2005

Broken Windows: New Evidence from New York City and a Five-City Social Experiment

Bernard E. Harcourt

Jens Ludwig

CHICAGO

PUBLIC LAW AND LEGAL THEORY WORKING PAPER NO. 93

BROKEN WINDOWS: NEW EVIDENCE FROM NEW YORK CITY AND A FIVE-CITY SOCIAL EXPERIMENT

Bernard E. Harcourt and Jens Ludwig

THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO

June 2005

Forthcoming, *University of Chicago Law Review*, Volume 73

Broken Windows:
New Evidence from New York City
& a Five-City Social Experiment[1]

Bernard E. Harcourt
Professor of Law
University of Chicago Law School
harcourt@uchicago.edu

Jens Ludwig
Associate Professor of Public Policy
Georgetown University
&
Faculty Research Fellow
National Bureau of Economic Research
ludwigj@georgetown.edu

[1] Special thanks to Jeffrey Fagan, Tracey Meares, Steven Messner, Robert Morrissey, and Robert Sampson for comments on earlier drafts; to Ella Delaney and Tim Ross at the Vera Institute for their assistance in assembling the data for New York City; to Stephen Schacht at NORC for comments and guidance regarding the data collection and analysis; as well as to Sarah Rose and Zac Callen for excellent research assistance.

Broken Windows: New Evidence from New York City & a Five-City Social Experiment

Abstract

In 1982, James Q. Wilson and George Kelling suggested in an influential article in the *Atlantic Monthly* that targeting minor disorder could help reduce more serious crime. More than 20 years later, the three most populous cities in the U.S.—New York, Chicago and, most recently, Los Angeles—have all adopted at least some aspect of Wilson and Kelling's theory, primarily through more aggressive enforcement of minor misdemeanor laws. Remarkably little, though, is currently known about the effect of broken windows policing on crime.

According to a recent National Research Council report, existing research does not provide strong support for the broken windows hypothesis—with the possible exception of a 2001 study of crime trends in New York City by George Kelling and William Sousa.

In this paper, we re-examine the Kelling and Sousa 2001 study and independently analyze the crime data from New York City for the period 1989–98. In addition, we present results from an important social experiment known as Moving to Opportunity (MTO) underway in five cities, including New York, Chicago and Los Angeles as well as Baltimore and Boston, which provides what is arguably the first truly rigorous test of the broken windows hypothesis. Under this program, approximately 4,800 low-income families living in high-crime public housing communities characterized by high rates of social disorder were randomly assigned housing vouchers to move to less disadvantaged and disorderly communities. The MTO program thus provides the ideal test of the broken windows theory.

Taken together, the evidence from New York City and from the five-city social experiment provides no support for a simple first-order disorder-crime relationship as hypothesized by Wilson and Kelling, nor that broken windows policing is the optimal use of scarce law enforcement resources.

Introduction

In 1982, James Q. Wilson and George L. Kelling suggested in an influential article in the *Atlantic Monthly* that targeting minor disorder—loitering, panhandling, prostitution, graffiti—could help reduce more serious crime.[2] The "broken windows" theory produced what many observers have called a revolution in policing and law enforcement.[3] Today, the three most populous cities in the U.S.—New York, Chicago and, most recently, Los Angeles—have all adopted at least some aspect of Wilson and Kelling's broken windows theory, primarily through more aggressive enforcement of minor misdemeanor laws, also known as "zero tolerance" policing.[4]

Despite the widespread policy influence of the 1982 *Atlantic Monthly* essay, remarkably little is known about the effects of broken windows. A number of leading researchers in sociology, law, and police studies—including Wesley Skogan at Northwestern, Robert Sampson at Harvard, Stephen Raudenbush at the University of Michigan, Anthony Braga at Harvard, and Jeffrey Fagan at Columbia, among others—have compiled datasets from different urban areas to explore the broken windows hypothesis, but the evidence remains, at best, mixed. In 2000, John Eck and Edward Maguire reviewed the empirical evidence and studies on broken-windows policing in their contribution to Alfred Blumstein's *The Crime Drop in America* (2000), and found

[2] James Q. Wilson & George Kelling, *Broken Windows: The Police and Neighborhood Safety*, Atlantic Monthly, Mar. 1982, at 29.

[3] *See generally* Bernard E. Harcourt, *Illusion of Order: The False Promise of Broke Windows Policing,* 2-4, 46-54 (2001).

[4] New York City Mayor Rudolph Giuliani first embraced quality-of-life policing in the mid-1990s, at a time when high crime rates began declining impressively in the City. Mayor Giuliani and his first police commissioner, William Bratton, traced their quality-of-life initiative directly back to the Wilson and Kelling essay. *See* Rudolph W. Giuliani and William J. Bratton, *Police Strategy No. 5: Reclaiming the Public Spaces of New York,* at 6 (New York: City of New York Police Department). The City of Chicago implemented an anti-gang loitering ordinance in the early 1990s that it vigorously enforced during the period 1993-1995 resulting in misdemeanor arrests of over 42,000 individuals (*City of Chicago v. Morales* 1999:49). In October 2002, Los Angeles Mayor James Hahn appointed William Bratton police commission on a platform that promised a broken-windows approach. According to news reports, "Mr. Bratton said his first priority after being sworn in on Oct.2 28 [2002] would be ending the smile-and-wave approach to crime fighting. He said he wanted policing based on the so-called broken-windows theory." *See* Charlie LeDuff, *Los Angeles Police Chief Faces a Huge Challenge*, NY Times (Oct 24, 2002); *see also* Tina Daunt and Megan Garvey, *Bratton Lays Out Ambitious Set of Goals for LAPD*, LA Times (Oct 4, 2002); Megan Garvey, *Bratton Is Planning a Clean Start; The police chief, who will be sworn in today, sees fighting graffiti as key to reducing crime*, LA Times, Metro Desk, p. 1 (Oct. 25, 2002).

that there is little evidence to support the claim that broken-windows policing contributed to the sharp decrease in crime during the 1990s.[5]

However, a recent report by a blue-ribbon panel commissioned by the National Research Council (NRC)—which is itself part of the National Academies of Science, chartered in 1863 by Congress to advise the federal government on scientific matters—suggests that there may be new evidence in support of the broken windows theory.[6] The NRC notes that "there is a widespread perception among police policy makers and the public that enforcement strategies (primarily arrest) applied broadly against offenders committing minor offenses lead to reductions in serious crime. Research does not provide strong support for this proposition … A recent study of New York [City] precincts, however, indicates a strong relationship between the rate of arrests for minor crimes and crime rates in precincts in New York (Kelling and Sousa, 2001).[7] Using a multilevel research design, the authors provide one of the first indications of a direct link between a generalized program of intensive enforcement and declines in more serious crime. While the study uses an innovative modeling approach to estimate this effect, limitations in the data available raise questions regarding the validity of the results…"[8]

The study by George Kelling and William Sousa, titled *Do Police Matter? An Analysis of the Impact of New York City's Police Reforms* and published by the Manhattan Institute in December 2001, shows that aggressive misdemeanor arrest policies in New York City account for the significant drop in crime during the mid- to late-1990s.[9] The 2001 Kelling and Sousa report has received significant media attention.

---

[5] John E. Eck & Edward R. Maguire, *Have Changes in Policing Reduced Violent Crime? An Assessment of the Evidence*, *in The Crime Drop in America* 228 (Alfred Blumstein & Joel Wallman eds., 2000); *see also,* Bernard E. Harcourt, *Reflecting on the Subject: A Critique of the Social Influence Conception of Deterrence, the Broken Windows Theory, and Order-Maintenance Policing New York Style*, 97 Mich. L. Rev. 291 (1998); Harcourt, *Illusion of Order*, supra note __.

[6] Wesley Skogan and Kathleen Frydl, editors, *Fairness and Effectiveness in Policing: The Evidence*. Washington, DC: National Academies Press 2004.

[7] The report is referring here to George L. Kelling & William H. Sousa, Jr., *Do Police Matter? An Analysis of the Impact of New York City's Police Reforms*, Manhattan Institute Center for Civic Innovation Civic Report No. 22 (2001).

[8] Skogan and Frydl, *Fairness and Effectiveness in Policing*, *supra* note __, at 229-30.

[9] The Kelling and Sousa report was issued with a simulcast editorial comment by the authors in the *New York Post.* "So what does all this mean?" Kelling and Sousa ask. "First, it means that New Yorkers should stop listening to critics who contend that police tactics matter little, if at all, in determining crime rates." These critics, the authors note, "have been parroting what is virtual dogma in criminal-justice circles, that crime is caused by 'root causes' such as racism, poverty and social injustice." In contrast, the authors declare, "This study places the 'root cause' theory of crime in serious jeopardy" George L. Kelling

In addition to being viewed as the only promising evidence by the NRC, the *Economist* reported on the study,[10] as did the *New York Times,*[11] the *Wall Street Journal*,[12] and the *Boston Globe*,[13] both of the latter in editorials, and the *Atlanta Constitution.*[14] For example, the *Wall Street Journal's* editorial page argued: "A brand new report from the indispensable Manhattan Institute chronicles these law-and-order achievements and explains what made them possible. . . . 'Do Police Matter?' also does a great public service in thoroughly refuting those media critics and political opponents of the Republican Mayor who've insisted for the past eight years that the NYPD had little if anything to do with the fall in crime. In this alternative universe, the city's drop in crime should be credited to low unemployment from a booming economy. Or the decline in crack cocaine use that had plagued the 1980s. Or the demographic reality that the proportion of young males—the most common offenders—to the general population had dropped. In fact, none of these alternative explanations stands up to scrutiny."[15]

An even more recent working paper distributed by the National Bureau of Economic Research, by economists Hope Corman and Naci Mocan, applies a slightly different empirical approach to data from New York City and claims to support the Kelling-Sousa conclusion. [16] Corman and Mocan analyze monthly time-series data for New York City as a whole and claim that the dramatic increase in misdemeanor arrest rates in New York during the 1990s is responsible for a large share of the city's drop in

---

& William H. Sousa, Jr., Editorial, *Tough Cops Matter*, N.Y. Post, Dec. 19, 2001, at 41. The *New York Post* carried its own editorial the same day, *It's the Cops, Stupid*, N.Y. Post, Dec. 19, 2001, at 42.

[10] *As New York's Inexperienced New Mayor takes Office, What Lessons Should He Draw From His . . .*, The Economist, Jan. 5, 2002

[11] Kevin Flynn, *Study Says a Slumping Economy Doesn't Mean Crime Will Rise*, N.Y. Times, Dec. 19, 2001, at 8.

[12] *New York's Finest,* Wall St. J., Dec. 27, 2001.

[13] *Boston Globe* editorial (2001) "Behind Giuliani's Jab." 29 December 2001, at A14.

[14] Colin Campbell, *New York a Blueprint for Cutting Atlanta Crime*, The Atlanta Constitution, Dec. 23, 2001, at 5F.

[15] *New York's Finest,* Wall St. J., Dec. 27, 2001. Even the *Courier-Mail*, the Queensland newspaper, reports on the 2001 study, reporting that "in precinct after precinct Kelling and Sousa found a similar pattern—as 'broken windows' policing was increased, violent crime declined." Ron Bruton, *Broken Windows' Plan Shatters Crime Theory*, Courier-Mail (Queensland), Jan. 5, 2002, at 24. Kelling and Sousa have also placed editorials in *The Australian*, *Turn Up the Heat and Beat Serious* Crime, The Australian, Oct. 3, 2002, *The Cincinnati Post*, *Ways of Policing Matter*, The Cinncinati Post, Jan. 7, 2002, at 8A, and *The Harrisburg Patriot*, *Broken Windows': Paying Attention to Neighborhoods Can Reduce Crime*, Harrisburg Patriot, Jan. 3, 2002, at A13.

[16] Hope Corman and Naci Mocan, *Carrots, Sticks and Broken Windows*, NBER Working Paper 9061 (July 2002).

crime over this period. So while Kelling and Sousa use variation across precincts over time in misdemeanor arrests and crime rates to identify the effects of the former on the latter, Corman and Mocan use city-wide variation over time to generate a similar finding. Moreover Corman and Mocan point to deterrence as the most plausible mechanism for this relationship, given that misdemeanor arrests typically result in either no jail time or short spells of incarceration.[17] The Kelling and Sousa study, together with the Corman and Mocan paper, are thus important contributions, representing the best existing evidence supporting the broken-windows hypothesis and the related (and widespread) broken-windows or zero-tolerance policing strategy.

