ABSTRACT

The Effect of *Ruiz v. Estelle* on Crime in Texas

Hunter A. Morris, M.S.Eco.

Mentor: Scott Cunningham, Ph.D.

I examine the effects of a 1980 Texas District Court ruling that ruled the conditions in the Texas prison system unconstitutional. In particular, I analyze its effects on crime and incarceration. I use a difference-in-differences estimation strategy to estimate these effects, and I implement the synthetic control method of Abadie, Diamond, and Hainmueller (2010) as a robustness check.

The Effect of *Ruiz v. Estelle* on Crime in Texas

by

Hunter A. Morris, B.A.

A Thesis

Approved by the Department of Economics

_____

Charles North, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Master of Science in Economics

Approved by the Thesis Committee

_____

Scott Cunningham, Ph.D., Chairperson

_____

James West, Ph.D.

_____

Jeanne Hill, Ph.D.

Accepted by the Graduate School
May 2020

_____

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

Introduction

Background

In 1980, Judge William Wayne Justice of the U.S. District Court for the Eastern District of Texas handed down his decision in *Ruiz v. Estelle*, in which he ruled that the conditions in Texas prisons constituted "cruel and unusual" punishment and were thus unconstitutional (Perkinson 2010). The case began in 1978 as a class-action lawsuit on behalf of David Ruiz, a jailhouse lawyer or "writ writer," and all prisoners in the custody of the Texas Department of Corrections (TDC) (Perkinson 2010). The court found the conditions in TDC's prisons to be abysmal. One doctor alone recalled seeing "hundreds" of cases of brutal abuse (Perkinson 2010). Cells were sometimes made to house four inmates at a time, with "matresses on the floor [and] tents in the yard" (Perkinson 2010). Ultimately, it was decided that the conditions the prisoners were experiencing violated their constitutional rights and that sweeping changes were necessary to remedy the deficiencies.

In consequence of this ruling, in 1981 TDC was put under court supervision and subject to numerous constraints related to quality of life in their prisons. Among these constraints, perhaps the most keenly felt were the overcrowding restrictions. In 1983, the director of TDC requested a massive budget increase in order to help the department meet the conditions imposed by the court (Perkinson 2010). Instead, the legislature decided to divert "tens of millions of dollars from regular prisons to probation, parole, and 'restitution centers,"', as well as to "[loosen] good-time provisions, [and eliminate] the governor's veto over paroles," (Perkinson 2010).

In this paper, I attempt to study the impacts of *Ruiz* and its consequent legislative intervention on the Texas criminal justice system. I therefore analyze five

incarceration-related outcomes: discretionary paroles, total incarceration, total releases, total admissions, and net flow of prisoners into or out of the system. In addition, I study the impact of the intervention on crime. I hypothesize that the constraints placed on TDC will lead to a large increase in paroles, a fall in total incarceration, and a fall in the net flow rate of prisoners. I also hypothesize that the increase in the population of potential criminals that results from these changes will lead to an increase in the crime rate. Because I am interested in the effects of the prison population constraints, I take the 1983 legislative intervention as the beginning of the treatment, rather than 1980 court ruling. This is because I believe that the population constraints became binding only after the legislature refused to increase TDC's funding. I take 1992 as the end of the treatment period, since that is the year in which Texas began to be released from court supervision. In the following chapters, I initially implement a difference-in-differences design to study these impacts. I note the problems with the usual mode of statistical inference in this setting and choose an alternative. I then implement the synthetic control method of Abadie, Diamond, and Hainmueller (2010), with an important modification due to Ferman and Pinto (2019b) and Doudchenko and Imbens (2016), as a robustness check.

Data

Data on crime outcomes for the years 1969–1992 come from the FBI's Uniform Crime Report (UCR). The UCR is a voluntary national survey of law enforcement agencies (city, state, and university police departments, sheriff's departments, etc.). In part I of the survey, each surveyed agency reports the number of four types of violent crime (murder and non-negligent manslaughter, forcible rape, robbery, and aggravated assault) and three types of property crime (burglary, larceny-theft, and motor vehicle theft) made known to them each month. Their estimates are based on "reports received from victims, officers who discover infractions, or other sources"

(United States Bureau of Justice Statistics 2020). These agency-level estimates can then be aggregated to form state-level estimates for each of these crimes. Although the raw UCR data are publicly available, I choose to make use of the state-level estimates from the Bureau of Justice Statistics' (BJS) online data tool, which are of course based on the raw UCR data. The reason for this is two-fold. First, it eliminates the risk of error associated with performing the aggregation manually. Second, the BJS data imputes crime estimates for those agencies with reporting problems. This imputation will potentially help mitigate any biases that would arise from omitting these problem states if the reporting problems turn out to be systematic.

I obtain data on prison outcomes for the years 1978–1992 from the BJS's National Prisoner Statistics program (NPS). The NPS is a yearly survey of state departments of correction which measures both stocks and flows of prisoners in each state's prison system. Stocks are measured as of the 31^{st} of December of the year in question, while flows are measured from January 1^{st} to December 31^{st} of the year in question. The outcomes that I am interested in are discretionary paroles (henceforth simply paroles), total number of prisoners in custody (henceforth incarceration), total releases, total admissions, and the operational capacity of the system. Total incarceration is not recorded for the years 1978–1982; however, admissions and releases are recorded for all years. I therefore subtract total releases from total admissions to obtain an estimate of the net flow of prisoners through the system. I am then able to use this estimate to impute total incarceration for the missing years. As we shall see, I am also interested in the net flow of prisoners as an outcome in its own right. Finally, the state of California does not report paroles for most of the period of interest. I therefore drop California from all estimation involving paroles.

State-level covariates for the years 1977–1992 come from the Census Bureau's Current Population Survey (CPS). The CPS is a monthly survey of households conducted by the U.S. Census Bureau and the Bureau of Labor Statistics. The survey is a

useful source of demographic and economic data. I use the CPS to obtain estimates of the median household income, divorce rate, unemployment rate, racial composition, educational attainment and age distribution of each state.

Finally, population estimates for the years 1969–1992 come from the National Cancer Institute's Surveillance, Epidemiology, and End Results Program (SEER). Though SEER obtains its yearly estimates of population from the U.S. Census Bureau, this data is more easily accessible as part of SEER.

It should be noted that Texas was not the only state to receive a shock of this kind to their prison system during the period of interest. In fact, nearly two dozen other states received had similar rulings handed down against them in federal court during this time. These include Alabama, Arizona, Arkansas, Colorado, Delaware, Florida, Georgia, Illinois, Iowa, Kentucky, Louisiana, Maryland, Mississippi, Missouri, New Hampshire, New Mexico, New York, Oklahoma, Ohio, Oregon, Pennsylvania, Rhode Island, and Tennessee, who all had their prison systems declared unconstitutional by federal courts during the period of interest. Since the estimation procedures used in this paper involve comparing Texas with *untreated* control states, these states do not serve as good controls. They will therefore be dropped in all the following analyses.

CHAPTER TWO

Difference-in-differences

Model

The first technique I will use to estimate the impact of these population constraints is the difference-in-differences (DID) method. The basic idea behind the DID estimator is this: in certain contexts, it may be reasonable to assume that the mean difference in the outcome of interest between the treated and the untreated groups, conditional on some covariates, is fixed over time in the absence of treatment. This crucial assumption is known as the "parallel trends" assumption. Given this assumption, estimation of the treatment effect proceeds as follows. We first estimate the conditional difference in means between the treated and control units for the pre-treatment period. Then the conditional difference in means for the post-treatment period is estimated. Any deviation in this estimate from the estimate for the pre-treatment period is then imputed to the treatment.

