PROACTIVE POLICING AND ROBBERY RATES ACROSS U.S. CITIES*

CHARIS E. KUBRIN
Department of Sociology
George Washington University

STEVEN F. MESSNER
GLENN DEANE
KELLY MCGEEVER
Department of Sociology
State University of New York, Albany

THOMAS D. STUCKY
School of Public and Environmental Affairs
Indiana University—Purdue University Indianapolis

KEYWORDS: proactive policing, violent crime, deterrence, endogeneity

In recent years, criminologists, as well as journalists, have devoted considerable attention to the potential deterrent effect of what is sometimes referred to as “proactive” policing. This policing style entails the vigorous enforcement of laws against relatively minor offenses to prevent more serious crime. The current study examines the effect of proactive policing on robbery rates for a sample of large U.S. cities using an innovative measure developed by Sampson and Cohen (1988). We replicate their cross-sectional analyses using data from 2000 to 2003, which is a period that proactive policing is likely to have become more common than that of the original study—the early 1980s. We also

* An earlier version of the article was presented at a workshop sponsored by the National Consortium of Violence Research (NCOVR). NCOVR has been supported by a grant from the National Science Foundation. We are grateful to the participants at the workshop for comments on the article, with special thanks to Robert J. Bursik, Jr., Thomas Bernard (deceased), and Paul Nieuwbeerta. We are also grateful to Debra K. Mack of the Criminal Justice Information Services Division of the FBI for providing data on offenses, arrests, and police employees from the Uniform Crime Reporting program. Direct correspondence to Charis E. Kubrin, Department of Sociology, George Washington University, Phillips Hall 409, 801 22nd St. NW, Washington, DC 20052 (e-mail: charisk@gwu.edu).

© 2010 American Society of Criminology

CRIMINOLOGY VOLUME 48 NUMBER 1 2010 57
extend their analyses by estimating a more comprehensive regression model that incorporates additional theoretically relevant predictors. Finally, we advance previous research in this area by using panel data. The cross-sectional analyses replicate prior findings of a negative relationship between proactive policing and robbery rates. In addition, our dynamic models suggest that proactive policing is endogenous to changes in robbery rates. When this feedback between robbery and proactive policing is eliminated, we find more evidence to support our finding that proactive policing reduces robbery rates.

During the past decade, crime rates have fallen dramatically in the United States, according to both official police statistics and victimization data. This pronounced decline was largely unexpected, but as Levitt (2004: 163) observed, “there has been no shortage of hypotheses to explain the drop in crime after the fact.” One explanation that has received considerable media attention is the possible effect of proactive, “quality-of-life,” or “broken windows” policing—a style that reflects strict enforcement of all laws and addresses minor infractions as well as serious crimes (Sampson and Cohen, 1988; Sherman, 1995). Indeed, a LexisNexis (Reed Elsevier, Inc., Albany, NY) search of articles in ten leading newspapers between 1991 and 2001 identified “innovative policing strategies” as the most frequently referenced explanation for falling crime rates after “increased reliance on prisons” (Levitt 2004: 164). Despite the widespread and often favorable public attention devoted to proactive policing, research on the efficacy of such strategies is surprisingly sparse.

One of the more sophisticated attempts to assess the deterrent effects of proactive policing is a study by Sampson and Cohen (1988) published more than 20 years ago. The following two features of this study are particularly noteworthy: 1) Their innovative measure of proactive policing, which focused on law-enforcement responses to disorderly conduct and driving under the influence as well as 2) their creative theorizing about the direct effects of proactive policing on offending rates in addition to any indirect effects via increased probability of apprehension. Sampson and Cohen found support for the deterrence perspective in their analysis of 171 large U.S. cities circa 1980. They reported a negative association between proactive policing and robbery rates in multivariate models, net of racial inequality, marital disruption, socioeconomic status, and racial composition indicators. A study by MacDonald (2002) approximately 15 years later reaffirmed Sampson and Cohen’s finding about the deterrent effect of proactive policing.

Our study builds on this literature in several ways. First, by employing a similar cross-sectional design, we replicated Sampson and Cohen (1988) to determine whether proactive policing is associated with lower robbery
rates using data for 2000–2003—a period during which proactive policing is likely to have diffused more widely across police agencies. Consistent with their study, we examined large U.S. cities with appreciable Black populations. Second, we significantly expanded their model specification. In the years since their study was published, a set of covariates has emerged as potent predictors of violent crime rates (Land, McCall, and Cohen, 1990). We, thus, determined whether any effect of proactive policing on robbery rates withstands controls for a composite measure of disadvantage.

Finally, we used panel data to assess the effect of proactive policing on robbery rates (Finkel, 1995; Voas, Olson, and Crockett, 2002). Although cross-sectional models identify contemporaneous associations, they cannot account for temporal ordering. In addition, unless properly instrumented, cross-sectional regression models are predicated on the assumption that the variables on the right-hand side of the equation are exogenous. This assumption is problematic when examining the impact of proactive policing because police practices are likely to be responsive to changes in crime levels. To anticipate the results, using a cross-lagged effects model (Finkel, 1995), our analyses indicated that proactive policing is endogenous to robbery rates, which suggests that reciprocal causal effects must be purged to quantify accurately the effect of proactive policing on robbery rates. We do so through an application of the Arellano–Bond dynamic panel-data estimator (Arellano, 2003; Arellano and Bond, 1991).

### POLICE STRENGTH, ARREST CERTAINTY, AND DETERRENCE

A common perception is that the police play a major role in preventing crime. Indeed, during times of rising crime rates, a frequent response is to call for more officers on the streets. Consistent with the deterrence perspective, it is assumed that a greater police presence will reduce crime rates because would-be offenders adjust their perceptions to the increased probability of arrest. Although a reduction in crime seems intuitively plausible, research on the deterrent effects of arrest certainty and police size has produced conflicting results. Some studies reported a deterrent effect (Levitt, 1997; Marvell and Moody, 1996; Sampson and Cohen, 1988; Tittle and Rowe, 1974; Wilson and Boland, 1978), others reported no relationship (Decker and Kohfeld, 1985; Weiss and Freels, 1996), and one study

---

1. In additional analyses, Sampson and Cohen (1988) examined age-race-disaggregated arrest rates and found the effect of aggressive policing on robbery is largest for adult and Black offenders. For our purposes, analyses were restricted to assessing the overall deterrent effect of proactive policing, which was determined with data on “offenses known to the police” rather than with arrest data.
even found a positive relationship between arrest certainty and crime (Jacob and Rich, 1980). It, thus, is not clear that more police equals less crime.

Critics of deterrence explain the conflicting results by arguing that police have little, if any, impact on crime because most police work is not devoted to crime reduction, and the most frequently employed police strategies are poor crime-prevention strategies (Marvell and Moody, 1996: 610). A related argument stated that changes in police practices might affect crime through causal linkages that do not increase arrest certainty because the police have minimal control over important factors such as the willingness of citizens to report a crime, offer information to the police, or identify suspects (Wilson and Boland, 1978: 369).

Methodological issues also might explain discrepant findings on the effects of proactive policing in previous research. First, scholars have attempted to capture arrest certainty with measures that reflect police presence, which include the size of the police force (Wilson and Boland, 1978), arrest–offense ratio or percentage of crimes cleared (Decker and Kohfeld, 1985; Tittle and Rowe, 1974), and number of moving violations or citations issued (Jacob and Rich, 1980; Wilson and Boland, 1978). But these measures have potentially serious drawbacks. With respect to clearance rates, Wilson and Boland (1978: 368) contended that “a crime ‘cleared by arrest’ is whatever the police say it is” and argued that such rates might vary substantially among police departments for reasons that have little to do with the objective probability of getting caught. Likewise, Jacob and Rich (1980: 113) concluded “the number of moving violations is not a good indicator of the type of police aggressiveness that might be related to the apprehension of robbers and the deterrence of robbery.” Finally, others contended that arrests are a measure of police responsiveness to crime rather than a primary source of deterrence (Decker and Kohfeld, 1985). From this perspective, arrests occur in response to crime rather than prevent it.

Decker and Kohfeld (1985: 439) raised an additional methodological issue and charged that the negative relationship between the arrest–offense ratio and the crime rates reported in studies might be spurious for several reasons. One reason has to do with the use of “offenses known” in the measurement of both variables. In the conventional measure of punishment certainty, the numerator of the crime rate (offenses known) is identical to the denominator of the measure of punishment certainty, which induces a nonlinear inverse relationship between crime and

---

2. In their research based on the Granger causality modeling to correct for simultaneity bias, Marvell and Moody (1996) found that police size has a significant negative impact on crime rates consistent with the deterrence perspective.
PROACTIVE POLICING AND ROBBERY RATES

the certainty of punishment measure. Decker and Kohfeld claimed that the negative correlation between arrest–offense ratios and crime rates might be an artifact of the construction of the rate measures, which they confirmed in their study. Others have noted this concern (Jacob and Rich, 1980: 114).

An additional methodological issue pertains to the nature of the causal relationship between police actions or police strength and crime rates. With respect to arrest certainty, scholars have noted that an inverse association could occur either because high arrest–offense ratios lead to lower crime rates or because higher crime rates cause lower arrest–offense ratios. The latter situation might emerge if increases in reported crime swamp police resources so that the rate at which criminals are arrested declines (Wilson and Boland, 1978: 369). With respect to measures of police-force size, the problem of reciprocal causation also results but with different implications. It is plausible that increasing crime rates will instigate the hiring of more police officers. Unless this positive reciprocal effect is controlled, estimates of the expected negative effect of police size on crime rates will be suppressed.

Research has attempted to address the endogeneity issue with varying identification strategies. Marvell and Moody (1996) used the Granger causality test on pooled city and state data across 20 years to determine the causal direction, if any, between police-force size and index crime rates. They found causation in both directions, although the impact of crime on the number of police is slight, whereas the impact of police size on most crimes is substantial. In his panel study of 59 U.S. cities from 1970 to 1992, Levitt (1997) also found support for deterrence. He used the timing of mayoral and gubernatorial elections as an instrumental variable in two-stage least squares (2SLS) analyses, arguing that it affected police-force size but did not directly influence crime rates. The results showed that increases in police size substantially reduce violent crime but have a smaller impact on property crime. Although these results and others (e.g., Wilson and Boland, 1978) supported a deterrence perspective, questions about methodological procedures remain (see McCrary, 2002).