In this article, we set out to re-analyze and assess the best available evidence from New York City about the effects of broken windows policing. We demonstrate that the pattern of crime changes across New York precincts during the 1990s that Kelling and Sousa attribute to broken windows policing is more consistent with what statisticians call mean reversion: Those precincts that received the most intensive broken windows policing during the 1990s are the ones that experienced *the largest increases in crime during the city's crack epidemic of the mid- to late-1980s*. Consistent with findings elsewhere from city-level data,[18] jurisdictions with the greatest increases in crime during this period tend to experience the largest subsequent declines as well. We call this Newton's Law of Crime: What goes up, must come down (and what goes up the most, tends to come down the most). For similar reasons we argue that the Corman and Mocan study is also unable to convincingly determine that broken windows policing is a causal contributor to crime rates in New York City.

Because our re-analysis of the New York data leaves us with a Scotch verdict—"not proven"—we then turn to data from a unique randomized experiment operated by the U.S. Department of Housing and Urban Development known as Moving to Opportunity (MTO), which provides a unique opportunity to test the original Wilson and

[17] As Corman and Mocan (2002, p. 14) note, only about 9% of misdemeanor arrests result in imprisonment, with an average sentence length of 27.5 days. So the expected prison time for a misdemeanor arrest is about 2.6 days.

[18] *See generally* Steven Raphael and Jens Ludwig, "Do Prison Sentence Enhancements Reduce Gun Crime? The Case of Project Exile," 251-286, in *Evaluating Gun Policy,* Jens Ludwig and Philip J. Cook, ed. Washington, DC: Brookings (2003).

Kelling broken windows thesis. MTO has been in operation since 1994 in five cities, including the three largest cities in the country that have adopted aspects of broken windows policing (New York, Chicago and L.A.) as well as Baltimore and Boston. Under MTO a total of around 4,800 low-income families living in public housing communities characterized by high rates of crime and social disorder were randomly assigned housing vouchers to move to less disadvantaged and disorderly communities. The random assignment of families to neighborhoods in MTO helps overcome the problem of determining the causal effects of neighborhood disorder on individual criminal behavior that plagues most previous studies in this literature.[19]

The implications of MTO for the ongoing debates about the broken windows theory have never yet been explored.[20] Yet the results from MTO suggest that moving people to communities with less social or physical disorder—the key intervening factor in the original Wilson and Kelling broken windows hypothesis—on balance *does not lead to reductions in their criminal behavior*. It is important to note that MTO changed multiple aspects of people's neighborhoods: MTO families moved to neighborhoods that were less disorderly, but also had fewer low-income families and more high-status households. MTO thus tests the combined effects of less disorder and increased affluence within a community, which is arguably the policy-relevant "treatment combination" for neighborhoods under the broken windows model because reductions in disorder, like other improvements in neighborhood amenities, should on average translate into increased neighborhood gentrification.

Taken together our examination of data from New York City and MTO provide no support for the idea that "broken windows" activities, including zero-tolerance policing or other measures designed to reduce the level of social or physical disorder within a community, represent the optimal use of scarce government resources.

---

[19] Because most people have at least some degree of choice over where they live and with whom they associate, previous *non*-experimental studies may confound the effects of neighborhood disorder and other characteristics on people's behavior with the effects of difficult-to-measure individual attributes that influence both their involvement with crime and their choice of residential neighborhood.

[20] While recent results of neighborhood effects on criminal behavior have been published in economics, *see* Jeffrey R. Kling, et al., *Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment*, Quarterly Journal of Economics 120 (2005), these findings are currently not widely known outside of that field and, as a result, their implications for broken windows has never been explored.

The paper is organized as follows. Part I of the paper locates the broken windows theory within the sociological and policy traditions and reviews preceding efforts to test the broken windows theory and the practice of broken-windows policing. Part II of the paper then presents our discussion of the evidence from New York City. Part III then presents our findings from the MTO experiment demonstrating that randomly assigning people to move to less disorderly communities does not yield the simple "less disorder, less criminal activity" result that broken windows policing predicts.

PART I. Locating the *Broken Windows* Theory

A. The Socio-Legal Theoretical Context

There is a long tradition within socio-legal research of studying visual cues of neighborhood disorder and exploring the relationship between those neighborhood characteristics and deviance. Prompted by a recurring observation of dramatic variations in crime rates across neighborhoods, the tradition grew over decades of research taking seriously the idea that there may be "neighborhood effects" on the production of crime: That is, arrangements in social space may significantly affect human behavior. This research tradition traces importantly to the early Chicago School of sociology—especially the monographs on neighborhoods and spatial settings, the Jewish ghetto,[21] the Italian "slum,"[22] the Near North side of Chicago,[23] taxi-dance halls,[24] and brothels[25]—and to the later social interactionist research of Irving Goffman, especially his study *Behavior in Public Places: Notes on the Social Organization of Gatherings,*[26] and others such as Albert Cohen[27] and Jane Jacobs.[28]

One of the most striking findings from the neighborhood effects research comes from the dramatic differences across neighborhoods in rates of crime and delinquency—

[21] L. Wirth, *The Ghetto* (1928).
[22] William F. Whyte, *Street Corner Society: The Social Structure of an Italian Slum* (1943).
[23] H. W. Zorbaugh, *The Gold Coast and the Slum* (1929).
[24] P. G. Cressey, *The Taxi-Dance Hall* (1932).
[25] W. Reckless, *Vice in Chicago* (1933).
[26] Ervin Goffman, *Behavior in Public Places: Notes on the Social Organization of Gatherings* (1963).
[27] Albert K. Cohen, *Delinquent Boys: The Culture of the Gang* (1955).
[28] Jane Jacobs, *The Death and Life of Great American Cities* (1961). As Andrew Abbott notes, "Chicago felt that no social fact makes any sense abstracted from its context in social (and often geographic) space and social time. Social facts are *located*." Andrew Abbott, *Of Time and Space: The Contemporary Relevance of the Chicago School.* 75 Social Forces 1149, 1152 (1997).

even across neighborhoods with similar concentrations of social disadvantage as measured by average rates of poverty, unemployment, familial and residential instability, and dependence on government benefit programs.[29] Robert Sampson and Stephen Raudenbush trace the rich intellectual history and the variations over time in neighborhood-effects research in their thorough paper, *Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods* (1999).[30]

A consideration of the research in this area suggests two lasting puzzles. The first focuses on locating sources of variation in crime across neighborhoods and identifies two leading candidates. First, differences in crime rates across areas could be due to unobservable individual characteristics related to the residents of the neighborhood, and thus the possibility of self-selection on the part of the individuals. Put differently, some neighborhoods may have more crime because they are home to a larger share of crime-prone people, although all of the individual attributes that predispose some people to engage in criminal activity are difficult to measure in social science datasets. A second explanation is that variation across areas in crime rates may be due to differences in social processes and conditions across neighborhoods, including disorderliness or informal mechanisms of social control. The notion of social disorganization pioneered by Clifford Shaw and Henry McKay[31] represented one effort to locate the answer to this first puzzle, at least in part, in mechanisms of informal social control and collective action—in identifying an agency of social control that could be disrupted by residential mobility and economic conditions. Sampson, Raudenbush, and Fenton Earls' Project on Human Development in Chicago Neighborhoods (PHDCN) research represents another answer focused on informal social processes, more specifically on the notion of "collective efficacy," which they define as "the linkage of cohesion and mutual trust with shared expectations for intervening in support of neighborhood social control."[32]

A second puzzle focuses on the issue of remedies. Even if the neighborhood-effects research suggests a causal relationship between, on the one hand, identifiable

[29] *See, e.g.,* Glaeser, Sacerdote and Scheinkman, 1996, Sampson, Raudenbush and Earls, 1997.

[30] *See* Robert Sampson and Stephen Raudenbush, *Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods*, American Jounral of Sociology 105(3):603-651 (1999).

[31] Clifford Shaw and Henry McKay, *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press (1942).

[32] Sampson and Raudenbush, *Systematic Social Observation*, *supra* note __ at 612-613.

social processes or neighborhood characteristics and, on the other hand, crime, does the causal explanation offer insight into what can be done to change things in a public policy sense? In this regard, the sociological theories have been relatively quiet, reflecting a general hesitation to move from the positive to the prescriptive.

It is within this rich research field that the "broken windows" hypothesis emerged in the early 1980s. Though first articulated and tested by Philip Zimbardo, a Stanford psychologist, in the late 1960s, the broken windows theory was most clearly articulated and popularized in James Q. Wilson and George L. Kelling's article titled *Broken Windows: The Police and Neighborhood Safety*, which appeared in the *Atlantic Monthly* in 1982.[33] "Disorder and crime are inextricably linked, in a kind of developmental sequence," Wilson and Kelling argued, so that efforts to reduce disorder might ultimately translate into reductions in criminal activity as well.[34] Minor social disorder—littering, loitering, public drinking, panhandling, and prostitution—as well as physical disorder—graffiti, abandoned buildings, and littered sidewalks—if tolerated in a neighborhood, produce an environment that is likely to attract crime. These forms of disorder signal to potential criminals that delinquent behavior will not be reported or controlled—that no one is in charge. To law-abiding citizens, these disorderly conditions signal the need to avoid the streets or even flee the neighborhood. One broken window, left unrepaired, invites other broken windows. These progressively break down community standards and leave the community vulnerable to crime. In this way, disorder breeds crime: "Such an area is vulnerable to criminal invasion. Though it is not inevitable, it is more likely that here," Wilson and Kelling wrote, "drugs will change hands, prostitutes will solicit, and cars will be stripped. That the drunks will be robbed by boys who do it as a lark, and the prostitutes' customers will be robbed by men who do it purposefully and perhaps violently."[35]

The broken windows theory thus addresses the first puzzle of the neighborhood-effects literature in a straightforward and provocative way: it is the variations in disorder in neighborhoods that explains the variation in crime, holding structural disadvantage constant. The real trigger is disorderliness itself. The theory was familiar to sociologists

---

[33] Wilson & Kelling, *supra* note __.
[34] *Id.* at 31
[35] *Id.* at 31-32

because of its proximity to theories of urban decay and social contagion. Urban sociologists interpreted the broken-windows hypothesis through the lens of urban decline: disorderliness, dilapidation, abandonment, and social disorder, such as prostitution, public intoxication and drug use, reflected and reinforced, in a cyclical manner, declining property values, residential instability, and the gradual decay of the urban neighborhood.[36] Philip Cook and Kristin Goss offer a closely-related interpretation focusing on a standard model of "social contagion."[37] From the contagion perspective, the broken-windows phenomenon reflects an information cascade: people with imperfect information about the risks and rewards of criminal activity may infer the net returns to crime from the social environment.[38] Information limitations are at the heart of the information cascade model. Here, the potential criminals do not know the probability of being detected in a neighborhood, but the lack of enforcement of minor crime and disorder fills this void and signals low enforcement. The characteristics of the local physical environment, which are themselves the product of the accumulated series of behaviors of local residents, thus communicate the statistical likelihood of being apprehended. They are a signaling mechanism that feeds into the calculus of whether to commit crime. This "contagion" interpretation offers a straightforward explanation of broken windows familiar to most sociologists and economists.[39]

As to the second puzzle—concerning the public policy prescriptions—the *Broken Windows* essay itself did not compel a particular policy outcome. From a policy perspective, the broken windows hypothesis is in principle consistent with a variety of potential policy levers, ranging from changes in policing to community organizing.

---

[36] *See* Wesley Skogan, *Disorder and Decline: Crime and the Spiral of Decay in American Cities* (1990); Gerald E. Frug, *City Making: Building Communities Without Building Walls* (1999).

[37] Philip J. Cook & Kristen A. Goss, *A Selective Review of the Social-Contagion Literature* (Sanford Institute of Public Policy Studies, Duke University, Working Paper, 1996).

[38] *Id.*

[39] For a discussion of the etiology of less-serious and more-serious crimes, see Michael Gottfredson and Travis Hirschi, *A General Theory of Crime,* Stanford, Calif.: Stanford University Press 1990. For a discussion of how "routine activities" across neighborhoods may affect criminal opportunities and outcomes, see Lawrence Cohen and Marcus Felson, "Social Change and Crime Rate Trends: A Routine Activity Approach," *American Sociology Review* 44:588-608 (1979), and Lawrence Cohen, James Kluegel, and Kenneth Land, "Social Inequality and Predatory Criminal Victimization: An Exposition and Test of a Formal Theory," *American Sociological Review* 46:505-24 (1981). Additional discussion of the "social disorganization" model of disorder and neighborhood effects on crime is provided by Robert J. Bursik, Jr., "Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects," *Criminology* 26:519-52 (1988).

Nevertheless, most policymakers seem to have understood the theory as implying what has come to be known as "broken-windows policing"—also known as "order-maintenance," "zero-tolerance," or "quality-of-life" policing. So for instance, in their 2001 study, George Kelling, the co-author of the original *Broken Windows* essay, and William Sousa suggest that the most effective way to address disorder and reduce crime is to increase the number of misdemeanor arrests.[40]

B. Testing the Broken Windows Hypothesis

To date, empirical testing of the broken windows theory has taken one of two forms. A first approach attempts to measure neighborhood disorder and crime, as well as other correlates of criminality, such as poverty and residential instability, in order to determine whether there are statistically interesting correlations between these variables. A second approach has focused on measures of broken-windows policing—for instance, rates of misdemeanor arrests—and conducts relatively similar statistical analyses on these variables in order, again, to identify significant correlations. We begin by reviewing the first approach.