In a linear-regression framework with one treated unit, DID can be estimated using the following two-way fixed-effects model:

$$Y_{st} = \beta_0 + \beta_1 D_{st} + \boldsymbol{\gamma} \boldsymbol{X}_{st} + \tau_t + \sigma_s + \epsilon_{st},$$

where Y_{st} represents the crime rate for state s at time t, D_{st} is a dummy for whether state s was treated in time t, β_1 is the parameter of interest, $\boldsymbol{\gamma}$ is a $1 \times k$ vector of parameters, $\boldsymbol{X}_{st}$ is a $k \times 1$ vector of relevant controls,[1] τ_t and σ_s are year and state fixed-effects, respectively, and ϵ_{st} is a transitory shock. As usual, unbiased estimation requires that $\mathbb{E}[\epsilon_{st}|D_{st}, \boldsymbol{X}_{st}, \tau_t, \sigma_s] = 0$. This is actually the parallel trends assumption in disguise, since it implies that, in expectation, the difference between units, conditional on treatment, covariates, and fixed-effects, would have been fixed

1. Time-variant controls are listed in the tables of coefficient estimates below.

in the absence of treatment. This assumption also implies that treatment cannot be correlated with state-specific, time-varying unobservables. However, arbitrary correlation is allowed between treatment and state-specific, time-invariant unobservables, or between treatment and time-variant unobservables that effect all states equally, since these are "differenced out" by the state and time fixed-effects, respectively.

This flexibility in unobservables is essential to my application. For example, it is almost certain that Texas received the treatment precisely because the conditions in its prisons were worse than those in the control states. It also plausible to think that the conditions in Texas prisons could be correlated with our outcomes of interest, especially the incarceration rate, since overcrowding is a large part of what made Judge Justice rule Texas prison conditions unconstitutional. As long as this difference in prison conditions if fixed over time, however, this will not bias the treatment effect estimates.

Inference

Issues with Cluster-Robust Standard Errors

To paraphrase Buchmueller, DiNardo, and Valletta (2011), while estimation is fairly straightforward in our application, inference is not. OLS standard error estimates rely on the homoskedasticity assumption in order to be consistent. It is very reasonable to suppose, however, that the variance of ϵ_{it} will differ across states. To correct for this, economists often employ "cluster-robust" standard error (CRSE) estimates, which are an extension of the White-Huber heteroskedasticity-robust standard error estimates. However, CRSE estimates are not reliable in the case of only one treated unit. To illustrate why this is so, I borrow a discussion from Ferman and Pinto (2019a), adjusting the notation slightly to conform with mine. To simplify the analysis, let us consider the case where covariates are excluded from the model. Let $s = 1$ be the treated state, and let there be N observed states. Let there also be T

observed time-periods, and let T_0 be the last time period before treatment. Then the DID estimator is given by

$$\begin{aligned}\hat{\beta}_1 &= \left[\frac{1}{T-T_0}\sum_{t=T_0+1}^{T} Y_{1t} - \frac{1}{T_0}\sum_{t=1}^{T_0} Y_{1t}\right] - \frac{1}{N-1}\sum_{s=2}^{N}\left[\frac{1}{T-T_0}\sum_{t=T_0+1}^{T} Y_{st} - \frac{1}{T_0}\sum_{t=1}^{T_0} Y_{st}\right] \\ &= \beta_1 + W_1 + \frac{1}{N-1}\sum_{s=2}^{N} W_s,\end{aligned}$$

where

$$W_s = \frac{1}{T-T_0}\sum_{t=T_0+1}^{T} \epsilon_{st} - \frac{1}{T_0}\sum_{t=1}^{T_0} \epsilon_{st}.$$

It follows that

$$Var(\hat{\beta}_1) = Var(W_1) + \left(\frac{1}{N-1}\right)^2 \sum_{s=2}^{N} Var(W_s),$$

since, in the clustering framework, ϵ_{st} is assumed to be independent across s. The cluster robust variance estimate is then proportional to

$$\widehat{Var(\hat{\beta}_1)_{cluster}} \propto \widehat{W}_1^2 + \left(\frac{1}{N-1}\right)^2 \sum_{s=2}^{N} \widehat{W}_s^2,$$

where the proportionality constant is a degrees-of-freedom correction. $\widehat{W}_s$ is the same as W_s with ϵ_{st} replaced by the regression residual $\hat{\epsilon}_{st}$. But $\widehat{W}_1 = 0$ by construction. To see why this is the case, recall that the fixed-effects model that we are estimating essentially fits a separate intercept for each state, each year, and for the treated state in the post-treatment period. By the OLS normal equations, the regression residuals must then sum to zero for each state, each year, and for the treated state in the post-treatment period. Thus the first term of $\widehat{W}_1$ must equal zero, which implies that the second term must also equal zero. Thus, $\widehat{Var(\hat{\beta}_1)_{cluster}}$ will systematically underestimate $Var(\hat{\beta}_1)$, which will lead to an over-rejection of the null hypothesis.

But how serious is this problem? MacKinnon and Webb (2017) run Monte-Carlo simulations that suggest that t-tests based on CRSE estimates when there is only one treated unit can reject a true null-hypothesis over 50% of the time for a nominal test size of 5%. Thus, we must seek a different inference strategy if we are to obtain a meaningful estimate of the statistical significance of our results.

Permutation Inference

P-values represent the estimated probability (under the null hypothesis and given that our modeling assumptions are correct) of observing a test statistic at least as extreme as the one actually observed. But there are multiple possible approaches to assigning this probability. The classic, sampling-based approach assigns this probability as the proportion of test statistics under the null hypothesis that a researcher would observe to be at least as extreme as the one actually observed if the researcher could take infinitely many random samples from the population of interest. But in experimental and quasi-experimental settings, sampling is not the only source of randomness in the data. Another, perhaps more fundamental source of randomness is the assignment of treatment. This leads me to the inferential strategy that I will be using throughout this paper: permutation inference. Under this paradigm, p-values are assigned as the proportion of test statistics under the null hypothesis that a researcher would observe to be at least as extreme as the one actually observed *if treatment had been assigned in different, equally likely ways.*

To make this more concrete, let us consider an example. Suppose a clinical trial of an experimental drug is being conducted on 20 patients, 10 of whom are randomly assigned the treatment, and the other 10 of whom are given a placebo. Let the difference in means along the outcome of interest then be calculated. In the permutation inference paradigm, the right-tailed p-value associated with this statistic could then be computed by repeatedly calculating the difference in means under the $\binom{20}{10} - 1$ other possible treatment assignments, taking as the p-value the proportion of statistics greater than or equal to the one actually observed.

Two things should be noted at this point. First, the data in this example are treated as fixed; no reference is made to hypothetical repeated samples. Second, it is essential that the counterfactual treatment assignments be *equally likely* with the one

that is actually observed. It is this assumption that allows us to calculate the p-value as a simple proportion.

There are several other important differences to consider between sampling-based and permutation inference. First, since permutation inference is done without reference to repeated sampling, results do not generalize to the population, but are only valid for the observed individuals. In our application this is not a problem, since we *observe* the population of interest. Second, permutation inference gives exact p-values in finite samples, whereas many sampling-based approaches give only asymptotically valid p-values.