In sum, recent research has attempted to assess the relationship between policing and crime, focusing primarily on measures of arrest certainty and the police-force size. Although some results supported the deterrence perspective, conclusions are difficult to draw because of methodological limitations. An alternative approach is to shift attention away from the arrest–offense ratio and measures of police size to consider more directly how the style of policing might affect crime because police activities are likely more important for reducing crime than police strength per se (MacDonald, 2002; Sherman, 1995).
The nature or style of policing is the central consideration in Sampson and Cohen's (1988) study. They built on the work of Wilson and Boland (1978) and Wilson and Kelling (1982) to examine whether proactive policing reduces crime. This policing approach, also referred to as “quality-of-life” or “broken windows” policing, relies on the professional model of policing (Wilson, 1968). It does not necessarily involve direct collaboration with the community and is, therefore, different from “community policing.” Instead, crime control occurs through strict enforcement of all laws (Sampson and Cohen, 1988; Sherman, 1995). A proactive strategy does not mean that the officer is antagonistic or hostile but that s/he maximizes the number of interventions in, and observations of, the community. The most common form of proactive enforcement involves field interrogations, whereby police engage in directed patrol efforts to question suspicious persons and enforce traffic violations (MacDonald, 2002: 595; Sherman, 1995).

Sampson and Cohen (1988) suggested two mechanisms through which proactive policing might reduce crime. The first pertained to the indirect effect of proactive policing on crime through arrest risk. As Wilson and Boland (1978: 373) argued, “[b]y stopping, questioning, and otherwise closely observing citizens, especially suspicious ones, the police are more likely to find fugitives, detect contraband (such as stolen property or concealed weapons), and apprehend persons fleeing from the scene of a crime.” Thus, proactive policing is hypothesized to affect crime rates by changing the actual probability that an arrest is made—in other words, by increasing the arrest–offense ratio.

The second means through which proactive policing might affect crime rates directly was by influencing prospective offenders’ perceptions of the apprehension probabilities for criminal behavior. Such an effect would result if would-be offenders believe their chances of being arrested for a crime have increased, even if they have not. Of course, this strategy rests on the assumption that changes in police behavior are visible to prospective offenders. Because direct information is generally not available, prospective offenders have little idea about the objective probability of arrests (Sampson and Cohen, 1988). Fisher and Nagin (1978: 388) note that “. . . unlike stock market prices, daily quotations of sanction levels are not available, and the information that is available derives from uncertain sources, including the criminal’s own experience, the experience of his

---

3. See Greene (2000) for similarities and differences between community and proactive policing and Muhlhausen (2002), Sabol (2005), and Zhao, Schneider, and Thurman (2002) for assessments of the impact of community policing programs.
peers, [and] news reports.” Moreover, arrests for many serious index crimes (e.g., robbery) are relatively rare, so potential offenders often do not witness index crime arrests. However, vigorous intervention by police on driving violations, drunkenness, and public disorder is a visible indicator of police activity in an area, according to Sampson and Cohen (1988). They suggested, therefore, that control of minor offenses might have an influence on crime rates, such as robbery, that is not strictly tied to changes in the objective arrest risk for the latter.

Interest in proactive policing has grown markedly since the publication of Sampson and Cohen’s (1988) study, in large part because of the highly favorable media coverage of the adoption of such a strategy in New York City (NYC) under the leadership of Police Commissioner William Bratton. Violent crime fell dramatically in NYC during (and after) Bratton’s tenure (Kelling and Coles, 1996: 110, 148–56). Despite favorable press coverage, however, the scholarly literature on the effectiveness of the NYC policing innovations has yielded mixed results (Bowling, 1999; Conklin, 2003; Harcourt, 2001; Kamen, 2000; Messner et al., 2007; Rosenfeld, Fornango, and Rengifo, 2007).

From this discussion, it is clear that the role of policing in reducing violent crime remains an open question. Moreover, an obvious drawback of much research in this area is its limited geographic coverage—findings reflect the experience of a single city. A major contribution of Sampson and Cohen (1988) was their introduction of a proactive policing measure based on arrest data that were collected as part of the Uniform Crime Reporting (UCR) program. They operationalized proactive policing in terms of the arrest ratio for two “public order” crimes recorded in the UCR—disorderly conduct and driving under the influence (DUI)—to the number of police officers in the jurisdiction. These data are available for a large number of cities. Moreover, this policing measure does not include terms that appear in typical dependent variables in either the numerator or the denominator (e.g., the number of violent offenses known).

More recently, MacDonald (2002) examined these issues in his study of large U.S. cities in the mid-1990s. Assessing the economic and political determinants of robbery and homicide rates in 164 cities, MacDonald (2002) found that although community policing has little effect on the control and decline in violent crime during the 1990s, proactive policing strategies, as operationalized by Sampson and Cohen (1988), had a significant inverse effect on violent crime rates and were related to reductions in violent crime across time. Specifically, police departments with a more aggressive enforcement of disorderly conduct and driving while under the influence exhibited lower levels of robbery and homicide, which is consistent with the deterrence perspective.
SPECIFYING A MORE COMPLETE MODEL

Although Sampson and Cohen (1988) and MacDonald (2002), among others, supported the idea that proactive policing is associated with lower violent crime rates, the analyses were limited in three important respects. First, research now has been considering the combined effects of social and economic predictors of crime. In many inner-city communities and as a result of macroeconomic changes that disproportionately have affected the urban poor, scholars have claimed it is the combination of poverty, unemployment, and family disruption that defines the socioeconomic context for residents (Sampson and Wilson, 1995; Wilson, 1987). They have posited that “concentration effects” contribute to social disorganization, which in turn leads to more crime and violence. Consistent with these arguments, researchers typically measure the multiple disadvantages that characterize areas by incorporating several measures into an overarching index of concentrated disadvantage. As Land, McCall, and Cohen (1990) noted, incorporating these covariates into an index has the added benefit of minimizing multicollinearity. Because this measure is associated positively with violent crime rates at “virtually all levels of analysis and time periods” (Land, McCall, and Cohen, 1990: 952), a rigorous assessment of the impact of proactive policing requires the inclusion of a comprehensive concentrated disadvantage measure.

Second, prior research on proactive policing and crime typically has not considered local politics, which is ironic given the intellectual history of this literature. Sampson and Cohen’s (1988) arguments in regard to aggressive policing were based on Wilson’s (1968) seminal book Varieties of Police Behavior, which identified three policing styles—the watchman, the legalistic, and the service.4 For Wilson, the adoption of these organizational styles was driven by the local political culture. Some research has shown police strategies and politics to be related (Crank, 1990; Langworthy, 1985; Slovak, 1986; Wilson and Boland, 1978); yet Langworthy (1986) noted minimal differences in police styles across political cultures and concluded that police strategies primarily are determined internally. Therefore, for many years, policing studies have ignored political effects.

Recently, however, the discussion of politics seems to have reentered policing research. Stucky (2005) found that cities with unreformed or “traditional city political systems” are likely to have more police per capita, net of other factors (see also Stucky, 2003). Similarly, Choi, Turner, and Volden (2002) found that cities with elected mayors were more likely to apply for Community Oriented Policing Services (COPS) grants from

4. The watchman style emphasizes order maintenance, the legalistic style emphasizes vigorous, professional law enforcement, and the service style emphasizes serving the public and aiding citizens in need.
the federal government. They argued this pattern was because of potential
credit claiming opportunities for mayors who want to be perceived as
doing something about crime. Politics, thus, seem to matter for policing,
although the direction of political effects might be opposite to that origi-
nally suggested by Wilson (1968). Wilson argued that aggressive policing
would be more likely in cities with reformed political systems because city
managers would hire “professional” police chiefs. Yet, as noted, research
finds the opposite—traditional cities had more police per capita and were
more likely to apply for COPS grants. We argue that this pattern exists
precisely because they are more susceptible to political pressure. If proac-
tive policing results from community pressure to “do something about
crime” beyond traditional strategies, then it seems most likely to occur in
places where community influence on government action is maximized
(i.e., cities with elected mayors, partisan elections, and district-based coun-
cil representation). Given these arguments, it is essential to consider the
effects of local politics on the style of policing and crime rates.

A third important research limitation that has followed in the tradition
of Sampson and Cohen (1988) is the failure to confront directly the vexing
problem of endogeneity, referenced earlier. Sampson and Cohen (1988:
171) correctly observed that their innovative measure of proactive policing
“avoids spurious correlations,” which could be produced by common
terms in the deterrence measure and the crime rate (e.g., offenses known).
However, they dismissed the possibility of simultaneity on purely a priori
grounds and argued that “police intervention in moving violations and dis-
orderly conduct is not causally determined by the crime rate but rather by
the dominant political culture and the professionalism of the police
department” (1988: 171). This claim is by no means obvious. It is possible,
for example, that increases in violent crimes such as robbery might stimu-
late a redeployment of police resources away from the less serious crimes
of moving violations and disorderly conduct to combat more serious
offenses. The observed negative association between proactive policing
and robbery, thus, might reflect a causal process opposite to the deter-
rence hypothesis. Alternatively, if rising robbery rates stimulate pressures
for the police to be “doing something” that is highly visible, then the effect
of increasing robbery rates on measures of proactive policing might be
positive. Under this scenario, any deterrent effect of proactive policing
might be suppressed in models that fail to account for endogeneity.

In sum, although some previous research is suggestive of the deterrent
effect of proactive policing, the substantive and methodological issues dis-
cussed preclude firm conclusions. In this vein, Weisburd and Eck (2004:
60) noted, “[m]any [police] tactics that are applied broadly throughout the
United States have not been the subject of systematic police research nor
have they been examined in the context of research designs that allow
practitioners or policy makers to draw very strong conclusions. American police research must become more systematic . . . if it is to provide solid answers to important questions of practice and policy.” In line with this claim, the goal of the current study was to determine whether the previously observed deterrent effects of proactive policing are robust across more fully specified and methodologically rigorous models.

