Counterproductive punishment

How prison gangs undermine state authority

Benjamin Lessing

September 2016

Rationality and Society, forthcoming
Counterproductive Punishment:
How Prison Gangs Undermine State Authority

Benjamin Lessing

September 30, 2016

Abstract

States' efforts to provide law and order can be counterproductive: mass incarceration policies, while incapacitating and deterring individual criminals, can simultaneously strengthen collective criminal networks. Prison gangs, by promising rewards and punishments inside prison to those who anticipate incarceration, can control criminal activity on the street. A formal model reveals that common crime-reduction policies, by making incarceration more likely and sentences harsher, can increase prison gangs' power over street-level actors. Leading cases from across the Americas corroborate these predictions: periods of sharply rising incarceration, partly driven by anti-gang laws, preceded qualitative leaps in prison-gang projection of power onto the street. Prison gangs use their capacity to project power not only for criminal governance, but to orchestrate violence—or intentionally curtail it—providing them critical leverage over the state. Thus, even if increased incarceration reduces crime rates, it may do so by strengthening prison-gang power at the expense of state authority.
1 Introduction

Providing law and order is an essential function of the state (Huntington 1968; O’Donnell 1993), and punishment is central to that provision. Threats of punishment undergird state authority at every turn, though ideally they are carried out only rarely: effective state authority implies “a certain minimum of voluntary submission” to the law (Weber 1947). When individuals challenge the state’s authority by disobeying its laws, the state exacts punishment, in part, as a means of restoring its authority (Foucault 1977).

How is punishment—incarceration, in the modern era—supposed to accomplish this? Most directly, by incapacitating captured lawbreakers and deterring potential ones, thus reducing crime and presumably shoring up the rule of law. A vast literature on individual-level effects like incapacitation and deterrence (e.g. Levitt 1996; Liedka et al. 2006; Raphael and Stoll 2009) finds that, at least over some range, increased incarceration reduces crime. On the other hand, important recent work reveals that incarceration, especially at high levels, can have crimogenic and other negative social, economic, and political effects on inmates (e.g. Gottschalk 2008; Lerman 2009; Weaver and Lerman 2010). Policy debates largely revolve around the relative weight of these positive and negative individual-level effects of incarceration.

Such work largely overlooks the collective effects of incarceration, which, I argue, can strengthen prison-based criminal organizations at the expense of the state. Outnumbered prison managers have long ceded partial authority to inmate groups (e.g. Jacobs 1978; Sykes 1958; Venkatesh and Levitt 2000); overcrowding can further foment gang control over aspects of prison life. Increasingly, though, sophisticated prison gangs are leveraging that control to project power onto the streets. From Los Angeles and El Paso to El Salvador and Brazil, they have organized street-level drug traffickers and neighborhood gangs into extensive and lucrative criminal networks (Lessing 2010). The most powerful prison gangs go further, enforcing self-styled social order within vast tracts of urban periphery, and orchestrating mass violence as a bargaining chip to force state concessions. This accumulation of non-state authority can undermine state authority, even though it sometimes leads to reduced crime and violence.
How do prison gangs organize street-level criminal activity? In a key contribution, Skarbek (2011), invoking Mancur Olson (1993), argues that prison gangs act as “stationary bandits”, providing disparate street gangs with criminal governance and taxing the resulting surplus. Skarbek's analytic emphasis, however, is on state absence: although he identifies street members’ anticipation of future incarceration as a necessary condition, his core argument is that retail drug markets constitute stateless spaces, in which prison gangs come to supply unmet demand for governance by traffickers. In this view, prison gangs belong to the analytic category of self-organizing, surplus-maximizing, non-state sources of social order, such as merchant guilds (e.g. Greif 1989; Milgrom et al. 1990), pirate constitutions (Leeson 2012), and embryonic versions of the state itself (Bates et al. 2002; Skaperdas and Syropoulos 1997; Tilly 1985).

Unlike these now-familiar examples, however, prison gangs arise and flourish not beyond the reach of the state’s coercive apparatus, but at its very core. This matters, because prison gangs’ ability to project power depends not only on state absence, but also crucially on state actions: arresting people and physically detaining them in facilities where prison gangs wield immense influence. A paradox of state punitive power emerges: the harsher, longer, and more likely a prison sentence, the stronger outside affiliates’ incentives to stay on good terms with imprisoned gang leaders, and hence the greater prison gangs’ coercive power over those who anticipate incarceration. This mechanism turns the logic of punishment on its head, presenting the state with a serious dilemma: how to punish gang affiliates without strengthening the gang? As an exasperated FBI agent said when her 10-year investigation of a Texas prison gang led to three life sentences, “I think I’ve made them stronger” (Sherman 2010).

This article argues that common state responses to crime like mass arrests and harsher sentences can increase prison gangs’ leverage on the street. After a background section on how prison gangs consolidate power within prison, the third section presents a formal analysis of prison-gang projection of power. This very simple model illustrates how ‘crackdowns’—increases in the likelihood of incarceration—can strengthen prison gangs’ power to tax and recruit street-level criminals if (1) gangs have sufficient capacity to reward or punish inside prison, and (2) arrests insufficiently target only those criminals who comply with prison-gang demands. Moreover, even a well-targeted crackdown can strengthen prison
gangs if it produces harsher prison conditions through, say, overcrowding. The model supplements Kleiman’s (2009) case for “swift and certain” punishment: the gang-strengthening effects of crackdowns can be offset by shorter sentences. It also offers a formal criterion for distinguishing street and prison gangs. Finally, incorporating elements of Becker’s (1968) seminal model, it distinguishes individual and collective effects of incarceration: crackdowns may reduce crime overall while still aiding gang recruitment.

The fourth section corroborates these formal predictions with empirical evidence from the Americas. I focus on three substantively important and well-documented cases, California’s Mexican Mafia (Eme), São Paulo’s Primeiro Comando da Capial (PCC), and El Salvador’s maras (MS13 and Barrio 18). In all three, massive increases in incarceration, driven in part by severe but poorly targeted anti-gang crackdowns, preceded unprecedented expansions of prison gangs’ power on the streets. In Nicaragua, by contrast, repression was targeted at the most violent groups, incarceration grew modestly, and prison-gang projection did not occur (Cruz 2011; Yashar 2012). Available evidence is thus consistent with the hypothesized causal mechanism; more systematic testing would require better data, distinguishing true prison-gang weakness from erroneous or deliberately downward-biased reports.

Having considered how the state enables prison gangs to project power, the fifth section examines how projection erodes state authority. Drawing together original fieldwork—prison visits and interviews with prison-gang members and officials in El Salvador and Brazil—with extant ethnographic and journalistic research, I identify two critical uses of prison-gang projection of power that go well beyond the criminal governance that Skarbek observes in U.S. gangs. First, prison gangs can establish “parallel power” over wider marginalized populations; second, they can orchestrate protest and violence—or intentionally curtail it—to gain bargaining leverage vis-à-vis the state. São Paulo’s PCC gang, for example, launched major terrorist attacks in 2006 that forced concessions from authorities (Hisayasu 2015), while also imposing a “criminal code of conduct” that sharply reduced homicide throughout the urban periphery. Such brazen tactics, while (still?) rare in the U.S., have helped prison gangs in Brazil and

---

1In El Salvador, “mara” has become associated exclusively with MS-13, with “pandillas” (gangs) increasingly used as a generic collective. Because “pandillas” and “gangs” are both so ambiguous in a comparative context, I use “mara” to refer to both groups, with apologies to Salvadoreans.
Central America become major political protagonists, surpassing insurgency as the foremost threat to state authority.

The implications are dire: state efforts to curb crime can inadvertently undermine state authority. Plunging crime rates may not reflect the restoration of state authority, but rather the organization of the criminal underworld into a powerful prison-based collective capable of holding state and society hostage to its demands. Herein lies the danger of studying only individual-level effects of incarceration like incapacitation and deterrence on easily measured outcomes like crime rates.

Instead, in the sixth section, I nest both individual and collective effects within a theoretical framework that considers the impacts of incarceration on state authority (Figure 1). A rough scholarly consensus, captured by the solid line, finds that the net individual effects of incarceration (reducing crime and thus restoring state authority) are large at low levels of incarceration, where basic social order is established, but diminish and may become negative at higher levels. The overarching claim of this article
is that incarceration can also have collective effects, depicted by the thick dashed line, that strengthen prison gangs at the expense of state authority. The total effect of incarceration on state authority, the vertical sum of individual and collective effects, can thus become negative, implying an inflection point beyond which further incarceration—even if crime-reducing—ultimately erodes the state’s legitimate authority.

2 Background

Lessing (2010) analyzes prison gangs’ growth in terms of three overlapping dynamics: consolidation of control over prison life; propagation throughout a prison system; and projection of coercive power beyond the prison walls. This article focuses on projection and the uses to which prison gangs put it; in this section, I briefly discuss consolidation—and the common state response of segregating incoming prisoners by gang affiliation—and propagation because together they define a prison gang’s “coercive jurisdiction”: the set of outside actors to whom it can credibly promise rewards and punishments. Such promises are the key to projecting power.

