Brown v. Plata (2011): Mental Health Outcomes and Prison Reform

Bennet Christensen

Leslie Myint, Advisor

Vittorio Addona, Reader

Paul Dosh, Reader

Department of Mathematics, Statistics, and Computer Science

May, 2024

Copyright © 2024 Bennet Christensen.

The author grants Macalester College Library the nonexclusive right to make this work available for noncommercial, educational purposes, provided that this copyright statement appears on the reproduced materials and notice is given that the copying is by permission of the author. To disseminate otherwise or to republish requires written permission from the author.

Abstract

In 2011 the Supreme Court ruled that California prisons were so overcrowded that they couldn't provide adequate health care. Subsequent research on the reductions in the prison population, known as Public Safety Realignment, focused on crime rates and how much money it would save the state instead of inmate wellbeing. This project tries to answer what should have been everyone's first question, which is: what effect did the court case have on the mental health of inmates? To do this, individual level survey data from the Bureau of Justice Statistics (BJS), namely the 2004 Survey of Inmates in State and Federal Correctional Facilities, and the 2016 Survey of Prison Inmates was used. The National Prisoner Statistics Program, also from the BJS was used to provide prison level population data. Additionally, to analyse changes in jails the time series data from California Board of State and Community Corrections was used in the form of the Jail Profile Survey. Informed by causal inference methodology, we used an ordinal logistic regression in a difference-in-differences framework to determine the effect of California's realignment policy on inmate mental health. We additionally used linear models under similar frameworks to add robustness to our findings, and compared results under different interpretations of mental health screener results. We then used a linear mixed effects model to look at the association between the realignment policy and mental health case rates in California Jails. We found that at best realignment moderately improved the mental health of inmates in prisons and at worst did not significantly improve mental health in prisons. However, for jails, realignment was associated with increased mental health case rates, indicating that some of the mental health burden was just potentially passed off from one institution to another.

Contents

Introduction

In 2009, the state of California had 171,275 individuals incarcerated in their prisons. By 2011, that number fell to 149,569. These numbers tell the most basic note of the story of California's "realignment," a series of lawsuits and legislation that reshaped criminal justice in the state. California is the most populous state in the union so it might be hard to have a frame of reference for how many people 171,275 actually is. In 2009, California had roughly 37 million residents, which means that about 0.5% of the state's population was facing incarceration in state prisons. The combined capacity of those state prisons was around 80,000, which meant that they were running at around 200% capacity. This overcrowdedness is not something new to 2009, instead being an ongoing legislative and judicial problem since the 1980s.

In 1990, Coleman v. Brown, a class-action case against the state on behalf of California inmates was filed. It alleged that the California Department of Corrections and Rehabilitation (CDCR) was violating the eighth amendment's protection against cruel or unusual punishment by providing insufficient mental health care, especially in relation to overcrowdedness. They alleged that there was not enough competent medical staff or medication to provide the care required by inmates, exacerbated by a whole set of inadequacies in medical service. In 1995 the district court presiding over that case ruled in the inmates favor, finding California to be in violation of the eighth amendment, and deciding that California should remedy the situation. The court also assigned a "special master" to routinely check in on the progress the CDCR was making. While the government, both state and federal, has no constitutional protections for receiving health care, it is important to remember that by imprisoning individuals, the state makes themselves responsible for the wellbeing of those they imprison. The precedent for this was set in 1976 in the supreme court case, Estelle v. Gamble [1].

In 2001 a similar case, Plata v. Brown, was filed, this time alleging that a lack of quality medical services was in violation of prisoners' eighth amendment rights, expanding the issue past mental health. This case alleged, much like the previous one, that there was insufficient competent staff to accommodate the medical needs of inmates, and that lack of proper, available medical care was leading to harm, and for some inmates, death. Both sides negotiated an injunction that required the CDCR to meet at the minimum, medical care good enough to not violate the eighth amendment.

In 2005, the same court that oversaw Plata v. Brown checked in on the progress made by the CDCR. They found that they had failed to significantly improve medical care or mental health care, which put them in violation of those previous decisions. No longer trusting the CDCR to fix their constitutional violations, the entire prison health care system was put under receivership, meaning an independent body was now in charge of bringing prison medical care up to standard, taking autonomy away from the CDCR.

In 2006, a three judge court was requested to decide whether the state should be forced to reduce its prison population. Following the request, then Governor of California, Arnold Schwartzenager, issued a state of emergency over the overcrowding of state prisons, citing increased risk of violence, riots, infectious disease, and damage to the prisons themselves. He instructed prisons to transfer inmates to other states to help ease the overcrowding in California. In 2009, after the court was finally assembled, it ruled that the state of California must reduce overcrowdedness down to 137.5% of capacity in two years.