DATA AND METHODS

To facilitate comparisons with Sampson and Cohen (1988), we applied analogous selection criteria to generate the sample—U.S. cities with a population of 100,000 or more with at least 1,000 Blacks in 2000.5 Missing data reduced the sample somewhat, which resulted in a sample size of 181 cities for our cross-sectional analyses. Our sample can be regarded as a reasonable reflection of contemporary large cities with appreciable Black populations, which represented 78 percent of the cities with the specified characteristics.6 Our assessment of dynamic processes necessitated multiple years of observation. Because our original cross-sectional analyses employed data from the 4-year period, 2000–2003, we created a panel-data structure for our investigation of endogeneity by extending backward across the preceding 4-year period, 1996–1999. This 8-year period gave us sufficient power to apply the Arellano–Bond dynamic panel-data estimator, which relied on identification using lagged levels of the dependent, predetermined, and endogenous variables as well as the differences of the strictly exogenous variables.7

The data were collected from the following five sources: 1) counts of robberies known to the police and city population totals as compiled in the Federal Bureau of Investigation’s (FBI) UCR program; 2) yearly arrest counts for DUI and disorderly conduct for agencies that submitted 12 complete months of data to the UCR; 3) police employee data as contained in the FBI’s Police Employee Master File; 4) demographic data as reported in the 1990 and the 2000 census; and 5) two databases on political

5. The minimum Black population requirement was necessary to measure racial inequality (see Blau and Blau, 1982).
6. A few cities require elaboration. First, Honolulu was not considered a place as defined by the U.S. Census, so for all census variables, we used Honolulu MSA data. Second, Las Vegas Metro PD did not have a corresponding census place, so we used Las Vegas urbanized area census data. The U.S. Census recognized Amherst Town, NY, and Ramapo, NY, as county subdivisions and not as places. Therefore, we used country subdivision data.
7. We used the 2000 sample as our selection criterion; therefore, a few places have populations less than 100,000 and/or less than 1,000 Blacks in years before 2000. Athens-Clarke County, GA, was a new geographic type in the 2000 U.S. Census. No comparable geography was available in the 1990 U.S. Census. Thus, it was not included in panel analyses.
PROACTIVE POLICING AND ROBBERY RATES

The UCR counts of robberies known to police for cities, along with corresponding population totals, were taken from files distributed by the Interuniversity Consortium for Political and Social Research (ICPSR) for all years except 2003. The 2003 UCR data were taken from the FBI's website (table 8; http://www.fbi.gov/ucr.htm#cisu). The FBI provided the arrest and police employee data in a personal communication. The census data on sociodemographic characteristics of cities were taken from the 1990 and 2000 census tapes. Information in regard to the city political system characteristics was obtained from the International City County Management Association (ICMA) and the National League of Cities’ Municipal Officials Database.

DEPENDENT VARIABLE

To generate the sample, we initially collected data on robbery “offenses known” for all cities with populations of 100,000 and greater that were in the UCR between 2000 and 2003. Robbery rates (per 100,000 population) were computed in the conventional manner for each year with available data using the robbery counts and totals from the UCR. To reduce the likelihood that reporting error might produce unstable estimates, we smoothed the data by aggregating the counts across the multiyear period. We imposed the selection criterion that a city must have at least 2 years of data to be included in the cross-sectional analyses. If a city reported for all years, then the robbery rate was based on the 4-year total. If the city reported for 3 of the 4 years, then the robbery rate reflected those 3 years.


9. We are grateful to Debra K. Mack, Chief of Programs Support Section of the FBI’s Criminal Justice Information Services Division, for providing the arrest and employee data and for clarification of these data.


11. Cities in some states (e.g., Florida and Illinois) did not report information to the UCR and were excluded from analyses.
and so on. The smoothed robbery rates were transformed to natural logarithm scale to reduce right skew in our cross-sectional analyses.

We used single-year calculations of the robbery rates rather than smoothed 4-year averages in the panel analyses (reported in tables 2 and 4, which appear in the Results section). Robbery rates were computed in the conventional manner between 1996 and 2003. We also constructed the 4-year average for 1996–1999 (reported in table 3, which also appears in the Results section) using the same rules described for the 2000–2003 average to estimate cross-lagged regressions, which will be explained later.

INDEPENDENT VARIABLES

The primary independent variable for the analysis was proactive policing. Following Sampson and Cohen (1988), this variable is a ratio measured as the sum of the number of arrests for driving under the influence and disorderly conduct that is divided by the number of sworn police officers. For the cross-sectional analyses, we computed this ratio for each year between 2000 and 2003, averaged the values for years with nonmissing data, and imposed the requirement of a 2-year minimum for data inclusion. Proactive policing ratios were transformed to natural log scale in our cross-sectional analyses. We used single-year proactive policing ratios for the 1996–2003 panel-data analyses (reported in tables 2 and 4 in the Results section) and the 1996–1999 average for the cross-lagged regression analysis (reported in table 3).12

In addition to examining the direct effect of proactive policing on robbery rates, Sampson and Cohen (1988) estimated an indirect effect via the robbery arrest–offense ratio. We computed an arrest–offense ratio by dividing the number of robbery arrests by the number of robberies known to the police. These ratios were averaged across as many years within the period with nonmissing data, with a 2-year minimum to be included in the cross-sectional analyses. The arrest–offense ratio for our dynamic panel analyses remained single-year measures for the analysis period of 1996–2003.13

In the interest of replication, we collected data on the same set of independent variables that appeared in Sampson and Cohen's (1988) models—

---

12. In 2002, proactive policing was problematic because only 20 valid cases were found. This issue was mostly because of missing data for the DUI and disorderly conduct arrests. As explained subsequently, the Arellano–Bond dynamic panel data estimator does not require complete data, so unbalanced data and missing observations no longer pose a threat to unbiased estimation. For all other years, missing data for proactive policing posed no serious problems.

13. Again, the year 2002 had the most missing data for the variable construction (robbery arrest information was missing); thus, only 51 cases for the arrest–offense ratio in 2002 were nonmissing.
PROACTIVE POLICING AND ROBBERY RATES

city population size (logged), median family income, percentage of the population aged 15 years and older that is divorced, the proportion of the population that is non-Hispanic Black, racial income inequality (ratio of the White per capita income to Black per capita income), and a binary variable that indicated cities located in the West as defined by the U.S. Census. To extend Sampson and Cohen’s (1988) models to include additional covariates, we collected data from the 2000 U.S. Census on measures commonly used in the literature, specifically poverty (percent of the population below the poverty line), unemployment (percent of the population aged 16 years and older who were unemployed), education (percent of the population aged 18 years and older who graduated from high school), female-headed households (percent households with children less than age 18 years, a female householder, and no husband present), residential instability (percent of the population aged 5 years and older who moved into a different house within the last 5 years), and percent young males (percent of the population that is male aged 15–24 years).

As noted, a growing body of literature has identified these covariates as robust predictors of violent crime rates (Krivo and Peterson, 2000; Land, McCall, and Cohen, 1990; Messner and Golden, 1992; Parker and McCall, 1999; Wadsworth and Kubrin, 2004). Many are highly correlated, however, and produce collinearity when individually included in models. To reduce collinearity, we conducted principal components analysis on residential instability, racial inequality, percent young males, percent divorced, percent poverty, percent unemployed, percent high-school graduates, percent female-headed households, median family income, and percent Black. Consistent with prior research (Land, McCall, and Cohen, 1990), a “disadvantage” component emerged from the results. The following variables loaded strongly on this component (factor loadings in parenthesis): percent female-headed households (.90), percent poverty (.92), percent unemployed (.90), median family income (–.89), and percent high-school graduate (–.75). With an Eigen value of 3.8, this component accounted for 76 percent of the variation in the construct. We computed the index using the corresponding factor scores as multipliers.

We collected the same information from the 1990 U.S. Census and applied linear interpolation–extrapolation to construct annual values of the sociodemographic variables between the 1990 and the 2000 U.S. Census for our panel-data analyses. Our 1990 disadvantage component was consistent with the 2000 measure. The corresponding loadings for the 1990 factor (in parenthesis) were as follows: percent female-headed households (.90), percent poverty (.91), percent unemployed (.90), median family income (–.89), and percent high-school graduate (–.75). We corrected for polarity to ensure the index accurately reflected disadvantage for the 2000 and 1990 measures.
income (–.84), and percent high-school graduate (–.84). The 1990 measure of disadvantage had an Eigen value of 3.9 and accounted for approximately 77 percent of the variation of the construct.

Information on the city political system characteristics was derived from two sources—the 2001 Form of Government (FOG) survey conducted by the International County/City Management Association and the National League of Cities (NLC) Municipal Officials Database, which was provided in a personal communication by John Miller. The index, ranging from 0 to 3, was composed of three elements relating to city political organization. The index increased by 1 for cities with mayor-council forms of government. It also increased by 1 if some or all city council members represented specific geographic areas. Finally, the index increased by 1 if city elections were partisan. Thus, the traditional government index is maximized at a value of 3 for areas with a mayor-council form of government, city council members that represent districts, and partisan local elections. Unfortunately, information on local political arrangements was missing for some cities in the sample. To reduce the number of missing cases, Internet research determined the nature of local political arrangements from official city websites. Cities were assumed to be mayor-council unless a city manager was listed on the website. Similarly, if city council member lists referred to districts, then they were considered to be district cities, and so on.

ANALYTIC FRAMEWORK

We began the analysis with a replication of Sampson and Cohen (1988), which entailed two cross-sectional ordinary least-squares (OLS) regressions with the arrest–offense ratio and the logged robbery rate serving as dependent variables using Sampson and Cohen’s set of covariates. Second, we expanded their model for predicting robbery rates to include additional key covariates. Third, we estimated panel-data models that allowed us to take advantage of the temporal order of the relationship between proactive policing (X) and robbery rates (Y).

One of the most challenging tasks for statistical modeling is making causal inferences from nonexperimental data. Panel data offer decided advantages over cross-sectional designs for the analysis of causal interrelationships among variables (Finkel, 1995). Cross-sectional data can provide evidence of covariation, but causal inference must be grounded in strong theory. In its absence, panel data enable the specification of models that satisfy the time precedence criterion for a causal effect to exist and can be

15. This raised potential causal order issues. It is unlikely, however, that crime in a particular year will lead to changes in the structure of city political systems. Moreover, city political system characteristics tend to be fairly stable across time.
used to control for the effects of outside variables that otherwise might produce spurious association. As discussed, prior theory offers few guidelines that concern the precise nature of the temporal process that might link proactive policing with robbery rates. Accordingly, we explored five types of panel-data models that captured the following hypothesized effects: 1) the contemporaneous effect of proactive policing on the level of robbery rates; 2) the effect of lagged proactive policing on the subsequent level of robbery rates; 3) a fixed-effects model of lagged proactive policing on the subsequent level of robbery rates; 4) the effect of changes in proactive policing on contemporaneous changes in robbery rates; 5) the effect of lagged proactive policing on the subsequent level of robbery rates with lagged robbery rates included as a predictor. Taken together, these models allowed us to examine the effect of proactive policing on robbery rates net of other variables and to control for endogeneity while allowing us to probe several important assumptions underlying models in previous research.