Successful gangs consolidate power inside prisons by eliminating or subjugating rivals, taking control of key aspects of prison life (including contraband flows), and securing the capacity to reward and punish inmates. While early stages may witness brutal violence among fledgling groups (e.g. Amorim 1993, 35; Blatchford 2008, 6; Salla 2007, 82) once a gang achieves primacy it can impose rules, indeed a whole social order, that reduce violence among members and sometimes the larger inmate population (Dias 2011). Welfare and public-goods provision can win a gang further prestige and loyalty.2 “Thanks to the PCC,” a São Paulo inmate explained, “The number of deaths fell.... Nobody kills without authorization.... There’s a steady supply of cell phones and drugs.... It hands out cesta básica3 .... Who does all this? Just the PCC” (de Barros 2006, 8). State agents can also benefit: corrupt guards may prefer

---

2Author interviews, four former Comando Vermelho and Terceiro Comando prison-gang members, Rio de Janeiro, August 2009. All interviews and translations of source material are the author’s.

3A standardized basket of food staples provided to the needy by charities or government agencies.
dealing with a single ruling gang, and honest officials can come to rely on prison gangs to maintain order (Biondi 2010, 74; Dilulio 1987, 134; Venkatesh and Levitt 2000, 435).

Conversely, organized violence against rivals has led administrators across cases to segregate inmates by gang affiliation, effectively handing gangs control over their designated wings or units. While this may save lives (and, officials pointed out to me, careers) it clearly reinforces consolidation (Knox 2012). Segregation also requires sorting incoming prisoners by affiliation, often by imperfect proxies like race and gang-turf geography. This has perverse effects, expanding gangs’ ‘coercive jurisdictions’ and putting weakly or un-affiliated first-time offenders under gang custody and tutelage (Skolnick et al. 1990, 24; Human Rights Watch 2004, 33; USAID 2006, 15).

Propagation—the spread of a gang to multiple facilities within a prison system—occurred in the U.S. and Brazil mainly via the transfer of core members to new facilities, often as part of a misguided “diesel therapy” strategy of breaking up gangs and isolating leaders (Boyd 2009, 997; Paixão 1987, 74). A distinct mechanism—deportation and subsequent incarceration—brought mara gangs from the United States where they were founded to Central America’s prisons. In all cases, norms, codes, and sometimes written statutes (Skarbek 2011, 712-714; Jozino 2004, 31) permit a decentralized structure that is robust to the isolation or elimination of individual leaders (Biondi 2010; Lima 1991), one reason that gang-abatement policies have not had decisive impacts (e.g. Cáceres 2009; Dias 2011, 173-174; Fleisher and Decker 2001; Knox 2012).

Consolidation and propagation allow prison gangs to make credible promises to reward or punish inmates throughout entire regional, state, or even national prison systems. The state-provided threat of incarceration then transmits prison gangs’ coercive power to street-level actors in the corresponding jurisdiction, as the model below makes transparent. Note, though, that this is a theory of how already consolidated prison gangs come to project power; initial formation and consolidation seem to depend on first-hand transmission of collectivist techniques through cohabitation (Lessing 2013), so that increased

---

5 Interviews, prison and police officials, Rio de Janeiro, July 8, August 26, 2009.
6 Author’s visit to and interview with the director of Neves Jail, Rio de Janeiro, August 29, 2009.
incarceration alone does not guarantee the emergence of sophisticated prison gangs within prison.

3 Modeling Projection of Power

Why do people on the streets obey the orders of prison-gang leaders who may spend the rest of their lives behind bars? An L.A. Sheriff’s Department sergeant testified that “the Eme controls the prisons and the [street] gangsters know that eventually they’ll end up in prison and be subject to sanctions and retribution if they don’t obey the Eme while they’re on the street” (Rafael 2007, 326). A former drug boss I interviewed in Rio put it even more simply: “Whatever you do on the outside, on the inside you’ll have to answer for it.”

As Skarbek (2011) notes, this logic depends on street-level actors’ belief that future incarceration is likely. Yet this likelihood, and hence the force that this logic carries, varies in accordance with state policy. So too does the severity—in both length and intensity—of imprisonment, which prison gangs ameliorate through protection and other member benefits. If state crackdowns make street-level actors more likely to go to prison, or make imprisonment harsher, those actors should have stronger motivation to stay in good standing with a prison gang that can protect them on the inside. At the same time, the “price” of such good standing may involve taking risky actions that increase actors’ chance of incarceration. Finally, state crackdowns alter the basic Beckerian calculus about whether to be a criminal in the first place: the point of mass-incarceration policies, presumably, is to make non-criminal outside options relatively more attractive and thereby reduce the crime rate as a whole.

To rigorously analyze these interactions, and clarify the conditions under which changes in state policy might strengthen prison gangs’ power on the street, I develop a simple formal model. It focuses on two sets of independent variables: the first concerns the prison gang’s ability to reward cooperation and punish defection, especially to ameliorate the pain of imprisonment; the second set concerns incarceration itself: the “certainty” (likelihood of incarceration) and “severity” (length of sentences, prison conditions, etc.) of punishment. While the first set are treated as parameters (i.e. relatively stable),

7 Author interview, August 17, 2009
the second are directly affected by state policy. I first examine “ideal-type” policies that affect severity and certainty independently, then turn to more realistic scenarios where increases in certainty result, through overcrowding and related channels, in concomitant increases in severity.

I test these policy effects on two outcomes, which we can think of as the intensive margin (the gang’s ability to impose taxes and other burdens on members) and the extensive (recruiting new members). For the intensive margin, the outcome of interest is the largest burden that can be imposed on outside affiliates (such as street gangs and/or drug retailers) as the price of gang membership or good standing.\(^8\) This burden can be a money tax, such as those charged by the Eme and the PCC on their street-level affiliates, but it might also involve taking risky actions, obeying prohibitions or other rules, or a combination of all three. Outside actors weigh these costs against the benefits of membership and the pain of gang retribution if they defect. Comparative statics reveal how carceral policy affects this tradeoff, taking the maximum feasible burden as a measure of prison-gang capacity to project power.

I then focus attention on the extensive margin, recruitment of “entry-level” criminals, possibly for orchestrated violent actions. A key example is the PCC’s recruitment of an impromptu army of non-member youth to throw bombs and set fires during the 2006 terror attacks. The goal was less to tax existing affiliates (who are likely to be established criminals) than to induce multiple non-members (likely low-level criminals or “borderline” individuals in the gangs’ social orbit) to take risky actions through a promise of “prison insurance” when and if incarcerated (Phillips 2006). Recruitment was probably on an “open call”, freelance basis, and individuals were not threatened with retaliation if they declined.\(^9\) This iteration therefore abstracts from taxation and reprisals. The outcome of interest here is the prison gang’s capacity to recruit “better” criminal types; I test the effect of different policies on this outcome as well as the general crime rate. Future work might seek to combine the intensive and extensive margins in a single model, but would likely require better empirical evidence of how gangs think about the trade-off between the two.

---

\(^8\)In many cases, outside actors remain affiliates in good standing, while true membership requires additional steps, often including “baptism” within prison (e.g. Biondi 2010). For the sake of parsimony, I elide this distinction here.

\(^9\)Interview, Head Researcher, São Paulo Office of the Public Prosecutor, August 2006.
Like most models, this one ignores many interesting questions in order to focus attention on a narrow slice of a complex reality. To explore changes in gangs’ capacity to project power on the street, I treat prison-gang power within prison—“consolidation”—exogenously: consolidation can vary (via comparative statics), but the model does not speak to how. The model also focuses on the decisions of street-level affiliates to obey or defy prison-gang edicts once issued, testing the effect of carceral policies on prison gangs’ capacity for projection. It does not formally explore gang leaders’ game-time decisions about when and how to employ that capacity, a fascinating question that would require a far more complex model of the political and strategic environment. Similarly, while the model characterizes tradeoffs states face when choosing policies, it says little about how states do or should evaluate them. Finally, this highly stylized formalization of the very complex decision, often made by vulnerable people, to join or obey violent gangs is intended only as a useful approximation; it surely ignores many important psychological and cultural nuances. All of these issues merit future inquiry; the model aims only to tractably explore the effects of pro-incarceration policies on prison gangs’ capacity to project power.

Previewing the results, increases in severity raise gangs’ power on the street whenever they are consolidated enough within prison to sufficiently mitigate the pain of incarceration. Crackdowns (i.e. increases in certainty) increase prison gangs’ power on the streets unless they are sufficiently “targeted”—falling more heavily on those who follow, as opposed to disobey, gang edicts. Roughly speaking, entirely untargeted crackdowns strengthen gangs that are stronger inside than outside prison, but weaken them otherwise; I suggest this as a useful conceptual distinction between prison- and street gangs. I then consider crackdowns that simultaneously produce overcrowding or otherwise worsen prison conditions; these require an even higher level of targeting to avoid inadvertently strengthening the gang. Finally, addressing the long-standing “certainty vs. severity” debate (e.g. Beccaria 1819; Donohue 2007), the model reveals that “swift and certain” punishments (Kleiman 2009) can offset the gang-strengthening effects of crackdowns with shorter sentences.

To study the extensive margin, I incorporate Becker’s classic model of crime by giving potential recruits a non-criminal outside option, since actors with little or no criminal history might realistically ‘go straight.’ This allows me to distinguish the ‘individual effects’ of policies (on the overall crime rate)
from the ‘collective effects’ (on prison-gang coercive power). Harsher punishment and insufficiently targeted crackdowns are found to simultaneously aid gang recruiting while lowering the expected utility of all criminals, thus reducing overall crime.

