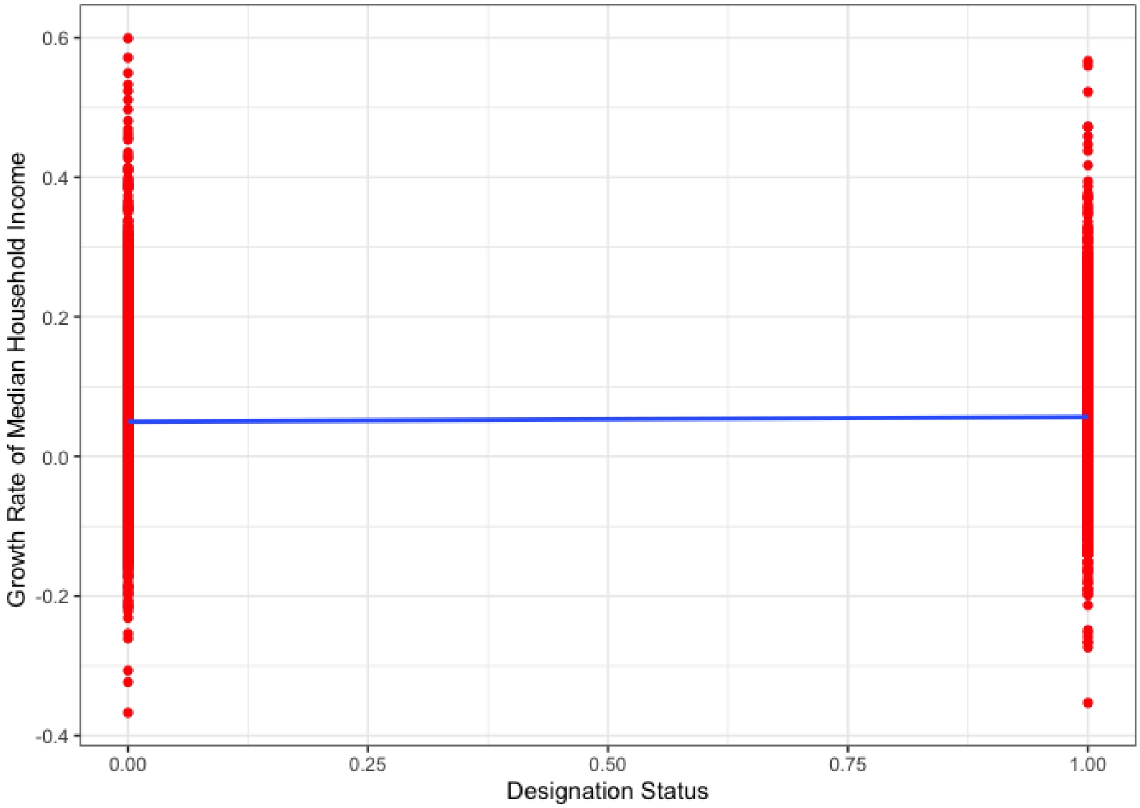
Figure 5: Contour Plot Demonstrating Sensitivity of Results



increases the risk unobserved confounding effects are present, and should further limit the inferences we can make from this result. The sensitivity of these results to unobserved confounding variables underscores the risk of overly aggressive interpretation of these regression results and highlights the importance of alternative methodologies.

Despite the conceptual problems inherent in using a regression to evaluate the OZ program, where treatment assignment is complex and uncertain, examining the regression diagnostics used here reveals valuable features of the data. A simple plot of the relationship between designation status and income growth (Figure 6) underscores the wide variation in growth rates between census tracts, and shows the need for a more complex model than what a simple uncontrolled regression can provide.

