

Policing Police Brutality: Evaluation of Models of Noncooperative Behavior by Subordinates

John Brehm, Duke University
Scott Gates, Michigan State University

February 24, 1993

Abstract

The incidence of police brutality is a problem of intrinsic interest for political scientists, as a manifestation of the problem of supervision of agents whose behavior is not always observable. The behavior of subordinates in public bureaucracies has drawn greater attention from a variety of social scientists. Two divergent perspectives have dominated this study, coming in the form of principal-agency models (e.g., Bianco and Bates 1990, Miller 1992) or organization theory (e.g., Brehm and Gates 1990, Feldman 1989, Perrow 1987, Simon 1945). These diverse models of supervision in public bureaucracies make divergent predictions about the likelihood of compliance. Our paper evaluates the predictions of several of these models with regard to the incidence of police brutality, using data collected in three cities during the mid-1960s. We develop criteria for evaluation of the models, and use these criteria to specify appropriate maximum likelihood estimators of the incidence of brutality.

An earlier version of this paper was presented at the Midwest Political Science Association meetings, April, 1992. Thanks to Doug Dion, Paul Gronke, Michael Shalev, Peter Fish, Peter Feaver, Jeffrey Hill, Paula Kearns, Mickey Quiñones, Jeffrey Riedinger, Brian Silver, and Robert Worden for comments.

POLICING POLICE BRUTALITY

Recent trials of police officers for beating black motorists in Los Angeles and Detroit has brought national attention back to a political problem of an earlier, more divisive era, which some of us may have thought had been left behind. Police brutality persists as a compelling public policy issue. Police brutality is an exceptional case of defection — defection against the public, against the law, against simple decency. The public policy problem of regulating brutality makes this a worthy subject of analysis *per se*. Further, analysis of brutality provides an opportunity to expand our theoretical understanding of general problems of supervision and control.

In this paper, we examine how the forces that influence defection vary across different forms of police brutality. We approach the issue of police brutality from the perspective of the organization, rather than strictly from the domain of criminal justice.¹ Our objective is to empirically evaluate general issues of organizational defection, cooperation, supervision, and control. These issues play a central role for the implementation of public policy. When considered in the context of police brutality these issues are salient indeed.

Our paper also addresses a common problem in the analysis of social phenomena, through the examination of rare events. Brutality is a rare event relative to all other police activities. Similar rare, but no less important, events happen elsewhere in politics — wars, coups, assassinations, and riots are just a few examples. In this paper, we develop several statistical models to analyze these rare but substantively important events. We utilize a three-pronged analysis of data collected by Reiss and his staff from three large American cities (Black and Reiss 1967, Reiss 1971). The first involves analyzing police brutality with simple probit models. The second approach utilizes a different analytical specification to deal with the problem of rare events, incorporating a Gompertz curve; we refer to

these as gompit models. Unlike the probit or logit models the gompit model is asymmetric, taking into account the relatively rare occurrence of brutality. Our third approach utilizes a censoring model where we differentiate police officers who contemplate abuse from those who do not contemplate abuse. To model this censoring, we borrow the selection bias model developed by Dubin and Rivers (1990). With this three-pronged analysis we hope to gain further understanding of rare dichotomous data.

Our paper proceeds as follows. We start with an explication of a set of propositions regarding the relationship between factors that are theoretically linked to different forms of police brutality. Next, we present our data set, our methods to analyze incidence of brutality, and some first results. Finally, we discuss the implications of our analysis for understanding police behavior in particular and supervision in general.

1 Explication of Propositions

A supervisor's basic task is to encourage compliance and discourage defection by subordinates. While the great preponderance of principal-agency models treat compliance and defection as a simple dichotomy, these are not discrete choices. Further, compliance and defection do not simply fall along a continuum. Forms of cooperative and non-cooperative behavior differ fundamentally. Brehm and Gates (1990) distinguish between two principal forms of defection, shirking and sabotage. A subordinate shirks when he does not work at full capacity. A subordinate engages in sabotage when he actively undermines the productivity of the workforce (engages in negative output). Clearly, police brutality is not a variety of shirking; nor can we classify it as a variant of organizational sabotage.² Nevertheless, brutality is a distinct form of behavior. Put in stronger terms, different forms of brutality are distinct. In the context of organization theory, defection can come in a variety of forms.

Over the decades, two distinct organizational theories regarding supervision and control have emerged. One camp focuses on the uses of rewards and punishments to create an incentive system that serves to motivate self-interested subordinates. This group is primarily composed of economists, and more recently some political scientists and sociologists, who have contributed to the principal-agent literature. The element common to all principal-agent models is that principals are unable to monitor agents' actions or information; the heart of these models involve setting a wage for an

agent without knowing the agent's effort (moral hazard) or ability (adverse selection).³ A second camp is rooted to a more organic perspective of organizational supervision and control, stressing the importance of leadership.⁴ This group is primarily composed of organizational psychologists, sociologists and political scientists (most of whom are considered to be organizational theorists). For the most part, there is little interaction between the two perspectives.

Recently some bridge-building has begun. Miller (1987, 1992), Bianco and Bates (1990), and Brehm and Gates (1990, 1993) each in different ways have started to link these two bodies of literature.⁵ Miller and Bianco and Bates argue that leadership plays a critical role in organizational control from a principal-agent perspective. Brehm and Gates focus on the importance of subordinate predispositions. In this paper, we extend this bridge-building by focusing on the nature of organizational defection in the context of police brutality.

There is a problem in comparing formal models and empirical analysis. The two approaches possess different assumptions and orientations. The solution to this problem, found with increasing regularity in the literature, is to use formal models to derive testable propositions whose verification while not proving the model *per se* do provide empirical support and strengthen the formal model's credibility.⁶ In this manner, we derive a set of propositions that follow from the basic conceptual premises of principal-agent and psychological approaches to the study of organizations. Statistical analyses of these propositions help us understand the nature of supervision and compliance in public agencies, and, in turn, provide some insight into the conceptual basis of these formal models.

In many respects our analysis reflects work done by Friedrich (1980). He also utilized the Black-Reiss data. Friedrich extended Reiss' analysis through a better developed set of propositions. We further extend this line of research by linking our propositions to two diverse approaches to the study of bureaucratic compliance, principal-agent theory and organization theory. We contribute even more so by applying more appropriate models and statistical analyses to these sets of questions regarding police brutality.⁷

1.1 Varieties of Defection

Questions of supervision and control inherently revolve around issues of compliance and defection (or cooperation and non-cooperation). Most works looking at this issue focus on one type of defection, shirking. These works rely on a simple work-shirk dichotomy. Brehm and Gates (1990), on the

other hand, argue that sabotage is a clear case of organizational defection qualitatively distinct from any variety of shirking (motivated either through policy motivations or for leisure). We posit here that police brutality constitutes another distinct form of defection.

Police brutality can be perpetrated in a variety of ways. The most obvious form of brutality is most literally defined as unnecessary force, but verbal brutality (threats and ridicule) exists, too: “the use of undue or unreasonable physical force in some aspect of police action.... Another level of definition for police brutality is verbal or psychological brutality” (Radelet 1986: 191-192). In this paper, we identify brutality as any excessive use of police authority in the officer’s interaction with a civilian.⁸ Unnecessary frisks, ridicule, threats, as well as excessive use of force constitute different forms of police brutality that we examine in this paper.⁹ In the pages that follow we explicitly examine whether different forms of brutality are statistically distinct from one another in their interaction with a common set of independent variables.

Defection Proposition

H1. Different forms of police brutality (unnecessary use of force, unnecessary frisking, verbal ridicule, and threatening behavior) exhibit different statistical relationships with a common set of independent variables.

1.2 Factors that Influence Defection

Several common themes emerge from models of supervision and compliance. Defection and compliance are functions of the payoffs of current or future interactions between subordinates and supervisors. We focus our attention here on four sets of independent variables: *organizational environment*, *work context*, *individual skills and attributes*, and *circumstances of the encounter*. We explicitly employ sets of variables from several levels of organizational analysis for a multi-level approach, where relationships are posited among variables which apply at two or more levels.¹⁰

1.2.1 Organizational Environment

We examine issues of political and organizational environment by differentiating police departments across cities. Reiss selected three cities, Boston, Chicago, and Washington, D.C. for his study, each representing a different type of police department. Chicago was chosen as representative of a “modern, bureaucratically organized department based on systems analysis and centralized command and control” (1971: xi-xii). Boston represented the traditional department with personalized adminis-

tration. "Washington was selected because the department was in the process of professionalizing the staff and moving toward modernization of its command and control system" (1971: xii). Each of these cities' police departments possess different organizational environments, which in turn affect the nature of supervision and control. We propose below that these different organizational environments, in turn, affect the propensity for police officers to engage in brutal activity.