### 3.1 Taxation (Intensive Margin)

Because the model is quite straightforward, I present the main results below somewhat informally; the appendix contains more formal versions of the propositions, as well as proofs and derivations.

The players are a gang $G$ (“she”) and an outside criminal actor, say a street-gang leader or autonomous drug dealer, $S$ (“he”). $G$ sets $M$ as the price of membership or good standing. $S$ then chooses whether to comply and carry out $M$ or defect and “go it alone”.¹⁰ Nature ($N$) then decides whether $S$ is imprisoned or not, and payoffs are realized. The payoffs reflect the assumption that $G$ rewards or punishes $S$ depending on whether he has complied or not.¹¹

---

³ MODELING PROJECTION OF POWER

---

¹⁰This setup denies $S$ the option to switch allegiance to a rival prison gang, because this does not appear to be a relevant option in many substantive settings. The Mexican Mafia is the dominant Latino prison gang in southern California, and Latino criminals are unlikely to successfully appeal to rival prison gangs for protection because of racial barriers (Skarbek 2014); strict gang loyalty in El Salvador makes switching from one mara to another basically unthinkable; and the PCC is hegemonic throughout São Paulo. Only in Rio is defection to another gang possible, and even there it has been rare.

¹¹Skarbek (2010) explores the conditions under which prison gangs’ promises to punish and reward are credible. For simplicity, I abstract from this question here and assume they are credible.
probability of imprisonment—his “certainty” of punishment in criminological terms—from $\pi$ (if $S$ defects) to $\tilde{\pi} \geq \pi$. These dimensions are independent: if $M$ is a pure money tax that police cannot easily observe, $\tilde{\pi}$ might be very close to $\pi$, while $\tau$ reflects the size of the tax. Conversely, if $M$ involves risky actions like assassination, we might have $\tilde{\pi} \gg \pi$, but, if $S$ doesn’t mind killing beyond the added risk, $\tau \approx 0$.

Focusing on $S$’s payoffs, let $j > 0$ measure severity of punishment, including both sentence length and harshness of prison conditions. If $S$ is imprisoned and has not complied, he suffers $-j$ in full. If he complied, though, he suffers only $-\frac{j}{\alpha}$ if he complied, where $\alpha > 1$ measures $G$’s capacity to ameliorate the pain of imprisonment. This formalization captures the idea, emphasized by interviewees, that gang protection and support improves total welfare more when sentences are long and conditions harsh. Examples include protecting inmates from violence, providing food and medicine when the administration does not, passing messages to and from prisoners in solitary, and aiding family members while incarcerated.

Let $y$ represent baseline profits from illicit activity that $S$ earns by “going it alone”. Since membership may permit efficiency gains from criminal cooperation, a collaborator receives $\beta y$ where $\beta \geq 1$. Assume that the gang punishes defectors, inflicting a cost of $\gamma \geq 0$ within prison and $\delta \geq 0$ on the outside. Finally, since $S$ represents a street-gang leader facing a risk of incarceration regardless of compliance with $G$, I do not include a non-criminal, “go straight” option here.

---

12The results do not depend on $\tilde{\pi} \geq \pi$; indeed, prison-gang membership might lower the risk of incarceration ($\tilde{\pi} < \pi$), perhaps by reducing gang violence or because prison gangs have connections with corrupt police or judges. Assuming so, however, ‘stacks the deck’ in favor of prison-gang projection of power, so I focus on the $\tilde{\pi} \geq \pi$ case.

13Interviews, four former Comando Vermelho and Terceiro Comando prison-gang members, Rio de Janeiro, August 2009; Directors of Guatemalan and Salvadoran Prison Systems, May 21-22, 2013. Modeling $\alpha$ as additive, so that jailed member suffer $-(j-\alpha)$, would imply, implausibly, that inmates get the same total relief from gang membership regardless of sentence length.

14Recruitment of low-level criminals and “borderline” individuals, where “going straight” is a more realistic option, is considered below. Dedicated criminals rarely go straight unless the state offers a combination of reduced sentences (lower $j$), amnesty (lower $\pi$), and protection from reprisals (lower $\gamma$ and $\delta$). Exploring the effects of such offers on prison-gang projection is an important avenue for further research.
This specification remains agnostic about what kind of gang $G$ is: street gangs may also try to tax and coerce street-level actors, while offering both benefits to members and credibly threatening punishment for disobedience. Indeed, scholars have found origin to be a decreasingly useful criterion for distinguishing gangs (e.g. Hunt et al. 1993) as members of gangs that formed on the street are incarcerated (yielding power on the inside: $\alpha$ and $\gamma$) and members of gangs that formed in prison are released but remain loyal (yielding power on the outside: $\beta$ and $\delta$). The model points to an alternative criterion: pro-incarceration policies have opposite effects on a gang’s capacity to project power depending on its relative power inside vs. outside prison, suggesting a novel conceptual distinction between street gangs and prison gangs.

The outcome of interest is the maximum burden or “utility tax” that $G$ can impose for any given risky action without provoking defection, $\tau^*(\tilde{\pi})$. To simplify exposition, I analyze $\tau^*$ directly; this is equivalent to assuming that $G$ wants to charge the highest tax possible.\footnote{Appendix A specifies such a utility function for $G$ and an equilibrium in which $G$ plays $\tau^*$ and $S$ complies. Of course, real-world gangs frequently punish defectors. An information asymmetry, say a distribution over types of $S$, can generate this kind of result, with the leadership choosing a $\tau^*$ which $S$ rejects with positive probability. None of the substantive findings would be affected by such a modification.} In practice, $G$ might have good reason for charging less, such as building morale or fending off competition. Even then, $\tau^*$ would capture how much “slack” $G$ has, i.e. how willingly or happily $S$ complies. In any case, since the focus here is on gangs’ capacity to project, not necessarily when and how they employ that capacity, analyzing the maximal imposable burden—and abstracting from $G$’s choice of an optimal tax rate—makes sense.

However, that burden ($M$) involves a mix of additional risk ($\tilde{\pi}$) and disutility ($\tau$); why focus only on $\tau$? Because, from $S$’s perspective, these are the only downsides of compliance, so there is a direct tradeoff between the two.

\textbf{Lemma 1.} \textit{The higher the incarceration risk of complying ($\tilde{\pi}$), the lower the maximum feasible tax ($\tau^*$), and vice versa: the higher $\tau$, the lower the maximum feasible risk of complying ($\tilde{\pi}^*$).}

\textit{If complying brings sufficiently little or no additional risk (i.e. if $\tilde{\pi}$ is very close to or less than $\pi$) then...}
Lemma 1 implies that we could fix either \( \tau \) or \( \bar{\pi} \), then solve for the maximum feasible value of the other. Since \( \bar{\pi} \) is probably discrete and bounded\(^{17}\) I focus on \( \tau^*(\bar{\pi}) \); this represents, for a given risky action, the maximum additional tax \( G \) could charge (or, if \( \tau^* \) is negative, how much \( G \) would have to pay \( S \) to induce compliance). Comparative statics on \( \tau^* \) thus reveal how changes in gang characteristics and state policy affect gangs’ coercive power over outside members.

In this simple setup, the relative payoff to complying is increasing in all aspects of the gang’s capacity to reward and punish; thus increases in \( \alpha, \beta, \gamma \) or \( \delta \) will all raise \( \tau^* \). This is a source of positive feedback: if the gang uses \( \tau^* \) to increase any of these parameters, its future coercive power will be even greater. Exogenous increases in outside profits (\( y \)) will also raise \( \tau^* \) whenever there is sufficient surplus from collective criminal activity relative to the increased risk of imprisonment it entails \( \left( \beta > \frac{1 - \pi}{1 - \bar{\pi}} \right) \).

As for the effects of carceral policy, it would be specious to assume that states directly and independently set \( j, \pi, \) and \( \bar{\pi} \); rather, I assume that policies affect these parameters jointly. Table 1 classifies policies into four broad types, based on these joint effects on severity (\( j' \)) and certainty (\( \pi' \) and \( \bar{\pi}' \)), which I test in turn.

**Table 1.** Policy interventions to be tested, defined by their effects on severity and certainty of incarceration.

<table>
<thead>
<tr>
<th>Policy Type</th>
<th>Effect on Severity: ( j' )</th>
<th>Effect on Certainty: ( \pi' ) and ( \bar{\pi}' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pure “Harshening”</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Pure Crackdowns</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Crackdowns w. Overcrowding</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>“Swift &amp; Certain” Crackdowns w. Offseting</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Pure “harshening” policies—including longer sentences, solitary confinement, reduced privileges, or worsening conditions—increase the pain of prison (\( j \)) but have no affect on \( S \)’s chances of incarceration. Crackdown-type policies—including anti-gang laws, increased arrests and prosecutions, and

\(^{16}\)Formal versions of all lemmata and propositions, as well as proofs, appear in Appendix A.

\(^{17}\)See Appendix A for further discussion.
mandatory sentencing—increase the likelihood of imprisonment. I first test “pure” crackdowns that, unrealistically, have no collateral effect on severity \((j' = 0)\); I then relax this assumption and test crackdowns that, perhaps inadvertently, result in harsher conditions through overcrowding, longer sentences, and related channels \((j' > 0)\). Finally, I consider “swift and certain” policies that offset increases in certainty with reductions in severity \((j' < 0)\), usually through shorter sentences. Comparative statics for each policy type are presented as parts of Proposition 1; the last three policies are represented graphically in Figure 3.

