2010

The Durability of Prison Populations

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

Recommended Citation
John Pfaff, *The Durability of Prison Populations*, 2010 U. Chi. Legal Forum 73 (2010)
Available at: http://ir.lawnet.fordham.edu/faculty_scholarship/720

The Durability of Prison Populations

John F. Pfaff †

Over the past thirty years the US prison population has exploded in a way unparalleled in American history or world experience. Since the mid-1970s it has quintupled in size, from just over 300 thousand inmates to more than 1.5 million. One out of every one hundred adults—and nearly one out of every twenty black males—is behind bars. The United States is home to under 5 percent of the world's population, but it warehouses approximately one out of every *three* of its prisoners.[1]

Such a sprawling penal system is expensive. By 2004, states were spending a total of nearly $40 billion per year to maintain their populations. Correctional spending generally makes up only about 2 percent of overall state budgets, but 10 to 20 percent of state discretionary spending. Even before the current recession, state legislators were beginning to look for ways to trim back prison populations and expenditures.[2] As state budgets have been eviscerated over the past year and a half, the need for reform grows all the stronger.

If policymakers are serious about reining in prison populations (and thus spending), it is essential that they understand

---

† Associate Professor, Fordham Law School. Thanks to the participants at the University of Chicago Legal Forum's conference on Crime, Criminal Law, and the Recession for very helpful comments. Sean Koehler provided outstanding research assistance. All errors are my own.

[1] See Pew Center on the States, *One in 100: Behind Bars in America 2008* 5, 35 (2008), online at http://www.pewcenteronthestates.org/uploadedFiles/8015PCTS_Prison08_FINAL_2-1-1_FORWEB.pdf (visited Oct 3, 2010); Roy Walmsley, *World Prison Population List*, 3 (King's College London International Center for Prison Studies 2007), online at http://www.kcl.ac.uk/depsta/law/research/icps/downloads/world-prison-pop-seventh.pdf (visited Oct 3, 2010). The one-in-three figure includes both prison and jail populations, but the United States' share of the world's prison populations (as opposed to its prison and jail populations) is likely similar.

[2] See, for example, Robin Campbell, *Dollars and Sentences: Legislators' Views on Prisons, Punishment, and the Budget Crisis* (Vera Institute of Justice 2003), online at http://www.vera.org/download?file=105/Dollars%2Band%2Bsentences.pdf (visited Oct 3, 2010); Daniel F. Wilhelm and Nicholas R. Turner, *Issues in Brief: Is the Budget Crisis Changing the Way We Look at Sentencing and Incarceration?* (Vera Institute of Justice 2002), online at http://www.vera.org/download?file=269/IIB%2BBudget%2Bcrisis.pdf (visited Oct 3, 2010).

what has powered growth over the past three decades. In another paper, I develop evidence suggesting that prison growth has not been driven by longer sentences but rather by increased admissions.[3] Specifically, I demonstrate that the typical prisoner does not spend much time in prison—the median inmate often serves under two years (and sometimes as little as six months), and 75 percent of all inmates are released within two to five years. A clear policy recommendation seems to flow from these results: state governments can quickly shrink prison populations simply by reducing new admissions, thus avoiding maneuvers with more potential political risk, such as granting early releases or reducing the official sentences for many crimes.

But the truth is a bit more complicated. Median and 75th percentile times to release can remain flat—or even fall—during periods in which prison population growth is being driven almost exclusively by a small number of inmates receiving increasingly longer sentences. The larger this cluster of long-serving inmates, the more slowly changes to admissions alone can shape future prison populations. This paper seeks to explore more carefully how important these long-serving inmates are to prison populations—and thus to provide a better understanding of the options available to policymakers who wish to meaningfully cut overall prison populations.

I have two goals here. The first is to assess the short-run importance of long-serving inmates to prison reform, which I do by measuring the extent to which these "durable" offenders limit the efficacy of short-term "admission-side" reforms.[4] To do this, I develop several counterfactual experiments using detailed inmate-level data from eleven states. In particular, I examine how much smaller prison populations would have been in 2001 had states enacted one of six reforms: reducing the number of admissions per year by 25 percent or 100 percent (that is, admitting no one) starting in either 1996 or 1999, or reducing the time served by all newly admitted prisoners by 25 percent starting in either 1996 or 1999.

My results indicate that such reforms would be relatively effective despite the presence of some long-serving inmates.

[3] John F. Pfaff, *The Myths and Realities of Correctional Severity: Evidence from the National Corrections Reporting Program on Sentencing Practices*, Am L & Econ Rev (forthcoming 2011).

[4] I use "admission-side reforms" to refer to changes in either the number of admissions or, less frequently, the time actually served by newly admitted inmates. This approach contrasts with "release-side reforms" such as early releases or furloughs, which change the time to be served by those admitted to prison before the reforms are adopted.

Reducing admissions by 25 percent leads to populations that are 15 to 20 percent smaller than they otherwise would have been within two years and 19 to 23 percent smaller than otherwise within five years. The 100 percent cut leads to population between 53 and 79 percent smaller than otherwise within two years and between 75 and 93 percent smaller than otherwise within five years. Reducing time served is also effective, though less so than admitting fewer inmates: reducing time served by 25 percent results in populations that are 7 to 16 percent smaller than otherwise within two years and 12 to 21 percent smaller than otherwise within five years.

The second goal of this paper is to measure the limits to admission-side reforms that are exposed by these counterfactuals. Note, for example, that a state can admit no prisoners for five years yet see its prison population drop to only 25 percent of what it otherwise would have been.[5] Durable inmates still exist, and they effectively create a "floor" below which admission-side reforms cannot push prison populations. To measure this floor, I extrapolate release trends into the future—always a risky endeavor, the results of which should be evaluated with some caution—to estimate how the floor decays over time. Based on various assumptions about trends in releases, my results suggest that admissions in eleven states from the late 1980s through 2002 created a cluster of inmates that will total between 25,500 and 56,000 in 2018—or between 84,000 and 184,000 nationwide, if my results can be extrapolated to the rest of the country.[6] The upper bound is half the total prison population of the late 1970s, but the less-extreme estimates imply that these long-serving inmates need not exert that much influence down the line.

This paper is organized as follows. Part I provides the motivation for this paper by detailing the current budgetary pressures forcing states to rethink their long reliance on incarceration. It also summarizes the short-sentence findings of my earlier work and explains in more detail the long-serving inmate problem. Part II then develops the models establishing the short-run responsiveness of state prison populations to changes in the size of admission cohorts and in the time actually served by those

[5] In other words, if a state's actual admission and release policies would have led to a prison population of 100,000 in 2001, the decision to admit no prisoners between 1996 and 2001 would have resulted in a prison population of 25,000 as of 2001. This is a number substantially less than 100,000, but also substantially greater than zero.

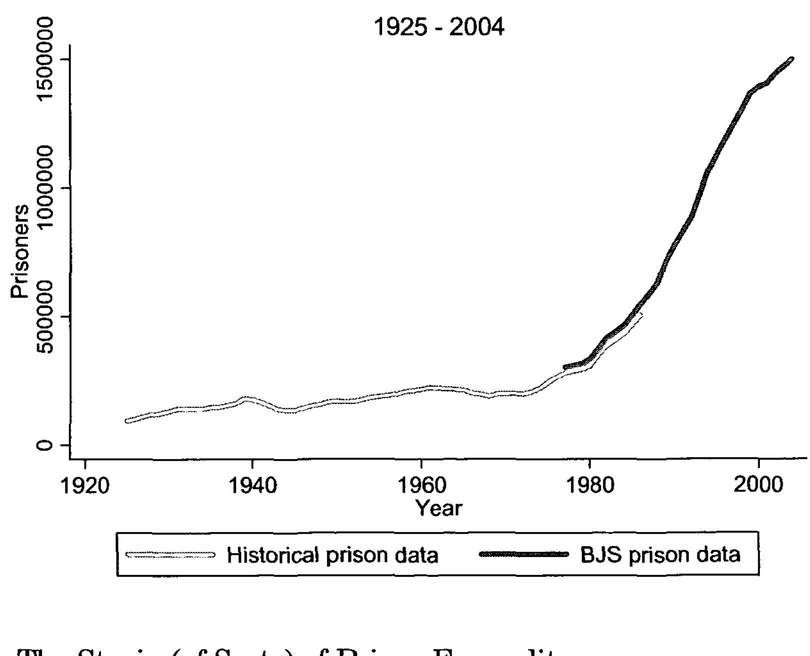
[6] Note that this floor is not the entire prison population, since some of those admitted between 2002 (when my data end) and 2018 will still be in prison in 2018 as well.

inmates. And Part III measures the floor that long-serving inmates admitted during the 1980s and 1990s impose on future prison populations. It also estimates how much these long-serving inmates contribute to the overall cost a given entering cohort imposes on state budgets. Finally, the Appendix provides a detailed discussion of the data and methods used in this paper.

I. PRISON BUDGETS, THE TYPICAL PRISONER, AND THE LONG-SERVING INMATE PROBLEM

Figure 1 shows the dramatic and unprecedented growth in US prison population over the past three decades. For fifty years the incarceration rate hovered around 100 per 100,000 only to soar to over 750 per 100,000 today. In this section I touch on three aspects of this meteoric rise: its cost, the fact that most prisoners serve relatively short sentences, and the theoretical implications of this latter fact for state efforts to rein in prison growth.

Figure 1. United States prison population



A. The Strain (of Sorts) of Prison Expenditures

As prison populations have climbed, so have the expenditures on maintaining them. States spent a total of $2.8 billion on corrections in 1977 and $39.3 billion in 2004; this represents a thirteen-fold increase in nominal dollars and a four-and-half-fold increase in real dollars (although per-prisoner expenditures have actually declined slightly in real terms).[7] The post-Internet-boom recession of 2000 encouraged state legislators to start thinking about how to pare back prison expenditures;[8] the financial pressures of the current credit crisis have only strengthened their desire to effect reform.

Yet the fiscal story is more complex than that suggested by the total rise in spending, and untangling it may help explain what is driving political calls for reform. There are two possible reasons policymakers want to cut back prison spending: as a short-run response to immediate fiscal pressures, or as part of a longer-run plan to alter more fundamentally how state resources are spent. Looking at the shares of state budgets allocated to corrections, plotted in Figure 2, casts some light on these political motives.

Figure 2 provides two ways of measuring correctional spending.[9] Figure 2A plots corrections' share of the overall budget, and Figure 2B plots their share of the discretionary budget. The results are perhaps surprising: despite a 450 percent increase in total real expenditures on corrections, the share of state budgets given to prisons has remained relatively low and flat. As a share of total spending, the median rose slightly, from 1.4 percent in 1977 to 2.6 percent in 2004;[10] almost no state ever dedicated more than 4 percent of its budget to corrections. In other words,

---

[7] Nominal expenditures grew from just under $11,000 per prisoner in 1977 to nearly $30,000 in 2004; $11,000 in 1977 was worth approximately $33,500 in 2004. All state budget figures come from the US Department of the Census's Survey of Government Finances dataset, online at http://www.census.gov/govs/www/financegen.html (visited Oct 3, 2010). Real values are calculated using the Bureau of Labor Statistics CPI deflator, online at http://www.bls.gov/data/inflation_calculator.htm (visited Oct 3, 2010). I use the prison population counts released by the Bureau of Justice Statistics, online at http://bjs.ojp.usdoj.gov/index.cfm?ty=pbdetail&iid=1763 (visited Oct 3, 2010).

[8] See, for example, Campbell, *Dollars and Sentences* at 3; Wilhelm and Turner, *Issues in Brief* at 2 (cited in note 2).

[9] For those unfamiliar with box plots: for each year the horizontal line in the middle of the box is the percent-of-budget value for the median state, the upper edge of the box is that for the 75th percentile state, and the lower edge is that for the 25th percentile state. The whiskers extend to 1.5 times the difference between the 75th and 25th percentile values (the "interquartile range"), and dots represent outlier values that fall outside the range of the whiskers. Particularly extreme outliers (such as the value near 10 percent in Figure 2A) may reflect errors in the data.

[10] The detailed state-government data needed to calculate the discretionary part of the budget are available only through 2004. The Census has released more-aggregated data through 2008; the median percent of total expenditures spent on corrections remained 2.6 percent that year. Unfortunately, data are not yet available for 2009, when the current crisis appears to have reached its peak.

the large spike in correctional expenditures closely tracks a more general growth in state revenue and expenditures.