Organizational Environment Propositions

- H2. Different political and organizational environments can be distinguished across cities which in turn influences the frequency of defection (unnecessary use of force, unnecessary frisking, verbal ridicule, and threatening behavior).

1.2.2 Work Context

The organizational psychology literature emphasizes the importance of the work context. Particularly relevant to a work context is the degree of cohesion within an organization. Behavioral conformity among subordinates is central to such notions of cohesion. The issue of conformity in a police agency is closely related to the notion of a police culture. A police culture is the informal organization that operates within a police agency (Brown 1981; Manning 1982; Van Maanen 1983). Brehm and Gates (1991) suggests that this conformity arises out of the interconnectedness of subordinates. This is consistent with Brown's observation:

The police culture demands of a patrolman unstinting loyalty to his fellow officers, and he receives, in return, protection and honor: a place to assuage real and imagined wrongs inflicted by a (presumably) hostile public; safety from aggressive administrators and supervisors; and the emotional support required to perform a difficult task (1981: 83).

Conformity among the subordinates toward policy varies by the network of associations among subordinates. In police organizations where brutality is considered to be acceptable among a network of subordinates, we would expect to see a greater propensity for brutality.¹¹ We examine the role of police conformity by looking at expressions of solidarity exhibited in police agencies. Organizational disaffection is also hypothesized to be related to the frequency of police brutality. As Reiss contends:

Police resort to harassment under conditions where they are caught between their own, or others', expectations that they control unlawful conduct, while other levels of the system thwart such enforcement by failing to treat their arrests of citizens seriously. Actually, one can predict that harassment will become police policy or unofficial practice whenever citizen influence compels the police to make arrests that are systematically disregarded by others in the criminal justice system. The harassment of juveniles, minorities, and even of those engaged in vice can be predicted, based on these factors (1971: 138).

Specific propositions regarding organizational cohesion and work context follow.

Work Context Propositions

H3. Subordinate conformity (solidarity) is related to the frequency of defection.

H4. Subordinate disaffection from an organization is positively related to the frequency of defection.

1.2.3 Individual Skills and Attributes

Another common feature of organizational psychology models are individual attributes. Of particular importance is the notion of professionalism. We posit here that a professional demeanor serves to limit defection. In the reforms to the police force in many cities (including Los Angeles), supervisors emphasized the importance of 'professionalism' as a way to regulate the behavior of officers.¹² This is very much in sync with those principal-agency approaches which regard adverse selection as a threat to cooperation by subordinates. Reiss (1971) is quite explicit about the professional status of police officers: "The police in America belong to one of the few occupations that includes all of the essential elements to qualify as a profession. They possess the power of coercive authority, and through their power to arrest and book for offenses, they control the fate of 'clients'" (pg. 123). One hypothesis is simply that professionalism decreases the probability of police brutality.

Other skills and attributes of the officers may affect brutality. Officers vary in their degree of intellectual involvement with the job. Some officers are reflective and alert while others are sluggish and unreflective. To the extent that brutality is a product of intense, ill-considered passions, we would expect that intellectual involvement should depress brutality.

Officers also vary in their human relations skills. Some officers may be antagonistic or abrasive, others will be responsive and concerned. If brutality arises because officers lose control, or because antagonism between officers and civilians escalates, then officers with better human relations skills should be less likely to resort to brutality.

Note that these three aspects of the officer's skills and attributes argue for different policies. It is entirely possible that police schools could train officers to become more professional. It may also be possible to train officers to improve their human relations skills. We are skeptical as to whether any program could improve officer's intellectual involvement with the job, short of extensive screening at the hiring stage.

Individual Skills and Attributes Propositions

- H5. Subordinate professionalism, intellectual involvement and human relations skills are negatively related to the frequency of defection.

1.2.4 Circumstances of Encounter

The specific circumstances of a police encounter may also influence the propensity for a subordinate officer to engage in brutal activity.¹³ Three features of the Rodney King case are quite relevant: the number of other officers, the differences of races, and the character of the alleged offense.

The number of officers might influence the probability that an officer would defect, but the direction of influence is quite unclear. If police officers resort to brutality because they need to assert control, we would expect that officers would be less likely to brutalize civilians with a large number of other officers at the encounter. Conversely, officers might be tempted to brutalize as a demonstration of power to their peers. We remain agnostic about the direction of the relationship between the number of officers and frequency of defection.¹⁴

A salient charge against the police officers, indeed the LAPD and its chief, is that the officers acted out of racial animosity against King. Charges of racism underlie the long-standing concerns about police brutality (Wilson 1968), (Reiss 1971). This is a testable hypothesis in principle; unfortunately, in the data at hand, the race of civilians is suppressed in the collection.

Finally, the type of alleged crime may also affect brutality. Officers investigating felonies (not all of which are violent) may be more likely to brutalize than officers investigating misdemeanors. We would not ordinarily expect brutality to occur when officers investigate relatively minor offenses. Officers investigating violent crimes (not all of which are felonies) may be more likely to brutalize civilians than officers who are investigating non-violent crimes.

Circumstances of Encounter Propositions

- H6. The number of officers present at an encounter is related to the frequency of defection.
H7. A felony crime is more likely to involve police brutality.
H8. A violent crime is more likely to involve police brutality.

1.2.5 Predispositions

Brehm and Gates (1990, 1993) demonstrate that defection varies by the attitudes of the subordinates toward a policy. In the context of a police agency, attitudes towards particular policies may

influence how officers carry out their duties (Friedrich 1980; Brown 1981). The frequency with which subordinates comply is positively related to subordinates' affect and initial responses to a policy. It follows that the frequency of defection is positively related to the subordinate's predisposition towards defection. In this specific case a police officer's attitude towards brutality is seen to influence a subordinate officer's propensity to engage in brutality.

Officers express a wide variation of likes and dislikes about their jobs (as would any set of employees). One might distinguish between likes and dislikes in terms of broad references. *Functional* likes and dislikes refer to how the officer feels about the duties of being a cop: an officer who dislikes arresting people expresses a functional dislike, while the officer who likes the idea of enforcing the law expresses a functional like. *Solidary* likes and dislikes address the officer's feelings about fellow officers.

A previous analysis of shirking behavior found that likes and dislikes strongly influenced the extent to which officers shirked (Brehm and Gates 1993). We expect a similar relationship to emerge in this analysis of police brutality, although we remain agnostic about the direction of the relationship. If Reiss' argument about the reasons for brutality holds true, then we would expect that officers who express a political dislike for the job would be more likely to brutalize than other officers.

Predisposition Propositions

- H9. A subordinate's predisposition towards brutality is positively related to the propensity of the individual to engage in brutality.
- H10. A subordinate's likes and dislikes about the job are related to the propensity of the individual to brutalize citizens.

From these theoretical observations, we now proceed to evaluate models of organizational control in the specific instance of police brutality. While we focus on supervision in a police agency, our interest extends to the general issue of supervision and control in any organization. We first describe the data that we use to examine this problem. The next step is to develop our three general analytical approaches for evaluating issues of supervision and control. The general task is to determine whether police brutality can be controlled by supervisors, by fellow subordinates, or that it is simply a very rare event.

2 Description of the Data

Over twenty years ago, Albert Reiss and his staff interviewed and observed officers in Washington, Boston, and Chicago in 1966 (Reiss 1971). This data set affords a unique opportunity to examine the behavior of 'street level bureaucrats' in their encounters with citizens. In addition to the observational data, we also draw upon the interviews with the officers about their feelings toward their jobs, their superior officers, and the public.

There are three separate data collections that we merge together in order to examine our propositions. Two of these are on the basis of observer's records: the 'runs' and 'encounters' files. The 'encounters' file treats each service-initiated¹⁵ police-citizen encounter as a single observation. We have recorded in the 'encounters' data the officer's attributes (e.g., rank, age, race, assignment, precinct), characteristics of the encounter (e.g., number of officers, whether the crime is a felony or violent), and the types of interactions (e.g., whether unnecessary force, unnecessary frisks, ridicule or threats were used). The 'runs' file represents an entire shift for the officer.¹⁶ Each run contains multiple encounters. It is relatively straightforward to merge the encounters and runs data.