**Proposition 1a (Increases in severity strengthen consolidated gangs).** Pure harshening policies (i.e. with no effect on certainty) will increase \(\tau^*\) whenever:

\[
\alpha > \frac{\pi'}{\pi} \tag{Condition C}
\]

In words, pure harshening policies strengthen gangs that are sufficiently consolidated within prison. Condition C pins down “sufficiently consolidated”, and has a natural interpretation. Complying makes an outside actor \(\frac{\pi'}{\pi}\) times likelier to be imprisoned; call this the risk differential. Whenever \(G\) can ameliorate the pain of prison by a factor larger than the risk differential, harsher punishment will increase its leverage over outside actors. If the price of membership involves little additional risk, as with a tax payment difficult for police to observe, then \(\frac{\pi'}{\pi} \approx 1\) and Condition C is easily satisfied; greater risk requires greater consolidation. As I show in the following section, prison gangs were quite consolidated when projection began in earnest, so I assume Condition C holds for the remainder of the analysis.

For crackdowns, a critical issue is how well targeted they are, specifically, what share of increased arrests falls on street-level actors who comply with gang edicts vs. those who defect. Formally, define a policy’s **degree of targeting** as its effect on \(\pi'\) relative to \(\pi\): \(\frac{\pi'}{\pi'}\). If the state could perfectly focus its additional repressive force solely on those street-level actors who comply with prison-gang orders, so that \(\pi' > 0\) and \(\pi' = 0\), then the degree of targeting would go to infinity. At the other extreme, a

\[\text{Pure harshening policies do not affect certainty at all, so } \pi' = \pi' = 0 \text{ and their degree of targeting is undefined: } \frac{0}{0}.\]
Figure 3. Effects of pro-incarceration policies (crackdowns) on gang’s capacity for projection. Shaded regions represent policies that increase the maximum tax the gang can impose without provoking defection ($\tau^*$).

blanket crackdown on street-level actors in general, making no attempt to distinguish those who obey gang edicts from those who defect, then $\pi' = \pi'$ and the degree of targeting $\frac{\pi'}{\pi}$ equals 1. Of course, the state could deliberately target defectors over compliers, so that $\pi' < \pi'$ and $\frac{\pi'}{\pi} < 1$; such a policy would naturally encourage compliance, stacking the deck in favor of my core claim, so I rule it out.

Figure 3 plots the space of possible crackdowns. Degree of targeting ($\frac{\pi'}{\pi}$) is on the vertical axis: higher values represent better-targeted crackdowns, and the origin is drawn at $\frac{\pi'}{\pi} = 1$, an untargeted crackdown. The horizontal dimension represents crackdowns’ effect on severity, $j'$. At the origin, crackdowns are “pure”, with no effect on prison conditions or sentencing. To the right are crackdowns that, perhaps unintentionally, also increase severity via overcrowding and similar mechanisms. To the left are policies that offset increased certainty with deliberate reductions in severity. The shaded region corresponds to those policies that raise $\tau^*$, i.e. “strengthen the gang” while policies in the non-shaded region lower $\tau^*$, or “weaken the gang”.

Empirically, many anti-gang sweeps poorly discriminate street-gang members from non-members, much less their compliance with prison-gang edicts. With that in mind, consider first the limiting case of a crackdown that is totally untargeted ($\pi' = \pi'$ and thus $\frac{\pi'}{\pi} = 1$) and is also “pure” in the sense that
it has no effect on severity \((j' = 0)\), shown as point \((a)\) in Figure 3.

**Proposition 1b (Pure, untargeted crackdowns strengthen “prison gangs”).** *Any crackdown with no effect on \(j\) that raises \(\pi\) and \(\tilde{\pi}\) equally will increase \(\tau^\ast\) if \(j - \frac{1}{\alpha} - (\beta y - y) + \gamma - \delta > 0\).*

In words, untargeted pure crackdowns increase \(G\)’s coercive power whenever it is stronger within prison than on the street. Intuitively, if the membership benefits deriving from “prison insurance” \((j - \frac{1}{\alpha})\) are greater than those from improved criminal profits \((\beta y - y)\), and if retribution is more likely inside prison than outside \((\gamma > \delta)\), then a higher chance of incarceration makes membership more valuable.

Proposition 1b suggests a useful criterion for differentiating prison- and street gangs: whether their total power to punish and reward is greater inside than outside prison. In other words, a prison-based criminal network can be defined as a gang whose outside coercive power is increased by an untargeted crackdown, like point \((a)\) in Figure 3. By this criterion, a street gang would be one whose capacity to impose burdens on members is weakened by a pure, untargeted crackdown, with the shaded region lying entirely below point \((a)\).

For a prison gang thus defined, an untargeted crackdown increases coercive power on the street. How targeted would a crackdown have to be to avoid strengthening the gang? We can define a *critical degree of targeting*, (denoted by \(\varphi^\ast\)) such that any crackdown whose degree of targeting is below this threshold, i.e.

\[
\frac{\tilde{\pi}'}{\pi'} < \varphi^\ast \quad \text{(Condition } T)\]

will strengthen the gang. Algebra reveals that the critical level of targeting for a crackdown \((\varphi^\ast)\) is a function of the model’s parameters as well as the crackdown’s effect on severity \((j')\). This can be seen in Figure 3, where \(\varphi^\ast\) appears as an upward sloping line.\(^1\)

\(^1\)If Condition \(C\) did not hold, this line would slope downward.
Proposition 1c (Better-consolidated gangs require more targeting). The critical level of targeting for crackdowns with no effect on severity is increasing in the “inside” parameters’ $\alpha$ and $\gamma$, and decreasing in the “outside” parameters $\beta$, $\delta$ and $\eta$.

The model’s main result is that insufficiently targeted crackdowns increase gangs’ coercive power. Condition $T$ and Proposition 1c pin down “insufficiently targeted”, revealing that the larger a gang’s capacity to reward and punish inside vs. outside prison, the more targeted crackdowns must be to avoid strengthening it. Graphically, the more consolidated a gang is within prison, the higher point $(b)$ is relative to $(a)$.

Thus far, the analysis has assumed that crackdowns only increase certainty. In reality, policies that increase incarceration rates are likely to increase severity as well, perhaps inadvertently, because of overcrowding or by introducing violent individuals into the prison system. With this in mind, consider the right half of Figure 3, where $j$ rises along with $\pi$ and $\tilde{\pi}$. Say the state implements what it thinks is a ‘just-sufficiently targeted’ pure crackdown, unaware of these unintended effects on severity. The state believes it is at point $(b)$, but the overcrowding effect means it is really at $(c)$, below $\varphi^*$ and inside the gang-strengthening region, in spite of the targeted nature of the crackdown. In general, the critical level of targeting is increasing in the size of this “overcrowding effect”:

Proposition 1d (Overcrowding increases the critical level of targeting). Say a pure crackdown with no effect on severity is just-sufficiently targeted, so that $\frac{\pi}{\pi'} = \varphi^*$. Any crackdown with the same degree of targeting but with a positive effect on severity (through, say, overcrowding) will increase $\tau^*$.

Finally, consider the left-hand side of Figure 3, which represents deliberate policies of offsetting increased likelihood of prison with less severe sentences. Kleiman (2009) offers important behavioral and game-theoretic arguments why ‘swift and certain’ punishment is a better deterrent than the status quo of long sentences for a small fraction of offenders. The present model suggests another advantage: if Condition $C$ holds and the gang is strengthened by increases in $j$, then a decrease in severity weakens the gang. For example, if outside actors face shorter sentences, the value of gang membership conditional
on incarceration is lower.\textsuperscript{20} This can be used to offset the gang-strengthening effects of an insufficiently targeted crackdown.

**Proposition 1e (Offsetting more certainty with less severity).** *Say a pure crackdown with no effect on severity is insufficiently targeted, so that \( \bar{\pi}' < \phi^* \). Then there exists \( k > 0 \) such that a crackdown with the same degree of targeting but \( j' < -k \) lowers \( \tau^* \). \( k \) is increasing in \( \pi' \) and decreasing in \( \bar{\pi}' \). \( k \) is decreasing in \( \alpha \) if*

\[
\frac{\bar{\pi}'}{\pi'} < \frac{j + y + \gamma - \delta}{j \bar{\pi} + \beta y}
\]  

\((C_K)\)

The term \( k \) indicates how large a reduction in severity is needed to offset an insufficiently targeted crackdown; in Figure 3, it is the distance from \((d)\) to \((e)\). Holding the impact on non-members \((\pi')\) constant, better-targeted crackdowns require smaller offsets.\textsuperscript{21} As for \( \alpha \), the more consolidated the gang, the larger the targeting gap, but, counterintuitively, the more efficacious any reduction in severity. Condition \( C_K \) holds whenever the latter effect predominates. Graphically, a rise in \( \alpha \) increases the slope of \( \phi^* \), reducing \( k \) for points below \( C_K \). Overall, the result suggests that Kleiman's approach is particularly apt when prison gangs are strong and targeting is difficult.