In our application inference will be conducted as follows. First, the estimated treatment effect will be calculated by estimating the model specified above. Then each of the control states will be designated as the "treated" state in turn, and the same DID model will estimated on each permutation of the data in turn.[2] Then the estimated treatment effects will be ranked, and the proportion of estimated coefficients that are greater than or equal to (or lesser than or equal to, depending on whether the estimated coefficient for Texas is above or below the median) multiplied by 2 will serve as our p-value. The multiplication by 2 serves to implement a two-tailed test.

Comparing the example case of the clinical trial above with our application reveals a weakness of permutation inference in our case. As was said before, permutation inference relies on the assumption that each of the counterfactual treatment assignments is equally as likely as the one actually observed. In the case of the clinical trial, this is a good assumption, since the researcher controls treatment assignment and can *ensure* that treatment is random and that every possible treatment assignment is equally likely. In our application, however, it is more difficult to say that it is equally likely that each of the other states could have been assigned the treatment, even if we condition on covariates and fixed-effects. This is because we are not study-

2. The treatment year will remain 1983 for all permutations

ing a generic, standard treatment that could have, in principle, been applied to the other states in some alternate universe. The treatment is a tailored, one-off policy intervention, unique to the place and time in which it was carried out.

Despite this, I will proceed to use permutation inference as my inference strategy. In this, I follow Buchmueller, DiNardo, and Valletta (2011) in their analysis of health insurance mandates on health coverage in Hawaii. The main reason for persisting in this imperfect inference strategy is that the p-values generated under this procedure are much more conservative than those based on cluster-robust standard errors, so permutation inference is at least an improvement over the alternative, in terms of over-rejection.

Another, non-stochastic interpretation of permutation-based p-values is possible, which is hinted at by Abadie, Diamond, and Hainmueller (2010) when discussing their preferred inferential procedure for synthetic control. Their procedure is closely related to permutation inference and will be discussed in detail in the next chapter. Their interpretation of p-values obtained under this procedure when treatment is not equally likely for all subjects under the null hypothesis is that the p-values represent how unusual the test statistic is relative to a "placebo distribution," the distribution of estimated test statistics for the untreated units. If the null hypothesis implies that the estimated effect for the treatment state should not be unusual or extreme, then the p-value is informative about the alternate hypothesis. However, the p-value under this interpretation is not a probability, but more of a heuristic.

Ultimately, there is no inference strategy that I know of that is unambiguously better than the one I have described here. Inference in DID with one treated unit is a notoriously difficult problem, one that statisticians and econometricians are still actively trying to solve satisfactorily.

Results

Before I give the formal results of my regressions, it may be enlightening to first present some preliminary graphical evidence. This evidence serves two purposes. First, it provides an informal test of the parallel-trends assumption. If the parallel trends assumption holds in our context, we should expect the outcomes of interest for Texas in the pre-treatment period to follow roughly similar trajectories to the mean outcomes of the control states. As can be seen in Figure 2.1, this is arguably the case. Second, this graphical evidence gives us a sense of how the outcomes evolve over time, which is lacking from the coefficient estimates. As can be seen by inspecting Figure 2.1, it does seem that Texas did see a spike in discretionary paroles and a drop in incarceration relative to the control group. In addition, there appear to be a large spikes in property and violent crime in the wake of the treatment. However, there does not seem to be a large change in the number of releases, the number of admissions, or the net flow of prisoners compared to the control group.

The regression results for incarceration outcomes are given in Tables 2.1 and 2.2, and the results for crime outcomes are given in Table 2.3. For each outcome I estimate conditional and unconditional models, since it is not clear that the demographic controls are actually necessary for unbiased estimation in this application. Since state-level covariates are available only for the period 1977-1992, but crime data is available going back to 1969, I estimate unconditional DID models of crime for both periods. It should be noted that, since there are only 26 or 27 control states (depending on which outcome is being analyzed), the highest level of significance that can be achieved under our inference procedure is 10%. In addition to the coefficient estimates and permutation-based p-values, I present the implied average percent change in the outcome of interest over the post-treatment period corresponding to each estimate. This is calculated by first subtracting the estimated coefficient from the corresponding

post-treatment average in Texas, then dividing the estimated coefficient by the result of this subtraction and multiplying by 100.

As expected, I estimate a large effect on paroles, an implied increase of 112.62% and 97.92% over the counterfactual for the unconditional and conditional model, respectively. In addition, these estimates have permutation based p-values of 8% and 15%, respectively. However, though the signs of the point estimates for all other incarceration outcomes are consistent with the hypothesis of large response by the Texas criminal justice system aimed at curbing the prison population, and while the magnitudes of these estimates are fairly large, none of them are statistically significant in either the unconditional or conditional models. In particular, given that the interventions were specifically designed to ease overcrowding, it is strange that the estimated change in incarceration rate is not statistically different from zero. If we were to take this at face value, it would lead us to conclude that Texas incarcerated just as many individuals as it would have if there had been no treatment. But this would amount to saying that the constraints opposed by the court and legislature were non-binding. This would be a strange conclusion, to say the least. Non-compliance is not a likely explanation, given that TDC was under court supervision for the entire treatment period.

Perhaps TDC responded by increasing their operational capacity, in which case they would be able to ease overcrowding without decreasing the incarceration rate over the counterfactual. Unfortunately, the NPS data on the operational capacity of the prison system is missing for 1978–1982. Consequently I cannot perform any formal statistical tests to determine if this is the case. I do, however, plot the operational capacity in Figure 2.2. It can be seen that the operational capacity initially remains stable in the wake of the treatment and then begins to rise sharply just before 1990. This suggests that perhaps Texas did need to decrease its incarceration rate initially, but then started increasing it again before the end of the treatment period as op-

erational capacity increased, thus wiping out any detectable effect. In fact, this is consistent with what we see in the incarceration panel of Figure 2.1.

It is also puzzling that our estimation fails to detect a significant increase in releases, given that discretionary paroles are a subset of releases. And even if the estimates were statistically different from zero, the estimated effect on releases is only one-third to one-half as large as that for paroles, suggesting a partial substitution away from other forms of release toward parole. The point estimate for the effect on admissions is also positive, which together with the positive estimate for releases, suggests an increase in gross flow of prisoners through the Texas prison system. However, the net flow point estimate is negative.

Given the failure to detect any significant impact on incarceration, the results regarding crime outcomes are striking. Both property and violent crime rates see large increases: 27.58–37.24% and 20.74–30.52% over the implied counterfactual, respectively. What is more, the estimates are all significant at the 10% level, except for violent crime in the unconditional model for the period 1969–1992. There are several ways we could take these seemingly dissonant results. As suggested above, it could be the case that an initial decrease in incarceration was followed up by a large increase in prison capacity that negated the initial impact. Interestingly, the spikes in property crime and violent crime that can be observed in Figure 2.1 seem to reverse direction about the same time that prison capacity and incarceration are both rising sharply. It could also be that *Ruiz v. Estelle* and its consequent legal and legislative interventions did somehow increase crime, but not through the hypothesized channel of decreasing the incarceration rate. For example, perhaps parolees are more likely to commit crime than other potential criminals. Then a sudden shock to the composition of the pool of potential criminals that increased the number of parolees would be sufficient to increase the crime rate, even if the total population of potential criminals did not increase. Or perhaps the bettering of conditions in the Texas prison system made the

prospect of going to prison less of a deterrent for potential criminals. Finally, it could be that all of my statistically significant coefficients are false rejections of the null hypothesis, perhaps due to incorrect modeling assumptions. For example, perhaps a Texas-specific, time-variant shock to prison conditions did increase the chance of treatment, biasing my estimates.