Several different types of brutality and harassment are recorded in this data set. These acts include the use of unnecessary force, unnecessary frisks, threats of violence, and ridicule. The frequency of these incidences of brutality in terms of the total number of citizen-police encounters is tiny: .2% for unnecessary force, .7% for unnecessary frisks, 1% for ridicule, and 6% for threats. These are our dependent variables for examining police brutality. We also utilize information regarding officers' attitudes towards physical abuse of suspects. This information is used to differentiate officers according to their propensity to be physically brutal.¹⁷

It should be immediately apparent to the reader that there is a significant risk of a 'Hawthorne effect' causing the officers to alter their behavior under the scrutiny of an observer. One would expect that the officers would react to being observed by restraining any inclinations they might have to engage in brutality. In fact, it is noteworthy that any brutality took place in front of these observers. For two reasons, we argue that it is well worthwhile to proceed with the analysis of these observations, despite the obtrusive nature of the measurement of our dependent variables. First, while our suspicions are that the officers suppressed non-cooperative behavior under observation, these suspicions are probably an exaggeration. In our study of police working and shirking, we found

a surprisingly high incidence of shirking despite the presence of an observer. Other analysts of police behavior (Worden 1989, Black 1980, Brown 1981) report that officers are relatively unconstrained by observers. Second, if this is a 'censored data' problem, the effect of the censoring on the dependent variable produces attenuation of the coefficients. In other words, the effect of the observers is likely to undermine the findings we report here by reducing the magnitude of the coefficients toward zero, not to exaggerate them. While the incidents of brutality are rare, if one considers all the police-public encounters, there are an amazing number of cases involving brutality.

The data also provides information specific to the encounter; this includes such information as the number of citizens present, the number of officers present, whether the crime is a felony or whether it was violent. As in our previous analysis of shirking, we also exploit answers to open-ended questions about the officer's likes and dislikes for their job. If the officer mentioned aspects of the job directly related to performance as a job officer (e.g., interest in law enforcement, making society better, possibility of false arrest), we code the reason as *functional*. If the respondent referred to his/her relations to other officers (such as to friends or family), we code the reason as *solidary*.

We lack several significant measures. The most conspicuous absence is of any measure of the supervisor's level of tolerance for defection or the level of sanction for defection. Nonetheless, we believe that this data set allows us to make several observations about police brutality, which, in turn, aids our understanding of organizational defection, supervision, and control.

3 General Models for Statistical Analysis of Non-cooperative Behavior

In this section, we outline three separate approaches toward analysis of the incidence of police brutality in several guises (unnecessary force, unnecessary frisks, ridicule and threats). Because each case represents a single police-citizen encounter, each of the dependent variables is dichotomous (e.g., unnecessary force used or not used). As long as we make two assumptions, the maximum likelihood function describing brutality is relatively simple. First, each observation must be independent (this assumption permits us to use the simple product of the density functions as the likelihood). This first assumption would be violated if brutality in any one instance (say earlier in a shift) were related to brutality in other instances (perhaps later in a shift). The occasions of brutality are sufficiently sparse that this assumption appears to be reasonable. Second, the process generating brutality must

be homogeneous (identically distributed) across police-citizen encounters (this assumption permits us to describe the probability of police brutality as a single parameter). This second assumption would be violated if the probability of brutality changed across the course of a shift. Both assumptions are reasonable, and quite commonly, if only implicitly, employed in dichotomous dependent variable models.

Using the two assumptions, we can describe the density function for the incidence of police brutality as a Bernoulli probability distribution:

$$\mathcal{L}(\pi|y_i) = \prod_{i=1}^N \Pr(y_i = \text{brutality}) \quad (1)$$

$$= \prod_{i=1}^N \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \quad (2)$$

This Bernoulli probability distribution is conventionally used for probit and logit analyses of dichotomous dependent variables. In more formal terms, we use a ‘link function’ (see McCullagh and Nelder 1983) to reparameterize π in terms of its ‘systematic component’:

$$\pi_i = F(\beta X_i) \quad (3)$$

Where the function $F()$ is *any* monotonic function of X , bounded between 0 and 1. In the probit case, we substitute the cumulative normal density function, $\Phi()$, for $F()$; in the logit case, we substitute the log-odds function, $\log(x/1-x)$, for $F()$. (Below, we suggest an alternative link function, the Gompertz curve, that may provide some additional power in this problem).

For each of the analyses in the next section, one analytical approach we take is to use the probit link function. This approach has the virtue of being statistically appropriate to the underlying theoretical problem, as well as conveniently implemented in a wide range of statistical packages. The coefficients reported in the probit analyses are expressed in Z -scores, and are interactive with respect to the values of the other variables (X) and parameters (β) in the link function. We employ ‘1st Differences’ as a simple means to gauge the effect of changes in the explanatory variables:

$$\mathcal{D} = \Phi(\beta X_{\max}) - \Phi(\beta X_{\text{mean}}) \quad (4)$$

Where Φ represents the cumulative normal density function, X_{mean} represents all independent variables at their respective means, X_{\max} represents all independent variables at their respective means *except* for the variable of interest, which is set to the maximum. In this way, the 1st Difference provides a measure of the maximum effect of each independent variable.

The probit link function, however, may not be the most appropriate function for analyzing brutality. The probit link is derived from the standard normal distribution, which is symmetric. The probability process generating brutality or the absence of brutality is most likely not symmetric: thankfully, the vast preponderance of encounters between police and citizens are not brutal. Given the rare occurrence of brutality in our data, one might argue that brutality is an especially exceptional circumstance, and that the probability of brutality must be driven far across some threshold before the encounter becomes brutal. Look at the generalized linear model ($X\beta$) embedded in the link function (3): one might hypothesize that the effect of a unit change in $X\beta$ in the negative (non-brutal) direction would be less than a unit change to $X\beta$ in the positive (brutal) direction.

[Figure 1 about here]

In order to address the possibility of asymmetry of the effects of $X\beta$ on π , we suggest an alternative link function — the Gompertz curve. Figure 1 displays the Gompertz curve and cumulative normal density function (probit) plotted for the same values of $X\beta$. The Gompertz curve is much slower to approach the upper asymptote (1) than the probit function. It is for this reason that some analysts suggest that the Gompertz curve is better suited to the analysis of ‘rare’ dichotomous events than probit (Holden and Overall 1987, McDonald and Butler 1990).¹⁸ The link function for the Gompertz curve is:

$$\pi_i = e^{-e^{-\beta X_i}} \quad (5)$$

(The Gompertz curve may be directly incorporated in the computation of 1st differences, substituting the above function for the cumulative normal density function, $\Phi()$, in equation (4) above.) Following the tag Goldberger (1964) assigned to ‘Tobin’s probit’ (tobit), we refer to the use of the Gompertz curve as a link function in the binomial distribution as a ‘gompit’ (in order to distinguish between the use of Gompertz curves in hazard rate analysis).

We expect that the gompits will identify more statistically significant independent variables than the probit analysis. Moreover, we expect that the 1st differences from the gompits will be larger in absolute value than those produced by the probits.

Both the probit and gompit methods treat the incidence of police brutality as independent events. In all of the probit and gompit analyses to follow, the constant term will be the most sizable and statistically significant ‘variable’ in the models. This makes sense: we know that the incidence of

brutality is extremely infrequent in our sample.

But it does beg the question of why brutality is so rare: is it because there are very few circumstances in which brutality is remotely possible? or is it because principals have been able to control the behavior of the agents? One probably would not expect brutality in most citizen-police encounters, perhaps the overwhelming majority of such encounters. The difference between these two questions is substantively critical: if the incidence of brutality is solely a function of the predispositions of the officers to act in a brutal fashion in certain encounters, then the principal-agent problem becomes irrelevant. We might expect brutality to become more frequent during crime waves, or as the recruitment patterns for officers change, or as peripheral circumstances change, but the characteristics of reward and punishment and of supervision might make very little difference. Conversely, the principal-agent problem of regulating police brutality is not a problem if the officers do not express an inclination to brutalize citizens.

[Figure 2 about here]

If we treat the principal-agent problem of regulating brutality as arising *only* in those circumstances where an officer has a predisposition to brutalize, then we transform the statistical problem from the simple Bernoulli distribution to a problem of selection bias in dichotomous models. Figure 2 displays the selection bias underlying the decision tree. The first branching in the tree discriminates between circumstances where a police officer is or is not predisposed to brutalize the citizen. The second branching in the tree is whether the officer, in fact, does brutalize the citizen or restrains him/herself. With both the probit and Gompit models, we fold together all those encounters where the officer does not wish to brutalize the citizen as well as the situations where the officer does want to brutalize the citizen, but holds back. Our analysis of brutality will be censored by the first stage, the question of the predisposition of the officers. The substantive importance of the censoring is that we will underestimate the effect of every variable on the incidence of brutality.

We have a crude measure of the predisposition of the officers in the record of whether officers avoided action because they felt such action was prohibited. The language of the codebook is instructive: "wanted to beat a little more on street but couldn't as others around"; "would like to bang the punks heads together"; "knock him out and make it easier for everyone"; "I'd like to show them." If we treat this first variable as a dichotomy, coded 1 for officers who used excessive force or expressed supervisory prohibition to using excessive force, 0 otherwise, we have a crude measure of

the first selection stage. The second, outcome stage, of this process is coded 1 for officers who used excessive force, 0 for officers who felt such actions were prohibited, and missing data for all others.