In 2011 the state of California appealed the decision to the supreme court in an attempt to avoid the mandate. In that supreme court case, Brown v. Plata, the highest court in the country sided with the previous court and upheld the release order [2]. After over 20 years of resisting improvements and solutions, the state of California had no options left but to reduce their prison overcrowding. That same year the state passed AB 109, called "Public Safety Realignment," that shifted around tens of thousands of inmates from prisons to jails. However, they would not reach sufficient population reductions until 2018, but this law made up the largest reduction [3].

Much of the literature has been focused on the costs associated with releases[4] [5], and how it would affect crime rates[6] [7]. Instead of focusing on those outcomes, this research focuses on mental health, one of the two problems being caused by overcrowdedness, as work has been done on some of the physical aspect of the health problem [8].

Data

Data on Prisons

Data on prisons in general, and on health conditions within prisons especially, is very limited, and rarely public. This analysis uses three data sets, the 2004 Survey of Inmates in State and Federal Correctional Facilities[9] (n=14499), the 2016 Survey of Prison Inmates[10] (n=20064), and the National Prisoner Statistics Program[11] (n=650). The surveys from 2004 and 2016 were used in the primary analysis, while the National Prisoner Statistics Program was used to select control states for the analysis.

Main Survey

For the statistical modeling of mental health in prisons, this study uses the 2004 Survey of Inmates in State and Federal Correctional Facilities and the 2016 Survey of Prison Inmates. These surveys are part of the same series that was renamed in 2016, hence the difference. They are cross sectional surveys collected via face-to-face interviewing. While the surveys are conducted in state and federal prisons, this study only uses data from the state prisons.

In order to preserve the anonymity of survey participants, precise information on where an inmate is incarcerated is not used in this analysis. With the exact facility not being known, the exact amount of overcrowding an inmate is facing is also not known. Not only that, but the exact state of incarceration is also not used. Instead, state of residence before incarceration is used as a proxy. The relationship between state of residence and state of incarceration is fairly high, so while using this proxy does misclassify a small number of inmates and introduce more error into our analysis, it should not invalidate results. While this data is accessible online, it is important to preserve anonymity of inmates, especially considering the vulnerability of the population.

The questionnaire was broken into about 5 sections, seen in the chart below. These are demographic variables, variables relating to the offenses and the sentencing process, socioeconomic variables, mental health variables, and physical health variables. We use the age, race, marital status, and veteran status variables from the demographic section, but there are additional military service questions that we opt to not use. As mentioned before, from the Offense and sentencing variables we use the state of imprisonment as a proxy for the state that an inmate is held. From the socioeconomic variables we use educational attainment. What allows this study to exist is the fact that the survey included mental health screener questions, specifically the Kessler 6 Screener. We do have some other information like prescription medication and access to therapists, but we opt for the mental health screeners give a broader picture of mental health conditions as a whole. We don't use any of the physical health questions because we experience a small data issue with most of them, and have a lot of problems with differences between reporting in 2004 vs 2016.

Variable Group	Used	Unused
Demographic	Age, Race, Marital Status, Veteran Status	Questions about length, branch, and discharge from military service,
Offense and Sentencing	State Imprisoned	number, type of, and new offenses, when, by who, were you sentenced by
Socioeconomic	Educational Attainment	Learning disabilities, citizenship, children, past employment, income
Mental Health	Mental Health Screeners	prescriptions, counseling/therapy, and hospital usage
Physical Health	None	Height, weight, sex, gender, sexuality, doctors visits, physical disabilities, health conditions

However, not all of these variables were nicely available in both studies. We construct a year variable, which just indicates which survey each response came from. Educational attainment and age are categorical in the 2016 survey, so numerical values provided in 2004 are coerced into the categories present in 2016. Race and marital status are roughly the same between the two surveys, and are categorical variables. This chart below provides an overview of the varaibles that are available in the data set and which ones were actually used.