Of course, all of these results rely on comparing the treated state with the simple average of the control states. But what if a different linear combination of states exists that would provide a better control group for Texas? This is the question that motivated Abadie, Diamond, and Hainmueller (2010) to introduce the synthetic control method. I will now proceed to discuss this method, its assumptions, and its relationship with the DID method. I will then use the synthetic control method to estimate the impact of the treatment on my outcomes of interest, in order to test the robustness of my previous results.

Table 2.1. Incarceration Outcomes (Difference-in-differences)

	Paroles	Incarceration	Releases	Admissions	Flow
			1978–1992		
Treatment	68.76*	−46.83	36.81	33.89	−2.93
	p = 0.08	p = 0.65	p = 0.43	p = 0.36	p = 0.93
Implied % Change	112.61	-15.13	21.75	21.8	-14.84
Controls	No	No	No	No	No
N	405	420	420	420	420

Note: All outcomes measured per 100,000 population per year. P-values are two-sided permutation based p-values. Time variant controls include median income, unemployment rate, percent male, white, black, asian, divorced, age 0-18, age 19-30, age 31-65, age 66+, high school dropouts, with high school diploma, with some college education, and with bachelor's degree. *p<0.1; **p<0.05; ***p<0.01

Table 2.2. Incarceration Outcomes (Difference-in-differences, cont'd)

	Paroles	Incarceration	Releases	Admissions	Flow
			1978–1992		
Treatment	64.23	−50.13	25.50	21.52	−3.98
	p = 0.15	p = 0.79	p = 0.43	p = 0.50	p = 0.65
Implied % Change	97.92	-16.02	14.12	12.83	-19.15
Controls	Yes	Yes	Yes	Yes	Yes
N	405	420	420	420	420

Note: All outcomes measured per 100,000 population per year. P-values are two-sided permutation based p-values. Time variant controls include median income, unemployment rate, percent male, white, black, asian, divorced, age 0-18, age 19-30, age 31-65, age 66+, high school dropouts, with high school diploma, with some college education, and with bachelor's degree. *p<0.1; **p<0.05; ***p<0.01

Table 2.3. Crime Outcomes (Difference-in-differences)

	Property	Violent	Property	Violent	Property	Violent
	1969–1992		1977–1992		1977–1992	
Treatment	1,824.39*	157.96	1,641.44*	137.93*	1,453.45*	116.03*
	p = 0.08	p = 0.22	p = 0.08	p = 0.08	p = 0.08	p = 0.08
Implied % Change	37.24	30.52	32.3	25.66	27.58	20.74
Controls	No	No	No	No	Yes	Yes
N	672	672	448	448	448	448

Note: All outcomes measured per 100,000 population per year. P-values are two-sided permutation based p-values. Time variant controls include median income, unemployment rate, percent male, white, black, asian, divorced, age 0-18, age 19-30, age 31-65, age 66+, high school dropouts, with high school diploma, with some college education, and with bachelor's degree. *p<0.1; **p<0.05; ***p<0.01

