

Local Labor Market Inequality in the Age of Mass Incarceration

The Review of Black Political Economy

1–35

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0034644620966029

journals.sagepub.com/home/rbp



Luke Petach¹  and Anita Alves Pena²

Abstract

We contend that the rise of mass incarceration in the United States can be framed through the lens of stratification economics, which views race- and class-based discrimination as a rational attempt on behalf of privileged groups to preserve their relative status and the material benefits which that status confers. Using the first (to our knowledge) local-level data set on incarceration rates by race, we explore the relationship between income inequality, poverty, and incarceration at the commuting zone level from 1950 to the present. Consistent with Michelle Alexander’s hypothesis that expansion of the penal system and the rise of “tough on crime” policy were efforts by privileged groups to drive a wedge into working-class political coalitions formed out of the Civil Rights Movement, we find that labor markets with greater inequality experienced larger increases in the overall incarceration rate. Furthermore, we find that *relative* rates of poverty play a key role in explaining differential effects of mass incarceration across race. Areas where White poverty rates were large relative to non-White poverty rates experienced no significant change in White incarceration, but an *expansion* of non-White incarceration. These findings have implications for policies related to economic and judicial systems.

Keywords

inequality, poverty, mass incarceration, stratification economics

Introduction

From 1970 to 2010, the number of persons incarcerated in state and federal prisons in the United States increased from 100 per 100,000 to nearly 500 per 100,000. To contextualize the magnitude of this increase, note that the average incarceration rate in

¹Belmont University, Nashville, TN, USA

²Colorado State University, Fort Collins, CO, USA

Corresponding Author:

Luke Petach, Belmont University, 1900 Belmont Boulevard, Nashville, TN 37212-3757, USA.

Email: luke.petach@belmont.edu

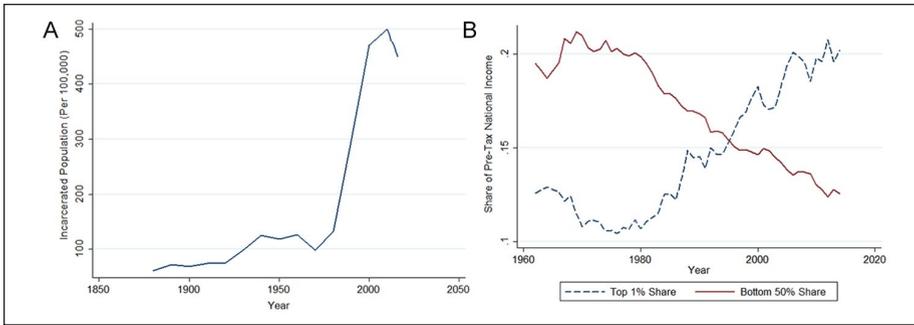


Figure 1. Inequality and incarceration. (A) Incarceration rate, 1880–2016. (B) Inequality, 1962–2014.

Source. Incarceration data from the Bureau of Justice Statistics. Inequality data from the World Top Incomes Database.

Note. Figure 1(B) plots data on pretax national income shares.

state and federal facilities from 1880 to 1970 was approximately 91 persons per 100,000, with single-year incarceration rates never exceeding 150. Addressing the expansion of the U.S. prison population after 1970, Alexander (2010) writes that “[t]he American penal system has emerged as a system of social control unparalleled in world history” (p. 8). Despite holding only 5% of the world’s total population, the United States now holds nearly 25% of the world’s prisoners (Darity et al., 2016).

The economic and social consequences of the post-1970 prison boom—particularly among low-income households and people of color—are well documented. Formerly, incarcerated individuals face worse job prospects, lower rates of wealth accumulation, a higher likelihood of engaging in future criminal activity, and increased instability in family life (Alexander, 2010; Darity et al., 2016; Liu, 2018; Pager, 2003; Western & Pettit, 2010). Despite a large body of research on the consequences of mass incarceration, little has been written by economists about the economic correlates of the prison boom.¹ This is surprising, because at least one well-known secular trend suggests itself as an explanatory factor: the rise in economic inequality in the post-1980 period. It is well established that nearly all the gains in income and wealth in the United States after 1980 were appropriated by households at the top of the income distribution (Piketty, 2014; Piketty et al., 2018; Saez & Zucman, 2016) and that the racial wealth gap is both persistent and significant (e.g., Hamilton and Darity, 2010, 2017). Figure 1 plots the share of pretax income going to both the top 1% and bottom 50% of the income distribution from 1962 to 2014, side-by-side with the incarceration rate. While the incarceration rate skyrocketed, the share of national income claimed by the bottom 50% of the distribution fell from approximately 22% to 12%, whereas the share claimed by the top 1% rose from 10% to nearly 21%.

The concomitant rise of inequality and incarceration is suggestive about the political economy of the prison boom. Alexander (2010) makes the case that mass incarceration arose as a means of reestablishing a racialized social order after the downfall

of Jim Crow during the Civil Rights Movement.² In particular, Alexander (2010) shows that “tough on crime” policy and the War on Drugs were used as a “racial bribe,” intended to drive a wedge into newly formed multiracial working-class political coalitions:

Just as race had been used at the turn of the century by Southern elites to rupture class solidarity at the bottom of the income ladder, race as a national issue had broken up the Democratic New Deal “bottom up” coalition—a coalition dependent on substantial support from all voters, white and black, at or below the median income. (p. 47)

The attempt of White elites to maintain and intensify the existing system of racialized social control by adopting policies expanding the penal system in response to perceived threats—the fall of Jim Crow, the formation of new working-class political coalitions—is the kind of behavior predicted by stratification economics. Stratification economics—in contrast to the neoclassical economic theory of discrimination, which characterizes discrimination as an irrationality that will eventually be priced out of the market—argues that race- and class-based discrimination represent a rational attempt by privileged groups to maintain their relative status and the material benefits, which that status confers (Darity, 2005; Darity et al., 2015). Viewed through this lens, the simultaneous rise of mass incarceration and inequality is not only not surprising, it is expected. As economic inequality increases, so do the returns to privilege. As a result, local elite have an incentive to increase their stratification efforts—including via more restrictive penal policy—potentially leading to higher incarceration rates in areas with greater economic inequality.

In this article, we expand on the understanding of mass incarceration as a system of racialized social control—a system that “depends for its survival on the tangible and intangible benefits that are provided to those who are responsible for the system’s maintenance and administration” (Alexander, 2010, p. 72)—by examining the relationship between incarceration rates and inequality in local labor markets. In doing so, we provide the first empirical test of the *Alexander Hypothesis*, which asserts that factors threatening the privileged economic status of Whites after 1970 are associated with the rise of Black incarceration. Using Census microdata, we construct the first (to our knowledge) complete data set on incarceration rates by race at the local labor market level, spanning the period from 1950 to 2010. Local labor market conditions have been increasingly acknowledged as a key determinant of an individual’s lifetime outcomes (Amior & Manning, 2018; Chetty et al., 2014), and are an important backdrop upon which to assess the correlates of mass incarceration. We validate our incarceration rate estimates using a recently released data set on prison population by county-of-commitment from the Vera Institute of Justice, which covers prison populations for some—but not all—counties in the United States from 1983 onward.

Using our estimates of the incarceration rate, we examine the relationship between local inequality, poverty, and incarceration at the commuting zone level from 1950 to 2010. We examine whether—and to what extent—the expansion of incarceration in the post-1970 period was larger in areas with greater local labor market inequality.

Consistent with the predictions of stratification economics and the Alexander Hypothesis, we find that labor markets with higher levels of inequality experienced larger increases in the overall incarceration rate. Furthermore, we find that *relative* rates of poverty play a key role in explaining the differential effects of mass incarceration across race. Areas where White poverty rates were large relative to non-White poverty rates experienced no significant change in White incarceration rates, but an *expansion* of non-White incarceration rates. By establishing a link between alternative measures of loss of privilege (inequality, relative poverty) and incarceration, our results provide strong empirical evidence in favor of the Alexander Hypothesis, asserting that the rise of incarceration in the United States beginning in the 1970s can be viewed as consistent with an effort (either explicit or implicit) by Whites, especially those of lower socioeconomic status, to maintain the privileges they experience from being White.³

These findings are striking and shed light on the *social* consequences of economic inequality. Our results suggest that the negative social consequences of rising inequality are disproportionately borne by non-White individuals. We argue that these distributional inequities have implications for public policies related to economic and judicial systems, including for our understanding of the “Formerly Incarcerated Reenter Society Transformed Safely Transitioning Every Person” or “First Step” Act of 2018—addressed in this special journal issue—as well as understanding and responding to the 2020 riots across the United States following the death of George Floyd at the hands of police.

The rest of the article is organized as follows. The section “Literature and Theoretical Considerations” discusses the existing literature and offers some theoretical considerations. The section “Data and Estimation Strategy” discusses the construction of the data and the estimation strategy. The section “Results” presents and discusses the results. The section “Discussion and Conclusion” concludes.

Literature and Theoretical Considerations

Mainstream economics suggests rising incarceration rates are a rational response to rising crime (Becker, 1968; Frost & Clear, 2018). Despite this strong theoretical prediction, empirical evidence in favor of the neoclassical view is weak. Recent research suggests that—if the goal of criminal punishment is to reduce the future supply of criminal behavior—mass incarceration is inconsistent with its *raison d’être*. Using quasi-random variation in punishment severity driven by the assignment of judges to 35,000 juvenile offenders, Aizer and Doyle (2015) find that incarceration leads to lower high school completion rates and higher *adult* incarceration rates, including for violent crimes. Much of the empirical literature addressing the “War on Drugs” is indicative of a battery of unintended consequences and a causal nexus seemingly unrelated to the purported goal of reducing crime (Baicker and Jacobson, 2007; Benson et al., 1995; Blumenson & Nilsen, 1998; Bluthenthal et al., 1999; Kassin et al., 2010; Koppl & Sacks, 2013; Lynch, 2012). To take one example, findings from this literature

suggest that incentives linked to changes in civil asset forfeiture laws result in increased drug arrest rates—regardless of the underlying amount of criminal behavior.

In contrast, a large sociological literature—examining both causes and consequences of incarceration—suggests alternative explanations for mass incarceration, including geographic inequality, variance in political and legal institutions (particularly those relating to policing), political rhetoric, and changes in discriminatory attitudes (Pettit & Gutierrez, 2018; Pettit & Western, 2004; Wakefield & Uggen, 2010; Western & Pettit, 2010; Western & Wildman, 2008; Yates & Whitford, 2009). Simon (2012) describes “racialized threat” as a prominent explanation for mass incarceration. This literature casts “the prison as the new urban ghetto and the policies that led to mass incarceration as the new Jim Crow” (Frost & Clear, 2018). Mass incarceration is, therefore, synonymous with attempts at social control of marginal populations and thus consistent with an economic theory centered on the role of group-based stratification, whereby collective self-interest prompts individuals to act in ways to maintain relative group position in the social hierarchy (Darity et al., 2015).