Another difficult problem in nonexperimental research is how to control statistically for unobserved differences that might produce spurious correlations. We might cast this issue as a question of making it possible to control for variables that have not or cannot be measured, or we might interpret this issue as a question of whether “to pool or not to pool” the data when we have repeated observations for the same units. Although it is true that all units differ, an assumption that the regression function is constant across time and space yields efficiency and parsimony. Yet, a pooled model can be misspecified greatly if this assumption is incorrect. From either perspective, a solution can be found using only within-unit, over-time differences, which essentially discards any information about differences between units. This class of regression models—known as fixed-effects models—allows each unit to serve as its own control so that the change in the outcome of each unit is solely a function of change in its independent variables. The essence of a fixed-effects model is to examine how within-unit, over-time variation in explanatory variables is related to over-time variation in the dependent variable. By discarding between-unit variation, fixed-effects models eliminate the source of variation likely to be confounded with unobserved characteristics of the units (Allison, 2005). In so doing, any correlations between the unobserved variables with the observed variables are inconsequential.

The simplest way to purge all between-unit variation, and retain only within-unit, over-time variation, is to allow each unit its own intercept by entering \( N - 1 \) dummy variables into the regression equation. This approach, known as the least-squares dummy variable (LSDV) estimation,
is computationally burdensome for large $N$. Fortunately, mean differencing within units, say, $y_{i,t}^* = y_{i,t} - \bar{y}_i; x_{i,t}^* = x_{i,t} - \bar{x}_i$, produces the same results (Greene, 2002)\textsuperscript{16}. It is also easy to show that first differencing, rather than mean differencing, also can eliminate unobserved variables. In the simple case of $T = 2$, we could write separate time 1 and time 2 equations as follows:

$$y_{i1} = \alpha_1 + x_{i1} \beta + z_i \gamma_1 + \epsilon_{i1} \quad i = 1, 2, ..., N \quad (1)$$

$$y_{i2} = \alpha_2 + x_{i2} \beta + z_i \gamma_2 + \epsilon_{i2}$$

If we consider $z$ to be unobserved variables (referred to as the unit effect, $u_i = z_i y$) and subtract the time 1 equation from the time 2 equation, then the difference model is expressed as follows:

$$y_{i2} - y_{i1} = \alpha^* + \beta (x_{i2} - x_{i1}) + (\gamma_2 - \gamma_1) z_i + \epsilon_i^* \quad (2)$$

where $\alpha^* = \alpha_2 - \alpha_1, \epsilon^* = \epsilon_2 - \epsilon_1$, and if we assume that $\gamma_1 = \gamma_2$, then $z_i \gamma_2 - z_i \gamma_1 = 0$. Unobserved (and time-invariant) variables were differenced out of the equation.

Yet another statistical issue in panel-data analysis concerns the merits versus the deficits of dynamic specifications (Achen, 2000; Allison, 1990; Beck and Katz, 2004; Keele and Kelly, 2006). This debate addresses the specification of a lagged dependent variable when modeling change. Although the protagonists in this debate are far from agreement, we accept the position that the issue of whether to include a lagged dependent variable comes down to the following theoretical question: “Does the past matter for the current values of the process being studied?” If one suspects that it does, then the inclusion of a lagged dependent variable is appropriate. Because prior crime levels, for many reasons, are likely to influence future crime levels, we believe this method is appropriate to approach panel-data models of crime.

If we combine the features of differencing to remove spuriousness (from unobserved variables) with the theoretical rigor of allowing past values to determine current ones, then we have a model that is well positioned to estimate unbiased effects of explanatory variables\textsuperscript{17}. Consider a model

\textsuperscript{16} In an excellent text, \textit{Fixed Effects Regression Methods for Longitudinal Data Using SAS}, Allison (2005) demonstrates the equivalence between LSDV—estimated fixed effects and “conditioning out” all stable characteristics of cross-sectional units via OLS estimation on within-unit mean deviation scores. Allison also presents a comprehensive comparison of the strengths and weaknesses of fixed-effects versus random-effects methods, which explicates a powerful “hybrid method” that addresses the limitations of each approach by blending the strengths of both.

\textsuperscript{17} Differencing also presents a solution to the difficulty that results when one tries to place the fixed-effects model in the context of a dynamic panel model because
containing a lagged dependent variable and an explanatory variable, \( x \), which is expressed as follows:

\[
y_{it} = \alpha + \rho y_{i,t-1} + \beta_1 x_{it} + u_i + \varepsilon_{it} \tag{3}
\]

where \( u_i = z_i y \) is referred to as the unit effect and \((u_i + \varepsilon_{it})\) is known as the composite-error term. The first difference transformation removes both the constant term and the unit effect (i.e., impact of unobserved variables, where \( u_i = z_i y \)) and is expressed as follows:

\[
y_{it} - y_{i,t-1} = \rho (y_{i,t-1} - y_{i,t-2}) + \beta_1 (x_{it} - x_{i,t-1}) + (\varepsilon_{it} - \varepsilon_{i,t-1}) \tag{4}
\]

or more simply as

\[
\Delta y_{it} = \rho \Delta y_{i,t-1} + \beta_1 \Delta x_{it} + \Delta \varepsilon_{it} \tag{5}
\]

The differenced lagged dependent variable will be correlated with the error term (the former contains \( y_{i,t-1} \), whereas the latter contains \( \varepsilon_{i,t-1} \)), but with the unit effect \( u_i \) swept out, we might construct instruments for the lagged dependent variable from the second and third lags of \( y \) in the form of differences or lagged levels. Depending on the strength of \( \rho \) and if \( \varepsilon \) is independent and identically distributed (i.i.d.), lags of \( y \) will be correlated highly with the lagged dependent variable (and its difference) but uncorrelated with the composite-error term (Anderson and Hsiao, 1982; Arellano and Bover, 1995; Baum, 2006: 233). This form of instrumental variable estimation in the context of the dynamic panel-data model is the essence of the Arellano–Bond estimator (Arellano and Bond, 1991).

Although the conventional panel design allows for modeling a temporal sequence, the successful estimation of causal effects in the presence of endogeneity is rarely unambiguous, even with panel data. The assumption that proactive policing (as well as all other independent variables) are strictly exogenous means that their observed values are uncorrelated with the regression error in a model that predicts robbery rates. If proactive policing is not strictly exogenous—if robbery rates influence policing—then the assumption for OLS to estimate the parameters of the structural model and their accompanying test statistics consistently is undermined, and the parameters are not identified.

It is relatively easy to show this. Consider our point of departure, equation 2, in which we estimated the impact of change (first difference) in proactive policing on the change (first difference) in robbery rate, with the following simple linear model using OLS:

\[
\Delta \text{Rob}rati = \beta \Delta PP_i = \Delta u_i \tag{6}
\]
where $\Delta \text{Robrati}$ is the first difference in robbery rate in locality $i$ between time $t$ and $t - 1$, $\Delta PP_i$ is the first difference in the measure of proactive policing, and $\Delta u_i$ is the error term. Assume, for the moment, that equation 6 is well specified but also that it captures only part of the picture—causality also runs from the change in the level of robbery rate to proactive policing. Here we can specify the following:

$$\Delta PP_i = \gamma \Delta \text{Robrati}_i + \Delta u_{p.pi}$$

(7)

which shows that the change in the level of proactive policing is influenced by robbery rates, wherein we expect $\gamma > 0$ because police agencies adopt the proactive policing innovation in response to increasing crime rates. If equation 6 is estimated by OLS, but the true set of relationships is captured by equations 6 and 7 together, then the estimated effect of proactive policing on robbery rates, $\beta$, will not be “consistent”—it will suffer from “endogeneity” or “simultaneity” bias—because the regressor $\Delta PP$ is endogenous in a system of simultaneous equations, which correlates it with the error term $\Delta u_i$ in equation 6; $\beta$ will be biased upward by the positive quantity $\gamma$. Indeed, if the reverse causality is strong enough (i.e., if $\gamma$ is large relative to $\beta$), we would infer that $\beta > 0$ even if the true impact of proactive policing on crime is negligible or negative.

In cross-sectional data, the estimation of reciprocal causal-effects data relies on incorporating instrumental variables that satisfy several restrictive assumptions about the relationship between these variables with $x$ and $y$ as well as the error terms in their respective equations. The temporal component of panel designs, however, gave us leverage to observe how prior values of $x$ influence future values of $y$ and vice versa. Accordingly, we proceeded to estimate cross-lagged regression models (Finkel, 1995) to elicit evidence, or lack thereof, of endogeneity between proactive policing and robbery rates.

Because our baseline replication of Sampson and Cohen’s (1988) cross-sectional analyses used averaged data from 2000 to 2003, and because we extended the panel-data structure for our investigation of endogeneity backward across the preceding 4-year period and averaged the scores from 1996 to 1999, we now can specify a two-wave design (wave 1 for 1996–1999 and wave 2 for the original 2000–2003 period). The cross-lagged effects model then can be shown as a two-dependent variable extension of a two-period conditional change- or static-score model (Finkel, 1995). The two structural equations can be written as follows:

$$\Delta y_{i2} = \beta_1 \Delta x_{i1} + \beta_2 \Delta y_{i1} + \Delta u_{i1}$$

$$\Delta x_{i2} = \beta_3 \Delta y_{i1} + \beta_4 \Delta x_{i1} + \Delta u_{i2}$$

(8)
PROACTIVE POLICING AND ROBBERY RATES

Visually the two-wave cross-lagged effects model has the form shown in figure 1, where $\Delta x_1$ and $\Delta x_2$ represent the average annual change in proactive policing at waves 1 and 2, respectively, and $\Delta y_1$ and $\Delta y_2$ are the corresponding average annual change in robbery rates.

**Figure 1. Two-Wave Cross-Lagged Effects Model**

The correlation between the wave 1 variables is given by $\rho_1$, and the correlation between the structural disturbances of wave 2 is captured in $\rho_{u1u2}$ (see table 3 in the Results section). Under assumptions that disturbance terms $\Delta u_1$ and $\Delta u_2$ have zero means, have constant variance, and are uncorrelated with the lagged endogenous variables $\Delta x_1$ and $\Delta y_1$, the cross-lagged model parameters are estimated consistently via OLS.