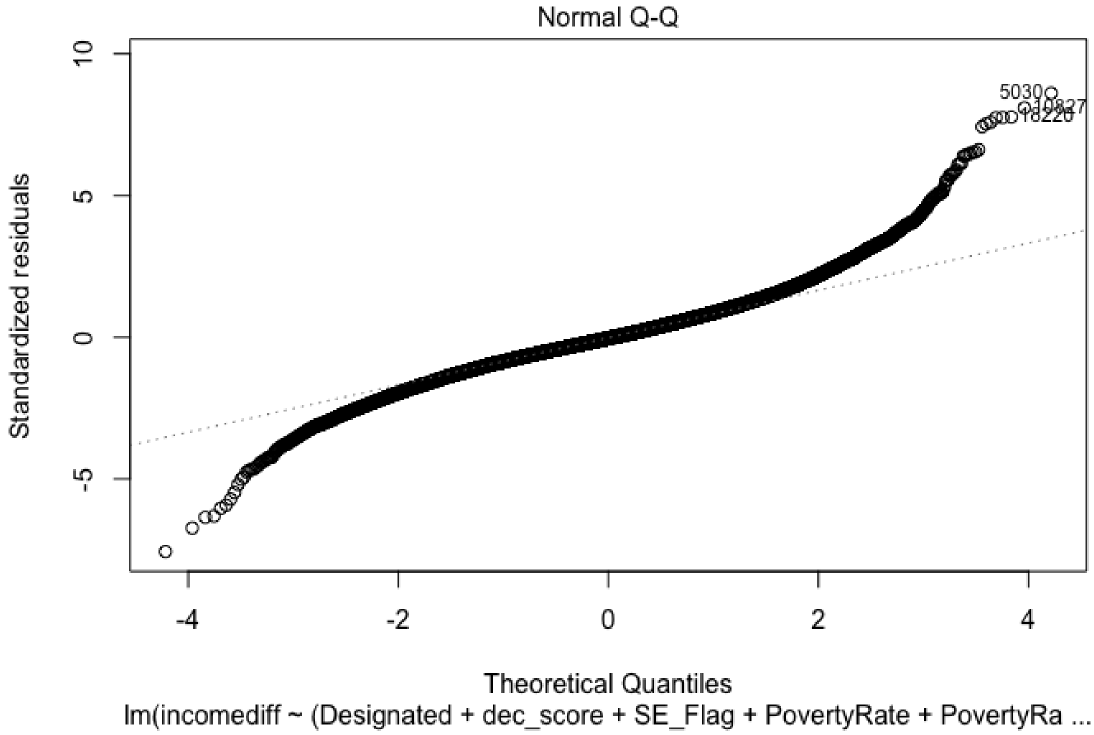
Figure 6: OZ Designation vs Growth Rate of Median Household Income



A Q-Q plot of the residuals of model 3 (all covariates with applicable log-transformations) shows that the residuals of our regression are heavy-tailed (Figure 7). This plot explicitly identifies three census tracts that are exceptional outliers, all with far higher income growth than the fitted model predicts.

These three census tracts (labeled by the Census as tracts 39061026300, 39049001600, and 39035104200) illustrate a critical assumption implicit in all of these analyses, and provide a useful example why the estimates of the causal effect of Opportunity Zone designation on growth in median household income presented in this paper must be interpreted with care. These census tracts are all located in the state of Ohio, and

Figure 7: Q-Q Plot



are all designated as Opportunity Zones. Census tract 39061026300 is a formerly industrial neighborhood near downtown Cincinnati, with the others similarly situated near the downtown areas of Columbus and Cleveland.

These census tracts are all areas that had low incomes and small populations at the beginning of the analytical period in 2016, but experienced large increases in both population and median household income by the end of the period in 2019. In tract 39061026300, for instance, the population rose from 147 to 1,580, and median household income rose from \$4,706 to \$18,088. These rapid changes appeared to be spurred by new residential development in areas previously dedicated to industrial use, leading to inflows of new residents from outside the census tract. [11] This suggests that, at least in these areas, increases in median household income are potentially driven by changes in the composition of the population in the relevant census tracts.