Figure 2. Corrections as % of budget

Figure 2A. Total budget

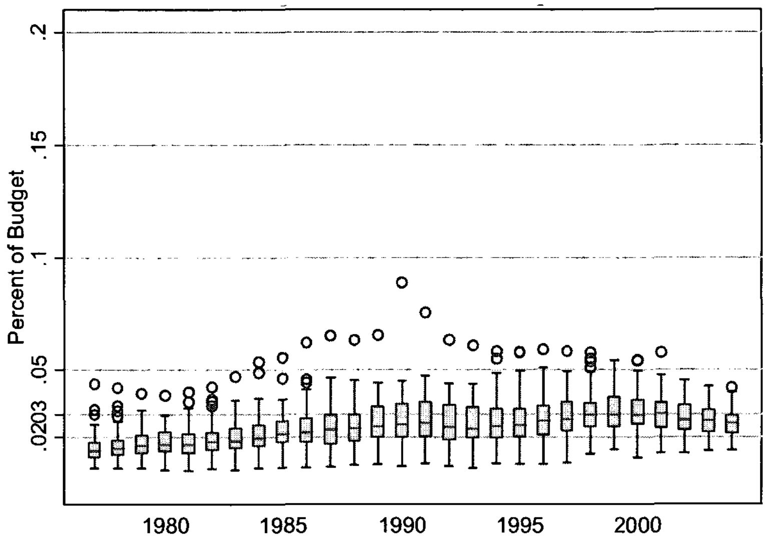
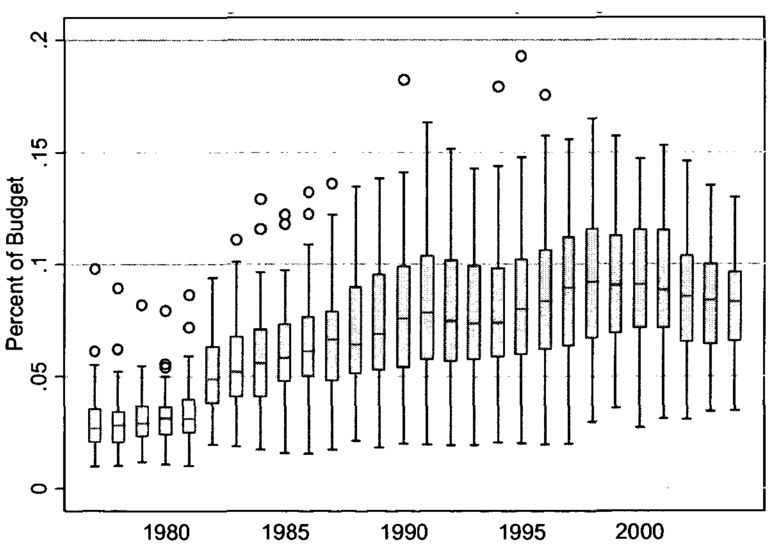


Figure 2B. Discretionary budget



As William Spelman has pointed out, however, looking at overall state budgets is somewhat deceptive.[11] A significant share of state spending is on mandated, nondiscretionary expenditures over which state officials have little year-to-year control, and so it is worth examining the share of the discretionary budget allotted to corrections. Following Spelman, I define the discretionary budget as the overall budget minus spending on public health, primary- and secondary-school education, transportation, debt payments, and certain insurance payouts. Figure 2B plots the results. On the one hand, the results differ from those in Figure 2A in a few noticeable ways: the share of discretionary spending on corrections by 2004 is substantially higher (median around 8.5 percent, with highs reaching over 14 percent), and there is a more appreciable jump in the early 1980s along with steadier upward growth during the early 1990s.[12] On the other hand, the two plots share at least one important common feature, namely that both remain relatively flat throughout much of the 1990s and 2000s.

This latter budgetary stability provides some hints about legislative desires for reform. The 2000 recession does not appear to have affected state financial conditions significantly (at least with respect to corrections), yet it led to calls for reducing correctional expenditures. This suggests that legislators may have viewed it as providing a convenient justification (or political cover) to enact more substantive reforms in criminal justice policy. And by the early 2000s, legislators certainly had valid reasons to shift spending priorities: crime rates had dropped since the early 1990s to the lowest levels in a generation, reducing the political salience of tough-on-crime stances (or soft-on-crime accusations), and freeing politicians to focus on other public harms.

Yet despite the calls for reform, the rate of prison population growth slowed only somewhat, and it certainly did not change direction. It has taken the financial pummeling of the current credit crisis, an event far more significant than the deflating of the Internet bubble, to force states to adopt more drastic measures to slow down or halt prison growth.[13] These obser-

---

[11] William Spelman, *Crime, Cash, and Limited Options: Explaining the Prison Boom*, 8 Criminol & Pub Pol 29, 63 (2009).

[12] The flatness of Figure 2A, and even that of Figure 2B (given the scale of increase in prison population), suggests that slack budgetary constraints may have played an important role in prison growth: it may not be a coincidence that the longest sustained growth in prison populations coincided with one of the longest economic expansions in American history.

[13] Jennifer Steinhauer, for example, discusses the extent to which state budgets have

vations imply that whatever the underlying desire for change, it takes serious financial strain to produce real reform.

The effect of the current recession on future prison populations is thus unclear. On the one hand, states may use it as an opportunity to enact far-reaching reforms that lead to smaller long-run prison populations. On the other hand, they may simply focus on short-run cuts that will ultimately allow prison populations to return to their pre-recession levels. Rather than try to read legislative minds and tea leaves, I focus on both issues: I examine how quickly states can cut prison populations in the short run as well as what kinds of barriers today's prison populations pose to reducing prison size further in the future.

In particular, I evaluate two options that policymakers have at their disposal: cutting back on the number of new admissions or reducing the time served by new admissions. I refer to these as "admission-side" reforms, as opposed to "release-side" approaches such as early releases and furloughs, which I do not consider here.[14] The effectiveness of these reforms depends critically on the durability of the current prison population. If the current population is highly durable—if next year's prison population roughly equals this year's even if no new prisoners are admitted—then admission-side changes will work very slowly. I thus measure the durability of prison populations and ask how quickly these admission-side approaches can reduce prison populations in both the short run (within two to five years) and the long run (fifteen to thirty years).

---

suffered during the credit crisis, and Hal Weitzman examines the recent pressures to reduce prison population sizes (as opposed, perhaps, to just the rate of growth). See Jennifer Steinhauer, *As the Economy Falters, So Do State Budgets*, NY Times 12 (Mar 17, 2008); Hal Weitzman, *US Jails Set to Empty as States Reduce Deficits*, Financial Times 9 (Nov 14, 2009).

[14] I put release-side options to the side for two reasons. First, their effectiveness is trivial to examine. If policymakers can release any number of current prisoners then they can essentially use that tool to set the prison population to any desired level. (A nontrivial question to ask is what limits exist on the number that can in fact be released early, but that is one beyond the scope of this paper.) And second, release-side options may be less politically viable. Early releases, for example, are likely attributable to a specific official (note that when Maurice Clemmons killed four Lakewood, Washington, police officers in 2009, journalists immediately seized on Mike Huckabee's 2000 decision to commute Clemmons's sentence), while the responsibility for failing to incarcerate an offender, or for doing so for less time, is much more diffuse. Furthermore, if people interpret the sentence imposed as a signal of the severity of the crime, then reducing a sentence can be seen as unjust leniency for an offender already deemed to have merited a particular sanction. Imposing a shorter or non-incarcerative sentence, however, takes place before the signal is created and, in fact, produces a lesser signal.

B. Seemingly Nondurable Prison Populations

At first blush, prison populations do not appear to be particularly durable. I recently calculated how many days the median, 75th percentile, and 90th percentile prisoner in each entering cohort[15] actually spend in prison for eleven states, from the late 1980s through the early 2000s.[16] It turns out that the typical prisoner does not spend that much time in prison. Median times to release are often under two years, and in big states such as California and Illinois they are as low as six months. Seventy-fifth percentile times to release are often under five years, and in some cases they are as low as one or two years. The results are perhaps surprising, given the frequent attention—in academic and media reports alike—to prisoners receiving very long sentences.

Such low numbers suggest that prison populations should decay quickly. Consider a state like California, where half of all admitted prisoners are released within six months and 75 percent within a year. Such rapid rates of release seem to imply that the prison population in, say, 2011 would be composed almost entirely of prisoners admitted in 2010 and 2011—and thus that such prison populations would respond quickly to changes in admission policies.[17]

But this need not be the case: low median and 75th percentile times to release can mask an influential cadre of long-serving inmates. And these long-serving inmates can restrict a state's ability to reduce its prison population via admission-side reforms in both the short and long runs. Understanding their importance is thus essential for any sort of reform.

Two simple examples illuminate the problem of long-serving inmates. The first points out a potential pitfall in drawing inferences from low median times served: the median can decline even when prison growth is driven almost entirely by inmates serving longer sentences. The second then demonstrates how this

[15] In other words, I look at all the prisoners who enter in a particular year and calculate how long it takes the median prisoner to be released. This is in contrast to how the Bureau of Justice Statistics often derives median times, which it does by computing the median time served by all prisoners released in a particular year.

[16] See Pfaff, *Correctional Severity* (cited in note 3). The eleven states are those listed in Table 2 below. These states possess approximately 30 percent of all prisoners in the United States; while underrepresenting the South, they reflect a rough cross-section of the US.

[17] In other words, if the prison population in 2011 is composed almost entirely of prisoners admitted since 2010, then reforms enacted at the end of 2009 would seem likely to have an immediate and almost complete effect on prison populations in 2011.

small cadre of long-serving inmates (should it exist) can function as a serious brake on prison reform, even when the median and 75th percentile prisoners leave quickly.

Example 1. Imagine a state that admits one prisoner convicted of a serious crime every year to serve a ten-year sentence. After ten years, the state's prison population reaches its equilibrium size of ten.[18] In some future year, though, the state decides to crack down on crime and enacts two reforms. First, it increases the sanction for the serious crime from ten years to twenty. Second, it starts incarcerating two minor offenders each year for one-year sentences. The new equilibrium prison population is now twenty-two.[19] Note that the median time served in each entering cohort drops from ten years to one year,[20] despite the fact that 91 percent (20/22) of the growth is due to longer sentences.[21]

Example 2. A state admits one hundred prisoners every year, of whom fifty serve six months, twenty-five one year, and twenty-five x years. Thus half of all prisoners are released within six months and 75 percent within a year, which approximates the actual release rates in California and Illinois. It is easy to show that the durability of prison populations, and thus the effectiveness of admission-side reforms, depends critically on x.

Table 1 reports the effect of three values of x: two years, five years, and twenty-five years. The first row provides the equilibrium population for each x.[22] The second and third rows consider a hypothetical "shocked population": how big would the prison population be if, after reaching equilibrium, the state simply admitted no one to prison for one year? The second row presents the absolute size of this shocked population, and the third row ("% Change (1)") the percentage change. The fourth row ("% Change (2)") gives the subsequent percentage change that would

---

[18] For the first ten years, the state admits one prisoner each year and releases zero. At the beginning of the eleventh year, it admits one prisoner but also releases one (the prisoner admitted ten years earlier), thus establishing an equilibrium population of ten.

[19] In any given year, there are twenty inmates serving twenty-year sentences (one with twenty years left, one with nineteen left, and so on) and two serving one-year sentences.

[20] The median time to be served of those currently in prison is, of course, twenty years. But the median times of both those entering and those leaving prison are just one year.

[21] Or perhaps 100 percent, if one considers the decision to admit two minor offenders for one year as a decision to raise their sentences from zero years to one.

[22] For example, if $x = 5$, then at any given point there are twenty-five inmates serving six months (since only half of the fifty are in prison on any specific day), twenty-five serving one year, and 125 serving five years (twenty-five with all five years left, twenty-five with four years left, and so on down to twenty-five with one year left).

take place were the state to admit no one to prison for a second consecutive year.