The two-wave cross-lagged effects model has wide applicability and represents the most basic model to estimate possible reciprocal effects. The coefficients of interest are the cross-lagged effects from the wave 1 variables to the wave 2 outcomes (i.e., $\beta_1$ and $\beta_2$) because these figures represent the estimate causal effect from each variable to the other. The cross-lagged model corresponds to the Granger test for causality in time-series analysis, which posits that a variable “Granger-causes” the other if any value of the first variable measured at time $t-1, t-2, \ldots, t-m$ has a significant effect on the second variable at time $t$, which controls for the prior values of the second variable (Finkel, 1995: 24–8). In the language of the fixed-effects model, the effect of $\Delta y_1$ on $\Delta x_2$, which controls for $\Delta x_1$, represents the effect of the average change (between 1996 and 1999) in robbery rates in the $i$th city on the average change in proactive policing in the subsequent period (between 2000 and 2003) in the same $i$th city after accounting for the impact of change in proactive policing during the prior
(1996–1999) period, and vice versa for the effect of $\Delta x_1$ on $\Delta y_2$, which controls for $\Delta y_1$. To anticipate the results, cross-lagged regression models suggest that proactive policing responds to prior change in robbery in exactly the manner described in equation 7 with $\gamma > 0$.

With evidence that proactive policing is endogenous to robbery rates, our final model, and our concluding inference, was derived from the Arellano–Bond dynamic panel-data estimator (Arellano, 2003; Arellano and Bond, 1991; Baum, 2006). The Arellano–Bond estimator was designed for situations that embody every aspect of our research problem (Roodman, 2009). First, our data were observed in relatively few time periods and were dominated by many more cross-sectional units, “small $T$, large $N$,” that is, a panel-data structure rather than a time-series cross-section structure (Beck and Katz, 2004). Second, we assumed that the data-generating process was dynamic and that the current realizations of the dependent variable were influenced by past realizations. Third, we allowed that our panel data arbitrarily have distributed fixed individual effects, which admitted that each unit had its own data-generating process rather than one shared process. This process is equivalent to the assumption that unobserved (and possibly unobservable) variables might produce spurious correlations. Fourth, we allowed that proactive policing was not strictly exogenous to robbery rates, which means that it is correlated with past and possibly current realizations of the error term. Other explanatory variables were assumed to be exogenous. Fifth, the only available instruments were “internal”—based on lags of the instrumented variables. Sixth, we assumed that the error variance (but not the first difference transform) might be heteroskedastic and autocorrelated within units but not across them. Finally, the data-generating process was suited to a linear functional relationship.

Appendix A reports relevant model diagnostics and the technical details of our use of the Arellano–Bond dynamic panel-data estimator. In non-technical and highly simplified terms, the procedure involved purging the unobservable individual effects by differencing, fitting a dynamic panel-data model by drawing instruments from within the data set with lags, and assessing the essential assumptions of no serial correlation and identification.

RESULTS

Sampson and Cohen (1988) hypothesized that proactive policing is likely to be related to robbery rates through the following two processes: an indirect effect via an increased risk of apprehension and a direct effect that reflects deterrence accompanying greater police visibility. The first equation (model I, in table 1) is relevant to the hypothesis of an indirect
Table 1. Cross-Sectional Regressions of Certainty of Arrest and Robbery Rates

<table>
<thead>
<tr>
<th></th>
<th>Certainty of Arrest</th>
<th>(log) Robbery Rate</th>
<th>(log) Robbery Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>β</td>
<td>b</td>
</tr>
<tr>
<td>(log) proactive policing</td>
<td>.000</td>
<td>.002</td>
<td>−.145*</td>
</tr>
<tr>
<td>(log) population</td>
<td>−.023*</td>
<td>−.153</td>
<td>.170*</td>
</tr>
<tr>
<td>Percent divorced</td>
<td>−.004</td>
<td>−.065</td>
<td>−.018</td>
</tr>
<tr>
<td>Western location</td>
<td>−.009</td>
<td>−.040</td>
<td>.230*</td>
</tr>
<tr>
<td>Racial inequality</td>
<td>−.016</td>
<td>−.058</td>
<td>.289*</td>
</tr>
<tr>
<td>Percent non-Hispanic Black</td>
<td>−.001*</td>
<td>−.224</td>
<td>.016*</td>
</tr>
<tr>
<td>Median income (in $1,000s)*</td>
<td>.002*</td>
<td>.217</td>
<td>−.032*</td>
</tr>
<tr>
<td>Disadvantage index</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Traditional government index</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Percent young males</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Percent moved</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>.563*</td>
<td>4.254*</td>
<td>2.199*</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.217</td>
<td>.764</td>
<td>.780</td>
</tr>
</tbody>
</table>

*Statistically significant for a two-tailed test at the .05 level.

In this equation, we regressed the arrest–offense ratio on proactive policing and control for the same variables specified in Sampson and Cohen’s analysis. Contrary to their findings, proactive policing had no effect on the arrest–offense ratio, which suggests that proactive policing does not reduce robbery rates by changing the actual probability that arrests are made. Considering Sampson and Cohen’s control variables, the results also indicated that certainty of arrest for robbery is lower in large cities and in cities with appreciable Black populations, and it is higher in more affluent cities as reflected in median income.

Turning to the prediction of robbery rates, model II is based on Sampson and Cohen’s (1988) specification. The evidence was consistent with the hypothesis of a direct effect of proactive policing on robbery rates. Proactive policing exhibits a significant negative association with robbery rates; cities with more arrests for DUI and disorderly conduct relative to the number of sworn police officers tend to have lower robbery rates, controlling for the structural characteristics in the model. With respect to the control variables, the percent of non-Hispanic Blacks in a city and racial inequality are associated positively with robbery rates. In addition, median income is associated significantly with robbery rates in an inverse direction, net of other factors. The results in model II also show that population
size is associated positively with robbery rates as is location in the Western United States. Contrary to Sampson and Cohen, the percent of people divorced in a city did not exert a significant effect on robbery in this model.

Model III in table 1 allows for more rigorous testing of the cross-sectional relationship between proactive policing and robbery rates. In this model, the measure of median income was incorporated in the disadvantage index, and measures of governmental structure, age and sex structure, as well as residential instability are included. Results reveal that the effect of proactive policing withstood the introduction of the additional controls. Population size and racial inequality continued to exhibit significantly positive effects. The dummy variable for “West” was no longer significant. The government index and the measures of age and sex structure as well as residential mobility also were not significant. Alternatively, as expected, the disadvantage index yielded a significantly positive coefficient. Indeed, it was by far the most powerful predictor of robbery rates, as indicated by the comparatively large standardized coefficient (.536).18

Turning to the panel analyses, we began by estimating conventional panel models that assumed all predictor variables were exogenous to robbery rates (see table 2). We considered five specifications. In the first, the robbery level (in log form) was regressed on contemporaneous values of predictor variables (model IV). In the second, capitalizing on the temporal ordering of cause and effect strengthened causal inference; accordingly, model V regressed (log) robbery rates on prior values (1-year lag) of the independent variables. Our next model (model VI) is a fixed-effects model.19 Differences in inference between models IV and VI revealed any pernicious effects of unit heterogeneity. Model VII expresses the outcome as a change score and estimates the effect of change in proactive policing on change in robbery rates, with other time-varying explanatory variables lagged. The final specification is a dynamic panel model in which robbery

---

18. Similar to Sampson and Cohen (1988), our results indicated the explanatory power for the models of robbery rates greatly exceeded that for arrest certainty, which is not surprising given the selection of control variables for the baseline model was informed by criminological theories of, and macrolevel research on, violent crime. Little theoretical guidance in the literature concerned the social structural determinants of arrest certainty.

19. We also fit random-effects panel models to account for possible unit heterogeneity (available upon request). Although fixed-effects models necessarily exclude variables with no within-unit variability, random-effects panel models permit the inclusion of indicator variables. However, in doing so, random-effects estimates use information both within and between units. The unobserved variables, therefore, must be assumed to be uncorrelated with all observed variables. Unless one makes this assumption, random effects do not really control for the effects of unobserved variables and fixed effects are preferable.
Table 2. Panel-Data Models Estimated Under the Assumption that All Predictor Variables Are Exogenous to Robbery Rate at Time $t$

<table>
<thead>
<tr>
<th></th>
<th>Model IV$^a$</th>
<th>Model V$^{ab}$</th>
<th>Model VI$^c$</th>
<th>Model VII$^{de}$</th>
<th>Model VIII$^{df}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(log)</td>
<td>(log)</td>
<td>(log)</td>
<td>$\Delta$</td>
<td>$\Delta$</td>
</tr>
<tr>
<td>Robbery Rate</td>
<td>$-0.088^*$</td>
<td>$-0.078^*$</td>
<td>$-0.039^*$</td>
<td>$-0.617^*$</td>
<td>$-0.018$</td>
</tr>
<tr>
<td>(log) proactive policing</td>
<td>$-0.088^*$</td>
<td>$-0.078^*$</td>
<td>$-0.039^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(log) population</td>
<td>$0.188^*$</td>
<td>$0.174^*$</td>
<td>$0.899^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divorce rate</td>
<td>$0.019^*$</td>
<td>$0.024^*$</td>
<td>$0.023^*$</td>
<td>$-0.844^*$</td>
<td>$-1.781^*$</td>
</tr>
<tr>
<td>Western location</td>
<td>$-0.300^*$</td>
<td>$-0.263^*$</td>
<td>$11.148^*$</td>
<td>$12.654^*$</td>
<td></td>
</tr>
<tr>
<td>Percent racial inequality</td>
<td>$0.282^*$</td>
<td>$0.293^*$</td>
<td>$-0.494^*$</td>
<td>$-8.884^*$</td>
<td>$-7.513^*$</td>
</tr>
<tr>
<td>Percent non-Hispanic Black</td>
<td>$0.012^*$</td>
<td>$0.013^*$</td>
<td>$-0.003^*$</td>
<td>$0.653^*$</td>
<td>$0.764^*$</td>
</tr>
<tr>
<td>Disadvantage index</td>
<td>$0.419^*$</td>
<td>$0.423^*$</td>
<td>$0.026^*$</td>
<td>$0.921$</td>
<td>$4.283$</td>
</tr>
<tr>
<td>Traditional government index</td>
<td>$0.067^*$</td>
<td>$0.079^*$</td>
<td>$3.840^*$</td>
<td></td>
<td>$5.705^*$</td>
</tr>
<tr>
<td>Percent young males</td>
<td>$-0.008$</td>
<td>$0.001$</td>
<td>$-0.237^*$</td>
<td>$-0.954$</td>
<td>$-1.801$</td>
</tr>
<tr>
<td>Percent moved</td>
<td>$0.000$</td>
<td>$0.004$</td>
<td>$-0.020^*$</td>
<td>$0.069$</td>
<td>$0.397$</td>
</tr>
<tr>
<td>Intercept</td>
<td>$2.154^*$</td>
<td>$1.933^*$</td>
<td>$-2.169$</td>
<td>$18.945$</td>
<td>$1.791$</td>
</tr>
<tr>
<td>rho</td>
<td>$0.704$</td>
<td>$0.743$</td>
<td>$0.972^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R$^2$</td>
<td>$0.960$</td>
<td>$0.950$</td>
<td>$0.050$</td>
<td>$0.050$</td>
<td>$0.950$</td>
</tr>
</tbody>
</table>

$N=1,180$  $N=1,035$  $N=1,180$  $N=834$  $N=1,029$

NOTES: Based on 180 panels (cities). Total number of observations in each model varies by lag structure and method of model estimation.

