Making the Law by Measuring its Enforcers: Quantitative Management in Policing *

Laurel Eckhouse †

Police make central decisions about which laws are enforced, and against whom. Front-line police work with minimal supervision, and have tremendous discretion. Police officers have far more information about the local environment than their supervisors; certainly far more than the voters who delegate the tasks of enforcing laws and maintaining order to police departments. In such a context, "as all police officers and many citizens recognize, discretion is inevitable," as much now as when Wilson advanced the argument in 1968. *How* police exercise this discretion, and how voters and supervisors can constrain it, is a classic principal-agent problem: supervisors must delegate control over decision-making to street-level officers who have more complete information and may have preferences that diverge from their managers'. (cite Miller 2005) Police, of course, are not alone: social workers, teachers, and other "street-level bureaucrats" face the dilemma of applying policies evenhandedly to wildly divergent situations while their managers lack information(cite Lipsky, Hess).

Since the early 1990s, violent crime has fallen substantially, but police contact remains high. What explains this paradox? With discretion comes the risk of shirking. In the New York Police Department, "cooping" – the slang term for sleeping on duty – has been a concern since the early 20th century, when Theodore Roosevelt rousted sleeping officers (cite about Theodore Roosevelt). Even recently, it was sufficiently widespread that the department banned officers from the best nap

*I am grateful to the National Science Foundation Graduate Research Fellowship Program, the Horowitz Foundation for Social Policy, and the Travers Department of Political Science at UC Berkeley for support for this research.

†Doctoral Candidate, Department of Political Science, University of California at Berkeley. eckhouse@berkeley.edu

spots (cite NYT piece). Shirking also dominated the academic literature on policing (cite Brehm and Gates, Allen 1982, Engel 2000).

Drawing on technological changes in the feasibility of data analysis, police departments developed managerial techniques for monitoring crime and arrest statistics, then moved to hold officers and supervisors accountable for productivity – measured in both reductions of violent crime, and in individual demonstrations of work. Wilson and Kelling, in 1982, argued that attention to minor crimes would lead to reductions in violent crime; the new public management encouraged public agencies, including police, to decentralize accountability and develop incentives for performance. In 1994, Bill Bratton pioneered Compstat in New York City. This system for visually displaying crime data and directing police activity aimed to change police behavior by measuring and encouraging the desired outcomes (**?**); similar systems have spread to education, city administration, and other bureaucracies. (cite new public management literature) Data-driven strategies for evaluation are a form of what I call *quantitative management*. Compstat and other forms of quantitative management are an institutional effort to solve one of the core issues in the bureaucratic politics literature: the information asymmetry between policymakers and bureaucratic agents (cite Miller 2005). Policymakers use quantitative management to gather information about policy implementers, prevent shirking, and align the incentives of street-level bureaucrats with those of policymakers by rewarding or punishing them for their performance. (cite McCubbins and Schwartz; McCubbins, Noll, Weingast 1987;)

Measurement has pitfalls, however. Journalists have uncovered examples of police departments reclassifying homicides, downgrading rapes to less serious crimes, and misclassifying assaults. (cite Rayman, Chicago Mag, LAT). In schools, cheating scandals have arisen in cities under pressure to meet test score targets. Data manipulation is not the only potential problem with quantitative targets, though: in places where Compstat and Broken Windows policing have been implemented, officers report pressure to meet quotas for arrests and tickets. Because departments track these productivity measures, officers express concern that if their arrest rates are low, they will face career consequences.

Using a formal model, I show that incentive changes designed to reduce shirking change the mix of tasks that street-level bureaucrats pursue, increasing incentives for police to pursue minor arrests rather than invest in long-term projects. This type of task substitution differs from the Brehm and Gates model of working, shirking, and sabotage: instead, it describes the ways that monitoring per se can influence the selection of different types of work. Goldstein and You provide empirical evidence that resource constraints mean that police departments must choose between routine enforcement and long-term work. However, supervisors face inherent challenges in simultaneously discouraging shirking and encouraging long-term work with inherently probabilistic payoffs, such as building community relationships and investigating serious crimes.

The model presents a two-player game: at the beginning of the game, the Supervisor sets a probability of monitoring. The Agent can choose one of three strategies: shirking ($shirk$), pursuing minor arrests ($arrest$), or investing in long-term work ($work$). The long-term work will $stall$ or $solve$ with some probability set by nature; if it stalls, the observable outcome from the Supervisor's perspective is the same as if the Agent had chosen $shirk$. The key result is that increasing the probability of monitoring (or the severity of punishment for shirking) decreases the Agent's payoffs for *both* $shirk$ and $work$, since sanctions apply to both $stall$ed cases and $shirk$ing behavior.

I test the observable hypotheses generated by this model with empirical data on the consequences of adopting Compstat, the quantitative management system for policing. I find evidence for task substitution: quantitative management is associated with a statistically significant increase in the share of arrests that are for minor offenses, as well as with an increase in the number of arrests overall. Bureaucratic incentives thus help explain the paradox of falling crime rates and high contact with the police: monitoring creates incentives to engage in activities with reliable, visible payoffs, even as serious crime has fallen. Some have argued that this increase is a function of a resource allocation shift from investigations of homicide and other serious crime to patrol and minor arrests (cite Leovy). I distinguish between the consequences of monitoring and the consequences of resource allocation by separately testing the effects of Compstat on the rates of clearances, which are created by separate units who cannot switch to routine work. I find evidence that Compstat has improved clearance rates, suggesting that task substitution rather than resource allocation is behind the increases in minor arrests.

