2011

The Myths and Realities of Correctional Severity: Evidence from the National Corrections Reporting Program on Sentencing Practices

John F. Pfaff
Fordham University School of Law, jpfaff@law.fordham.edu

The Myths and Realities of Correctional Severity: Evidence from the National Corrections Reporting Program on Sentencing Practices

John F. Pfaff, *Fordham Law School*

Send correspondence to: John F. Pfaff, Fordham Law School, 140 West 62nd Street, New York, NY 10023, USA; Tel: 212-636-7661; Fax: 212-636-6899; E-mail: jpfaff@law.fordham.edu.

The forces driving U.S. prison growth are poorly understood. This article examines one factor that has received insufficient attention: changes in time served. It demonstrates that time served has not risen dramatically in recent years, even declining in some jurisdictions. It also shows that time served is fairly short: median release times are approximately one to two years. Thus, admissions practices, not longer sentences, appear to drive prison growth. This article also examines whether time served varies across different types of inmates. Young, Hispanic, and violent offenders appear to serve longer sentences; race and sex appear to be of minor importance. (*JEL* K14, K42)

Thanks to Bernard Black, Shawn Bushway, Jeffrey Fagan, Sam Gross, Katherine Litvak, Anup Malani, Tom Miles, J.J. Prescott, Dan Richman, Margo Schlanger, two anonymous referees, and workshop participants at the 2008 Conference on Empirical Legal Studies, Fordham Law School, the University of Michigan Law School, and the University of Texas Law School for helpful comments. Thanks also to Christopher Pece of the U.S. Census Bureau for providing me with access to the Annual Survey of State and Local Government Finances, and to Damien Lyons and Christina Schweikert of Fordham's Computer Science Department for access to a sufficiently powerful computer. Anne Boustead, Ryan Kaupelis, Sean Koehler, and Satinder Singh provided valuable research assistance. All remaining errors are my own.

American Law and Economics Review
doi:10.1093/aler/ahr010
Advance Access publication September 15, 2011

1. Introduction

The scope of incarceration in the United States is hard to underestimate. Approximately 1.6 million people are currently in prison, an incarceration rate of ∼700 per 100,000, or nearly 1 out of every 130 Americans (and about one out of every twenty black men).[1] The United States possesses only 5% of the world's population, but it houses one-third of its prisoners (Walmsley, 2007), and cash-strapped states are spending nearly $40 billion per year to incarcerate them. This is a new development: The U.S. incarceration rate was a steady 100 per 100,000 people for nearly fifty years, from 1920 (when the statistics were first gathered) through 1970, at which time there were only 300,000 people in prison. The subsequent quintupling of prison populations is an event previously unseen here or abroad.

Yet equally remarkable is how little we actually understand about causes of this buildup. Researchers have posited a wide array of insightful theories, and in some cases they have mustered anecdotal or historical evidence in corroboration. But there is scant rigorous empirical evidence, and what work has been done has primarily taken the total stock of prisoners as the dependent variable. As I argue in Pfaff (2008), this stock variable is shaped by a host of concurrent and lagged factors that cannot be easily modeled,[2] and it is ultimately unclear if these efforts have produced any well-identified estimates.

The flow variables of admissions and releases are substantially more conducive to empirical analysis. A few empirical studies have looked at admissions (see Marvell, 1995; Jacobs and Helms, 1996; Marvell and Moody, 1997, 1996; Sorensen and Stemen, 2002; Listokin, 2003; Nicholson-Crotty, 2004; Stucky, Heimer, and Lang, 2005),[3] but none has rigorously examined releases. This article starts to fill this gap, as part of a broader empirical

1. The recent report by the Pew Center on the States' (2008) Public Safety Performance Project puts the incarceration rate at 1 out of every 100 adults. The rate is 1 out of 130 Americans of all ages.

2. The prison population in 1995 is a function of criminological, economic, political, and demographic conditions in 1995, 1994, 1993, and so on into the past. The models done so far, however, essentially regress, say, unemployment in 1995 on the stock of prisoners in 1995, and it is thus unclear what the coefficient really means.

3. Note that all but Listokin (2003) suffer from substantial identification problems (Pfaff, 2008).

project examining trends in American prison populations over the past thirty years. Despite the claims and anecdotes—in media and academic writings alike—about the severity of contemporary sentencing policy, we have at best only a weak understanding of what patterns in time served actually look like. This article provides a rich descriptive account of time served in prison, demonstrates that prison growth has been driven primarily by changes in admission policy rather than release policy, and examines how inmate characteristics such as race, ethnicity, sex, age, and conviction offense influence time served.

To date, investigators have made only rough examinations of release policy, looking at releases through crude approximations of time served, such as dividing the number of prisoners admitted in a given year by that year's total prison population (see, e.g., Blumstein and Beck, 1999).[4] Taking advantage of the National Corrections Reporting Program (NCRP), a highly detailed—and underused—data set, I can map (to the day) how the time served by individual inmates changes over a ten- to twenty-year period in a sample of eleven states. As a result, I am able to paint a much more detailed and nuanced picture of the time spent in prison and its effect on overall prison growth.

The results stand in stark contrast to conventional wisdom, which often argues that prison growth has been driven by increasingly punitive, "tough-on-crime" legislatures passing harsher and harsher sentencing laws—such as mandatory minimums, truth-in-sentencing (TIS) laws, two- and three-strike laws, and parole abolitions—even as crime rates entered their long decline in the 1990s. Media accounts of sentencing focus heavily on shockingly long sentences for relatively minor crimes. Academics have similarly emphasized longer sentences: Blumstein and Beck (1999), for example, argue that 57.7% of the growth in prison population between 1980 and 1996 was due to longer sentences.

4. Thus, if 10,000 people were admitted in 1990 and the total prison population was 100,000, this formula returns an average sentence length of ten years. This would be an accurate estimate were the prison population in equilibrium, but during transition periods it is biased.

The results from the NCRP tell a much different story. In many states, the median time served has declined over much of the 1990s; so too has the 75th percentile of time served, and even the 90th percentile in some cases. As I demonstrate below, it is undoubtedly clear that the United States has become more punitive over the past thirty years, but rarely with respect to sentence length actually served. It is our willingness to incarcerate in the first place, not to keep people in prison once admitted, that appears to have been the fundamental engine of prison growth since the late 1980s and early 1990s.[5]

It is important, however, not to overstate this point. First, the rise in admissions has been driven at least in part by the incarceration of increasingly minor offenders, who likely serve shorter sentences; after all, prison populations grew during the longest recorded crime drop in American history. Comparing time served by those admitted in 1988 to time served by those admitted in 1998 is thus challenging since the two groups are not identical—a greater fraction of those admitted in 1998 would be serving short(er) sentences even if sentencing policy had grown more severe. Controlling for this shift in the distribution of offenders yields less dramatic results, but ones that still suggest time served has frequently remained relatively constant or even declined.

Second, other countervailing trends may mask the effect of increasingly punitive practices. Some states, for example, parole prisoners quickly but violate them back aggressively—the median prisoner in California serves only ~180 days, but California violates nearly 30% of its parolees back to prison in any given year (accounting for ~43% of all parolees violated back in the United States).[6] If a large number of inmates returns to prison at least once, is the median time 180 or 360 days? Also, capacity may limit time served, so were admission rates to decline, total population may decline

5. Because of limitations in the data, I cannot compare the relative importance of admissions to that of time served prior to the late 1980s (no state in my sample provides reliable data prior to 1987, and in two cases not before 1992). Figure 1b, however, suggests that increasing time served may have played a more important role in prison growth around the late 1980s.

6. This is an approximation—180 days comes from the 2002 NCRP data, while the parole data come from the 2006 BJS survey of parole practices (Glaze and Bonczar, 2007).

more slowly—with more capacity available, states may force prisoners to serve longer fractions of their (longer) sentences.

And third, aggregate state-level quantiles may mask important differences in time served across various types of prisoners. To explore this issue more carefully, I develop offender-level survival models to examine the effects of race, ethnicity, sex, age, and (broadly defined) offense type on time actually served.[7] My aggregate results appear to overstate time served by nonviolent and by elderly inmates, and they understate time served by the young and by the Hispanic. No other trait—including race—appears to significantly influence time served.

The results in this article shed new, and important, light on U.S. penal policy. First, they indicate that our attention to sentencing[8] matters is at least partially misplaced. While institutions such as the Families Against Mandatory Minimums and the Justice Kennedy Commission (2004) challenge sentencing regimes such as mandatory minimums, most of the growth in prison population has come through admissions, an area that receives substantially less attention; and the one facet of admissions that receives perhaps the most attention—the incarceration of low-level drug offenders—is not as important as that attention suggests.[9] Tough-on-crime activists likewise focus on passing increasingly draconian sentences that appear to have little effect on actual outcomes. Reformers on both sides are looking in the wrong place.

Second, these results suggest that, despite its great size, the U.S. prison population need not be particularly stable. Rising admissions drive the current growth in prisons, and admission rates—unlike release rates—can

7. Though employment status or, say, education could affect time served as well, the five traits listed here are the only five consistently reported in the NCRP. Particularly disappointing is the NCRP's failure to provide data on prior convictions: Though officially included in the data, the entry is blank for every inmate in every state in every year.

8. To be clear, I use "sentencing" to refer to choices about how long to incarcerate someone already heading to prison, not the decision whether to incarcerate in the first place; that falls under "admissions." The line, however, is not always clear—admission can be thought of as raising time served from zero to something greater than zero.

9. While the number of drug offenders in prison, and their share of the total prison population, has risen over the years, their importance peaked at ~20% of the nation's total population. The incarceration of low-level violent and property offenders thus appears to be much more important.

change immediately.[10] Again, though, it is important not to overstate this point. A small cadre of very long-serving inmates could exert a disproportionate effect on both the size and the rate of change of the prison population. As a result, a one-year admissions freeze in, say, California, where the 75th percentile time to release is one year, need not result in the prison population falling by 75%.[11]

This article proceeds as follows. Section 2 provides a detailed look at time served by inmates in those states providing sufficient data to the NCRP. Section 3 then develops two counterfactuals that demonstrate more sharply the substantially larger role changes in admissions, as compared with changes in releases, have played in driving overall prison population sizes. Section 4 addresses the relative importance of changes in parole violation policies. And Section 5 constructs a more rigorous empirical model to estimate the extent to which the general results given in Section 2 vary across different types of inmates. The Appendix discusses the data and the technical aspects of the empirical models in more detail.

2. Prison Release Trends in the United States

Our understanding of sentence lengths and release policy is driven more by anecdote and crude empiricism than by extensive examinations of the data. In general, precise release behavior is not readily observable in much

10. Changes in sentencing policy generally operate with a lag. Reducing the sentence for burglary from ten years to six years will have no effect for six years (ignoring parole), while changing the admissions rate for burglary reduces the prison population instantly.