### 3.2 Recruitment (Intensive Margin)

I now turn to the question of recruiting outside actors to take risky actions. The street-level actor \( S \) is now replaced by a continuum of potential recruits, indexed by their expected income from “go-it-alone” criminal activity \([y, \bar{y}]\); in this setting, \( y \) measures individuals’ “criminal talent”. I assume that \( G \) is interested in recruiting low-level foot soldiers for organized criminal actions, and focus on its capacity to recruit ‘better’ criminal types (i.e. higher \( y \)). As with the PCC’s reliance on non-member youth to carry out terror attacks in 2006,\textsuperscript{22} I assume recruits receive only “prison insurance”, but no

\textsuperscript{20} Shorter sentences could also weaken gangs by giving them less time to learn incoming inmates’ type, or to socialize new recruits (e.g. Biondi 2010, 98), channels beyond the scope of this model.

\textsuperscript{21} But note that two crackdowns with the same \( \bar{\pi}'/\bar{\pi} \) can have different values for \( k \). The slope of \( \phi^* \) in Figure 3 depends on \( \pi' \).

\textsuperscript{22} Interview, Head Researcher, São Paulo Office of the Public Prosecutor, August 2006.
cash or additional criminal rents: $\tau = 0$ and $\beta = 1$. Also, since there are many potential recruits, $G$ does not punish defection, which here just means non-recruitment: $\gamma = \delta = 0$. Thus the cost to an individual recruit of compliance ($C$) is simply an increase in the risk of incarceration from $\pi$ to $\tilde{\pi}$; in exchange, $G$ provides in-prison benefits, captured by $\alpha$. As above, I treat $\tilde{\pi}$ as exogenous, and show that requiring more risky actions result in poorer recruitment. Finally, since potential recruits have relatively low criminal profiles, assume all $y$ have an outside option to ‘go straight’ ($O$) worth $y_o$, with $\pi_o = 0$. The choice between non-gang crime ($D$) and legality ($O$) is thus equivalent to Glaeser’s (1999) simplified version of Becker’s (1968) canonical model of crime.

An individual $y_i$ is recrutable if she prefers $C$ to both $D$ and $O$. Formally, define the relevant cutpoints as:

$$y^* : y_i < y^* \iff C > D; \quad y^C : y_i < y^C \iff O > C; \quad y^D : y_i < y^D \iff O > D$$

For $y^*$ to be positive, and hence for anyone to be recrutable, Condition $C$ must hold: $\alpha > \frac{\pi}{\pi_o}$. Since the risk differential $\tilde{\pi}$ here is substantial, Condition $C$ is non-trivial: recruitment via prison-insurance is only viable for well-consolidated gangs able to provide significant in-prison benefits.

\begin{align*}
&M = \text{Action(s) required for membership} \\
&\tilde{\pi} = \text{Risk of jail if carry out } M \\
&\pi = \text{Baseline risk of jail} \\
&j \in \mathbb{R}^+ = \text{Pain of incarceration} \\
&y_o \in \mathbb{R}^+ = \text{Expected licit income for all } i \\
&y_i \in [\underline{y}, \overline{y}] = i’s \text{ expected criminal income} \\
&\alpha \in [1, \infty) = \text{Inside reward (“prison insurance“)}
\end{align*}

Figure 4. Game Tree: Recruitment (Extensive Margin)

I make two further assumptions. First, since all three actions are observed empirically, I focus on regions of the parameter space where each is taken along some portion of the interval $[\underline{y}, \overline{y}]$. Algebra reveals that $y^* > y^C \iff y^C < y^D$, so a sufficient and necessary condition for all three actions being taken is $\underline{y} < y^D < y^* < \overline{y}$. Thus $y^*$ represents the highest recrutable type, offering a measure of

\begin{footnote}
Glaeser has a distribution over $y$’s legal wages and a constant criminal wage; I reverse this, permitting assessment of the quality of criminals the gang can recruit. Comparative statics of policies on the crime rate are unaffected.
\end{footnote}
G’s ability to recruit.\(^{24}\) Second, I assume that \(G\) does not recruit low types (below \(y^D\)) who would not otherwise be involved in crime.\(^{25}\) This implies that \(y^D\) measures, inversely, overall participation in criminal activity, i.e. the crime rate, as in Becker/Glaeser. Thus comparative statics on \(y^D\) and \(y^*\) capture, respectively and roughly, the individual vs. collective (i.e. crime-rate vs. gang-strengthening) effects of carceral policy.

To analyze the model, note first that gang control over prison life aids recruiting (\(\frac{\partial y^*}{\partial x} > 0\)) but has no effect on overall crime (\(\frac{\partial y^D}{\partial x} = 0\)), while increases in the outside option reduce the crime rate (\(\frac{\partial y^D}{\partial y_*} > 0\)) but have no effect on recruiting strength (\(\frac{\partial y^*}{\partial y_*} = 0\)), a consequence of the assumption that \(G\) does not recruit the very lowest criminal types. Turning to the effects of policy interventions in Table 1:

**Proposition 2a (Increased severity reduces crime but aids recruitment).** *All pure harshening policies increase \(y^D\) and, assuming Condition C holds, raise \(y^*\).*

Harsher sentences unambiguously make criminals worse off with respect to non-criminals, lowering the overall crime rate; at the same time, they increase the gang’s ability to recruit as long as it is sufficiently consolidated to make prison-insurance recruitment viable.

**Proposition 2b (Crackdowns reduce crime).** *Any policy with \(\pi' > 0\) raises \(y^D\).*

As long as crackdowns are not exclusively targeted at gang collaborators, they will always reduce overall crime. However, they also aid recruiting if not sufficiently targeted:

**Proposition 2c (Untargeted ‘pure’ crackdowns always aid gang recruitment).** *For crackdowns \(\rho_C^P\) with no effect on severity, the critical level of targeting is \(\phi^* = \pi^*\).*

\(^{24}\)For expositional simplicity, I focus on \(y\)’s choices in a decision-theoretic framework. As above, \(y^*\) as equilibrium outcome can be easily generated by assuming that \(G\)’s utility is increasing in \(y\).

\(^{25}\)Formally, I assume that the volume of people \(G\) seeks to recruit, \(N_R\), is small compared to the total pool of recruitable actors: \(N_R < \int_{y_D}^{y_D^*} F(\cdot) \, dy\). This is sounder than the alternative assumptions \(\int_{y_D}^{y_D^*} F(\cdot) \, dy < N_R < \int_{y_C^*}^{y_C^*} F(\cdot) \, dy\) or \(\int_{y_C^*}^{y_C^*} F(\cdot) \, dy < N_R\), which imply that prison gangs recruit actors who would otherwise ‘go straight’. Theoretically, it is unclear how \(G\) could observe such actors’ \(y_i\); empirically, recruits seem to have some criminal experience. Most results hold under alternative assumptions, though their interpretation is less clear.
Since $\frac{\pi}{\pi'} > 1$, it immediately follows that any untargeted policy (such that $\frac{\pi'}{\pi} = 1$) with no effect on severity will aid recruiting. Even a targeted crackdown, if it affects the risk differential by a factor less than the differential itself, increases $G$’s ability to recruit. This suggests diminishing returns: targeting raises the risk differential, which in turn makes further crackdowns more likely to be counterproductive.

As for overcrowding, Proposition 1d holds here as well: a seemingly sufficiently targeted policy, if it inadvertently worsens prison conditions, will still strengthen gangs. However, overcrowding will also increase the crackdown’s crime-reducing effect, since $\frac{\partial y}{\partial j} > 0$. Also as before, the gang-strengthening effect of an insufficiently targeted crackdown can be offset with a reduction in severity:

**Proposition 2d (Offsetting ‘more certainty’ with ‘less severity’).** Say a ‘pure’ crackdown with no effect on severity is insufficiently targeted, so that $\frac{\pi'}{\pi} < \varphi^*$. Then there exists $k \in \mathbb{R}^+$ such that a crackdown with the same degree of targeting but $j' < -k$ lowers $y^*$. $k$ is increasing in $\pi'$ and decreasing in $\pi$, and decreasing in $\alpha$.

Again, better targeting requires smaller offsetting reductions in severity. In this case, increases in $\alpha$ do not affect the “targeting gap”, and so unambiguously reduce the size of offset needed. Such an offset will necessarily reduce deterrence, but crime will still fall if the targeting gap is not too large, or, ironically, if $\alpha$ is high enough. This suggests that when prison gangs are very strong, taking a “swift and certain” approach to punishment is a particularly appropriate strategy.

### 4 Empirical Evidence

The model predicts that prison gangs’ capacity to project power grows under the following conditions:

- **C:** Prison gangs are sufficiently consolidated to effectively reward compliance and punish defection within prison;

- **P:** Pro-incarceration policies increase the certainty (i.e. likelihood of incarceration) and / or severity of punishment (i.e. length and harshness of sentences, overcrowding, etc.);

---

26The caveat of note 21 still applies, however.
**T:** Additional arrests are insufficiently targeted at those street-level actors who comply with prison-gang edicts vs. those who defect.\(^{27}\)

I check this prediction against three substantively important and well-documented positive cases—California (roughly 1980-2002), São Paulo (1990-2012), and El Salvador (1990-2012). I show that Conditions \(C\), \(P\), and \(T\) were met, and that these periods indeed witnessed important expansions in prison-gang projection of power onto the streets. I focus on positive cases because rarely can one distinguish the true absence of prison gangs from a lack of data, particularly since gangs and authorities alike may seek to hide or downplay gang power. Nonetheless, I do briefly explore one well-documented negative case, Nicaragua. I then address these case-selection issues and related obstacles to strong causal inference and larger-\(n\) analyses. Improving data for more systematic testing is a critical avenue for further research.