I also test the hypotheses advanced by both critics and advocates of Compstat: that quantitative management leads to data manipulation, and that it reduces serious crime. I find evidence for data manipulation: adopting quantitative management is associated with a significant increase in classifications of rape as 'unfounded', a way of manipulating rape statistics (cite). Placebo tests for the share of assaults and murders classified as unfounded confirm that this is not an artifact of data reporting.

Some might argue that increases in minor arrests merely fulfill the policy goals of Broken Windows policing (cite Wilson and Kelling). Compstat adopters do not share this belief, however: they describe using quantitative management because they hope it will help reduce serious crime and allow more control over department operations (cite Weisburd et al). These goals are not supported by the data: quantitative management is not associated with a decline in the incidence of index crimes. Recent evidence suggests that public disorder is not, contra Wilson and Kelling, the precursor to violence: instead, violence and disorder follow private, intra-community disorder (cite Sampson and O'Brien). The result is that Compstat adoption is associated with data manipulation and minor arrests, without being associated with a drop in serious crime; however, Compstat also contributes to improved clearance rates for high-profile cases, like murders.

The unintended consequences of bureaucratic accountability measures in policing should matter to policymakers and scholars. Increases in misdemeanor arrests have important social and political consequences, even when they result in neither convictions nor jail time. They bring more people under state surveillance and put those arrested at risk for future warrants or harsher sentencing (Kohler-Hausmann); even without jail time, those arrested lose jobs and suffer economic losses (cite bail article, Howell 2009). As Eric Garner's death while under arrest for selling loose cigarettes shows, contact with police carries the risk of serious consequences even when the crime in question is minor. Carceral contact leads to important political spillover effects: less voting, fewer requests for help from city government, less civic engagement (Lerman and Weaver x3, Burch, Clear). Arrests, warrants, and incarceration disrupt social relationships and economic options, generating more serious crime (cite Goffman, Pager). The burden of these minor arrests falls most heavily on black and Latino communities, and especially on young men: this has not gone unnoticed, and the distributive and procedural justice concerns damage the credibility of the criminal justice system and reduce cooperation with the law (cite Fagan, Tyler and Wakslak, others). However, most political science scholarship on the carceral state has focused on the feedback effects of the carceral state, like voting (cite Meredith and Morse; Lerman and Weaver; Green; Uggen and Manza), or on qualitative and historical analyses of specific policies (cite Weaver: Frontlash, Embedding Crime Policy; Murakawa).

My research suggests that local institutions and bureaucratic governance play a critical role in shaping the distribution of police contact, and therefore of enforcement of the law. Most research on the expansion in carceral contact over the last several decade locates the causes either in the attributes of individual police officers or in changes in state and federal criminal law (cite Weaver, Murakawa, Eberhardt, Glaser, Twersky-Glazner, Fielding and Fielding). These theories, however, struggle to explain the wide variation in carceral contact between cities. Local policy differences provide more leverage for reformers: trying to change the attitudes of individual police relies on replacing officers – a challenging feat, considering the wide variety of potential attributes to consider and the strong union protections of police – or changing the attitudes of existing police with poorly supported tools (cite Paluck and Green).

This research does not suggest that police should give up on quantitative management entirely. Collecting and publishing data on police activity is key to transparency. Rather, police departments – and other agencies using quantitative management – should be attentive to the incentive problems that monitoring quantitative results can produce. Supervisors need to understand how police activities are chosen, carried out, and recorded in thorough qualitative detail; supervisors and managers need to be aware of the risks of task substitution and data manipulation. Quantitative management and monitoring are well-suited to *stopping* behaviors: used with appropriate attention to data quality, they can reduce shirking and the number of police stops (cite Mummolo). They are less suited, however, to encouraging hard-to-monitor behaviors where the quality of the work is especially important, like solving crimes, defusing conflicts, and building civic trust.

1 Monitoring Probabilistic Work

What happens when supervisors try to deter shirking, but stalled work is observationally equivalent to shirking? I use a two-player game to model the principal-agent dynamics involved. At the outset of the game, the Supervisor sets the probability of monitoring, m. The Agent then chooses to $shirk$, $work$, or $arrest$. Choosing $work$ has an immediate cost c in effort; the case will $solve$ with probability p and otherwise $stall$, leaving the Agent no observable accomplishment. This strategy models the choice to pursue long-term, probabilistic work: investigations, community relationship building, and other police work with uncertain payoffs.

Choosing $arrest$ yields a guaranteed payoff x, net of the cost of the strategy to the Agent. This strategy corresponds to the decision to pursue routine arrests. Choosing $shirk$ has no direct cost to the Agent, and yields a benefit of s. After the Agent chooses a strategy, the Supervisor monitors with the probability chosen at the outset. If the Agent has chosen $shirk$, or if $work$ has $stall$ed, the Supervisor imposes a penalty z; otherwise, the Agent keeps the payoff. The sequence of play and the payoffs are summarized in Figure 1 and Table 1.