Table 1. Population durability for various long-serving rates

	Value of x		
	2	5	25
Eq. Population	100	175	550
Shock Pop.	25	100	475
% Change (1)	75%	43%	14%
% Change (2)	100%	25%	5%

As Table 1 makes clear, even in the presence of low median and 75th percentile times to release, prison populations can be quite durable. If x equals 5, for example, a one-year freeze results in a drop of only 43 percent, even though 75 percent of all those admitted the year before leave prison that year. If x equals 25, the results are even more dramatic, with a drop of only 14 percent.

Note, too, that the long-serving inmates can cause the effect of a reduction in admissions to *decline* over time. For an x of 5, the effect of a second-year freeze is barely half that of the first year freeze (25 to 43 percent), and it is two-thirds smaller when x equals 25 (5 to 14 percent).[23] The intuition is straightforward. The reforms quickly remove all the short-serving inmates, leaving behind only the slowly depreciating long-serving inmates.[24]

There are a few important points to take away from these examples. First, despite the presence of low medians and other quantiles, state prison populations may nonetheless be home to a durable core of prisoners. Second, these prisoners may make admission-side reforms almost impossible (as when x equals 25). And third, when the number of long-serving inmates is sufficiently low, states may be able to accomplish a fair amount in the short run with admission-side reforms but then face diminishing returns in the longer runs (such as when x equals 5).

As the next two sections make clear, the third point seems to best explain what we see in the data. States can reduce their

[23] For x = 5, the second year of admissions freeze sees the prison population drop by only 25, from 100 to 75; for x = 25, also a drop of only 25, but from 475 to 450.

[24] As this core depreciates, the percent effect will gradually creep up again, since (by the design of this example) the number leaving each year is fixed but the number remaining declines. Thus the value of 100 percent for x = 2.

current populations rather quickly using admission-side reforms, but the effectiveness of such policies plateaus. In the long run, then, there is a core of prisoners admitted between the late 1980s and early 2000s—perhaps on the order of 100,000—that may remain in prison through 2020 in the absence of release-side reforms. Part II turns its attention to the short-run responsiveness of prison populations, and Part III to the long run.

II. THE SHORT-RUN RESPONSIVENESS OF PRISON POPULATIONS

To examine how quickly state prison populations would respond to admission-side reforms, I use data from the National Corrections Reporting Program (NCRP) to simulate prison population growth under varying assumptions. The NCRP is an inmate-level dataset that allows me to calculate (to the day) how long offenders spend in prison; I use it to construct model prison populations and subject them to various admission-side shocks, measuring how the populations change in response. My results suggest that states can realize substantial declines in prison populations within one or two years solely by adjusting the number of new inmates admitted or (to a lesser degree) the time these new entrants serve.

The Bureau of Justice Statistics began the NCPR in 1983, and while approximately forty states participate in the program today, I focus here on the eleven states that provide the most reliable data for sufficiently long periods of time: California, Colorado, Illinois, Kentucky, Michigan, Minnesota, Nebraska, New Jersey, South Dakota, Virginia, and Washington State.[25] The most obvious limitation to this sample is that it underrepresents the South. Otherwise, these states possess roughly the expected number of prisoners (22 percent of the states and 31.5 percent of the nation's prisoners in 2002; 20 percent and 19.5 percent, respectively, without California), and they are relatively representative of the country as a whole when it comes to political ideology, demographics, and so on.[26] Table 2 lists the first year for which each state provided complete data and the fraction of the nation's prison population each state housed in 2002.[27]

---

[25] As I explain in the Appendix, I consider a state's data sufficiently "reliable" in a particular year if the number of admissions and releases reported by the NCRP are within 10 percent of those values as given by the Bureau of Justice Statistics' National Prison Statistics program. A "sufficient" number of years is at least eleven.

[26] Pfaff, *Correctional Severity* (cited in note 3), evaluates the representativeness of this sample in more detail in its Appendix.

[27] For each state the time series runs through 2002, the last year of NCRP data

Table 2. Included states

State	First Year of Data	% of Nat'l Pris. Pop.
California	1988	13%
Colorado	1992	2%
Illinois	1990	3%
Kentucky	1988	2%
Michigan	1987	4%
Minnesota	1989	1%
Nebraska	1990	0.3%
New Jersey	1992	2%
South Dakota	1991	0.2%
Virginia	1987	3%
Washington	1987	1%

Note: First year of data refers to the first year of reliable data, not necessarily the first year the state participated in the NCRP. The national population percentages are for 2002.

The Appendix discusses the NCRP, and how I address its limitations, in more detail. For our purposes here, what matters is that it provides the date of entry and exit for each inmate (or just the date of entry for those not released by the end of 2002) as well as the inmate's official sentence (specifically, the total maximum sentence imposed) and most serious conviction offense.[28] I know the exact number of inmates released from each annual entering cohort, but due to errors in the NCRP, I have only (decent) approximations of the number admitted each year who are not released by 2002.

I use the NCRP data to create ersatz prison populations that I can then manipulate. I start by calculating how many prisoners each entering cohort contributes to each subsequent year's prison population: of those admitted in, say, 1990, how many are still in prison at the end of 1990? 1991? 1992? And so on through the end of 2002. This allows me to build up prison populations that look nothing like real prison populations in the early years but closely mimic them by the mid to late 1990s. I then examine how

available when I began this paper. Note that the NCRP provides data only on state prisoners, not federal prisoners housed in those states.

[28] There are, of course, other relevant sentences, such as the minimum, but the total maximum is the sentence that is reported most consistently. Shortcomings in the NCRP also force me to cluster inmates' offenses into fifteen categories and the sentence imposed into sixteen. The Appendix provides a more in-depth explanation of these issues. Despite its limitations, the NCRP remains the richest multistate prisoner dataset available.

these later-year prison populations respond to changes in admission policies.

A concrete example can clarify. Table 3 provides data from a hypothetical state whose data start in 1990. In that year, the state admits 1,000 inmates, 500 of whom are still incarcerated at the end of 1990, 200 at the end of 1991, 100 at the end of 1992, and none by the end of 1993. In 1991, the state admits 1,500 inmates, 750 of whom are still incarcerated at the end of that year, 300 at the end of 1992, and so on. The value of 1,400 in the starred cell reflects the fact that by the end of 1992, admissions since 1990 have contributed a net total of 1,400 inmates to the state's prison population (the 100 remaining from 1990, the 300 remaining from 1991, and the 1,000 remaining from 1992). This "Total" row is my ersatz prison population.

Table 3. Simulation example

Year	Admits	Prisoners remaining at end of year:				
		1990	1991	1992	1993	1994
1990	1000	500	200	100	0	0
1991	1500		750	300	150	0
1992	2000			1000	400	200
1993	2500				1250	500
1994	3000					1500
Total		500	950	1400 (*)	1800	2200

Note that "Total" does *not* report the actual prison population, nor are changes in the "Total" row equal to changes in the actual prison population. The real prison population in 1991, for example, is larger than 950 since that number does not account for any prisoners admitted prior to 1990 who are still in prison at the end of 1991. Similarly, the change in "Total" from, say, 1991 to 1992 reflects only the effect of post-1990 changes; some of those admitted prior to 1990 are released during this time as well, but their departure is unrecorded. Thus "Total" underestimates the size of the population but overestimates the rate of growth.

In effect, the "Total" row calculates the state's prison population by assuming that it is equal to zero when a state's data start. This is done solely as a concession to limitations in the

NCRP data,[29] but it nonetheless yields informative numbers. Though a highly inaccurate measure of total population at the start of a state's time series, the "Total" measure closely approximates the true population by the mid to late 1990s (providing further evidence that most prisoners serve fairly short terms).[30] Table 4 compares the values of my ersatz prison populations to their real counterparts (taken from the Bureau of Justice Statistics' National Prisoner Survey) in 1996, 1999, and 2001.

Table 4. Comparing ersatz to real prison populations

St	1996			1999			2002		
	Real	Ersatz	%	Real	Ersatz	%	Real	Ersatz	%
CA	146049	135248	92	163067	152413	93	159444	143446	90
CO	12438	9819	79	15670	13800	88	17448	15902	91
IL	38852	36416	94	44660	43400	97	44348	43726	99
KY	12910	12257	95	15317	14667	96	15424	15628	101
MI	42349	37701	89	46617	45324	97	48849	47838	98
MN	5158	5968	116	5969	6968	117	6606	7863	119
NE	3287	3670	112	3688	4675	127	3937	4936	125
NJ	27490	23087	84	31493	28080	89	28142	25151	89
SD	2063	1819	88	2506	2368	94	2790	2646	94
VA	27655	27767	100	29789	29881	100	31662	32112	101
WA	12527	11983	96	14590	14152	97	15159	14735	97

Note – Percentages measure how close the ersatz populations come to the real. The "Real" values are taken from the Bureau of Justice Statistics' National Prisoner Survey (NPS).

As is immediately clear, even by 1996 most ersatz prison populations are quite close to their real levels, and all are within at least 89 percent of the NPS by 2001. In some cases, my ersatz prison populations are *larger* than the NPS estimates, even though the latter values include prisoners not in the former (that is, those admitted before the start year given in Table 2 and not yet released). These overages, however, are always small in absolute value; such discrepancies are to be expected, given the notorious noisiness of criminal justice statistics.

Though my ersatz populations closely approximate the real populations, my design does introduce a potentially important source of selection bias. The prisoners missing from my populations are, by definition, those serving the longest sentences—the

[29] By definition, data before the first available year are unreliable.

[30] In other words, my approach assumes that California's prisons are empty at the start of 1988 (and Colorado's at the start of 1992, Illinois's at the start of 1990, and so on), and that the ersatz 1988 population is just the number of prisoners admitted in 1988 and not released by the end of that year. Such a measure clearly underestimates the real population, since it ignores all those admitted before 1988 who are still in prison at the end of 1988. The 1995 estimate, however, is closer to the real value since the importance of the pre-1988 inmates declines every year (as more and more of them are released between 1988 and 1995), and the 1999 estimate is thus closer still.

very inmates whose influence I am attempting to measure.[31] I account for this bias below; it does not appear to exert a significant effect.

To explore the durability of prison populations, I use my ersatz populations to effectively recreate the thought experiment given in Table 1 above: How quickly would prison populations decline if states permanently reduced their rates of admissions? I shock the size of the admission class in some year, say 1999, by simply dropping some percent of the entering cohort from my sample and comparing population growth in the shocked and unshocked scenarios.

A simple example, built on that given in Table 3, illuminates how this works. Assume I build up the ersatz prison population until 1992, at which point I shock all future admission cohorts, randomly reducing their numbers by 10 percent. Thus in the shocked scenario the state would admit 1,000 prisoners in 1990; 1,500 in 1991; but only 1,800 in 1992; 2,250 in 1993; and 2,700 in 1994. And of the 1,800 admittees from 1992, I would observe 900 of them remaining at the end of 1992, 360 at the end of 1993, and 180 at the end of 1994; the patterns for the 1993 and 1994 entering cohorts would follow suit.[32] My shocked ersatz prison population would thus be 500 in 1990 and 950 in 1991 (since no changes have yet occurred), 1,300 in 1992 (compared to 1,400, a 7.1 percent reduction), 1,635 in 1993 (compared to 1,800, a 9.2 percent reduction), and 1,980 in 1994 (compared to 2,200, the full 10 percent reduction).[33]

I consider two types of shocks here: to the number admitted and to time served. In other words, states could attempt to reduce future prison populations either by admitting 10 percent fewer inmates each year or by keeping the number of admissions unchanged but requiring inmates to serve sentences that are 10 percent shorter in practice.[34] Both are admission-side reforms,

---

[31] In other words, the only prisoners unaccounted for in California's ersatz prison population in 2002 are those who were admitted before 1988 and thus have served at least fourteen years.

[32] Approximately. I do not cull observations based on time actually spent in prison, but a uniformly random 10 percent reduction will, on average, reduce each cluster of time served by 10 percent.

[33] Because my example does not have any long-serving inmates (100 percent are released within three years), a 10 percent reduction in admissions leads to a prison population 10 percent smaller in three years.