11. Consider the following example. A state admits 100 prisoners per year: 25 serve three-year terms, 25 serve one-year terms, and 50 serve six-month sentences. The median time served is six months, and the 75th percentile one year, as in California. In equilibrium, the prison population at any given point is 125. If the state imposed a one-year ban on admissions, total population would immediately drop to 50, a 60% decline (despite 75% of all prisoners being released within one year). The longer the right tail, the smaller the effect of the one-year freeze: If the 25 long timers served four years, the one-year freeze would result in a 43% drop from 175 to 100; if they served ten years, a 25% drop from 300 to 225. (Conversely, if they served two years, the drop would be the full 75%.) I examine this issue more fully in Pfaff (2010), and I find that such long-serving inmates do not impose too strong a barrier to real reform.

of the data available on the correctional system. The Bureau of Justice Statistics (BJS) reports the number of inmates released each year in every state, but these aggregate values provide no information on the time served by those released, and time-served estimates such as Blumstein and Beck's (1999) rest on troubling assumptions. To overcome this problem, this article uses data gathered by the NCRP to develop substantially more precise measures of time served. The NCRP is an offender-level data set, in which participating states submit data on each offender as he enters and leaves prison.[12] As a result, I can calculate to the day the exact time served by each inmate released from prison, and I can count the number of offenders admitted in a particular year who are still serving time in a future year. The NCRP also provides demographic, offense, and sentencing information on each inmate. As I explain in the Appendix, the NCRP is not without its limitations, but it provides the clearest picture of releases available at the national level.

The NCRP began gathering data in 1983 and continues to this day; this article uses data through 2002, the most current release available when it was started. I use data from the eleven states that consistently reported data for sufficiently long periods of time during the sample period: California, Colorado, Illinois, Kentucky, Michigan, Minnesota, Nebraska, New Jersey, South Dakota, Virginia, and Washington. Together, they contain ~30% of the total prison population in United States.

This section presents the trends in median, and 75th and 90th percentile times to release.[13] It also examines the challenge changing admissions policies pose for interpreting these trends. Subsequent sections use the NCRP to (1) develop two counterfactuals to demonstrate more clearly the generally minor

12. These data are on state prisoners held in state-run prisons. Thus, a federal inmate housed in a state prison does not appear here (nor does a federal prisoner in a federal prison), nor do any convicted inmates serving time in local jails.

13. Looking beyond the 90th percentile would be informative, but it is hard to do: In all but one state in my sample, for example, the 99th percentile prisoner from any entering cohort has yet to be released. That life sentences and death sentences alone comprise ~3% of all violent sentences, and that violent offenders make up ~50% of all inmates, suggests that a core of offenders serves very long sentences. Pfaff (2010) examines this issue in more detail.

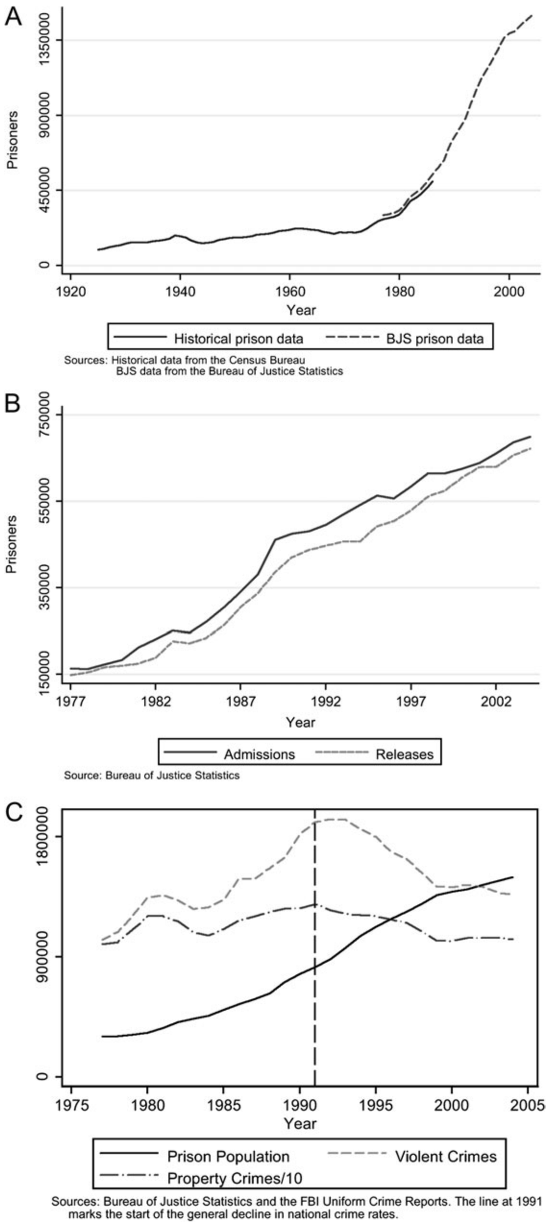

Figure 1. (a) Prison population 1925–2004. (b) Admissions and release 1977–2004. (c) Prison population and crime 1977–2004.

effect of changes in time served on prison population sizes, (2) consider the possible importance of parole and probation violations, and (3) explore the importance of inmate-specific traits on time served.

Even before turning to detailed release data, it is easy to demonstrate that the conventional wisdom about increasing sentencing severity faces some challenges. Figure 1a provides the startling, but well-known, image of prison growth in the United States, with the placid flatness of the 1920s to the 1970s giving way suddenly to the steady and unrelenting growth of the past thirty years.

It is somewhat disappointing that Figure 1b does not have the same fame. This figure decomposes the surging part of Figure 1a into its two primary pieces: admissions and releases.[14] Increasingly severe sentencing practices should manifest themselves in a widening gap between admissions and releases. The story told in Figure 1b is mixed. The growing gap during the late 1980s is consistent with increased sentencing severity. But Figure 1b also suggests that sentencing has not become additionally punitive since then, and if anything has becomes somewhat less harsh.[15] This is surprising. Zimring (2001), for example, defines the 1980s as a "lock-'em-up" period, in contrast to the "throw-away-the-key" approach of the 1990s. And many seemingly tough sentencing reforms—such as three-strike laws and TIS laws—were widely adopted in the 1990s. But despite these legislative reforms, actual sentencing outcomes appear to have remained relatively stable, at least for the last ten or fifteen years.[16] Thus, even before delving into the more detailed NCRP data, we should not expect to see a profound shift in recent sentencing outcomes—although as I explain more

14. For any year, the gap between the two lines in Figure 1b equals the amount by which the line in Figure 1a rises.

15. As will soon become apparent, my results differ from those by Blumstein and Beck (1999), who allocate nearly 60% of the prison growth since the 1970s to longer sentences. That I focus primarily on the late 1980s and the 1990s—solely due to limitations in the NCRP—may explain our disagreement since their 60% is not disaggregated by period (although the 1980s and the 1990s appear to be distinctly different phases). Another partial explanation may be the underrepresentation of the South in my sample due to poor reporting to the NCRP by Southern states (which, if true, points to a limitation of drawing inferences from aggregate national-level data).

16. Some sentencing reforms may have been adopted for purely symbolic reasons. Dhammika Dharmapala, Nuno Garoupa, and Joanna Shepherd, "Sentencing Guidelines, TIS Legislation, and Bargaining Power" (working paper, 2006), for example, point out that numerous states adopted TIS laws after abolishing or restricting parole. Such TIS laws thus appear to be at least somewhat redundant or symbolic.

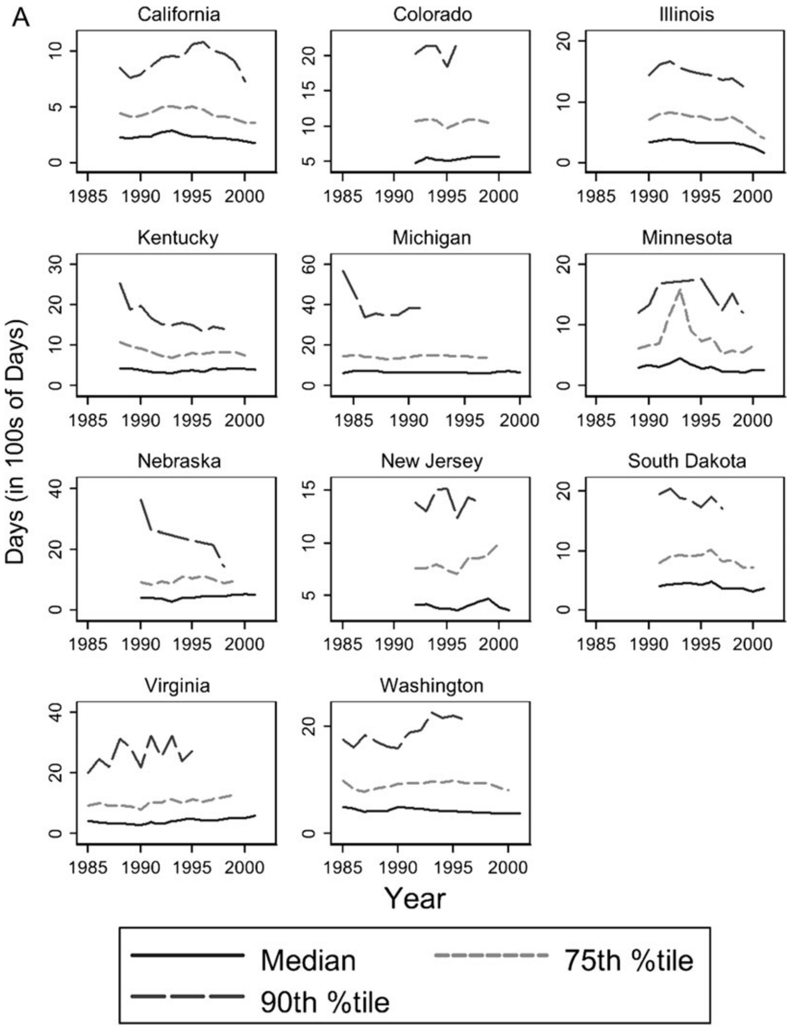

Figure 2. Days to release: (a) median and 75th and 95th percentiles; (b) 10th and 25th percentiles.

fully below, the trends in Figure 1c complicate this story somewhat: During the 1990s, prison populations rose while crime rates fell, suggesting that the severity of the average and marginal prisoner declined during this period.

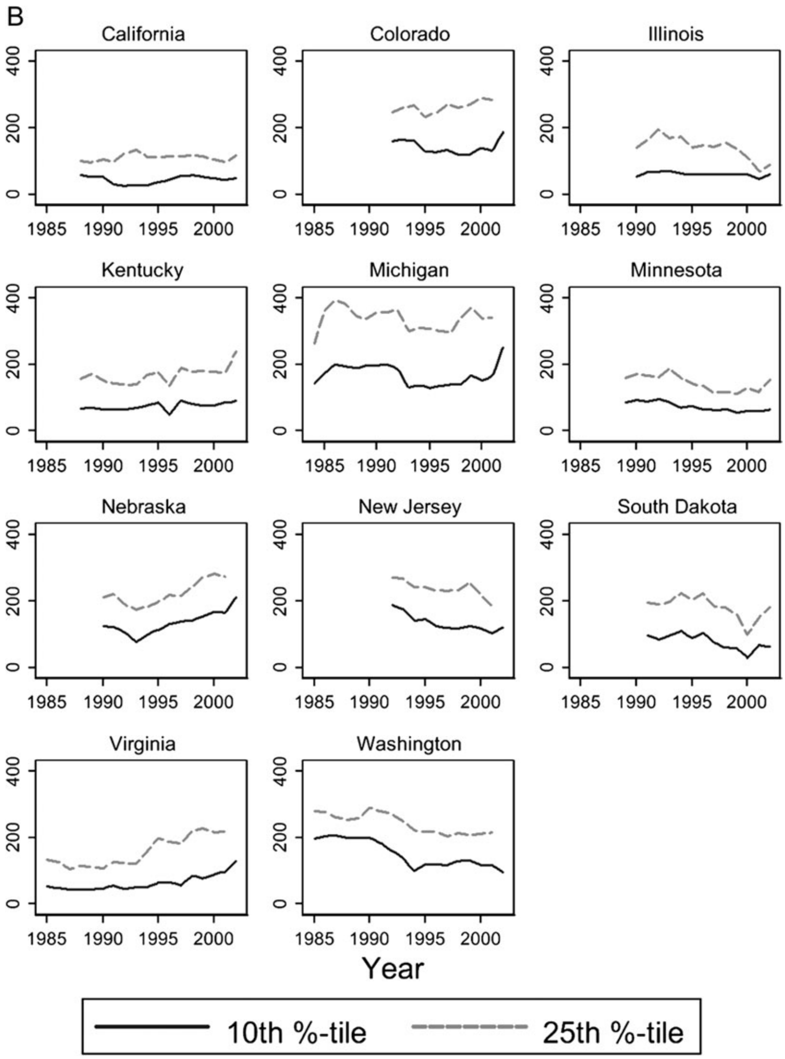

Figure 2. Continued.