The composite mental health score is also a constructed variable, coming from different variables from the mental health screeners. Specifically, questions about nervousness, hopelessness, restlessness, depression, everything being an effort, and worthlessness were used. These questions are all based on the Kessler Psychological Distress Scale, which is commonly reported on a five point scale, with responses ranging from none of the time to all the time. However, in 2004 respondents were instead given a more comprehensive set of questions, including the six from the Kessler Psychological Distress Scale, but only given the option of answering yes or no. For this reason all scores have been converted to be a 1 for yes or a 0 for no. This forces a loss of nuance as the five point scale must also be converted to a binary. Answering “none of the time” or “a little of the time” is always interpreted as a 0 and “most of the time” or “all of the time” is always interpreted as 1, but when it comes to the response of “some of the time” the interpretation is less clear. To account for bias that could be introduced by different interpretations of “some of the time,” the modeling is done twice, once under the assumption that “some of the time” is 1, which will henceforth be referred to as the lenient interpretation, and once under the assumption that “some of the time” is 0, which will henceforth be referred to as the strict interpretation.

This difference in interpretation may not seem that important to the potential results of this study, but it is important to understand the distribution of answers to the mental health screeners in 2016. In four of the six questions, responding "some of the time" is the second most common answer, and in the remaining two it is the third most common. Not only that but it typically has as many responses as "all of the time" and "most of the time combined," so how "some of the time" is interpreted has a large sway on the proportion of the inmate population identified as having poor mental health. This makes it important to be thorough in examining the implications of each interpretation, even if intuition points more towards one interpretation.

Controls

The National Prisoner Statistics Program is a program by the Bureau of Justice Statistics that collects state level year-end prison population count data from. It is accessed through the Corrections Statistical Analysis Tool. It has year-end population numbers for each state’s prison system. These counts are initially stored in wide format, with a row for each state and a column for each year, from 1983 to 2019. After pivoting and reducing the time frame down to 2004 to 2016, the same years as the prison surveys, the data set is left with 650 observations. An additional variable is created for each year, comparing each year's population size to that state's population size in 2004. These changes in populations since 2004 are then used to select control states for the study.

Data on Jails

All of this data lets us answer questions about those that remained in the California prison population, but what about those who were moved out of prisons in order to reduce the crowding? For those who were put on parole or released early, we can't track without further invading their privacy, but for those placed in jail not only does data already exist, but we are also additionally interested in these people because they face a very similar overcrowding issue. The California Board of State and Community Corrections (BSCC) has been conducting a quarterly Jail Profile Survey [12] for the last fifty years. Each county records data about their jails and submits this data to the BSCC which then publishes the data all together. With the data being monthly for each individual jail, we can view it as longitudinal data.

So what does all this data look like? The version of the data set used in this analysis spans from 2002 to 2019, with 12,945 total observations from 56 different jails. However, jail policy has seen additional changes that have changed jail crowdedness. In 2014 the state legislature passed AB 2499 which amended the state penal code to expand programs that provided inmates the opportunity to serve their sentence under house arrest instead of in jail. Additionally variables of interest weren't reported for the whole time period, so our time period of interest is from 2007 to 2014, which reduces our total observations down to $5,568$. There are 53 different variables, namely the reporting jurisdiction, variables for when the data was collected, demographic counts of who was held in that jail that month (men, women, non-sentenced, sentenced, misdemeanor, security level), variables specific to rules or laws, and variables about health in the jails.

This last set of variables is the most important to us. Specifically, we are interested in the number of new mental health cases opened in the last month. This gives us an idea about the volume of mental health problems in any given jail. Unfortunately it does lack some nuance as not all mental health cases that are opened are equally pressing, and it doesn't preserve as much information about past mental health cases, so our observations at our earliest time point are probably underestimating the severity of mental health, as we don't have information about the number of preexisting mental health cases. Some other options would be mental health cases opened on the last day of the month, inmates receiving psych medication on the last day of the month, and inmates assigned to mental health beds on the last day of the month, but these aren't preferred because they all only provide information about the last day of the month, which isn't necessarily representative of the entire month.

There are some additional considerations with working with the new mental health cases. Only 22 of the 56 jails have complete reporting of this variable. This reduces our observations down to 2,112. Additionally, these jails vary in size significantly, with the Los Angeles county jail being about three times larger than the next largest jail. With that being said, just using a count of new mental health cases is not super informative. To adjust for size, we divide new mental health cases by the total average daily population of the jail for that month, and are left with a mental health case rate. We can then model how the policy influenced case rates in the California jails with available data.

Methods

Causal Inference

The ideal but almost always impossible scenario for causal inference studies is to be able to observe the exact same population of interest (California inmates in this scenario) as they receive the treatment (reducing prison overcrowdedness) and as they don't. In the absence of being able to observe California without the policy change, we must turn to some sort of control group that hopefully approximates what would have happened if our treatment group hadn't received treatment. With control groups, the golden standard would be randomized control trials in an experimental setting, but for policy research, especially on a state-wide level, this is not possible, especially post hoc. Instead, this study implements a method called difference-in-differences.