[34] If earlier releases or fewer admissions increase the crime rate (either through higher recidivism or first-offense rates) and states respond to the higher crime rates by increasing their incarceration rates, then my results may overstate the effectiveness of admission-side reforms. How significant this effect may be, however, is unclear. The

and both have arguments in their favor. Cutting sentence length, for example, aligns well with the research showing that the certainty of punishment is a greater deterrent than its severity; cutting time served may thus not significantly reduce deterrence.[35] And it may be politically more tenable to impose lesser sentences on many offenders than no prison sentences at all on a smaller number of criminals.[36] On the other hand, there is evidence that even short incarcerations can impose important personal and social costs, so reducing the number of people entering prison in the first place may yield important dividends.[37] As I show below, shocks to the number admitted reduce prison populations faster than comparable reductions in time served.

I start with the number-admitted shocks. I develop four counterfactuals here: 100 percent reductions (total admissions freezes) starting in either 1999 or 1996 and continuing through 2001, and 25 percent reductions starting in either 1999 or 1996 and again running through 2001.[38] Clearly, no state would ever completely stop admitting people to prison, but such a scenario creates a hypothetical comparable to the illustrative example given in Table 1; it thus provides a clear way to shed light on the effect of long-serving inmates on short-run reductions. The 25 percent cut reflects a more plausible, if still somewhat extreme, situation. Starting the cuts in 1996 grants me a longer view, but starting the cuts in 1999 minimizes the importance of changing policy over time.[39] For reasons I explain in the Appendix, I end

---

relationship between length of sentence and recidivism appears to be week (see, for example, Bureau of Justice Statistics Special Report, *Recidivism of Prisoners Released in 1994* (DOJ 2002), online at http://bjs.ojp.usdoj.gov/content/pub/pdf/rpr94.pdf (visited Oct 3, 2010)), and while the best evidence suggests that decreasing incarceration increases crime on average (see, for example, Steven Levitt, *The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation*, 111 Q J Econ 319 (1996)), it is no longer clear that that is true on the margin, given the size of the US prison population.

[35] Consider Erling Eide, *Economics of Criminal Behavior*, in 5 *Encyclopedia of Law and Economics* § 8100 at 345 (Edward Elgar 2000).

[36] Note, though, that implementation could be challenging. Control over time served is highly diffuse, spread across prosecutors, defense attorneys, and judges (during plea bargains and trials), as well as parole boards (after sentencing).

[37] See, for example, John H. Laub and Robert J. Sampson, *Turning Points in the Life Course: Why Change Matters to the Study of Crime*, 31 Criminol 301 (1993).

[38] I spread the 25 percent cut randomly across the entire cohort. In an earlier version of the paper I ran an alternate specification in which I cut 25 percent of the total population, but restricted the cuts to nonviolent offenders (if nonviolent offenders made up 60 percent of the entering cohort, I reduced their numbers by 41.67 percent for a total cohort decline of 25 percent). The results were almost identical to those of the random 25 percent reduction, so I omit the nonviolent approach here.

[39] State sentencing policies have changed over the years; the closer to the present I

my sample in 2001 rather than 2002, so the 1999 cuts allow me to see what happens over two years, the 1996 cuts over five.

Figure 3 graphs the 100 percent cuts and Figure 4 the 25 percent; Table 5 summarizes the findings. In each graph, the solid line is the unshocked ersatz prison population and the dashed line breaking away from it the shocked population. In Figure 4, the gray dashed line in the lower part of each graph is the "target" line, namely the unshocked population reduced by 25 percent.[40] The faster the counterfactual line converges on the target line, the less durable state prison populations are—or phrased differently, the less important are long-serving inmates.[41]

Figure 3 makes it clear that population freezes result in rapid declines in prison populations. Outside of Michigan, freezes starting in 1999 result in populations over 60 percent smaller by 2001; in six of the eleven states, over 70 percent smaller. The 1996 results are even more dramatic: within five years, four states see reductions of over 90 percent, three over 80 percent, and (basically) all over 75 percent.[42] Look back at Table 1. Though the comparison is imperfect, these results suggest a world akin to that in which $x < 5$, that is, in which "long-serving" inmates often serve fewer than five years.

That said, even total prison population freezes confront a hard, durable core of prisoners. After five years of admitting no prisoners, states still retain between 10 percent and 25 percent of their prisoners, and the rate of decline appears to level off in the counterfactual scenarios, as predicted by the hypothetical given in Table 1. Part 0 addresses this issue in depth.

Figure 4 presents the 25 percent cut, and the results here are even sharper. In the 1999 example, populations are about 17 to 19 percent smaller within two years (except, again, for Michigan, at 13 percent). In the 1996 example, declines are between 19 and 23 percent within five years—often almost the full 25 percent reduction. In many cases, reducing the size of the entering

shock the data, the more likely my results are to track current outcomes. There is nothing significant about 1999 except, given that my sample period ends in 2001, it provides me with more than one post-shock year to observe.

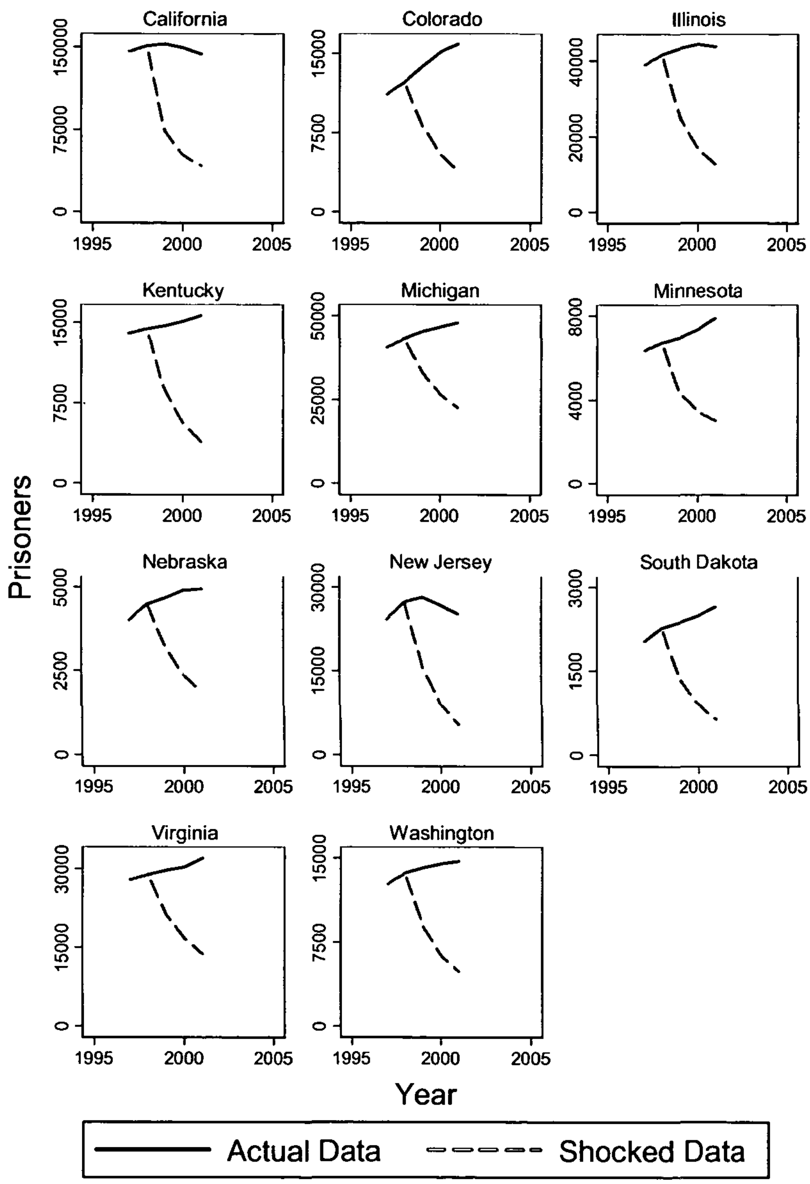
[40] Such a line is not necessary in Figure 3, since the target is zero—the x axis.

[41] If prison populations are not durable at all—if all prisoners serve one year, so the total population in a year is just equal to that year's admissions—then a 25 percent reduction in admissions leads immediately to prison population 25 percent smaller than it otherwise would have been. The more durable the prison population—the more inmates from previous (unshocked) years remain in prison—the smaller the immediate effect.

[42] Michigan, a seemingly harsh state, reports a reduction of 74.9 percent.

cohort by 25 percent is enough to flatten or even reverse the rate of growth.[43]

Figure 3A. 100% shock in 1999



43 Of course, if states insist on increasing their admission cohorts every year, albeit at a level 25 percent lower than they otherwise would have chosen, then populations will ultimately continue to rise, but at a 25 percent-lower level. This can be observed in some instances in the 1996 counterfactual (perhaps most vividly with New Jersey).