Figure 8: Residuals Versus Fitted Values

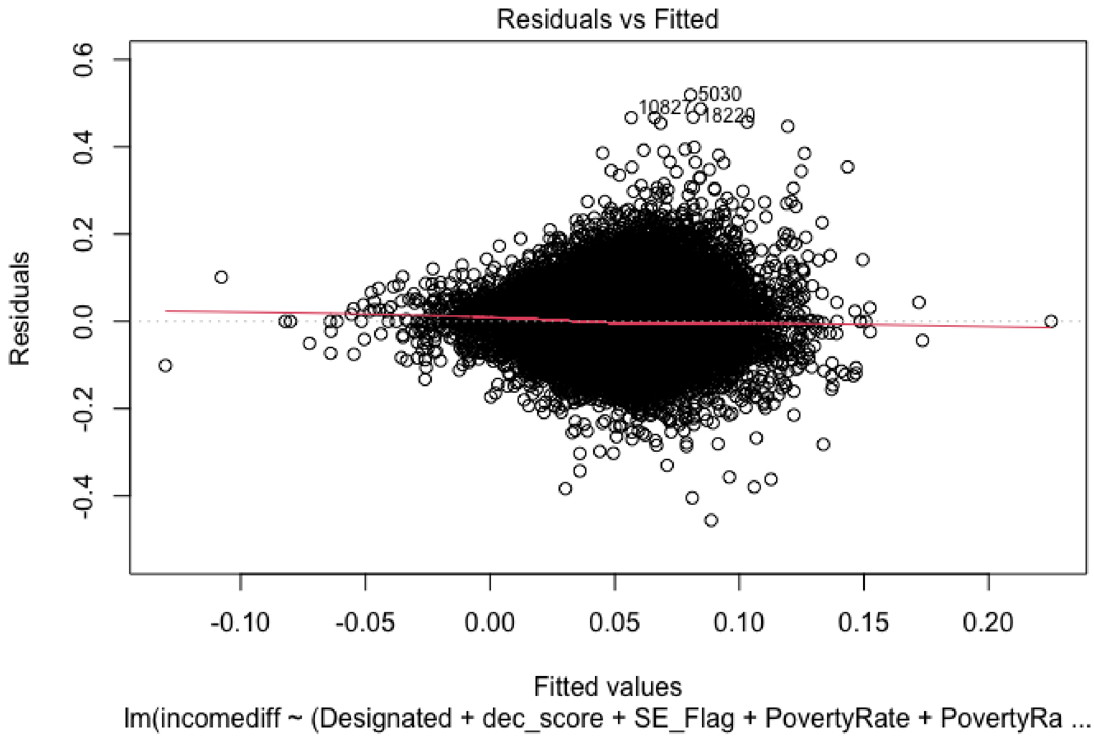
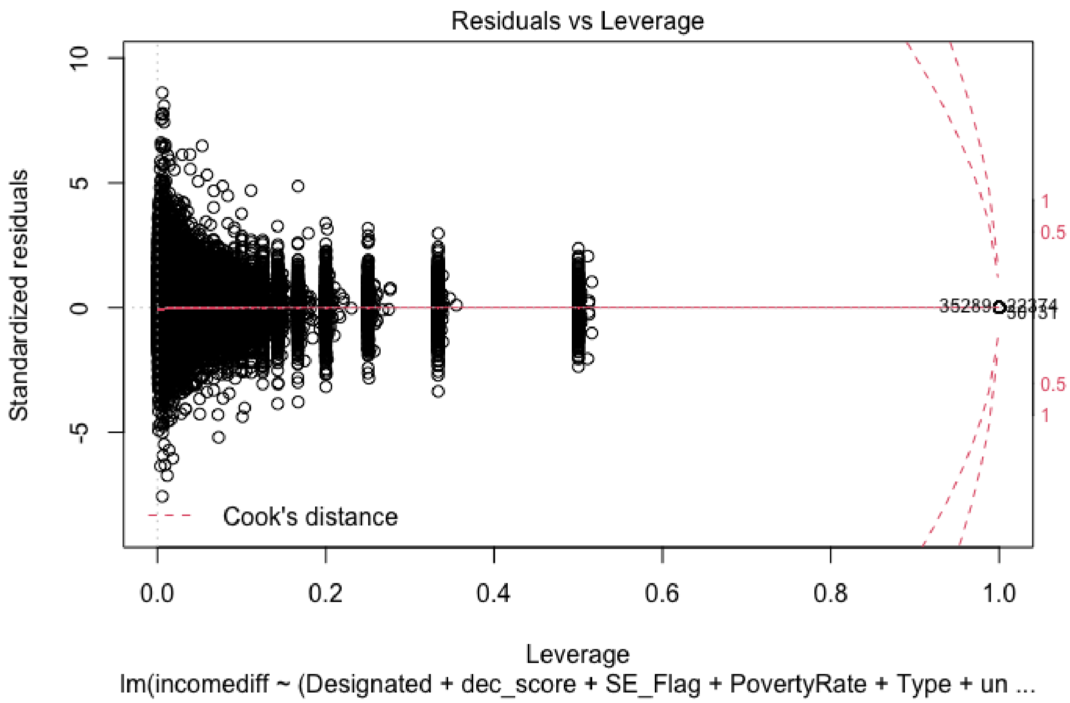


Figure 8, a plot of the residuals against fitted values for regression model 3 (with transformed covariates), reveals no notable heteroscedasticity, but does show the same outlying observations discussed above.

An examination of observations with high leverage (Figure 9) shows several census tracts with an extraordinarily high Cook's distance. However, removal of the observations with the highest leverage (those with a Cook's distance of greater than 0.6) leads to no change in the coefficient for Opportunity Zone designation relative to the version of Model 3 which includes the full population of eligible census tracts.

Figure 9: Leverage Points



2.3.2 Difference in Differences

An alternative method - difference-in-differences - would seem to be a natural tool to determine the effectiveness of the Opportunity Zones program, comparing income levels for treated and untreated areas both before and after the Opportunity Zone designation process in order to estimate the effect of the program on household income. Because the designation process occurred in a single year (functionally simultaneously), it is possible to simply compare the incomes of designated census tracts and eligible-but-not-designated ones in 2016 (before the Opportunity Zone designation process) and 2019 (after the designation process) to establish the causal effect of the program.