Dubin and Rivers (1990) develop the mechanics for handling selection bias in logit and probit models. The basic log-likelihood function resembles the Bernoulli log-likelihood:

$$\mathcal{L}(\beta_1, \beta_2, \rho) = \sum_{i=1}^n y_{1i}(y_{1i} \log(\Pr(\text{brutal})) + (1 - y_{1i}) \log(\Pr(\text{restraint}))) \\ + (1 - y_{2i}) \log(\Pr(\text{no brutality})) \quad (6)$$

Where $\Pr(\text{brutal})$ is the joint probability that officers want to brutalize the citizen *and* actually brutalize that citizen, $\Pr(\text{restraint})$ is the joint probability that officers want to brutalize the citizen *but* restrain themselves from following through, and $\Pr(\text{no brutality})$ is the probability that the officer does not want to brutalize the citizen. The variables are y_1 , incidence of brutality given predisposition to brutalize, and y_2 , predisposition to brutalize. We seek the parameters β_2 , the effect of the variables affecting predisposition to brutalize, β_1 , the effect of the variables affecting the officers' following through on the predisposition, and ρ , the correlation between the error terms in the two stages of the decision tree.

Dubin and Rivers (1990) offer two different link functions for the probabilities in (6). In this paper, we use the probit link functions. The joint cumulative density function for the error terms (u_1, u_2) to both stages of the decision tree (Figure 2)¹⁹:

$$F(u_1, u_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \times \int_{-\infty}^{u_2} \int_{-\infty}^{u_1} e^{-\frac{1}{2(1-\rho^2)}(u_1^2 - 2\rho u_1 u_2 + u_2^2)} du_1 du_2 \quad (7)$$

The probabilities in (6) are:

$$\Pr(\text{brutal}) = F(\beta_1' x_1, \beta_2' x_2, \rho) \quad (8)$$

$$\Pr(\text{restraint}) = \Phi(\beta_2' x_2) - \Pr(\text{brutal}) \quad (9)$$

$$\Pr(\text{no brutality}) = 1 - \Phi(\beta_2' x_2) \quad (10)$$

The model may be estimated in the appropriate maximum likelihood package (e.g., SST, Gauss, SHAZAM).

This same kind of selection process is likely to be at work with all the manifestations of police brutality. Due to the nature of the data at hand, we are only able to estimate this selection bias model with the use of unnecessary force. The record of 'prohibited actions' does not include mentions

of the other forms of brutality, and it is the basis for our coding of the first stage of the decision tree. Were one to consider a replication of Reiss' original data collection, we would strongly encourage the development of measures of restraint for the alternate forms of brutality.

Armed with these estimation techniques, our next enterprise is to produce estimates of the effect of various characteristics of the precinct, the officer, supervision, and the alleged crimes on the incidence of brutality. We turn to an examination of the probit and gompit estimates for all four forms of brutality, followed by an examination of the selection bias approach to the use of unnecessary force.

4 Probit and Gompit Results

In this first section of results, we examine the probit and gompit models of incidence of police brutality for each of the four manifestations of brutality (unnecessary force, unnecessary frisk, ridicule, threat). Our selection of variables replicates a previous analysis of shirking by police officers (making use of a different part of the same data set) (Brehm and Gates 1993). Coding for these variables is as we describe in section 2 of this paper. Tables 1–4 display the probit and gompit results for the same model applied to each of the four forms of police brutality (unnecessary force, unnecessary frisks, ridicule, and threats).

[Tables 1–4 about here.]

With these analyses, we want to attend to both the statistical and substantive significance of the coefficients. The first observation one might make stems from the magnitude of the constant term, which is substantially negative in all cases. Hence, the naive hypothesis that brutality does not occur clearly overwhelms the data. One traditional measure of the 'fit' of the model is a comparison of how well the model predicts cases relative to the modal value. The size of the constant term relative to the substantive coefficients means that all of these models lead to predictions that are identical to the naive hypothesis.

An alternative measure of goodness of fit is to compare the log-likelihood for the substantive model against the null model: minus two times the ratio of the log-likelihoods is distributed as a χ^2 with 11 degrees of freedom. In each of the four models, the χ^2 is quite significant. Unfortunately, the striking skewness of the dependent variable in the four models means that this log-likelihood ratio test is approximately equal to the number of cases in the sample (cf McCullagh and Nelder

1983). Surely one would not regard sample size as a measure of goodness of fit! Furthermore, this log-likelihood ratio test cannot discriminate between the probit and gompit estimates: the link functions do not affect the likelihood. Indeed, one of the virtues of maximum likelihood estimation is invariance to reparameterization, and so the log-likelihoods for probit and gompit models will be identical (with the occasional slippage due to rounding error).

Several of the formal models predict conformity among subordinate behavior. Such principal-agent models as Bianco and Bates' production-by-teams model (1990) and organizational theory models as Brehm and Gates' imitation model (1990) imply that we will see convergent behavior across subordinates in the same workforce. There is also strong evidence from social psychology that individuals in conditions of uncertainty look toward 'like others' in order to identify appropriate behavior.²⁰ And with our analysis of police shirking, we found striking convergence among officers in the same city in the levels of shirking. Does the same pattern persist for the four varieties of brutality?

The answer varies somewhat depending whether one looks at the probit or gompit analyses. The probit coefficients for the officer's city in the analyses of unnecessary force and unnecessary frisks are not statistically distinguishable from zero for both Boston and Chicago. Furthermore, only the probit coefficient for officers working in Chicago is statistically significant in the model for incidence of threats. (Both probit coefficients are significant at $p < .05$ in the model for incidence of ridicule). However, at least one city coefficient in the gompit model of each form of brutality is statistically significant, and usually both. Boston and Chicago officers are more likely to engage in unnecessary force or employ threats, while less likely to use unnecessary frisks or ridicule, than the Washington, D.C. officers. The first differences for all variables for both probit and gompit estimates are all small. (Because the constant term is substantially negative, the base probability on which we evaluate the probit and gompit estimates is very small, and hence only small changes are possible with the first differences). Nonetheless, these first differences are substantively important: when one considers the hundreds of thousands of encounters between police and citizens across the country each year, first differences of .8% (as in use of ridicule by Boston officers) refer to hundreds of cases of brutality). Tables 1-4 thus confirm hypothesis H2, differences by city.

The differences between the probit and gompit estimates leads to a second general observation: the gompit estimates are more likely to be statistically significant and have first differences that

are greater in absolute value than the probit estimates. How should the analyst interpret this observation? On what grounds would we prefer the gompit estimates over the probit estimates? (Simply being able to produce significant findings is hardly adequate). The risk in this circumstance lies in the nature of hypothesis testing: we are never in the position, formally, of 'accepting' the null hypothesis, only of 'failing to reject' the null. In the current circumstances, the risk of falsely failing to reject the null is the risk of a 'false negative,' or failing to find an effect that would become apparent with a larger sample size. In public policy terms, the question is one of where we appropriately direct our prevention efforts. If our goal is to constrain police behavior such that brutality is not possible, then we would prefer to use statistical methods that detect important factors that other methods miss. If our goal is to tolerate 'minor' deviations by officers, and to encourage active use of appropriate police behavior, then we should prefer more stringent statistical tests. This is an active aspect of the policy debate over the rights of suspects and the authority of police, and a debate that can not be settled in the present paper.²¹

The imitation and enhanced principal agent models of subordinate compliance Brehm and Gates (1990) make the strong claim that the predispositions of the subordinates are what determine compliance. Likewise, the problem of 'adverse selection' in the principal-agent context emphasizes the predispositions of subordinates prior to entering into the game. In our analysis of shirking by police officers, we were surprised to discover that human relations skills, professionalism and intellectual style did *not* affect the incidence of shirking.

The results of the present analyses of incidence of the different forms of brutality largely bear out observations about shirking seen in Brehm and Gates (1993). In general, in neither the probit or gompit estimates of the models for the four forms of brutality, the coefficients on human relations skills, intellectual style and professionalism are not distinguishable from zero. There are some important exceptions to this general rule, and worthy of comment. Professionalism is positively related to both the incidence of unnecessary frisks and of use of ridicule. In other words, the more professional officers were more likely to engage in these forms of unprofessional behavior. One explanation might be that the more professional officers learn to resort to these lesser forms of brutality instead of the more violent forms. However, one would then expect that professionalism would be negatively associated with unnecessary force or threats, which is not the case, at this level of analysis. (As the reader will see in the next section, this pattern reverses in the censored probit

approach).