But while Figure 1b is provocative, we can attack this issue more rigorously. Figure 2a provides times to release (in days) of the median, 75th percentile, and 90th percentile prisoners; Figure 2b completes the distribution by plotting these times for the 10th and 25th percentile prisoners.[17] Focusing for now on Figure 2a, two features stand out. First, the typical inmate does not spend a significant amount of time behind bars.[18] Median times to release are on the order of about one to two years, whether in relatively moderate states such as Minnesota or punitive ones like Michigan. In the big states, such as California and Illinois, the median time to release is well under a year, reaching lows of 179 days in California and 168 days in Illinois.

Seventy-fifth percentile times to release are often on the order of four to five years, though in some cases they are substantially lower—around two years in Minnesota and Illinois, and about one year in California. In most states in the sample, the seventy-fifth percentile time to release has remained relatively constant, although Virginia and New Jersey have seen it rise; Virginia is one of the few states in the sample to see its median time to release grow as well. These results are robust to breaking down trends by major crime categories (violent crimes, property crimes, and drug offenses) and even by offense subcategories.

At one level, these figures undermine the conventional wisdom about punitive penal policy on the release side. Despite the plethora of tough-on-crime sentencing laws passed in the 1980s and 1990s, time actually served—the metric that really matters—has not changed much in recent years. And the time that is served does not appear to be particularly long for the typical inmate.

However, two seemingly contradictory patterns in the 1990s complicate matters. As Figure 1c makes clear, while prison population rose with rising crime rates from the late 1970s through the end of the 1980s, the constantly rising prison populations of the 1990s occurred during a time of steadily

17. I focus on quantiles here since the mean time served by an entering cohort cannot be calculated until all its prisoners have been released. Moreover, quantiles are less influenced by outliers, and time served has a long, thin right tail. In Pfaff (2010), I turn my attention to averages and the role of these long-serving inmates.

18. Although using BJS data, my results are often lower than those reported in BJS reports that use the NCRP. I believe the difference arises from different frames of reference. The BJS computes the mean time served by those released in a given year (regardless of the year admitted), while I focus on the time served by those admitted in a given year (regardless of the year released). That I restrict my focus solely to states that provide complete data may also explain some of the difference (if the missing observations in other states are not randomly distributed).

declining crime rates. If more people are going to prison as fewer people commit crimes, increasingly marginal offenders must be entering our prisons at increasing rates. Thus, even if states aggressively applied tougher sentencing laws, the median could fall.[19] Controlling for this selection effect yields less dramatic results, but it does not qualitatively undermine the claim that release policy has not grown noticeably more severe over the past ten to fifteen years.

Figure 2b sheds light on this issue. If increasingly marginal offenders are driving prison growth (and thus masking longer stints being served by more serious inmates), the 10th and 25th percentiles of time served should decline over the sample period. Such a pattern exists, to some degree, only in New Jersey, South Dakota, and Washington. A slightly more complex way to address the issue entails looking at trends in three factors. The marginal offender hypothesis is more likely to be correct if we observe (1) a rising share of short-term offenders (e.g., those serving less than one year), (2) declining time served by those short-termers, and (3) rising trends in time served by those serving longer sentences (more than one year). Though not reported here, these trends reinforce Figure 2b's claims about New Jersey, South Dakota, and Washington. But they also suggest that the marginal offender effect may be particularly strong in California, and that it is present, to a much lesser degree, in Illinois (which appears to be locking up more low-level offenders but not imposing longer sentences on more serious criminals) and Michigan. The marginal offender effect appears to be absent in Colorado, Kentucky, Minnesota, Nebraska, or Virginia, but then these states do not exhibit much of a decline in time served in Figure 2a.

Though only a rough means of addressing the marginal offender effect, these results imply that to the extent time served appears to decline (as

19. A simple thought experiment makes this clear. A state locks up one serious offender for ten years each year, so the median time served is ten years. Later, the state continues to incarcerate the one serious offender, but it raises his sentence to twenty years, and it starts to incarcerate two minor offenders every year for one-year terms each. The median sentence length will fall to one year, despite the fact that the total equilibrium prison population rises from 10 to 22, with 83% (10/12) of that growth due to the longer sentences for the serious offenders—and with all offenders facing increased sentences (since the minor offenders also see their sentences rise, from zero years to one year).

opposed to stagnate), some of the decline is ephemeral. What appears to be a decrease in time served is partially the result of (1) locking up more and more minor offenders—who previously would not have been incarcerated at all—for very short periods and (2) locking up more severe offenders for longer. Even so, few states exhibit strong upward swings in the time served by long-serving inmates, suggesting that the increased willingness to incarcerate minor offenders has not been matched by an increased willingness to lock up more serious offenders longer.[20] Despite the attention paid to three-strike laws, TIS laws, and the like, the predominant locus of increased punitiveness over the past ten to fifteen years still appears to be at the low end, not the high.

3. Two Counterfactual Experiments

The above results indicate that time served has remained relatively low and stable, but they do not clearly illuminate the weak relationship between time served and prison population growth. To better measure how admissions and release policies have shaped prison populations, I develop two counterfactuals in this section. Moving beyond simple quantiles, I calculate the percent of prisoners admitted in one year that is released in each future year, and I track how these year-by-year distributions move over time. With these numbers, I derive two informative comparisons. First, I compare the actual net increase in prisoners in each state from the first year of available data to 2002 to the hypothetical net increase that would have occurred had the state released prisoners at the same rate every year as it did in the first year. Call this the "fixed release rates" counterfactual. Second, I compare the actual net increase to the hypothetical net increase that would have occurred had the state admitted the same number of prisoners in each subsequent year as it did in the first year. Call this the "fixed admissions" counterfactual.[21]

20. In other words, effect (1) is stronger than effect (2).

21. In an earlier version of this article, I computed several other counterfactuals, such as fixing only the distribution of offenses, fixing the distribution of offenses and release rates, and fixing the distribution of offenses, the release rates, and the size of the admissions class. These additional counterfactuals do not provide much additional insight beyond that captured in the two I present here.

The values in these counterfactuals require some explanation. My goal is to understand what prison populations would have looked like had the rates of releases or the number of admissions not changed over the sample period. To do this, I first calculate how many extra prisoners each entering cohort contributes to each subsequent year: How many prisoners admitted in 1988 are still in prison at the end of 1988, of 1989, of 1990, and so on. I then recalculate these numbers in one of two ways—by assuming that the distribution of release rates is constant across cohorts or that the entering size is constant across cohorts.

In other words, any given point on the "actual data" curve is equal to:

$$\mathrm{pop}_{t_1} = \sum\nolimits_{t=t_0}^{t_1} \left(1 - \sum\nolimits_{t=t_0}^{t_1-t_0} p_t(\tau)\right) \mathrm{admit}_t. \tag{1}$$

Here, $p_t(\tau)$ is the fraction of prisoners[22] admitted in year t released τ years after admission. The quantity is summed over all the years between the t_0, the first year for which data are available for that state, and the observation year t_1. pop thus measures the net increase in prison population between t_0 and t_1. The fixed release rates counterfactual sets $p_t(\tau) = p_{t_0}(\tau)$ for all t (not τ), and the fixed admissions counterfactual freezes admit_t at admit_{t_0} for all t.

A concrete example can make this clearer. Assume that of all the prisoners admitted in year t, 50% are released that year, 30% in year $t + 1$, 10% in year $t + 2$, and 10% in year $t + 3$ (so no prisoner serves more than three years). Then, for each admission year t, $p_t(0) = 0.5$, $p_t(1) = 0.3$, $p_t(2) = 0.1$, $p_t(3) = 0.1$, and $p_t(\tau) = 0$ for all $\tau \geq 4$. Table 1 provides a hypothetical example in this state over five years. The value of 200 in the first row of data for 1991 indicates that by the end of that year, only 200 of the 1,000 prisoners admitted in 1990 remain in prison (since 500 are released in 1990 and 300 in 1991). The total of 950 for that year implies that by the end of 1991, prison admissions since 1990 have added a net 950 inmates to the system. In other words, 950 = (1 – 0.5 – 0.2) (1,000) + (1 – 0.5) (1,500). This is Equation (1).

22. I allow the fraction to vary across five broad crime categories (violent, property, drug, miscellaneous, and unknown) and sum across categories to generate a single total population number.

Table 1. Hypothetical Releases

		Remaining Prisoners				
Year	Admits	1990	1991	1992	1993	1994
1990	1,000	500	200	100	0	0
1991	1,500		750	300	150	0
1992	2,000			1,000	400	200
1993	2,500				1,250	500
1994	3,000					1,500
Total		500	950	1,400	1,800	2,200

Note what Table 1 does not tell us. The "Total" row for 1991 does not give the actual prison population in 1991, nor does it give the change in the actual prison population since the start of time series (here, 1990). Instead, it indicates how admissions and releases since 1990 have changed the prison population. The prison population in 1991 is surely larger than 950 since some prisoners admitted in 1989 are still serving time that year. And part of the real change between 1990 and 1991 is due to prisoners from 1989 and 1988 being released. The "Total" row picks up neither of these forces.

In effect, the "Total" row assumes that the prison population equals zero when the data start and calculates what the prison population would be in subsequent years. This is done solely as a concession to limitations in the NCRP data,[23] but it nonetheless yields informative numbers. This is particularly true in later years, where—given the relatively short terms actually served by prisoners in the United States—the "total" values converge on the actual sizes of prison populations; by 2002, for example, very few pre-1990 prisoners are in the system.[24]

The fixed release rates counterfactual, then, compares two calculations of Equation (1). The first "true" calculation uses the values of $p_t(\tau)$ in the actual data. The second "counterfactual" calculation assumes that the release

23. By definition, data before the first available year are unreliable, so it is impossible to accurately estimate the release percentages $p_t(\tau)$ for those years.

24. In fact, for the states considered here, the net number of prisoners added by 2002 from the first year of data range from 88% (and often >95%) to a bit >100% of the total BJS prison counts for 2002. (Values >100% clearly reflect counting differences between the BJS and the NCRP or errors in the data.) This demonstrates again the relatively short sentences served by a vast majority of prisoners.