1. *Disorder and Crime*

Early on, many proponents of the broken-windows hypothesis pointed to the research of Wesley Skogan, especially his monograph *Disorder and Decline: Crime and the Spiral of Decay in American Neighborhoods* (1990), and argued that it empirically verified the broken-windows theory.[41] Skogan's book, *Disorder and Decline*, addressed the larger question of the impact of neighborhood disorder on urban decline, but in a section of the book, Skogan discussed the broken windows hypothesis, ran a regression of neighborhood disorder on robbery victimization, and concluded that "'Broken windows' do need to be repaired quickly."[42] Many observers interpreted this as an endorsement of the broken-windows theory and accepted Skogan's view of the evidence. George Kelling, co-author of *Broken Windows*[43] and of a book entitled *Fixing Broken*

---

[40] Kelling & Sousa, *supra* note __.
[41] Skogan, *supra* note __.
[42] *Id.* at 75.
[43] Wilson & Kelling, *supra* note __.

Windows,[44] contended that Wesley Skogan "established the causal links between disorder and serious crime—empirically verifying the 'Broken Windows' hypotheses."[45] Dan Kahan at Yale similarly argued that "[t]he work of criminologist Wesley Skogan supplies empirical support for the 'broken windows' hypothesis."[46] Subsequent work by one of the co-authors of this article, however, has cast some doubt about what conclusions can properly be drawn from Skogan's analysis.[47]

A few years later, Ralph Taylor of Temple University conducted research in sixty-six neighborhoods in Baltimore using longitudinal data, and attempted to determine the relationship between neighborhood crime and what he terms social and physical "incivilities"—panhandlers, public drunks, trash graffiti and vacant lots, among other things. What he found was that, while certain types of incivilities were associated with crime or urban decay, others were not. In his book, *Breaking Away from Broken Windows*, Taylor concludes from his data that different types of incivilities may require different policy responses. "Researchers and policy-makers alike," Taylor writes, "need to break away from broken windows per se and widen the models upon which they rely, both to predict and to preserve safe and stable neighborhoods with assured and committed residents."[48]

One of the most comprehensive and thorough studies of the broken windows theory to date is Robert Sampson and Stephen Raudenbush's 1999 study. Their study grows out of the PHDCN and is based on systematic social observation: using trained observers who drove a sports utility vehicle at five miles per hour down every street in 196 Chicago census tracts, and randomly selecting 15,141 street sides, they were able to collect precise data on neighborhood disorder. With regard to the disorder-crime nexus, Sampson and Raudenbush found that disorder and predatory crime are only moderately correlated, but that, when antecedent neighborhood characteristics are taken into account,

---

[44] George Kelling & Catherine Coles, *Fixing Broken Windows: Restoring Order and Reducing Crime in Our Communities* (1996).

[45] *Id.* at 24.

[46] Dan Kahan, *Social Influence, Social Meaning, and Deterrence*, 83 Virginia Law Review 349, 369 (1997); *see also* Dan Kahan, *Between Economics and Sociology, 95* Michigan Law Review 2477, 2488 n.62 (1997).

[47] Harcourt, *Illusion of Order*, *supra* note __, at 59–78.

[48] Ralph B. Taylor, *Breaking Away From Broken Windows: Baltimore Neighborhoods and the Nationwide Fight Against Crime, Guns, Fear, and Decline, at 22* (2001).

the connection between disorder and crime "vanished in 4 out of 5 tests—including homicide, arguably our best measure of violence."[49] They nevertheless suggest that disorder may have indirect, neighborhood effects on crime by influencing "migration patterns, investment by businesses, and overall neighborhood viability."[50]

On the basis of their extensive research, Sampson and Raudenbush conclude that "[a]ttacking public order through tough police tactics may thus be a politically popular but perhaps analytically weak strategy to reduce crime."[51] As an alternative to the broken-windows theory, Sampson and Raudenbush suggest that disorder is of the same etiology as crime—being, so often, forms of minor crime—and that both crime and disorder have the same antecedent conditions. "Rather than conceive of disorder as a direct cause of crime, we view many elements of disorder as part and parcel of crime itself."[52] Thus, "a reasonable hypothesis is that public disorder and predatory crimes are manifestations of the same explanatory process, albeit at different ends of a 'seriousness' continuum."[53]

2. *Studies of Aggressive Misdemeanor Arrest Policing*

Another strand of research, focusing on studies of aggressive arrest policies, was also brought to bear on the broken-windows hypothesis. Here too, James Q. Wilson sparked the debate, primarily with his 1968 book on the *Varieties of Police Behavior*, and his research with Barbara Boland on the effects of police arrests on crime.[54] Wilson and Boland hypothesized that aggressive police patrols, involving increased stops and arrests, have a deterrent effect on crime.

---

[49] Robert J. Sampson & Stephen W. Raudenbush, *Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods,* 105 American Journal of Sociology 603, 637 (1999).

[50] *Id.*

[51] *Id.* at 638

[52] *Id.* at 608

[53] *Id.* Sampson and Raudenbush have a more recent study showing that neighborhood racial composition affects people's perceptions of neighborhood disorder (Sampson and Raudenbush 2004). They conclude as a result that order maintenance may not be helpful because it affects actual but not perceived disorder (2004:337). For a study of disorder and youth crime in Canada, see John Hagan and Bill McCarthy, *Mean Streets: Youth Crime and Homelessness*. New York: Cambridge University Press (1997).

[54] James Q. Wilson and Barbara Boland, "The Effect of the Police on Crime," *Law & Society Review* 12: 367–390 (1978); James Q. Wilson and Barbara Boland, "The Effects of the Police on Crime: A Response to Jacob and Rich," *Law & Society Review* 16: 163–169 (1981).

A number of contributions ensued, both supporting and criticizing these findings, but, as Robert Sampson and Jacqueline Cohen suggested back in 1988, the results were "mixed."[55] There have been strong contributions to the literature, such as the 1999 study led by Anthony Braga, titled "Problem-Oriented Policing in Violent Crime Places: A Randomized Controlled Experiment," published in *Criminology*.[56] But still, most of this research is unable to distinguish between the broken windows hypothesis and more traditional explanations of incapacitation and deterrence associated with increased police arrests, presence, contact and surveillance. The problem is somewhat endemic to the design of these studies. As Sampson and Cohen conclude with regard to their own work, "[i]t is true that our analysis was not able to choose definitely between the two alternative scenarios."[57]

In this vein, Jeffrey Fagan and Garth Davies test, in their research titled *Policing Guns: Order Maintenance and Crime Control in New York*, whether quality-of-life policing in New York City contributed to the reduction in lethal violence in the late 1990s. They analyze precinct crime rates from 1999 and try to determine whether these crime rates can be predicted by the amount of stop-and-frisk activity that occurred in the precinct in the preceding year. Based on their research, Fagan and Davies find that "[f]or both violence arrests broadly and homicide arrests specifically, there is no single category of citizen stops by police that predicts where crime will increase or decrease in the following year."[58] When they examine homicide fatalities, they observe different effects by type of stop and by victim race. "Stops for violence are significant predictors of reductions in both gun homicide deaths and overall homicide deaths, but only among Hispanics."[59] In contrast, for African-Americans, no type of arrests predicts homicide victimization a year later; and for whites, the results are not reliable because of the low white homicide victimization rate.

---

[55] Robert J. Sampson and Jacqueline Cohen, "Deterrent Effect of the Police on Crime: A Replication and Theoretical Extension," 22 *Law and Society Review* 163, 166 (1988).

[56] Anthony A. Braga et al., "Problem-Oriented Policing in Violent Crime Places: A Randomized Controlled Experiment," *Criminology* 37 (1999): 541–580; *see also* Anthony A. Braga, *Problem-Oriented Policing and Crime Prevention*. Monsey, NY: Criminal Justice Press 2002.

[57] Sampson and Cohen, *supra* note __, at 185

[58] Jeffrey Fagan & Garth Davies, *Policing Guns: Order Maintenance and Crime Control in New York*, *in* Guns, Crime, and Punishment in America (Bernard E. Harcourt ed., 2003).

[59] *Id.*

Why is it that there may be effects for Hispanics, but not for African-Americans? Fagan and Davies suggest that it may have to do with what they call "stigma saturation" in black communities: when stigma is applied in ways that are perceived as too harsh and unfair, it may have reverse effects. They write, "When legal control engenders resistance, opposition or defiance, the opportunity to leverage formal social control into informal social control is lost. The absence of crime control returns from OMP policing may reflect just such a dynamic among African Americans, who shouldered much of the burden of OMP."[60]

The final and most recent contribution to this literature is Steve Levitt's 2004 *Journal of Economic Perspectives* review essay, in which Levitt argues that policing practices probably do not explain much of the crime drop in the 1990s because crime went down everywhere, even in places where police departments did not implement new policing strategies. Instead, Levitt attributes the massive period effects on crime throughout the U.S. during the 1990s to some combination of increased imprisonment, increases in the number of police, the ebbing of the crack epidemic that started in many big cities in the mid-1980s, and the legalization of abortion in the U.S. during the early 1970s.

PART II. New York City's Experience

In this section we discuss the most recent studies on broken-windows policing in New York City, both the Kelling and Sousa (2001) study and the evidence presented by Corman and Mocan (2002). We argue that the Kelling and Sousa (2001) analysis has limitations that ultimately render it uninformative about the causal effects of broken windows policing practices. We also show that the Corman and Mocan (2002) analysis cannot support the claim that broken windows policing activities are responsible for declines in crime.

A. The Kelling and Sousa (2001) Study

The study by George Kelling and William Sousa (hereafter KS) fits in the larger tradition of studies of aggressive arrest policies discussed earlier. The goal of their study

[60] *Id.*

is the "systematic attempt to statistically parse out the relative contributions of police actions, the economy, demographics, and changing drug use patterns on crime" in New York City. The major problem with previous studies, they argue, is that those studies lack an adequate comparison group for New York City: previous research has either used an unsuitable comparison, such as other cities, or failed to use any comparison at all. The key insight in this study, Kelling and Sousa suggest, is to simulate comparison groups by treating the city as 75 separate and comparable entities. "Rather than one city," they explain, "we view New York as 75 separate entities, corresponding to the 75 police precincts."[61]

The research design, then, is to statistically compare the relationship between violent crime and four dependent variables—broken-windows policing, economic indicators, young male population shifts, and the decline in crack cocaine consumption—in the 75 precincts of New York City. They find a strong negative relationship between precinct-level misdemeanor arrests and violent crime. In what follows we re-examine these NYC results using a wide variety of alternative statistical approaches. Our efforts to obtain, replicate and extend their data are discussed in detail in Appendix A.

Replicating the KS results is complicated in part by the fact that in neither the KS Manhattan Institute report nor Sousa's dissertation do the authors spell out the exact estimating equations for their analysis. Nor does their Table 4, which presents their key results, show the number of observations used to generate their estimates (to give some sense for how the analysis is structured). Nevertheless, from reading over the discussion in KS and in Sousa's dissertation it would appear that they are estimating a two-level hierarchical linear growth model, of the sort discussed in Chapter 6 of Raudenbush and Bryk (2002).[62] If we let level 1 in this model represent time (subscripted by t) and level 2 represent precincts (subscripted by i), we believe that the two-level linear growth model that they are estimating is given by the following equations:

---

[61] Kelling & Sousa, *supra* note __ at 1, 4. In 1994, a precinct was divided in two, resulting in 76 precincts existing today. To maintain consistency over the studied period, the authors use the original 75 precincts.

[62] Stephen W. Raudenbush and Anthony S. Bryk, *Hierarchical Linear Models: Applications and Data Analysis Methods*, Second Edition (2002).