Alexander’s (2010) account in *The New Jim Crow*, which we stress here, is perhaps the closest to a stratification-based approach to mass incarceration. In particular, Alexander’s work suggests a clear theoretical link between a loss of White privilege associated with the fall of Jim Crow, and the advent of the war on drugs, “tough on crime” rhetoric, and mass incarceration. In this article, we present two tests of the Alexander Hypothesis: the idea that mass incarceration arose as part of an effort by low socioeconomic-status Whites to maintain their position of social and economic privilege.

First, we examine the link between local labor market inequality (measured by the commuting zone-level Gini coefficient) and incarceration. As inequality rises in the post-1970 period, Whites have more to lose from Black competition in the labor market—especially in formerly segregated Southern states, where such competition was previously legally restricted. Although rising inequality might be seen to reduce threats to White status, this is likely true only in the case where (a) there are few poor Whites and (b) poor White households are not inclined to form working-class political coalitions with other racial groups. To the extent that poor Whites are forced into labor market competition with Blacks (and so face a loss of privilege), and rich Whites are faced with the specter of effective redistributive political action on behalf of multiracial working-class coalitions, Whites of varying socioeconomic status have an incentive to support the “tough on crime” politics associated with mass incarceration as a means to protect their position of privilege in the social hierarchy. Thus, we expect increases in post-1970 incarceration in response to rising inequality to have origins within local labor markets as opposed to at higher levels of geography.

Second, we examine the relationship between White relative poverty—measured as the ratio of the White poverty rate to the non-White poverty rate—and mass incarceration. As White poverty rises—relative to Black poverty—this reveals a loss of White privilege. We expect increases in relative poverty to positively affect incarceration for at least two reasons. First, if Alexander’s (2010) description of mass incarceration as a

Table 1. Summary of Theoretical Discussion and Empirical Expectations.

Hypothesis	Summary/Expectation
Theory: Alexander Hypothesis	The rise of Black incarceration in the United States after 1970 can be viewed as an effort by White individuals—especially those of lower socioeconomic status—to maintain the privileges they experience from being White. Factors associated with a loss of White privilege positively affect incarceration.
Empirical test: Alexander Hypothesis 1	An increase in post-1970 labor market inequality positively affects the local incarceration rate.
Empirical test: Alexander Hypothesis 2A	An increase in post-1970 White relative poverty positively affects the local non-White incarceration rate.
Empirical test: Alexander Hypothesis 2B	An increase in post-1970 White relative poverty has no effect on (or reduces) the White incarceration rate.

“racial bribe” is correct, both poor Whites—due to loss of privilege from increased labor market competition from non-Whites—and rich Whites—acting on class interest—have a stratification-based incentive to adopt stricter penal policy as relative poverty rises. Second, to the extent that greater poverty relates to an increase in fear of crime on the part of Whites, this fear can once again be used to justify alterations in laws (stiffer minimum sentences) and policing practices (i.e., stop and frisk), leading to mass incarceration. We therefore expect increases in post-1970 relative poverty to be associated with higher rates of incarceration within local labor markets. In addition, if Alexander’s “racial bribe” theory is correct, increases in relative poverty should have no impact on (or potentially reduce) White incarceration. Table 1 summarizes the above theoretical discussion and our empirical hypotheses.

Data and Estimation Strategy

Data

Measuring the incarceration rate. The main dependent variable in our study is the commuting zone–level incarceration rate. Data on commuting zone–level incarceration rates are constructed using the Integrated Public Use Microdata Series (IPUMS) Decennial Census Microdata for the years 1950 to 2010 (Ruggles et al., 2019). We focus on commuting zones because of our interest in the relationship between local labor market inequality and incarceration outcomes. Commuting zones are clusters of counties characterized by strong commuting ties, and were initially defined by Tolbert and Sizer (1996) with the express purpose of constructing a geographic unit that best captured the notion of a local labor market. Commuting zones have been shown to be an appropriate geographic level in which to analyze questions of long-run well-being (e.g., Chetty et al., 2014; Chetty & Hendren, 2018). Our sample frame consists of the

722 commuting zones making up the contiguous United States. We focus on the population of working-aged individuals between 15 and 64 years.

In each sample year, IPUMS classifies all housing units as falling into one of three categories: households, group quarters, or vacant units. Institutionalized individuals—including those incarcerated in state or federal prisons or local correctional facilities—are classified as residing in group quarters. For the years 1950 to 1980, the Census specifically designates whether an individual was housed in a correctional institution. For each of these years, we estimate the number of incarcerated individuals in a given commuting zone by totaling the number of individuals housed in correctional institutions. We aggregate the individual-level microdata to commuting zones following Autor and Dorn (2013).⁴

From 1990 onward, IPUMS indicates whether an individual residing in group quarters was institutionalized but does not specify correctional institutions in detail. Thus, for the 1990 to 2010 samples, we apply an adjustment factor based on the proportion of institutionalized individuals residing in correctional institutions in 1980. We calculate a unique correction factor for each state-by-race group—defined as the share of the institutionalized population belonging to each subgroup that was housed in a correctional institution in 1980. We then multiply this correctional factor by the total number of institutionalized individuals in a commuting zone to arrive at our estimate for the years 1990 to 2010. Figure 2 presents a map of our commuting zone-level incarceration rates—defined as the number of incarcerated individuals per 100,000—for the years 1950 and 2010.

In Appendix A, we validate our incarceration rate measure using data from the Bureau of Justice Statistics and a recently released data set on incarceration by county of admission from the Vera Institute of Justice. Our estimates of commuting zone-level incarceration rates match aggregate trends and are externally validated by the Vera data.⁵ Given that our measure of incarceration stands up to a reasonable degree of external validation, we now examine changes in the spatial pattern of incarceration over time. First, it is clear from Figure 2 that the geographic area that experienced the largest expansion in incarceration rates between 1950 and 2010 was the Southern United States—in particular, the Gulf states including Texas. This pattern is not surprising. It is well known that the Civil Rights movement faced heavy political backlash in the South, a stronghold area of Jim Crow. At the national level, Kuziemko and Washington (2018) show that the exodus of Southern Whites from the Democratic Party between 1958 and 1980 is entirely explained by the exodus of “racially conservative” Whites following the Civil Rights initiatives of the 1960s. Expansion of the prison population in these states is consistent with Alexander’s (2010) description of mass incarceration as a response by local elites to a loss of social control.

However, despite higher *overall* incarceration rates in Southern states, *relative* rates of incarceration—the ratio of non-White incarceration to White incarceration—are higher in Northern and Midwestern states. Figure 3 maps the relative incarceration rate across commuting zones for the year 2010. This figure suggests that—to the extent that incarceration rates expanded in the South, they did so for both Whites and

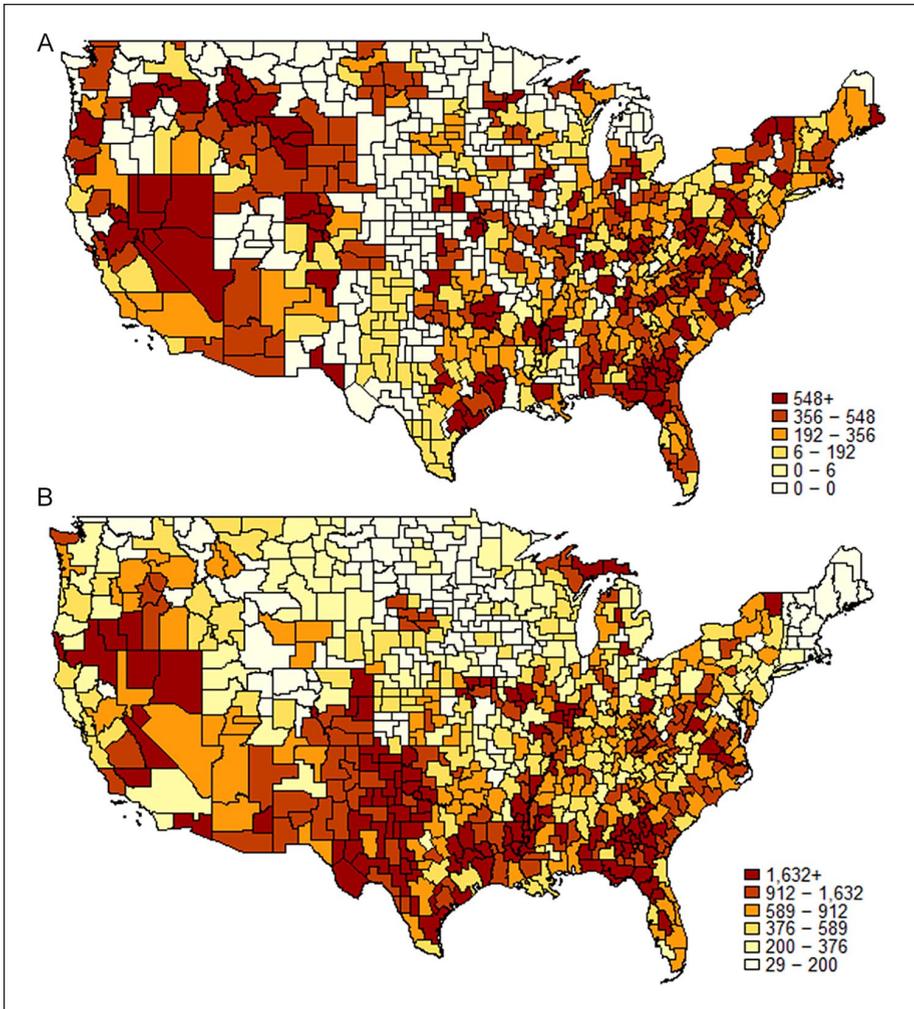


Figure 2. Mapping the incarceration rate. (A) Commuting zone map of incarceration rates, 1950. (B) Commuting zone map of incarceration rates, 2010.

Source. Incarceration estimates constructed using the IPUMS/Census microdata. Commuting zone crosswalks obtained from Autor and Dorn (2013).

Note. Figures plot the number of incarcerated persons per 100,000 population at the commuting zone level.

non-Whites. In contrast, expansion of the penal system in Northern states appears to have concentrated largely on non-White populations. Insofar as Alexander's (2010) hypothesis focuses on the disproportionate impact of the prison boom on people of color, Figures 2 and 3 suggest it is a Northern story.