Table 2. Fixed Release Rates Counterfactual Example

	Real Data		Counterfactual Data	
	$t = 1988$	$t = 1989$	$t = 1988$	$t = 1989$
$p_t(0)$	0.45	0.40	0.45	0.45
$p_t(1)$	0.30	0.25	0.30	0.30

Notes: t denotes year of admission, $p_t(0)$ denotes the percent admitted in t who were released in less than a year, $p_t(1)$ denotes the percent admitted in t who were released after serving more than one year but fewer than two.

patterns never changes after the first year. For example, assume that the data start in 1988 and, in the real data, 45% of all prisoners admitted in 1988 are released in 1988 and an additional 30% in 1989. However, for those admitted in 1989, sentencing is tougher, and only 40% released in 1989 and 25% in 1990. Table 2 compares the real and counterfactual versions.

The fixed admissions counterfactual in turn allows the release rates [i.e., $p_t(\tau)$] to always take on their real-data values, but it fixes the size of each entering cohort at the size obtained in the first year of available data. It does, however, allow the distribution of offense types to change with each cohort. In other words, assume that a state's real data start in 1990. In that year, it admits 10,000 people to prison, 45% of whom are violent offenders and 55% of whom are property offenders; in 1991, it admits 15,000 people, 48% of whom are violent offenders and 52% of whom are property offenders. In the fixed admissions counterfactual, the fictitious entering cohort in 1991 numbers only 10,000 (fixed at 1990s level), but 48% of these are violent offenders and 52% property offenders (taking on the actual 1991 distribution).[25]

Figures 3a and b compare the actual and counterfactual outcomes. Two features immediately stand out. First, fixing release rates has almost no effect on the net change in prison population. And second, fixing admission rates leads many prison systems to steady if not shrinking populations.

Start with Figure 3a. If punishments have become more severe over the sample period, the solid "actual" curve (the real data) should rise above the dashed "fixed release rates" curve (the counterfactual data). And while that pattern generally holds, the effects are slight. Actual prison populations are usually

25. The shares of offenders in each class are relatively stable, so little turns on the changes in admissions composition, though the coarseness of the categories may mask important intracategory changes.

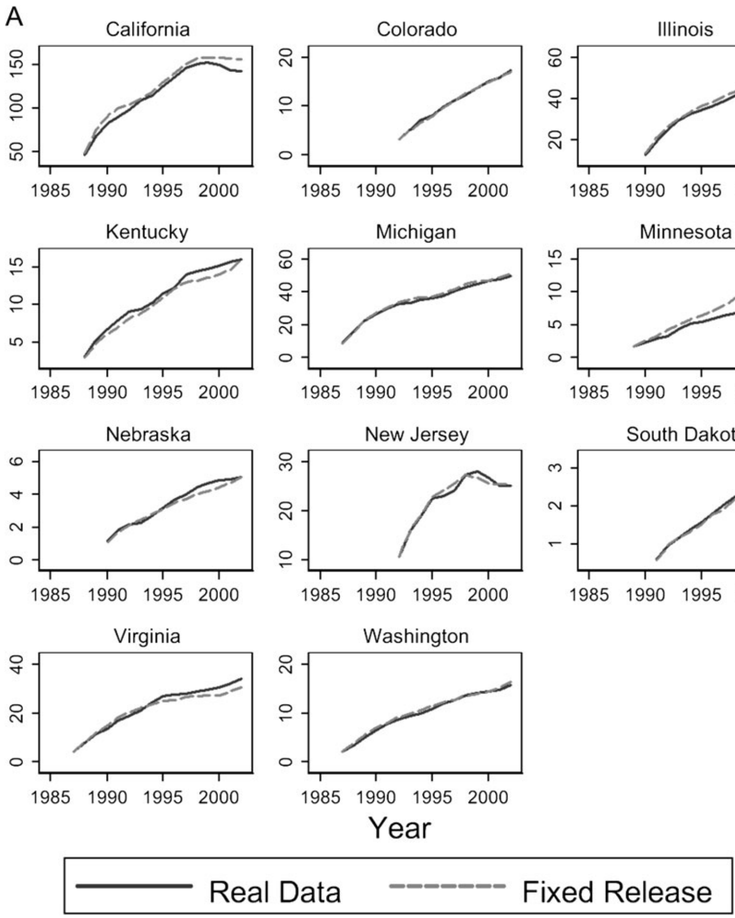

Figure 3. Growth rates: (a) fixed release rates; (b) fixed admission size.

<10% or 11% (and often ~1%) larger than the counterfactual populations. Only Virginia witnesses a dramatic divergence, with the real population growing 24% faster, seemingly tied to Virginia's sentencing guidelines (which were adopted in 1992 and heavily amended in 1995). In a few states, the real populations are smaller than the counterfactual ones. Illinois, for example, added 21% fewer prisoners than had it kept its 1990 policies in place.

Figure 3b presents a fundamentally different picture. In almost all the states under consideration, had admission rates remained constant, prison population would have leveled off quickly. In fact, Figure 3b understates the extent to which growth would have slowed, if not reversed, since it does

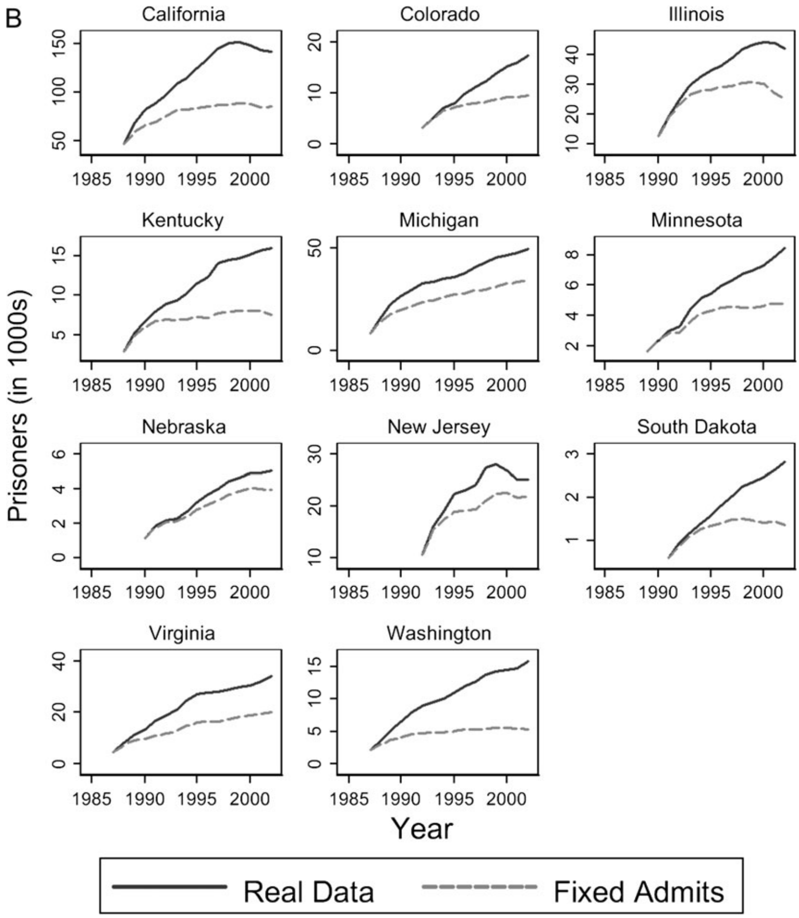

Figure 3. Continued.

not account for the release of prisoners admitted before the first year of data. Thus, release rates in most states are such that prison populations would have quickly reached equilibrium for a given level of admissions. To the extent that prison populations have grown since 1990 or so, the pressure appears to come from the admissions, not release, side of the ledger.

One potential objection to the results in Figure 3 is the different starting times: Perhaps trends in release times or the distribution of entering inmates differed between the late 1980s (when my data for states such as California and Virginia start) and the mid-1990s (when my data for states such as Colorado and New Jersey start). To check the robustness of Figure 3's

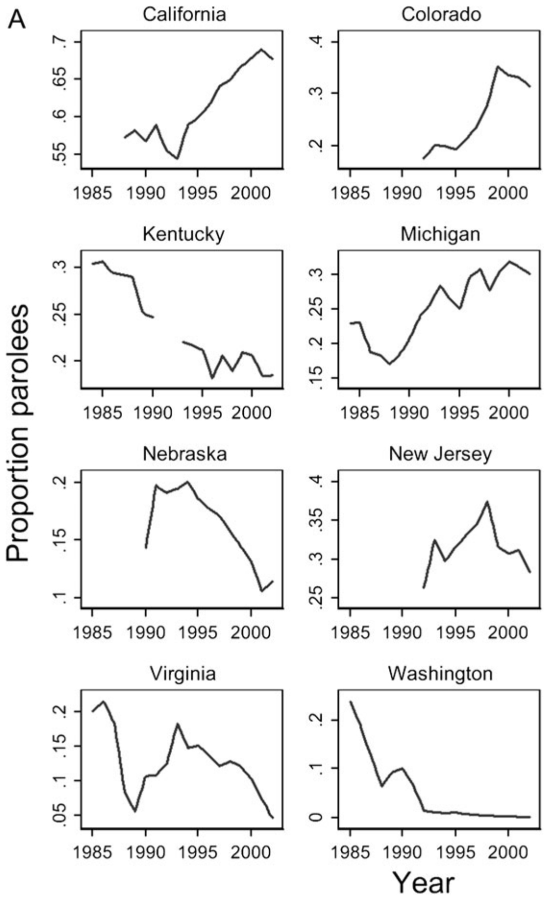

Figure 4. (a) Parolees as percent of admits. (b) Parolee movements: releases and returns.

results, I reran both counterfactuals starting every state in 1992, immediately after the nationwide crime drop of the 1990s began, thus providing a common frame of reference. By and large, the results are nearly identical. The only noticeable differences were that longer sentences played an even lesser role in California, Illinois, Washington, and especially Minnesota[26]; and that

26. In other words, in these four states, the counterfactual fixed release rate prison populations were larger than the real-data prison populations, implying that release policies had become more lenient over time. Such an effect is already visible for Illinois in Figure 3a, albeit at a lesser level than that observed in the 1992-start counterfactuals.