(1) $VC_{ti} = \pi_{0i} + \pi_{1i} A_t + \varepsilon_{ti}$

(2) $\pi_{0i} = \alpha_{00} + \alpha_{01} MA_i + \alpha_{02} X_i + r_{0i}$

(3) $\pi_{1i} = \alpha_{10} + \alpha_{11} MA_i + \alpha_{12} X_i + r_{1i}$

where

VC_{ti} = violent crimes in precint (i) in year (t)

A_t = time (1989, 1990, …, 1998)

MA_i = precinct (i)'s average misdemeanor arrests over the sample period

X_i = average value of other covariates for precinct (i) over sample period

The empirical setup that is being estimated by KS is easier to see by substituting equations (2) and (3) into (1) to get the reduced-form estimating equation (4):

(4) $VC_{ti} = \beta_1 + \beta_2 MA_i + \beta_3 A_t + \beta_4 MA_i{*}A_t + \varepsilon_{ti}$

We can replicate the key coefficient in their analysis (β_4 or, equivalently, α_{11}) as shown in the first row of Table 1, where we estimate equation (4) measuring all of our variables in precinct counts (rather than per capita rates) and do not weight by precinct population.[63] Note that as shown in Table 1, these estimates are not very sensitive to decisions about whether to weight by precinct population or not, or to work in per capita crime and arrest rates rather than counts. Note also that the coefficient and standard error for the effects of misdemeanor arrest rates on the time slope in violent crimes—which is the key estimate of interest—is identical to what is reported in KS, their Table 4, although our point estimates for the intercept terms have a slightly different scaling.

Kelling and Sousa conclude from these results that broken windows is a highly effective crime-fighting strategy. The bottom line: "The average NYPD precinct during the ten-year period studied could expect to suffer one less violent crime for approximately every 28 additional misdemeanor arrests made." This, Kelling and Sousa

[63] Note that we can also reproduce the point estimate and standard error for the time slope in their unconditional model (-131, se=10 or 11, Manhattan Institute Table 3), which is just a regression of violent crimes against time (this regression has N=750, not weighting by precinct population, works in precinct crime counts not rates). All of these results from f:\research\broken_windows2\stata\jens_regs_jan2105.do

suggest, offers "the most-definitive possible answer to the question of whether police mattered in New York City during its intense crime-drop."[64]

Our conclusion from these results is somewhat different, and points in the direction of mean reversion. Any study of the influences on American crime patterns during the past 20 years is complicated by the massive period effects that have generated dramatic year-to-year changes in crime across the country. The increase in crime rates was particularly dramatic from the mid-1980s through the early- to mid-1990s, which is thought to have been driven largely by the growth in crack cocaine use and involvement of firearms in the new street markets for crack.[65] Using city-level data, Steven Raphael and Jens Ludwig show that those cities that experienced the largest increases in crime during this period subsequently also experienced the largest crime drops. [66] A natural concern is to worry that the same process may be at work at the neighborhood or police precinct level as well.

Figure 1 suggests that crime patterns across New York precincts that KS attribute to the effects of broken windows policing can be explained by mean reversion: Broken windows policing (as measured by misdemeanor arrests) was conducted most intensively in New York within the city's most violent neighborhoods, which are the areas that experienced the largest increases in violent crime during the 1980s and the largest declines in violent crime during the 1990s. Panel A shows that at the start of the KS panel (1989) precincts with higher violent crime rates also have higher rates of misdemeanor arrests. That is, the regression line relating violent crime and misdemeanor arrests in 1989 has a positive slope, consistent with Kelling and Sousa's own findings (top panel of their Table 4). Panel B shows that the most violent precincts in 1989 also experienced the largest increase in misdemeanor arrests from 1989–98. Panel C shows that the

[64] Kelling & Sousa, *supra* note __ at 1.

[65] Alfred Blumstein, "Youth Violence, Guns, and the Illicit-Drug Industry." *Journal of Criminal Law and Criminology*, 86: 10–36 (1995); Philip J. Cook and John H. Laub, "After the Epidemic: Recent Trends in Youth Violence in the United States," National Bureau of Economic Research, Working Paper 8571 (2001) (available at http://www.nber.org/papers/w8571).

[66] Steven Raphael and Jens Ludwig, "Do Prison Sentence Enhancements Reduce Gun Crime? The Case of Project Exile," 251-286, in *Evaluating Gun Policy,* Jens Ludwig and Philip J. Cook, ed. Washington, DC: Brookings (2003).

neighborhoods with the highest violent crime rates in 1989 experience the largest declines in such crimes from 1989–98.

Why do precincts with unusually high initial crime rates experience unusually large declines in crime thereafter? Mean reversion is a good candidate—Panel D shows that, as is true with city-level crime data, those police precincts with the largest increases in crime during the crack epidemic have the largest declines thereafter. Most criminologists believe that this increase in violent crime was driven by the crack cocaine epidemic and attendant violence in the crack market, which began to ebb during the early 1990s—hence those places where crack served to drive violent crime to unusually high levels at the height of the epidemic would be expected to experience the largest subsequent declines as the influence on violence from crack use and distribution begin to wane.

The KS analysis seems particularly susceptible to confounding from mean reversion because their model basically relates *changes* in violent crimes (each precinct's linear trend in violent crime over the 1989–98 period) against the *levels* of misdemeanor arrests (average arrests from 1989 to 1998). Put differently, their analysis throws away all of the over-time variation in misdemeanor arrests across precincts from 1989 to 1998, and simply relates variation in the linear trend in violent crime rates across precincts to variation in the average number of misdemeanor arrests over this period.[67] The level of misdemeanor arrests are strongly related to the initial level of violent crimes, as suggested by Figure 1, which may lead to a spurious association between misdemeanor arrests and violent crimes in their study.[68]

---

[67] In this sense their two-level linear growth model is set up in a fashion analogous to Raudenbush and Bryk's 2002 example on p. 167, relating changes in student's test scores measured four times each year over several years with the total hours of instruction the child received. But the time trend in the key treatment variable of interest in the policing example seems to matter much more than in the schooling example offered by Raudenbush and Bryk.

[68] The problem of relating levels against changes can be illustrated with a simple hypothetical example:

Precinct	Year	MA	VC
1	1989	150	500
1	1990	100	400
1	1991	50	300
2	1989	75	500
2	1990	50	475
2	1991	25	450

Precinct 1 has a higher mean number of misdemeanor arrests over the sample period than does precinct 2 (100 versus 50), and also experiences a larger decline in violent crimes per year (100 per year compared to

Table 2 presents the results of a more formal analysis that seems to implicate mean reversion. The first column of Table 2 presents estimates for the parameters in equation (5), where the change in violent crimes within a precinct for the period 1989 to 1998 is regressed against the average misdemeanor arrests within that precinct over the entire 1989 to 1998 period. This simple model is based on the same intuition as the HLM linear growth model of KS, although the key difference is that our dependent variable is the actual change in violent crimes from 1989 to 1998 for each precinct rather than each precinct's estimated linear trend in violent crimes over this period. (The choice by KS to fit a linear trend through these violent crime counts for each precinct is itself a bit puzzling given that Appendix Figure 1 in our paper and Figure 1 in their Manhattan Institute report show a non-linear trend in such crimes in New York over this period, first increasing for a few years and then declining thereafter). The average number of misdemeanor arrests within these precincts has a strong negative relationship with the change in violent crime rates over this period, as with the basic results presented by KS.

(5) $\Delta VC_i = \lambda_1 + \lambda_2 MA_i + v_i$

The remaining columns of Table 2 show that controlling for either the precinct's 1989 violent crimes or change from 1984 to 1989 in violent crimes reduces the coefficient on the average misdemeanor arrest variable by more than two-thirds. The reason is suggested by Figure 1: The average number of misdemeanor arrests over the 1989–98 period is highest in those precincts that experienced the largest increases in crime from 1984–89 and had the largest number of violent crimes in 1989. Statistically relating the average number of misdemeanor arrests from 1989–98 with the decline in violent crimes over this period without controlling for differences across precincts in the run-up in violent crime they experienced during the crack epidemic mistakenly attributes the influence of these initial conditions and subsequent mean reversion to the average number of misdemeanor arrests. Unfortunately none of the proxies for crack, including the borough-level measure of cocaine-related hospital discharges used by Kelling and

---

only 25 in precinct 2). The Kelling-Sousa model applied to these data would suggest a negative relationship between misdemeanor arrests (MA) and the time trend in violent crime (VC) across precincts – more misdemeanor arrests, less crime. However regressing changes against changes – the change over time in violent crimes against the change in misdemeanor arrests – would yield the opposite conclusion.

Sousa, seem to adequately capture the influence of crack markets and use on crime. For example the cocaine proxy used by KS does not have a statistically significant relationship to violent crime rates in their own analysis (see KS, Table 4), nor is this variable statistically significant when included in our own models (and by implication does not change any of the other results shown in our Table 2, either).

In contrast to the weak explanatory power of the KS proxy for crack-related violence—admittedly an extremely difficult phenomenon to quantify—the final column of Table 2 shows that controlling for the set of detailed precinct-level covariates in our dataset yields an estimated relationship between the change in violent crime and the 1989–98 average number of misdemeanor arrests that is about 10% as large as the baseline estimate and no longer statistically significant. These covariates include measures of structural disadvantage (such as the percent of the precinct that is poor, receiving public assistance or has less than a high school degree), demographics (percent of the precinct in their peak offending ages, percent households headed by a female, percent black), measures of physical disorder (percent housing units that are vacant), and police manpower assigned to the precinct.[69]

(6) $\Delta VC_i = \lambda_1 + \lambda_2 \, \Delta \, MA_i + v_{ti}$

Now suppose we instead use the within-precinct over-time variation in the data by relating *changes* in violent crime rates from 1989 to 1998 to *changes* over this period in misdemeanor arrests, as in equation (6). The results from this analysis, shown in Table 3, suggest that if anything, increases in misdemeanor arrests are accompanied by *increases* in violent crime—more misdemeanor arrests, more crime. While the positive relationship between changes in misdemeanor arrests and changes in violent crime is somewhat sensitive to the model specification, there is no evidence from this first-difference model for a *negative* relationship between changes in misdemeanor arrests and violent crime.

[69] The police manpower variable is potentially problematic because some arrests within a precinct might be made by law enforcement officers that are officially assigned to different areas, although our results are not sensitive to excluding this variable. Adding just a control for the percent of the precinct's population that is black to the baseline model in the first column of Table 2 reduces the coefficient on average misdemeanor arrests from -.30 to -.28. Including the Kelling and Sousa measures of cocaine-related hospital discharges and borough-level unemployment rates has little effect on the results shown in Table 2.

The expectation that violent crime should decline in response to an increase in misdemeanor arrests is the key empirical prediction of the argument that broken windows policing is effective. While the Kelling-Sousa analysis does not directly test this prediction, our own analysis shown in Table 3 demonstrates that the data are not consistent with the idea that stepped-up zero tolerance policing reduces crime.

B. The Corman and Mocan (2002) Study

But even putting aside these precinct comparisons, for many observers, the massive drop in New York City's crime rate during the 1990s—coincident with the onset of broken-windows policing in the City—alone provides compelling proof for the efficacy of this policing strategy. Corman and Mocan's analysis provides a more formal version of this same insight, by analyzing monthly time-series data for New York City as a whole. Controlling for city-wide measures of New York's unemployment rate, real minimum wage, incarceration rate, police manpower, number of 14–16 year olds and lagged values of monthly crime rates, they find a negative relationship between city-wide misdemeanor arrest rates and city-wide robbery and motor vehicle theft rates. They do not find a relationship between the former and other types of crime. While Corman and Mocan's time series uses data from 1970 to 2000, graphs of their data suggest that the relationship between misdemeanor arrests and crime would appear to be driven by the unusually large increase in misdemeanor arrests that occurred in New York during the mid- to late-1990s.[70]

What can we conclude about the causal effects on crime of broken windows policing—at least as measured by misdemeanor arrests? Research designs that rely on time series data for a single jurisdiction (in their case, New York) typically provide weak power to rule out alternative explanations for the patterns observed in the data. For example, consider just one candidate counter-explanation, what we term the "Broken Yankees Hypothesis" (BYH). When the New York Yankees do well, violence should decline through the strengthened social ties that develop by the bonding that occurs among the city's residents at local bars and restaurants, with much of the city's attention

[70] Corman and Mocan (2002, Figure 11) show that after hovering between 9,000 and 14,000 between 1982 and 1994, the number of misdemeanor arrests in New York City nearly doubled from 1994 to 2000.

focused on a single, shared goal. When the Yankees do poorly residents may be less likely to aggregate together for a common purpose in communal settings, and moreover the team's poor performance may even spur dissension among New Yorkers over the causes of these failures.

While Corman and Mocan were not willing to share their monthly time-series data with us, we were able to construct on our own an annual time series for New York measuring crime rates and a reasonable proxy for the operational mechanism behind the Broken Yankees Hypothesis, defined as the cumulative number of World Series championships dating back to 1921.[71] Figure 2 provides what appears to be some empirical support for the BYH: the strong performance of Billy Martin's Yankees teams during the late 1970s coincides with a drop in homicides, but even more striking is the massive decline in homicides that accompanies the consistent excellence of Joe Torre's squads beginning in the late 1990s. A time-series regression of the homicide rate against the BYH index and lags of the murder variable frequently yields a negative and statistically significant coefficient (and even controlling for lagged values of robbery to proxy for other criminogenic characteristics), although we note that the magnitude of the point estimate and standard error is somewhat sensitive to the choice of lag length.