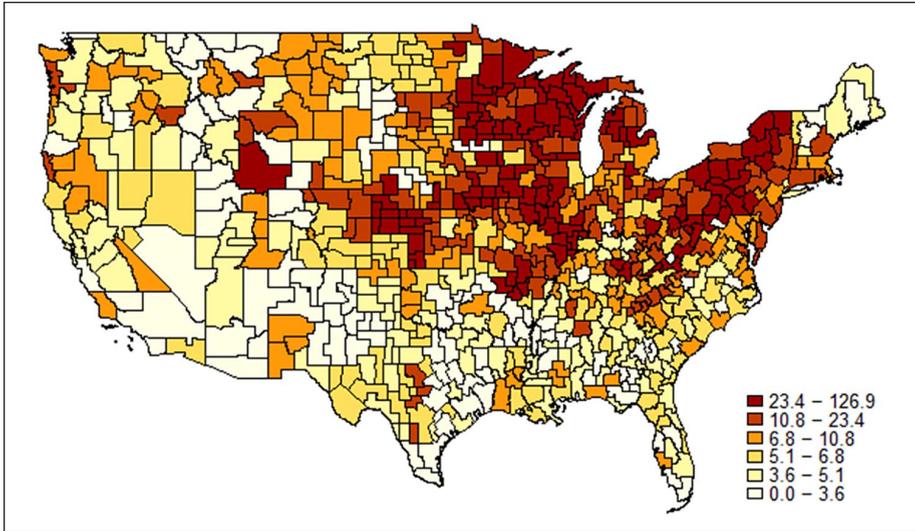


Figure 3. Commuting zone map of relative incarceration rates, 2010.

Source. Incarceration estimates constructed using the IPUMS/Census microdata. Commuting zone crosswalks obtained from Autor and Dorn (2013).

Note. Figure plots the ratio of the non-White incarceration rate to the White incarceration rate at the commuting zone level.

Measuring loss of privilege. In our evaluation of the Alexander Hypothesis—asserting a link between the loss of White privilege and the rise of mass incarceration—we adopt two measures of perceived loss of privilege: inequality and relative poverty. Two mechanisms link these variables to incarceration. First, as income inequality increases in the post-1970 period (captured by an increase in the Gini coefficient), Whites have more to lose from Black competition in the labor market (brought on by the civil rights movement, affirmative action, etc.). In response to this threat, more restrictive penal policy (e.g., minimum sentencing guidelines) may be adopted. Second, as White poverty rises relative to Black poverty, a loss of privilege is revealed, invoking a potentially similar response.

Commuting zone-level measures of inequality are constructed using IPUMS Census Data (Ruggles et al., 2019). For each commuting zone, we compute the Gini coefficient across households using total household income, which includes the total pretax personal income or losses from all sources for all persons residing in a given household over the previous year. Due to top coding in the IPUMS data, our estimates of the Gini coefficient will understate the true extent of inequality. However, what is important for our analysis is not necessarily the level of inequality itself, but variation in inequality across commuting zones and over time, both of which should be preserved even in the presence of top coding. Figure 4 presents a map of the Gini coefficient across commuting zones for the year 2010. The map shows a clear spatial pattern,

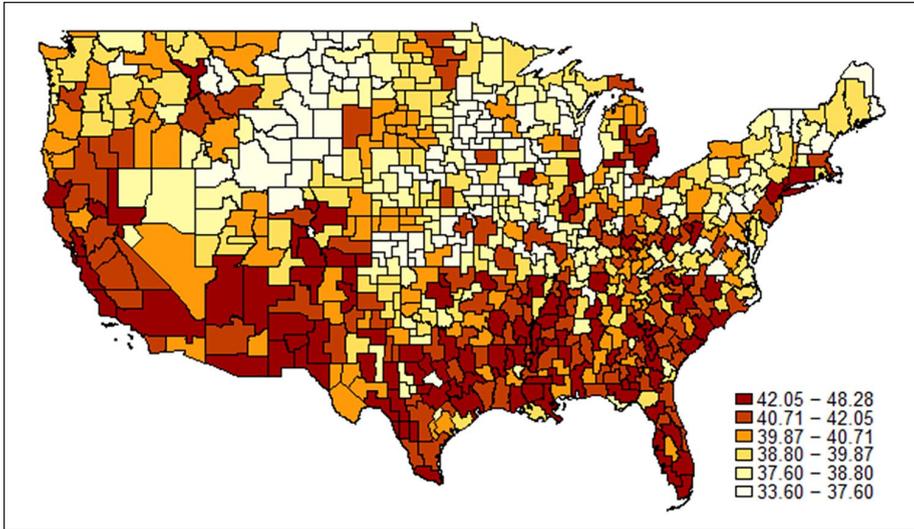


Figure 4. Commuting zone map of Gini coefficient for household income, 2010.

Source. Gini coefficient estimates constructed using the IPUMS/Census microdata. Commuting zone crosswalks obtained from Autor and Dorn (2013).

Note. Figure plots the Gini coefficient for household income at the commuting zone level.

indicating higher average levels of inequality in Southern commuting zones. Jointly, Figures 2(B) and 4 offer prima facie support for the idea that inequality and incarceration may be correlated across space, suggesting an overlap between high inequality and high incarceration rate commuting zones, especially in the South. Figure 5 formalizes this insight, plotting the sample correlation coefficient for inequality and incarceration in the period after 1970 for all commuting zones. For many commuting zones, this correlation is strongly positive, taking on values more than .8.

To measure relative poverty, we take the ratio of the White poverty rate in a commuting zone to the non-White poverty rate in a commuting zone. An increase in this variable is informative about changes in the relative economic status of Whites in a given labor market. Sample means for both the commuting zone Gini coefficient and White relative poverty are presented in Table 2.

Additional controls

Crime statistics. To evaluate the relationship between inequality and incarceration in a way that addresses Alexander's (2010) claim that mass incarceration arose as an attempt by privileged Whites to reestablish social control after the fall of Jim Crow, it is important to interrogate alternative mechanisms that might explain the relationship between inequality and incarceration. The most obvious of these mechanisms is an increase in the rate of criminal activity in response to rising inequality. If rising incarceration rates are simply a response to an increase in criminal activity, there

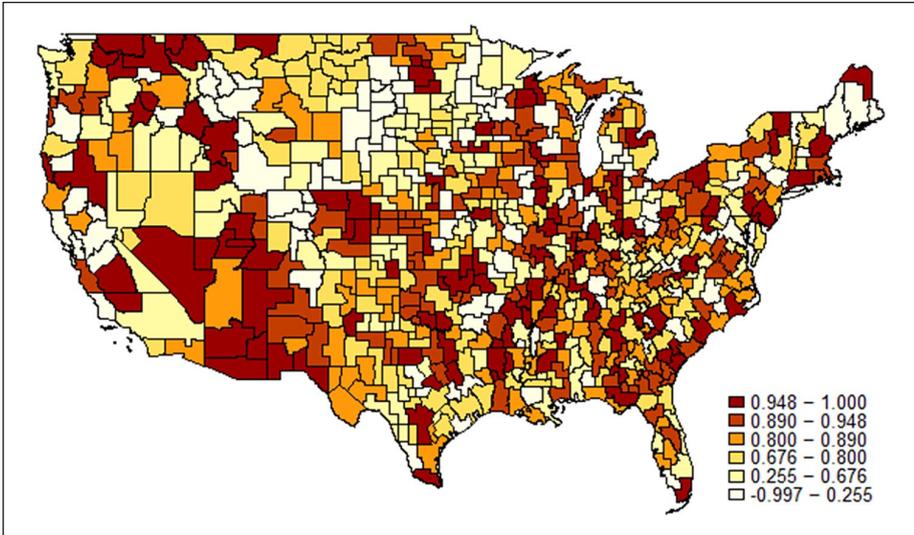


Figure 5. Inequality and incarceration correlation coefficient map, 1980 to 2010.

Source. Gini coefficient estimates constructed using the IPUMS/Census microdata. Incarceration estimates constructed using the IPUMS/Census microdata. Commuting zone crosswalks obtained from Autor and Dorn (2013).

Note. Figure plots the sample correlation coefficient for the Gini coefficient for household income and the overall incarceration rate at the commuting zone level for the post-1970 period.

would appear to be little evidence for a stratification-based explanation of mass incarceration. Thus, in our analysis, we include a measure of the violent crime rate as an additional control variable. We calculate the crime rate in each commuting zone as the number of violent crimes per 100,000 population, using commuting zone data on the number of violent crimes from the Federal Bureau of Investigation's Uniform Crime Reporting (UCR) Statistics (Kaplan, 2019).⁶

Local-level economic and demographic indicators. Sample means for the main variables of interest and control variables are presented in Table 2.⁷ Additional variables include commuting zone population shares by race, the share of residents in a commuting zone with at least a bachelor's degree, employment shares by industry (with industries defined according to the 1990 Census Bureau industrial classification scheme),⁸ the unemployment rate, population, and decadal population growth. The table presents sample means for the entire 1950 to 2010 period.

Estimation Strategy

To assess (a) the impact of inequality on the post-1970 increase in incarceration across commuting zones and (b) whether non-White incarceration rates rose more in areas

Table 2. Sample Means, 1950–2010.

Variable	M	SD	Source
Incarceration rate (per 100,000)	426.19	493.96	IPUMS
White incarceration rate (per 100,000)	249.22	298.48	IPUMS
Non-White incarceration rate (per 100,000)	2,048.14	3,778.18	IPUMS
Crime rate (per 100,000; 1960–2010)	444.43	318.09	UCR
Gini coefficient	35.9	5.13	IPUMS
White poverty rate (%)	12.4	8.6	IPUMS
Non-White poverty rate (%)	30.0	15.33	IPUMS
Relative poverty rate (ratio)	0.42	0.25	IPUMS
White population share (%)	79.85	15.3	IPUMS
Black population share (%)	10.04	8.73	IPUMS
Hispanic population share (%)	7.01	10.5	IPUMS
Asian population share (%)	2.06	3.2	IPUMS
American Indian population share (%)	0.45	1.43	IPUMS
Other race population share (%)	0.58	0.85	IPUMS
Bachelor's degree (%)	19.38	13.8	IPUMS
Unemployment rate (%)	4.5	2.2	IPUMS
Household income	US\$47,939.22	US\$19,153.06	IPUMS
Commuting zone population (persons)	2,823,229	3,798,615	BEA
Decadal commuting zone population growth (%)	13.9	15.7	BEA
Manufacturing emp. share (%)	19.8	10.08	IPUMS
Retail emp. share (%)	16.48	1.9	IPUMS
Agriculture emp. share (%)	4.06	6.2	IPUMS
<i>N</i> = 5,054			

Note. Sample means weighted using commuting zone population. UCR = Uniform Crime Reporting, BEA = Bureau of Economic Analysis.