Difference-in-Differences

Acknowledging that we can't find a state with identical starting conditions to California, we instead try to find states with similar trends in mental health. This is much easier to find than an identical state, but also imposes additional assumptions onto our model.

Conceptual Overview

Ideally a control will have as close to identical starting conditions to the treatment group as possible. However, in policy settings this is extraordinarily difficult to find. Instead of looking for identical starting conditions we instead look for identical starting trends. Specifically, the difference-in-difference model not only requires these parallel trends in outcome between the exposed group and the control group, but it additionally assumes that if the treatment had not been implemented they would have continued to have parallel trends. This is to say, in order to use this model, you assume that the outcome of interest would have change at the same rate in both groups had they both not received treatment. Not only that, but that change is assumed to be linear. While we would ideally check the trend in mental health between our treatment group and our control group, only having one time point before treatment prevents us from doing this. Instead we have to settle for the trend of overcrowdedness, which we assume to be a causal pathway to mental health.

Ordinal Logistic Model

Specifically, our difference-in-difference model is built on a generalized linear model (GLM) with a logit link function. Our outcome variable, the mental health score, despite being represented numerically, is actually categorical, which is why we rely on logistic regression. However, our model has to be a bit more complex because the categories are ordered. We care specifically about movements from a lower score to a higher score.

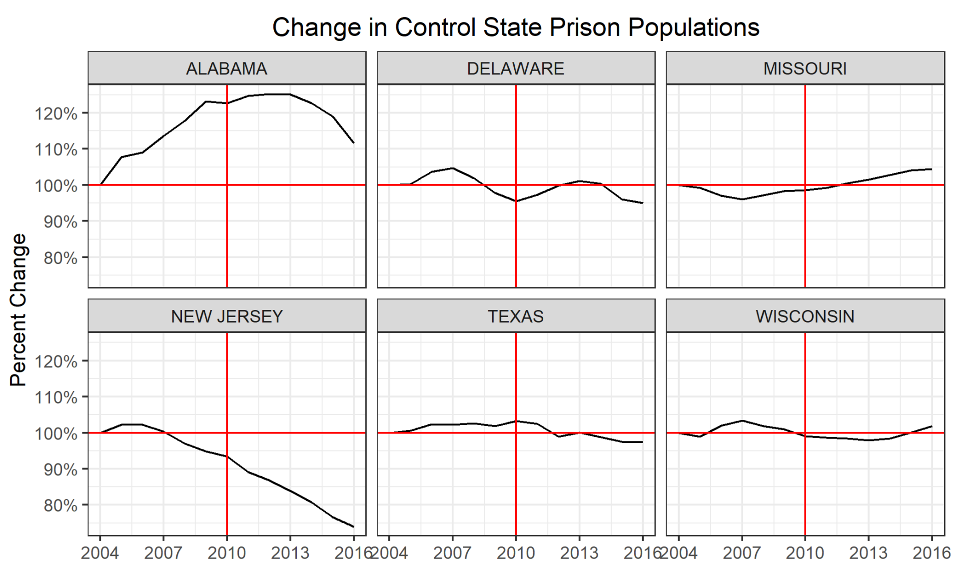
To account for this ordering we dichotomize our outcome variable. We look at the log odds of a group of worse mental health scores compared to a group of better mental health scores. On one end we have this "some vs none" scenario, where we compare those who had any indication of poor mental health, answering at least one screener question yes, compared to those who answered no to every single question. We continue to shift our thresholds until we are comparing the worst mental health, answering all six screeners yes, to any lower number of yeses. We then run a model for each comparison. This allows us to compare the average treatment effect in the treated at each level of mental health scores.

Controls

In order to satisfy model assumptions, namely parallel trends, and establish a baseline of comparison to account for national level changes, control states are necessary. Ideally this selection process would look for parallel trends in mental health scores, our outcome variable, but with only one set of observations on either side of the policy date, this is not possible. We already fundamentally assume some causal pathway between overcrowding and mental health, which is supported by the very supreme court decision we are trying to study in the first place. Instead of establishing a parallel trend in mental health, we instead establish a parallel trend in state prison populations, a setting with much more available data.

Specifically, we chose states that have maintained a stable prison population, similar to California during the aughts according to data from the National Prisoner Statistics Program. States that had changing prison populations due to state level policy changes could confound results, and are thus avoided. Take for example Alabama and New Jersey. From 2004 to 2011 Alabama dramatically increased their prison population, and thus have a competing policy that would make it a poor control state. Likewise, New Jersey dramatically decreased their prison population and thus would be a poor control because they also should have change in their mental health trend brought on by their reduction in prison population. This is problematic because we wouldn't be able to discern whether the relative change in mental health was caused more by the change in the bad control state, or the change in California as a result of realignment.