While our simple empirical example is not intended to provide a rigorous test of the Broken Yankees Hypothesis, it does serve to highlight the vulnerability of single-city time series findings to counter-explanations. An equally or perhaps even more plausible counter-explanation for New York City's crime pattern during the 1990s comes from the dramatic period effects that caused crime to decline almost everywhere throughout the U.S. during this period, even in cities that did not adopt innovative policing strategies.[72]

[71] These data come from the Yankees web site: http://newyork.yankees.mlb.com/NASApp/mlb/nyy/history/championships.jsp

[72] Levitt 2004. Levitt argues that crime declined throughout the U.S. during the 1990s due to some combination of increased spending police, increased incarceration, the ebbing of the crack epidemic that is widely thought to have caused violent crimes to increase during the late 1980s, and legalization of abortion during the early 1970s. While we find Levitt's explanation persuasive, accepting the specific bundle of causal factors implicated by Levitt is not crucial to our argument for a skeptical interpretation of Corman and Mocan's findings. One need only accept Levitt's observation that crime dropped everywhere over this period to accept the importance of common period effects in understanding crime drops during the 1990s.

Part III: Evidence from the Moving to Opportunity Experiment

Suppose that we could design the ideal social experiment to test the effects of disorder *alone* on criminal behavior. We would start with a sample of people who were at high-risk for criminal offending, and were living in very socially disordered communities. We would then randomly assign some of these families, but not others, to neighborhoods that were less disorderly—ideally, much less disorderly, so that the "treatment dose" that families experience from neighborhood moves would be large enough to yield statistically detectable impacts on behavior. In this idealized experiment we would then wish to follow participants for many years, measure their involvement in criminal activity in different ways (for example with both self reports and administrative arrest records) as well as characteristics of their neighborhoods, and be careful to minimize sample attrition.

The Moving to Opportunity (MTO) experiment, launched in 1994 by the U.S. Department of Housing and Urban Development, conforms in every way to the parameters of the ideal experiment described above. In what follows we provide a review of the effects of MTO on criminal offending by program participants about 5 years after random assignment, and discuss their implications for ongoing debates about broken windows policies.[73]

We show that MTO succeeds in moving families to neighborhoods that are characterized by much lower levels of both physical and social disorder—arguably a more relevant "treatment indicator" for measuring the broken-windows policing hypothesis compared to more indirect policy levers such as misdemeanor arrests that may or may not succeed in reducing disorder. However, we also show that the findings from MTO are not consistent with the idea that changes in neighborhood disorder is enough to change criminal activity.

---

[73] These results are reported in greater technical detail in Kling, et al., *supra* note __, and Jens Ludwig, et al., Neighborhood Effects on Crime Over the Life Cycle (Georgetown University Public Policy Institute, Working Paper, 2005).

A. Background on MTO

Sponsored by the U.S. Department of Housing and Urban Development (HUD), MTO has been in operation since 1994 in five cities: Baltimore, Boston, Chicago, Los Angeles, and New York. Eligibility for the program was restricted to low-income families with children in these five cities, living within public or Section 8 project-based housing in selected high-poverty census tracts.[74] The approximately 4,600 families who volunteered for the program from 1994 to 1997 were randomly assigned into one of three groups. The *Experimental* group was offered the opportunity to relocate using a housing voucher that could only be used to lease a unit in census tracts with 1990 poverty rates of 10 percent or less.[75] Movers through MTO were required to stay in these tracts for at least one year. Experimental group families were also provided with mobility assistance and in some cases other counseling services as well. Families assigned to the *Section 8* group were offered housing vouchers with no constraints under the MTO program design on where the vouchers could be redeemed. Families assigned to the *Control* group were offered no services under MTO, but did not lose access to social services to which they were otherwise entitled such as public housing.

Because of random assignment, MTO yields three comparable groups of families living in very different kinds of neighborhoods during the post-program period. This random assignment helps overcome the self-selection problem that is very likely to plague most previous studies of "neighborhood effects" in general or "broken windows" in particular.

The results summarized below from Kling, Ludwig and Katz (2005) and Ludwig, Kling and Hanratty (2005) measure the delinquency and criminal behavior of youth in MTO using two main sources: survey data and administrative arrest records. Adults were also surveyed but they were not asked about criminal behavior, so we can only measure adult criminal activity using official arrest records. Information on potential mediating

[74] Section 8 project-based housing might be thought of as essentially privately-operated public housing (Olsen 2003). The U.S. Department of Housing and Urban Development contracts with private providers to develop and manage housing projects that include units reserved for low-income families.

[75] Housing vouchers provide families with subsidies to live in private-market housing. The subsidy amount is typically defined as the difference between 30 percent of the household's income and the HUD-defined Fair Market Rent, which equals either the 40th or 45th percentile of the local area rent distribution.

processes that could lead to these outcomes comes from the surveys as well as administrative data on local-area crime rates.[76]

The families in the main survey sample enrolled in the MTO demonstration from 1994 to 1997. At the time of enrollment, the head of household completed a baseline survey which included information about the family as well as some specific information about each child. Descriptive statistics for the baseline characteristics of youth and adults are shown in Table 4. Overall about two-thirds of MTO participants are black, with the program populations in Chicago and Baltimore almost entirely black and an even mix between black and Hispanic in the other sites. MTO households are quite poor, with around three-quarters having been on welfare at baseline. One quarter of household heads had their first child before the age of 18, and only a little more than half of all heads had a GED or high school diploma. Around three-quarters of households report gangs and drugs as the first or second most important reason they enrolled in the MTO program, while around one-half report access to better schools as one of their top two reasons. Eligibility for the MTO program was limited to families in public housing or Section 8 project-based housing located in some of the most disadvantaged census tracts in the five MTO cities and, for that matter, in the country as a whole.

Consistent with random assignment of families to MTO groups, Table 4 shows that there are no statistically significant differences across MTO groups in the fraction of male or female adults or youth who have ever been arrested prior to random assignment or for other baseline characteristics. These results together with those presented elsewhere suggest that MTO random assignment was in fact random.[77]

Of the families with youth in the survey sample (15–20 at the end of 2001), 44 percent of those in the experimental group and 57 percent of those in the Section 8 group complied with treatment (that is, relocated through MTO). These moves lead to substantial differences across treatment groups in neighborhood attributes, as seen in Table 5. Four years after random assignment the average census tract poverty rate (from the 2000 Census) for families assigned to the Section 8 group was 18% lower than that of the Control group, while families assigned to the Experimental group had average census

---

[76] For more detail on these data sources see Appendix A of Kling, Ludwig and Katz (2004).
[77] Kling, et al., *supra* note __.

tract poverty rates 24% below those of Controls. Assignment to either the Experimental or Section 8 groups reduces local-area (police precinct) violent crime rates by 13–15% compared to Controls, with proportionally smaller effects on property crime rates. Given the changes in tract poverty rates induced by MTO, it is surprising that the program engenders so little residential integration with respect to race. The average family in all three MTO groups lives in a census tract where the large majority of residents are also members of racial or ethnic minorities.

The bottom panel of Table 5 presents results from surveys of MTO adults conducted from 4–7 years after random assignment about their perceptions of physical and social disorder within their neighborhoods, as well as the quality of local policing. Adults assigned to the Experimental or Section 8 groups are less likely than Controls to report that neighbors would fail to get involved if local youth were truant or engaging in delinquency (spray painting graffiti). The next row shows that adults in the Experimental and Section 8 group also report less physical disorder as well compared to the reports of adults in the Control group, as measured by the fraction that report that graffiti is a problem in the neighborhood.

Orr et al. (2003) demonstrate that MTO reduces a wide variety of other self-reported measures of neighborhood social and physical disorder as well for both the experimental and Section 8 groups relative to controls, including 20–30% increases in the fraction who feel safe in their neighborhood at night, one-quarter reductions in the share who saw drugs in their neighborhood the past 30 days, 10–15% declines in the share who report problems with litter, trash, graffiti, or abandoned buildings in the neighborhood, 15–25% declines in the share who report problems with public drinking or groups of people hanging out in public spaces, and 10–25% increases in the share who are satisfied or very satisfied with their neighborhoods.[78]

The last row highlights the potential problems with the key explanatory variable used in the Kelling and Sousa (2001) study, namely the police precinct misdemeanor arrest rate. These data are available for New York but not the other MTO sites. The final row of Table 5 shows assignment to either the Experimental or Section 8 groups

[78] Larry Orr, et al., Moving to Opportunity for Fair Housing Demonstration Program: Interim Impacts Evaluation Exhibit 3.5, 66 (Washington D.C.: Office of Policy Development and Research, US HUD, 2003).

substantially *reduces* the local misdemeanor arrest rate compared to the neighborhoods in which the Control group resides.[79] Yet the survey data reported by the MTO participants reveal that Experimental or Section 8 assignment also reduces social and physical disorder. This fact reinforces the notion that there are many ways to reduce disorder within a community beyond stepped-up policing against minor crime, and measures of zero-tolerance policing such as misdemeanor arrests need not be very informative about variation across neighborhoods in actual disorder.

MTO enables us to rigorously test what happens to individuals' criminal behavior when they move to neighborhoods characterized by what broken windows theory predicts should be of greatest relevance—disorder.[80] Of course as Table 5 shows, MTO also induces changes in a variety of other characteristics of the communities in which program participants live, including lower crime rates, fewer low-income residents and more residents with high levels of schooling or occupation in high-status jobs. Findings from MTO thus provide a test of the combined effects of reducing community disorder together with increasing neighborhood affluence, the sort of combined neighborhood changes that we would expect in normal circumstances: When government policies reduce neighborhood disorder, an important local amenity, we would expect gentrification to occur to some degree and so change the socio-economic composition of the neighborhood somewhat.

B. Effects of MTO on Criminal Behavior

Analysis of arrest records and survey data suggests that moving to a less disadvantaged, less disorderly neighborhood on net does not reduce criminal behavior for MTO program participants. While some sub-groups do respond to moves to less

---

[79] This finding is consistent with our analysis above demonstrating that the highest levels and largest increases in misdemeanor arrests in New York City during the 1990s were in the highest crime (and so presumably most disadvantaged) police precincts.

[80] Ideally we would wish to complement the survey-based measures of social and physical disorder obtained from MTO adults with measures for systematic social observation (SSO) of the sort pioneered by the PHDCN research team (Sampson and Raudenbush 1999). Such data were not collected as part of the MTO evaluation for cost and other reasons, although fortunately PHDCN research shows that, at least for Chicago neighborhoods, measures of disorder from SSO and surveys are highly correlated (Sampson and Raudenbush 1999, Table 3, p. 625). SSO measures of disorder are also highly correlated with neighborhood structural disadvantage (Sampson and Raudenbush 1999, Table 2, p. 624). The fact that various measures of disorder and structural disadvantage are all highly correlated means that MTO provides a test for the causal effects of changing all of these neighborhood attributes simultaneously.

disorderly neighborhoods by reducing their involvement in criminal behavior, most notably female youth, these effects are offset by increases in anti-social behavior among other sub-groups. Nothing in broken windows theory or most other models of neighborhood effects suggests that such influences on criminal behavior should be strongly contingent on people's demographic characteristics. So at the very least broken windows is not a complete explanation for how communities influence criminal behavior, since even if the broken windows mechanism is at work for MTO participants other behavioral processes seem to dominate for at least some sub-groups. Moreover for policy purposes what is most relevant is the impact of neighborhood disorder on the overall offending rate, and MTO provides fairly strong evidence that for at least this population there is no net reduction in crime or other anti-social behaviors.

The first row of Table 6, adapted from Ludwig, Kling and Hanratty (2005), summarizes the main MTO finding: When we pool youth and adults, using data for both males and females, and compare overall arrests across MTO groups, we find no statistically significant differences in arrest rates for people who live in neighborhoods with quite different levels of physical and social disorder. The intent-to-treat (ITT) estimates compare the average number of arrests for everyone assigned to the Experimental versus Control group or Section 8 versus Control, regardless of whether the family has moved through the MTO program.[81] The estimates for the effects of treatment-on-the-treated (TOT) are essentially equal to the ITT estimates divided by the fraction of families in the Experimental group (or Section 8 group, for the Section 8-Control estimate) that relocate through the MTO program (Bloom, 1984).[82]

[81] These across-group differences are calculated with regression-adjustment for a series of baseline survey characteristics such as household head race, age, educational attainment and employment status, as well as indicators for pre-random assignment arrests. Because of random assignment, regression adjustment for these characteristics has little effect on the point estimates for the across-group differences but helps improve the precision of our estimates (that is, reduce the standard errors) by accounting for residual variation in the outcome measures of interest. We calculate robust standard errors that are adjusted for the clustering of adult and youth participants within the same households. The estimates also use weights to account for changes in the random assignment probabilities over time during the course of the MTO demonstration.