with larger concentrations of poor White households—in a fashion consistent with Alexander's (2010) description of mass incarceration as a “racial bribe”—we exploit variation in both incarceration rates and inequality across time and across commuting zones using a two-way fixed effects approach:

$$\text{Incarceration}_{ct} = \beta_1 \times \text{treat}_c + \beta_2 \times \text{After}_t^{1970} + \beta_3 \times (\text{treat}_c \times \text{After}_t^{1970}) + \mathbf{X}_{ct}^T \beta + \lambda_c + \delta_t + \epsilon_{ct}, \quad (1)$$

where $\text{Incarceration}_{ct}$ is the incarceration rate in commuting zone c in year t . treat_c is a continuous variable intended to measure perceived loss of privilege at the commuting zone level, giving the average value of either the household income Gini coefficient or the ratio of White-to-non-White poverty for the post-1970 period. After_t^{1970} is a dummy variable taking a value of 1 after 1970—which we identify as the start of the period of mass incarceration—and 0 otherwise. The data are decennial, where each year t represents a decade, with t running from 1950 to 2010. The regression

coefficient on the interaction between treat_c and $\text{After}_t^{1970} - \beta_3$ —is our parameter of interest, giving an estimate of the average impact of inequality or relative poverty on incarceration after 1970. The treatment effect is interpreted as the impact of a unit increase in a given loss of privilege measure on post-1970 incarceration, as the variable After_t^{1970} is nonzero only for observations following the year 1970. Our treatment variable is constructed using the post-1970 average of inequality or relative poverty,⁹ as our interest is in the relationship between the simultaneous rise of inequality and incarceration, which occurs almost entirely after 1970 (with pre-1970 levels of inequality not necessarily predicting the extent of the post-1970 increase). Rather than dividing the sample into discrete treatment and control groups, our method assumes all labor markets are treated—in that the post-1970 average Gini coefficient is positive for all labor markets—but that the “dosage” or intensity of the treatment varies across commuting zones. As an alternative, Appendix B presents results from a discrete treatment setting where the treatment and control groups are identified by splitting the sample into “high” and “low” inequality groups. The results from the discrete treatment specification are qualitatively similar to those presented in the body of the article.

The remaining variables are \mathbf{X}_{ct}^T —a vector of controls, λ_c —a commuting zone-specific fixed effect, δ_t —a year-fixed effect, and ϵ_{ct} —an idiosyncratic error term. In addition to Equation 1, we estimate the following flexible specification, which allows the treatment effect to vary by year:

$$\text{Incarceration}_{ct} = \beta_1 \times \text{treat}_c + \delta_t + \sum_t (\delta_t \times \text{treat}_c) + \mathbf{X}_{ct}^T \beta + \lambda_c + \epsilon_{ct}, \quad (2)$$

where the interaction between δ_t and treat_c gives the impact of a unit increase in the Gini coefficient or relative poverty rate on incarceration in year t .

In addition to controlling for the impact of crime on incarceration as a possible channel through which inequality affects prison populations, it is important to control for labor market outcomes that might otherwise be correlated with both inequality and incarceration. To address this concern, the above specification adjusts for time-varying local levels of unemployment, education, income, population, population growth, and industrial composition, as well as demographic controls for population shares by race. Thus, although our treatment variable—average inequality in the post-1970 period—is not randomly assigned to commuting zones, we believe that we are able to adjust for most of the confounding backdoor paths that could bias our estimate of the impact of inequality on incarceration. Nonetheless, we are cautious about overemphasizing the causal nature of the estimates obtained from the above specification, emphasizing their reduced-form nature and focusing on the bigger picture implied by the relationships we observe.

A requirement of the difference-in-differences-type specifications in Equations 1 and 2 is that the treatment and control groups satisfy the parallel trends assumption. This assumption requires that trends in the dependent variable across the treatment and control groups be parallel in the pretreatment period. If this assumption is violated, our

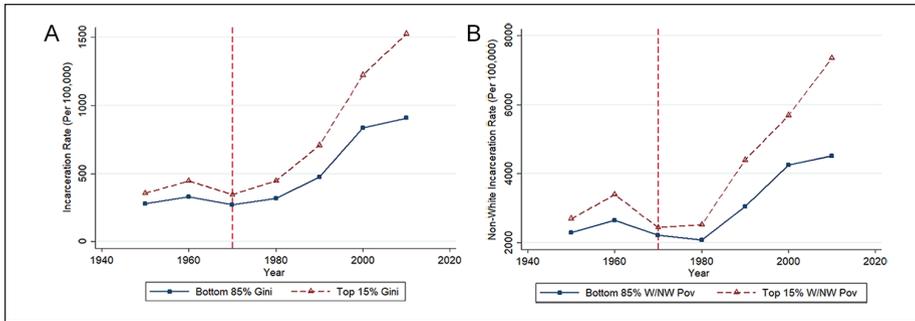


Figure 6. Incarceration trend plots by level of inequality and relative poverty. (A) Incarceration by Gini. (B) Non-White incarceration by ratio of White-to-non-White poverty.

estimate of β_3 will be biased. In Equation 2, this amounts to requiring that the interaction term between the inequality variable and year-fixed effect in pretreatment years be statistically insignificant and close to zero. A formal test of the parallel trend assumption in the framework of Equation 1 can be implemented via a placebo test that restricts the treatment window to the pretreatment period and assigns an alternate “placebo” treatment date. In section “Results,” we implement both tests. Insight can be gained prior to formal hypothesis testing by plotting the outcome variable across treatment and control groups. Because we have a continuous variable, we divide the sample into “high treatment” and “low treatment” groups, where commuting zones in the top 15% of the distribution of the treatment variable are included in the former. Figure 6 plots trends in the overall incarceration rate as well as the non-White incarceration rate for “high” and “low” treatment values of inequality and relative poverty, respectively. The figure is suggestive of parallel trends across groups in the pretreatment period, as well as being indicative of a divergence in outcomes across the two groups following the start of mass incarceration.

Results

Inequality

Table 3 presents results from our initial estimates of Equation 1. We first look at the effect of overall inequality—as measured by the average commuting zone household income Gini coefficient in the post-1970 period—on commuting zone-level incarceration rates. Standard errors are clustered at the state level. All regressions include controls for population shares by race, total population, population growth, average household income, unemployment, average education, and employment shares by industry. Column 2 adds time-fixed effects. Column 3 adds both time- and commuting zone-fixed effects. Column 4 weights observations by the average commuting zone population over the sample period. Column 5 adjusts for the violent crime rate.

Table 3. Estimation Results, Inequality and Overall Incarceration.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Incarceration						
$(Gini_c \times Aftcr_t^{970})$	65.94*** (20.79)	94.85*** (28.05)	65.83*** (15.59)	61.98*** (14.80)	56.42*** (16.69)	33.16*** (11.96)	48.31*** (16.92)
Crime _{it}	—	—	—	—	0.210** (0.103)	0.114* (0.059)	0.264** (0.122)
N	5,054	5,054	5,054	5,054	4,332	4,320	4,296
Demographics	Y	Y	Y	Y	Y	Y	Y
Industry	Y	Y	Y	Y	Y	Y	Y
Economy	Y	Y	Y	Y	Y	Y	Y
CZ FE	N	N	Y	Y	Y	Y	Y
Year FE	N	Y	Y	Y	Y	Y	Y
Vera weights 1	N	N	N	N	N	Y	N
Vera weights 2	N	N	N	N	N	N	Y
Population weights	N	N	N	Y	N	N	N

Note. Dependent variable is incarceration rate (per 100,000). Inequality is measured by the commuting zone-level household income Gini coefficient. All regressions include controls for population shares by race, total population, population growth, average household income, unemployment, average education, and employment shares by industry. Column 2 adds time-fixed effects. Column 3 adds both time- and commuting zone-fixed effects. Column 4 weights observations by the average commuting zone population over the sample period. Column 5 adjusts for the violent crime rate. Columns 6 and 7 use alternative weighting schemes based on the similarity between the IPUMS/Census incarceration rate measure and the Vera Institute incarceration rate measure, where commuting zones with IPUMS/Census incarceration rate estimates closer in magnitude to the Vera Institute estimates are given higher weights. Standard errors in parenthesis, clustered at the state level.

* $p < .10$. ** $p < .05$. *** $p < .01$.

CZ = Commuting Zone, FE = Fixed Effects.

Columns 6 and 7 use an alternative weighting scheme based on the similarity between the IPUMS/Census incarceration rate measure and the Vera Institute incarceration rate measure, where commuting zones with IPUMS/Census incarceration rate estimates closer in magnitude to the Vera Institute estimates are given higher weights. The alternative weighting schemes are constructed as follows. In Column 6, observations are weighted according to the sample mean of the inverse of the absolute value of the difference between the IPUMS/Census incarceration rate estimate and the Vera Institute incarceration rate estimate. In Column 7, observations are weighted by the absolute value of the within-commuting zone correlation between the IPUMS/Census measure and the Vera Institute measure. Note that neither of these weighting schemes require observations for the Vera Institute measure in all years, and thus allow for the inclusion of a wider range of commuting zones than would be available in the Vera data in any given year. The weighting schemes adjust for differences between the alternative measures of incarceration, as discussed in Appendix A.

The results suggest a statistically significant positive relationship between inequality and the rate of incarceration after 1970. Averaging across all columns, a one-unit increase in the average post-1970 household income Gini coefficient in a commuting zone is associated with an approximately 61-point increase in the incarceration rate (an increase of 61 people per 100,000). For a commuting zone at the sample mean post-1970 incarceration rate of 684.94, this represents a nearly 10% increase. This result appears robust to the inclusion of a wide variety of controls, including commuting zone- and year-fixed effects.¹⁰ Importantly, this result is robust to the inclusion of the violent crime rate, which suggests that the increase in incarceration after 1970 was not simply a rational response to rising crime. Holding the crime rate constant, an increase in commuting zone inequality nonetheless translates into higher incarceration rates after 1970. However—consistent with the neoclassical approach—the commuting zone-level crime rate has a statistically significant, positive effect on the incarceration rate.

Even if we consider the specification with the smallest estimated effect magnitude (33.16)—Column 6, which accounts for differences in measured incarceration (due to the recorded location of prisoners) between the Census data and the Vera Institute data—the impact of inequality on incarceration is nonetheless highly statistically and economically significant, suggesting the positive relationship between inequality and incarceration is not driven merely by measurement error in the Census incarceration data.

Table 4 extends our estimates of Equation 1 over different race and ethnicity categories. We focus on differences in our parameter estimates across White non-Hispanic, non-White non-Hispanic, Hispanic, and Black non-Hispanic race/ethnicity groupings.