Figure 1: Game

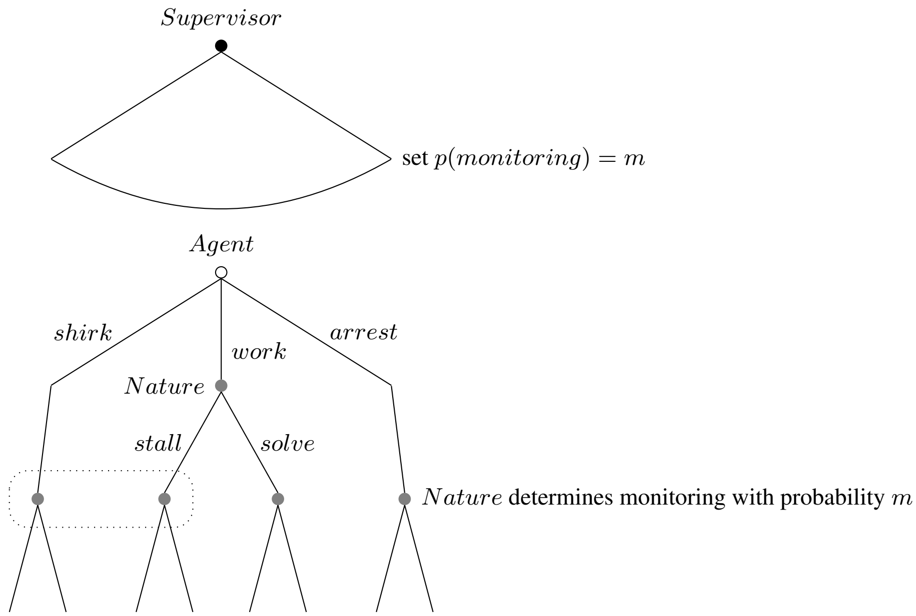

Table 1: Payoffs ($Agent$, $Supervisor$)

$Agent$ strategy	$arrest$	$work$	$shirk$
payoff	(x, a)	$(yp - e - (1-p)mz, bp)$	$(s - mz, c)$

Suppose the Supervisor's payoffs are such that $bp > a > c$: that is, the Supervisor would prefer the Agent to $work$, but would prefer $arrest$ to $shirk$. Increasing the probability of monitoring, m, is formally equivalent to increasing the severity of punishment for choosing $shirk$. As mz rises, though, the value of $work$ relative to $arrest$ declines. The payoff for $arrest$ is not changed by monitoring, but the payoff for $work$ declines by $(1-p)z$ multipled by the change in the probability of monitoring. Similarly, the payoff for $shirk$ declines even more steeply, by zm. Unless Agents have very strong exogenous preferences to $work$ – that is, $yp - e$ is higher than x by at least $(1-p)mz$ – increasing monitoring will not only decrease shirking, but decrease working.

An Agent who is indifferent between $work$ and $arrest$ without monitoring – that is, for whom $yp - e = x$ – will prefer $arrest$ in the presence of monitoring. However, if $s > yp - c$ and $s > a$, the Supervisor must increase mz to induce the Agent to choose $work$ or $arrest$. Unless the payoff of $work$ is already higher than the payoff for $arrest$ by at least $(1-p)mz$, the Agent will choose $arrest$.

This model is agnostic with respect to the Agent's preferences in the absence of monitoring. Indeed, police departments are likely to be filled with individuals with varying values and costs for different kinds of work, depending on their personal preferences, career goals, and beliefs about policing. Rather, it shows that Supervisors face an inevitable conflict between deterring $shirk$ing and promoting probabilistic $work$, which is observationally indistinguishable from $shirk$ing when it does not $solve$.

Probabilistic work is common in many types of bureaucracies. Should teachers focus on long-term capacity building and classroom community development, which occurs at an uneven rate, or teach test strategies by rote? How should social workers allocate their time? Bureaucratic principals who want to deter shirking must reckon with the risk that they will lead agents to substitute routine, reliable work for long-term, probabilistic projects.

1.1 Observable Implications

- As monitoring increases, shirking should decline as the value of $shirk$ falls. That is,

$$\frac{d}{dm}(shirk) = \frac{d}{dm}(s - mz) = -z < 0 \tag{1}$$

While no data measures shirking, recent ethnographers of black neighborhoods in major cities write that ”times have changed” since the days of the 1980s, when ethnographers reported that police left black neighborhoods largely to their own devices. Now, warrant sweeps at workplaces and neighborhood gathering spots are common, as are street stops. The risk of arrest for minor offenses is high in many cities, especially for young Black and

Latino men. (cite Goffman; Anderson; Gelman Fagan and Kiss, Goel)

- As monitoring increases, the absolute number of $arrest$s for minor offenses should increase. As monitoring increases, the expected utility of $arrest$ for the Agent is constant as monitoring increases.

$$\frac{d}{dm}(arrest) = \frac{d}{dm}x = 0 \tag{2}$$

  Thus, for individual Agents who prefer $shirk$ to $arrest$ by an amount $\leq z$ there will be a point at which the payoff from $arrest$ is greater than the payoff from $shirk$.

  The payoff for $work$ also falls as monitoring increases, as follows:

$$\frac{d}{dm}(work) = \frac{d}{dm}(yp - e - (1-p)mz) = -(1-p)z < 0 \tag{3}$$

  Since the payoff for $work$ and $shirk$ fall with increased monitoring, an increase in monitoring should lead more agents to pursue $arrest$, with the observable implication of an increase in routine arrests.

- If monitoring causes task substitution, the share of tasks completed that indicate the choice to $arrest$ should increase. As the value of $arrest$ rises relative to $work$ and $shirk$, Agents will shift their strategies.