Figure 3B. 100% shock in 1996

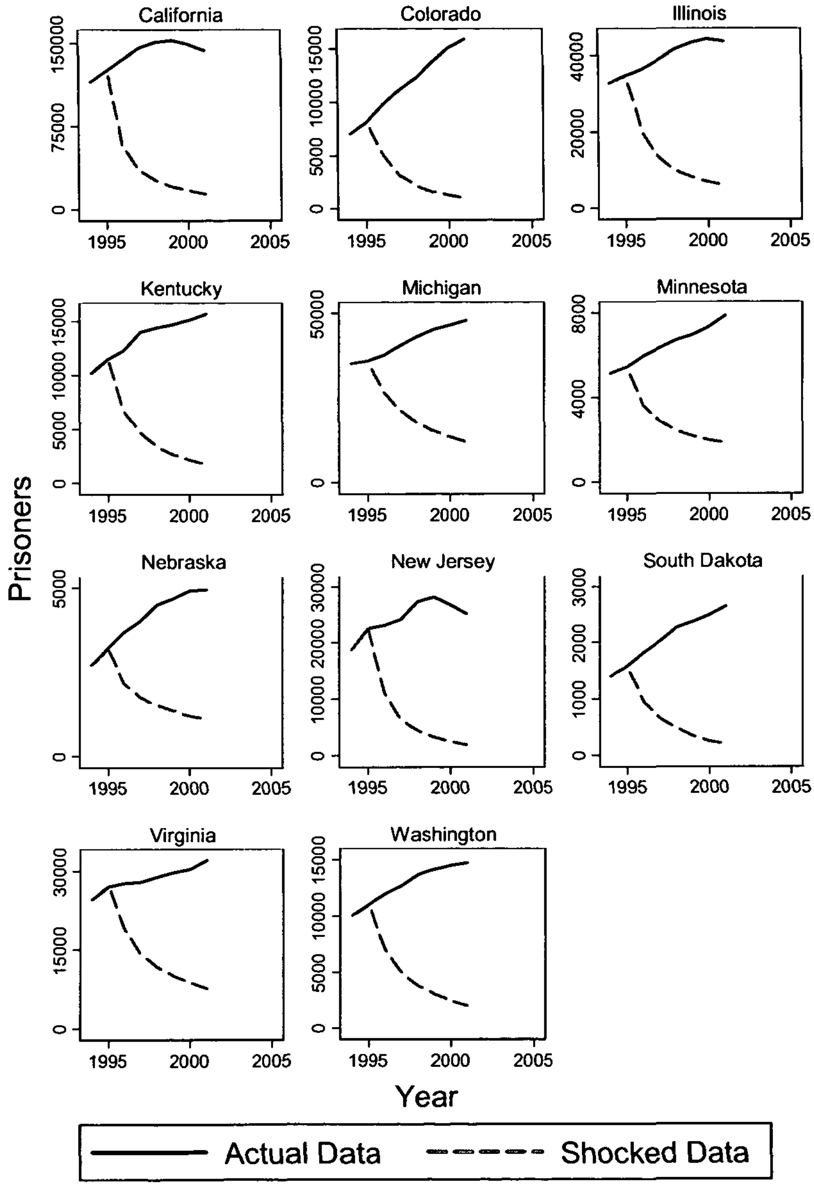


Figure 4A. 25% shock in 1999

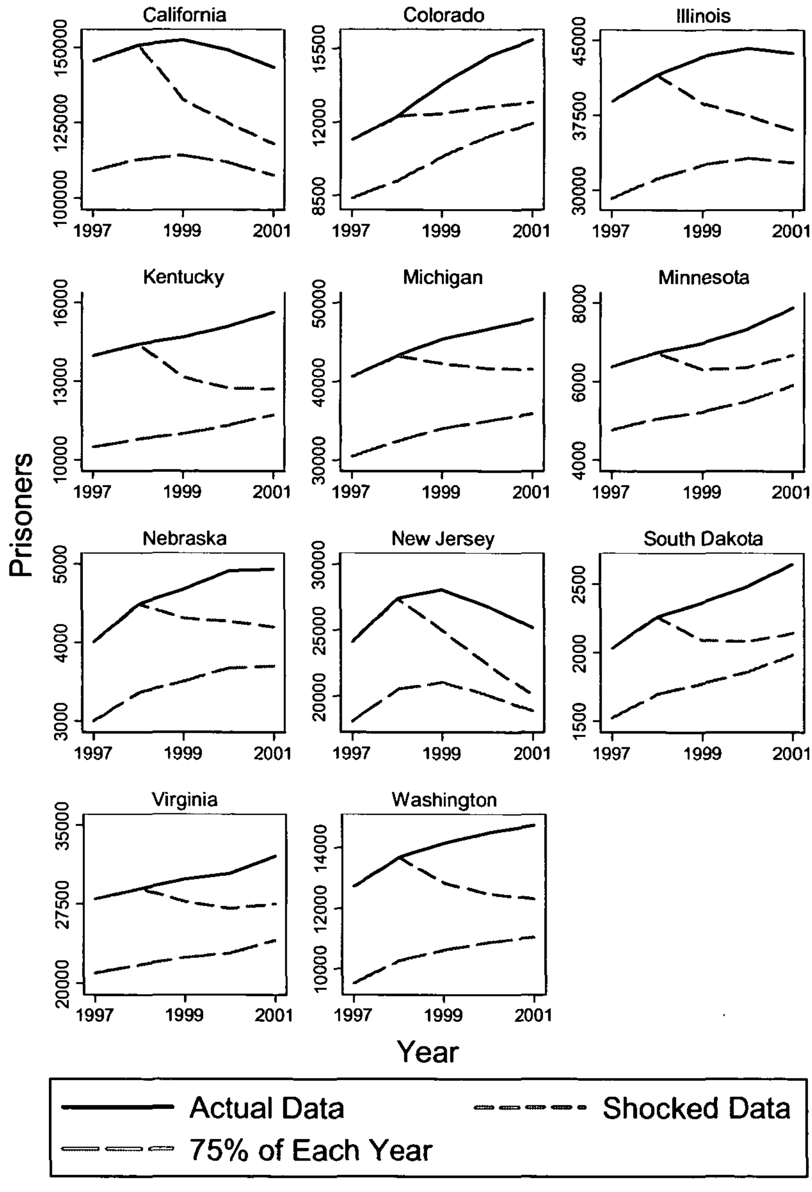


Figure 4B. 25% shock in 1996

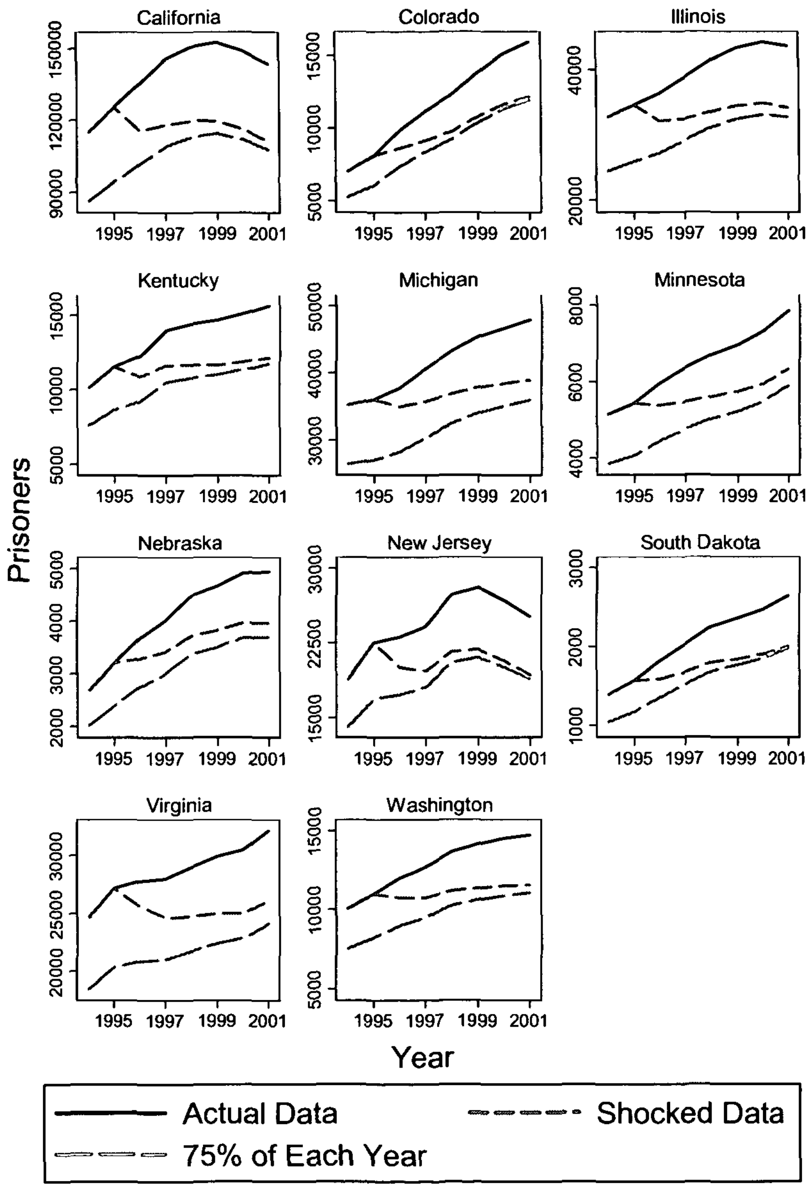


Table 5. Summary of counterfactual results

State	100%: 1999	100%: 1996	25%: 1999	25%: 1996	1999-10%	1996-10%
California	71%	90%	18%	23%	16%	21%
Colorado	76%	94%	19%	23%	17%	21%
Illinois	71%	86%	18%	22%	16%	20%
Kentucky	75%	89%	19%	22%	17%	20%
Michigan	53%	75%	13%	19%	12%	17%
Minnesota	62%	76%	15%	19%	14%	18%
Nebraska	62%	78%	15%	20%	14%	18%
New Jersey	79%	93%	20%	23%	18%	21%
S. Dakota	76%	93%	19%	24%	17%	22%
Virginia	57%	76%	14%	19%	13%	17%
Washington	67%	86%	16%	22%	15%	20%
Values in cells are the percentage difference between the real and counterfactual ersatz populations as of the end of 2001.						

Table 5 summarizes the results of Figures 3 and 4, but it also includes two additional columns, "1996-10%" and "1999-10%." These columns modify the results for the 25 percent shock to account for the selection problem discussed above, namely that the offenders not included in my ersatz prison populations are those most likely to have very long sentences (and thus perhaps most likely to remain in prison during the counterfactual period). To estimate an upper bound on the importance of these offenders, I assume that the actual prison population in 2001 is 10 percent larger than that given in my data—that is, I assume that offenders admitted before the first year of data who are still in prison at the end of 2001 raise my ersatz prison populations by 10 percent in that year.[44] Table 4 indicates that this is an overly cautious upper bound, and thus that this correction will overstate the importance of self-selection (or understate the real effectiveness of the 25 percent shock).[45]

[44] In other words, assume that in 2001 my ersatz prison population for a particular state is 100,000, with a 25 percent counterfactual population of 80,000 (20 percent smaller). My selection correction adds an additional 10 percent, or 10,000, inmates to my ersatz population, raising it to 110,000. It also adds these 10,000 inmates to the counterfactual population, raising it to 90,000 (since these additional 10,000 were, by assumption, admitted before the reforms and remain in prison at the end of the sample period). Thus the decline in the prison population falls from 20 percent to 18.2 percent.

[45] In theory, I could augment my ersatz prison population by the difference between the ersatz and "real" values in Table 4. That some ersatz populations are greater than their real counterparts, however, suggests that the ersatz and NPS values are not wholly

Accounting for the selection problem this way, however, does not alter my basic findings about the nondurability of prison populations. By definition the effect of the admissions-side reforms must be smaller, but the impact is slight: for the 1999 counterfactual the effect of reform is smaller by approximately one to two percentage points (from a baseline of about 15 to 19 percent), and for the 1996 counterfactual by about two percentage points (from a baseline of about 20 to 23 percent).[46] Even the unobserved long-serving inmates do not seem to present much of a barrier to states reducing prison populations by reducing admissions.

Turn now to reductions in time served. As Table 6 points out, cutting time is somewhat less effective than reducing the entering cohort by the same percentage. This Table reports two counterfactuals, each imposing a 25 percent reduction in the time actually served by all inmates, one starting in 1996 and the other in 1999, and it compares the results to the analogous cuts in the number admitted.[47] A 25 percent cut in time served in 1999, for example, produces declines that are between 12 to 49 percent smaller than a 25 percent cut in admissions; for the 1996 cut, the effects are between 3 and 34 percent smaller.[48] While less effective than reductions in the size of the entering cohort, shortened sentences still yield reductions in prison populations that are not trivial—7 to 16 percent for cuts in 1999 and 12 to 21 percent for cuts in 1996. These smaller effects may be justifiable if cuts to time served are more normatively appealing than similar reductions in the number admitted; this is, however, a profoundly challenging question outside the scope of this paper.[49] Moreover, in the long run a 25 percent cut in time served must

---

comparable.

[46] For example, the 1999 counterfactual for California that does not account for the selection effect reports a decline in expected prison population of 18 percent; the counterfactual that does account for selection, a decline of 16 percent. Similarly, for the 1996 counterfactual, the uncorrected results report a 23 percent drop, the corrected results a 21 percent drop.

[47] There is no need to consider a 100 percent reduction in time served, since that is identical to a 100 percent reduction in the number of admissions, which is covered by Figure 4.

[48] The 10 percent check for self-selection, not reported here, returns similar results as in Table 5: the effects are similarly one to two percentage points weaker for both the 1996 and 1999 cuts (though from slightly lower baselines).

[49] There are different pragmatic issues here as well. If being incarcerated causally increases the lifetime risk an inmate faces for future incarceration, then reducing the number admitted may have a more powerful long-term downward push on prison populations than an otherwise-identical reduction in time served. Such dynamics, while of great importance, are beyond the focus of this Article.

yield the same result as a 25 percent cut in the size of the entering cohort;[50] the 1996 results suggest that this convergence can occur rapidly in some states.

Table 6. 25% shock to time served

State	1999	1996	1999-no.	1996-no.
California	15%	19%	18%	23%
Colorado	12%	17%	19%	23%
Illinois	16%	21%	18%	22%
Kentucky	14%	20%	19%	22%
Michigan	7%	12%	13%	19%
Minnesota	11%	15%	15%	19%
Nebraska	11%	16%	15%	20%
New Jersey	15%	21%	20%	23%
S. Dakota	13%	19%	19%	24%
Virginia	7%	12%	14%	19%
Washington	11%	17%	16%	22%

Values in cells are the percentage decline as of the end of 2001 for each of the counterfactual scenarios. "1999-no." and "1996-no." restate the results for the 25 percent cut in *admissions* given in Table 5.

To summarize: If politicians wish to reduce prison populations quickly in response to the present financial crisis, admission-side reforms appear to be up to the task. They may not be as rapidly effective as release-side policies—a state could reduce its prison population by 25 percent in a single day by releasing 25 percent of its prisoners—but they come close, and they may be far more politically viable. Moreover, admission-side reforms can lead to long-run reductions more efficiently than repeated release-side actions.[51] Prison populations certainly contain some long-serving inmates, but states can accomplish a lot within the constraints they create. That said, these long-serving inmates do

[50] This follows directly from the associative property of multiplication. In the long run, a state's prison population is simply $\sum_y$[(number serving y years)×y]. A 25 percent cut to admissions changes this equation to $\sum_y$[(0.75×number serving y years)×y], and a 25 percent cut to time served to $\sum_y$[(number serving y years)×(0.75×y)]. These are identical quantities. The two methods have different transition paths to equilibrium, however, and thus do not yield immediately identical results.