The results in Table 4 suggest that a unit increase in the average household income Gini coefficient in the post-1970 period has a larger impact on non-White, non-Hispanic incarceration than White incarceration, where the magnitudes of the coefficients suggest the bulk of the effect is driven by the impact of inequality on incarceration of Black individuals. A unit increase in the average post-1970 Gini results in an increase in incarceration of approximately 429 persons per 100,000 for non-White, non-Hispanic individuals. The coefficient for Black, non-Hispanic individuals is of similar

Table 4. Estimation Results, Inequality and Incarceration by Race.

Variable	(1)		(2)		(3)		(4)	
	White, non-Hispanic	Non-White, non-Hispanic	White, non-Hispanic	Non-White, non-Hispanic	Hispanic	Black, non-Hispanic	Hispanic	Black, non-Hispanic
$(Gini_c \times After_t^{1970})$	26.76*** (9.122)	429.4** (167.3)	91.34 (109.5)	418.4 (330.8)				
Crime _{ct}	0.0663 (0.0432)	0.392 (0.545)	0.570 (0.443)	-0.817 (1.267)				
White _{ct}	-9.030** (4.241)	22.34 (80.16)	30.13 (26.07)	29.16 (105.0)				
N	4,332	4,332	4,280	4,303				
Demographics	Y	Y	Y	Y	Y	Y	Y	Y
Industry	Y	Y	Y	Y	Y	Y	Y	Y
Economy	Y	Y	Y	Y	Y	Y	Y	Y
CZ FE	Y	N	Y	N	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y

Note. Dependent variable is incarceration rate (per 100,000) by race/ethnicity category. Inequality is measured by the commuting zone-level household income Gini coefficient. All regressions include controls for the White population share, total population, population growth, average household income, unemployment, average education, and employment shares by industry. Standard errors in parenthesis, clustered at the state level.

* $p < .10$. ** $p < .05$. *** $p < .01$.

CZ = Commuting Zone, FE = Fixed Effects.

magnitude—although less precisely estimated. In contrast, the same increase in the Gini coefficient increases the incarceration rate for White individuals by only 27 persons per 100,000, an effect smaller in magnitude than the average overall effect estimated in Table 2. These results add to the growing body of work suggesting that the *social* consequences of inequality are not borne equally across race/ethnicity groups, with emphasis on the negative social impact of economic inequality on minorities and people of color.

In addition to the main results presented in Table 4, it is worth drawing attention to the estimated regression coefficients on the crime rate and White population share. In every specification, the coefficient on the crime rate variable is positive but statistically insignificant, weakly consistent with the neoclassical hypothesis. Second, the White population share variable has a statistically significant negative impact on White, non-Hispanic incarceration rates. This result suggests that commuting zones where a large share of the population is White incarcerate White individuals at lower rates. In contrast, the effect of the White population share is positive, but statistically insignificant, for all non-White groups.¹¹ Thus, the racial composition of a commuting zone appears to have an independent effect on incarceration. In the next subsection, we extend this dimension of the analysis to explore whether or not differences in the economic standing of various race/ethnicity groups have differential impacts on incarceration rates by race in a fashion consistent with Alexander's (2010) description of tough on crime policies as a "racial bribe."

Relative Poverty

Table 5 presents regression coefficients from estimating Equation 1 using relative poverty—measured as the ratio of the White poverty rate to the non-White poverty rate in a given commuting zone—as the variable of interest. Specifically, $RelPov_c$ measures the White poverty rate as a percentage of the non-White poverty rate, such that a unit increase in $RelPov_c$ implies a 1% increase in the White poverty rate relative to non-White poverty. The results in Columns 1 and 3 of Table 5 indicate that as the White poverty rate increases—as a percentage of the Black poverty rate—there is no effect on post-1970 incarceration for either White non-Hispanic or Hispanic groups. In contrast, an increase in White relative poverty is associated with statistically significant, positive increases in both non-White, non-Hispanic incarceration rates and Black, non-Hispanic incarceration rates. A 1% increase in the White relative poverty rate after 1970 corresponds to an increase in the incarceration rate of 35.55 points (35.55 persons per 100,000) for the entire non-White, non-Hispanic group, and an increase of 76.10 points (76.1 persons per 100,000) for Black, non-Hispanics. This result suggests that the spatial incidence of mass incarceration for non-Whites is greater in local labor markets with a larger *relative* concentration of poor Whites—the group that Alexander (2010) describes as a target of a "racial bribe" intended to disrupt working-class political coalitions after the demise of Jim Crow. Although this result does not provide direct confirmation of the Alexander Hypothesis—in that the reduced-form estimates presented here cannot prove discriminatory intent—it adds to the body of evidence in

Table 5. Estimation Results, Relative Poverty and Incarceration by Race.

Variable	(1)		(2)		(3)		(4)	
	White, non-Hispanic		Non-White, non-Hispanic		Hispanic		Black, non-Hispanic	
$(\text{RelPov}_c \times \text{After}_t^{1970})$	-0.577 (0.874)		35.55*** (8.588)		-11.93 (8.865)		76.10*** (21.83)	
Crime _{ct}	0.0868* (0.0787)		0.865 (0.553)		0.609 (0.446)		-0.203 (1.337)	
White _{ct}	-10.68*** (4.558)		2.603 (81.10)		22.92 (29.09)		16.48 (99.19)	
N	4,332		4,332		4,280		4,303	
Demographics	Y		Y		Y		Y	
Industry	Y		Y		Y		Y	
Economy	Y		Y		Y		Y	
CZ FE	Y		N		Y		Y	
Year FE	Y		Y		Y		Y	

Note: Dependent variable is incarceration rate (per 100,000) by race/ethnicity category. RelPov_c measures the White poverty rate as a percentage of the non-White poverty rate. All regressions include controls for the White population share, total population, population growth, average household income, unemployment, average education, and employment shares by industry. Standard errors in parenthesis, clustered at the state level.

* $p < .10$. ** $p < .05$. *** $p < .01$.

CZ = Commuting Zone, FE = Fixed Effects.

support, indicating that incarceration of non-Whites expanded more in areas with higher relative White poverty rates.

Additional Sensitivity Checks

We begin the sensitivity analysis by extending the main specification to Equation 2, which allows year-by-year interactions between year-fixed effects and the variable of interest. This specification allows for flexible dynamics in measuring the impact of inequality and/or relative poverty on incarceration in the post-1970 period by varying the treatment effect by year. Table 6 presents the results. We report results from specifications including the crime rate variable, although we also run each of the regressions without it. We find no meaningful difference in the estimated regression coefficients on our variables of interest in either case. Columns 1 to 4 present results using average post-1970 Gini coefficient as the variable of interest. Columns 5 to 7 present results using the average post-1970 ratio of the White poverty rate to the non-White poverty as the variable of interest.

Table 6 supports the results found previously: Higher post-1970 inequality is associated with greater incarceration rates across all race/ethnicity groups (although the effect is statistically insignificant when Black, non-Hispanics, and Hispanics are considered separately), but greater relative poverty for Whites is only statistically significantly associated with incarceration for non-Whites, and for Black, non-Hispanics in particular. The results in Table 6 also shed new light on the temporal dimension of the relationship between inequality, poverty, and mass incarceration. The estimates suggest that the effect of local labor market inequality on incarceration occurs predominantly after 1980. This finding is consistent with the overall trajectory of incarceration depicted in Figure 1, which—despite indicating an initial uptick in incarceration between 1970 and 1980—illustrates that the lion's share of the increase in incarceration in the post-1970 period occurs between 1980 and 2010.

Finally, Table 6 lends initial support for the parallel trends assumption required for difference-in-differences. For difference-in-differences estimates to be valid, trends in the outcome variable must be parallel across treated and untreated groups in the period prior to the treatment. In this case, with a continuous treatment indicator variable (local labor market inequality), what is required is that trends in incarceration be the same across commuting zones for varying levels of inequality for all years prior to 1970. In Table 5, this amounts to requiring the interaction term featuring the 1970 year-fixed effect and the variable of interest be statistically insignificant and close to zero. This condition is satisfied across all columns. The regression coefficient on the variable of interest is statistically insignificant and small in magnitude in all cases when interacted with the 1970 year-fixed effect. However, to test the parallel trends assumption more formally, we nonetheless implement an additional placebo test across “high” and “low” treatment groups.

Table 7 presents the results of the placebo test. To implement the test, we separate the sample into “high” and “low” treatment groups—commuting zones in the top 15% of the sample distribution for inequality or relative poverty are counted as “high”

Table 6. Estimation Results, Year-by-Year Effects.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Overall	White, non-Hispanic	Non-White, non-Hispanic	Black, non-Hispanic	White, non-Hispanic	Non-White, non-Hispanic	Black, non-Hispanic
$(Gini_c \times \delta_{1970})$	3.790 (18.91)	4.275 (13.73)	-99.43 (235.6)	411.4 (391.6)	—	—	—
$(Gini_c \times \delta_{1980})$	0.734 (14.94)	5.966 (11.67)	150.9 (159.2)	493.7* (289.5)	—	—	—
$(Gini_c \times \delta_{1990})$	36.57* (18.71)	18.43 (14.60)	369.7* (216.7)	647.4 (463.6)	—	—	—
$(Gini_c \times \delta_{2000})$	116.7*** (35.49)	51.78** (21.44)	517.0 (310.0)	715.6 (525.8)	—	—	—
$(Gini_c \times \delta_{2010})$	148.5*** (36.44)	61.49*** (19.63)	628.5** (304.3)	834.3 (538.0)	—	—	—
$(RelPov_c \times \delta_{1970})$	—	—	—	—	-0.458 (0.790)	0.0268 (14.25)	-4.602 (24.29)
$(RelPov_c \times \delta_{1980})$	—	—	—	—	-1.051 (0.807)	2.038 (12.98)	1.478 (24.64)
$(RelPov_c \times \delta_{1990})$	—	—	—	—	-0.586 (1.166)	44.13** (20.57)	77.71 (48.27)
$(RelPov_c \times \delta_{2000})$	—	—	—	—	-0.429 (1.413)	43.88*** (15.47)	121.5*** (34.76)
$(RelPov_c \times \delta_{2010})$	—	—	—	—	-1.252 (1.402)	57.30*** (18.35)	104.3** (42.68)
N	4,332	4,332	4,332	4,303	4,332	4,332	4,303
Demographics	Y	Y	Y	Y	Y	Y	Y
Industry	Y	Y	Y	Y	Y	Y	Y
Economy	Y	Y	Y	Y	Y	Y	Y
CZ FE	N	N	Y	Y	Y	Y	Y
Year FE	N	Y	Y	Y	Y	Y	Y

Note. Dependent variable is incarceration rate (per 100,000) by race/ethnicity category. All regressions include controls for the White population share, total population, population growth, average household income, unemployment, average education, violent crime rate, and employment shares by industry. Standard errors in parenthesis, clustered at the state level. * $p < .10$. ** $p < .05$. *** $p < .01$. CZ = Commuting Zone, FE = Fixed Effects.