- Tasks not substitutable with $arrest$ will not be governed by this dynamic. For example, felony investigations are conducted by dedicated squads who do not have the option of switching to $arrest$. In these cases, increased monitoring will increase the choice of $work$. This can be modeled by simply omitting $arrest$ from the game tree. While monitoring decreases the payoff for $work$ as well, the payoff for $work$ falls more slowly than the payoff for $shirk$, as shown below:

$$-z < -(1-p)z < 0 \tag{4}$$

Note that task substitution may still occur if Agents can choose among tasks with different probabilities of $solve$. Increases in monitoring decrease the value of $work$ by $-(1-p)z$, so as p increases the expected cost of monitoring for the agent decreases.

2 Data on Policing Data

I test this model with a data set on Compstat adoption. Since Bratton and the NYPD developed it, Compstat has become the professional norm among police departments; adopters describe reducing crime and gaining additional control over field operations as among the central reasons for using Compstat. Willis et al. identify these programs as incorporating specific, measurable objectives; regular meetings using data to assess progress towards those objectives and evaluate strategies; and accountability for middle managers. “These features are most visible in the NYPDs twice-weekly Compstat ‘Crime Control Strategy Meetings,’ during which precinct commanders appear before the departments top echelon to report on crime in their districts and what they are doing about it Crime analysts collect, analyze, and map crime statistics to spot trends and help precinct commanders identify underlying factors that explain crime incidents. Top administrators use this information to quiz precinct commanders on the crime in their beats and to hold them responsible for solving the problems. Failure to provide satisfactory responses to these inquiries may lead to stern criticism or removal from command.”**?** The key features of Compstat, both in New York City and elsewhere, including mapped data collection, regular meetings, and shifts in both authority and responsibility for data to the managers of geographic subunits. These meetings are aimed at inducing managers to reduce the incidence of major crimes, but quantitative management in policing also involves monitoring data about the performance of individual officers.

Compstat is thus an ideal test case for this model. It is a form of increased monitoring, driven by the twin goals of reducing serious crime and increasing control over operations (which includes reducing shirking). It has been widely adopted in the United States, but temporal variation in diffusion means that the effects of adoption can be distinguished from overall time trends. How does quantitative management change the practice of policing?

Previous evaluations of the consequences of quantitative management have focused on individual cities. Reporters and ethnographers have described how departments supervise police officers' day to day activities and the collection of crime statistics. However, this is the first nation-wide evaluation of Compstat's effects on police work.

Data Collection

I collected data from newspapers, police department websites, and other public sources on the adoption of CompStat or similar quantitative management techniques by police departments in 55 of the 68 largest cities in the United States. We coded cities as adopting quantitative management when they began using the combination of mapping software and accountability meetings. In practice, adoption of quantitative management was not difficult to identify: most departments (or the journalists reporting on them) compared quantitative management initiatives to CompStat and described the similarities to New York's pioneering program. Of the departments excluded, most used CompStat but the team was unable to determine a date of adoption; only three departments had not adopted CompStat or a similar mapping/accountability program.

Sociologists and criminologists studying CompStat have raised important questions about whether CompStat and other quantitative management techniques actually influence internal organization. Jeffers, observing CompStat implementation in a Midwestern citys police department, reports that experienced officers did not change their policing strategies despite major changes in performance evaluation policy (**?**). Weisburd et al. argue that departments adopting CompStat generally do not engage in the wholesale organizational overhaul required for true data-driven accountability among officers (**?**). Bratton's consulting organization agrees, finding in recommendations to the city of Oakland (CA) that the Compstat Process as previously practiced in Oakland was more of a report or a presentation by a captain than the system of vigorous strategic oversight." (**?**) Ironically, a consulting firm headed by Patrick Harnett, who worked for Bratton in New York, had made substantially similar recommendations seven years earlier, which were partially adopted and then later abandoned (**?**). While mapping does not require complex institutional changes, the organizational reforms are less likely to be fully implemented. This data collection effort does not attempt to measure the implementation of quantitative management or accountability; to the extent that this biases the estimates, though, it should work against finding effects of quantitative management.

I merge this data with data on arrests from the Federal Bureau of Investigations Uniform Crime Reporting Program from 1990 to 2013. I operationalize major crimes as crimes reported in Part 1 of the Uniform Crime Reporting guidelines, which are also sometimes called index crimes. These include homicide, manslaughter, forcible rape, robbery, aggravated assault, larceny/theft, arson, and motor vehicle theft. I operationalize minor crimes as crimes reported under Part 2 of the Uniform Crime Reporting guidelines, which includes such arrests as those for drug charges, simple assault, fraud, receiving or possessing stolen property, vandalism, weapons charges, prostitution, and quality of life offenses like drunkenness and disorderly conduct. Eight of the departments in question have missing UCR records over a large portion of this period, and are therefore dropped, for a total of 47 departments over 23 years, or 1081 unique department-year observations. Appendix A lists the cities included in the analysis.

I also use Federal Bureau of Investigation data on the number of offenses known to police and the number of clearances (i.e. how many crimes result in an arrest or other clear designation of the perpetrator).[1] Many analyses of crime focus on arrests as a proxy for the number of crimes; as I

[1] Arrests can also be cleared by "extraordinary" means if the police identify the responsible party as someone who is unavailable to arrest: for example, someone who has died since the incident.

show above, though, these may vary separately, and their differences provide an important source of information.