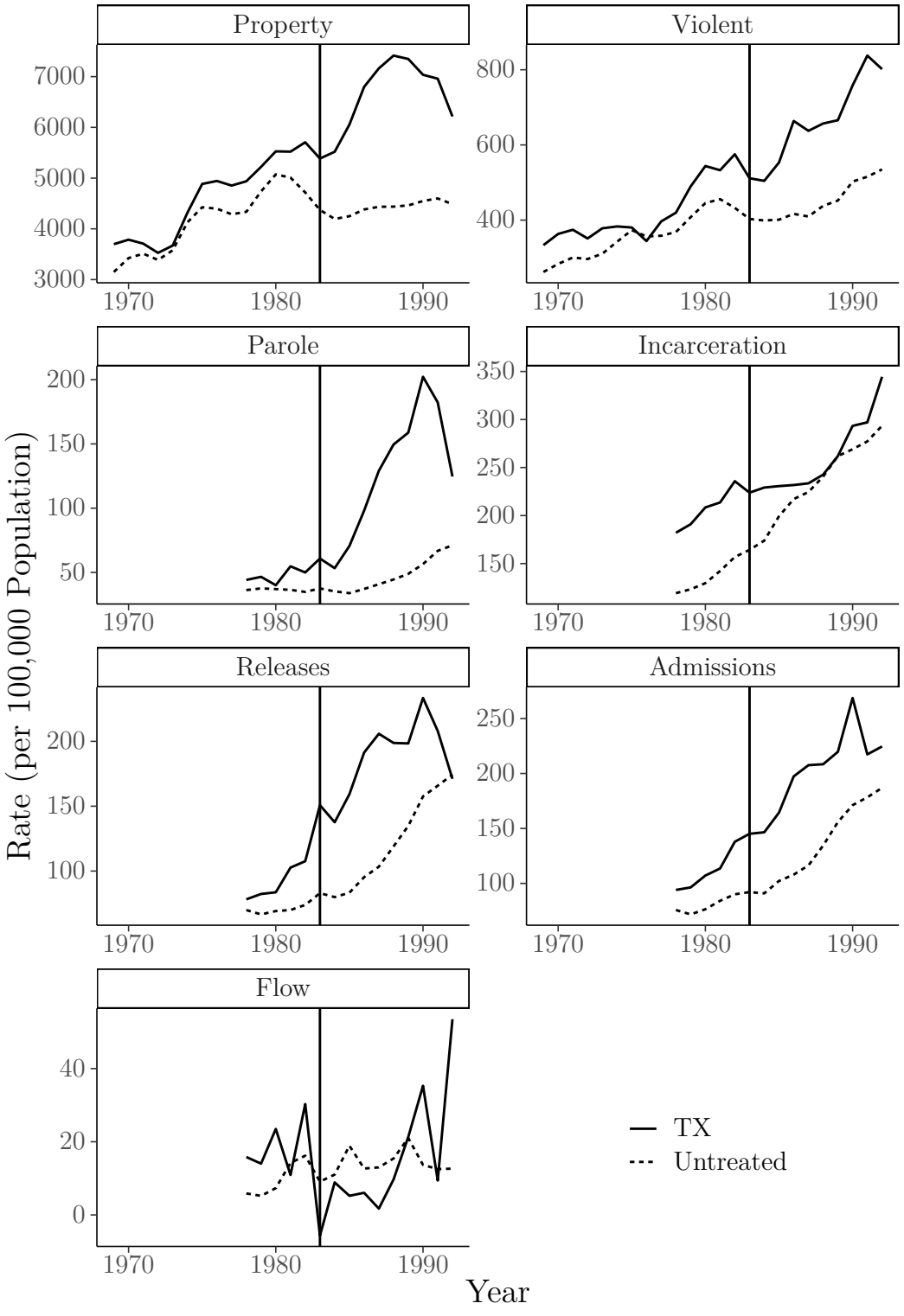

Figure 2.1. Trend Plots

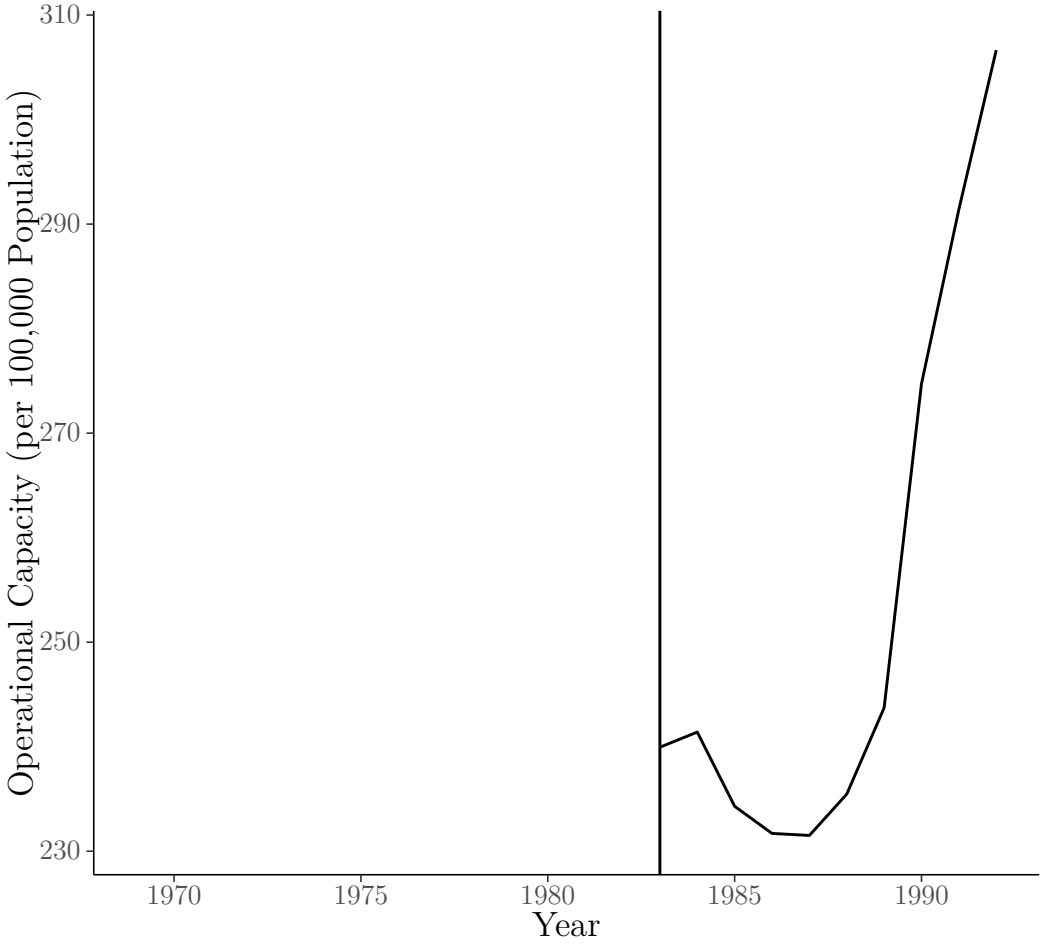

Figure 2.2. Capacity of Texas Prison System

CHAPTER THREE

Synthetic Control

In this section, I will first describe the synthetic control estimator and the conditions under which Abadie, Diamond, and Hainmueller (2010) show it to be approximately unbiased. Then I will discuss important developments by Ferman and Pinto (2019b) regarding the estimator's asymptotic properties. Next I will draw attention to an important difference in modeling assumptions between DID and synthetic control, which I believe has been under-discussed in the literature. Finally, I will present the results of synthetic control estimation for my particular application.

The Estimator

In case-studies such as ours, it is often the case that the average of the untreated units do not provide the ideal control group. However, it may be the case that some other convex combination of control units better resembles the treated unit along the relevant dimensions than the simple average. This convex combination is called a "synthetic control."

To make the above more formal, suppose the outcome of interest in the absence of treatment can be modeled by the following factor model:

$$Y_{it}^N = \delta_t + \boldsymbol{\theta}_t \boldsymbol{Z}_i + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \nu_{it},$$

where Y_{it}^N is the outcome of interest for unit i at time t if untreated, $\boldsymbol{Z}_i$ is a $r \times 1$ vector of *time-invariant* observable covariates, $\boldsymbol{\theta}_t$ is a $1 \times r$ vector of parameters, $\boldsymbol{\mu}_i$ is a $F \times 1$ vector of unobserved factors, $\boldsymbol{\lambda}_t$ is a $1 \times F$ vector of unknown factor-loadings, and ν_{it} is a mean-zero transitory shock. Abadie, Diamond, and Hainmueller (2010) prove that, if a convex combination of untreated units can be found with weights

$\boldsymbol{W}^* = \begin{bmatrix} w_2^* \\ w_3^* \\ \vdots \\ w_N^* \end{bmatrix}$ such that, for treated unit $i = 1$

$$Y_{1t} = \sum_{i=2}^{N} w_i^* Y_{it} \tag{3.1}$$

and

$$\boldsymbol{Z}_1 = \sum_{i=2}^{N} w_i^* \boldsymbol{Z}_i \tag{3.2}$$

for $t \leq T_0$, then, given certain regularity conditions, the bias of the estimator,

$$\hat{\alpha}_t = Y_{1t} - \sum_{i=2}^{N} w_i^* Y_{it}, \qquad\qquad t \geq T_0$$

is bounded by a function that goes to zero as T_0 increases. The expression $\sum_{i=2}^{N} w_i^* Y_{it}$ is the synthetic control in year t. The intuition behind the above result is that the only way equations 3.1 and 3.2 hold over a large number of pre-treatment periods is if

$$\boldsymbol{\mu}_1 \approx \sum_{i=2}^{N} w_i^* \boldsymbol{\mu}_i.$$

The condition that the linear combination of untreated states must be convex prevents extrapolation outside the support of the data.

In practice, however, it is almost never the case that weights can be found such that equations 3.1 and 3.2 hold exactly. Consequently, Abadie, Diamond, and Hainmueller (2010) intuit that it should be sufficient to find weights such that these conditions hold approximately. They then reccomend the following implementation. Let $\boldsymbol{Q}_1$ be a $J \times 1$ vector of potentially relevant covariates and optionally some lags of the outcome of interest for the treated unit. Let $\boldsymbol{Q}_0$ be the corresponding $J \times (N-1)$ matrix for the control units. Let V be some $J \times J$ symmetric, positive semi-definite

matrix. Then, choose $\boldsymbol{W}^*$ such that

$$\boldsymbol{W}^* = \arg\min_{\boldsymbol{W}} \sqrt{(\boldsymbol{Q}_1 - \boldsymbol{Q}_0\boldsymbol{W})^\top \boldsymbol{V}(\boldsymbol{Q}_1 - \boldsymbol{Q}_0\boldsymbol{W})}. \tag{3.3}$$

Now obviously our choice of $\boldsymbol{W}^*$ will depend on our choice of $\boldsymbol{V}$. Abadie, Diamond, and Hainmueller (2010) recommend choosing $\boldsymbol{V}$ such that the pre-treatment root-mean-squared-prediction-error for the outcome variable is minimized.