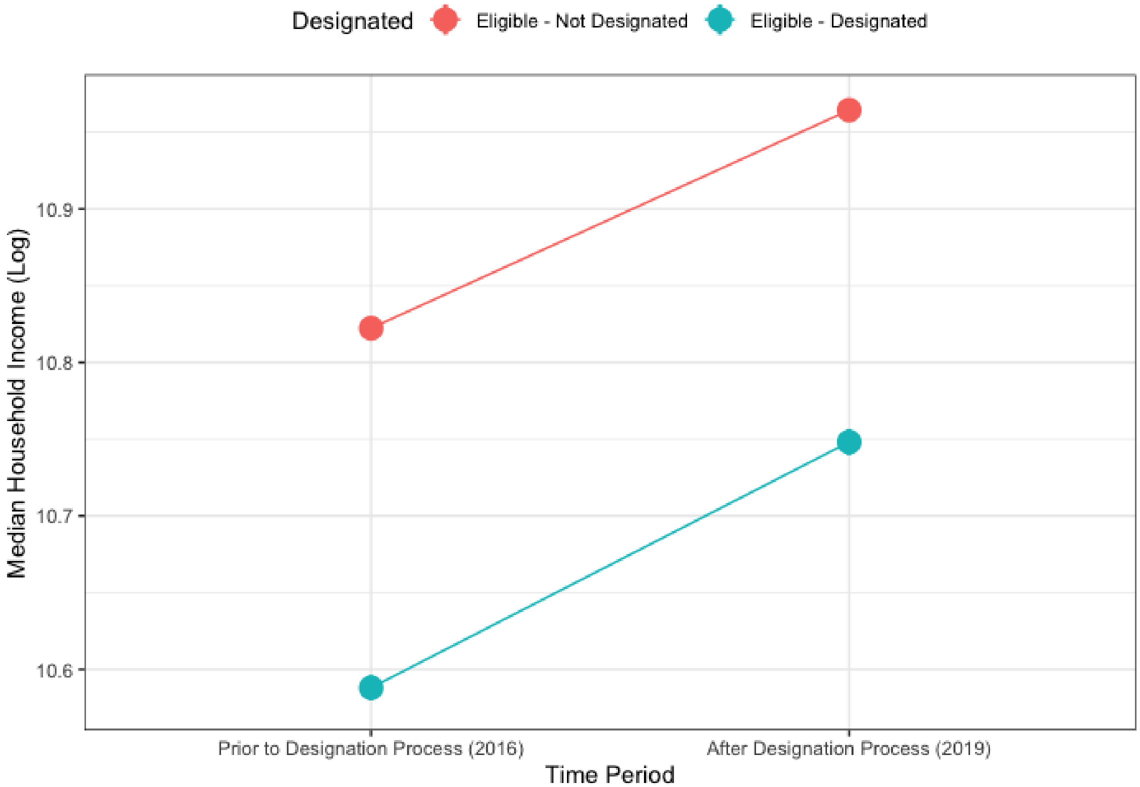
In this approach, one need only assume that, in a counterfactual scenario without the Opportunity Zones program, the areas that were actually designated as Opportunity Zones would have had the same expected rate of growth as areas that were eligible for the program but not designated.

However, for the difference-in-differences regression to provide a valid estimate of the causal effects, the assumption of parallel trends must hold. In the case of this analysis, there it a good chance that it does not. Crucially, there was no uniform process for selecting the tracts that ultimately received the Opportunity Zone designation, and in many cases it appears that areas designated as Opportunity Zones would have received large amounts of outside investment even without the designation. Because of this, designated census tracts would have had higher income growth than eligible-but-not-designated ones in a counterfactual scenario where the Opportunity Zone program was never implemented, and a difference-in-differences analysis will overestimate the average treatment effect of the program on designated census tracts.

The difference-in-difference results suggest that designation as an Opportunity Zone has a positive and statistically significant effect on income growth for designated census tracts relative to an eligible but non-designated comparison group.

The results of this analysis are described in Table 6 and can be visualized in the accompanying Figure 10. In this diagram, the slope of the blue line (tracts designated as Opportunity Zones) is slightly higher than the slope of the orange line (eligible tracts that did not receive the designation). The higher slope of the blue line corresponds to faster growth of median income, and implies a positive effect for the Opportunity Zones program on treated tracts. The left scale is log of median income, and absolute differences in values these logged values correspond approximately to

Figure 10: Difference in Differences



percentage changes in income.

Table 6: Difference in Differences Regression Results

Model	Effect Size	T-Statistic	P-value
No covariates	0.6%	2.88	0.004

The 0.6% effect size suggests that, in areas designated as Opportunity Zones, median household incomes grew 0.6% more per year between 2016 and 2019 than areas that were eligible but not designated. As with the simple regression presented above, this effect implies a fairly substantial practical effect. For the median household in the median designated tract an increase in income growth rate of this size is the equivalent of an annual income around $710 higher than it would have been had the

tract not been designated as an Opportunity Zone.

2.3.3 Stability Controlled Quasi-Experiment

However, there is another approach to estimating the Average Treatment Effect on the Treated (ATT) using a technique introduced by Hazlett.[17] This technique replaces the parallel trends assumption used in difference in differences with an assumption of δ:the counterfactual shift in outcomes given no treatment, which in this case is the annual rate of growth in median household income between 2016 and 2019. This is not something that the researcher using this technique is expected to know precisely, but is something that can be placed within reasonable bounds. For instance, it is unlikely that designated census tracts would have had income growth averaging 10% per year absent the OZ program, or that incomes for designated would have fallen overall had the OZ program not existed. This assumption acknowledges that the assumptions of parallel trends used in difference-in-differences are unrealistically strict, and that under most circumstances the populations being analyzed would not have behaved so nicely in a counterfactual scenario.