2.1 Methodology

I use fixed effects regression to test the hypotheses described below. Including city-level fixed effects means that coefficient estimates capture variation *within* rather than *between* cities, so these results are not biased by unobserved differences between cities. I also include year fixed effects, because both crime and arrests have substantial time-linked variation. I test the regressions with and without demographic covariates (percent of population that is black and percent of population that is white) at the county level.

The estimated equation is shown below:

$$Y_{it} = \beta X_{it} + \alpha_i + \gamma_t + \epsilon \tag{5}$$

Y is the dependent variable for each analysis. X_{it} is the variable for Compstat adoption, as well as a matrix of control variables described for each individual analysis. α_i is the agency effect, and γ_t represents the year effect.

3 Hypotheses and Results

I test four empirical hypotheses. The first two arise directly from the model. First, increasing monitoring by adopting Compstat increases both the absolute number and the share of arrests that are for minor crimes. Second, to distinguish between task substitution and resource allocation, I test the effects of Compstat on clearances, which are driven by the activities of organizationally separate units. Third, I test for data manipulation. Finally, I test the argument made by Compstat advocates: that serious crime will decline as a result of improved police performance and reduced shirking.

The results confirm the model's predictions: adopting Compstat increases both the absolute number of arrests for minor offenses, and the share of arrests that are for minor offenses. The Compstat-associated increase is over 20,000 arrests yearly, nearly 42% of the mean number of arrests. Increased monitoring thus leads to tremendous increases in exposure to carceral contact for minor offenses. As described above and in the conclusions section, such misdemeanor arrests have substantive consequences for individuals' political participation, financial situations, and life chances, even when charges are dismissed or defendants are sentenced only to brief jail terms. (White, Lerman and Weaver, Goffman, Kohler-Hausmann)

Compstat also improves clearance rates, especially for homicides. Most felony clearances are the work of dedicated investigative squads rather than officers engaged in street patrol, so this provides evidence that increases in minor arrests are not driven by increased resources for patrol. When task substition is not an option, monitoring can increase desired behavior as well as (per Mummolo) decreasing unwanted behavior.

Pressure on data has been linked to manipulation both in journalistic investigations of policing and in other bureaucratic contexts, such as education. (cite LA Times, Chicago Mag, Rayman) I develop a strategy for identifying manipulation in rape statistics: the share of reported rapes reclassified as unfounded. I find a clear link between Compstat adoption and increases in unfoundedness in rape statistics.

Finally, I test the main reason agencies give for adopting Compstat: the hypothesis that it reduces serious crime. While arrests for serious offenses increase, I find no change in the number of serious incidents.

3.0.1 Increases in minor arrests

While officers, teachers, and other bureaucrats are evaluated on the results they are intended to achieve, it is often quite difficult for service providers to influence student scores or crime directly. Instead, supervisors collect data on the activities of the bureaucrats they manage. With data collection in place, supervisors are able to reward agents for carrying out tasks which are frequently performed, reliably accomplished, and easily measured. Work leading to major arrests is a probabilistic activity: officers need to devote substantial resources without guaranteed payoff. Preventive policing and activities that build relationships with communities are even more difficult to reward, since they leave no measurable trace of the officers activity (**?**). In policing, these qualities are characteristic of (inevitably highly discretionary) enforcement of minor crimes. Warrant sweeps and arrests for minor crimes such as casual drug use are easily performed and documented. As the model descibes, quantitative management in policing puts pressure on police to increase their visible enforcement activity. Rayman describes a particularly vivid example of this in his reporting on the NYPD tapes scandal, in which a rookie officer taped meetings at which patrol officers were instructed to meet minimum arrest and citation quotas. This leads to the first hypothesis: adopting quantitative management will lead to an increase in arrests for minor crimes, both in absolute terms and in relative terms. Since quantitative management includes monitoring of officer and unit productivity, we would expect that officers would demonstrate productivity by engaging in routine policing with reliable, visible results. Thus, both the *number* and the *share* of arrests that are for minor, consensual crimes should increase.

I operationalize minor crimes by examining crimes reported under Part 2 in the Uniform Crime Reporting guidelines, which includes such arrests as those for drug charges, simple assault, fraud, receiving or possessing stolen property, vandalism, weapons charges, prostitution, and quality of life offenses like drunkenness and disorderly conduct. [2] In cities that adopt Compstat, arrests for

[2]These crimes create disorder, which many citizens would prefer to see ameliorated. Using these arrests to measure 'minor' crimes is not a claim that they are not important to city residents. Rather, these are exactly the types of arrests

Part 2 offenses increase, both in absolute terms (controlling for population) and as a share of the total number of arrests. Adopting Compstat is associated with over $20,000$ additional Part 2 arrests per year ($p < 0.001$, or an increase of 1.3 percentage points ($p < 0.05$) in the share of Part 2 arrests. The reported occurrence of Part 2 crimes increases, as do arrests for these minor crimes. Quantitative management is not merely a tool which enforces accountability and prevents agents from shirking; it has substantive effects on the character of the policies being implemented.

Table 2: Effect of Compstat on Number of Part 2 Arrests

	Part 2 arrests		
	(1)	(2)	(3)
Compstat	20,471***	20,294***	20,176***
	(6,339)	(5,737)	(5,729)
Part 1 incidents		0.82***	0.80***
		(0.06)	(0.06)
Agency	Yes	Yes	Yes
Year	Yes	Yes	Yes
Constant	No	No	No
Demographics	No	No	No
Population	Yes	No	No
Observations	1,081	1,081	1,081
Note:	*p<0.1; **p<0.05; ***p<0.01		

Distinguishing between task substitution and resource allocation.