[51] A one-time change that results in fewer annual admissions—such as establishing a diversion program to drug courts—can lead to a long-run reduction in the prison population, while release-side reductions have to be implemented each time the state wishes to reduce the size of the prison population.

impose some limits on how far admission-side reforms can go, and the next section examines just how durable these long-serving inmates are.

III. THE LONG-RUN RESPONSIVENESS OF PRISONS

There is an ambiguity lurking in Figure 3. On the one hand, the rapid declines in prison populations following the admissions freezes imply that there are not that many long-serving inmates. On the other hand, even after a four-year freeze as many as 25 percent of all prisoners may remain behind bars, suggesting that a relatively small number of long-serving inmates can still impose a meaningful floor on the effectiveness of admission-side reforms.[52] In this section I estimate the size of this floor and its rate of decay over time. I also consider how significant the long-serving inmates are to the overall cost an entering cohort imposes on state budgets; after all, an inmate who serves three years is three times as expensive to house as one who serves one year, so a small number of long-serving inmates can consume a disproportionate amount of resources.

The importance of long-serving inmates has often received significant attention. The Sentencing Project, for example, recently released a report—provocatively titled *No Exit*—pointing out that approximately one in eleven prisoners (over 140,000) in the US are currently serving life sentences, with almost a quarter (about 41,000) of these being sentences of life without parole.[53] And life sentences reflect only a portion of the very long sentences handed down by our criminal justice system. The eleven states I examine here, for example, sentenced nearly two offenders to terms of over twenty years for every one given a life or death sentence during the sample period.[54]

Yet as Figure 5 makes clear, focusing on the sentence imposed, rather than time served, can overestimate the number of

---

[52] An example can make clear how this floor operates. According to my data, at the end of 2002 California prisons contained 143,446 inmates who had been admitted since 1988 and had not yet been released. Assume that 10,000 of these inmates will not be released until 2010. Then even a total admissions freeze starting in 2003 could not by itself push the population below 10,000 until 2010.

[53] Ashley Nellis and Ryan S. King, *No Exit: The Expanding Use of Life Sentences in America* 7–9 (The Sentencing Project 2009), online at http://sentencingproject.org/doc/publications/publications/inc_noexitseptember2009.pdf (visited Oct 3, 2010).

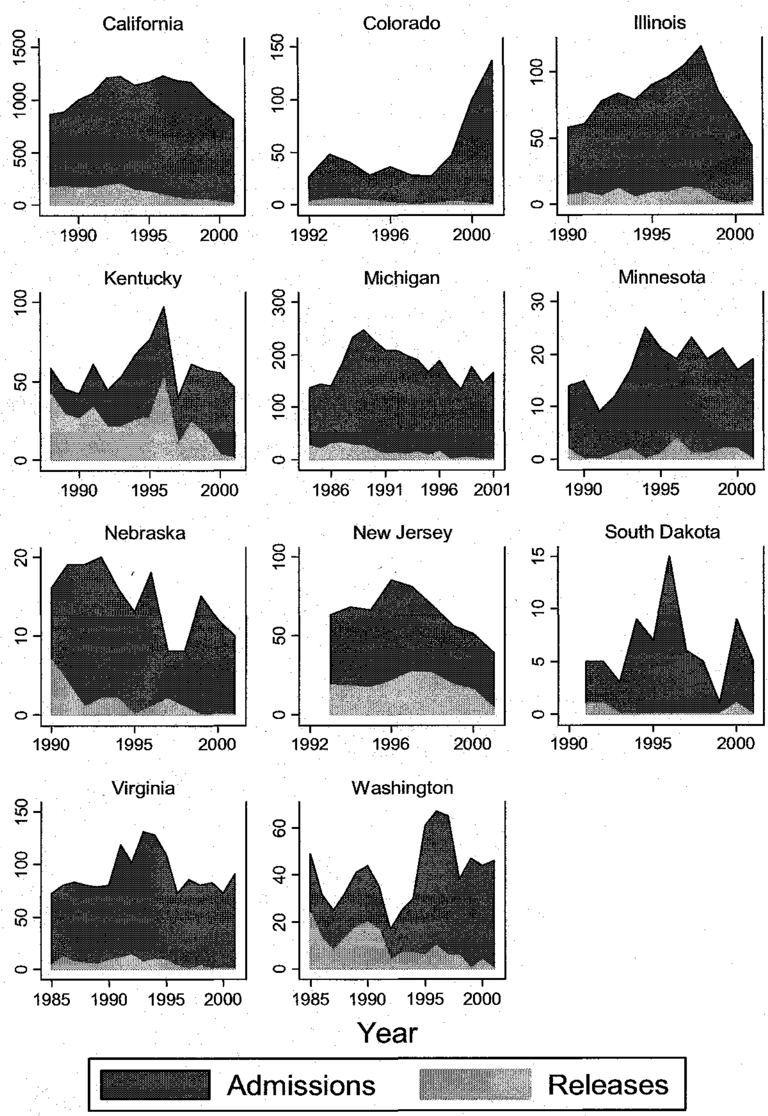
[54] Specifically, states imposed 39,531 life or death sentences, 43,388 non-life sentences of over twenty-five years, and 25,823 sentences of twenty to twenty-five years. States sentenced another 58,893 to terms of fifteen to twenty years, so more than three inmates received sentences of at least fifteen years for every one receiving a life or death sentence.

long-serving offenders. Figure 5 plots the number of inmates in each entering cohort sentenced to some form of life sentence (life, life plus years, life without parole, or death[55]) along with the number of such inmates from that cohort released by 2002. To be cautious, I restrict "released" to mean some form of supervised release, and thus I exclude releases such as transfers or even deaths and suicides.[56] Though most inmates sentenced to life terms serve long sentences, an appreciable number of them have been released by 2002. In states such as Kentucky and Washington, as many as 50 percent of inmates with life sentences are released within ten or fifteen years; more common release rates are on the order of 15 to 20 percent within that time. Even "life," in other words, often does not mean life.

[55] Classifying death as a *long* sentence may be somewhat controversial, but of the states in my sample only Virginia executes a meaningful number of its death-row inmates (105 between 1976 and 2009). California has executed thirteen, Illinois (which is now effectively a non-death penalty state) twelve, Kentucky three, Nebraska three, South Dakota one, and Washington four. See Death Penalty Information Center, *Number of Executions by State and Region Since 1976*, online at http://www.deathpenaltyinfo.org/number-executions-state-and-region-1976 (visited Oct 3, 2010). Thus in most states a death penalty sentence acts as a form of life (or at least a very long) sentence.

[56] The NCRP often does not define its terms clearly, so it is likely that I am excluding both transfers within a state's system, which clearly do not reduce a state's prison population, and transfers outside of that system—such as deportations—which do.

Figure 5. Releases from life terms



The cautionary tale told in Figure 5 shapes how I address the durability of the unreleased, long-serving inmates. While I have data on the sentence imposed on every offender, what matters more is how long each of these offenders is expected to actually spend in prison, and the sentence imposed may not be that reliable a signal. I thus attempt to estimate how long the inmates in my ersatz prisons who have not been released by the end of 2002 are expected to stay in prison. I know how many of these inmates there are, and I know the sentence each has received; I then predict the rate at which these prisoners will be released over the future years.

To make these predictions, I need to extrapolate, which is always a risky procedure. In the Appendix, I explain in detail how I devise my extrapolated values, but the intuition is straightforward. Take California as an example. The data for California start in 1988, and for that admission cohort I know the percent of inmates released by the end of each year for the next fourteen years (from 1988 through 2002).[57] For the 1989 admission cohort, I know a little less: the percent of inmates released each year over the next thirteen years. And so on. For the 2002 admission cohort, I know the least, namely the percent released by the end of that year. My basic strategy is to estimate, say, the percent of the 2002 entering cohort released in 2003 (one year after admission) using data from 1988 through 2001—all years for which I observe the percent of inmates released one year after admission.

Figure 6 provides a simplified concrete example, using a hypothetical state whose data start in 1997. For that year, I have complete data on the percent released for five years (1997 through 2002); for 1998 the percent for four years (1998 through 2002); and so on. These are the dark gray boxes. I then employ a multiple imputation technique to predict the values in the unshaded boxes. In other words, I use the data from 1997 and 1998 to predict the percent of those admitted in 1998 who will be released five years later (in 2003). At the most extreme, I use the data from 1997 through 2001 to predict the percent of those admitted in 2002 who will be released between 2003 and 2007.

With data starting (at the earliest) in 1997 and ending in 2002, I cannot extrapolate more than five years past the date of admission.[58] For example, for the 2002 cohort I can estimate the release rate through 2007, and for the 2001 cohort through 2006. I can also estimate the number of inmates remaining five years

57 A different way to conceptualize the problem would be to focus on years served, not calendar years: to ask how many admitted in 1988 are released within one year of being admitted (which could be in either 1988 or 1989), two years of being admitted (which could be in either 1989 or 1990), and so on. In an earlier version of this paper I designed the model this way, but for technical reasons years-served approach requires additional, untestable assumptions. Since the results are nearly identical either way, I focus only on the calendar-year approach here.

58 In my real data, I have longer time series. Illinois, for example, provides enough data to extrapolate thirteen years into the future. I can always extrapolate at least eleven years into the future.

after admission, but I have no information on how these inmates will leave over time. In other words, I can estimate (and, in the case of 1997, know exactly) how many inmates are in the light gray boxes—how many will not leave prison until at least the sixth calendar year after admission—but I do not know the rate at which they will subsequently depart.[59]

Figure 6. Extrapolation example

Entry year	Years after admission year						
	0	1	2	3	4	5	6+
1997	1997	1998	1999	2000	2001	2002	2003
1998	1998	1999	2000	2001	2002	2003	2004
1999	1999	2000	2001	2002	2003	2004	2005
2000	2000	2001	2002	2003	2004	2005	2006
2001	2001	2002	2003	2004	2005	2006	2007
2002	2002	2003	2004	2005	2006	2007	2008

Note: For each entry year, the dark gray boxes contain the release years for which I have actual data, and white boxes the years that require extrapolation. Extrapolation is not possible past the years in the light gray boxes.

To handle these long-lasting inmates, I need to make (unverifiable) assumptions about the rate at which they depart. I start by estimating an extreme upper bound. I assume that any prisoner not released by the end of the extrapolation period (the end of 2002 for the 1997 entry cohort or the end of 2007 for the 2002 entry cohort) is serving a true life sentence and will not be released for another fifty years.[60] I also develop two more-moderate rules. The first rule assumes that the inmates remaining at the end of the extrapolation period leave uniformly over the next twenty years; this rule is likely biased towards overestimating the durability of long-serving inmates, given the short median and 75th percentile times reported in Part 0 above.[61] The second

[59] Thus, for 2005 I have an estimate for the percents released that year from the 2000, 2001, and 2002 entry cohorts (since 2005 is within five years of admission, and thus within the range of extrapolation). For the 1999 entry cohort, I have an estimate of how many prisoners are still in prison at the start of 2005, but I do not know how many of them leave that year or any subsequent years. And for the 1997 and 1998 entry cohorts I know even less, since my imputation does not allow me to see past 2003 (for 1997) or 2004 (for 1998).

[60] For the 1997 entry cohort, these inmates leave during 2052; for the 2002 entry cohort, during 2057.

[61] Inmates admitted in 1997 and still in prison at the end of 2002 thus leave uniformly between 2003 and 2023; for the 2002 cohort, between 2008 and 2028.

rule attempts to exploit the information I have on sentences imposed. One inmate admitted in 2002 and still in prison in 2008 may have received a ten-year sentence and another inmate a twenty-five year sentence; it is unlikely that each inmate is equally likely to be released in any subsequent year. I thus create two rates of release: one from the end of the extrapolation period to the end of the sentencing period, and another for any inmates who may be in prison after their official sentence has appeared to expire.[62] Thus, for inmates admitted in 2002 with ten-year sentences, I impose one release rate from 2008 to 2011 (when the sentence would expire), and another from 2011 forward for any inmates who appear to still remain in prison.[63] Specifically, I assume that 0.5 percent of the entering cohort departs each year up until the sentence maximum, and then the remainder decay uniformly over the next five years, though all are released within fifty years.[64]

I do not assume that any of my rules accurately captures the real release rates that we should expect to see in the data. My goal here is to estimate the robustness of extrapolations of this sort and to measure the range of possible outcomes we can reasonably expect to witness in the future. Any extrapolation must be viewed cautiously, and my findings here are broad estimates at best.