In the time period between the turn of the century and realignment, California prisons never varied more than 5.38% from their 2004 populations. Only four states were as stable from the 2004 to 2016, the year of our second survey. Those states are Delaware, Missouri, Texas, and Wisconsin, and are the states used as controls for this study. Only using these control states and California, our sample sizes are reduced to 4592 observations in 2004, and 4817 observations in 2016.



This method isn't perfect, as it fails to capture other policy changes within these states, such as the Texas Justice Reinvestment act of 2007. While this policy was popular, federal data shows that instead of greatly shrinking prison populations in the state, it mostly just held the prison population constant [13]. As we see with the data from the National Prisoner Statistics

Program, none of these policy changes were large enough to dramatically change inmate populations as much as California did before even enacting realignment. It is still important to be mindful of the fact that there can not be a perfect set of controls for the precise conditions that would have been present in California prisons, had they not gone through their realignment and that these conditions and policies in our control state could skew the treatment effect found in this analysis.

Confounding

An additional consideration with our difference-in-difference is how to account for confounding variables, especially considering that the demographic composition of the prisons changes over time. For privacy reasons, and due to the fact that people are always finishing and starting their sentences we can't match individuals' covariate values from 2004 to 2016. Additionally, because realignment shifted individuals out of prisons, and this shift wasn't equitable across demographic groups, our demographic covariates are time varying and potentially divergent when comparing our treatment group to our controls. Not only are confounding, but they violate our parallel trends assumption if we don't adjust for them. To make this adjustment we include the interaction between year and each covariate, to account for the changing relationship between our covariates and outcome over time[14] [15].

Linear Model

Traditionally, difference-in-difference models rely on linear regression. Assumptions in logistic models like proportional odds can complicate the interpretation of our average treatment effect on the treated, so to improve the robustness of our results, some additional methods from statistics literature were considered[16].

One approach for categorical outcomes in difference-in-differences is to leverage our parallel trends and assume that they are additive in nature. We dichotomize our mental health scores, much like in logit setting, but instead use a linear regression to compare the probabilities of the having relatively worse mental health instead of the odds. Additionally, our assumptions with a linear model are more straight forward given the assumptions we already made by using a difference-in-differences framework.

Another approach is to leverage the fact that the outcome variable is ordinal, and that there should be additive parallel trends and run a linear model of the outcome without dichotomizing. The data is not continuous, but there is an assumed linear trend. This model is advantageous because it allows the average treatment effect in the treated to be viewed as numerical changes in mental health scores, which are far more interpretable than changes in odds or probabilities.

Linear Mixed Effects Model

Linear Mixed Effects models are models that account for correlation between observations. They do this by including fixed effects for covariates assumed to not be correlated, and random effects for variables with correlation that needs to be accounted for [17]. This method is important to this study for handling the jail data. The jail data has repeated observations for different jails. These repeated observations are correlated with the other observations from the same jail, so they require a random effect. For the dates after our our treatment, and month, included to account for potential seasonality in observations, fixed effects are used instead.

Results

Prison Results

In total four ordinal model variations were run, using a linear model and a generalized linear model under the lenient interpretation of the mental health screener and under the strict interpretation of the mental health screener.

Ordinal Logistic Model

	Ordinal Logistic Model		
	Threshold	ATT	95% CI
Lenient Interpretation	score ≥ 1	0.68	(0.55, 0.82)
	score ≥ 2	0.71	(0.59, 0.86)
	score ≥ 3	0.65	(0.54, 0.80)
	score ≥ 4	0.75	(0.59, 0.94)
	score ≥ 5	0.78	(0.59, 1.03)
	score ≥ 6	0.62	(0.39, 0.99)
Strict Interpretation	score ≥ 1	0.75	(0.62, 0.91)
	score ≥ 2	0.79	(0.64, 0.97)
	score ≥ 3	0.80	(0.62, 1.02)
	score ≥ 4	0.88	(0.64, 1.20)
	score ≥ 5	1.02	(0.69, 1.51)
	score ≥ 6	0.88	(0.47, 1.63)

In this table, our estimate of interest, the number reported in this chart, is the average treatment effect on the treated (ATT), which for the ordinal logistic models is the log odds ratio between receiving the treatment and not receiving the treatment. To be more specific, these points here indicates the policy effect in California after realignment. It being less than 1 indicates that the odds of some mental health problems after realignment are lower now than if California had not undergone realignment. In the model this estimate represents the interaction between treatment group and year, specifically representing observations that are in the treatment group after the policy came into effect. For interpretability, these log odds ratio have been exponentiated.