Leovy, writing about homicides in Los Angeles, argues that a focus on patrol and arrests has shifted police focus away from solving homicides and other serious crimes, and toward efforts to reduce disorder (**?**). This suggests an alternate hypothesis for explaining the above increase in Part 2

that are most discretionary for police, and for which the 'correct' level of arrestable offenses is least clear.

Table 3: Effect of Compstat on Share of Part 2 Arrests

	Share of Part 2 arrests			
	(1)	(2)	(3)	(4)
Compstat	0.01302***	0.0131***	0.00979**	0.00979**
	(0.00444)	(0.00438)	(0.00428)	(0.00428)
Part 1 incidents per capita			−0.671***	−0.671***
			(0.0913)	(0.0914)
Demographics	No	Yes	No	Yes
Population	No	No	No	No
Observations	1,023	1,023	1,023	1,023

Note: *p<0.1; **p<0.05; ***p<0.01
All regressions include year and agency fixed effects.

arrests: if departments have allocated more resources to arrests, clearance rates should decline. In fact, I find that clearance rates rise after Compstat adoption. This effect is most pronounced for homicides.[3] This suggests that departments adopting Compstat have used separate incentives to boost murder clearances – also a clear, legible, publicly visible data point.

I test this using Federal Bureau of Investigation data on offenses known to police and clearances, as well as data on arrests and Compstat adoption. I find that departments that adopt Compstat see an increase in the number of Part 1 arrests, and in the overall number of murder arrests. The share of Part 1 offenses and homicides that are listed as cleared also increases when departments adopt Compstat. Clearances for assault, rape, and manslaughter do not see significant changes, however. This is consistent with reactions to quantitative management in other areas: murder clearance rates are among the most well-publicized performance figures, and make a clear top priority for departments concerned about public reaction. Compstat adoption is associated with 5.1 additional murder arrests ($p < 0.001$), controlling for the number of murders and manslaughters,

[3]Clearance rates for rapes also rise, but as I show in the next section, this may be an artifact of data manipulation.

and an increase of 5.3 percentage points in the share of murders that are cleared ($p < 0.001$). However, it is associated with only a 1.6 percentage point increase in the total share of Part 1 incidents that are cleared, suggesting that Compstat changes departmental focus to the most public numbers. Clearing homicides, however, is a particularly significant activity: Leovy and others argue that when the state does not punish killers, citizens seek vigilante justice in the form of retributive murder.

Table 4: Effect of Compstat on Clearance Rates

	Clearance Rates				
	Murder	Overall	Assault	Rape	Auto Theft
	(1)	(2)	(3)	(4)	(5)
Compstat	0.062***	0.018***	−0.001	0.031**	0.008
	(0.019)	(0.005)	(0.011)	(0.016)	(0.005)
Agency	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Constant	No	No	No	No	No
Demographics	Yes	Yes	Yes	Yes	Yes
Population	Yes	Yes	Yes	Yes	Yes
Observations	898	898	898	897	898

Note: *p<0.1; **p<0.05; ***p<0.01

Data manipulation

When people's careers and livelihoods are affected by data they produce but the reported activity is difficult to change, they have an incentive to manipulate the data by changing reporting patterns. In policing, this may lead major crimes to be underreported, because supervisors are rewarded for having reduced the levels of serious crimes; at the same time police are expected to demonstrate productivity through specific, quantifiable actions, leading them to over-enforce and thus over-

Table 5: Effect of Compstat on Murder Arrests, Murder Incidents, and Death Incidents

	Murder Arrests		Murders		Death Incidents	
	(1)	(2)	(3)	(4)	(5)	(6)
Compstat	51.843***	51.419***	−4.188	−4.900	−3.850	−4.555
	(11.784)	(11.742)	(13.411)	(13.311)	(13.456)	(13.359)
Number of murders	0.346***	0.344***				
	(0.031)	(0.031)				
Agency	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Constant	No	No	No	No	No	No
Demographics	No	No	No	No	No	No
Population	Yes	Yes	Yes	Yes	Yes	Yes
Observations	980	980	1,081	1,081	1,081	1,081

Note: *p<0.1; **p<0.05; ***p<0.01
All regressions include year and agency fixed effects.

report minor violations.

Media accounts of the New York, Chicago, and Los Angeles police departments have described police downgrading of serious crimes. In New York, officers under pressure to post weekly crime reductions reported sexual assaults as criminal trespassing or other misdemeanors. Over a two-month period, the man committing these assaults became increasingly confident that he would not be caught, and increasingly bold in his choice of methods and targets. Indeed, the pattern of attacks was not discovered until he was apprehended and confessed, at which point the detective questioning him looked through the precinct's complaints and found the misclassified incidents **?**.

This single case exemplifies the real-world consequences of data manipulation. Communities are robbed of core police services – public safety and access to formal justice – when officers downgrade complaints. In Chicago, police manipulated the number of robberies and, more dramatically, the number of homicides, typically considered the benchmark statistic that departments will be unable to alter. The Chicago Police Department hid bodies in the data by reclassifying suspicious deaths as "noncriminal death investigations" so they would not contribute to the city's murder rate; these noncriminal deaths included one where a woman's body was found in an abandoned warehouse, tied to a chair and gagged **?**. A Los Angeles Times investigation found that the LAPD misclassified nearly 1200 violent crimes as minor incidents (**?**).