Figure 7 plots the basic results. By design, most of my models produce a cadre of inmates who spend many years in prison—under the twenty-year decay rule no state runs out of prisoners[65] until 2032 at the earliest, and under the more complex 0.5 percent rule not until 2051. But there are only a few of these highly

[62] Several of my sentence categories are ranges, such as "eleven to fifteen years." For my purposes here, I focus solely on the maximum of the range, which overestimates the durability of inmates. Efforts to develop a more sophisticated measure were thwarted by errors in the NCRP.

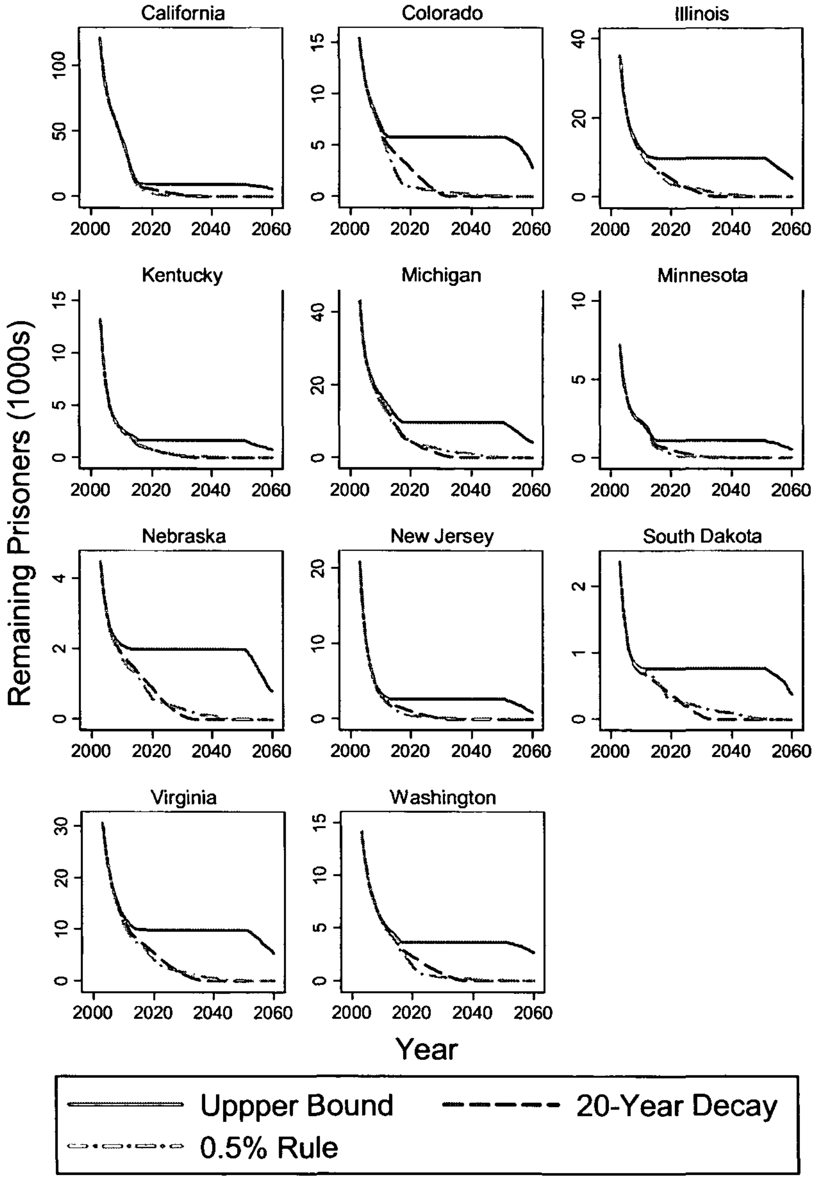
[63] Why might inmates with ten-year sentences serve more than ten years? Malfeasance in prison is one possible explanation. So too is an inmate receiving additional time for a separate offense, the sentencing for which was not completed until after he had already started serving time in prison (due, say, to an ongoing appeal). Unfortunately, data error (such as incorrect information on the sentence imposed, for example) may play a role as well.

[64] For example, assume that my hypothetical state admits 1,000 prisoners for assault with ten-year sentences in 2002, and that according to my extrapolated data 450 remain at the start of 2008 (the end of my extrapolated data). My rule states that 0.5 percent of the original 1,000—or five inmates—leave each year between 2008 and 2011 (for a total of twenty inmates), when the term expires. The remaining 430 leave uniformly over the years 2012 through 2017.

[65] Recall that what is being measured here is not when the total prison population would hit zero but when the pool of inmates admitted during my sample period (the late 1980s/early 1990s through 2002) would finally depreciate to zero.

durable inmates, and overall durability is much less. Consider the number of inmates remaining fifteen years after my sample ends, in 2018. Under the upper-bound model there are 56,015 inmates left across all eleven states; under the twenty-year decay model, 33,935; and under the 0.5 percent model, 25,507. My ersatz prison populations totaled approximately 305,000 in 2003, so in percentage terms between 8 and 18 percent of those still in prison at the end of 2002 would remain as of 2018.

Figure 7. Rate of decay



What are the national implications? Assuming that the experiences in these states (which hold just over 30 percent of US prisoners) are nationally representative, my results suggest that of all the prisoners admitted from the late 1980s/1990s to the early 2000s, between 84,000 and 184,000 would remain in prison nationwide by 2018. The upper bound—built on unrealistic assumptions—is nothing to scoff at, representing about half the size of the entire US prison population in the 1970s. But the less extreme values suggest that, left untouched (that is, subjected to no early-release policies), the pool of inmates admitted during the 1990s and early 2000s will decay rather rapidly on its own.

* * *

I want to conclude this section by addressing one last long-run question: what fraction of the total costs imposed by an entering cohort is produced by the longest-serving inmates? There are important political-economy implications to its answer. The larger the share of costs borne by the future, the tougher it may be to enact meaningful reform.

A simple example makes the problem clear. If housing a prisoner for one year costs the state $10,000, then there are at least two ways an entering cohort can cost a state $250,000 (ignoring present-value discounting for simplicity): by admitting twenty-five prisoners for one year each or by admitting one prisoner for a twenty-five year term. The first option is more costly to political incumbents—the governor, the legislators, the mayors—because they bear the full financial cost of the twenty-five incarcerations, and thus the full opportunity and political costs as well. Not so in the second case, in which much of the spending is deferred. Today's politicians reap the short-run benefits being "tough on crime," but they push much of the costs (both financial and political) onto future administrations.

I thus measure how much of an entering cohort's "lifetime" costs are incurred in the short-run as compared to the long run. For two of my release rules (the twenty-year decay rule and upper bound[66]) I estimate the total number of bed-years that each entering cohort is expected to consume and thus the total cost that this consumption will impose on state budgets.[67] I then derive how much of that total is accumulated during that cohort's first year and first five years (the short run) as well as after ten years and after twenty (the long run).[68] Figure 8 presents the

[66] The results for the twenty-year and the 0.5 percent rules are almost identical, so I omit the latter to conserve space.

[67] Assuming that the annual real cost of incarcerating an inmate is flat, then looking at bed-years is identical to looking at real costs. As noted in Part 0, real per-prisoner costs appear to have declined slightly, which implies that my bed-year analysis will overstate the importance of long-serving inmates to the total real cost imposed by an entering cohort.

[68] An example can make this clear. Consider a state that admits ten prisoners in 1990. Three (A, B, and C) spend one year in prison, two (D and E) spend two years, three (F, G, and H) spend seven years, and two (I and J) spend fifteen years. The total number of bed-years this cohort consumes is fifty-eight: (3x1) + (2x2) + (3x7) + (2x15) = 58. During the first year, ten bed-years are consumed—not just those for A, B, and C, but one year

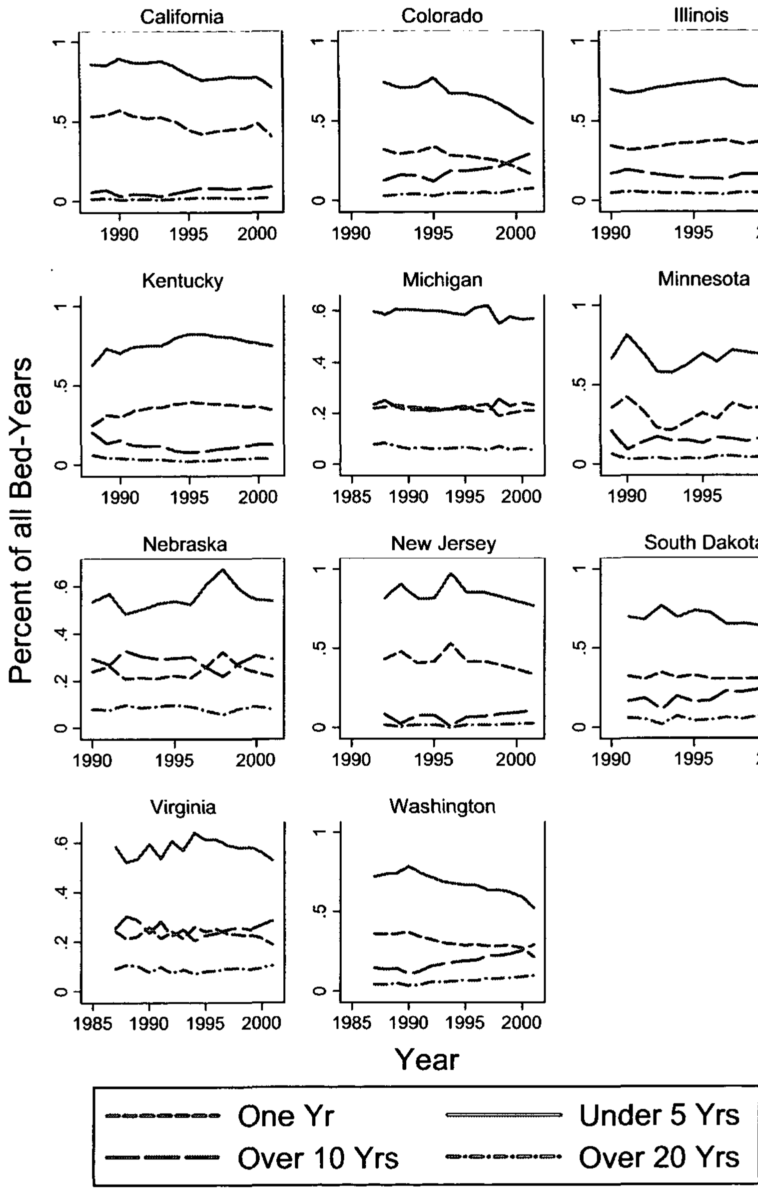
results for the less-extreme twenty-year rule, which suggest that much of the cost of an entering cohort is incurred in the early years of its admission. More than half the lifetime cost is consistently borne during the first five years, with highs often reaching 75 percent and above. In some cases, half the total costs are incurred during the very first year, implying that a majority of the financial punch is felt by the incumbent administrations. Even in the upper-bound results, which by construction substantially overestimate the impact of long-serving inmates, about 40 to 50 percent of lifetime costs are incurred during the first five years in most of the eleven states. That said, the costs pushed onto future policymakers are not trivial. For the twenty-year extrapolation rule, nearly 20 percent (and in some cases as much as 30 percent) of a cohort's total cost is incurred at least ten years after admission in most of the states, and over 5 percent (and in some cases over 10 percent) at least twenty years after admission. These percentages rise to over 40 percent and over 33 percent, respectively, for the upper-bound rule.

Thus some cost shifting to future administrations is clearly taking place. But the policymakers responsible for the sentences imposed on a given cohort nonetheless bear a substantial portion of the lifetime demands they are imposing on state budgets, suggesting that the partially deferred costs of long-serving inmates need not seriously impede reform (though plenty of other political pressures may).

---

each for D through J as well, since an inmate imposes the same costs on the state budget during the first year after admission whether he is serving one year or twenty. During the first five years, the cohort consumes a total of thirty-two bed-years: A, B and C contribute one each, D and E two each, and F through J five each. Looking at the long run, sentences of more than ten years contribute a total of ten bed-years to the total: I and J each consume five bed-years after ten years. So of the total cost, 10/58 (or 17 percent) occurs during the first year, 32/58 (or 55 percent) occurs during the first five years (to partition fully, years two through five contribute 22/58, or 38 percent, of the cost), and only 10/58 (or 17 percent again) is pushed more than ten years into the future.