[82] The TOT estimate will be an unbiased estimate of the effects of treatment on the treated if random assignment is truly random, and if assignment to the treatment group has no effect on those who do not move through MTO. This second assumption may not be literally true, since the counseling services and search assistance offered to treatment families may influence later mobility patterns or other youth behaviors even among families that do not relocate through MTO. The disappointment of searching but failing to find an apartment may also affect non-movers in the treatment groups. If the effects of treatment-

The second row shows that for females the effects on arrests of assignment to the Experimental or Section 8 rather than Control group are negative but not statistically significant, while for males the across-group differences are positive and not quite significant at the conventional cutoff level. The remaining panels of Table 6 disaggregate the results by crime type. For females, the treatment-control group differences in arrests are negative (albeit not significant) for violent, drug and other crimes, but not for property crimes. Males assigned to the Experimental group experience more property-crime arrests than do those assigned to the Control group.

Heterogeneity in people's responses to moving to a less disorderly, less disadvantaged neighborhood arises with respect to age as well as gender. Figure 3 from Ludwig, Kling and Hanratty (2005) shows average arrest rates for MTO participants in each of the three MTO groups by age at the end of 2001, where each panel shows results separately by crime type and gender. These results come from re-estimating the intent-to-treat estimates with an interaction between the treatment indicator variables and a cubic polynomial in age, and then presenting the predicted values of arrests-by-age for each group implied by the parameter estimates. The eight panels of Figure 3 taken together suggest that on balance moving to a less disadvantaged, less disorderly neighborhood has more beneficial (or less detrimental) effects on younger compared to older MTO participants.[83] In national data most crime seems to be committed by adults, even though offending rates per year are higher for teens,[84] so the detrimental effects on adults are not as encouraging as one might like from a policy perspective.

---

group assignment are substantially smaller for those who do not move through MTO compared to those who do, our TOT estimates will approximate the effects of MTO moves on those who move through the MTO program. Mechanically, we calculate TOT estimates using two-stage least squares where we use indicators for random assignment outcomes as instruments for indicators for MTO treatment take-up.

[83] One concern with these results stems from the use of official arrest data, which capture the combined effects of the behavior of both MTO participants and local criminal justice agencies. Variation in the probability of arrest (P) across neighborhoods will affect the likelihood that a criminal event (C) results in arrest (A), with A = P×C. Above we showed that compared to adults assigned to the Control group, those in the Experimental or Section 8 group report that local police are more responsive to calls for service. If responsiveness of police to 911 calls is positively correlated with the probability of arrest, so that the probability of arrest is higher in more affluent areas, then our analysis of arrest data may understate any effects of the MTO experimental and Section 8 treatments that reduce criminal behavior and overstate any effects that lead to an increase in criminal offending.

[84] For example in 1998, 81.3% of all people arrested in the United States for any crime were ages 18 or older at the time; the figures for violent and property crime equal 82.8 and 65.2 percent, respectively. See U.S. Department of Justice, Federal Bureau of Investigation, *Crime in then United States, 1997*,

Additional evidence to suggest that moving to a less disorderly, less disadvantaged community does not on net reduce criminal behavior comes from the self-reported survey data collected for program participants. Survey data on youth reveal no statistically significant differences across groups (for either males or females) in self-reported arrests or delinquency, and an increase in self-reported problem behaviors among males in the Experimental compared to Control groups (Kling, Ludwig and Katz, 2005).[85]

The sharp gender difference in youth responses to moving to a less disorderly, less disadvantaged neighborhood do not appear to be driven by different responses by males and females to the stress and disruption of moving *per se,* in part because in the first few years after random assignment experimental males experience fewer violent-crime arrests compared to controls.[86] The gender difference in effects—also found in recent MTO research on education, substance use, mental health, and physical health[87]–seems to reflect differences in how males and females respond to similar neighborhoods. Boys and girls in the same randomly assigned treatment groups move into similar types of neighborhoods, and within families, brothers and sisters respond differentially to the same mobility patterns.[88]

Washington, DC: US Government Printing Office, 1998, pp. 232-3. For evidence of differential offending rates by age see Figure 3 in the present paper.

[85] Comparing the control group's mean self-reported arrest rate with what is implied by the administrative records suggest that the former are susceptible to considerable under-reporting. Whether this is also true for the behavior problems index, which reveals a positive Experimental-control difference for male youth, is not clear. Of course misreporting would have to be systematically different across groups in order to affect the estimate for across-group differences in behavior problems.

[86] Previous studies of the Baltimore, Boston and New York sites that use the exogenous variation in neighborhoods induced by MTO within individual demonstration sites on balance yield evidence consistent with the view that moving to less distressed communities reduces anti-social behavior by youth, at least in the short run (1 to 3 years from random assignment). In the Boston site, boys in the experimental and Section 8 groups exhibit about one-third fewer problem behaviors compared to controls in the short run [Katz, Kling and Liebman 2001]. For the Baltimore site, official arrest data suggest that teens in both treatment groups are less likely than controls to be arrested for violent crimes. These short-run impacts are large for both boys and girls, but not statistically significant when disaggregated by gender [Ludwig, Duncan and Hirschfield 2001]. Short-term survey data from the New York site reveals no statistically significant differences across groups in teen delinquency or substance use [Leventhal and Brooks-Gunn 2003]. 5-year data for MTO reveal that there were short-term declines in violent criminal offending for males in the experimental versus control groups in every site except for New York, which then dissipated over time, which suggests that changes over time in the effects of neighborhood mobility, rather than idiosyncracies of the Boston or Baltimore sites, is the way to reconcile the short-term and medium-term results from MTO.

[87] *See* Kling and Liebman (2004).

[88] *See* Kling, Ludwig and Katz (2005).

The findings from MTO suggest that either declines in community disorder do not translate into reductions in individual criminal behavior or, at the very least, that any effects on criminal activity from less disorder are outweighed by the countervailing effects from increased neighborhood socio-economic status. These results would seem to suggest that any policy intervention that reduces disorder may not reduce people's criminal behavior if such changes are also accompanied by gentrification that alters the composition of neighborhoods in a fashion analogous to what the Experimental or Section 8 families experience in MTO.

PART IV: Conclusion

When Wilson and Kelling proposed the idea of broken windows in the early 1980s many academic researchers were skeptical about the ability of police activities to reduce crime. But since that time, a new body of empirical literature has, convincingly in our view, demonstrated that increased police spending does indeed reduce crime,[89] and that targeting police resources against the highest-crime "hot spots" can also help prevent criminal activity.[90] Outside of perhaps a few remaining university departments and some Berkeley coffee shops, the notion that "police matter" is (or at least should be) widely accepted. The key scientific and policy question behind the Kelling and Sousa analysis is thus whether asking police to focus on minor disorder crimes as in broken windows policing yields more pronounced reductions in violent crime than does having police focus on violent crimes directly. Our analysis provides no empirical evidence to support the view that shifting police towards minor disorder offenses would improve the efficiency of police spending and reduce violent crime.

We have set out, in this paper, not only to assess the best available evidence for the broken windows theory—George Kelling and William Sousa's 2001 study—but also to rethink the research design most appropriate to studying the broken windows hypothesis. We demonstrate that the pattern of crime changes across New York City precincts during the 1990s that Kelling and Sousa (2001) attribute to broken windows policing is equally consistent with mean reversion: Those precincts that received the most

[89] Levitt, 1997, 2002.
[90] Sherman, 2002.

intensive broken windows policing are the ones with the largest increases and levels in crime during the city's crack epidemic. Consistent with findings elsewhere from city-level data,[91] jurisdictions with the greatest increases in crime during this period tend to experience the largest subsequent declines as well. The data from MTO experiment reveal that moving to a less disorderly, less disadvantaged community on balance does not appear to reduce criminal behavior among the MTO program population. If disorder does affect crime, any such effects are small enough to be dominated by whatever pernicious effects on people's criminal behavior may arise from increases in neighborhood socio-economic status, as would be expected to occur in normal circumstances as neighborhoods with declines in disorder begin to gentrify.

When asked in January 2004 whether the broken-windows theory had ever been empirically verified, James Q. Wilson reportedly told the *New York Times*: "People have not understood that this was a speculation."[92] The theory was not based on empirical data, Wilson emphasized. "We made an assumption that a deteriorating quality of life caused the crime rate to go up."[93] As to whether that assumption is right, Wilson states, still in 2004: "I still to this day do not know if improving order will or will not reduce crime."[94] As Wilson noted in a different interview, "God knows what the truth is."[95]

Yet, understanding the ability of a broken-windows policy to affect disorder and crime is important for both legal and scientific purposes. The notion that broken windows policing might reduce crime is plausible because many of the behavioral mechanisms underlying this policing strategy are at least in principle consistent with existing models of social contagion.[96] Since the Almighty has so far resisted the temptation to publish in scholarly journals, our results help answer Wilson's question in the interim. Our bottom line is that there appears to be no good evidence that broken-windows (or zero-tolerance) policing reduces crime, nor evidence that changing the desired intermediate output of broken-windows policing—disorder itself—is sufficient to change criminal behavior.

---

[91] Raphael and Ludwig, 2003.

[92] Dan Hurley, *On Crime as Science (a Neighbor at a Time)*, N.Y. Times, Jan. 6, 2004, at F1.

[93] Patricia Cohen, *Oops, Sorry: Seems That My Pie Chart is Half-Baked*, N.Y. Times, April 8, 2000, at B7.

[94] Hurley, *supra* note __.

[95] Cohen, *supra* note __

[96] *See, e.g.* Cook & Goss, *supra* note __.

Table 1
Replicating Kelling and Sousa's Multi-Level Model with a Reduced-Form Single Equation Model

Model specification:	Coefficient on MA	Coefficient on MA*A
Counts, not pop weighted	72.68 (5.94)	–.036 (.003)
Counts, pop weighted	70.06 (13.20)	–.035 (.007)
Rates, not pop weighted	509.95 (0.27)	–.255 (.0001)
Rates, pop weighted	139.02 (76.56)	–.070 (.038)

NOTES: Standard errors in parentheses. Each row in Table 2 represents the results from estimating a separate regression of the form $VC_{ti} = \beta_1 + \beta_2 MA_i + \beta_3 A_t + \beta_4 MA_i{*}A_t + \varepsilon_{ti}$ where VC = violent crimes for precinct (i) in year (t), MA = misdemeanor arrests for precinct (i) in year (t), and A = year (ranging from 1989 to 1998). See text for additional details.

Table 2
The Effects of Model Specification and Mean Reversion in the Kelling-Sousa Analysis: Regressing Crime Changes Against Arrest Levels

Dependent variable = Precinct change violent crimes, 1989–98

Explanatory variables:	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Avg. misdemean arrests, 1989–98	–.303** (.035)	–.221** (.023)	–.079** (.019)	–.082** (.022)	–.101** (.019)	–.031 (.024)
Violent crime 1989			–.546** (.029)	–.524** (.057)	–.528** (.048)	–.576** (.055)
Change violent crimes 1984–89		–1.338** (.124)		–.069 (.162)	–.053 (.137)	–.097 (.140)
Chg. Manpower, 1989–98					4.070** (.763)	3.786** (.944)
Other covariates?	N	N	N	N	N	Y
N	75	74	74	74	74	74
R-squared	.504	.811	.915	.914	.939	.970

F:\research\broken_windows2\stata\jens_meanreversion_april605.do
Other covariates include change 1989–98 in poverty, racial and age composition of the population, percent households headed by females, public assistance, vacant housing.