**Condition \(C\): Prison gangs could reward and punish within prison.** In southern California’s prisons, the Eme first rose to prominence in the 1960s, and by the 1980s had significant ability to punish and reward inmates (Diaz 2009, 128-130; Skarbek 2011). São Paulo’s PCC, founded in 1993, rapidly consolidated and propagated through the state prison system in the following years; by 2001 it was powerful enough to instigate its first “mega-rebellion”: simultaneous mutinies in 30 prisons, involving 28,000 prisoners (Dias and Salla 2013, 397). In El Salvador, the arrival in the 1990s of *mara* members deported mostly from California—where they were themselves taxed and governed by the Eme prison gang\(^ {28}\) (Valdez 2011, 28-29)—turned a vast collection of small, local youth gangs into *clikas* (‘cliques’) of the prison-based MS and M18 *mara* franchises (Cruz 2010). By the early 2000s, officials were segregating prisons by gang (Cruz 2010, 391), giving each *mara* a consolidated base of operations.

**Condition \(P\): Certainty and severity of punishment increased.** In all three cases, mass incarceration policies produced, over roughly two decades, a steady increase in incarceration rates from about 100 per 100,000 residents to over 400 (Figure 5). These periods of “carceral hyperinflation” (Wacquant

---

\(^{27}\)Conditions \(C\) and \(T\) are stated formally in Section 3; Condition \(P\) is implicit in the analysis.

\(^{28}\)Interview, deported former Barrio 18 leader, San Salvador, July 7, 2015.
2009, 120) saw explosions in prison populations and severe overcrowding. Recidivism rates indicate that outside criminal actors anticipate future incarceration. Severity of punishment also increased, due to longer sentences (Zimring et al. 2001), harsher custodial practices such as U.S. Supermax prisons and Brazil’s Special Disciplinary Regimes (Caldeira 2004; Salla 2007), and general worsening of conditions due to overcrowding.

<table>
<thead>
<tr>
<th>n-Fold Increase in Prison Population</th>
<th>6.50</th>
<th>8.11</th>
<th>5.05</th>
<th>2.62</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inmates per Bed</td>
<td>1.96</td>
<td>1.87</td>
<td>3.25</td>
<td>1.28</td>
</tr>
<tr>
<td>[Year]</td>
<td>[2007]</td>
<td>[2012]</td>
<td>[2012]</td>
<td>[2010]</td>
</tr>
<tr>
<td>Recidivism</td>
<td>67.5%</td>
<td>70%</td>
<td>&gt; 50%</td>
<td>&gt; 90% (maras)</td>
</tr>
</tbody>
</table>

Figure 5. Incarceration rates and key events that revealed prison gangs’ capacity to project power. Sources: see Appendix B.

**Condition T**: The degree of targeting was low. Carceral expansion was driven in part by intensified law enforcement against drug trafficking and other gang-related activities. In São Paulo, such efforts were not targeted at criminal organizations, but rather on peripheral areas generally, while inequities in the criminal justice system guaranteed that incarceration was far more likely for poorer defendants, even for relatively minor crimes (IBA 2010, 8-10). California’s Street Terrorism Enforcement and Pre-

---

Pfaff (2011, 514) finds no increase in median sentences in the 1990s, but notes that for California his results “understate time served by the young and by the Hispanic” as well as violent offenders.
vention (STEP) Act and El Salvador’s Mano Dura were harsh anti-gang initiatives that, studies revealed, consistently failed to distinguish street-gang members from non-members (e.g. Cruz 2010, 390-391; Gilmore 2007, 214-225; Van Hofwegen 2009). As such, these initiatives almost certainly failed to specifically target those street-gang members who obeyed prison-gang dictates.

**Predicted Result: Prison Gangs’ Capacity to Project Power Increased.** Directly measuring projective capacity is difficult, because it is not always employed, and when it is, prison gangs and street affiliates may conceal their activities and contacts. However, when it is employed, the resulting revelatory events—including abrupt changes in the structure of street-level criminal markets, establishment of control over peripheral areas, and orchestrated violence / protest—provide reliable observations of prison gangs’ accumulation of power over street-level agents. Figure 5 shows, for each case, one early and salient example:

- September 1993: Mexican Mafia leaders convoke mass public gatherings of California’s Sureño street gangs to impose the “Eme Edict”, a system of “complete vertical integration” involving a ban on drive-by shootings, loyalty to the Eme, and a tax on drug profits (Rafael 2007, 36-39).30

- May 2006: São Paulo, the world’s third-largest city, is held hostage when the PCC instigates synchronized riots in some 90 prisons and street-level attacks on hundreds of civilian and police targets, bringing the city to a standstill for days, until authorities made concessions.

- September 2010: Imprisoned leaders of El Salvador’s MS-13 and M-18 maras join forces to induce—via threats of mass violence by street-level affiliates against city busses—a transportation strike that shuts down the capital for three days, demanding improved prison conditions and the veto of an anti-gang law (Wolf 2012, 86).

These events constitute “smoking-guns”: they could not have occurred if prison gangs lacked the projective power to induce outside affiliates to pay taxes, submit to dictates, and take highly risky actions.31 As measurements, they are likely lower bounds, since gangs need not have “maxed out” their

---


31 Smoking-gun tests are sufficient but not necessary conditions for inference; thus the absence of revelatory events cannot be taken as evidence that prison gangs lack power on the streets.
accumulated projective power; gangs may also have began projecting power in smaller, less ostensible ways earlier. Moreover, these were not one-offs: subsequent events, including the imposition of violence-reducing strictures in São Paulo and El Salvador discussed below, show that projective capacity held steady or grew. The empirical record is thus consistent with the theoretical model and my core claim.

Reverse causality, however, is an obvious concern: perhaps increased projection of power causes the rise in incarceration rates, not vice versa. Two considerations suggest that this is not the case. First, increases in incarceration rates began many years prior to the first clear observations of projection shown above. True, some low-level projection, such as piecemeal criminal governance may have occurred earlier. However, as discussed below, criminal governance generally reduces crime, at least crimes that are likely to lead to increased police repression and, hence, incarceration. In any case, these early periods did not witness marked increases in crime. On the contrary, in Sao Paulo, violent crime was falling throughout the 2000s, while in El Salvador the period of sharpest increase in incarceration (2005-09) saw homicide rates fluctuate and other crime rates drop (UNDP 2009, 72). For reverse causality to seriously threaten my claim, it would need to be the case that projection both went unobserved for years and caused increased incarceration without dramatically affecting crime rates. This is less plausible than my claim: that incarceration rates rose for political reasons and, over several years, caused the increase in gangs’ capacity to project.32

Further comparative evidence comes from Nicaragua, which shares several factors often blamed for the rise of the maras: a history of civil war, easy availability of firearms, widespread poverty and unemployment, and a long-standing presence of neighborhood gangs. Yet the maras made no inroads into Nicaragua, its native gangs never developed into prison-based criminal networks, and its homicide rate remains far lower than its northern neighbors (Cruz 2011; Yashar 2012). Certainly, one important factor was the relative lack of returning mara deportees from the United States. Cruz, however, argues that this explanation is easily “oversold” (2011, 139), drawing attention to equally dramatic differences

32That the early revelatory events shown above all occurred as incarceration rates reached the 350-400 / 100,000 range is suggestive, though hardly conclusive given the small number of cases.
in anti-gang policies. Nicaraguan officials, many of them former insurgents, pioneered a preventive approach (Rocha 2010, 33) that sought to de-stigmatize non-violent youth gangs rather than criminalize them. Unlike El Salvador’s untargeted mano dura sweeps, Nicaragua first categorized groups, then directed repression only at the more serious organized crime outfits, while bringing vulnerable youth into community programs (Cruz 2011, 140-143). This avoided large increases in incarceration, and suggests that repression was also better targeted (in the sense of Condition T). Comparing case studies of Nicaragua with El Salvador, Guatemala, and Honduras, Cruz (2011, 155) concludes that “the mechanism that perhaps most facilitated gang organization and recruitment” in the latter three “was the simultaneous incarceration of thousands of youth gang members and wannabes.”

Unfortunately, such comparative leverage is rare, because negative cases cannot generally be distinguished from missing data. For most countries, including high-incarceration authoritarian regimes like Russia and China, there is little reliable information about prison gangs one way or the other. Absence of evidence of prison-gang power is not evidence of its absence, however, because both gangs and authorities have incentives to deny or downplay such power, and even honest measurement is likely to be biased downward. Even in the United States, prison-gang secrecy and official “gang denial” severely hamper detection and assessment (Fleisher and Decker 2001, 3; Fong and Buentello 1991, 66-7; Knox 2012); in Latin America, reliable information is even rarer (Macaulay 2007, 630).