Because no single specific δ assumption can be used with certainty, a stability controlled quasi-experiment is used in order to evaluate the estimated ATT for a range of counterfactual growth rates. As a result, the output from this analysis does not consist of a single effect size estimate, but instead is a set of potential effect-size/counterfactual growth rate pairs, which must be independently evaluated for plausibility using domain knowledge and by conferring with outside experts.

When implementing this approach, this paper adopts Hazlett's potential outcomes structure where Y represents the outcome of interest, D is a treatment indicator indicating whether a given census tract was designated as an Opportunity Zone, and T indicates the time period - with T_0 being prior to Opportunity Zone designation in 2016 and T_1 being after designation in 2019.

In addition to these values, we also incorporate the proportion of the population treated in $T = 0$ and $T = 1$. In this scenario, our calculation is simplified by the fact that no census tracts are designated as Opportunity Zones at $T = 0$, but this need not necessarily be the case.

$$\delta \equiv E[Y(0)|T = 1] - \mathbb{E}[Y(0)|T = 0]$$

In its original application to the evaluation of medical outcomes, the SCQE is used to make inferences about the effect of the introduction (or expanded use) of new treatments under circumstances where random assignment cannot be assumed. By comparing the distribution of outcomes both before and after the change in use, and developing an assumption about the counterfactual shift in the distribution of outcomes, an estimate of the ATT can still be developed. Specifically, this approach uses the Law of Iterated expectations to separate the observed and unobserved portions of the counterfactual outcome.

$$\begin{aligned}
E[Y(0)|T = 0] &= E[Y(0)|T = 1] - \delta \\
&= E[Y(0)|D = 1, T = 1]Pr(D = 1|T = 1)+ \\
&\quad E[Y(0)|D = 0, T = 1]Pr(D = 0|T = 1) - \delta
\end{aligned}$$

After rearranging to identify $E[Y(0)|D=1,T=1]$ in terms of observables:

$$\begin{aligned} E[Y(0)|D=1,T=1] &= \frac{E[Y(0)|T=0] - E[Y(0)|D=0,T=1]Pr(D=0|T=1) + \delta}{Pr(D=1|T=1)} \\ &= \frac{E[Y|T=0] - E[Y|D=0,T=1]Pr(D=0|T=1) + \delta}{Pr(D=1|T=1)} \end{aligned}$$

This can then be rearranged to identify the ATT.

$$\begin{aligned} ATT &= E[Y(1)|D=1,T=1] - E[Y(0)|D=1,T=1] \\ &= E[Y|D=1,T=1] - \frac{(E[Y|T=0] - E[Y|D=0,T=1]Pr(D=0|T=1) + \delta)}{Pr(D=1|T=1)} \end{aligned}$$

In this equation, $E[Y(0)|T=1]$ is the expected median family income in these census tracts had the Opportunity Zones program never been implemented, while $E[Y(0)|T=0]$ is the median family income in 2016, before the Opportunity Zones program was implemented (because this is pre-treatment, $E[Y(0)|T=0] = E[Y(1)|T=0]$)

The underlying assumption of SCQE, which differentiates it from difference-in-differences, is the value of δ. In SCQE, δ is assumed, and the effect size of the intervention is evaluated for a given value.

When analyzing the effect of economic policies, determining possible counterfactual growth rates (and thus the relevant δ assumption) is exceptionally challenging because these growth rates are influenced by many factors and are not often stable over time. For this analysis, one must rely on a number of methods to evaluate potential counterfactual growth rates: speaking with economists who are experts in the field, analyzing the relative growth rates of designated and non-designated census tracts and their stability prior to the designation of areas as Opportunity Zones, and looking at areas with income that places them just above the threshold for Opportunity Zone designation.

For the SCQE results, the estimated effect size changes based on the assumed value of δ used. Table 6 presents the δ assumption thresholds at which a positive and negative statistically significant disparity is reached, as well as the δ assumption at which there is no estimated causal effect.