Table 7. Placebo Test.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Overall	White, non-Hispanic	Non-White, Non-Hispanic	Black, non-Hispanic	White, non-Hispanic	Non-White, non-Hispanic	Black, non-Hispanic
$(\text{HighGini} \times \text{After}_t^{1960})$	28.42 (46.76)	63.57 (38.45)	157.2 (455.4)	589.2 (1,271.5)	—	—	—
$(\text{HighPov} \times \text{After}_t^{1960})$	—	—	—	—	46.06 (61.26)	494.1 (866.8)	1,342.0 (1,999.6)
N	2,166	2,166	2,153	2,075	2,166	2,153	2,075
Demographics	Y	Y	Y	Y	Y	Y	Y
Industry	Y	Y	Y	Y	Y	Y	Y
Economy	Y	Y	Y	Y	Y	Y	Y
CZ FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y

Note. Dependent variable is incarceration rate (per 100,000) by race/ethnicity category. All regressions include controls for the White population share, total population, population growth, average household income, unemployment, average education, and employment shares by industry. Standard errors in parenthesis, clustered at the state level.

* $p < .10$. ** $p < .05$. *** $p < .01$.

CZ = Commuting Zone, FE = Fixed Effects.

Table 8. Oster (2019) Bias-Adjusted Coefficients.

Inequality on overall incarceration	Inequality on non-White, non-Hispanic incarceration	Relative poverty on non-White, non-Hispanic incarceration
$\beta_{\text{bias adjusted}} = 58.79$	$\beta_{\text{bias adjusted}} = 462.49$	$\beta_{\text{bias adjusted}} = 34.70$
Original estimate: Table 3, Column 3	Original estimate: Table 4, Column 2	Original estimate: Table 5, Column 2

treatment, as in Figure 6—and then reestimate Equation 1 using a placebo treatment date (1960), restricting the sample to the period prior to the treatment (all years before—and including—1970).

The parallel trends assumption is satisfied in every case. For both the inequality treatment and the relative poverty treatment, the trend in incarceration for the high treatment group is statistically insignificantly different from the low treatment group in the period prior to the prison boom.

A final robustness check concerns the degree to which selection on unobservables influences our coefficient estimates. Recognizing that assignment into the treatment category (i.e., the level of inequality) is potentially nonrandom, identifying the effect of mass incarceration on commuting zones of differing levels of inequality will—at minimum—require parallel trends in the pretreatment period. The results in Table 7 suggest this requirement is satisfied. However, we can further assess the degree to which selection on unobservables influences our results by applying the test developed by Oster (2019).

Oster (2019) shows that information about the degree of selection on *observables* can be used to bound coefficient estimates in the presence of selection on unobservables. In particular, Oster (2019) shows that under the assumption of equal selection—that the degree of selection on unobservables is proportionate to the degree of selection on observables—the bias-adjusted regression coefficient is given by:

$$\beta_{\text{bias adjusted}} = \beta_{\text{long}} - (\beta_{\text{short}} - \beta_{\text{long}}) \frac{R_{\text{max}}^2 - R_{\text{long}}^2}{R_{\text{long}}^2 - R_{\text{short}}^2},$$

where β_{long} and R_{long}^2 are the regression coefficient and R^2 estimate obtained from a regression including a full battery of controls, β_{short} and R_{short}^2 are the regression coefficient and R^2 estimate from a restricted regression including only the inequality terms, and R_{max}^2 is the maximum R^2 . Oster (2019) suggests $R_{\text{max}}^2 = 1.3 \times R_{\text{long}}^2$ as a conservative estimate for R_{max}^2 , as only 45% of nonrandomized results surveyed by Oster survive this threshold.

Table 8 presents the bias-adjusted coefficients for three of our key results: the effect of inequality on overall incarceration; the effect on non-White, non-Hispanic incarceration; and the effect of relative poverty on non-White, non-Hispanic incarceration. The adjusted coefficients suggest that selection on unobservables is not a primary

driver of our results. The estimated effect of inequality on overall incarceration and the effect of relative poverty on non-White, non-Hispanic incarceration are only slightly attenuated. In contrast, the effect of inequality on non-White, non-Hispanic incarceration is slightly elevated. Taken in conjunction with the results in Table 7, Table 8 suggests our estimates are robust to concerns about selection on unobservables.

Discussion and Conclusion

In this article, we explore the extent to which labor market inequality is associated with incarceration at the local level between 1950 and 2010. Using decennial IPUMS/Census microdata to construct estimates of the commuting zone incarceration rate—verifying our estimates with recently released data on incarceration by county of admission from the Vera Institute of Justice—we examine the impact of inequality on the overall incarceration rate in a commuting zone, as well as the varying impact of inequality on incarceration by race. Although income inequality is associated with higher rates of incarceration for all race and ethnicity groups (although not always in statistically significant fashion), the effect is largest for non-White, non-Hispanic individuals. Furthermore, when we examine the relationship between relative poverty and incarceration, we find a strong, positive relationship between White relative poverty (measured as the ratio of the White poverty rate to the non-White poverty rate) and non-White incarceration, particularly of non-Hispanic Blacks. We find no relationship between White relative poverty and White incarceration. Our results have several important implications.

First, our results are broadly consistent with Alexander's claims in *The New Jim Crow* that mass incarceration and the "War on Drugs" arose following the Civil Rights Movement as attempts by White elites to reestablish the system of racialized social control that was disrupted when Jim Crow was dismantled. The positive relationship between inequality and incarceration fits within the theoretical framework for understanding racial inequality coming out of stratification economics (Darity, 2005; Darity et al., 2015), which suggests that race- and class-based discrimination represent a rational attempt by privileged groups to maintain their relative status and the material benefits, which that status confers. To the extent that rising economic inequality confers greater rents to people in positions of privilege, local elites have an incentive to increase their stratification efforts—including via more restrictive penal policy. The positive relationship between local incarceration rates for non-Whites and White relative poverty is consistent with Alexander's (2010) argument that at least one of the political motivations for the rise of tough-on-crime policy after the Civil Rights Movement was to function as a racial bribe that attracted poor Whites away from newly formed, otherwise diverse, working-class political coalitions. Our results constitute strong empirical evidence in favor of the Alexander Hypothesis, which asserts that the rise of incarceration in the United States beginning in the 1970s can be viewed as an effort by White males, especially those of lower socioeconomic status, to maintain the privileges they experience from being White. This result, however, does not come without caveat. Given the reduced-form nature of our empirical model, we

cannot directly speak to the role of intention—an important part of the stratification narrative. Thus, although our results are consistent with the Alexander Hypothesis, additional evidence regarding the degree to which intent drives the mechanisms behind the Alexander Hypothesis is desirable. In this sense, our article represents only a first step in the empirical evaluation of the stratification approach to mass incarceration.

Still, we argue that our findings of linkages between inequality and incarceration by race matter independently of whether discrimination is implicit or explicit. Recent work has recognized implicit bias as an impactful mechanism for understanding economic outcomes (e.g., Bertrand et al., 2005; Guryan & Charles, 2013; Jost et al., 2009). Bertrand et al. (2005) go as far as concluding,

However we test for it, implicit discrimination is not useful simply as a subtle alternative interpretation. If it is a powerful driver of discriminatory behavior, it should reshape the way we understand discrimination and alter our available spectrum of remedies. (p. 97)

Further highlighting complexity, Small and Pager (2020) add the concept of “perceived” discrimination as consequential for policy.

Whether or not there is explicit intent behind the relationships between inequality and disproportionate incarceration by race that we document, our results speak to the uneven social burden of rising economic inequality and the importance of positional concerns in determining who bears that burden. Although both Whites and non-Whites experienced an increase in incarceration as a result of rising inequality during the post-1970 prison boom, the magnitude of that increase—and the impact of inequality as a mechanism determining the incidence of that increase—were disproportionately born by racial and ethnic minorities. To the extent that our results on the relationship between relative poverty and non-White incarceration—although reduced form—reflect behavior motivated by race-based positional concerns on behalf of poor Whites, they suggest that (a) comparisons across groups may matter as much for positional concerns as comparisons across individuals and (b) Robert Frank’s (2005) claim that “positional externalities cause large and preventable welfare losses” is an understatement, given that additional group-based social multiplier effects are involved.

Finally, our results suggest links between economic inequality, race, and the criminal justice system, offering important insight for both those who would undertake future criminal justice reform efforts and those who work on economic policy. Our results are relevant for the starting point discussions initiated with the First Step Act of 2018 and for the importance of continued “steps” toward effective and equitable prison and criminal justice reforms. Our results are also relevant for discussions of appropriate policy responses to cases of police brutality directed at Black men such as Rodney King and George Floyd and the associated public protests.

We argue that heterogeneities affect the intersections of economic inequality and policing and prison outcomes. Reforms that aim to reverse course on mass incarceration are unlikely to have long-term success in the absence of actions that (a) assuage the identity-driven positional tensions, which led poor Whites to lend their support to the policies and politicians initially responsible for mass incarceration and (b) address

inequality so as to eliminate the rents associated with positions of racial and economic privilege. As long as hierarchies based on race and economic status confer benefits to those at the top, there will be an incentive for individuals to support institutions that maintain and enforce the existing pattern of resource allocation. Rather than issues to be addressed separately, economic inequality and racial inequality—including racial disparities in the criminal justice system—must be treated as part of the same problem.

Appendix A

External Validation of Incarceration Rate Data

We validate our incarceration rate estimates by examining (a) whether our estimates replicate the movement in incarceration rates over time implied by the aggregate data and (b) whether our estimates are suitably representative at the local level. To proceed with the validation exercises, we make use of two alternative measures of incarceration. First, we use the aggregate incarceration rate—measured by the total number of people held in state or federal prisons or local correctional facilities—per 100,000 people compiled by the Bureau of Justice Statistics (BJS). The BJS data are compiled from several sources including the Annual Probation Survey, the Annual Parole Survey, the Annual Survey of Jails, the Census of Jail Inmates, and the National Prisoner Statistics Program. As our Census-based estimate of the incarceration rate focuses on the working-age population, we make use of the BJS data from 1980 onward, for which aggregate incarceration rates for those 18 and older are calculated separately, allowing for a closer comparison between the two rates.