Under these conditions, naïve empirical testing can generate spurious results, as repeated intelligence failures surrounding the PCC illustrate. Officials denied the PCC’s existence prior to the 2001 “mega-rebellion” (Salla 2007, 81), then only a year later declared that crackdowns and harsher prison policies had rendered it “a failed and dismantled organization” (Simas Filho and Rodrigues 2003, 2). Even specialists who acknowledged the PCC’s strength within prison doubted its power on the streets (Adorno and Salla 2007, 9). In reality, between 2002 and 2006, while São Paulo’s prison population doubled, the PCC consolidated internal control (Dias 2011), and built a network of outside cells whose efficacy the 2006 attacks soon revealed. Yet on the eve of the attacks, the best available sources would have reported that prison-gang capacity for projection was very low; the resulting lack of correlation

---

33 Interview, Head Researcher, São Paulo Office of the Public Prosecutor, May 9, 2005.
between these (erroneous) measures and sharply rising incarceration would appear as strong evidence
that incarceration has zero or even negative effect on prison-gang projection of power. Such a finding
would be proven entirely spurious the following day, once PCC street attacks had brought São Paulo to
a halt, and the positive correlation between incarceration and prison-gang projection became apparent.

To avoid the extreme measurement error and spurious correlation that can occur when zeros cannot
be distinguished from missing data, I focus here on clearly positive cases. The tradeoff is a reduction in
inferential leverage: the available evidence demonstrates the plausibility of the model’s prediction that
mass incarceration strengthens prison gangs, but can make no claims to strong causal identification.
More rigorous testing will require reliable, comparable assessments of street-level prison-gang power
(or lack thereof) over space and time; I discuss potential steps forward in the conclusion. Meanwhile,
in such data-poor environments, the “modeling dialogue” (Myerson 1992, 64) is especially helpful in
clarifying concepts and focusing empirical research on relevant phenomena. In this case, as Proposition
2 showed, crime rates are no measure of prison-gang power, since the same policies that reduce crime
may strengthen prison-gang power on the street.

5 Implications: The Possibility of Counterproductive Punishment

If prison gangs used projection of power only to predate on street gangs, then increased incarceration
might merely raise prison gangs’ relative criminal income. However, contemporary prison gangs use
projection of power in ways that are problematic for state authority, even if they sometimes reduce crime
rates. This poses a double threat to inferences about the effects of increased incarceration. First, we may
inadvertently attribute some reduction in crime to individual effects (incapacitation and deterrence)
that was in fact due to increased organization of criminal activity by prison gangs. Second, we may be
implicitly equating crime reduction with improved rule of law and state authority, failing to account
for the slow, relatively hidden expansion of criminal authority at the expense of the state’s.

This section addresses these two concerns. First, I categorize the uses to which prison gangs have
put projection of power. Then, I conceptually distinguish their effects on crime (sometimes positive,
sometimes negative) from their effects on authority (usually negative). Finally, I draw these strands together in the framework of Figure 1, conjecturing that beyond some point, state efforts to bolster the rule of law through increased incarceration become counterproductive.

5.1 Uses of Projection of Power

The case of Rio’s Comando Vermelho (CV) is useful for theory-building but not for testing: incarceration data is not available for 1970-90, when it first took control of Rio’s prisons, then projected outward to dominate the city’s *favelas* and retail drug trade. I discuss three key uses of projection practiced by the CV, then show how the Eme, PCC, and Salvadoran *maras* have engaged in similar actions.

**Organization of local criminal activity:** Local illicit markets, especially retail drug markets, tend to be fragmented and unstable. Street gangs and small operators rarely establish thoroughgoing control beyond small pieces of home turf (Hagedorn 1994; Skolnick et al. 1990), despite significant investments in arms and soldiers (Levitt and Venkatesh 2000). Yet as Skarbek (2011) has shown, there is a potential surplus to be extracted by any group capable of providing criminal governance. Rio de Janeiro’s CV used a code of mutual-aid among its members to systematically oust or subdue incumbent drug retailers from a majority of the city’s favelas in the 1980s (Amorim 1993; Lima 1991), then hold that territory in the face of decades of extreme police repression. Comparing four Brazilian cities, Lessing (2008) found this level of drug-market organization unique to Rio, and plausibly due to the CV’s prison-based governance structure.

**Parallel Power:** Prison-gang authority can extend to entire peripheral regions and populations, providing order, justice, and other public goods, and effectively supplanting state authority. In Rio de Janeiro, an entire generation of favela residents has been born and raised under the armed dominion of prison-coordinated drug syndicates, while the state’s presence was largely limited to intermittent, corrupt, and highly lethal police invasions (e.g. Arias 2006; Leeds 1996). As a founding CV member explained, “We catechize the favela residents and show them that the government cannot help them or see their side of things. So we give food, medicine, clothes, textbooks.... We even resolve domestic disputes; there can’t be trouble or else the police will enter” (Amorim 1993, 162).
Orchestrated Protest and Violence as a Bargaining Chip: These tactics work both inside and outside prison. The CV—whose founding members watched the leftist militants they were housed with successfully protest their way to amnesty—regularly organized hunger strikes and petitions (Lima 1991), often coercing the larger inmate population into adherence. The CV has also instigated prison riots, often in multiple prisons simultaneously, as a means of pressuring or punishing officials. On the outside, the CV has frequently induced its foot soldiers in favelas to carry out city-wide shutdowns of businesses, burn busses, and machine-gun public buildings and police stations (e.g. Penglase 2005), usually to pressure officials to slacken carceral policies.

Projection of Power in Tested Cases: Like the CV, California’s Eme and São Paulo’s PCC have both used their coercive power to organize street-level drug markets. Yet whereas the Eme’s power is limited to areas dominated by southern Californian latino gangs (Skarbek 2011), the PCC operates throughout the entire urban periphery of São Paulo as wholesaler, tax collector, and arbiter of disputes among myriad small-scale retailers (Feltran 2010; Hirata 2010, 289). It has imposed a violence-limiting lei do crime (“criminal code of conduct”) (Telles and Hirata 2009, 53) through an astonishing system of trials, via cell-phone conferencing, before a jury of jailed PCC elders (Feltran 2010).

In El Salvador (as well as Guatemala and Honduras) the maras organized extortion rackets, perhaps because retail drug markets were too small. Leaders introduced professionalized hierarchies, stricter and savvier codes of outside behavior (e.g. prohibiting gang tattoos that made members easy targets for anti-gang enforcement) (Cruz 2010, 390-392; Savenije 2009; Wolf 2012, 86-87) and a system of prison-coordinated and -taxed extortion (Fogelbach 2010, 439) of businesses and public transportation known as la renta (the rent) (Aguilar and Carranza 2008, 23). Mara leaders explicitly attribute this shift in structure and behavior to increased incarceration under anti-gang measures:

“Before [the Mano Dura mass-incarceration policy] began it was different. We hadn’t gotten to seeing things collectively. The system has united us… like it or not, we cannot look at things...
individually, because they haven’t treated us individually, nor have they pursued or locked us up individually” (Cruz 2010, 393).

In terms of parallel power, the Eme has made minor efforts, coordinating offensives by affiliated Sureño gangs against black residents in Los Angeles (United States of America v. Rios et al. 2011) and Norteños governed by Eme rival La Nuestra Familia in central California (Reynolds and Sánchez 2003; Valdez 2011, 32). *Maras*, by contrast, play a dominant role in neighborhoods throughout El Salvador. The PCC has relentlessly expanded its presence throughout São Paulo’s urban periphery since 2000 (Biderman et al. 2014), and its dispute-resolution and order-provision services now extend to a broad population poorly served by state institutions (Feltran 2010; Hirata 2010). As one detective noted: “[T]he PCC is now judging small-claims cases, even domestic disputes. It’s clogging up our wiretaps, which capture fewer and fewer [serious crimes]” (Redação Terra 2008).

Finally, in their use of violence and its curtailment as a political bargaining chip, the PCC and the Salvadoran *maras* are unparalleled. The 2006 PCC attacks, more than just a destructive affront to state authority, were an effective political cudgel: they not only forced concessions in carceral policy, but helped defeat PCC antagonist Gerardo Alckmin, then-governor of São Paulo and architect of its mass incarceration policies, in his 2006 bid to unseat President Lula da Silva. When I asked what the PCC gained from their attacks, São Paulo’s former DA for Organized Crime told me, “Power, in the political arena. Now they must always be taken into consideration; everyone is afraid.” Meanwhile, the halving of São Paulo’s homicide rate since 2000 is widely attributed to the PCC’s ‘criminal code of conduct’ (e.g. Biderman et al. 2014; Feltran 2010), providing its leaders with additional leverage.

El Salvador’s *maras* followed their 2010 show of force with a March 2012 prison-brokered truce that produced a stunning 60 percent drop in the national homicide rate—testifying to imprisoned leaders’ control over street-level behavior. Though the government initially denied any role in the truce, top *mara* leaders were returned from isolation to low-security prisons and allowed cell phones, among other concessions (Economist 2012). Once the homicide drop became undeniable, the government began to take credit, inviting security ministers from Guatemala and Honduras to discuss exporting the Sal-

---

37 Interview, September 1, 2009.
vadoran ‘experiment’ (Membreño 2012a). Gang leaders later ‘deepened’ the truce, curbing extortion in exchange for reduced police patrols (Membreño 2012b). The efficacy of the truce gave maras important political leverage as well: bargaining for jobs programs, they threatened an increase in violence were the truce to collapse (Economist 2014). Indeed, violence surged in 2014 as a new government was elected that disowned the truce and any future negotiations. Mara leaders then reached a new truce on their own, again halving homicide rates, aiming, perhaps, to gain leverage against a hostile administration (Pachicho 2015). In July 2015, they returned to violent strategies, inducing another transportation strike (Malkin 2015).