Reporters, officers, and citizens in these accounts argue that pressure to produce quantitative results led to data manipulation. Eterno and Silverman find evidence that NYPD officers frequently observe such behavior (**?**). Similar effects have been observed in education. After No Child Left Behind made test scores a critical measure of school and district success, teachers and administrators were implicated in cheating scandals. Without national analysis, though, it is difficult to determine whether this type of data manipulation is commonly associated with adopting quantitative management. Measuring this type of manipulation in quantitative data is challenging. However, investigations in Chattanooga and other mid-sized cities suggest a way: police can alter rape statistics – unlike numbers for assaults, homicides, and other violent crimes – by designating rapes ‘unfounded’, a category which is reported in the Federal Bureau of Investigation’s Uniform Crime Reports. Yung’s comprehensive survey of problems in rape statistics summarizes the evidence: “ Several small studies indicate that police classify numerous ordinary rape complaints (often involving intoxicated or confused victims) as unfounded. Professor Jeffrey Bouffards 2000 study of one unnamed police department found that 27.9% of cases were classified as ‘unfounded.’ The newspaper accounts of police practices in Baltimore, New Orleans, and St. Louis illustrate how the ‘unfounded’ designation served as a convenient technique to make many rape complaints disappear from crime statistics. Indeed, police in Baltimore turned the UCR exception into a verb by openly stating that they had ’unfound’ a rape complaint.” (**?**). FBI reports on unfounded crimes therefore offer a national source of data on potentially manipulated rape statistics.

I find that adopting Compstat is associated with an increase of 1.9 percentage points ($p < .01$) in the share of rapes that are designated unfounded. Since the mean share of rapes reported unfounded is 6.8%, adopting Compstat is associated with a 28% increase in the share of rapes that are reported as unfounded. I also use auto thefts and homicides as a placebo test, to confirm that the increase in unfoundedness in rape complaints does not arise from improved record-keeping. Auto thefts and homicides are the most difficult statistics to manipulate: the former, because victims insist on reporting for insurance reasons, and the latter because concealing homicides involves the difficult task of concealing murder victims. Neither auto thefts nor homicides are more likely to be considered unfounded with the adoption of Compstat.

The second two hypotheses test other arguments for and against quantitative management.

3.0.2 No effect on serious crime

Departments adopt Compstat because they hope quantitative management will help them reduce serious crime, and because they hope it will improve departmental control over field operations(**?**).

Table 6: Effect of Compstat on Share of Rapes Declared Unfounded

	Share Unfounded					
	Rape		Auto Theft		Murder	
	(1)	(2)	(3)	(4)	(5)	(6)
Compstat	0.020***	0.019***	0.005	0.003	−0.006	−0.007
	(0.006)	(0.006)	(0.003)	(0.003)	(0.005)	(0.005)
Demographics	No	Yes	No	Yes	No	Yes
Population	No	No	No	No	No	No
Observations	900	900	900	896	900	896

Note: *p<0.1; **p<0.05; ***p<0.01
All regressions include year and agency fixed effects.

Indeed, advocates for Compstat focus largely on the benefits for serious crime reduction in their descriptions of the benefits. (cite Compstat Paradigm) Similarly, proponents of other forms of quantitative management argue that the key outcomes for which bureaucracies are responsible – like educational achievement and public safety – will be most effectively reached with a bureaucracy focused on quantitative targets.

I operationalize serious crime as Part 1 crimes from the Federal Bureau of Investigation's Uniform Crime Reports. These are commonly called 'index crimes' and are typically used as the main measure of serious crime in a city. These include homicide, manslaughter, forcible rape, robbery, aggravated assault, larceny/theft, arson, and motor vehicle theft. If reformers are correct that Compstat effectively reduces serious crime, cities that adopt Compstat should see a reduction in index crimes. I find no such reduction, regardless of whether I examine index crimes as a whole or consider murders, manslaughters, rapes, or other specific serious crimes. The coefficient estimates are close to zero, in dramatic contrast to the estimates for changes in Part 2 crimes.

Table 7: Effect of Compstat on Number of Serious Incidents and Part 1 Arrests

	Part 1 Incidents		Part 1 Arrests	
	(1)	(2)	(3)	(4)
Compstat	214.51	-113.89	2,374.60**	2,359.77**
	(3,282.05)	(3,193.72)	(1,005.90)	(1,006.19)
Number of Part 1 incidents			0.22***	0.22***
			(0.01)	(0.01)
Agency	Yes	Yes	Yes	Yes
Demographics	No	Yes	No	Yes
Population	Yes	Yes	Yes	Yes
Observations	1,081	1,081	1,081	1,081

Note: *p<0.1; **p<0.05; ***p<0.01
All regressions include year and agency fixed effects.

4 Significance

These findings suggest that quantitative management influences policy implementation by street-level bureaucrats, but not always in the ways policymakers might expect. Harnetts report on the city of Oakland argues for adopting quantitative management in a bid to reduce major crimes.(**?**) However, these results suggest that quantitative management does not effectively reduce the incidence of major crimes. Instead, its most consistent effect is to increase arrests for minor crimes, and it may also lead to data manipulation by police. This raises important related questions about the spread of quantitative management to other policy areas, and the potential effects on the decisions of policy implementers and agencies. While advocates for these reforms see them as mechanisms for accountability and transparency, they have distinctive policy consequences that in certain cases actually reduce citizens' ability to monitor and evaluate state activity. Quantitative management in bureaucracy can effectively reduce shirking, but supervisors must be attentive to the problems of task substitution and data manipulation.