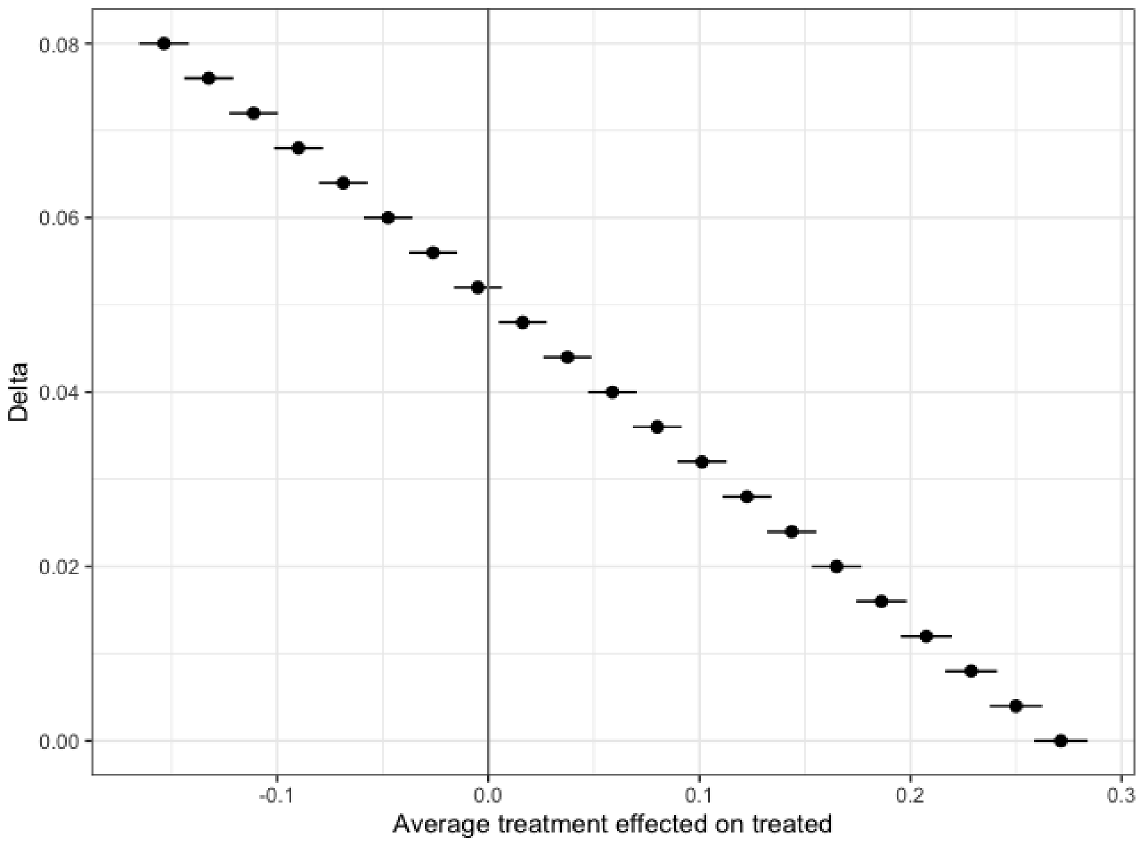
Table 7: Stability Controlled Quasi-Experiment Scenarios

Assumption Required About Mean Growth, Absent Treatment (δ)	Result Justified By Assumption
Counterfactual income growth $\geq$ 5.3%	Opportunity Zone designation has a statistically significant negative effect on income growth
Counterfactual income growth = 5.1%	Opportunity Zone designation has exactly no effect on income growth
Counterfactual income growth $\leq$ 4.9%	Opportunity Zone designation has a statistically significant positive effect on income growth

For the Opportunity Zones program to have a statistically significant positive effect on income growth, income growth rates for a counterfactual scenario without the program must have been 4.9% or lower. This corresponds to the lower-right corner of the accompanying Figure 11, which displays ATT estimates under a variety of counterfactual income growth assumptions.

Is it plausible that eligible census tracts would have had an income growth rate of 4.9% or lower without the Opportunity Zone program? There are a few ways to evaluate this possibility. From 2013 to 2015 (entirely before the Opportunity Zone program, and in fact before the TCJA was even introduced to congress), census tracts that would become eligible for OZ designation had an annual income growth rate of

Figure 11: Potential ATT Estimates from SCQE



3.4% - much lower than the 4.9% δ assumption threshold for a statistically significant positive treatment effect. Of course, it is extremely risky to use income growth rates from an earlier time period as a counterfactual assumption about income growth rates in a later period, but this shows that a δ assumption of 4.9% or lower is at least in the range of plausibility.

Table 8: Income Growth in Designated Opportunity Zones and Eligible-But-Not-Designated Census Tracts

	Average Income Growth 2013-2015	Income Growth 2016-2019
Designated	3.3%	5.6%
Not Designated	3.4%	5.0%
Overall	**3.4%**	**5.1%**

Historically, poor areas have consistently grown slower than wealthier ones.[18] This suggests a potential "upper bound" to counterfactual income growth; areas that were eligible for the Opportunity Zone program would not have grown faster than wealthier areas had the program not existed. Areas that were just barely ineligible for the Opportunity Zone program (i.e. the 5% of the tracts with a incomes just above the eligibility threshold) had income growth of 4.8% from 2016-2019. Taking this upper bound as the counterfactual growth rate would imply an ATT of between 0.2% and 2.5%. This is in line with the other results presented above, and suggests that the Opportunity Zone program raised median household incomes by anywhere from $300 to up to as much as $3,100 in the median treated census tract.