$^a$Prais–Winsten (iterative) estimation of regression coefficients with panel-corrected standard errors under the assumption that disturbances are heteroskedastic and common first-order autocorrelation (shown as $\rho$) within panels.

$^b$All predictor variables are in 1-year lags.

$^c$Unit heterogeneity removed via fixed effects ($y_{it} = \bar{y}_i - \bar{y}_t; x_{it} = \bar{x}_i - \bar{x}$); effects of Western location and traditional government index are excluded because of no interunit variability.

$^d$OLS estimation of regression coefficients with panel-corrected standard errors under the assumption that disturbances are heteroskedastic (no autocorrelation).

$^e$Change-score ($\Delta$) model wherein $y_{it} = y_{it} - y_{i,t-1}$ and proactive policing is first differenced ($pp_{it} = pp_{it} - pp_{i,t-1}$); all other predictor variables are in 1-year lags.

$^f$Static-score/regressor variable model.

$^*$Indicates that the $z$ ratio is significant at $p < .05$.

Robbery rates at time $t$ are regressed on robbery rates at time $t - 1$ as well as all other right-hand-side variables measured at time $t - 1$. Model VIII is the “conditional change/static-score” or “regressor variable” model (Allison, 1990; Finkel, 1995). It is important to note that model VIII is still a
model of change in robbery rates, but it differs from the specification in model VII because it assumes current robbery rates were influenced, in part, by past ones.

Table 2 evidences the impact of unit heterogeneity. Interestingly, proactive policing and population were the only explanatory variables that retained their signs and significance when we compared model IV with the fixed-effects model (model VI). By using only within-unit (over-time) covariation, we found that divorce rate, percent non-Hispanic Black, and disadvantage were no longer statistically significant. Percent young males and percent that moved shifted from nonsignificance to being significant and negative, whereas the sign for racial inequality was reversed. Clearly, unobserved characteristics of cities affected model estimates.

Our model estimation also gave us perspective on autoregressive error. Estimates of rho in models IV–VI were strong and positive and ranged from 0.7 to 0.97. This source of correlated error virtually is eliminated in our change models (VII and VIII); indeed, model specifications that set rho to zero (table 2) easily are preferred more than those that estimate this additional parameter.21

At first glance, the panel analyses results seem to yield inconsistent evidence about the deterrent effect of proactive policing on robbery rates. The coefficients for proactive policing in the equations predicting robbery were significant and negative, as in the cross-sectional analyses. A high level of proactive policing in a given year was associated with lower robbery rates, both contemporaneously and in the following year. However, the coefficients for proactive policing on change in robbery rates (including model VIII) were all nonsignificant. Thus, no consistent evidence in the conventional panel models of table 2 supported the claim that levels of proactive policing or changes in levels of proactive policing reduce crime.

However, before we dismissed the hypothesis that proactive policing deters robberies, we needed to bring evidence to bear on the assumption that proactive policing is exogenous to robbery rates. If a feedback loop exists in a structural model, then each of the variables in the loop will be correlated with the error term in the equation in which they are one of the independent variables, which means that the OLS coefficients will be biased estimates of the effects of the variables in the loop. Again, panel data offered a compelling format for assessing reciprocal causation. Table

---

Supplemental analyses showed that \( \rho = -0.09 \) in model VII specification and that \( \rho = 0.018 \) in model VIII specification.
PROACTIVE POLICING AND ROBBERY RATES

Table 3. Cross-Lagged Model to Assess Reciprocity

<table>
<thead>
<tr>
<th></th>
<th>ΔRobbery Rate</th>
<th>ΔY_i</th>
<th>ΔProactive Policing</th>
<th>ΔX_i</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>β</td>
<td>b</td>
<td>β</td>
</tr>
<tr>
<td>Lagged Δ proactive policing ΔX_i</td>
<td>1.230</td>
<td>.017</td>
<td>.039</td>
<td>.036</td>
</tr>
<tr>
<td>Lagged Δ robbery rate ΔY_i</td>
<td>.159</td>
<td>.122</td>
<td>.003*</td>
<td>.148</td>
</tr>
<tr>
<td>R²</td>
<td>.015</td>
<td>.022</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ρ_1</td>
<td>−.071</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ρ_e1u2</td>
<td>−.037</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>180</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTES: ρ_1 is the estimated correlation between Δproactive policing and Δrobbery rates at wave 1. ρ_e1u2 is the estimated correlation between the structural disturbances of the wave 2 equations. *Indicates that the z ratio is significant at p < .05.

The regression estimates shown in table 3 support our suspicion that proactive policing is endogenous to robbery rates. Of the four path coefficients estimated in our cross-lagged regression, only the effect of lagged change in robbery rates on subsequent change in proactive policing attained statistical significance. The sign is positive, and the standardized effect size is the largest among the four coefficients, which suggests that police agencies likely adopt more proactive policing strategies in response to increasing crime rates.

Given our results thus far, we closed our inquiry by estimating a dynamic panel-data model of the effect of proactive policing, which treats this model as an endogenous regressor and the other as an exogenous predictor, which include the arrest–offense ratio, on robbery rates using the Arellano–Bond estimator. In table 4, we report two sets of coefficients.

Dynamic models (regression on a lagged dependent variable) portray the time path of the dependent variable in relation to its past values. A justification for the inclusion of the lagged dependent variable often is tied to the concept of the stability of social systems (Coleman, 1968). A causal system is stable if it will approach a fixed equilibrium point in a future time period where the values of y for each case will be constant (Finkel, 1995: 9). Because most systems analyzed in empirical research have not yet reached equilibrium but are moving toward this state, the effect of the lagged dependent variable on Δy often is interpreted as a proxy for causal paths that link prior values of y to its future realizations through variables omitted from the model. The regression effects for the explanatory variables, therefore, are interpreted generally in two ways. In their raw form, they represent the short-term effect on y, or Δy, across the panel waves. If we remove the adjustment of y to its future equilibrium state, then we
Table 4. Arellano–Bond Dynamic Panel-Data Estimation of Effects of Exogenous and Endogenous Predictors of Change in Robbery Rates Between Time $t$ and Time $t - 1^a$

<table>
<thead>
<tr>
<th></th>
<th>Model XI$^b$</th>
<th>Long-Run Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robbery rate</td>
<td>.138</td>
<td>.862$^b$</td>
</tr>
<tr>
<td>Proactive policing</td>
<td>$-21.851^a$</td>
<td>$-25.339^a$</td>
</tr>
<tr>
<td>Population (1,000s)</td>
<td>-.147</td>
<td>-.170</td>
</tr>
<tr>
<td>Divorce rate</td>
<td>$-17.614$</td>
<td>$-20.426$</td>
</tr>
<tr>
<td>Percent racial inequality</td>
<td>99.942</td>
<td>115.899</td>
</tr>
<tr>
<td>Percent non-Hispanic Black</td>
<td>$-2.691$</td>
<td>$-3.121$</td>
</tr>
<tr>
<td>Disadvantage index</td>
<td>$-12.557$</td>
<td>$-14.562$</td>
</tr>
<tr>
<td>Percent young males</td>
<td>$-3.307$</td>
<td>$-3.835$</td>
</tr>
<tr>
<td>Percent moved</td>
<td>5.307</td>
<td>6.154</td>
</tr>
<tr>
<td>Certainty of arrest</td>
<td>2.965</td>
<td>3.439</td>
</tr>
<tr>
<td>Intercept$^d$</td>
<td>7.072$^a$</td>
<td>—</td>
</tr>
</tbody>
</table>

$^a$Based on 172 panels (cities) contributing between 1 and 6 years (average observations/panel = 3.59) of nonmissing observations.

$^b$Standard errors adjusted for within-unit heteroskedasticity.

$^c$Coefficient of adjustment.

$^d$The estimated intercept is an estimate of the coefficient on a time trend; it is the average linear change in robbery rate and has no separate long-run function.

$^*$ indicates that the Wald test statistic is significant at $p < .05$.

have an alternative interpretation as the “long-run” effects of $x$ on the equilibrium value of $y$. The long-run effect of a covariate usually is defined to be the sum of the current and lagged coefficients divided by 1 minus the sum of the lagged coefficients on the dependent variable. This latter quantity is referred to as the coefficient of adjustment (Johnston, 1972: 300–3). Both interpretations are reported in table 4. In the column labeled “$b$,” we give the short-term effects; the long-run effects are reported in the subsequent column.$^{22}$

The evidence clearly points to the deterrent effect of proactive policing on robbery. Specifically, in the short run, proactive policing reduces robbery rates by an average of almost 22 points (robberies per 100,000 population) for every additional increase in arrests for DUI and disorderly conduct per police officer, net of other effects and after accounting for its
endogeneity to robbery rates. This deterrent effect is estimated to bring a 25-point reduction when the system reaches its equilibrium state.

None of the coefficients (short or long run) for our exogenous variables attained statistical significance. Aside from certainty of arrest, this result was expected because the panel scores are the result of linear interpolation/extrapolation from observed scores in the 1990 and the 2000 U.S. Census. This near-stability eliminates almost all information because the first difference (change) is constant across time. As such, little methodological justification exists to make inferences or interpret their effects on the change in robbery rates. Rather, these variables are treated as controls to condition our investigation of the impact of proactive policing more.

SUMMARY AND DISCUSSION

The purpose of this research has been to revisit and extend an innovative study that reports an intriguing finding of considerable theoretical and practical significance. In their pioneering 1988 study, Sampson and Cohen reported a significant negative relationship between proactive policing and robbery rates. This finding has been interpreted as supporting a principal claim of deterrence theory, namely, that vigorous enforcement of laws that prohibit these forms of disorder increases police visibility and, thereby, acts as a deterrent against serious violent crimes. In terms of causal modeling, however, the evidence that we bring to bear in support of the desired impact of proactive policing is far more compelling than that from previous efforts.

We have reassessed the impact of proactive policing on robbery rates using more recent data for a comparable sample of cities. Our results have replicated the earlier findings of a negative association between proactive policing and robbery rates when an analogous model is estimated. We also have assessed the robustness of this relationship by expanding the set of structural control variables included in the equation. The results have revealed that the negative relationship initially observed by Sampson and Cohen (1988), and subsequently reaffirmed by MacDonald (2002), is indeed robust, at least for the sample under investigation. Proactive policing retains its statistically significant negative association with robbery rates in the more fully specified model.