The rationale behind this procedure is very intuitive. We would like to choose the weighting vector $\boldsymbol{W}^*$ that minimizes the distance between the treated unit and the synthetic control in pre-treatment outcomes and in covariates simultaneously. But it is clear that in this minimization we will have to prioritize some covariates over others. Ideally, we would like to prioritize matching those covariates which are most predictive of the outcome. That is the function of the weighting matrix $\boldsymbol{V}$; it allows us to weight different covariates more heavily than others in calculating the distance which is being minimized in equation 3.3.

However, despite the fact that this procedure was originally designed to match on covariates *and* pre-treatment outcomes, many researchers have declined to match on covariates altogether, and have instead set $\boldsymbol{Q}_1 = \boldsymbol{Y}_1$ and $\boldsymbol{Q}_0 = \begin{bmatrix} \boldsymbol{Y}_2 & \boldsymbol{Y}_3 & ... & \boldsymbol{Y}_N \end{bmatrix}$, where $\boldsymbol{Y}_i$ is the $T_0 \times 1$ vector of pre-treatment outcomes for unit i. In this case, the ideal weighting matrix $\boldsymbol{V}$ is the $J \times J$ identity matrix, and the optimization problem reduces to

$$\boldsymbol{W}^* = \arg\min_{\boldsymbol{W}} \frac{1}{T_0} \sum_{t=1}^{T_0} [Y_{1t} - \sum_{i=2}^{N} w_i^* Y_{it}]^2.$$

While it is intuitive to suspect that this version of the estimator should still be approximately unbiased as long as equation 3.1 holds (since the term $\boldsymbol{\theta}_t \boldsymbol{Z}_i$ can simply be pushed into the $\boldsymbol{\lambda}_t \boldsymbol{\mu}_i$ term), this was not proven rigorously until Botosaru and Ferman (2019). The function that bounds the bias in this case will still go to zero as T_0 increases, but these bounds will generally by wider than if both equations 3.1 and 3.2 hold.

Important Developments

Notice that the proof for the approximate unbiasedness of the synthetic control estimator relies on exact equality in equations 3.1 and 3.2. As mentioned above, Abadie, Diamond, and Hainmueller (2010) intuit that, as long as $\boldsymbol{W}^*$ can be found such that equations 3.1 and 3.2 hold *approximately*, and as long as T_0 is large enough, then the synthetic control estimator will still be approximately unbiased. Ferman and Pinto (2019b) show, however, that this intuition is not, in general, correct. It instead turns out that the exact equality of these conditions is in many cases essential to the desirable qualities of the estimator. In the case with imperfect fit, Ferman and Pinto (2019b) show that, if treatment is correlated with unobserved heterogeneity, then the synthetic control estimator will be asymptotically *biased* as $T_0 \to \infty$. The reason is because, under these conditions, $\sum_{i=2}^{N} w^* \boldsymbol{\mu_i}$ is inconsistent for $\boldsymbol{\mu}$.[1] Thus, the synthetic control estimator, as implemented by Abadie, Diamond, and Hainmueller (2010), is asymptotically biased even in cases where DID is unbiased, since, in DID, treatment is allowed to be correlated with time-invariant heterogeneity.

Despite these shortcomings, Ferman and Pinto (2019b) argue that synthetic control need not be abandoned entirely. Instead, they propose an alternate version of the estimator, which they attribute to Doudchenko and Imbens (2016), which involves demeaning the outcome variable by its pre-treatment mean for each unit (similarly to a fixed-effects estimator) and estimating a synthetic control by minimizing pre-treatment RMSPE only, without matching on covariates. This estimator, they show, will be unbiased (asymptotically, as well as in finite samples) so long as treatment is not correlated with *time-varying* heterogeneity. This is similar to the identifying assumption for DID. The synthetic control estimator is, however, asymptotically more efficient than DID (Ferman and Pinto 2019b). Therefore, this is the version of the synthetic control estimator that I will be using in my empirical application.

1. The reason for this inconsistency is beyond the scope of this paper.

Inference

Abadie, Diamond, and Hainmueller (2010) propose the following inference procedure. First, synthetic controls are estimated for all units, treated and untreated. Next, the root-mean-squared-prediction-error (RMSPE) of the synthetic control relative to the actual pre-treatment and post-treatment outcomes are calculated for each unit. Finally, the ratio of post-treatment RMSPE to pre-treatment RMSPE is calculated for each unit, and these ratios are ranked. If this ratio for the treated state is extreme relative to the others, then the null hypothesis is rejected. Under the hypothesis of equally likely treatment assignment, permutation p-values can be calculated as the proportion of pre/post-treatment RMSPE ratios greater than or equal to that of the treatment state.

The rationale behind this procedure is very simple. If a treatment effect is present, the post-treatment RMSPE of the treated unit should be very large relative to that of the untreated units. However, if the pre-treatment fit for the synthetic control is poor, we might expect to find a large post-treatment RMSPE by default. Dividing the post-treatment by the pre-treatment RMSPE corrects for this latter possibility.

To harmonize the statistical inference in this section with that in the last, I decline to use this RMSPE-based test statistic in favor of estimates of the treatment effect size. However, to correct for the possibility of a poor synthetic control fit for a control state causing a spuriously large treatment effect estimate for that state, I drop all states from the comparison that have a pre-treatment RMSPE of more than two times that of Texas.

Before proceeding to our results, I would like to first discuss an important difference between DID and synthetic control that I believe has been under-discussed in the literature.

Comparison with Difference-in-Differences: A Digression

It is important to note that the data-generating mechanism is seemingly modeled quite differently for synthetic control than it is for the DID estimator. Recall that the outcome variable under no treatment in the DID context is modeled as

$$Y_{st}^{N} = \beta_0 + \boldsymbol{\gamma}\boldsymbol{X}_{st} + \tau_t + \sigma_s + \epsilon_{st}$$

The linear-factor model allows for more flexibility in the unobservables than the standard linear data model. This is because, in the standard model, time-specific and individual-specific effects are constrained to be additive, whereas in a linear-factor model they are allowed to interact. If one wishes to constrain the interactive fixed-effects to be additive, one can simply set $\boldsymbol{\lambda_t} = \begin{bmatrix} \tau_t \\ 1 \end{bmatrix}$ and $\boldsymbol{\mu_i} = \begin{bmatrix} 1 & \sigma_{s=i} \end{bmatrix}$. This insight prompted Abadie, Diamond, and Hainmueller (2010) to claim that the synthetic control estimator is a generalization of the DID estimator. This is not strictly true, however, because, although the unobservables have more flexibility under this factor structure, the structure of the observed covariates is restricted.[2] Notice that in the factor model, the observables $\boldsymbol{Z}_i$ are not indexed by time. Therefore, any observed covariates must be fixed in time. At first, this may seem a crippling constraint. This assumption is not as restrictive as it may at first appear, however. It is actually possible, with a few assumptions, to write the fixed-effects model as a factor model with the structure laid out above. First let

$$\boldsymbol{\lambda_t} = \begin{bmatrix} \tau_t \\ 1 \end{bmatrix}$$

and

$$\boldsymbol{\mu_i} = \begin{bmatrix} 1 & \sigma_{s=i} \end{bmatrix}$$

2. Abadie, Diamond, and Hainmueller (2010) note that the synthetic control estimator is still approximately unbiased under more general settings than the linear-factor model. They also give an AR-1 model as one of these other possible settings. My contribution is to explicitly lay out how certain common models can be converted to a linear-factor model, showing that synthetic control is still valid for data generating mechanisms that do not at first appear to conform to the assumptions of the method.

as above. Then impose,

$$\boldsymbol{X}_{st} = \boldsymbol{\alpha}_t \boldsymbol{Z}_i + \boldsymbol{\rho}_{st},$$

where $\boldsymbol{\alpha}_t$ is a $k \times r$ matrix, $\boldsymbol{\rho}_{st}$ is a $k \times 1$ vector of stochastic errors and $s = i$. Finally set

$$\boldsymbol{\gamma}\boldsymbol{\alpha}_t = \boldsymbol{\theta}_t,$$

$$\delta_t = \beta_0,$$

and

$$\epsilon_{st} = \boldsymbol{\gamma}\boldsymbol{\rho}_{st} + \nu_{it},$$

where $s = i$. Under these conditions, the two models are equivalent.