Taken in concert with the earlier difference-in-differences analysis, this would suggest that Opportunity Zones had a positive effect on incomes in the treated census tracts, but with a wide range of potential effect sizes. An effect size of over 2% seems implausibly large. In evaluating other research, there is limited apparent evidence that vast swathes of the US were being hurled out of poverty thanks to their designation as Opportunity Zones (which is what an effect of that size would imply). However, the plausible estimates of the SCQE and the estimate derived from difference-in-differences align well, suggesting that Opportunity Zones have increased income growth in treated areas, and that this increase in growth rate is in the range of 0.2 to 2.0 percent.

2.3.4 Regression Discontinuity

An alternative form of analysis is also available using this data: a regression discontinuity. The Opportunity Zones program was structured so that each state faced a hard cutoff for Opportunity Zone eligibility (80 percent of state median income), so the the exact income level for eligibility changed based on a given state's median income. Some rural tracts with high levels of out-migration had slightly higher income thresholds - 85 percent of state median income rather than 80 percent - but these tracts are excluded from this analysis. This threshold provides an instrument to identify the effect of eligibility for the Opportunity Zone program on income growth rates by comparing census tracts just below the threshold of eligibility with tracts just above it.

However, this approach has limited utility because of the patterns of Opportunity Zone designation. As discussed above, census tracts designated as Opportunity Zones have lower median income on average than tracts that are eligible but do not receive the designation. This effect is such that it is highly uncommon for tracts near the eligibility cutoff to be designated as Opportunity Zones. As a result, what might be a clean discontinuity corresponding to the eligibility threshold does not actually appear upon closer inspection of the data.

When performing F tests to establish the suitability of potential discontinuity windows, it is found that there are no appropriate windows surrounding the eligibility window with sufficient representation of designated tracts for a useful estimate of the relationship between eligibility and income growth.

3 Discussion

This analysis shows the applicability of SCQE to the evaluation of economic policies. It also reveals the core challenge to interpreting SCQE results when applied to these causal questions: determining appropriate counterfactual assumptions. Unlike in medical research, where many illnesses take independent courses for each individual case, economic outcomes are far harder to predict, and plausible counterfactual scenarios far harder to establish. Because the utility of the SCQE methodology is fundamentally dependent on the accuracy of the δ assumption used, effective results depend on the careful determination of this assumption. Based on my research, it seems likely that a δ of less than 4.9% is plausible, which leads to the conclusion that Opportunity Zones have had a statistically significant effect on growth in median household income in treated tracts.

When it comes to the effect of the Opportunity Zones program itself, these results suggest that the implementation of this program has increased the growth rate of household incomes in designated census tracts under the most plausible counterfactual growth rate assumptions, and that this increased growth rate has increased the income of the median household in these areas by anywhere between $900 and $3,000 per year during the 2016-2019 study period.

For the 12.4 million households who live in Opportunity Zones, the difference-in-difference regression results suggest that their designation has increased annual household income by approximately \$8.8 billion over the 2016 to 2019 period. Is this worth it? Estimates of the total cost of the Opportunity Zone program vary. Initial estimates suggested a total cost (in the form of foregone capital gains tax revenue) of \$1.6 billion over 10 years.[8] Subsequent analyses have estimated a higher cost, with the Joint Committee on Taxation estimating a cost of \$8.2 billion over 5 years.[14] If these estimates of cost are accurate, the results presented here suggest that the Opportunity Zone program has been a cost-effective way to increase income within the targeted census tracts. Even under the highest cost estimate, each dollar in foregone tax revenue increases annual income for the median household in the treated tracts by \$5.36. This estimate is driven by the difference-in-differences analysis, where the parallel trends assumption likely does not hold. But even under the most conservative δ assumption considered plausible for SCQE, the Opportunity Zones program has been cost-effective, raising incomes by more than its cost in tax revenue. Of course, the federal government's foregone revenue from Opportunity Zones also comes with an opportunity cost, but evaluating this is beyond the scope of this analysis.

3.1 Limitations

Several other factors could also affect this conclusion. First, the data used here is not able to differentiate between changes in median income within census tracts that are a result of changing incomes for households living in those census tracts for the entire 2016-2019 period and changes in income that are a result of changing census tract composition. For instance, it is possible that in low-income census tracts designated as Opportunity Zones, poorer households are more likely to move away from the census tracts, and richer households are more likely to move in from other areas.