We also examine the relationship between the circumstances of the encounter and the incidence of brutality. We know the number of officers on the scene, and the type of alleged crime to be investigated. Here, there are some strong patterns across all four forms of brutality. In general, felonies did not encourage more brutality. Although the probit and gompit coefficients are positive for all but incidence of threats, they do not meet conventional statistical significance levels. Violent crimes are positively associated with use of unnecessary force, ridicule, and threats, but negatively associated with incidence of unnecessary frisks. The number of officers at the scene is positively associated with all four forms of brutality. (Recall that we were agnostic in our hypotheses about the direction of the relationship between number of officers and the incidence of brutality).

The maximum number of officers in the data collection was 99. Obviously, any encounter between police and citizens with 99 officers is a very unusual situation. If we were at this extreme, the probability of threats increases by about 10%, unnecessary force by 1.3%, unnecessary frisks by 3% and ridicule by 6%. What are we measuring with the number of officers? We are probably accounting for two phenomena simultaneously: the presence of others observing the action of the officer, as well as the likelihood that a very unusual situation is at work. The present analysis can not discriminate between these sharply divergent reasons for the magnitude of the first difference.

In an analysis of shirking by police officers, Brehm and Gates (1993) found that the officers' likes and dislikes for being an officer strongly affected the extent to which they shirked. Officers who expressed a functional dislike for the job (e.g., disliked the duties of being an officer) shirked 72% more than the average officer. Conversely, officers who expressed a solidary like for the job (e.g., enjoyed the company of the other officers) shirked 63% less frequently. In the context of incidence of brutality, one common explanation (see Reiss, above) is that officers resort to brutality when they feel stymied by the courts.²² If so, we should expect that those officers who express dislikes about the politics of being an officer would be more likely to resort to brutality.

Although singularities in the overlap among the likes and dislikes with the incidence of brutality prevent us from including all the likes and dislikes officers expressed for their jobs, the present study does include the effect of solidary and functional likes on the four forms of brutality. Solidary and functional likes are *negatively* associated with unnecessary force, unnecessary frisks, and threats. The interpretation is that officers who like their jobs for reasons associated with connections to

fellow officers or for the duties of being a police officer are less likely to engage in these more violent forms of brutality. The patterns reverse for the incidence of ridicule: officers who express functional likes for the job are more likely to ridicule citizens than those who do not.

These four models provide a glimpse at the factors influencing the incidence of police brutality in these three cities in the 1960s. Several points should be clear. First, defection is not an amorphous black box of subordinate behavior; there are distinctly different forms of defection, which are governed by very different factors. Even though the preponderance of formal models treats non-cooperative behavior as one, uniform response, there are clearly major differences in the empirical explanations among different forms of defection, only one category of non-cooperative behavior in its own right. Shirking, sabotage, and brutality as different forms of defection differ, as do the varied manifestations of brutality, in terms of explanatory factors. In fact, we can obtain a direct statement of the non-interchangeability of the different modes of brutality via the definition of conditional probability. The percentage of encounters between police and citizen that include either unnecessary frisks, ridicule or threats is 7.32%. The percentage of encounters where unnecessary force is used and any of the other forms also occurs is .05%. Therefore, the probability of unnecessary force given an occurrence of one of the other modes of brutality is $.0005/.0732$, or .7%. While all forms of police brutality are relatively rare, far rarer still are occasions when the forms of brutality overlap. The variation in propensity for the different forms of brutality by city confirms H1, that different modes of brutality are influenced by different factors. We take this finding as indirect evidence of either different incentives among the officers for different forms of defection, or different structure of the principal-agency game itself.

We may take this to refer to differences in the preferences of subordinates, differences in the payoffs for the game, or differences in the cognitive processing of the decisions to work, shirk or sabotage. In any application of formal models to empirical data, there will be some slippage. We must consider the varied forms of defection, and not inappropriately lump together all non-productive behavior.²³

Secondly, the forms of brutality are rare, but still important. In the hundreds of thousands of interactions between police and citizens across the country every year, tiny frequencies make for a substantial number of cases. The probit and gompit analyses provide some clues as to the causes of police brutality: conformity among police forces, absence of professionalism, absence of solidary

allegiances, type of assignment and number of other officers. We would clearly prefer to have better measures of the principal-agency features of the police forces, but these initial findings are still useful. The next question is to ask why is brutality so rare? Changing the question brings us to the domain of censored data, and much closer to the problem of principal-agency.

5 Censored Probit Estimates for Unnecessary Force

This section treats the incidence of unnecessary force as a censored data problem. That is, we examine the incidence of unnecessary force *given* a predisposition by the officer to use unnecessary force. The two stages of the censoring problem (see Figure 2) are 1) does the officer express a desire to brutalize the citizen? and 2) if so, does the officer follow through on that desire and actually use unnecessary force?²⁴

We need a set of independent variables to predict the first stage in this selection process.²⁵ Three measures appear to do reasonably well in predicting whether an officer expresses a desire to brutalize a citizen: *authoritarian* demeanor, *number of officers* at the encounter, and whether the alleged crime was *violent*. The top half of Table 5 reports the probit estimates for the selection process.

[Table 5 about here]

A couple of points are worth noting in the selection model. The authoritarian demeanor of the officer is far and away the most important predictor of a predisposition to brutalize. This variable comes from the encounter data set itself, and is probably endogenous to some degree with respect to brutality (i.e., observers might classify officers who brutalize citizens as authoritarian, by definition). The coefficient on violent crime is also interesting: as one might expect, officers investigating a violent crime have a predisposition that (marginally) favors brutality over officers who are not investigating violent crime. And the coefficient on the number of officers suggests that officer's predisposition to brutalize civilians is enhanced by the presence of other officers, not diminished.

The lower half of the model replicates the model in table 1, reporting both the probit and censored probit coefficients. There are sharp and substantively significant differences between the standard probit results and the censored probit estimates.²⁶

Unlike the standard probit model, both coefficients for the officer's city are now statistically significant. Again, the first differences remain somewhat small: officers in Boston were 1% and

those in Chicago 2% more likely to use unnecessary force given a predisposition to brutalize.

Some of the more important substantive findings from the censored probit approach appear in the effect of the officer's attributes (human relations skills, intellectual style, and professionalism). Professionalism appears to have an important effect on the officer's decision to follow through on a predisposition to brutalize, and in the direction one might anticipate. More professional officers were less likely to use unnecessary force. Note that this is opposite to the findings in Tables 1-3, where more professional officers were more likely to resort to unnecessary force, unnecessary frisks, or ridicule than less professional officers. The implication we draw from the censored probit is that we were confounding a desire to brutalize suspects with an actual incidence of brutality. Similarly, officers with better human relations skills were also less likely to use unnecessary force. Here, the 1st differences are still small, but theoretically significant in their implications. In contrast, officers who were more intellectually engaged were more likely to brutalize civilians than those who were less engaged. Whereas the simple probits tended to disconfirm H5, that professionalism reduces brutality, the censored probits confirm.

The circumstances of the encounter between police and citizen significantly affected the likelihood of use of unnecessary force. Officers investigating a violent crime were about 7% more likely to follow through on their predispositions to brutalize, by the first differences. Officers who were investigating a felony were also slightly more likely to follow through on predispositions to brutalize. But the most striking effect is of the number of officers. If there were the maximum number of officers (99) on the scene, and the officer under observation expressed a predisposition to brutalize, then it is almost a certainty that the officer followed through with that predisposition: the 1st difference is 84%, meaning an 84% increase in the probability of unnecessary force over the average. In Figure 3, we display the predicted probability of unnecessary force per officer (holding the other variables constant at their mean). With twelve other officers on the scene, the probability of unnecessary force given a predisposition to brutalize was about 33%. (We take these results as confirmation of H6-H8). Again, we are unable to separate the effects of social proof (due to the presence of like others) from the admittedly unusual circumstances where there would be so many officers.

[Figure 3 about here.]

Finally we turn to the effect of the officer's solidary and functional likes for the job. Unlike the simple probit coefficients, both solidary and functional likes are positively related to the officer's

vast number of such encounters, police brutality is far too common. Consider the use of unnecessary force, which occurs in less than 1% of all the encounters in our data. By dint of the hundreds of thousands of encounters each year, we could infer several hundred, if not several thousand, incidences of unnecessary force each year. Such rare, but substantively significant, events pervade many aspects of politics. Our paper develops two distinct approaches. With the gompit analysis, we exploit the asymmetry in the occurrences of brutality to a gain purchase on the possible magnitude of effects of explanatory variables. The gompit analysis is agnostic about the reasons for the rarity of brutality. Our second approach argues that we have to ask why brutality is so rare. If one proceeds from the assumption that brutality is rare because these incidences when officers conceive of brutality as an option are rare, then we arrive at the censored probit method.