Figure 8. Timing of consumption



CONCLUSION

As the current financial crisis continues to force states to wrestle with deteriorating fiscal conditions, prison expenditures increasingly look like a promising place to cut spending. With crime rates still at historic lows, the political salience of crime has surely weakened, providing politicians with a window of opportunity. These historic lows further suggest that reducing prison populations may also be good criminological policy, since the marginal admittee to prison may no longer be cost justifiable.

In this paper, I have sought to measure the feasibility of reform. In particular, I have asked whether today's policymakers can quickly restrict or even reverse prison growth by relying solely on forward-looking, admission-side reforms—admitting fewer prisoners or imposing shorter sentences on those admitted—without turning to policies such as early releases for those already incarcerated. The effectiveness of such approaches turn critically on how quickly those already in prison will leave in future years. My results indicate that admission-side reforms may in fact be quite effective. Prison populations do not appear to be particularly durable, so yesterday's decisions do not shape tomorrow's prison populations as strongly as may have been thought. That is not to deny the existence of a durable core of long-serving inmates—such prisoners clearly appear in my data—but rather to suggest that states can accomplish much despite the constraints these prisoners impose.

APPENDIX

In this Appendix I discuss in more detail the nature of the NCRP data and how I handled its limitations, and I explain more fully how I generated the counterfactuals used in Part II and the imputed data used in Part III.

A. The Structure of the NCRP

The Bureau of Justice Statistics began compiling the NCRP in 1983, and it continues to gather data to this day; this paper uses data through 2002, the most recent year available when it was begun. Participation has always been voluntary, and while only thirteen states provided data in 1983, approximately forty do so now. Officials in participating states fill out a card for each inmate when he enters prison, and they complete a separate card upon his release. The NCRP thus provides offender-level data on the exact dates an inmate enters and leaves prison, as well as demographic and offense information. For this paper, I focus solely on the primary conviction offense and the total maximum sentence imposed.[69]

Though it is the most detailed centralized source of information on prison inmates available, the NCRP suffers from some substantial flaws. First, there are significant discrepancies between the NCRP and other sources of data. In many instances, the total number of admissions or release entries in the NCRP for a particular state varies greatly from the same number reported in the BJS's National Prisoner Survey dataset—at times by as much as 75 or 100 percent. The eleven states listed in Table 2 are the only ones that provide a sufficiently long string of consecutive years (at least eleven) that do not deviate too much (generally by less than 10 percent) from the corresponding NPS data.

Second, and more problematic, the NCRP's admissions and release files are wholly separate, with no variable that links an observation in the admissions file to that inmate's subsequent release entry. 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 it does significantly complicate my ability to determine how many prisoners from a particular admission cohort remain unreleased at the end of 2002.

If the NCRP contained no reporting flaws, calculating these remaining inmates would be trivial. The admissions data would tell me, for example, how many thirty-five-year-old black males convicted of arson and sentenced to eight years entered prison in California in 1995 (say, one hundred). And the release files for 1995 through 2002 would tell me how many thirty-five-year-old black males convicted of arson and sentenced to eight years who entered prison in California in 1995 were released in each subsequent year (say, ten each year). The number of these inmates still in prison at the end of 2002 would simply be twenty: the one hundred admissions minus the eighty releases (ten released per year for eight years). The NCRP contains so much demographic data that I would be able to precisely identify which inmates in each admission file never show up in the subsequent release files and who are thus still in prison at the end of 2002.[70] In short, I

[69] The focus on the maximum, as opposed to the minimum, is in part simply a concession to data. The NCRP includes a variable for minimum sentence imposed, but it is blank for a large percentage of entries; total maximum sentence is reported for almost every offender.

[70] Even a state with a prison system as large as California's likely does not admit too

could merge the admission and release files to create an integrated dataset.

Unfortunately, flaws in the NCRP make this impossible. The key problem is that in many cases it appears that over 100 percent of a certain type of inmate in a specific entering cohort are released by 2002: a state may admit one hundred inmates of a particular type in 1990 but appear to release 120 such prisoners over the next thirteen years. In numerous cases, prisoners who do not exist in the admissions files appear in the release files—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. The reasons for these differences cannot be determined.[71]

There is no (apparent) overrelease problem when looking at unconditional total counts.[72] But the magnitude of the errors tends to grow with the level of detail in the data. Conditioning on just the primary conviction offense and the total maximum sentence imposed, as I do here, yields overcounts for some states that run as high as 620 inmates.[73] And these are the errors that remain after I take other steps to minimize the overcount problem: I do not condition on the NCRP's 186 offense codes and continuous measure of sentence imposed, but rather use less granu-

many prisoners with the identical dates of admission, races, sexes, ethnicities, birthdays, educations, crimes of conviction, sentences imposed, credits granted for time already served, and so on.

71 Since the offense variable measures the offense of conviction, such discrepancies cannot be the result of subsequent crimes committed while in prison (for which there is separate entry).

72 The total number admitted to a particular cohort is always greater than the total number of that cohort released over all subsequent years. For example, the 1990 admission file reports that Illinois admitted 17,971 prisoners that year, and the release files for 1990 through 2002 provide information on 17,822 releases from that cohort. I say "apparent," though, because that fact does not guarantee that the number of releases is not an overcount, just that whatever overcount takes place is sufficiently small that it does not lead to an impossible outcome.

73 In other words, the 1988 admission file claims that Michigan admitted 134 prisoners with an offense code of 35 and a sentence range code of 12, yet the release files from 1988 to 2002 contain a total of 754 releasees who are allegedly offense-35/ sentence-12 offenders admitted in 1988, an overcount of 620.

lar measures that reduce the number of offense codes to sixteen[74] and the measure of sentence imposed to fifteen clusters.[75]

These errors in the data preclude me from directly merging the admission and release files. To get around this, I create "phantom" observations for the unreleased members of a cohort (initially assuming there is no overcount problem) and attach them to the release file. A simple example makes this process clear. Assume that Colorado admits one hundred offense-4/sentence-8 inmates in 1994, and that it releases seventy-five of these inmates between 1994 and 2002. The release files provide the complete data on the seventy-five released inmates, and I then create twenty-five "phantom" inmates for those not yet released (who I know are offense-4/sentence-8 and admitted in 1994, though I do not know more than that). I thus have in effect an integrated dataset. The one missing piece of data for the phantom inmates that can be important is the day and month of admission; I randomly assign these dates to the phantom inmates.[76]

All that is left is to correct for the overcount problem. The overcounts do not reflect departures that did not happen, nor the failure to account for admissions that did happen—they are coding errors, inmates who are classified one way at entry and another way at exit.[77] It is impossible, however, to correct these errors directly. Instead I develop a more indirect solution.

---

[74] The sixteen are: murder and other killing offenses (including assault with intent to kill), kidnapping, sex offenses, robbery, assault (including hit-and-runs 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).

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

[76] In the data, the day and month of admission are uniformly distributed across the month and year of admission, respectively, so I draw my fake days and months from a uniform distribution.

[77] For example, according to the wholly aggregated data, California admitted 97,259 inmates in 1990, and between 1990 and 2002 released 96,974 of this cohort, leaving 285 in prison as of the end of 2002. Disaggregating by offense and sentence range codes yields 195 offense/sentence pairs, and the number of inmates remaining in prison at the end of 2002 across these 195 pairs ranges from 774 to −126 (that is, the data indicate that the number of releases for a particular offense/sentence pair exceed the number of admissions by 126). The average number remaining per pair, though, is 1.461538, and 195×1.461538 = 285, the same number remaining from the disaggregated data.

A simple example illuminates the problem and demonstrates 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 remain in prison at the end of 2002. My methodology above would start with the 850 observations from the release files and then create 50 phantom offenders 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.[78] The total remaining in prison is the same: the state has released 850 (520 + 330) of the 900 (500 + 400). But things get tricky when I try to create my phantom observations. My data claim that I should create −20 phantoms of type A and 70 phantoms of type B. In theory, I would like to correct this by turning 30 erroneous phantom Bs into phantom As, but I do not actually know that the number misclassified is 30. I know that it is at least 20, but I have no more information than that. I cannot create "negative" phantoms, and simply ignoring type As (creating zero phantoms of this type) while creating 70 phantom type Bs leads me to overstate 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 five hundred. I then reduce the number of type B phantoms from seventy to fifty—by the number of overcounts of type As in the data. So I end up with zero type-A phantoms and fifty type-B phantoms. Of course, this is likely not the true correct value (in my example, I produce ten too few type As), but it is the best I can do with the data available, and it is likely a close enough approximation.

In practice, I have more than two categories—for California, for example, I have 195 categories in 1990. I thus take a more aggregate approach. I total the number of overages within a state/year and then randomly delete that many phantoms from the non-overage offense/sentence pairs. For example, in California in 1990, 100 of the 195 offense/sentence pairs yield overages, ranging from −1 to −126, for a total of 1,702 excess releases. The remaining 95 pairs have remainders ranging from 0 to 774, for a

---

[78] I am assuming here that the errors are on the release end, not in admissions. I have no evidence either way, and there does not appear to be any ex ante bias to choosing one assumption about the location of the errors over any other.

total of 1,987 phantoms. I thus randomly delete 1,702 of these 1,987 phantoms, leaving me with 285 phantom observations, exactly as required.[79] With this adjustment, I have a complete set of observations. I have real data on all those released, and I have phantom observations on all those admitted but not yet released. I then use these data to produce my ersatz prison populations.

B. More on the Short-Run Extrapolations

Part II above explains the number-of-admissions counterfactuals in detail. One feature of the time-served counterfactual, however, requires further explanation. Unlike the counterfactuals that shocked the number of admissions, those that shocked the time served require looking into the future. Consider the counterfactual that reduced the time served by 25 percent starting in 1996. An inmate admitted in 1996 and—prior to the shock—destined to serve 7 years would not leave until 2003, which is outside the range of the data. With the 25 percent reduction, however, he would leave prison in 2001, after only 5.25 years in prison. I thus need to estimate how many such out-of-range offenders exist and "pull" them out of the future. To do this, I use the extrapolated data that I develop in Part III (and which I discuss in more detail in the next part of this section).

Using such extrapolations, of course, introduces extra uncertainty into the results, since the reliability of the counterfactual values turns on the reliability of the extrapolation. It is for this reason that I examine the period from 1999 to 2001, rather than from 2000 to 2002—no extrapolated values are needed to run the 1999 shock counterfactual through 2001. The only offenders who need to be "pulled" from outside the range of 1999 to 2001 are some of those released in 2002, and I have real data on the number of such releasees and the time they actually served.

C. Extrapolating the Long-Run Models

To calculate how long offenders admitted during my sample period may endure into the future requires extrapolation. To do this, I use the multiple-interpolation package that is part of the Stata 11 release. Figure 6 above explains the basic intuition; here I just touch on some of the more technical aspects of my approach. My data allow me to calculate the percentage of each en-

[79] Recall from note 77 above that by 2002 California had released 96,974 of the 97,259 inmates it admitted in 1990, implying that 285 from that entering cohort remained in prison at the end of 2002.

tering cohort that is released in the year of admission, the next year, and so on, for some number of years: for the 1988 cohort, for example, I can compute fifteen years of release rates, for the 1989 cohort fourteen years, etc., leading to a triangular grid of values akin to that given in Figure 6. Before running the imputation program, I convert the per-period release rates into cumulative release rates. I then run fifty multiple-imputation iterations to fill in the bottom half of the triangle and average across all fifty to generate my estimated release rates.

These estimates require some cleaning. In some cases, the estimated values "backtrack": for a particular cohort the cumulative percent released twelve years out, for example, may be estimated as being less than that eleven years out, an impossibility. In these cases I set the backtracking value equal to its predecessor, effectively assuming in such cases that no one was released in that year.[80] I also cap the cumulative probabilities at one. With these corrections, I am able to estimate the results shown in Figure 6.

[80] Note that my data are disaggregated by offense and sentence range, so I am not assuming that a state released no prisoners for a year, but rather that no prisoners from the backsliding offense/sentence pair were released.