This is compatible with the assertion by some individuals that Opportunity Zones are “Gentrification Bombs”, which have been specifically targeted at areas that are on the cusp of significant population transitions, accompanied by new investment and development.[10]

This assertion of cost-effectiveness also assumes that there are no “spillovers” between treated and untreated areas. In reality, there are plausible mechanisms by which the Opportunity Zone designation may have either increased or decreased incomes in untreated areas. For instance, if an Opportunity Zone designation leads to more investment in a given census tract, this investment might be expected to raise incomes even in areas beyond the boundaries of the designated tract. On the other hand, the Opportunity Zone designation might also reduce the amount of investment in nearby non-designated areas. For instance, consider a shopping mall developer capable of serving a large area from an individual development. After the Opportunity Zone designation process, this developer may preferentially select areas that are in Opportunity Zones, and nearby eligible-but-not-designated areas may experience a decline in investment. That may or may not lead to a corresponding decline in incomes, depending on the degree to which corresponding wage effects spill over across tract boundaries.

3.2 Further Research

Future research could address these issues in several ways. First, a geospatial element could be incorporated. Instead of simply evaluating areas designated as Opportunity Zones against those eligible-but-not designated, an analysis could instead match areas near designated Opportunity Zones with similar areas further away - including areas that were above the eligibility threshold.

Another major question regarding the performance of Opportunity Zones remains unresolved in the current literature: whether the program has actually increased investment in targeted areas. Under the proposed causal pathway used in this paper, the Opportunity Zone designation lead to increased investment, which then results in higher wages. If it were the case that Opportunity Zones did not lead to changes in investment patterns, it would enable the rejection of the entire putative pathway. Future research could also use administrative data or information on construction or building permits to directly measure changes in investment and evaluate the plausibility of this pathway.

There is a potential counterargument to this, which is that the slightly-too-rich areas often border areas designated as Opportunity Zones. If the tax incentives provided by Opportunity Zones end up drawing marginal investment away (reducing the amount of investment in these slightly-too-rich areas), then the observed income in the too-rich areas would actually be lower than counterfactual income, and the "upper bound" for counterfactual growth would actually be too low.

REFERENCES

[1] American Community Survey, 2016 And 2019 American Community Survey 5-year Estimates, Table B06009.

[2] American Community Survey, 2016 And 2019 American Community Survey 5-year Estimates, Table B17026.

[3] American Community Survey, 2016 And 2019 American Community Survey 5-year Estimates, Table B19001.

[4] American Community Survey, 2016 And 2019 American Community Survey 5-year Estimates, Table B19113.

[5] American Community Survey, 2016 And 2019 American Community Survey 5-year Estimates, Table B23025.

[6] American Community Survey, 2016 And 2019 American Community Survey 5-year Estimates, Table C17002.

[7] Ernsthausen, Jeff and Elliott, Justin, one Trump Tax Cut Was Meant To Help The Poor. A Billionaire Ended Up Winning Big. — ProPublica, june 19, 2019 https://www.propublica.org/article/trump-inc-podcast-one-trump-tax-cut-meant-to-help-the-poor-a-billionaire-ended-up-winning-big.

[8] Joint Committee On Taxation, Estimated Revenue Effects Of The Conference Agreement For H.r. 1, The "tax Cuts And Jobs Act", Pg. 6, Https://www.jct.gov/publications.html?func=startdown&id=5053.

[9] The New York Times Editorial Board, Opinion | Opportunity Zones — For Billionaires, November 16, 2019.

[10] Smith, Edward, To Understand Why Some Black Men Support Trump, Start With Ice Cube, October 17, 2020, Https://www.latimes.com/california/story/2020-10-17/ice-cube-black-men-trump-biden-voter-cwba.

[11] Smith, Lisa, West End Residents Concerned About Gentrification With Stadium Development, February 8, 2019 https://www.wcpo.com/news/transportation-development/move-up-cincinnati/west-end-residents-concerned-about-gentrification-with-stadium-development. Section: Move Up Cincinnati.

[12] A Trump Tax Break To Help The Poor Went To A Rich GOP Donor's Superyacht Marina.

[13] The Urban Institute, Opportunity Zones, Https://www.urban.org/policy-centers/metropolitan-housing-and-communities-policy-center/projects/opportunity-zones.

[14] U.s. Congress, Joint Committee On Taxation, Estimates Of Federal Tax Expenditures For Fiscal Years 2020-2024, Pg. 32, Https://www.jct.gov/publications/2020/jcx-23-20.

[15] Chen, Jiafeng And Glaeser, E. The (Non-) Effect Of Opportunity Zones On Housing Prices, December 2019. Series: Working Paper Series.

[16] Frank, Mary Margaret And Hoopes, J. What Determines Where Opportunity Knocks? Political Affiliation In The Selection Of Opportunity Zones, June 2020.

[17] Hazlett, C. Estimating Causal Effects Of New Treatments Despite Self-selection: The Case Of Experimental Medical Treatments. Publisher: De Gruyter Section: Journal Of Causal Inference, March 2019.

[18] Li, Huiping And Campbell, H. A. F. S. Residential Segregation, Spatial Mismatch And Economic Growth Across US Metropolitan Areas, Pg. 17. 2642–2660. Publisher: SAGE Publications Ltd, October 2010.