Second, we use a recently released data set on incarceration rates by county of admission from the Vera Institute of Justice.¹² The Vera Institute data compile county-level estimates of prison populations using the BJS National Corrections Reporting Program. These data are then merged with county jail population estimates from the Census of Jails and the Annual Survey of Jails to construct a national data set allowing the examination of jail and prison populations at the county level. The Vera prison estimates are available from 1983, and the Vera jail estimates from 1970. The data cover the working-age population 15 to 64 years. An important limitation of the Vera data concerns the lack of availability for many counties in certain years. This limitation results from the voluntary nature of the National Corrections Reporting Program. As an example, in 1983—the first year of the Vera prison data—county-level prison populations are only available in 20 states, and not every county within each state is represented. This means that the overall geographic coverage of the Vera data is less than that which is achievable with the IPUMS/Census data. By our estimate, the average number of commuting zones available in the Vera data on a decennial basis over the 1980 to 2010 period—after aggregating from the county level—is approximately 562.¹³

Despite these shortcomings, the Vera data nonetheless constitute an important source of external validation for the IPUMS/Census estimates, primarily because of the difference in where prisoner location is recorded in the two sources. It is well known that the Census reports an individual's location based on their "usual residence." The residence

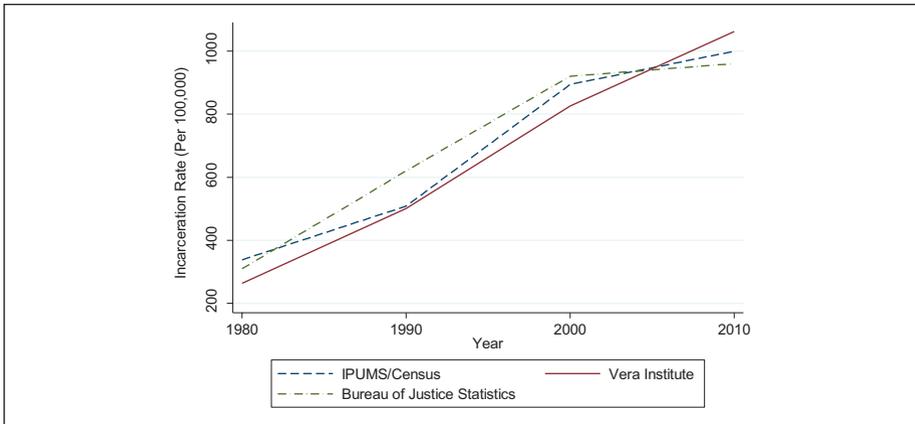


Figure A1. Data validation, comparison of aggregate trends.

Note. The figure presents estimates of the total number of incarcerated persons—including individuals in state or federal prisons and local correctional facilities—per 100,00 population, from each alternative data source.

rule results in the location of incarcerated persons coinciding with the location of the facility they are in, rather than their “true” home address outside of prison. In contrast, the Vera data estimate prison population by county of admission, rather than facility location. At the state level, this distinction should not make much difference. Federal prisoners make up only a small proportion of the overall incarcerated population, such that state-level totals should be roughly the same whether the admission location or holding location is used in the count. However, as geographic detail is refined, the distinction between admission location and holding location becomes increasingly important. At low levels of geographic aggregation, counts will potentially differ greatly depending on which location rule is used. The Vera data thus allow us to make a crucial check of whether the IPUMS/Census data are suitably representative at the commuting zone level. If commuting zones are at a sufficiently high level of geographic aggregation—such that the impact of using alternative location rules is muted—the Vera Institute data and the Census data should be highly correlated.

We perform a validation exercise by comparing aggregate trends in incarceration across our three data sources from 1980 onward. Specifically, we compare the aggregate totals from the BJS with annual averages across commuting zones from the IPUMS/Census data and Vera data. Figure A1 presents the results of this comparison.

The trends presented in Figure A1 indicate that both the IPUMS/Census data and the Vera Institute data successfully replicate the aggregate incarceration trends in the BJS data. The ex ante correlation between all three measures at the aggregate level over the 1980 to 2010 time period is close to .98.

Next, we assess the extent to which the IPUMS/Census data are suitably representative at the commuting zone level. To do this, we adopt two approaches. First, we present a binned scatterplot of the IPUMS/Census data against Vera data to examine

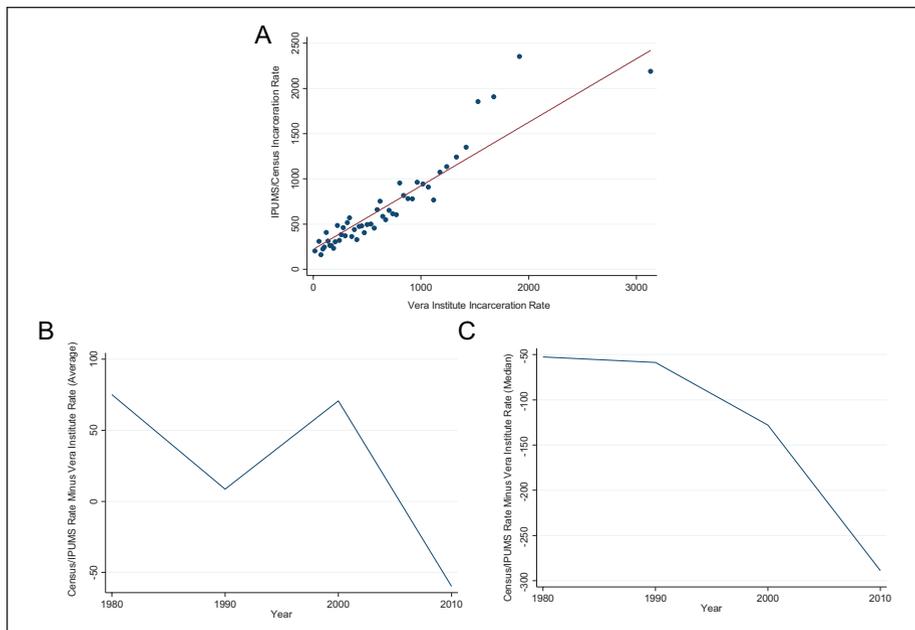


Figure A2. Data validation, local representation. (A) IPUMS/Census rate versus Vera Institute rate. (B) Average difference. (C) Median difference.

Note. Figure A2(A) presents a binned scatterplot of the IPUMS/Census incarceration rate against the Vera Institute incarceration rate. Figure A2(B) and A2(C) present the average and median differences between the Census rate and the Vera rate, respectively.

the relationship between the two across all time periods. Second, we calculate the difference between the IPUMS/Census data and the Vera data for each commuting zone, and plot the change in the median and average of this difference over time across all commuting zones to assess whether there are serially correlated measurement differences between the two. Figure A2 presents each of these plots.

Figure A2(A) provides prima facie evidence that the Census incarceration measure is highly representative at the commuting zone level, indicating that the Census incarceration measure and the Vera incarceration measure are highly correlated across commuting zones. However, despite the strong, positive correlation between the two measures, there are nonetheless differences between the Census incarceration estimates and the Vera incarceration estimates over time. For example, whereas Figure A2(B) indicates that the average difference between the two estimates was positive up until 2000, Figure A2(C) shows that the median difference has been negative over the entire time period in which both estimates were available. Furthermore, although the average difference between the two measures is small relative to the sample average incarceration rate (the average difference over the whole sample is 23 people per

100,000), the median has been increasing in absolute value (becoming increasingly negative) over time. These results suggest that although the Census measure may underestimate the incarceration rate in most commuting zones, there are nonetheless large outliers in the tails of the distribution, which drive the average up, possibly due to the location of state and federal prison facilities. In the “Results” section, we adopt two alternative weighting schemes as sensitivity checks on our estimation to address the differences in measured incarceration rates between the Vera data and the Census data. We find that these weights make little qualitative difference to our overall results. The combination of these tests with the strong *ex ante* correlation between the Census data and the Vera data leads us to conclude that the Census data are acceptably representative at the commuting zone level.

Appendix B

Discrete Treatment Assignment

Table B1 presents results from an alternative version of our main regression specification where commuting zones are divided into discrete treatment and control groups. For simplicity, we apply a cutoff equal to the post-1970 sample mean for both the household income Gini coefficient and relative poverty. All commuting zones with post-1970 values above the sample mean are included in the treatment group for a given loss of privilege variable. All other commuting zones are included in the control group. Our modified estimating equation is thus

$$\text{Incarceration}_{ct} = \beta_0 + \beta_1 \text{Treat}_{ct} (1 = \text{high}, 0 = \text{low}) + \beta_2 \text{After}_t^{1970} + \beta_3 \left(\text{Treat} \times \text{After}^{1970} \right)_{ct} + \text{Controls} + \epsilon_{ct}.$$

The results are consistent with those presented in the body of the article. Commuting zones in the high inequality group experienced a larger increase in the incarceration rate—relative to the control group—with the largest effect for non-White non-Hispanic incarceration. Similarly, commuting zones above the sample mean level of White relative poverty experienced a larger increase in non-White non-Hispanic incarceration (although the estimated coefficient is statistically insignificant) with a large, statistically significant effect on Black incarceration. Interestingly, commuting zones in the treated group for White relative poverty appear to have experienced a decline in the White, non-Hispanic incarceration rate, suggestive of further unequal treatment. Nonetheless, we emphasize that the continuous treatment specifications presented in the body of the article are preferred. An increase in inequality or relative poverty may affect incarceration regardless of whether either variable exceeds the sample mean, and the inclusion of plausibly treated units in the control group is likely to bias the regression coefficients in Table B1 downward.

Table B1. Discrete Treatment Assignment, Results.

Variable	(1)		(2)		(3)		(4)	
	Overall		White, non-Hispanic		Non-White, non-Hispanic		Black, non-Hispanic	
GiniTreat × After ¹⁹⁷⁰	196.2** (86.61)		71.80** (34.14)		1,449.9** (619.7)		689.3 (1,182.9)	
N	4,332		4,332		4,332		4,303	
PovertyTreat × After ¹⁹⁷⁰	-70.86 (68.39)		-47.04* (23.59)		376.2 (511.5)		2,229.4* (1,321.9)	
N	4,332		4,332		4,332		4,303	

Note. Dependent variable is incarceration rate (per 100,000) by race/ethnicity category. All regressions include controls for race, total population, population growth, average household income, unemployment, average education, violent crime rate, and employment shares by industry. Standard errors in parenthesis, clustered at the state level.

* $p < .10$. ** $p < .05$. *** $p < .01$.

Appendix C

Extended Summary Statistics

Table C1 presents extended summary statistics for key variables in high- and low-inequality commuting zones before and after the treatment date. The summary statistics suggest that commuting zones with above average post-1970 inequality had higher initial incarceration rates, levels of inequality, and Black population shares. This is consistent with existing literature (e.g., Derenoncourt, 2019) that shows localities with higher initial Black population shares saw both greater inequality and greater mass incarceration. Inclusion of time-varying population shares by race in our empirical model, as well as commuting zone-level fixed effects, are intended to address this fact.

Table C1. Extended Summary Statistics.