The most significant of our substantive findings is that professionalism appears to work as a safeguard against brutality. If the question is not why officers become angry enough to brutalize civilians, but why they follow through, professionalism appears to aid in officer's self-restraint. Unfortunately, the magnitude of the effect of professionalism, as seen in this data, appears to somewhat marginal. There appear to be significant variations in the incidence of brutality by city, suggesting that the officers arrive at some level of conformity in response to police-citizen encounters. Unfortunately, we have no real way of assessing whether this conformity stems from solidarity with peers or supervisory tolerance for brutality.

Finally, we look towards the events in Los Angeles and Detroit, and ask which of our findings illuminates the situation. Of course, our data are from a much earlier time period and in different cities. The number of officers at the scene of the Rodney King beating suggests that the beating was inevitable *given that the officers had a predisposition to brutalize*. We do not know whether the officers arrived at the scene with such a predisposition. Nor do we know whether the effect of the number of officers comes because of the unusual nature of situation with so many officers present, or because so many fellow officers pressured the police into brutality.

¹A rich tradition of applying organization theory to issues of criminal justice exists; Wilson (1968), Manning (1977), and Brown (1981) constitute prominent examples. Black and Reiss (1967), Reiss (1971), and Friedrich (1980) also employ organizational factors in their analysis of police brutality. We reanalyze the Black and Reiss data, more explicitly drawing from organization theory and utilizing more appropriate statistical techniques.

²We are sorely tempted to regard brutality as negative output — surely, from the citizen's point of view, we do not condone criminal behavior by the police. But several scholars of police behavior (see Reiss 1971) as well as many of the officers themselves, regard police brutality as the exercise of unusual means to accomplish the social goal of punishing criminals. In this manner, police brutality could be actually condoned by superiors and may constitute compliance rather than defection. Unfortunately, we have no data to evaluate supervisors' attitudes towards police brutality. For the purposes of evaluation of principal-agency in this paper, we assume that brutality is illegal and not condoned by police departments.

³The central concern of principal-agency models is the extent to which the parameters of supervision influence the incidence of defection. Do rewards and punishments affect compliance? Does the supervisor's ability to target incentives or to closely monitor subordinates determine the subordinate's likelihood of cooperation? Prominent early principal-agent works include Alchian and Demsetz (1972), Holmstrom (1979), and Grossman and Hart (1983). Similarly, in an earlier era, Fayol (1916), Gulick (1937), and Taylor (1911) pondered the optimum span of control, the chain of command, and division of responsibilities. These concerns are obviously pertinent to police brutality. We are limited by our data in examining these questions. Rewards and punishment play an integral role in any principal-agent model. Unfortunately, the data set we use does not provide adequate information which can be used to analyze the relationship between the distribution of rewards and benefits and the propensity for subordinates to defect. In the case of police brutality, missing information regarding punishments is especially unfortunate.

⁴See for example Graen, Novak, and Sommerkamp (1982), Graen and Scandura (1987), Kaplan (1984), Mintzberg and McHugh (1985), Perrow (1987), Schweiger, Anderson, and Locke (1985), and Yukl (1981). For examples of pioneering work see Barnard (1938), Mayo (1945), and Simon (1945).

⁵Miller (1992) explicitly serves to link the two camps along a number of dimensions. Others who have produced formal work utilizing more anthropological variables include Kreps (1990) and

Koford and Miller (1991).

⁶While the amount of empirical analysis of formal models varies considerably, attempts to integrate these two analytical approaches is becoming more common. Examples of a few studies that have examined public bureaucracies includes Weingast and Moran (1983), Wood (1988), and Wood and Waterman (1991).

⁷Friedrich only uses OLS in his analysis. Interestingly, our analysis supports many of the propositions that he was unable to confirm.

⁸See Carter who similarly defines brutality as the abuse of authority (Barker and Carter 1991: 197-217).

⁹We do not examine the use of deadly force, an issue which has attracted a great deal of attention in many communities.

¹⁰See Rousseau (1985) for her discussion of compositional, multi-level and cross-level organizational research.

¹¹Barker's study of police deviance in a southern city is particularly relevant. Fellow officers perceived police brutality to be less deviant than sleeping on duty, sex on duty, police perjury, or drinking on duty. Corresponding to this finding, they responded that they would be more likely to report a fellow officer for sleeping on duty than for the use of excessive force. Drinking on duty was considered to be the most deviant (Barker 1978). Also see Matza (1964) and Rubinstein (1973).

¹²Moe (1984) writing about organizations in general discusses the special problems of adverse selection facing a police force. To overcome these problems, police academies are designed to screen, train, and evaluate police recruits so as to foster professionalism. Also see Bittner (1971), Manning (1977: 193-197), and Sykes (1985).

¹³A large number of studies have examined the influence of situational factors on police behavior. See for example Black (1980), Smith and Visher (1981), and Worden (1989).

¹⁴A related issue is whether an incident of brutality stems from malfeasance (a wanton desire to defect) or misfeasance (unintentional violence resulting from incompetence) (Frye 1986: 467). Either problem theoretically could be addressed through appropriate training.

¹⁵At this time, we have not included records of citizen-initiated encounters.

¹⁶The 'runs' file constituted the data set for our previous analysis of shirking. As we note in that paper (Brehm and Gates 1993), there are many serious difficulties in working with the 'runs'

data. Chief among the difficulties was that the attitudinal data did not cleanly merge with the data recording what happened on each shift. We construct the combined runs-attitudinal data set through straightforward merges of officer's records where possible, and imputation where a merge is not conclusive. Further details on the procedure for merging appear in the original paper.

¹⁷We lack any comparable measure for other types of harassment or brutality.

¹⁸These same works suggest further alternative link functions to the Gompertz curve for analysis of rare events, including the cumulative density functions for the Weibull and 'J' distribution. For the purposes of the present discussion, the Gompertz curve is a useful illustration. Note also that the Gompertz curve is the cumulative density function for Gumbel's extreme value distribution (Devroye 1986, pg. 414). Nagler (1992) advocates using the Burr-10 distribution to generate a scobit estimator as an alternative to logit and probit for models with dichotomous dependent variables.

¹⁹This cumulative normal density function is actually slightly different from that used by Dubin and Rivers, adding a 2 to the interaction term.

²⁰See Cialdini (1984) for a comprehensive review of the 'social proof' models of persuasion.

²¹Note that the goodness-of-fit measures are also of no help in resolving a choice between probit or gompit link functions. Both the probit and gompit functions are reparameterizations of the underlying probability π_i of brutality. One property of maximum likelihood estimation is invariance to reparameterization (see King 1989), and we should obtain almost exactly the same likelihood for the probit and gompit estimates across the same data. A desirable statistical test, although not feasible in the current data, is a 'goodness-of-link' test. Pregibon (1980) develops one such goodness-of-link test, but only for link functions within the same 'family' of links (e.g., variations on the logit link function). However, we are comparing two link functions that are wholly different, and are not amenable to such tests.

²²See Klockars (1980) for a similar discussion.

²³See Brehm and Gates (1993) for a similar argument.

²⁴As we comment above, we are doubly limited in this analysis. We would like to be able to replicate this analysis for the incidence of unnecessary frisks, ridicule and threats, but we have no indicators that the officers restrained themselves in these manifestations of brutality. Further, our indicator for the first stage is not exactly what we would like — we can't get inside the heads of officers on the beat twenty-five years ago. We have the observer's record of whether the officer

expressed a desire for some action, but felt prohibited against following that action.

²⁵Unlike the Heckman (1976) or Achen (1986) two-stage approaches to selection bias, the Dubin-Rivers estimator uses full information maximum likelihood, estimating both stages simultaneously. We refer to the 'first stage' only to make it clear that predispositions censor our observations of brutality in the second stage.

²⁶One might also want to draw some provisional interpretations about ρ , the correlation in the error terms u_1, u_2 in the selection (predisposition to brutalize) and outcome stages (actually following through on the predisposition). There is some amount of controversy about how to interpret ρ : Dubin and Rivers' (1990) article interprets the sign of ρ to denote the direction of the relationship between the first and second stage. In this context, the negative ρ suggests that predisposition to brutality counters following through. In each of the subsequent tables, ρ is positive, or that predisposition to brutalize varies in the same direction as following through.

7 References

- Alchian, Armen and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization," *American Economic Review*, 62: 777-795.
- Achen, Christopher H. 1986. *The Statistical Analysis of Quasi-Experiments*. Berkeley: University of California Press.
- Barker, Thomas. 1978. *Journal of Police Science and Administration*. 6: 264-272.
- Barker, Thomas and David L. Carter. 1991. *Police Deviance, Second Edition*. Cincinnati: Anderson Publishing.
- Barnard, Chester. 1938. *The Functions of the Executive*. Cambridge, MA: Harvard University Press.
- Bianco, William and Robert Bates. 1990. "Cooperation By Design: Leadership, Structure, and Collective Dilemmas," *American Political Science Review*, 84: 133-148.
- Bittner, Egon. 1970. *The Function of the Police in Modern Society*. Washington, D.C.: U.S. Government Printing Office.
- Black, Donald. 1980. *The Manners and Customs of the Police*. New York: Academic Press.
- Black, Donald, and Albert J. Reiss. 1967. "Patterns of Behavior and Citizen Transaction," in President's Commission on Law Enforcement and Administration of Justice, *Studies in Crime and Law Enforcement in Major Metropolitan Areas*, Field Studies III, Vol. II, Sec. I

(Washington: U.S. Government Printing Office).