In addition, we have explored potential implications of endogeneity by using a dynamic model from panel data whose practical performance has been thoroughly studied in the econometric literature. The results of our dynamic panel-data analyses have been supportive of those obtained in the OLS regressions, although model diagnostics unambiguously locate

---

23. The region dummy and traditional government index are time-invariant and, thus, “partialed out” of the regression.
proactive policing as endogenous with robbery rates. Interestingly, our results have not provided additional evidence of Sampson and Cohen's (1988) finding of an indirect effect of proactive policing through the arrest–offense ratio. Thus, the effect of proactive policing on robbery rates seems to be through generalized perceptions of greater law-enforcement activity, regardless of the probability that arrests are made.

Although these results are highly suggestive, we acknowledge limitations associated with our analyses. Our model specification is extensive and includes most of the “usual suspects” in macrolevel studies of violent crime, but causal inference with nonexperimental data is inherently precarious. The adequacy of our solution for simultaneity bias also is open to question. The use of the intertemporal covariance matrix of the validity errors to the instrumental variables requires strong assumptions that ultimately depend on a sound theoretical understanding of the causal processes involved. Unfortunately, existing theory offers little guidance as to whether the implied lag structure of this approach is appropriate. We, thus, do not claim to have resolved the long recognized, but still vexing, problem of disentangling causal order in the study of policing and crime. Our analyses do indicate, however, that more efforts along these lines are warranted, which include studies that systematically examine various lag structures using longitudinal data that encompass longer time spans.

A final challenge for future research is to “deconstruct” the notion of proactive policing. We have commended Sampson and Cohen (1988) for developing an innovative measure that has a certain degree of face validity and has the considerable benefit of being based on data collected as part of the UCR program, thereby, permitting analyses for relatively large samples of cities. However, some ambiguity is associated with the measure. It is intended to reflect more than what it does on the surface—the enforcement of laws against disorderly conduct and DUI. It presumably captures a more general style of policing, which results in a highly visible police presence.

But what specific styles of policing, in fact, are captured by the proactive policing indicator? The terms “proactive,” “quality-of-life,” or “zero tolerance” policing have been used loosely to encompass a wide array of enforcement approaches (Taylor, 2001), which might have different consequences for crime. Moreover, the proactive policing measure could serve as a proxy for other policing variables associated with the prevention of robbery, such as increased training, or the use of technologies, such as geographic information systems. Unfortunately, the data necessary to “unpack” proactive policing into the elements of policing most relevant to robbery are not as widely available as the data used to construct the current measure. It, thus, is not possible here to identify the precise mechanisms that might be at work, but the results of our analyses, in conjunction
with earlier studies, provide ample grounds for more exploration into the connection between distinctive policing styles and violent crime rates. With greater knowledge of such mechanisms, it might be possible to design randomized experiments or quasi-experiments that would overcome the limitations inherent in any effort to make causal inferences on the basis of statistical modeling of correlational data (see Berk, 2005).

We close with a comment on policy implications. Policy makers sometimes express frustration with sociological analyses of crime on the grounds that the primary variables under consideration rarely are amenable to direct manipulation (for a classic statement, see Wilson, 1975). Policing, in contrast, is a factor that should be responsive, at least to some degree, to deliberate decisions about social policy. The identification of deterrent effects of particular styles of policing, thus, holds considerable promise for crime control. We caution, however, against a narrow approach to the general question of what kinds of policing strategies should be promoted. Policing might be important for crime control, but in a democratic society, it is ultimately part of a larger institutional framework dedicated to criminal justice. As Manning (2005) observed, democratic policing aims to be fair, just, and humane, in addition to providing for the security of the population. Thus, whatever the deterrent effects of proactive or any other policing style might prove to be, policy decisions need to be informed not only by considerations of crime control but by the fundamental values of a democratic society.

REFERENCES

Achen, Christopher. 2000. Why lagged dependent variables can suppress the explanatory power of other independent variables. Paper presented at the Annual Meeting of the Society for Political Methodology, Los Angeles, CA.


Baum, Christopher F. 2006. *An Introduction to Modern Econometrics Using Stata*. College Station, TX: Stata Press.


PROACTIVE POLICING AND ROBBERY RATES


Messner, Steven F., Sandro Galea, Kenneth J. Tardiff, Meliss Tracy, Angela Bucciarelli, Tinka M. Piper, Victoria Frye, and David Vlahov. 2007. Policing, drugs, and the homicide decline in New York City in the 1990s. *Criminology* 45:385–413.


PROACTIVE POLICING AND ROBBERY RATES


StataCorp. 2003b. *Stata Statistical Software: Release 8.0*. College Station, TX: Stata Corporation.


Charis E. Kubrin is an associate professor of sociology at George Washington University and a research affiliate at the George Washington Institute of Public Policy. Her research focuses on neighborhood correlates of crime, with an emphasis on race and violent crime. A new line of research examines the intersection of music, culture, and social identity, particularly as it applies to hip-hop and minority youth in disadvantaged communities. In addition to publications in professional journals, Charis is coeditor of Crime and Society: Crime, 3rd edition, and coauthor of Researching Theories of Crime and Deviance and Privileged Places: Race, Residence, and the Structure of Opportunity.
PROACTIVE POLICING AND ROBBERY RATES  91

Steven F. Messner is the Distinguished Teaching Professor of Sociology at the University at Albany, State University of New York. His research focuses on social institutions and crime, understanding spatial distributions and trends in crime, as well as crime and social control in China. In addition to his publications in professional journals, he is coauthor of Crime and the American Dream and Perspectives on Crime and Deviance and is coeditor of Theoretical Integration in the Study of Deviance and Crime, and Crime and Social Control in a Changing China.

Glenn Deane is an associate professor of sociology at the University at Albany, State University of New York. His research activities include modeling the interrelationship between population and environment, multiple race identifications, and intrafamily dynamics.

Kelly McGeever is a PhD candidate in the Department of Sociology at the University at Albany, State University of New York. Her research interests include communities and crime as well as GIS applications and the spatial analysis of crime. Her dissertation tests NIMBY arguments about the negative effects of group quarters, such as halfway houses, on neighborhood quality. Her work has appeared in Criminology and the Journal of Quantitative Criminology.

Thomas D. Stucky is an interim director of Indiana University’s Center for Criminal Justice Research and an associate professor of criminal justice in the School of Public and Environmental Affairs at Indiana—Purdue University at Indianapolis. His scholarly interests include the intersection of politics and criminal justice, specifically the relationship between politics and crime—policing at the city and state levels as well as trends in imprisonment and correctional spending. He also is interested in the continuing development of the systemic social disorganization theory as well as the relationship among land use, the physical environment, and crime. He has authored two books and journal articles that appear in Criminology, Journal of Quantitative Criminology, Journal of Research in Crime and Delinquency, and Justice Quarterly, among others.
Appendix A. Technical Details on the Arellano–Bond Dynamic Panel Estimator

Instrumental variable (IV) estimation provides one solution to the problem of an endogenous explanatory variable. This solution requires at least one variable correlated with proactive policing (“instrument relevance”) and is uncorrelated with $u_t$, the error term in the robbery rates prediction equation (“instrument validity”). Because IV estimation relies on satisfying restrictive assumptions, it can fall short. Statistical guidelines are complicated, potentially misleading, and interlaced with decision making (with all the pitfalls of model-selection criteria). For example, instrument relevance implies strong instruments. The consequences of weak instruments are particularly damaging to the IV approach, but finding strong instruments is not an easy undertaking and adding instruments comes with a cost—the finite sample bias of the IV estimator increases with the number of instruments. From the other perspective, testing the instrument validity assumption relies on the validity of orthogonality conditions that correspond with the instruments not being tested (Baum, 2006: 200–2).

IV estimation in panel data was originally proposed by Anderson and Hsiao (1981, 1982) as an alternative to the fixed-effects model to sweep out individual (unit) effects in a model with a lagged dependent variable. The Anderson–Hsiao estimator begins by handling the unit effects by first differencing and then using an instrumental variable correlated with the lagged first-differenced dependent variable but not the differenced error term. In addressing the criticism that the instrument proposed by Anderson and Hsiao is weak, Arellano and Bond (1991) suggested a much more efficient estimator that uses all available lags at each observation as instruments. It is essential to their use of the intertemporal covariance matrix of the errors and to the validity of the instrumental variables that the errors in fact are serially uncorrelated. If this condition is not met, then the estimators lose consistency.

Because of the critical essence of the lack of serial correlation, Arellano and Bond (1991) paid close attention to testing the specification and offered both a direct test for $m$-order serial correlation based on the differenced residuals and Sargan tests (closely related to a Hausman test) of overidentifying restrictions (1991: 281–3). As general specification tests, the Sargan tests have a dual utility; when used in conjunction with the direct test for serial correlation, the Sargan tests can help shed light on the causal designation of right-hand side (RHS) variables because the Sargan

---

24. Presentation of the material in this appendix follows closely Baum (2006) and documentation of the XTABOND procedure in the *Stata Cross-Sectional Time-Series Reference Manual* (StataCorp, 2003a: 15–33). We used the XTABOND procedure to estimate the models presented herein (StataCorp, 2003b).
tests can be used to exploit differences in the overidentifying restrictions based on treatment of RHS variables (1991: 291).

The Arellano–Bond estimator fits a dynamic panel-data model as follows:

\[ y_{it} = \sum_{j=1}^{p} \alpha_j y_{i,t-j} + x_{it} \beta_1 + w_{it} \beta_2 + v_i + \epsilon_{it} \quad i = 1, \ldots, N; \ t = 1, \ldots, T_i \]  

(A.1)

where the \( \alpha_j \) are parameters to be estimated, \( x_{it} \) is a \( 1 \times k_1 \) vector of strictly exogenous covariates, \( \beta_1 \) is a \( k_1 \times 1 \) vector of parameters to be estimated, \( w_{it} \) is a \( 1 \times k_2 \) vector of predetermined or endogenous covariates, \( \beta_2 \) is a \( k_2 \times 1 \) vector of parameters to be estimated, \( v_i \) are the random effects that are i.i.d. across the panels (cities) with constant variance \( \sigma_v^2 \), and \( \epsilon_{it} \) are i.i.d. across the whole sample with constant variance \( \sigma_e^2 \).