Lenient Interpretation

Under the lenient interpretation, holding demographic factors constant, the ordinal logistic model found realignment had a moderate but significant improvement on inmates with less extreme mental health problems. The ATT when comparing those that answered yes to one or more of the questions to those who answered no to every question was 0.68, which is to say that their odds of being in the worse mental health group was 0.68 times lower than the odds of those who didn't benefit from realignment. The 95% confidence interval for the ATT does not include one so this is a significant result. This general trend of odds ratios that are less than one continues to be the case up until comparing those with mental health scores of five and above to those with mental health scores lower than five. For this model, the confidence interval for the ATT included one, so there is not a significant difference in the ratio of odds between those who are in the treatment group and those who aren't at this level of mental health. However, when comparing those who answered yes to struggling with all six things covered by the mental health screener to everyone else, the model once again find the ratio between the odds of those who received treatment compared to those who didn't to be less than one, with significance as the confidence interval does not include one.

Strict Interpretation

Under the strict interpretation of mental health scores the impact of realignment appears to be weaker, which is to be expected as more individuals are classified as having poorer mental health under this interpretation. However, When comparing the odds of answering one or more questions yes to answering no to every question, and answering two or more questions yes to answering no more than one question yes, the post-realignment California inmates had odds ratios of 0.75 and 0.79 respectively, indicating a reduction in odds of having worse mental health. However, for those with worse mental health, when comparing scores of three and above to those with scores below three there is no longer a significant improvement in the odds of having better mental health, as the odds ratio for those who benefited from realignment are not significantly different from one. This pattern continues for all higher mental health comparisons, indicating that under the strict interpretation, the improvement gained from realignment was only significant for those with relatively low mental health screener scores.

Dichotomized Linear Model

	Linear Model		
	Threshold	ATT	95% CI
Lenient Interpretation	score ≥ 1	-0.079	(-0.121, -0.038)
	score ≥ 2	-0.082	(-0.127, -0.037)
	score ≥ 3	-0.091	(-0.134, -0.047)
	score ≥ 4	-0.048	(-0.085, -0.010)
	score ≥ 5	-0.027	(-0.058, 0.003)
	score ≥ 6	-0.022	(-0.042, -0.003)
Strict Interpretation	score ≥ 1	-0.073	(-0.117 , -0.029)
	score ≥ 2	-0.030	(-0.072, 0.013)
	score ≥ 3	-0.019	(-0.056, 0.018)
	score ≥ 4	-0.002	(-0.033, 0.029)
	score ≥ 5	-0.002	(-0.026, 0.022)
	score ≥ 6	-0.004	(-0.018, 0.010)

In this table, our estimate of interest is still the ATT, but for the dichotomized linear model it is the change in probability of being in the poorer mental health group between those who experienced realignment and those who did not. This estimate still represents the interaction between treatment group and year. These probabilities are reported on a 0 to 1 scale, not a 0 to 100% scale.

Lenient Interpretation

The linear model under the lenient interpretation had very similar results. Inmates in post realignment California had a 0.048 percentage point lower chance of answering four or more questions as yes, which is to say that realignment significantly improved individuals chances of having improved mental health for individuals with good to moderate mental health. Much like with the logistic version of the model, this relationship vanishes when looking at the chance of answering five or more questions as yes, with the confidence interval, (0.058, 0.003) including 0. However, the improvement in mental health reappears when looking at the chance of answering six or more questions as yes, with an estimated 0.022 percentage point lower chance of answering yes to every screener question.

Strict Interpretation

The linear model has an even weaker indication of improvement under the strict interpretation. The post-realignment California inmates had a 0.073 percentage point lower chance of answering more than one question yes, compared to answering none of the questions yes, but at all higher mental health score comparisons there was no significant relationship between realignment and mental health scores.

Linear Model

When running the linear model without dichotomizing we continue to see fairly similar results. Under the lenient interpretation our ATT is -0.35, which indicates that the treatment improved mental health scores by 0.35. This result was significant, with a 95% confidence interval of (-0.52, -0.18). However, under the strict interpretation realignment is not associated with an improvement in mental health scores, with the 95% confidence interval including 0 (-0.28,0.01).

Summary of Prison Results

Across all models, inmates in California prisons post realignment were more likely to not indicate poor mental health than their counterparts in control states. However, across the board at very high levels of poor mental health, specifically answering five or more of the mental health screener questions as yes, realignment was not associated with an improvement in mental health. For the more moderate cases of poor mental health, the interpretation of answering “some of the time” plays a part in the story.