Brehm, John and Scott Gates. 1990. "Supervision and Control," paper presented at the 1990 American Political Science Association Meetings, San Francisco, California.

Brehm, John and Scott Gates. 1993. "Donut Shops and Speedtraps: Evaluating Models of Supervision on Police Behavior." *American Journal of Political Science* 37: forthcoming May, 1993.

Brown, Michael K. 1981. *Working the Street: Police Discretion and the Dilemmas of Reform*, New York: Russell Sage Foundation.

Cialdini, Robert. 1984. *Influence: The New Psychology of Modern Persuasion*, New York: Quill Books.

Devroye, Luc. 1986. *Non-Uniform Random Variate Generation*. New York: Springer-Verlag.

Dubin, Jeffrey A. and Douglas Rivers. 1990. "Selection Bias in Linear Regression, Logit and Probit Models," *Sociological Methods and Research*, 18:360-390.

Fayol, Henri. 1949, 1916. *General and Industrial Management*. Trans. C. Storrs. London: Pitman.

Friedrich, R.J. 1980. "Police Use of Force: Individuals, Situations, and Organizations," *Annals of the American Academy of Political and Social Sciences*. 452: 82-97.

Fyfe, James J. 1986. "The Split-Second Syndrome and Other Determinants of Police Violence," *Violent Transactions*, edited by A. Campbell and J. Gibbs. New York: Basil Blackwell.

Goldberger, Arthur. 1964. *Econometric Theory*. New York: Wiley.

- Goldstein, Herman. 1977. *Policing in a Free Society*. Cambridge, Mass.: Ballinger.
- Graen, Gary Bearnard, Novak, M. and Sommerkamp, P. 1982. "The Effect of Leader-Member Exchange and Job Design on Productivity and Job Satisfaction: Testing a Dual Attachment Model," *Organizational Behavior and Human Performance*, 30: 109-131.
- Graen, Gary Bearnard and Terri A. Scandura. 1987. "Toward a Psychology of Dyadic Organizing," *Organizational Behavior*, 9: 175-208.
- Grossman, Sanford and Oliver Hart. 1983. "An Analysis of the Principal Agent Problem," *Econometrica*, 51: 7-45.
- Gulick, Luther. 1937. "Notes on the Theory of Organization." In *Papers on the Science of Administration*, eds. L.Gulick and L. Urwick. New York: Institute of Public Administration.
- Heckman, James J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models," *Annals of Economic and Social Measurement*, 5/4:475-492.
- Holden, Kritina L. and John E. Overall. 1987. "Tests of Significance for Differences in Counts of Rare Events in Two Treatment Groups," *Educational and Psychological Measurement*, 47:881-892.
- Holmstrom, Bengt. 1979. "Moral Hazard and Observability," *Bell Journal of Economics*, 10: 74-91.
- Kaplan, Robert E. 1984. "Trade Routes: The Manager's Network of Relationships," *Organizational Dynamics*, 12: 37-52.
- Klockars, Carl B. 1980. "The Dirty Harry Problem," *Annals of the American Academy of Polit-*

ical and Social Sciences. 452: 33-47.

Koford, Kenneth J. and Jeffrey B. Miller, editors. 1991. *Social Norms and Economic Institutions*. Ann Arbor: University of Michigan Press.

Krehbiel, Keith. 1991. *Information and Legislative Organization*. Ann Arbor: University of Michigan Press.

Kreps, David M. 1990. "Corporate Culture and Economic Theory," in *Perspectives on Positive Political Economy*, edited by James E. Alt and Kenneth A. Shepsle. New York: Cambridge University Press.

Lipsky, Michael. 1980. *Street-Level Bureaucracy*. New York: Russell Sage Foundation.

Manning, Peter K. 1977. *Police Work*. Cambridge, Mass.: MIT Press.

Manning, Peter K. 1982. "Organizational Work: Structuration of Environments," *British Journal of Sociology*, 33: 118-132.

Matza, D. 1964. *Delinquency and Drift*. New York: John Wiley.

Mayo, Elton. 1945. *The Social Problems of an Industrial Civilization*. Cambridge, MA: Harvard University Press.

McCullagh, P. and J.A. Nelder. 1983. *Generalized Linear Models*, 2d Edition. London: Chapman and Hall.

McDonald, James B. and Richard J. Butler. 1990. "Regression Models for Positive Random Variables," *Journal of Econometrics*, 43:227-251.

- Miller, Gary J. 1992. *Managerial Dilemmas: The Political Economy of Hierarchy*, New York: Cambridge University Press.
- Mintzberg, Henry and Alexandra McHugh. 1985. "Strategy Formation in an Adhocracy," *Administrative Science Quarterly*, 30: 160-197.
- Nagler, Jonathan. 1992. "Scobit: An Alternative Estimator to Logit and Probit for Models with Dichotomous Dependent Variables." Paper presented at the Midwest Political Science Association meetings, Chicago.
- Perrow, Charles. 1987. *Complex Organizations: A Critical Essay, Third Edition*, New York: Random House.
- Pregibon, Daryl. 1980. "Goodness of Link Tests for Generalized Linear Models," *Applied Statistics*, 29:15-24.
- Radelet, Louis A. 1986. *The Police and the Community, Fourth Edition*, New York: Macmillan Publishing.
- Reiss, Albert J. 1971. *The Police and The Public*, New Haven: Yale University Press.
- Rousseau, Denise M. 1985. "Issues of Level in Organization Research: Multi-level and Cross-level Perspectives," *Research in Organizational Behavior*, 7: 1-37.
- Rubinstein, Jonathan. 1973. *City Police*. New York: Farrar, Straus, and Giroux.
- Schweiger, David M., Carl R. Anderson, and Edwin A. Locke. 1985. "Complex Decision Making: A Longitudinal Study of Process and Performance," *Organizational Behavior and Human Decision Processes*, 36: 245-272.

- Simon, Herbert A. 1945. *Administrative Behavior*. New York: Free Press.
- Smith, Douglas A. and Christy A. Visser. 1981. "Street-Level Justice: Situational Determinants of Police Arrest Decisions," *Social Problems* 29: 167-178.
- Stein, Robert M. 1990. "The Budgetary Effects of Municipal Service Contracting: A Principal-Agent Explanation," *American Journal of Political Science* 34: 471-502.
- Sykes, Gary W. 1985. "The Functional Nature of Police Reform: The 'Myth' of Controlling the Police," *Justice Quarterly*. 2: 52-65.
- Taylor, Frederick W. 1911. *The Principles of Scientific Management*, New York: W.W. Norton.
- Van Maanen, John. 1983. "The Boss: First-Line Supervision in an American Police Agency," in M. Punch, ed. *Control in the Police Organization*, Cambridge, MA: MIT Press.
- Weingast, Barry R. and Mark J. Moran. 1983. "Bureaucratic Discretion or Congressional Control?: Regulatory Policymaking by the Federal Trade Commission," *Journal of Political Economy* 91: 765-800.
- Wilson, James Q. *Varieties of Police Behavior: The Management of Law and Order in Eight Communities*. Cambridge, MA: Harvard University Press.
- Worden, Robert E. 1989. "Situational and Attitudinal Explanations of Police Behavior: A Theoretical Reappraisal and Empirical Assessment," *Law and Society Review*, 23: 667-711.
- Wood, B. Dan. 1988. "Principals, Bureaucrats, and Responsiveness in Clean Air Enforcements," *American Political Science Review* 82: 213-234.
- Wood, B. Dan, and Richard W. Waterman. 1991. "The Dynamics of Political Control of the