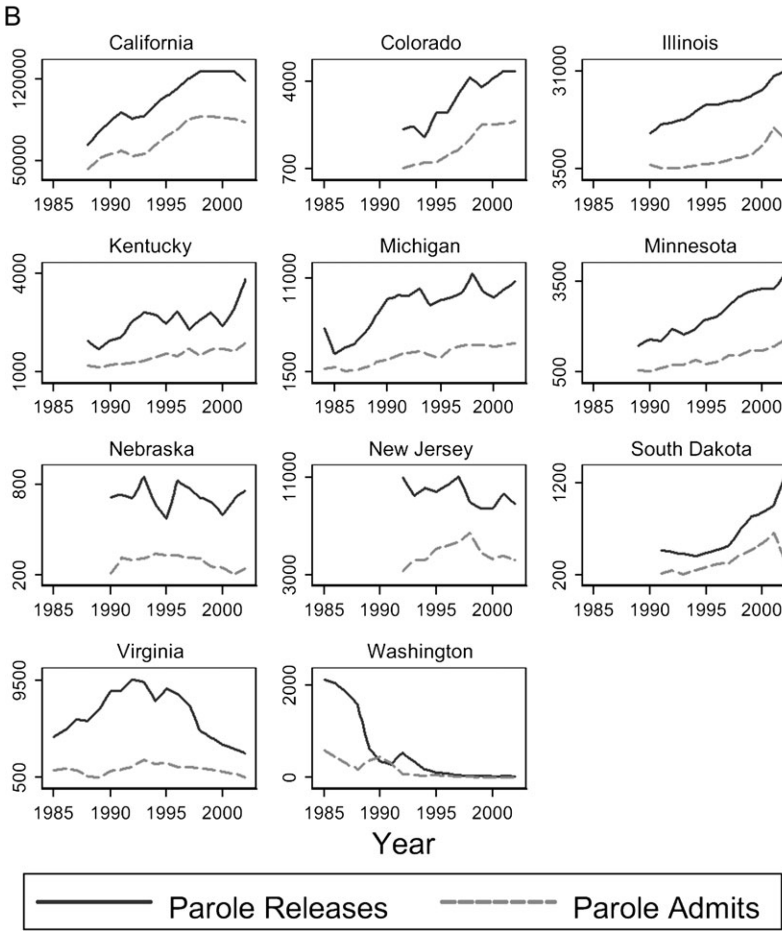

Figure 4. Contined.

growing admission cohorts played a lesser role in Michigan and Virginia, where the real and fixed admissions counterfactual data are effectively indistinguishable.[27] The results in Figure 3 thus appear to tell a robust story

27. In Virginia, the fixed admissions prison population was actually larger than the real population, but just barely so. The fixed release counterfactuals (both in Figure 3a and in the 1992-start version) for both Michigan and Virginia imply that little changed in either state with respect to releases as well. Taken together, these results suggest that neither admission nor release policy in either state has changed much since the early 1990s.

about the source of prison growth since the late 1980s—both before and after the post-1991 reversal in crime trends.

4. The Problem of Violations

Although the counterfactuals indicate that admissions are driving growth, the nature of the admissions shapes our interpretation of this result. If states are violating parolees and probationers back more aggressively than in the past, then the seemingly short times served seen above may be deceptive: If a large number of prisoners return quickly to prison due to (perhaps technical) violations, then their effective sentence is much longer than the individual stints they serve, and so the seeming unimportance of releases may be overstated. Figures 4a and b shed light on this issue.

Figure 4a plots the fraction of each admission cohort entering prison due to either a voluntary or mandatory revocation of parole.[28] It suggests that while increases in parole violations may explain some of the decline in apparent time served, they likely do not explain much. Outside of California, relatively few admittees—rarely >30%—are returning parolees (in California the rate neared 70% by 2002), though this fraction rose in several states over the sample period, usually from ~20% to 30%. In a handful of states, especially Virginia and Washington, the fraction of parolees plummeted; these declines, however, are not due to dramatic improvements in reentry programs but rather to restrictions on, or the abolition of, parole (so parolees simply no longer exist).[29]

Figure 4b poses an even more serious problem for the argument that parole violations drive prison growth. With only a few exceptions, the number of returning parolees is never larger than the number of parolees leaving prison. Moreover, the trends in parolees returning roughly track those in parole releases, and in many states the gap between releases and returns remains constant, or even grows slightly. Figure 4b thus suggests

28. Each state relies almost exclusively on either voluntary revocations or mandatory ones, but never both. It is unclear whether this reflects differences in reporting or in state policy choices.

29. Probation violations appear to play almost no role at all, never rising above ~1% (except for a small, and quickly vanishing, spike to 6% in South Dakota). In many cases, no inmates in a given cohort are entering due to probation violations.

that a parole-driving claim may invert the direction of causality. As the prison population grows, the number of parolees grows with it, and a roughly constant fraction of those paroled violate back to prison. It is not clear that it is good accounting to then claim their return as causing prison growth.[30] The rising number of parolees may at some level be a symptom, not a cause, of prison population growth.

Some caution, however, is needed when considering these results. First, the NCRP only records the primary reason for admission. If an offender violates his parole or probation and, in the process, is liable for a new substantive crime and is admitted for that new crime, he does not show up in the NCRP statistics as a parole readmission. If courts are more willing to incarcerate violators for new offenses than otherwise similarly situated nonparole offenders, then parole violations influence prison growth in a manner invisible to the NCRP. Also, the constancy of parole violations may be surprising during a period when the severity of the marginal offender was declining, assuming lesser offenders violate parole less often. The constant ratios in Figure 4b may thus understate the importance of changing parole practices. Offsetting this effect, however, is the possibility that parolees may be violated back—even on technical grounds—when the triggering act is also a separate crime. Violations are easier than new charges for police and prosecutors to process, so they may use violations as procedural shortcuts for dealing with new offenses. In this case, the simple number of returning parolees overstates the importance of violations.

5. The Determinants of Sentence Length

My results so far can be summarized as saying: "The typical prison sentence is short, and thus long prison sentences do not appear to drive recent prison growth." But it is worth asking "typical *for whom*?" The sections above provide a highly aggregated view of penal outcomes, and these broad patterns may mask important underlying variation in times served by different types of inmates: Black prisoners, for example, may serve systematically

30. For example, if 20% of all parolees are violated back to prison, then an increase in parolees from 1,000 to 2,000 increases violations from 200 to 400, but the net effect of parole is a 800-prisoner reduction in the prison population (from 1,000 – 200 = 800 to 2,000 – 400 = 1,600 prisoners on parole).

longer sentences than otherwise-identical nonblack inmates. This section briefly considers how important such masking may be by examining how likelihoods of release vary based on inmates' races, ethnicities, sexes, ages, or offenses.[31] I find that the already surprisingly short aggregate sentence lengths overstate time served by nonviolent offenders (or, conversely, understate that served by the violent). They also overstate time served by women and the elderly (those over the age of forty years), and understate that served by Hispanics and the young (those under the age of twenty-one years). The results are somewhat more complex for race, though they suggest that whatever effect it may have is weak.

The basic model I use here is a Cox proportional hazard model, which in its most general form is:

$$h_{is}(t) = h_0(t|s)\exp(\beta\text{trait}_i + \Gamma X_i + \Phi Z_{st}). \tag{2}$$

Here, $h_{is}(t)$ is the hazard rate, the probability that offender i in state s is released in year t, conditional on having stayed in prison until t in the first place. $h_0(t|s)$ is the baseline hazard, which may be conditioned on the state in question either through stratification or a shared frailty (random effects) parameter.[32] Trait$_i$ is the particular individual-level characteristic under observation for inmate i, X_i the vector of all other individual-level traits (which are fixed for each inmate), and Z_{st} a vector of (generally) time-varying, state-level controls.

I examine the effect of twelve dichotomous relationships on the relative likelihood of release: whether the defendant is black (compared to either all other inmates or all other non-Hispanic inmates); Hispanic; female; young (under twenty-one years when entering prison), old (over forty years), or middle-aged (between twenty-one and forty years); convicted for a violent crime, a property crime, or a drug crime; and whether the defendant is

31. In an earlier version of this article, I also examined whether there were differences in the trends in the likelihood of release, but no effects appeared, with one exception—Hispanics appear to have faced declining likelihoods of release over the sample period, though the effect size is small. I omit this discussion here.

32. Under stratification, $h_0(t|s) = h_{0s}(t)$; under shared frailty, $h_0(t|s) = h_0(t)\alpha_s$, where α_s is a state-level random effect.

a young black male (compared to either older, black male inmates or young, nonblack males). To reduce the effect of outlier observations and to minimize the risk of model sensitivity, I use a nonparametric matching algorithm to balance my data before applying the Cox model (see, e.g., Abadie, Diamond, and Hainmueller, 2007; Ho et al. 2007).[33] The Appendix discusses in more detail both the intuition behind matching and how I accomplish it; for our purposes here, it should just be noted that matching requires that I use dichotomous treatment variables and that I examine each one in a separate regression.

I run four different versions of Equation (2) for each of the twelve sets of traits (for a total of forty-eight regressions), varying the exact set of control variables and assumptions about how to model state-level heterogeneity. In all cases, standard errors are clustered at the state level. Table 3 lists the basic elements of the four specifications, and Figure 5 provides the results. The "hazard" here is release from prison, so a ratio greater than one implies a greater likelihood of release in a given period—prisoners, in other words, want to experience the hazard. Since these are ratios, the critical question is whether the results differ from one, not zero: A hazard ratio of 1 implies that the two groups in question face the same hazard (i.e., that the particular trait has no apparent effect).

These graphs provide both the point estimate and the upper and lower bounds on the 95% confidence interval.[34] The results are defined such that the hazard ratio refers to the trait given in the title of the figure: In Figure 5.1, for example, a value greater than one implies that black prisoners are more likely to be released in any given period than otherwise identical nonblack prisoners. The numbers under the results correspond to the model numbers in Table 3.

These results suggest that that the aggregate picture provided in the previous sections does mask intergroup variation in some settings. Hispanics, for example, are $\sim$7% less likely to be released in any given period than non-Hispanics. And (not surprisingly) violent offenders are $\sim$20% less

33. Note that matching is not done in place of the Cox model but rather constitutes a preliminary cleaning process. The matching protocol eliminates certain types of problematic outliers, and the matched data are then used conventionally in the Cox regressions.

34. I only provide the forest plots here, rather than tables of results for all forty-eight regressions. The forest plots provide all the relevant information much more concisely, and they make it substantially easier to see broad patterns within and across results. Complete results are available on request.

Table 3. Basic Model Specifications

Model	Individual Controls	State Controls	Heterogeneity Controls
1	Yes	No	No
2	Yes	Yes	No
3	Yes	No	Stratify
4	Yes	No	Shared frailty

Notes: Individual controls are black, sex, age, offense, sentence, and credit for time served. Time-varying state controls are per capita crime rates, percent young black male, percent below the poverty line, citizen and state ideologies, government revenue shortfall, and per capita prison admissions. Time-invariant state controls are whether the state had a TIS law or sentencing guidelines in the inmate's admission year. Stratify and shared frailty are defined at footnote 32.

likely to be released than nonviolent offenders with similar sentences and demographic traits[35]; phrased differently, nonviolent offenders serve even shorter sentences than those suggested above, and violent offenders longer. Conversely, women are $\sim$13% more likely to be released than similarly situated men, and older defendants $\sim$7% more likely.

The effect of race on time served requires a bit more attention. Black and nonblack inmates appear to face roughly the same likelihood of release (see Figure 5.1), implying that the aggregate results above apply equally to both types of inmates. And this result appears to be relatively robust. First, the similarity between black and nonblack outcomes does not seem to be an ephemeral result of blacks making up a disproportionate number of increasingly marginal offenders.[36] I condition on offense and sentence length, for example, and the share of blacks in each entering cohort remains relatively constant (by about $\pm$5%). I am thus comparing similarly situated offenders, and the racial composition of the various "similarly situated" cohorts seems relatively stable.

Second, the equivalent likelihoods are not influenced by the lower release likelihood of Hispanics. Hispanics comprise $\sim$30% of my observations, and almost all Hispanics in my data self-identify as white. Thus, the overall similarity between whites and blacks could arise in part from non-Hispanic

35. The high hazard rates for property and drug offenders likely reflect the low hazard rate of violent offenders: If violent offenders are less likely be released in any period, by definition nonviolent (i.e., property and drug) offenders are more likely.