Jail Results

The linear mixed effects model implemented (n=2112) looked at mental health case rates as a function of the jurisdiction as a random effect to account for the correlation between observations from the same jail, the month of the observation to control for any seasonality as a fixed effect, and the time period, either before or after the realignment as a fixed effect, to analyse the association between the policy and mental health in jails. What it found was that holding the individual jail and time of year constant, realignment was associated with a 1.45% increase in mental health case rates. The t-value for the time period coefficient was 5.405. It is contentious as to what distribution these t-values for the model follows, and it is often unclear how many degrees of freedom should be used so reporting a p-value is not necessarily accurate [18]. However, given the size of the sample we can assume these t-values are approximately normal, and observe that because $5.405 > 1.96$ this result is significant at the $\alpha = 0.05$ level. Specifically, the 95% confidence interval for our estimate would be (1.44%, 1.46%).

Discussion

This analysis does suggest that California Public Safety Realignment did improve inmate mental health, at least for inmates with better mental health to begin with, the most prevalent group. However, for the most vulnerable individuals, those with poor mental health, realignment potentially didn't provide any improvement, depending on how the 2016 mental health screener is interpreted. Additionally, it is important to note that just because something is better, doesn't mean that the issue is solved. As noted throughout this paper, the decreases in prison populations was primarily achieved by shifting inmates from prisons to jails, which this analysis suggests was associated worsening mental health conditions in said jails. For those who were part of the shift it is entirely plausible that they received no benefit, moving from one overcrowded facility to another, facing additional bureaucracy in the move without having any improvement in living conditions.

Data Limitations

There are a few issues within the data that limit the results of the study. While privacy, especially in a setting as sensitive as this is important, not having information about what prison an inmate is specifically incarcerated at does weaken the results. Not all prisons are equally crowded, or have equal mental health care capabilities. Additionally, not all prisons saw the same level of population reduction, the court decision required that the prison system as a whole reached 137.5% capacity, not that every individual prison was reduced to that capacity, so it would be helpful to know how the worse mental health outcomes were distributed, as to better match each prisons' change in crowdedness with each prisons' change in mental health. It is still not unreasonable to assume that because the prison system as a whole saw the reduction, the differences in mental health average out and we can look at system wide changes in mental health. While prisons were selected randomly, with their probability of being selected being weighted by their size, the Bureau of Justice Statistics aimed for an equal number of surveys from each prison that was selected, which could make the sample non-representative of the state system. However, without being able to see which prisons were selected, and if there were any patterns of mental health specific to any prison that was selected, it is hard to verify whether or not this is true.

Only having two years to bound the policy is also limiting to the study. Especially with a difference-in-differences model, the study would ideally have more observations across time. Ideally the parallel trends assumption is verified through comparing outcome levels in the pre-treatment time period, but with only one year of pre-treatment data it is impossible to find a trend. This is why this study instead uses data about general prison population to verify parallel trends instead. Not only that, but timing of the Survey of Inmates in State and Federal Correctional Facilities in 2004 is also not ideal because of how far it is from treatment in 2011. While the crowding in 2011 was within 1% of what it was in 2004, there is fluctuation within that time period that could be important to account for. In general, having data from time periods closer to the treatment is preferable in causal inference studies because the more time has passed the more chance there is that some other event has happened that effected mental health. Ideally for a study like this, if there was data right before and right after treatment something like interrupted time series could be used instead of difference-in-differences, which would impose fewer assumptions on the model.

Societal Context

An additional difficulty in analyzing the effectiveness of this policy is legal interpretation. The Eighth Amendment of the constitution prohibits "cruel and unusual punishment," in this case inadequate medical and mental health care, but in interpreting the law the courts decided that overcrowding was the core cause of the problem, and in the final decision the state of California was required to reach a certain level of overcrowding, not a certain level of health quality. Giving a quantifiable, unambiguous goal is good for enforcement purposes because it makes it clear whether the state of California is complying or not, but it also reinforces the premise that the main problem is overcrowding, and not some other structural factor. Not only that, but the level of overcrowding that California was required to reach being over 100%, 137.5% to be precise, implies that overcrowding isn't inherently a cruel punishment. There is precedent for this, with Rhodes v. Chapman in 1981 ruling that double celling in Ohio prisons wasn't unconstitutional. In the eyes of the court, this was a remedy to the constitutional violation, but in trying to look past the overcrowding to the actual mental health of inmates, whether or not the constitutional violation was actually remedied is much more unclear. This analysis is limited by the fact that legal interpretation is already challenging without trying to quantify a psychological phenomenon for an entire population. It is thus hard to conclusively say whether or not this policy remedied the constitutional violation.