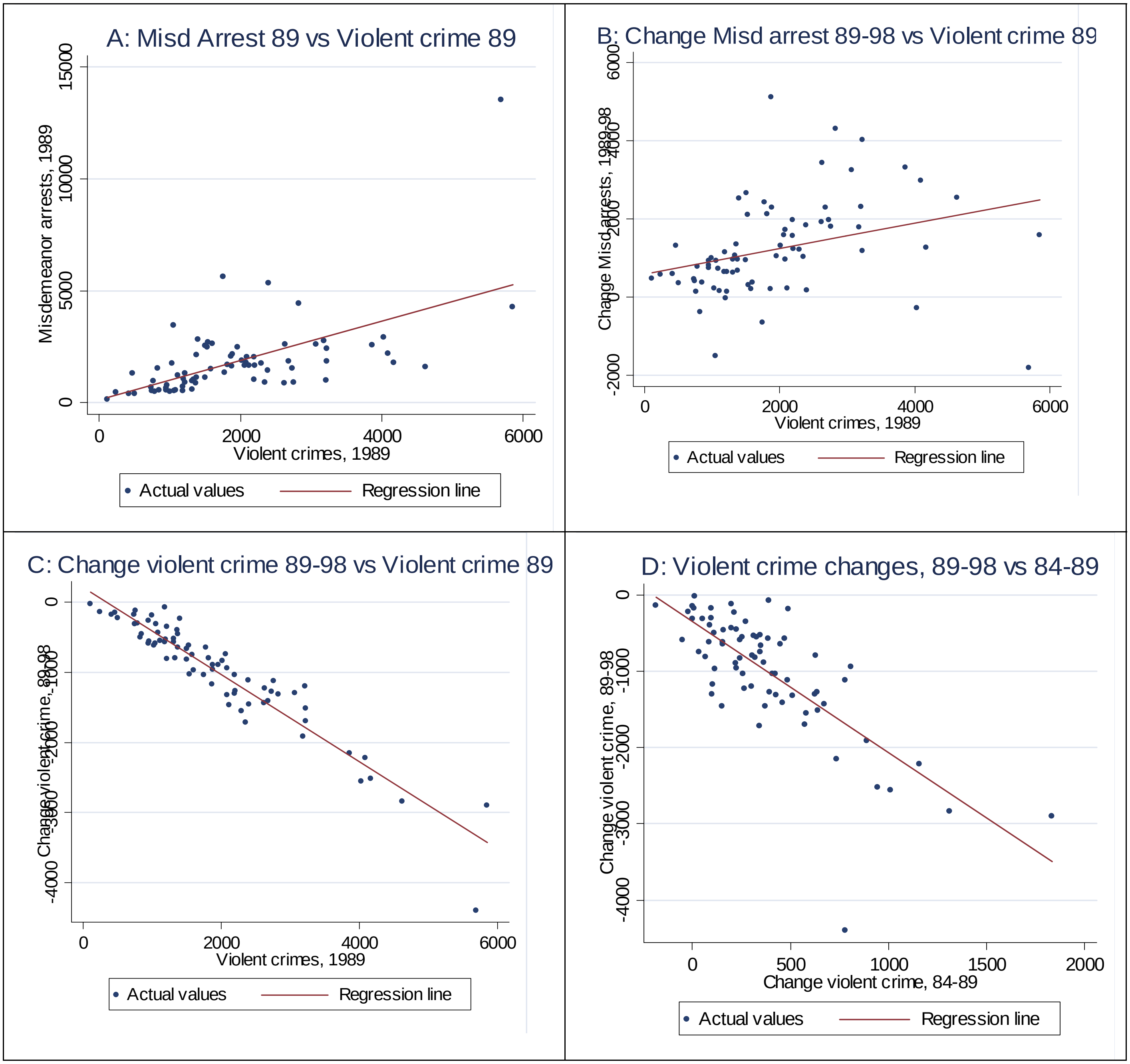
Table 3
The Effects of Model Specification and Mean Reversion in the Kelling-Sousa Analysis: Regressing Crime Changes Against Arrest Changes

Dependent variable = Precinct change violent crimes, 1989–98

Explanatory variables:	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Change avg. misdemean arrests, 1989–98	–.086 (.074)	.046 (.051)	.114** (.022)	.114** (.022)	.094** (.025)	.004 (.030)
Violent crime 1989			–.660** (.023)	–.710** (.039)	–.716** (.039)	–.625** (.041)
Change violent crimes 1984–89		–1.762** (.183)		.214 (.133)	.243* (.133)	–.013 (.127)
Chg. Manpower, 1989–98					1.412 (.963)	3.326** (1.065)
Other covariates?	N	N	N	N	N	Y
N	75	74	74	74	74	74
R-squared	.018	.561	.924	.926	.928	.969

F:\research\broken_windows2\stata\jens_meanreversion_april605.do
Other covariates include change 1989–98 in poverty, racial and age composition of the population, percent households headed by females, public assistance, vacant housing.

Figure 1: Misdemeanor Arrests and Violent Crime in NY Precincts, 1989–98



F:\research\broken_windows2\stata\jens_figure2_march305.do

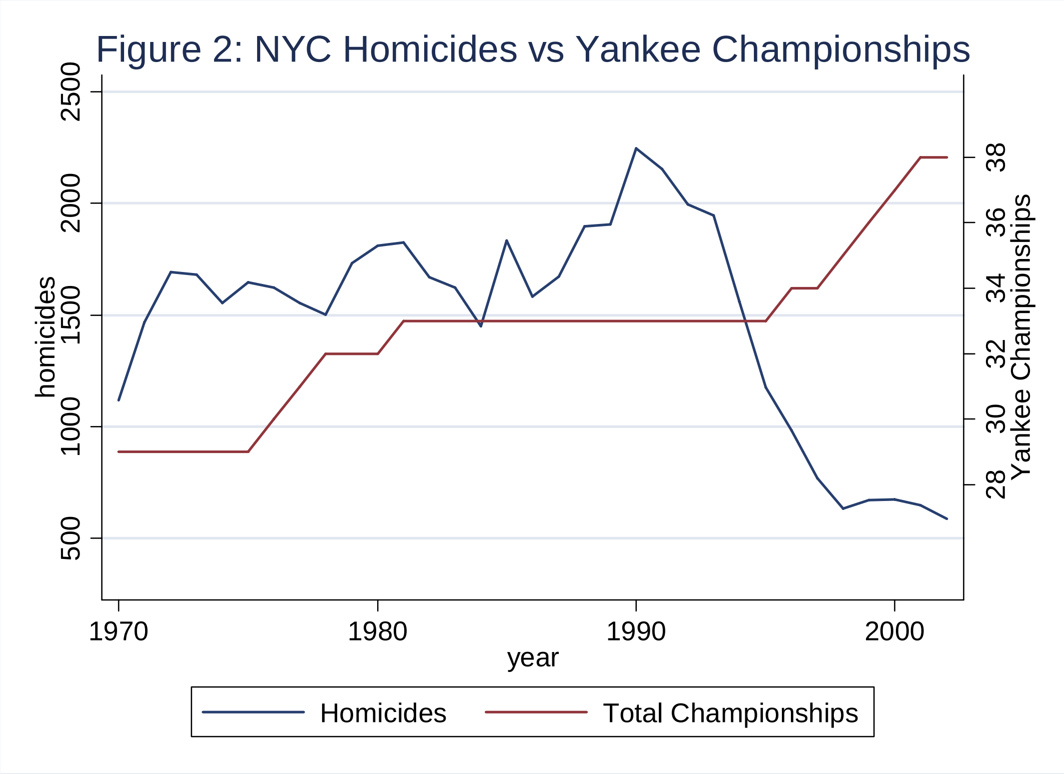


Figure 2: NYC Homicides vs Yankee Championships
homicides
500
1000
1500
2000
2500
Yankee Championships
28
30
32
34
36
38
1970
1980
1990
2000
year
Homicides
Total Championships

Table 4
Baseline Descriptive Statistics for MTO Adult and Youth Samples

	FEMALES Exp	S8	Control	**MALES** Exp	S8	Control
ADULTS						
Black Hispanic	.650 .294	.646 .297	.657 .298	.359 .505	.364 .494	.386 .487
MTO site: Baltimore Boston Chicago LA NYC	 .150 .229 .209 .155 .257	 .162 .223 .209 .149 .257	 .147 .221 .210 .158 .264	 .039 .211 .149 .304 .297**	 .071 .192* .128 .351 .259	 .051 .287 .131 .345 .185
HH on AFDC at baseline	.739	.752	.756	.579	.586	.491
Moved because: Drugs, crime Schools	 .767 .468	 .755 .521**	 .783 .465	 .739 .469	 .755 .577	 .764 .489
Age at end of 2001	38.96	39.40	39.13	43.00	43.39	44.84
Any pre-RA arrest	.258	.231	.260	.375	.423	.354
Missing admin arrest data	.038	.054	.035	.056	.048	.057
N	1,483	1,013	1,102	224	153	166
YOUTH						
Black Hispanic	.647 .296	.606 .318	.640 .304	.609 .329	.605 .333	.612 .339
MTO site: Baltimore Boston Chicago LA NYC	 .168 .187 .210 .165 .270	 .138 .192 .215 .185 .271	 .140 .216 .203 .199 .242	 .151 .166 .220 .195 .269	 .154 .200 .209 .189 .248	 .139 .189 .205 .196 .270
HH on AFDC at baseline	.732	.744	.749	.743	.706	.727
Moved because: Drugs, crime Schools	 .807 .460	 .732 .524	 .782 .483	 .780 .511	 .760 .549	 .791 .505
Age at end of 2001	19.05	18.90	18.90	19.02	18.86	18.96
Any pre-RA arrest	.062	.041	.048	.147	.122	.131
Missing admin arrest data	.057	.048	.055	.059	.063	.061
N	966	651	716	988	691	739

Source: Ludwig, Kling and Hanratty (2005).
NOTES: * = Difference with control mean statistically significant at 10 percent cutoff.
** = Difference with control mean statistically significant at 5 percent cutoff.

Table 5
Effects of Moving to Opportunity Random Assignment on Community Disorder and Other Neighborhood Characteristics

	Control—all	Exp—all	Exp—movers	S8—all	S8—movers
Neighborhood characteristics, 4 yrs after randomization (All 5 MTO sites)					
Avg tract poverty rate	41.68	31.70	19.24	34.38	28.49
% in tract w/ pov 0 – .2	13.43	33.75	65.05	21.91	29.74
% in tract w/ pov .2 – .4	33.73	33.90	26.81	42.46	51.61
% in tract w/ pov > .4	52.83	32.35	8.14	35.63	18.65
Avg tract black	53.93	53.27	41.29	52.05	50.79
Avg tract minority	89.27	84.20	73.94	87.83	85.08
Violent crime rate	235	204	128	200	201
Property crime rate	513	491	373	463	508
Adult survey reports on neighborhood in 2002 (All 5 sites):					
Neighbors would not likely do something about truant children	.65	.53	.43	.57	.58
Neighbors would not likely do something about spraying or graffiti	.47	.36	.26	.41	.40
Problem in neighborhood with graffiti	.48	.38	.19	.40	.32
Problem in neighborhood with police not coming when called	.33	.22	.11	.27	.23
Misdemeanor arrest rate, 4 yrs after randomization (NY site only)	6838	5294	3587	5758	4428

Note: Panel on adult survey reports from Kling, Ludwig and Katz (2005), for adults with youth ages 15–25 at end of 2001.

Table 6
Effects of MTO Random Assignment on Arrests to Youth and Adults

	Controls	E-C ITT	E-C TOT	S8-C ITT	S8-C TOT
All crimes					
All	1.123	0.031	0.072	0.015	0.027
		[0.053]	[0.120]	[0.060]	[0.104]
Females	0.759	–0.049	–0.111	–0.036	–0.062
		[0.047]	[0.107]	[0.051]	[0.089]
Males	1.994	0.225	0.513	0.136	0.237
			[0.131]	[0.298]	[0.147]
[0.256]					
Violent crimes					
All	0.322	–0.017	–0.039	–0.001	–0.002
		[0.017]	[0.040]	[0.022]	[0.038]
Females	0.23	–0.024	–0.055	–0.029	–0.051
		[0.016]	[0.035]	[0.018]	[0.031]
Males	0.54	0	0	0.063	0.109
		[0.042]	[0.096]	[0.054]	[0.094]
Property crimes					
All	0.329	0.056	0.128	0.031	0.054
		[0.026]	[0.060]	[0.027]	[0.048]
Females	0.257	0.025	0.057	0.016	0.028
		[0.030]	[0.068]	[0.030]	[0.052]
Males	0.502	0.131	0.3	0.066	0.115
		[0.050]	[0.115]	[0.052]	[0.091]
Drug crimes					
All	0.26	0	0	–0.024	–0.041
		[0.024]	[0.055]	[0.025]	[0.044]
Females	0.136	–0.032	–0.073	–0.018	–0.031
		[0.020]	[0.045]	[0.022]	[0.037]
Males	0.558	0.077	0.176	–0.036	–0.062
		[0.063]	[0.143]	[0.065]	[0.113]
Other crimes					
All	0.212	–0.008	–0.018	0.009	0.016
		[0.016]	[0.036]	[0.018]	[0.031]
Females	0.136	–0.018	–0.041	–0.005	–0.009
		[0.014]	[0.032]	[0.015]	[0.026]
Males	0.394	0.016	0.037	0.043	0.075
		[0.039]	[0.090]	[0.048]	[0.083]

NOTES: Source—Ludwig, Kling and Hanratty (2005). Sample consists of 2731 males and 6402 females, which reflects a pooled sample of youth 15–25 at the end of 2001 plus MTO adults. The gender disparity in the sample arises because most MTO households are headed by a single female, and so there are far more female than male adults in the sample. The youth sample is gender balanced. Standard errors in bracket