These results also highlight a critical methodological issue for scholars studying crime, legal policy, and the state. Crime statistics are jointly produced by whatever criminal behaviors individuals engage in, the enforcement activity of police officers, the reporting behavior of citizens, and the reporting behavior of officers. Researchers using arrest data should treat it as a product of state activity rather than as a proxy for underlying behavior, and should be attentive to the problems created by policy changes around data collection.

Quantitative management alters the relationship between the distribution of criminal behavior and crime statistics by rewarding officers for particular data outcomes, giving them incentives to change enforcement or reporting if they are not able to change criminal behaviors. This is particularly true for Part 2 arrests, where it is well established that legal consequences are not evenly distributed across the population engaged in illegal activity and where there is often no complainant. Black Americans are much more likely to be arrested for marijuana use than whites, even though African Americans are not more likely to use marijuana. (**?**) These results shed important light on how policy changes can influence enforcement patterns, and therefore the distribution of the implemented law.

The effects of quantitative management on policing have long-term effects, both practical and political. Political scientists and legal scholars have found that contact with the carceral state has substantial effects on political engagement and participation, as well as on views of state legitimacy and procedural fairness. Sociologists have found that contact with the criminal justice system or with police leads to distinctive attitudes (Goffman) and behaviors; criminal records, meanwhile, make individuals less able to obtain secure employment. Lerman and Weaver find that stop and frisk is linked to reduced willingness to contact local government non-emergency hotlines (**?**), while Burch finds that living in a neighborhood with high criminal justice contact reduces voting and volunteering (**?**). These normative consequences of increases in Part 2 arrests are deeply troubling. Such feedback effects change who votes, who participates, who is able to seek legal enforcement, and ultimately who has access to both justice and legal power.

Together, the increase in Part 2 arrests and the downgrading of serious crimes have particularly important effects for equal access to the law in the context of selective increases in enforcement of minor crimes. Individuals who feel targeted by law enforcement officers become reluctant to seek law enforcement assistance in resolving serious issues, making them potential targets of violence and leaving them to seek extralegal means for dispute resolution (**??**). At the same time, data manipulation means that crime victims lose access to the formal justice system because the incidents they experience are literally not counted. In American cities, neighborhoods with high levels of violent crime are the same neighborhoods with elevated arrest rates for minor crimes, and where many serious incidents go unsolved (**??**). This combination is not a coincidence. Increased enforcement of minor crimes separate citizens from the state; when that alienation combines with the difficulty of getting police services after a violent incident, people turn to extralegal sources of justice, which only increases the distance between these citizens and the state.

References

Bernstein, David, and Noah Isackson. "The Truth About Chicago's Crime Rates." *Chicago Magazine* http://www.chicagomag.com/Chicago-Magazine/May-2014/Chicago-crime-rates/.

Bratton 2013. "Oakland Crime Reduction Project: Bratton Group Findings and Recommendations." Technical report, Bratton Group, 2013.

Burch, Traci. *Trading democracy for justice: Criminal convictions and the decline of neighborhood political participation*. University of Chicago Press, 2013.

Edwards, Ezekiel, Will Bunting, and Lynda Garcia. "The War on Marijuana in Black and White." Technical report, American Civil Liberties Union, New York, NY, 2013.

Eterno, John, and Eli B. Silverman. *The crime numbers game : management by manipulation*. Boca Raton, FL: CRC Press, 2012.

Goffman, Alice. *On the run : fugitive life in an American city*. University of Chicago Press, 2014.

Harnett 2006. "Crime Fighting in Oakland: an Assessment of the Oakland Police Department." Technical report, Harnett Associates, 2006.

Jeffers, Greg. *Proactive Policing in a Majority Black, Urban Community: Restructuring Patrols, Police-Citizen Interactions, and Residents' Perceptions of the Police*. Ph.D. thesis, Indiana University, Bloomington, 2014.

Leovy, Jill. *Ghettoside*. Spiegel & Grau, 2015.

Lerman, Amy E., and Vesla Weaver. "Staying out of sight? Concentrated policing and local political action." *The ANNALS of the American Academy of Political and Social Science* 651, 1: (2014) 202–219.

Moskos, Peter. *Cop in the hood: my year policing Baltimore's eastern district*. Princeton University Press, 2009.

Poston, Ben, and Joel Rubin. "LAPD Misclassified Nearly 1200 Violent Crimes as Minor Offenses." *Los Angeles Times* .

Rayman, Graham A. *The NYPD Tapes: A Shocking Story of Cops, Cover-ups, and Courage*. Macmillan, 2013.

Venkatesh, Sudhir Alladi. *Off the books: The underground economy of the urban poor*. Harvard University Press, 2006.

Weisburd, David, Stephen D. Mastrofski, Rosann Greenspan, and James J. Willis. *The growth of Compstat in American policing*. Police Foundation Washington, DC, 2004.

Willis, James J., Stephen D. Mastrofski, and David Weisburd. "Making sense of COMPSTAT: A theory-based analysis of organizational change in three police departments." *Law & Society Review* 41, 1: (2007) 147–188.

Yung, Corey Rayburn. "How to Lie with Rape Statistics: America's Hidden Rape Crisis." *Iowa Law Review* 99, 1197.