Arellano and Bond (1991) derived a generalized method of moments (GMM) estimator, known as the Arellano–Bond dynamic panel-data estimator, for \( \alpha_j, j \in \{1, \ldots, p\} \), \( \beta_1 \) and \( \beta_2 \) using lagged levels of the dependent variable as well as the predetermined and endogenous variables, as well as differences of the strictly exogenous variables. First differencing equation A.1 removes \( v_i \) and produces an equation estimable by instrumental variables. Note also that \( x_{it} \) and \( w_{it} \) might contain lagged independent variables and time dummies, but time-invariant covariates, like WEST and TGI in our analyses, are partialed out of the estimation by differencing.

Our OLS regressions assumed that all variables on the RHS, which included proactive policing, are strictly exogenous. In the language of simultaneous-equation models, variables are termed endogenous if their values are determined within the model and predetermined if their values are determined outside the model, with endogenous variables regarded as stochastic and predetermined variables treated as nonstochastic. Determined variables are divided into the following two categories: exogenous and lagged endogenous. This latter classification might seem counterintuitive given the preceding definitions, but because the value of a lagged endogenous variable is known at the current time, it is regarded as nonstochastic, therefore, predetermined (cf. Johnston, 1972). To elaborate, a variable \( x_{it} \) is said to be strictly exogenous if \( E[x_{it} \epsilon_{is}] = 0 \) for all \( t \) and \( s \), for panel \( i \) in time \( t \). In many cases, this assumption is not tenable. Intuitively, if the error term at time \( t \) has some feedback on the subsequent realizations of \( x_{it} \), then \( x_{it} \) is a predetermined variable. Endogenous variables differ from predetermined variables in that the former allows for correlation between the \( x_{it} \) and the \( v_i \) at time \( t \), whereas the latter does not. Formally, an independent variable is predetermined if \( E[x_{it} \epsilon_{is}] \neq 0 \) for \( s < t \), but \( E[x_{it} \epsilon_{is}] = 0 \) for \( s \geq t \), and endogenous if \( E[x_{it} \epsilon_{is}] \neq 0 \) for \( s \leq t \), but \( E[x_{it} \epsilon_{is}] = 0 \) for \( s > t \).
The Arellano–Bond estimator uses differences as instruments for strictly exogenous variables and levels lagged one or more periods as instruments for predetermined variables. Endogenous variables are treated similarly to the lagged dependent variable, with levels of endogenous variables lagged two or more periods serving as instruments. In addition, the Arellano–Bond approach easily handles missing observations in the interior of the panels. A simple balanced panel with no missing values and no predetermined or endogenous variables (aside from the lagged dependent variable) needs to be considered first:

\[ y_{it} = y_{i,t-1} \alpha_1 + y_{i,t-2} \alpha_2 + x_{it} \beta + v_i + \epsilon_i \]  
(A.2)

after first differencing equation A.2, then

\[ \Delta y_{it} = \Delta y_{i,t-1} \alpha_1 + \Delta y_{i,t-2} \alpha_2 + \Delta x_{it} \beta + v_i + \Delta \epsilon_i \]  
(A.3)

The first three observations are lost to lags and differencing. Because \( x_{it} \) contains only strictly exogenous covariates, \( \Delta x_{it} \) will serve as its own instrument in estimating the first-differenced equation A.3, assuming that \( \epsilon_i \) are not autocorrelated. For each \( i \) at \( t = 4 \), \( y_{i1} \) and \( y_{i2} \) are valid for the lagged variables. Similarly, at \( t = 5 \), \( y_{i1}, y_{i2}, \) and \( y_{i3} \) are valid instruments. Continuing in this fashion, an instrument matrix with one row for each time period being instrumented is obtained as follows:

\[
Z_t = \begin{pmatrix}
  y_{i2} & y_{i3} & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \Delta x_{i5} \\
  0 & 0 & y_{i1} & y_{i2} & y_{i3} & \ldots & 0 & 0 & 0 & \Delta x_{i6} \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots \\
  0 & 0 & 0 & 0 & \ldots & 0 & y_{i2} & \ldots & y_{i,T-2} & \Delta x_{iT}
\end{pmatrix}
\]

Because the number of lags of the dependent variable \( p \) is 2, \( Z \) has \( T - p - 1 \) rows and \( \sum_{m+k_i} \) columns, where \( k_i \) is the number of variables in \( x \). The extension to other lag structures is readily apparent, and unbalanced data and missing observations are handled by dropping rows for which no data are available and placing zeros in columns where missing data occur. For predetermined variables, levels lagged one or more periods serve as valid instruments, endogenous variables are treated similarly to lagged dependent variables, and lagged levels of two or more periods serve as valid instruments.

Although Arellano and Bond (1991) derived one-step and two-step GMM estimators, they recommended using the one-step estimator for inference on the coefficients. The two-step Sargan test, a test of overidentifying restrictions, however, might be better for inference on model specification because the one-step Sargan test overrejects in the presence of heteroskedasticity. When the number of orthogonality conditions \( r \) exceeds the number of parameters \( k \), testable restrictions are implied in
the econometric model. Estimation of $\beta$ in the linear regression, $y = x'\beta + u$, sets to zero $k$ linear combinations of the $r$ sample orthogonality conditions $b_N(e)$. When the model is right, $r - k$ linearly independent combinations of $b_N(\hat{\beta})$ should be close to zero.

The main result here is that a minimized optimal GMM criterion scaled by $N$ has an asymptotic chi-square distribution with $r - k$ degrees of freedom (Arellano, 2003: 192–3). A statistic of the following form:

$$N^{-1}(\hat{\beta}) = N b_N(\hat{\beta})^{1/2} b_N(\hat{\beta})$$

is called a Sargan test statistic. The Sargan test statistic for the two-step Arellano–Bond estimator is as follows:

$$S_2 = \left( \sum_{i=1}^{N} \hat{e}_i^* Z_i \right) A_2 \left( \sum_{i=1}^{N} Z_i' \hat{e}_i^* \right)$$

where $\hat{e}_i = y_i^* - X_i^* \hat{\beta}$ are the two-step residuals for the differenced $y_i$ and $X_i$, and $\hat{\beta}$ is the two-step estimator of the $K \times 1$ vector of coefficients, $Z_i$ is as defined, and $A_2 = \sum_{i=1}^{N} Z_i' G_i Z_i$, with $G_i = \hat{e}_i^* \hat{e}_i^*$ estimated by the one-step residuals.

Additionally, Sargan tests are sensitive to the instrument set specified in the $Z_i$ matrix (Arellano and Bond, 1991: 291, table 4). Contradictory assessment to the direct test for serial correlation might reflect the failure of the exogeneity assumption, which suggests that alternative forms of the instrument matrix $Z_i$ should be considered. Our postestimation assessment of model specification, two-step Sargan test of overidentifying restrictions ($\chi^2(33) = 44.78; Pr > \chi^2 = .08$) supported the adequacy of our instrumentation and the appropriateness of our assignment of proactive policing as endogenous and all other explanatory variables, which included certainty of arrest, as strictly exogenous.

Arellano and Bond (1991) also derived a direct test for $m$-order serial correlation based on the differenced residuals. Our assumption of no autocorrelation was evidenced in both the one-step GMM test ($z = -.60; Pr > z = .55$) and the two-step GMM test that the average autocovariance in the residuals of order 2 is 0 ($z = -.39; Pr > z = .69$). In short, model diagnostics overwhelmingly favor the regression estimates reported in table 4.

As noted in the body of the article, our research design entailed relatively few time periods with observations for a much larger number of cross-sectional units, “small $T$, large $N$,” that is, a panel-data structure rather than a time-series cross-section structure (Beck and Katz, 2004).

25. See Arellano and Bond (1991) for derivation of the second-order form and StataCorp (2003a, 2003b) for the $m$-order form.
Although various analysts often make much or little of the distinction between temporally and serially dominated data sets (with $T > N$ and $N > T$), for our purposes, the critical issue is whether $T$ is large enough to do the requisite averaging across time and whether it is large enough to address some econometric issues (Beck and Katz, 2004). The size of $T$ tells us a lot about which potential econometric problems might be serious for the data analyzed. Although no universally accepted cutoff level exists, “panel” studies almost invariably have single-digit $T$’s (3 being a common value), whereas “large $T$, small $N$” data sets (often referred to as time-series—cross-section (TSCS) data) commonly have $T$’s of 20 or more (Beck and Katz, 2004).

Although TSCS and panel data might share a common notation, they differ, and the properties of alternative estimators depend not so much on the relative size of $T$ and $N$ but on the absolute sizes of each (cf. Bruno, 2005). For example, since the seminal paper by Nickell (1981) in which it is shown that the LSDV estimator is not consistent for finite $T$ in models with a lagged dependent variable, several consistent instrumental variable and generalized method of moments estimators (GMM/IV) have been proposed in the econometric literature as an alternative to LSDV (see Anderson and Hsiao, 1982; Arellano and Bond, 1991). GMM/IV estimators can handle different numbers of instruments for each observation by using all available lags at each observation as internal instruments with different lag structures for strictly exogenous, predetermined, and endogenous variables.

More recently, Bruno (2005) argued that a weakness of GMM/IV estimators is that their properties hold when $N$ is large, so they can be severely biased and imprecise in panel data with a small number of cross-sectional units. But GMM/IV estimators can generate moment conditions prolifically, which can cause several problems in finite samples (Roodman, 2009). First, because the number of elements in the estimated variance matrix of the moments is quadratic in the instrument count, it is \textit{quartic} in $T$. In addition, a large instrument collection can overfit endogenous variables. In the extreme, consider that in 2SLS, if the number of instruments equals the number of observations, then the $R^2$’s of the first-stage regressions are unity, and the second-stage results match those of (biased) OLS. This bias is present in all instrumental variables regression and becomes more pronounced as the instrument count rises. Although little guidance is available from the literature on how many instruments are too many, it is clear from simulation studies that the number of instruments should not outnumber the cross-sectional units ($N$) and that small $T$ is preferable to large $T$ (Roodman, 2009; Ruud, 2000).
Clearly, no panacea is available in analyzing $N \times T$ structured data. The Arellano–Bond dynamic panel-data estimator has all the notable limitations that apply to fixed-effects models. A fixed-effects analysis cannot conclude anything about the interunit effects of the independent variables because such effects have been removed. Thus, the ability of this approach to remove the bias of relevant omitted variables relies on the assumption that those variables are fixed across time. And even when omitted variables are stable, discarding between-unit variation can yield standard errors that are considerably higher than those produced by methods that use both within- and between-unit variation (Allison, 2005). The methods we applied in this article embraced the perspective that data analysis must depend both on the relevant theory and on several informed choices that seem most defensible given practical and logistical considerations.