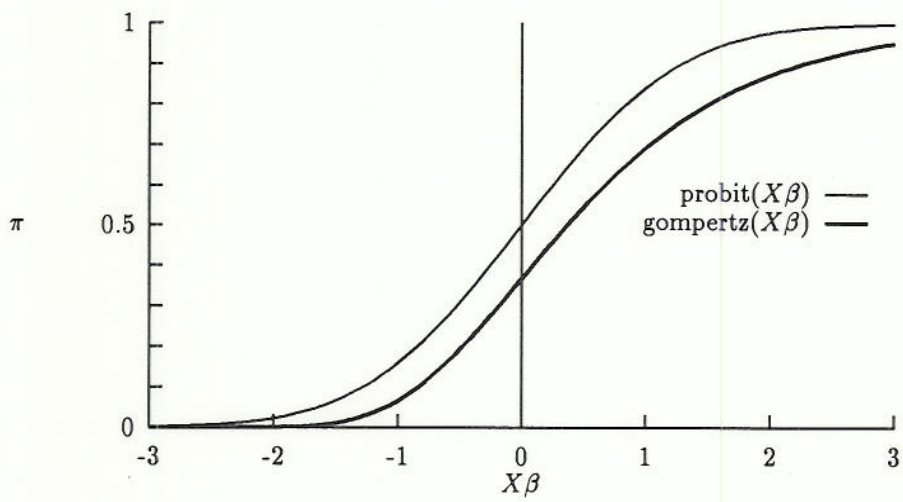


Figure 1: Comparison of Gompertz and Probit Link Functions

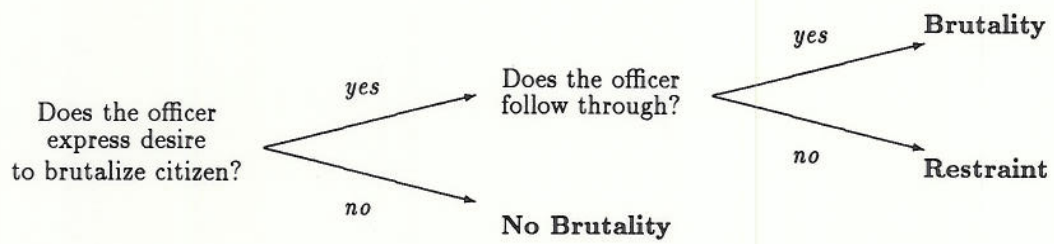


Figure 2: Police Brutality as a Censored Data Problem

Table 1: Probit and Gompit Estimates of Use of Unnecessary Force by Police Officers as a Function of the City of Residence, 1966 Police Behavior Data

	Probit Estimates		Gompit Estimates	
	<i>B</i>	<i>D</i>	<i>B</i>	<i>D</i>
Constant	-2.71		-1.71	
Boston	.72	.0044	.39*	.0046
Chicago	.77	.0034	.42*	.0035
Human Relations Skills	-.17	-.0002	-.10	-.0003
Intellectual Style	.02	.0000	-.02	.0001
Professional Dimension	.03	.0000	.01	.0000
Felony	.29	.0013	.16	.0015
Violent	.43	.0017	.23*	.0018
Number of Officers	.01	.0250	.004	.0250
Solidary Likes	-.80	-.0006	-.45*	-.0006
Functional Likes	-.80	-.0009	-.45*	-.0008

*=statistically significant at $p < .05$; *D* = 1st Differences.

There are 3507 observations in each analysis. $\chi^2 = 4774$ with 15 d.f. for both probit and gompit models.

Table 2: Probit and Gompit Estimates of Use of Unnecessary Frisks by Police Officers as a Function of the City of Residence, 1966 Police Behavior Data

	Probit Estimates		Gompit Estimates	
	<i>B</i>	<i>D</i>	<i>B</i>	<i>D</i>
Constant	-2.57		-1.65	
Boston	.28	.0067	-.16*	-.0030
Chicago	-.08	-.0011	-.05	-.0009
Human Relations Skills	-.01	-.0002	.02	.0000
Intellectual Style	-.04	-.0005	-.03	-.0005
Professional Dimension	.22	.0022	.12*	.0015
Felony	.33	.0102	.19*	.0078
Violent	-.14	-.0022	-.07	-.0015
Number of Officers	.008	.0692	.006	.0298
Solidary Likes	-.34	-.0025	-.21*	-.0019
Functional Likes	-.26	-.0038	-.15	-.0028

*=statistically significant at $p < .05$; *D* = 1st Differences.
 There are 3507 observations in each analysis. $\chi^2 = 4564$
 with 15 d.f. for both probit and gompit models.

Table 3: Probit and Gompit Estimates of Use of Ridicule by Police Officers as a Function of the City of Residence, 1966 Police Behavior Data

	Probit Estimates		Gompit Estimates	
	B	\mathcal{D}	B	\mathcal{D}
Constant	-3.10		-1.99	
Boston	-1.09*	-.0110	-.62*	-.0077
Chicago	-.51*	-.0069	-.29*	-.0050
Human Relations Skills	.11	.0019	.06	.0014
Intellectual Style	-.008	-.0001	-.04	-.0009
Professional Dimension	.28	.0036	.16*	.0027
Felony	.01	.0003	.05	.0018
Violent	.12	.0031	.08	.0025
Number of Officers	.008	-.0014	.007*	.0095
Solidary Likes	.18	.0021	.11	.0016
Functional Likes	.54	.0197	.30*	.0135

*=statistically significant at $p < .05$; \mathcal{D} = 1st Differences.

There are 4280 observations in each analysis. $\chi^2 = 4411$
with 15 d.f. for both probit and gompit models.

Table 4: Probit and Gompit Estimates of Use of Threats by Police Officers as a Function of the City of Residence, 1966 Police Behavior Data

	Probit Estimates		Gompit Estimates	
	<i>B</i>	<i>D</i>	<i>B</i>	<i>D</i>
Constant	-1.17		-.75	
Boston	.13	.0126	.09*	.0122
Chicago	.21*	.0169	.14*	.0163
Human Relations Skills	.05	.0029	.03	.0029
Intellectual Style	-.16*	-.0098	-.11*	-.0103
Professional Dimension	-.05	-.0020	-.03	-.0021
Felony	-.13	-.0124	-.09	-.0130
Violent	.35*	.0361	.25*	.0366
Number of Officers	.007	.1669	.007*	.2201
Solidary Likes	-.23*	-.0086	-.17*	-.0089
Functional Likes	-.21	-.0161	-.15*	-.0168

*=statistically significant at $p < .05$; *D* = 1st Differences.
 There are 3507 observations in each analysis. $\chi^2 = 3260$
 with 15 d.f. for both probit and gompit models.

Table 5: Probit and Censored Probit Estimates of Use of Unnecessary Force by Police Officers as a Function of the City of Residence, 1966 Police Behavior Data

	Probit Estimates		Censored Probit Estimates	
	<i>B</i>	<i>D</i>	<i>B</i>	<i>D</i>
	Selection Model			
Constant			-2.45*	
Authoritarian			.51*	
Violent			.26*	
Number of Officers			.004	
	Outcome Model			
Constant	-2.71		-2.64	
Boston	.72*	.0044	.09*	.0114
Chicago	.77	.0034	.18*	.0195
Human Relations Skills	-.17	-.0002	-.28	-.0214
Intellectual Style	.02	.0000	1.09*	.1420
Professional Dimension	.03	.0000	-.42*	-.0225
Felony	.29	.0013	.03	.0051
Violent	.43	.0017	.53*	.0792
Number of Officers	.01	.0250	.05*	.9103
Solidary Likes	-.80*	-.0006	.11*	.0059
Functional Likes	-.80*	-.0009	.15*	.0200

*=statistically significant at $p < .05$; *D* = 1st Differences.
 There are 3507 observations in each analysis. $\chi^2 = 4774$
 with 15 d.f. for both probit and gompit models.

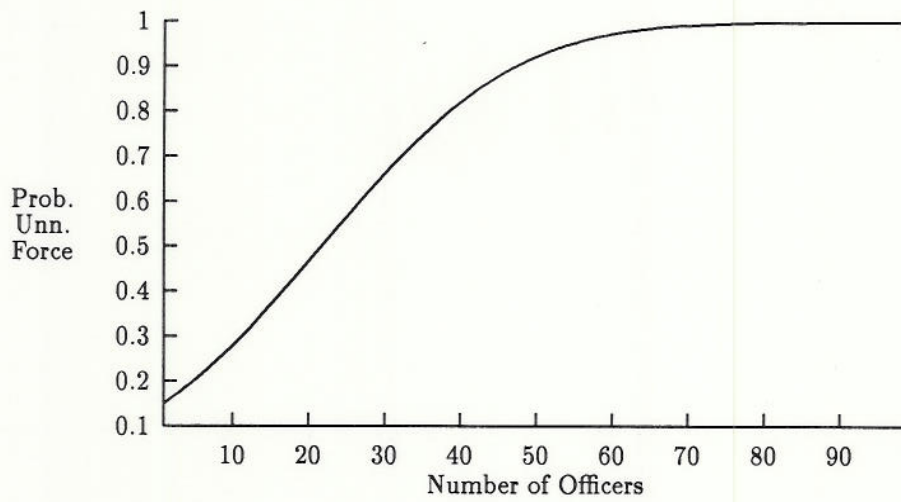


Figure 3: Predicted Probability of Unnecessary Force per Officer at the Scene, Given a Predisposition to Brutalize