Variable	< Mean Gini				≥ Mean Gini			
	Before 1970		After 1970		Before 1970		After 1970	
	M	SD	M	SD	M	SD	M	SD
Incarceration	281.85	378.4138	424.8195	455.9037	301.9224	303.3102	527.5611	583.1911
Gini coefficient	29.63697	3.428918	35.96181	3.229826	32.43195	4.191346	39.33157	3.725801
% White	93.7273	7.022331	86.87081	10.13246	84.41029	9.777339	70.11293	15.30585
% Black	5.471013	6.904076	7.530693	8.002505	10.99993	9.07725	12.19483	8.585876
N	1,101		1,468		1,065		1,420	

Note. "Mean Gini" identifies the average post-1970 commuting zone-level household income Gini coefficient. "Incarceration" gives the Census incarceration rate per 100,000.

Acknowledgments

The authors thank fruitful conversations with colleagues following several conference and seminar presentations of earlier drafts, including at the American Economic Association Annual meetings, the Eastern Economic Association Annual meetings, and at Georgia State University's Andrew Young School of Policy Studies.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Luke Petach  <https://orcid.org/0000-0003-4626-0986>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. A search of recent issues of the *American Economic Review* revealed no articles with titles or abstracts containing the phrase “mass incarceration.” Only four papers with titles or abstracts containing the word “incarceration” were found.
2. We argue that the persistence and significance of the racial wealth gap points to the success of these efforts. Evidence from the literature supports this conjecture. Bayer and Charles (2018) show that once nonworkers (i.e., individuals who are incarcerated) are accounted for, the Black–White earnings gap is essentially unchanged from its 1950 level.
3. Although our results are consistent with the Alexander Hypothesis, the reduced-form nature of our empirical design precludes us from speaking directly to the *intention* of dominant social groups, a caveat that we discuss throughout this article.
4. For 1960, which is left out of the Autor and Dorn (2013) sample, we use the commuting zone crosswalk provided by Evan Rose, made available here: <https://ekrose.github.io/resources/>
5. To the extent that differences between our estimates and the Vera Institute data remain, the “Results” section weights observations according to the similarity between the two samples to address measurement error–induced bias in the regression estimates.
6. The violent crime rate is a commonly available measure of criminal activity that is correlated with both other crime rates—including drug crime—and the level of law enforcement. See, for example, Resignato (2010).
7. Appendix C presents additional summary statistics for treatment and control groups before and after the “intervention” year.
8. To economize on space, we report sample means for only select industries. Full data available upon request.
9. For a given commuting zone the variable treat_c is set equal to the post-1970 commuting zone average of the loss of privilege variable in all years, and is, therefore, constant—similar to the way a treatment indicator variable would be set equal to 1 in all years for treated units in a difference-in-differences setting with discrete treatment assignment. However, unlike a setting with discrete treatment assignment, treat_c varies across commuting zones—even within the “treated” group.
10. Inclusion of commuting zone–fixed effects should minimize concerns about the location of prison facilities and/or private prison dynamics, to the extent that the presence of such facilities is largely time invariant.
11. The interpretation of this coefficient is an increase in the White population share *relative to the omitted category*—in this case, all non-White groups. When additional population shares by race are included in the regression, such that the omitted category shifts to Black, non-Hispanics, the effect of the White population share on the incarceration rate of non-White, non-Hispanics becomes statistically significant.
12. <http://trends.vera.org/incarceration-rates>
13. This number may overstate the reliability of the Vera data, as it is likely that some commuting zone totals are underestimated due to missing totals in the underlying county-level data. Reporting to Vera is voluntary, and not all states and counties participate in the program. This selection is avoided in the Census data.

References

- Aizer, A., & Doyle, J. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *Quarterly Journal of Economics*, *130*(2), 759–904.
- Alexander, M. (2010). *The new Jim Crow: Mass incarceration in the age of colorblindness*. The New Press.
- Amior, M., & Manning, A. (2018). The persistence of local joblessness. *American Economic Review*, *108*(7), 1942–1970.
- Autor, D., & Dorn, D. (2013). The growth of low skill service jobs and the polarization of the U.S. labor market. *American Economic Review*, *103*(5), 1553–1597.
- Baicker, K., & Jacobson, M. (2007). Finders keepers: Forfeiture laws, policing incentives, and local budgets. *Journal of Public Economics*, *91*(11–12), 2113–2136.
- Bayer, P., & Charles, K. K. (2018). Divergent paths: A new perspective on earnings differences between Black and White men since 1940. *The Quarterly Journal of Economics*, *133*(3), 1459–1501.
- Becker, G. S. (1968). Crime and punishment: An economic approach. In N. G. Fielding, A. Clarke, & R. Witt (Eds.), *The economic dimensions of crime* (pp. 13–68). Palgrave Macmillan.
- Benson, B. L., Rasmussen, D. W., & Sollars, D. L. (1995). Police bureaucracies, their incentives, and the war on drugs. *Public Choice*, *83*(1–2), 21–45.
- Bertrand, M., Chugh, D., & Mullainathan, S. (2005). Implicit discrimination. *American Economic Review: Papers and Proceedings*, *95*(2), 94–98.
- Blumenson, E., & Nilsen, E. S. (1998). Policing for profit: The drug war's hidden economic agenda. *The University of Chicago Law Review*, *65*(1), 35–114.
- Bluthenthal, R., Lorvick, J., Kral, A., Erringer, E., & Kahn, J. (1999). Collateral damage in the war on drugs: HIV risk behaviors among injection drug users. *International Journal of Drug Policy*, *10*(1), 25–38.
- Chetty, R., & Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *The Quarterly Journal of Economics*, *133*(3), 1107–1162.
- Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, *129*(4), 1553–1623.
- Darity, S. (2005). Stratification economics: The role of intergroup inequality. *Journal of Economics and Finance*, *29*(2), 144–153.
- Darity, S., Hamilton, D., & Stewart, J. (2015). A tour de force in understanding intergroup inequality: An introduction to stratification economics. *Review of Black Political Economy*, *42*(1–2), 1–6.
- Darity, S., Hamilton, D., & Zaw, K. (2016). Race, wealth and incarceration: Results from the National Longitudinal Survey of Youth. *Race and Social Problems*, *8*(1), 103–115.
- Derenoncourt, E. (2019, December 31). *Can you move to opportunity? Evidence from the great migration* [Working Paper]. https://www.dropbox.com/s/l34h2avpjomylr/b/derenoncourt_2019.pdf
- Frank, R. (2005). Positional externalities cause large and preventable welfare losses. *American Economic Review*, *95*(2), 137–141.
- Frost, N. A., & Clear, T. R. (2018). Theories of mass incarceration. In J. Wooldredge, & P. Smith (Eds.), *The Oxford handbook of prisons and imprisonment* (pp. 104–124). Oxford University Press.

- Guryan, J., & Charles, K. K. (2013). Taste-based or statistical discrimination: The economics of discrimination returns to its roots. *The Economic Journal*, 123(572), F417–F432.
- Hamilton, D., & Darity, W. A., Jr. (2010). Can “baby bonds” eliminate the racial wealth gap in putative post-racial America? *The Review of Black Political Economy*, 37(3–4), 207–216.
- Hamilton, D., & Darity, W. A., Jr. (2017). The political economy of education, financial literacy, and the racial wealth gap. *Federal Reserve Bank of St. Louis Review*, 99(1), 59–76. <http://dx.doi.org/10.20955/r.2017.59-76>
- Jost, J. T., Rudman, L. A., Blair, I. V., Carney, D. R., Dasgupta, N., Glaser, J., & Hardin, C. D. (2009). The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore. *Research in Organizational Behavior*, 29, 39–69.
- Kaplan, J. (2019). *Jacob Kaplan's concatenated files: Uniform crime reporting program data: Offenses known and clearances by arrest, 1960-2017*. Inter-University Consortium for Political and Social Research. <https://doi.org/10.3886/E100707V11-12733>
- Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior*, 34(1), 3–38.
- Koppl, R., & Sacks, M. (2013). The criminal justice system creates incentives for false convictions. *Criminal Justice Ethics*, 32(2), 126–162.
- Kuziemko, I., & Washington E. (2018). “Why did Democrats lose the South? Bringing new data to an old debate.” *American Economic Review*, 108(1), 2830–2867.
- Liu, S. (2018, December 11). *Incarceration of African American men and the impacts on women and children* [Working Paper]. https://www.econ.queensu.ca/sites/econ.queensu.ca/files/sitianJMP_paper.pdf
- Lynch, M. (2012). Theorizing the role of the “war on drugs” in US punishment. *Theoretical Criminology*, 16(2), 175–199.
- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business and Economic Statistics*, 37(2), 187–204.
- Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology*, 108(5), 937–975.
- Pettit, B., & Gutierrez, C. (2018). Mass incarceration and racial inequality. *American Journal of Economics and Sociology*, 77(3–4), 1153–1182.
- Pettit, B., & Western, B. (2004). Mass imprisonment and the life course: Race and class inequality in US incarceration. *American Sociological Review*, 69(2), 151–169.
- Piketty, T. (2014). *Capital in the twenty-first century*. Harvard University Press.
- Piketty, T., Saez, E., & Zucman, G. (2018). Distributional national accounts: Methods and estimates for the United States. *Quarterly Journal of Economics*, 133(2), 553–609.
- Resignato, A. (2010). Violent crime: A function of drug use or drug enforcement? *Applied Economics*, 32(6), 681–688.
- Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Jose, P., & Sobek, M. (2019). *IPUMS USA* (Version 9.0, data set). IPUMS. <https://doi.org/10.18128/D010.V9.0>
- Saez, E., & Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2), 519–578.
- Simon, J. (2012). Mass incarceration: From social policy to social problem. In J. Petersilia, & K. R. Reitz (Eds.), *The Oxford handbook of sentencing and corrections* (pp. 23–52). Oxford University Press.

- Small, M. L., & Pager, D. (2020). Sociological perspectives on racial discrimination. *Journal of Economic Perspectives*, 34(2), 49–67.
- Tolbert, C., & Sizer, M. (1996, September). *US commuting zones and labor market areas: A 1990 update* (USDA Economic Research Service Staff Paper, Number 9614). <https://usa.ipums.org/usa/resources/volii/cmz90.pdf>
- Wakefield, S., & Uggen, C. (2010). Incarceration and stratification. *Annual Review of Sociology*, 36, 387–406.
- Western, B., & Pettit, B. (2010). Incarceration and social inequality. *Daedalus*, 139(3), 8–19.
- Western, B., & Wildman, C. (2008). Punishment, inequality, and the future of mass incarceration. *University of Kansas Law Review*, 57, Article 851.
- Yates, J., & Whitford, A. B. (2009). Race in the war on drugs: The social consequences of presidential rhetoric. *Journal of Empirical Legal Studies*, 6(4), 874–898.