Since it was necessary to impose conditions on on both models to make them equivalent, it is apparent that neither is a generalization of the other. But the most important takeaway from the above exercise is that minimal structure was needed to convert time-varying covariates to non-time-varying ones, and vice-versa. The results above can be used to illustrate how several common models for covariates can be written with a linear-factor structure with time-varying coefficients but time-invariant variables.

Suppose $\boldsymbol{X}_{st}$ can be modeled by an AR-1 process:

$$\boldsymbol{X}_{st} = \boldsymbol{\phi}\boldsymbol{X}_{s(t-1)} + \boldsymbol{\omega}_{st},$$

where $\boldsymbol{\phi}$ is a $k \times k$ diagonal matrix of coefficients. Then, by recursive substitution

$$\begin{aligned}
\boldsymbol{X}_{st} &= \boldsymbol{\phi}^2 \boldsymbol{X}_{s(t-2)} + \boldsymbol{\phi}\boldsymbol{\omega}_{s(t-1)} + \boldsymbol{\omega}_{st} \\
&\vdots \\
&= \boldsymbol{\phi}^{t-m} \boldsymbol{X}_{sm} + \boldsymbol{\phi}^{t-m-1}\boldsymbol{\omega}_{s(m+1)} + \ldots + \boldsymbol{\phi}\boldsymbol{\omega}_{s(t-1)} + \boldsymbol{\omega}_{st}.
\end{aligned}$$

Take

$$\boldsymbol{\alpha}_t = \boldsymbol{\phi}^{t-m},$$

$$\boldsymbol{Z}_i = \boldsymbol{X}_{sm},$$

and

$$\boldsymbol{\rho}_{st} = \boldsymbol{\phi}^{t-m-1}\boldsymbol{\omega}_{s(m+1)} + \ldots + \boldsymbol{\phi}\boldsymbol{\omega}_{s(t-1)} + \boldsymbol{\omega}_{st}.$$

Now we are able to write our standard linear model as a linear-factor model with time varying coefficients and time invariant variables. Of course, we are not limited to modeling $\boldsymbol{X}_{st}$ as AR-1. The important thing is that $\boldsymbol{X}_{st}$ must be a function of some variables at a fixed point (or set of fixed points) in time, such that it is sufficient to know the value of of those variables at that (those) point(s) in time in order to know the value of $\boldsymbol{X}_{st}$ in expectation. However, it is not necessary to explicitly model $\boldsymbol{X}_{st}$, only to assume that it has a certain, fairly flexible structure.

Of course, such an explanation could also give a possible interpretation of the interactive fixed-effects themselves. This is especially relevant in my application, since I will not be matching on covariates, thus pushing any covariates that I *could* have matched on into the interactive fixed effects.

This section was meant to show that assuming a linear factor structure on the data generating mechanism is not very restrictive compared to the standard DID framework. With this assurance in hand, we will proceed with the results of our estimation.

Results

The results for synthetic control estimation can be found in Table 3.1 for incarceration related outcomes and in Table 3.2 for crime outcomes. As can be seen, paroles see a statistically significant increase. The estimates for property and violent crimes also remain positive and statistically significant at the 10% level, and are of similar magnitude to the corresponding DID estimates, although the property crime estimate is a little smaller.

We also still observe the same puzzling pattern of a statistically significant increase in paroles without a correspondingly significant increase in releases or decrease in incarceration. However, the estimates for incarceration outcomes are potentially not as reliable as the estimates for crime outcomes. This is because the number of observed pre-treatment periods for incarceration-related outcomes is only 5, whereas the number for crime outcomes is 15. Since the synthetic control estimator relies on asymptotic approximation as $T_0 \to \infty$, the estimates for incarceration-related outcomes are potentially biased. I am inclined to believe, despite the lack of significance, that Texas must have decreased their incarceration over the counterfactual, simply because the behavior of TDC in requesting a budget increase does not make sense if they did not anticipate it being a binding constraint. However, one could also tell a similar story to the one laid out in the last chapter, where parolees are more likely to commit crimes than other members of the potential criminal population. Again, it could also be the case that better conditions in Texas prisons made the prospect of prison less of a deterrent. If we accept the incarceration story, the results presented here amount to moderate support of the idea of the incapacitation effects of prison on criminals. If, on the other hand, we accept the story about paroles in the absence of a decrease in incarceration, our results could perhaps be taken as circumstantial evidence of an "anti-reformatory" effect of prisons on prisoners, such that those who have been through the prison system are more likely to commit crimes than those who have not. Finally, it could be that my results provide evidence that Texas prisons were more effective as deterrents to crime before the treatment than after.

Table 3.1. Incarceration Outcomes (Synthetic Control)

	Paroles	Incarceration	Releases	Admissions	Flow
Treatment	70.73*	−37.29	3.11	−4.03	−18.88
	p = 0.09	p = 0.48	p = 0.45	p = 0.67	p = 0.16
Implied % Change	119.68	-12.43	1.53	-2.09	-52.93

Note: P-vaules are calculated as stated in the text. All outcomes measured per 100,000 population per year. *p<0.1; **p<0.05; ***p<0.01

Table 3.2. Crime Outcomes (Synthetic Control)

	Property	Violent
Treatment	848.90*	135.34*
	p = 0.09	p = 0.08
Implied % Change	14.45	25.06

Note: P-vaules are calculated as stated in the text. All outcomes measured per 100,000 population per year. *p<0.1; **p<0.05; ***p<0.01