This isn't to say that it was a useless policy, its good that there was reductions in prison populations. Not only that, looking past the immediate years following realignment, additionally policies were put into place to further reduce jail populations, like AB 2499 in 2014 that expanded home detention programs and brought jail populations down to their pre-realignment levels. While in the short term the problem was just shifted from institution to institution, there is some hope that with the lowering of jail populations, realignment did somewhat improve inmate mental health.

Future Work

Looking towards the future, there is a lot more work to do. Admittedly much more challenging to do, especially this far after the policy, this project could be vastly improved by including a more human perspective. Direct interviews with individuals impacted by realignment, both those who have since been released and those who are currently incarcerated, would provide perspective that no amount of survey questions can provide. This project is inherently quantitative, but a qualitative approach would be have also been productive.

While the Brown v. Plata decision explicitly did not open the door for individuals in other states to sue for decrowding, another next step should be looking for other states that could potentially benefit from a similar policy. Overcrowding is not a problem unique to California, and while the policy wasn't perfect, it still helped people. Having the knowledge gained from this analysis and the hindsight of a state already implementing the policy, other states can also look towards reducing their overcrowding while also implementing improvements along the way to ensure the problem isn't being shifted from institution to institution.

Bibliography

[1] Committee on Causes and Consequences of High Rates of Incarceration, Committee on Law and Justice, Division of Behavioral and Social Sciences and Education, National Research Council, Board on the Health of Select Populations, and Institute of Medicine, *Impact of incarceration on health*. Washington, D.C., DC: National Academies Press, Aug. 2013.

[2] "Brown, et al. v. plata, et al., 563 u.s. 493," 2011.

[3] M. Schlanger, "Plata v. brown and realignment: Jails, prisons, courts, and politics," *Harv. C. R. -C. L. L. Rev*, vol. 48, no. 1, pp. 165–215, 2009.

[4] T. Chase-Wagniere, "The perfect storm: Brown v. plata and california's financial crisis," *Southern California Review of Law and Social Justice*, vol. 22, no. 3, pp. 345–375, 2013.

[5] W. D. Ball, ""a false idea of economy"," *Ann. Am. Acad. Pol. Soc. Sci.*, vol. 664, pp. 26–42, Mar. 2016.

[6] J. Crodelle, C. Vallejo, M. Schmidtchen, C. M. Topaz, and M. R. D'Orsogna, "Impacts of california proposition 47 on crime in santa monica, california," *PLOS ONE*, vol. 16, pp. 1–27, 05 2021.

[7] J. Sundt, E. J. Salisbury, and M. G. Harmon, "Is downsizing prisons dangerous?," *Criminology & Public Policy*, vol. 15, no. 2, pp. 315–341, 2016.

[8] C. L. Rowe, A. Hubbard, and J. Ahern, "California's public safety realignment act and prisoner mortality," *PLoS One*, vol. 18, p. e0284609, Apr. 2023.

[9] United States. Bureau of Justice Statistics, "Survey of inmates in state and federal correctional facilities, [united states], 2004," 2019.

[10] United States. Bureau of Justice Statistics, "Survey of prison inmates, united states, 2016," 2023.

[11] United States. Bureau of Justice Statistics, "National prisoner statistics, [united states], 1978-2019," 2021.

[12] California Board of State and Community Corrections, "Jail profile survey," 2023.

[13] M. Gottschalk, "No star state: What's right and wrong about criminal justice reform in texas," *Seattle J. Soc. Just*, vol. 19, 2021.

[14] B. Zeldow and L. A. Hatfield, "Confounding and regression adjustment in difference-in-differences studies," *Health Serv Res*, vol. 56, pp. 932–941, May 2021.

[15] J. Ohrnberger, E. Fichera, and M. Sutton, "The relationship between physical and mental health: A mediation analysis," *Social Science & Medicine*, vol. 195, pp. 42–49, 2017.

[16] J. A. Graves, C. Fry, J. M. McWilliams, and L. A. Hatfield, "Difference-in-differences for categorical outcomes," *Health Serv Res*, vol. 57, pp. 681–692, Feb. 2022.

[17] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *J. Stat. Softw.*, vol. 67, no. 1, 2015.

[18] S. G. Luke, "Evaluating significance in linear mixed-effects models in R," *Behavior Research Methods*, vol. 49, pp. 1494–1502, Aug. 2017.