36. Assume blacks admitted for Crime 1 serve three years, compared to two for whites; and blacks admitted for Crime 2 served four years, compared to three for whites. Average time served by whites and blacks may appear similar if blacks are increasingly admitted for Type 1 crimes.

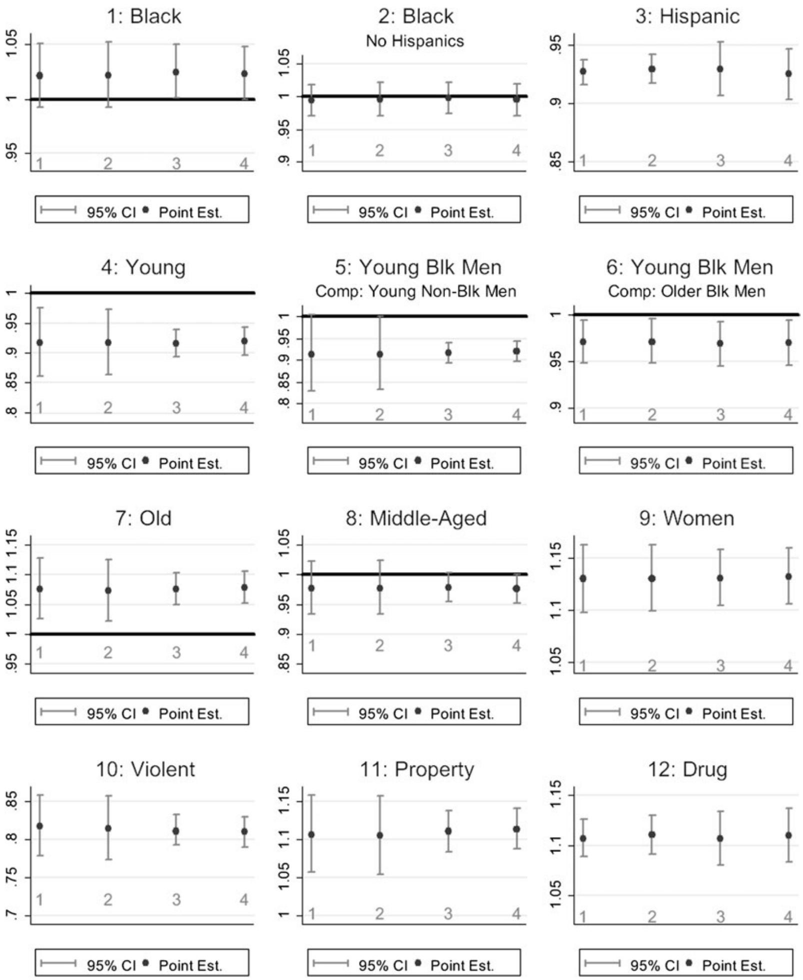

Figure 5. Relative hazard ratios. CI, confidence interval.

whites being released earlier than blacks, and blacks earlier than Hispanic whites. Figure 5.2, however, which eliminates all Hispanics from the sample, suggests that this is not the case.

Race may play a role, however, in shaping the longer sentences served by young inmates. That young males—and young black males in particular—tend to serve longer sentences is well documented (Steffensmeier, Ulmer, and Kramer, 1998; Spohn and Holleran, 2000; Ulmer and Johnson, 2004; Kurlychek and Johnson, 2004; Bushway and Piehl, 2007), and a similar effect appears in my data. To determine whether race or age drives it, I look at two hazard ratios: that between young black males and young nonblack males (Figure 5.5), and that between young and old black males (Figure 5.6). As the two figures indicate, age seems to explain the "young black male" penalty more than race.

Two possible limitations with the results in Figure 5 deserve attention. First, California represents ∼40% of the data, so it may overly influences the results. To examine this possibility, I rerun Models 1 and 2 without California. No differences in results are statistically or numerically significant. California thus does not appear to be driving the result here.[37]

Second, there is a potential source of post-treatment bias lurking in these models. One of the control variables is the imposed sentence. The models thus ask if, for example, a black 25-year-old male arsonist sentenced to ten years with no credit for time served spends as many days in prison as a white 25-year-old male arsonist sentenced to ten years with no credit for time served. This is certainly a valid question. But perhaps we are concerned that the real source of the problem is that a black 25-year-old male arsonist is more likely to get a ten-year sentence than a similarly situated white defendant. Conditioning on sentence length masks this effect. Thus, I rerun Models 1, 2, and 4, as well as those dropping California, without the sentence length term. The hazard for young black males compared to young nonblack males falls somewhat (from ∼0.97 to 0.91), suggesting that race may influence the sentence imposed[38]; otherwise, the results are relatively stable.[39]

6. Conclusion

The dramatic boom in American prison populations over the past thirty years is a striking and oft-mentioned fact, but our understanding of its causes remains surprisingly weak. This article seeks to shed light on the issue by

37. For space considerations, I omit the results here, but they are available upon request.

38. How race plays a role remains unclear. Racial bias is of course a possibility, but note that the NCRP does not provide data on prior records, so if young black men have longer criminal histories that influence sentencing, this effect would be picked up by the "race" variable.

39. Again, I omit the results here, but they are available upon request.

probing the role of changes in time served. Increasingly severe sentencing regimes are often cited as important forces in the growth of U.S. prison populations, but there has been little effort to examine this claim rigorously.

The results in this article indicate that the story is quite complex. In most states in my sample, the median time served has been either relatively flat or declining since the late 1980s or early 1990s. Some of this leveling or decline, however, appears to be due to states locking up increasingly marginal offenders who would likely have received no term of incarceration in earlier years: By serving short sentences, they pull down the median. When examining the effect in as much detail as the NCRP permits, it appears that it is a key issue in California, a lesser issue in states like Illinois and Washington, and not much of one in many of the other states in the sample.

Nonetheless, these results make clear that the primary engine driving prison growth—at least over the past ten to fifteen years—has been changes in admissions, not time served. These two concepts are not entirely distinct since the decision to incarcerate can be thought of at some level as a decision to not release immediately at Time 0. But it is nonetheless useful to separate out whether prison growth is being driven by (1) locking up those who otherwise would have gone to prison for much longer terms, (2) locking up those who otherwise would not have gone to prison for long terms, (3) locking up those who otherwise would not have gone to prison for short terms, or some combination of these three. The results from this article suggest that reason (3) is the key force.

Appendix

This appendix discusses the data and models in more detail. The first part provides a closer look at the NCRP data and the steps taken to clean it. It also explores the external consistency of the NCRP and our ability to extrapolate from the states included in this study to the United States as a whole. The second part fleshes out the data and methods used in Section 5.

The NCRP

Basic Structure of the Data

The BJS began gathering data for the NCRP in 1983, with data available through 2002 when work on this article began. Participation in the program is voluntary, and the number of states contributing data has fluctuated over the years, though nearly forty participate now (up from thirteen in 1983). Participating states fill out an information card for every inmate entering and leaving prison, and they provide offender-level data not only on the exact date of entry into and exit from prison but, among other things, on each inmate's race, sex, age, educational background, conviction offenses, sentences imposed, nature of incarceration (new commitment, parole violation, etc.), subsequent malfeasance while in prison (and its sentencing implications),[40] and type of release (completion of sentence, parole, etc.). It is the most detailed centralized source of information on prison inmates available.[41]

Unfortunately, the NCRP suffers from several noticeable shortcomings. Most immediately troubling, there are significant discrepancies between the NCRP and other sources of data. For many states, the total number of admission or release observations in the NCRP often does not match that given in the BJS's annual National Prisoner Survey. In some cases, the NCRP numbers differ by as much as 75% or 100%. As a result, I use only those states that have a sufficiently long string of consecutive years (at least eleven) that do not deviate too much (generally by <10%) from the corresponding BJS data. Table A1 lists the states I use here and the years that are included. In the end, I have 164 state-year pairs, containing information on over three million inmates, and on over eight million inmate-year observations.

The second significant shortcoming of the NCRP is that the admission and release data sets are wholly separate, with no way to link an observation in an admissions file to that inmate's corresponding observation in

40. Note that because the NCRP operates only at entry and exit, it provides no information on misbehavior by those admitted in a given year but not yet released. This variable is thus quite hard to use since (unobserved) bad behavior may be responsible for some inmates not yet being released.

41. Some entries theoretically available prove not to be. For example, the NCRP includes a binary variable indicating whether the inmate has any prior felony convictions, but it is blank for every observation in every state in every year of data. Other variables, such as education, are not always left blank but are reported with enough infrequency to make them hard to use. Despite these limitations, the NCRP remains the most comprehensive data set of its kind.

Table A1. Years Available for Reporting States

State	Start Year	End Year
California	1988	2002
Colorado	1992	2002
Illinois	1990	2002
Kentucky	1984	2002
Michigan	1984	2002
Minnesota	1989	2002
Nebraska	1990	2002
New Jersey	1992	2002
South Dakota	1991	2002
Virginia	1985	2002
Washington	1984	2002

a release file. This does not restrict my ability to calculate time served by those released since the release file contains both admission and release dates for each inmate. But to calculate either the time served quantiles or the hazard ratio, I need to know both the time served by each released prisoner and how many members of each entering cohort remain in prison at the end of the sample period. The separate admission and release files, combined with numerous reporting errors, complicate this latter calculation.

Ultimately, I must create ersatz observations for the right-censored unreleased prisoners. To do this, I use the admissions files to determine how many inmates of a certain type were admitted in a particular year, I use all the subsequent release files to calculate how many of those prisoners were eventually released, and then I create dummy entries to make up the difference. The ideal approach is the following: Assume that in 1990, California admits 100 twenty-five-year-old black male inmates for second-degree assault with sentences of ten years and no credit for prior time served. The release files for 1990 through 2002 indicate that ninety-five inmates meeting these criteria—admitted in 1990, twenty-five years old, black, and so on—have been released. I then create five dummy inmates with these traits, giving me the complete set: for ninety-five inmates I know the exact day of entry and release, and for the five dummy observations I create fake admission dates (uniformly distributed across the year, in keeping with the real admission data) and know that they have not been released as of December 31, 2002. With these dummy inmates, I effectively create an integrated admissions and release file, allowing me to easily compute quantiles and hazards.

In theory, I could identify the exact admittees not yet released—and thus not have to create ersatz alternatives—by conditioning on every characteristic provided in the NCRP. Reporting errors in the NCRP, however, render this approach impossible, and ultimately I condition on only six variables: race, sex, age, most serious offense, sentence imposed, and credit for time served.[42] Moreover, I have to adjust even these categories to account for reporting flaws in the NCRP.

The key problem is that it often appears that $>100\%$ of inmates of a particular type in a specific admission cohort are subsequently released: A state may admit 100 inmates of a particular type in 1990 but appear to release 120 such prisoners from the 1990 cohort over the next thirteen years. In numerous cases, prisoners who do not exist in the admissions file appear in the releases file—for example, no prisoner is admitted in California in 1988 for crime code 13—yet several prisoners released in subsequent years from the 1988 entry cohort are listed as having committed code 13 crimes. These errors appear to be coding errors: These are not inmates who were never recorded as entering but are subsequently being released, but rather they are inmates classified one way at the time of admission and labeled differently at release.[43] Not surprisingly, these errors grow the more disaggregated the data become, so to minimize this problem I cluster many of the variables. I narrow the NCRP's 186 offense codes down to sixteen broad categories,[44] its continuous measure of age to six categories,[45] its continuous

42. In some cases, I condition on a seventh, ethnicity. However, due to reporting errors in the NCRP, I must drop Kentucky and Michigan from my sample when doing so (a vast majority of observations in those two states are missing data on ethnicity), so I condition on ethnicity only where necessary.