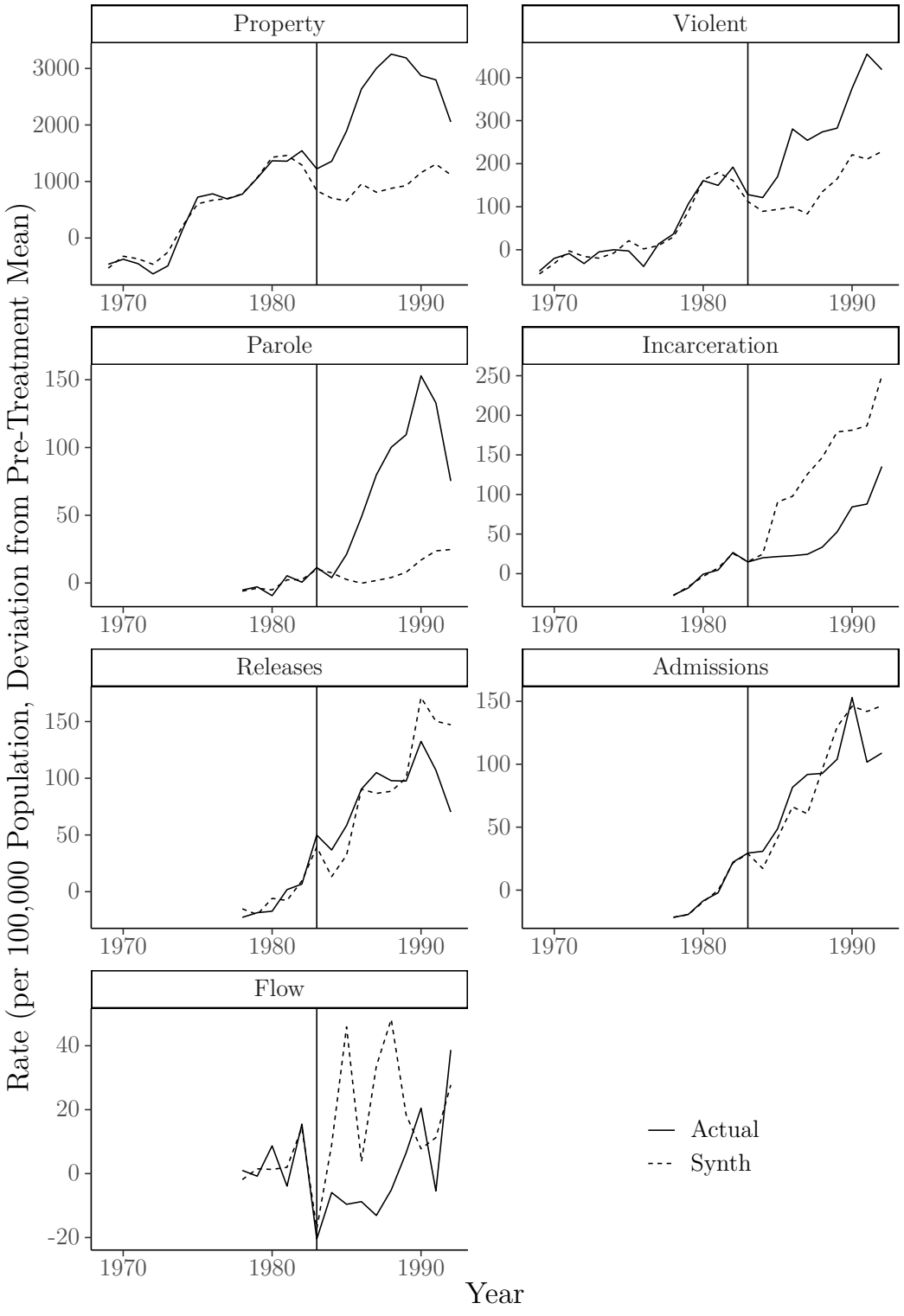

Figure 3.1. Actual vs. Synthetic Texas

CHAPTER FOUR

Conclusion

In conclusion, I find suggestive evidence that *Ruiz v. Estelle* and its subsequent legislative intervention did, in fact, substantially increase the number of discretionary paroles in Texas, as well as both the property and violent crime rates. However, no decrease in incarceration was detected. I am inclined to believe that the failure to reject the null hypothesis of no effect with regard to total incarceration represents Type II error. However, I have also provided two potential explanations for how crime could increase without a decrease in incarceration: namely, the "anti-reformatory" effect of prisons and the lessening deterrent effect of better prison conditions.

There are a number of potential concerns with my results, however. If treatment is correlated with time-varying unobservables, my estimates of the treatment effect could be biased. For example, if it turns out that there was a Texas-specific shock to prison conditions that brought about the ruling, then my estimates would no longer be valid. I have not, however, identified a suitable proxy for prison conditions such that I could address this concern by controlling for them. Controlling for prison conditions would also only unbias the DID estimates, since the version of the synthetic control estimator I use in this paper cannot directly control for time-varying covariates. Finally, the permutation-based inference that I employ is problematic in case-studies. Nevertheless, I believe that I have provided sufficient evidence to warrant additional investigation into this question.

APPENDICES

APPENDIX A

Synthetic Control Weights

Table A.1. Synthetic Control Weights (Paroles)

State	Weight
SD	0.130
VA	0.110
WI	0.340
WY	0.420

Table A.2. Synthetic Control Weights (Incarceration)

State	Weight
DC	0.170
NC	0.530
VT	0.240
WY	0.050

Table A.3. Synthetic Control Weights (Releases)

State	Weight
AK	0.920
CA	0.080

Table A.4. Synthetic Control Weights (Admissions)

State	Weight
AK	0.220
CT	0.070
DC	0.040
NV	0.560
VA	0.110

Table A.5. Synthetic Control Weights (Flow)

State	Weight
DC	0.340
MI	0.120
NC	0.530

Table A.6. Synthetic Control Weights (Property Crime)

State	Weight
DC	0.020
NC	0.830
SC	0.140

Table A.7. Synthetic Control Weights (Violent Crime)

State	Weight
DC	0.080
NJ	0.420
NC	0.060
SC	0.110
VT	0.330

BIBLIOGRAPHY

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." Journal of the American Statistical Association 105(490): 493–505. doi: http://amstat.tandfonline.com.ezproxy.baylor.edu/loi/uasa20.

Botosaru, Irene, and Bruno Ferman. 2019. "On the role of covariates in the synthetic control method" [in en]. Publisher: Oxford Academic, The Econometrics Journal 22, no. 2 (May): 117–130. Accessed March 14, 2020. doi:10.1093/ectj/utz001. https://academic.oup.com/ectj/article/22/2/117/5303851.

Buchmueller, Thomas C., John DiNardo, and Robert G. Valletta. 2011. "The Effect of an Employer Health Insurance Mandate on Health Insurance Coverage and the Demand for Labor: Evidence from Hawaii." American Economic Journal: Economic Policy 3 (4): 25–51. https://www.jstor.org/stable/41330440.

Bureau of Justice Statistics. 1960-2014. "Uniform Crime Report: UCR Data Online." United States Department of Justice. https://www.ucrdatatool.gov/index.cfm.

Doudchenko, Nikolay, and Guido W. Imbens. 2016. "Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis." Publisher: National Bureau of Economic Research, Inc, NBER Working Papers: 22791.

Ferman, Bruno, and Cristine Pinto. 2019a. "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity." Review of Economics and Statistics 101 (3): 452–467. http://ezproxy.baylor.edu/login?url=http://search.ebscohost.com/login.asp?direct=true&db=ecn&AN=1785220&site=ehost-live&scope=site

Ferman, Bruno, and Cristine Pinto. 2019b. Synthetic Controls with Imperfect Pre-Treatment Fit. arXiv: 1911.08521 [econ.EM].

MacKinnon, James G., and Matthew D. Webb. 2017. "Wild Bootstrap Inference for Wildly Different Cluster Sizes." Journal of Applied Econometrics 32 (2): 233–254. doi:10.1002/jae.2508. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.2508. https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2508.

Perkinson, Robert. 2010. Texas Tough: The Rise of America's Prison Empire. New York: Metropolitan Books.

Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D030.V7.0.

Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) Research Data (1975-2016), National Cancer Institute, DCCPS, Surveillance Research Program, released April 2019, based on the November 2018 submission.

United States Bureau of Justice Statistics. 2020. About the UCR Data Tool: Methodology. Accessed March 27. https://www.bjs.gov/ucrdata/data/methoducrdatatool.doc.

United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. National Prisoner Statistics, 1978-2016. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2018-09-14. https://doi.org/10.3886/ICPSR37003.v1.