5.2 Projection of Power as a Collective Threat to State Authority

Disentangling the perniciousness of prison-gang power from its sometimes salutary effect on crime requires careful conceptualization. Canonical notions of state authority center on the ability to lay down rules that “subjects” ultimately submit to voluntarily (Weber 1947). Coercion may be needed to establish authority (Tilly 1985), and assure citizens that free-riders will be punished (Levi 1989), but order, once imposed, is self-reinforcing, in part due to the surplus it produces (Olson 1993). Myerson (2009) further identifies authority with Schelling’s (1960) notion of “focal points”: the ability to get people to focus on one of multiple potential equilibria. By definition, once we all believe that everyone else will play according to a given equilibrium, nobody has reason to defect—an apt game-theoretic translation of Weber’s “voluntary submission”.

A distinction can now be drawn: individual challenges to state authority consist in non-submission to its rules, i.e. law-breaking. Collective challenges consist in laying down one’s own rules, establishing non-state authority at the expense of the state’s. Individual crime, in Myerson’s terms, is a defection from the state-selected equilibrium of law and order; its out-of-equilibrium nature is evidenced by the fact that crimes often surprise or shock us. As crime rates rise, law and order loses its focal power, and new equilibria such as vigilantism—or even widespread looting—take hold. Incarcerating criminals can restore state authority by reducing individual crime: by re-establishing a common expectation that others will respect the law (or be incapacitated from breaking it again).
Such individual affronts to state authority, even in aggregate, are conceptually different from collective threats. Of course, criminal groups also engage in crime and chaos-causing. But when they impose their own rules and social order, producing a shared surplus, they are establishing their own (criminal) form of authority. “Collaboration is voluntary,” a Salvadoran mara leader explained, echoing Weber, when I asked how leaders got so many street members to observe the prison-negotiated truce: “Backed by discipline, but voluntary.”

Even if some types of non-state authority need not come at the expense of state authority—such as within corporations, religions, and families—this could hardly be true for criminal groups. In practice, ethnographic evidence (e.g. Feltran 2010) reveals precisely the “brown areas” of limited state penetration and overlapping authority that O’Donnell (1993) identified as detrimental to democratic consolidation and the rule of law.

How then does increased punishment affect collective threats to state authority? For groups like insurgencies, mafias, and street gangs, mass incarceration probably reduces the ability to impose non-state authority. The model shows how this logic is inverted in the case of consolidated prison gangs: crackdowns increase coercive power on the street, reinforcing their ability to impose criminal authority. Regardless of whether this criminal authority is exercised in ways that increase or reduce street-level crime and violence, the effect on state authority is likely to be negative.

5.3 Counterproductive Punishment

Returning to Figure 1, the individual effects of incarceration on state authority—essentially the opposite of its effects on crime—are relatively easy to measure. The collective, authority-eroding effects of incarceration are far less visible. The causal pathway is neither obvious—it is modeled here for the first time, to the best of my knowledge—nor immediate: from consolidation to projection can take years or decades. Moreover, measurement error may worsen as gangs grow powerful, accumulate corrupting illicit rents, and give even honest officials incentives to prevent accurate assessments of prison-gang power.

38 Interview and visit, Quetzaltepeque Prison, El Salvador, May 24, 2013.

39 Above, I suggested defining ‘street gang’ as those whose leverage is reduced by untargeted crackdowns.
Another asymmetry concerns “returns to scale”. Theory predicts, and empirical results mostly confirm, that the marginal crime-reducing effect of punishment is decreasing (Kleiman 2009; Useem and Piehl 2008), and may even become negative beyond some point (Chen and Shapiro 2007; Gaes and Camp 2009; Liedka et al. 2006). For this reason, the solid line in Figure 1 is drawn as convex, with a confidence interval at its right tail that crosses zero.

In contrast, the marginal collective effects of punishment may be increasing over the relevant range (indicated by the concavity of the dashed line in Figure 1). First, as the model shows, the stronger prison gangs grow within prison, the more crackdowns increase their power outside; there will be positive feedback if gangs use the ensuing surpluses to further consolidate. Second, as incarceration becomes part of the life course for some groups (Pettit and Western 2004), prison-gang initiation may become a rite of passage, and obtaining “prison insurance” a widespread norm. Finally, the focal-point effects that help explain how prison-gangs quickly expand their authority from core members to larger constituencies are likely to be tipping-point phenomena: even modest increases in coercive power could have decisive effects. For example, in Los Angeles, the MS first brazenly opposed Eme taxation, then, as reprisals escalated, not only relented but amended its name (to MS-18) to proudly proclaim its Eme affiliation (Valdez 2011, 28-29). A former Rio drug lord also described a tipping point when I asked him about free members’ loyalty to imprisoned leaders: “[A rebellious lieutenant] could try a ‘coup d’état’... but it’s very rare.... Out of 30 employees, half would kill him, you can be sure... and you only need one.”

The logical consequence of the foregoing is the inflection point of Figure 1, beyond which additional incarceration undermines overall state authority, even if crime rates fall. In this view, prison gangs do not just produce hidden costs that can lead to higher-than-optimal incarceration. Rather, they may fundamentally limit states’ capacity to restore authority through punishment, no matter the price.

---

40 Kleiman (2009) emphasizes, however, that well-planned “dynamic concentration” policing may have tipping-point effects, and hence increasing returns.

41 The letter M (for Eme) is the eighteenth letter of the alphabet.

42 Author interview, Rio de Janeiro, August 17, 2009.
This may sound fanciful, but consider the radical (and expensive) experiments in mass incarceration carried out in California, São Paulo, and El Salvador: how much legitimate authority have they really produced? In the latter two, prison gangs can now use the threat of debilitating violence to force authorities into negotiation. California may lie further to the left in Figure 1: the stark racial cleavages that define U.S. gangs may prevent the emergence of “universal” and hegemonic gangs like the PCC, who are better able to carry out and reap the benefits of direct confrontation (Bailey and Taylor 2009). Nonetheless, U.S. prison gangs have organized riots and protests; the retail drug trade thrives; and in those street-level criminal markets, if not the broader communities they operate out of, order is maintained as much by prison-based criminal networks as by the state. For officials, it may seem expedient to leave the core state function of establishing order to illegal groups, who work “cheap”. Ultimately, though, such an abdication may erode state authority and constitute a net loss for society.

6 Conclusion

My core claim is simple: the more likely a potential criminal is to go to prison, and the longer he will stay if he goes, the more he is willing to “pay” for good standing with a prison gang that can protect him on the inside. I formally analyze this logic, distinguishing the individual effects (on overall crime) of mass incarceration policies from the collective (on prison-gang power). The model predicts that harsher sentences and poorly-targeted crackdowns strengthen prison gangs that already hold power within prison. These predictions fit an initial empirical assessment: in three prominent cases, mass incarceration policies and anti-gang sweeps not only failed to eliminate consolidated prison gangs, but seem to have fomented qualitative leaps in their projective capacity. These new uses of projection, I showed, constitute serious threats to state authority, though they can have varied effects on crime rates. The implication is that increased incarceration can strengthen criminal authority at the expense of the state, even if crime falls. Finally, if gang-strengthening effects have increasing returns—plausibly due to the tipping-point nature of authority—then state efforts to restore law and order through increased punishment may end up only eroding overall state authority. Like public torture before it (Foucault
mass incarceration could become a self-defeating form of punishment.

These are big claims; further research is needed to fully test them. Theoretical work should focus on prison gang leadership, to better understand why, when, and how it deploys its projective power. Empirically, better measures of projection of power are needed. Public attacks and protests provide critical snapshots of prison-gang strength, though their timing is endogenous. In the interim, researchers can triangulate among observations that suggest prison-gang projection: structural changes in criminal markets, sudden decreases in street-gang violence or its concentration into large-scale battles between consolidated networks, and targeted violence against officials in response to changes in carceral policy. Ultimately, though, our best observations of projection are ethnographic, and hence single-case, making empirical generalization problematic. Cooperatively designing and aggregating the results of parallel in situ studies could produce more systematic assessments.

Despite a rich tradition of prison ethnography (e.g. Sykes 1958), and increasing attention to the political and social consequences of a burgeoning “carceral state” (e.g. Gottschalk 2008; Simon 2007), prison gangs remain understudied relative to the potential threat they pose. Most of the time they lurk far from public view, a problem that officials, political leaders, and even scholars might prefer to ignore. Yet in the starkest cases, they have transformed the prison system—in theory the core of the coercive apparatus—into an operational headquarters’ for organized criminal defiance of state authority. Perhaps recommendations based on prison-gang research (Skarbek 2014) will yield more effective anti-gang programs; thus far such efforts have had only marginal impact (e.g. Knox 2012). And since incarceration has become, in the modern era, the punishment upon which all state coercion ultimately rests, what these cases suggest is a fundamental limit to state power.

References


REFERENCES


Lessing, Benjamin. 2008. “As Facções Cariocas em Perspectiva Comparativa [Rio’s Drug Syndicates in Compar-