43. The evidence for this claim is indirect: There is no apparent "overrelease" problem at the unconditional level. In other words, the total number of prisoners admitted in a particular year is never less than the number released from that cohort over all future years, suggesting that problem is misclassification, not the release of unrecorded admissions.

44. The sixteen are as follows: murder and other killing offenses (including assault with intent to kill), kidnapping, sex offenses, robbery, assault (including hit-and-run and child abuse), burglary, arson, theft and associated offenses (including trafficking, distributing, or receiving stolen goods), drug trafficking, drug possession, persistent felony violators, unknown offenses, and four "other" categories (other violent, other property, other drugs, and other [lesser] crimes that do not fit easily into the violent/property/drug taxonomy).

45. The six are under sixteen, sixteen to under eighteen, eighteen to under twenty-one, twenty-one to under thirty-one, thirty-one to under forty-one, and forty-one and over.

measure of sentence imposed to fifteen categories,[46] and its continuous measure of credit for time served to five categories.[47] I keep the NCRP's binary classification of sex but reduce its 5-fold classification of race to a binary black/not-black measure. As a result, I have potentially 28,800 ($16 \times 6 \times 15 \times 5 \times 2 \times 2$) different "types" of offenders (and dummy offenders) per state per year.[48]

These coarser categories curtail the overcount problem, but they do not eliminate it completely. A simple example demonstrates more concretely the problem that remains and how I correct it. Assume that in 1990, a particular state admits 900 inmates, and between 1990 and 2002 it releases 850 of them, so 50 from the 1990 cohort remain in prison at the end of 2002. My methodology above would add fifty ersatz offenders to the release files to create a full set of prisoners. Now assume I condition on some trait, resulting in two groups of inmates, A and B. According to the admissions file, the state admits 500 of Type A and 400 of Type B in 1990. According to the release files, however, between 1990 and 2002 the state releases 520 of Type A and 330 of Type B; assume the unseen true values are 490 and 360, respectively.[49] The total remaining in prison is the same: The state has released 850 (520 + 330) of the 900 (500 + 400) admitted prisoners. But things get tricky when I try to create my ersatz observations. My data claim that I should create –20 dummies of Type A and 70 dummies of Type B. In theory, I would like to correct this by turning thirty erroneous ersatz Bs into ersatz As, but I do not actually know that the number is thirty. I know that at least twenty inmates are misclassified (since I have twenty "impossible" releases for Type A), but I have no more information than that. I cannot create "negative" dummies, and simply ignoring Type

46. The fifteen are less than a year; one year ranges from one to two years through nine to ten years; ten to fifteen years, fifteen to twenty years; twenty to twenty-five years; over twenty-five years but not life; and all life and death sentences. In general, at least 80% of all sentences for all categories of crimes are below ten years, and 90% are below twenty.

47. The five categories are no credit, less than one year, one year to three years, three years to five years, and more than five years.

48. For the models looking at the effect of being Hispanic, the total rises to 57,600 since Hispanic/non-Hispanic is now an option as well.

49. I am assuming here that the errors are on the release end, not the admission. I have no evidence either way, and there appears to be no *ex ante* bias to choosing one assumption about the location of the errors over the other.

As (i.e., creating zero dummies of this type) while creating seventy ersatz Type Bs overstates the size of the prison population.

My solution, then, is the following. I constrain the number of "overage" inmates released to the number admitted: The number of Type As released is restricted to 500. I then assume that all of the "excess" releases—the twenty extra Type A releases—should have been Type Bs, so I reduce the number of ersatz Type Bs by twenty. I thus end up with zero Type A dummies (since 500 out of 500 are assumed released) and fifty Type B dummies (since 350—the 330 in the data and the 20 extra Type As—of 400 are assumed released). These are, of course, not the true values (in my example, I produce ten too few ersatz Type As and ten too many ersatz Type Bs), but it is the best I can do with the data available, and it is likely a close enough approximation. In practice, of course, I have more than two categories—in fact, I can have as many as 28,800 or 57,600. I thus take a more aggregate approach. I total the number of overages within a state/year and then randomly delete that many dummies from all the non-overage categories.[50] With this correction, I have a complete data set that approximates the real data.

External Validity

My sample here consists of eleven states, which house ~30% of the nation's inmates. It is thus essential to understand how generalizable the findings here are to nonreporting states. Since the NCRP is a voluntary program, there is no reason to assume that the reporting states are merely a random sample of the entire nation, particularly since some states have opted out after opting in.

Figure A1 plots quantile-quantile comparisons of key demographic, economic, and political traits of the included states against those of the excluded states. As always, the more the points fall along

50. Consider an example when conditioning on just two traits, the conviction offense and sentence imposed. In California in 1990, 100 of 195 offense-sentence pairs yield excess releases, ranging from 1 to 126, for a total of 1,702 "overages." The remaining 95 pairs have remainders ranging from 0 to 774, for a total of 1,987 ersatz observations. I thus randomly delete 1,702 of these 1,987 dummies, leaving me with 285 ersatz observations across the 95 non-overage pairs. Looking at the unconditional data, by the end of 2002 only 285 inmates remain from the 1990 cohort, exactly the same as the number of ersatz observations I created.

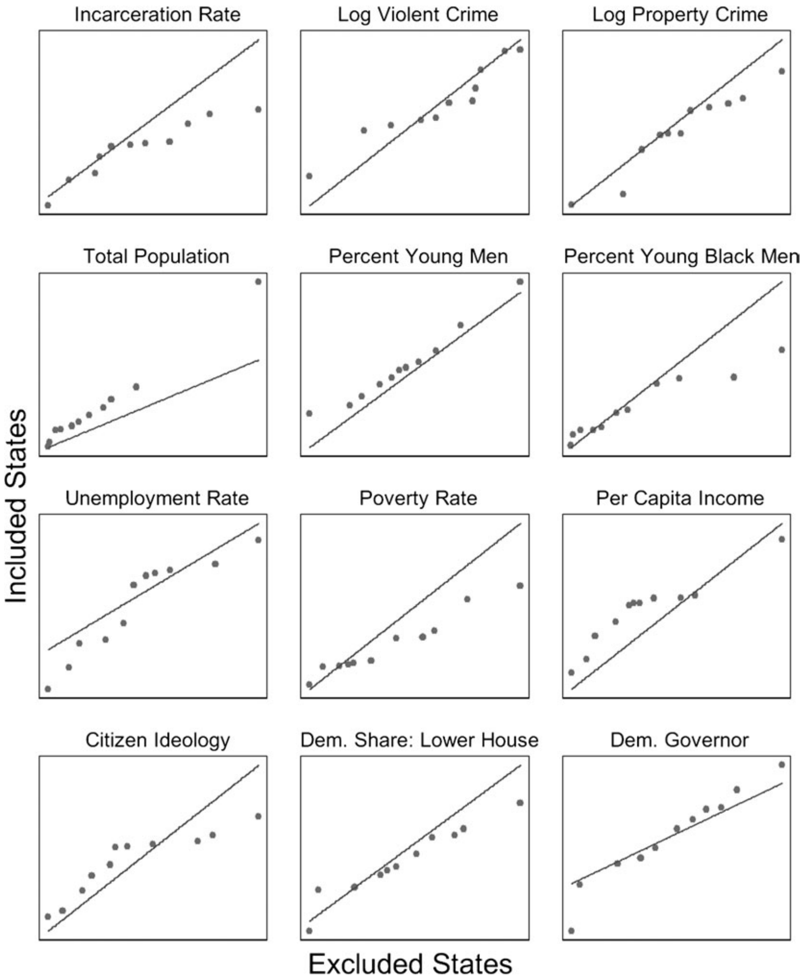

Figure A1. Quantile-quantile plots.

the forty-five degree line, the more similar the two samples are. The results are mixed. In general, the included states have lower property crime rates, larger populations, more young men, lower unemployment and poverty rates, higher per capita incomes, and more conservative citizens and lower houses. Yet there are several categories where the two sets of states are fairly similar, including violent crime, percent young black (for most of the observations), and Democratic governors. And for many categories where there are systematic differences, they do not appear to be large. Even incarceration rates, which seem substantially lower in the included states, are actually fairly similar but for one outlier.

Figure A1 thus indicates not only that some care should be taken when extrapolating to the rest of the country but also that the included states do not differ radically from other states. Cautious extrapolation may thus not be out of place.

Empirical Models

Section 5 sets forth the basic Cox regression used here, but this section covers three issues in more detail: the matching protocol, the sources of the data, and the summary statistics.

The Basics of Matching

The intuition behind non-parametric matching is straightforward. The goal is to convert observational data into something that looks more experimental, by ensuring that the "treated" and "control" groups look as similar as possible. When "trait" in Equation (2) is black/not-black, for example, the matching algorithm takes each black inmate (where black is viewed as the treatment) and looks for a nonblack inmate (the control) admitted in the same state and year with identical values for sex, age, offense, sentence, and time credit. If no such nonblack inmate can be found, the black inmate is dropped.[51] As Ho et al. (2007) demonstrate, using matched data in a parametric or semiparametric model (such as the Cox model) leads to substantially more stable, less model-dependent results than when using unmatched data.[52]

51. For example, consider an inmate admitted in 1996 in California who is a black male with an age score of 4, an offense code of 12, a sentence code of 11, and a credit score of 0. In the model considering race, the matching protocol looks for a white male with identical (4,12,11,0) traits. If none is found, the black observation is dropped. Thus, a (black, X) inmate is in the matched data only if the protocol can find a (white, X) inmate with identical X data. Once matched, the data are handled by the Cox model like any other data—matching is not used in place of a parametric method. Because of the discrete nature of my data, I keep only those observations with exact matches.

52. This matching approach is not equivalent to using a propensity score. As Ho et al. (2007) make clear, the nonparametric nature of their matching technique confronts directly the core limitations of propensity scores.

I must do more than simply match my data, however. The NCRP is an extensive data set: The core data set for the eleven included states contains over three million offender observations and over eight million offender-year observations.[53] To make estimation computationally feasible, I need to take a sample of my data. Ideally, I want take a 1% sample of the data and thus have ~30,000 offender observations per regression. However, pure random sampling after matching undoes the matching since random deletion may eliminate only one half of a matched pair. But sampling before matching is not feasible since it often results in data sets with substantially fewer than 30,000 observations: A match that may have been possible with the full data set becomes impossible if one of the matching observations is deleted before matching. To solve this problem, I perform the match on the full data set, and I then sample by paired observations (either keeping both pairs in a match or eliminating both), setting the sample weight such that each sampled data set has ~30,000 observations.[54] The summary statistics for the matched data sets are provided in Table A2.

Sources of Data

All offender-level data come from the NCRP between 1985 and 2002. Population data are from the U.S. Census Bureau's Population Estimates, available online at http://www.census.gov/popest/archives/, and poverty data are from the Census Bureau's Historical Poverty Tables, available online at http://www.census.gov/hhes/www/poverty/histpov/perindex.html. Crime data are from Uniform Crime Reports, and prison admission data are from the BJS. Citizen and state ideologies

53. In other words, an offender admitted on August 1, 1988, and released on June 3, 1990, provides three offender-year observations (August 1, 1988, to December 31, 1988; January 1, 1989, to December 31, 1989; and January 1, 1990, to June 3, 1990). This is done to accommodate year-varying state controls.

54. Since some traits are easier to match on than others, the matched data sets for various traits were often of different sizes. In some cases, the matched data included an asymmetry: When matching on black, for example, there may have been fifty blacks inmates of type *X* and fifty-five white inmates of type *X* in the matched data. In these cases, I would randomly delete five of the white inmates so I had fifty pairs to draw from; the randomly deleted white inmates were indistinguishable with respect to time served from those retained.

Table A2. Summary Statistics

	Black	Sex	Under 21	Over 40
Original observations	29,626	30,130	29,810	30,486
Total observations	63,827	59,021	80,804	62,581
Black	0.5 (0.500)	0.431 (0.495)	0.445 (0.497)	0.407 (0.491)
Sex	0.091 (0.288)	0.5 (0.500)	0.030 (0.170)	0.094 (0.291)
Age[a]	4.649 (0.848)	4.755 (0.767)	0.5[b] (0.500)	0.5[b] (0.500)
Offense[a]	26.657 (12.672)	28.824 (11.851)	24.611 (12.473)	27.886 (12.933)
Sentence[a]	4.139 (2.976)	3.628 (2.700)	4.952 (3.468)	4.158 (3.131)
Credit[a]	2.256 (1.298)	2.020 (1.267)	1.567 (1.102)	2.317 (1.331)
% Young black men	0.016 (0.006)	0.016 (0.007)	0.018 (0.007)	0.051 (0.006)
Poverty rate	13.522 (2.785)	13.363 (2.814)	12.655 (2.778)	13.401 (2.672)
Citizen ideology	52.538 (6.561)	51.876 (7.537)	51.861 (7.929)	52.410 (6.872)
Government ideology	51.411 (22.325)	51.379 (23.274)	45.462 (21.553)	55.027 (25.239)
Per capita crime rate	0.049 (0.012)	0.048 (0.011)	0.0481 (0.011)	0.046 (0.011)
Prison admit rate	0.003 (0.001)	0.003 (0.001)	0.002 (0.001)	0.003 (0.001)
Per capita state revenue shortfall	0.337 (0.693)	0.341 (0.742)	0.313 (0.655)	0.282 (0.822)
TIS law	0.602 (0.489)	0.650 (0.477)	0.554 (0.497)	0.723 (0.448)
Guidelines	0.887 (0.316)	0.880 (0.325)	0.797 (0.402)	0.902 (0.298)

	Between 21 and 40	Violent	Property	Drugs
Original observations	30,268	29,468	29,478	29,594
Total observations	68,170	71,570	61,750	61,754
Black	0.416 (0.493)	0.397 (0.489)	0.363 (0.481)	0.467 (0.499)
Sex	0.069 (0.253)	0.046 (0.210)	0.098 (0.297)	0.102 (0.303)
Age[a]	0.5[b] (0.500)	4.505 (0.861)	4.630 (0.840)	4.666 (0.837)
Offense[a]	26.824 (12.831)	0.5[b] (0.500)	0.5[b] (0.500)	0.5[b] (0.500)
Sentence[a]	4.402 (3.238)	4.893 (3.223)	3.892 (2.765)	3.790 (2.501)
Credit[a]	2.068 (1.312)	2.226 (1.230)	2.197 (1.298)	2.154 (1.299)
% Young black men	0.016 (0.007)	0.016 (0.006)	0.016 (0.006)	0.016 (0.006)

Poverty rate	13.170 (2.731)	13.474 (2.783)	13.557 (2.753)	13.521 (2.792)
Citizen ideology	52.249 (7.230)	52.348 (6.666)	52.012 (6.987)	52.482 (6.734)
Government ideology	51.894 (24.447)	52.121 (23.232)	50.849 (22.710)	52.197 (22.965)
Per capita crime rate	0.047 (0.011)	0.049 (0.011)	0.049 (0.011)	0.049 (0.011)
Prison admit rate	0.003 (0.001)	0.003 (0.001)	0.003 (0.001)	0.003 (0.001)
Per capita state revenue shortfall	0.293 (0.772)	0.314 (0.723)	0.329 (0.699)	0.342 (0.726)
TIS law	0.667 (0.471)	0.617 (0.486)	0.598 (0.490)	0.647 (0.478)
Guidelines	0.869 (0.338)	0.876 (0.330)	0.871 (0.335)	0.913 (0.281)
	Young Black Male versus Other Young Male	Young Black Male versus Non–Young Black Male	Hispanic	Black (Hispanic omitted)
Original observations	29,992	29,984	30,366	29,526
Total observations	84,903	84,274	61,657	59,682
Black	0.5 (0.500)	N/A	0.005 (0.073)	0.5 (0.500)
Hispanic	N/A	N/A	0.5 (0.500)	N/A
Sex	N/A	N/A	0.064 (0.245)	0.101 (0.301)
Age[a]	N/A	3.740 (0.992)	4.577 (0.808)	4.713 (0.833)
Offense[a]	23.403 (12.366)	26.024 (12.086)	27.318 (13.436)	27.073 (14.465)
Sentence[a]	4.911 (3.364)	5.349 (3.558)	3.405 (2.113)	3.706 (2.575)
Credit[a]	1.555 (1.084)	1.651 (1.148)	2.212 (1.304)	2.257 (1.287)
% Young black men	0.018 (0.007)	0.021 (0.007)	0.014 (0.004)	0.016 (0.007)
Poverty rate	12.773 (2.796)	12.100 (2.588)	14.234 (2.562)	13.633 (2.823)
Citizen ideology	52.323 (7.393)	52.503 (7.842)	53.381 (3.682)	52.601 (5.093)
Government ideology	45.322 (20.852)	40.964 (19.366)	57.216 (22.340)	53.353 (21.831)
Per capita Crime rate	0.049 (0.011)	0.048 (0.010)	0.051 (0.012)	0.049 (0.011)
Prison admit rate	0.002 (0.001)	0.022 (0.008)	0.003 (0.001)	0.003 (0.001)
Per capita state revenue shortfall	0.311 (0.630)	0.301 (0.614)	0.306 (0.755)	0.350 (0.724)
TIS law	0.547 (0.498)	0.543 (0.498)	0.650 (0.477)	0.670 (0.470)
Guidelines	0.816 (0.387)	0.771 (0.420)	0.996 (0.064)	0.952 (0.214)

Note: values in parentheses are the standard errors.

[a]Unless otherwise indicated, values reflect the categorical values described in footnotes 44–47 (i.e., the value of age can be 1, 2, 3, 4, 5, or 6, depending on which range the offender's age falls within).

are from Berry et al. (1998). Government revenue shortfall comes from the Annual Survey of State and Local Government Finances, and it is computed as the difference between a state's total revenue and total expenditure in a given year. TIS laws and guidelines come from the sources given in Pfaff (2006).

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2007. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," Working paper, NBER.

Berry, William D., Evan J. Ringquist, Richard C. Fording, and Russell L. Hanson. 1998. "Measuring Citizen and Government Ideology in the American States, 1960-1993," 42 *American Journal of Political Science* 327–48.

Blumstein, Alfred, and Allen J. Beck. 1999. "Population Growth in the US Prisons, 1980–1996" in Michael Tonry and Joan Petersilia, eds., *Prisons*. Chicago: University of Chicago Press.

Bushway, Shawn D., and Anne Morrison Piehl. 2007. "The Inextricable Link between Age and Criminal History in Sentencing," 53 *Crime and Delinquency* 156–83.

Dhammika, Dharmapala, Nuno Garoupa, and Joanna Shepherd. 2006. "Sentencing Guidelines, TIS Legislation, and Bargaining Power," working paper.

Glaze, Lauren E., and Thomas P. Bonczar. 2007. Probation and Parole in the United States, 2006. Washington: Bureau of Justice Statistics.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference," 15 *Political Analysis* 199–236.

Jacobs, David, and Ronald E. Helms. 1996. "Toward a Political Model of Incarceration: A Time-Series Examination of Multiple Explanations for Prison Admission Rates," 102 *American Journal of Sociology* 323–57.

The Justice Kennedy Commission. 2004. *Report to the ABA House of Delegates*. Washington: American Bar Association.

Kurlychek, Megan C., and Brian D. Johnson. 2004. "The Juvenile Penalty: A Comparison of Juvenile and Young Adult Sentencing Outcomes in a Criminal Court," 42 *Criminology* 485–517.

Listokin, Yair. 2003. "Does More Crime Mean More Prisoners? An Instrumental Variables Approach," 46 *Journal of Law and Economics* 181–206.

Marvell, Thomas B. 1995. "Sentencing Guidelines and Prison Population Growth," 85 *Journal of Criminal Law and Criminology* 696–709.

Marvell, Thomas B., and Carlisle E. Moody. 1996. "Abolishing Parole: The Long-Term Impacts on Prisons and Crime," 34 *Criminology* 107–28.

———. 1997. “Age-Structure Trends and Prison Populations,” 25 *Journal of Criminal Justice* 115–24.

Nicholson-Crotty, Sean. 2004. “The Impact of Sentencing Guidelines on State-Level Sanctions: An Analysis Over Time,” 50 *Crime and Delinquency* 395–411.

Pew Center on the States. 2008. *One in 100: Behind Bars in America 2008*. Washington: Pew Charitable Trusts.

Pfaff, John F. 2006. “The Continued Vitality of Structured Sentencing following *Blakely*: The Effectiveness of Voluntary Guidelines,” 19 *UCLA Law Review* 202–307.

———. 2008. “The Empirics of Prison Growth: A Critical Review and Path Forward,” 98 *Journal of Criminal Law and Criminology* 101–79.

———. 2010. “The Durability of Prison Populations,” 2010 *University of Chicago Legal Forum* 73–115.

Sorensen, Jon, and Don Stemen. 2002. “The Effect of State Sentencing Policies on Incarceration Rates,” 48 *Crime and Delinquency* 456–75.

Spohn, Cassia, and David Holleran. 2000. “The Imprisonment Penalty Paid by Young, Unemployed Black and Hispanic Male Offenders,” 38 *Criminology* 281–306.

Steffensmeier, Darrell, Jeffrey T. Ulmer, and John Kramer. 1998. “The Interaction of Race, Gender, and Age in Criminal Sentencing: The Punishment Cost of Being Young, Black, and Male,” 36 *Criminology* 763–97.

Stucky, Thomas D., Karen Heimer, and Joseph B Lang. 2005. “Partisan Politics, Electoral Competition and Imprisonment: An Analysis of States Over Time,” 43 *Criminology* 211–48.

Ulmer, Jeffrey T., and Brian Johnson. 2004. “Sentencing in Context: A Multilevel Analysis,” 42 *Criminology* 137–77.

Walmsley, Roy. 2007. World Prison Population List. 7th ed. London: King’s College London International Centre for Prison Studies.

Zimring, Franklin E. 2001. “Imprisonment Rates and the New Politics of Criminal Punishment,” 3 *Punishment and Society* 161–66.