Figure 3: **MTO Treatment Effects on Lifetime Arrests by Age and Gender**

Males

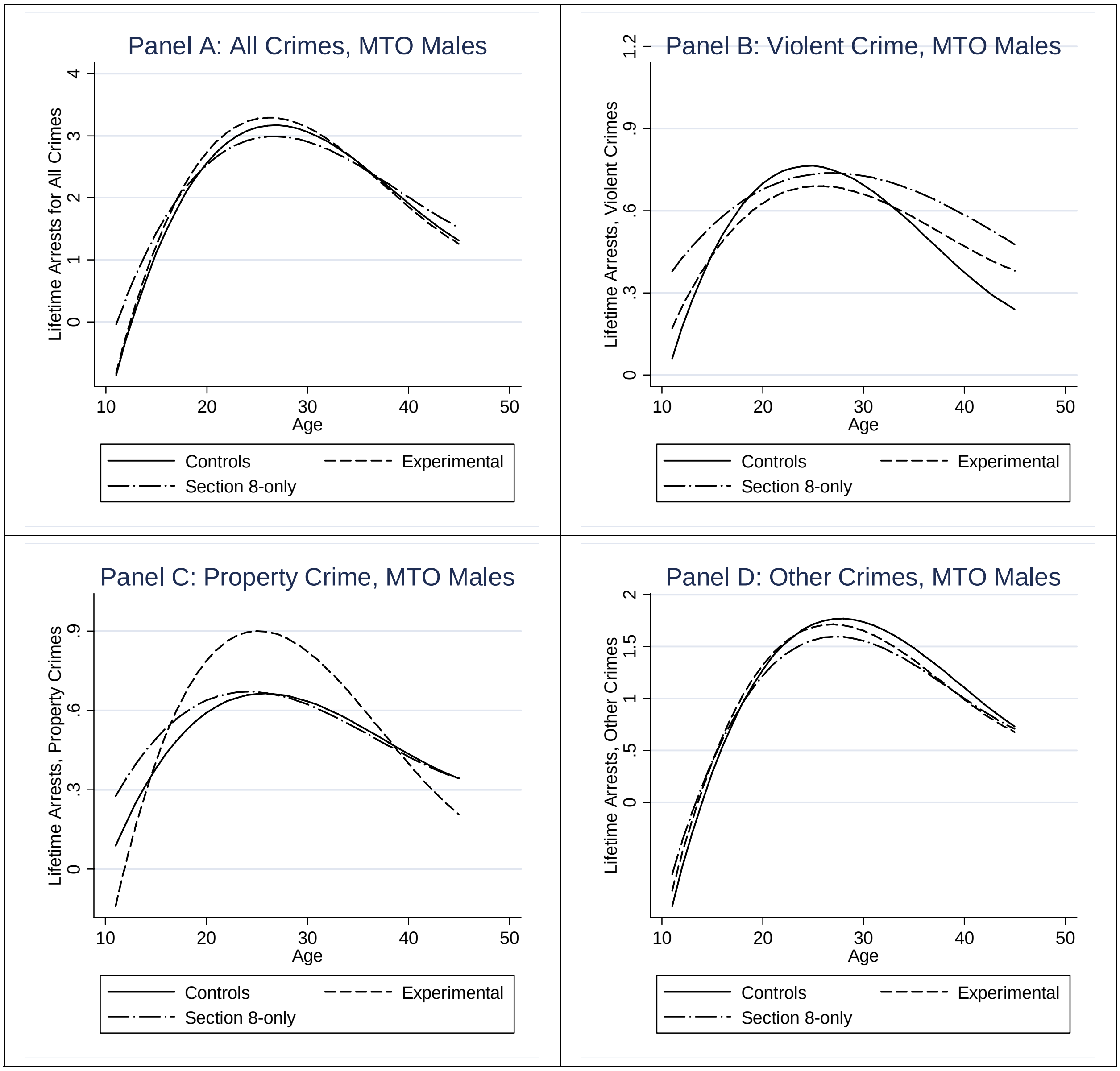
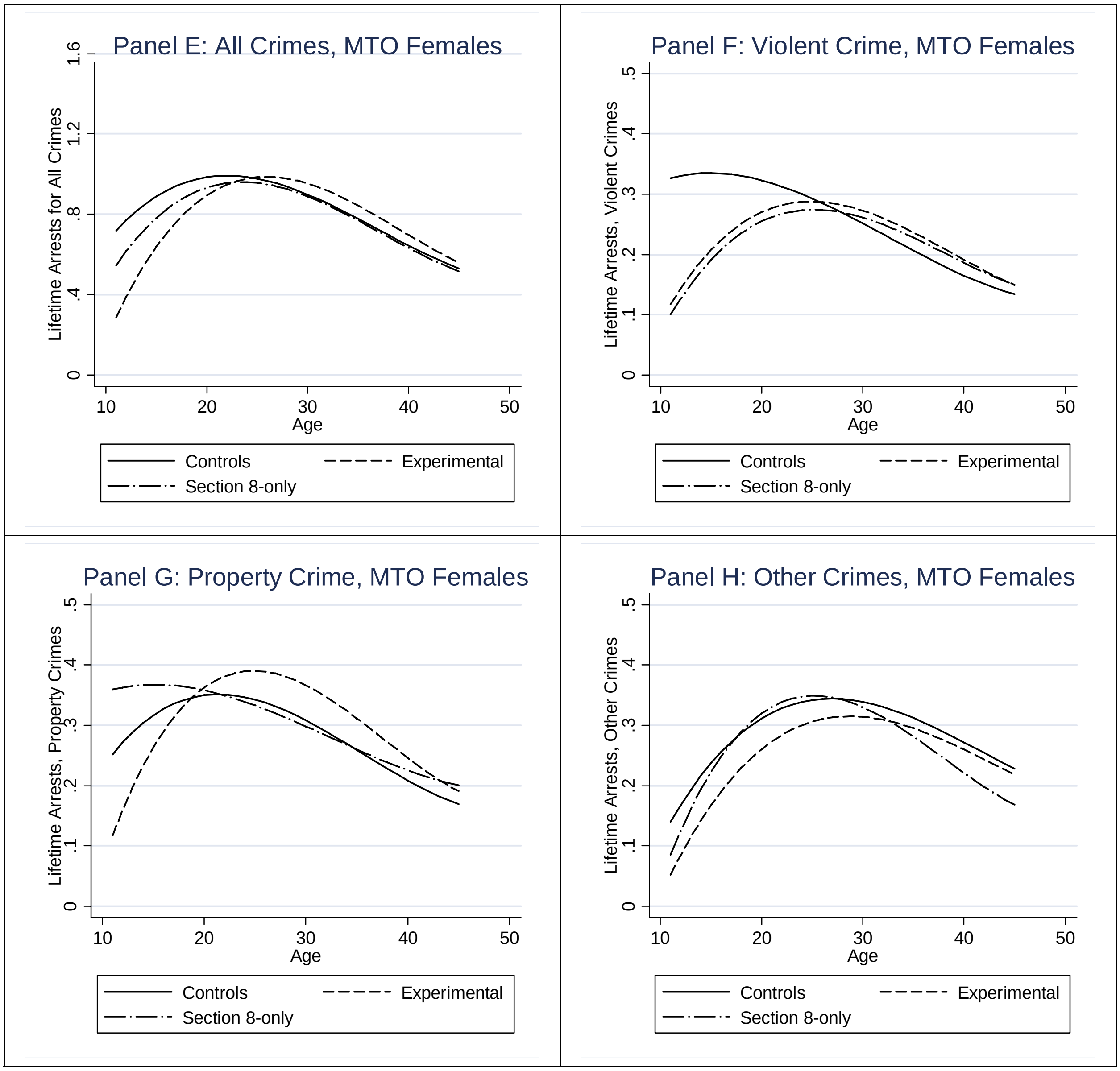


Figure 3, Continued

Females



Programming note: J:\mtopanels\adult_figure2.do

Appendix A: New York City Data

As noted earlier, Kelling and Sousa refused to share their data with us. Fortunately we have been able to obtain the same crime and arrest data from the NYPD used by Kelling and Sousa (2001) as their key dependent and explanatory variables. To measure broken-windows policing, KS use precinct-level reports of total misdemeanor arrests. To measure violent crime, KS use precinct-level reports of four violent offenses (murder, rape, felonious assault, and robbery). In all cases, KS use data from 1989 to 1998. We have these data from 1989 through 2000, and so have the option of examining whether the results are sensitive to the inclusion of additional year's worth of precinct-level information. We also have precinct-level reports for other types of crime, including property offenses, which enables us to explore the pattern of BROKEN WINDOWS POLICING effects across crime types.

One challenge for the KS study and for ours as well is that data on important potential confounding factors is not readily available for NYC at the precinct level. To proxy the effect of the crack epidemic, they use borough-level reports of hospital discharges for cocaine-related episodes. To proxy the number of young males, they use precinct-level school enrollment data. To measure unemployment, they use borough-level gross unemployment data. Whether data measured at the level of New York's five boroughs adequately captures variation in social and policy conditions across the city's 76 separate precincts is an open question. Moreover the hospital discharge data by its nature cannot distinguish between the prevalence of crack use from powered cocaine consumption. The standard concern in the case of poorly measured explanatory variables is attenuation—bias towards zero in the coefficients for these covariates. Some evidence for this concern comes from the fact that the control variables for young males and borough cocaine consumption used by Kelling and Sousa have limited explanatory power in their model (Table 4, KS).

We have also obtained the measures used by KS to capture variation across precincts in the drug problem and economic conditions. Specifically, we have obtained borough-level data on the number of unemployed people from the New York State Department of Labor. We have also obtained data on hospital discharges for drug-related causes from the New York State Department of Health, Bureau of Biometrics.

In addition, however, we attempt to improve upon the KS dataset in part by incorporating census tract-level measures of socio-demographic characteristics, taken from the 1990 and 2000 decennial censuses. (Data for the inter-censal years are linearly interpolated). Because census tract and police precinct boundaries do not perfectly overlap in New York City, we have geocoded both tract and precinct boundaries, and then aggregate tracts up to the precinct level by assuming that the population of tracts that cross precinct boundaries are distributed across precincts proportionately to the tract's land area.[97] We use these census data to calculate measures of each precinct's age distribution, poverty rate, female-headed households, fraction of adults with different levels of educational attainment, median income, and welfare receipt. To measure physical signs of disorder we also control for the fraction of housing units in the precinct that are vacant. Put differently, compared to the data used by KS our dataset includes a much richer set of socio-demographic covariates measured at the precinct rather than some much larger unit of analysis.

Finally, we also incorporate into our dataset a measure of the number of police officers assigned to each precinct in each year by the NYPD. One important conceptual concern with the KS study is whether their key explanatory variable of interest—the misdemeanor arrest rate—captures the effects of changes in how police resources are deployed or instead simply reflects increased police presence. This counter-explanation for the KS findings is of some concern because, as Kelling and Sousa note, from 1994 to 1999 the size of the NYPD force increased by about one-third (2001:19).

Descriptive statistics from our dataset on the key dependent and explanatory variables closely match those reported by KS and by Sousa's doctoral dissertation. For example in Sousa's Table 5–2, the mean number of misdemeanor arrests per precinct for the 1989–98 period is 2247, with a standard deviation of 1968; in our dataset the mean is equal to 2245 with a standard deviation of 1958.[98] Appendix Table 1 repeats this

[97] Suppose for example that census tract 1 lies entirely within precinct A, tract 2 lies entirely within tract B, but 25% of the land area of tract 3 is in precinct A while 75% of the land area of tract 3 is within precinct B. Let X_i be some population characteristic for tract (i), such as percent poor, and let P_i represent the population of tract (i). In this case we calculate percent population poor in precinct A as $(P_1*X_1+.25*P_3*X_3)/(P_1+.25*P_3)$.

[98] Note that following what is apparently the procedure used by KS and Sousa, these means are calculated without weighting by precinct population, and are equal to the raw number of arrests within each precinct, not arrest rate per precinct resident.

comparison for 1989, 1993 and 1998, and again shows that our figures and theirs are quite close.

Appendix Table 1
Average Misdemeanor Arrests per Precinct, Selected Years

	1989	**1993**	**1998**
Harcourt-Ludwig dataset	1754	1795	3034
Kelling-Sousa dataset	1811	1779	3034

Notes: The Kelling-Sousa figures are taken from Table 2 of their (2001) Manhattan Institute report. These figures are mean misdemeanor arrests per precinct, calculated without weighting by precinct population.

Finally, while KS do not report the mean violent crime rate for their dataset over the entire 1989–98 period (the sum of murder, rape, robbery, and felonious assault), their Figure 1A reports the total number of violent crimes for New York City as a whole by year. In our dataset these figures equal 144,375 in 1989, in 1993 it is 131,310, in 1995 it is 97,170, and in 1998 it is 70,725. Each of these numbers, and the overall trend shown in the top panel of our Appendix Figure 1, match closely the numbers represented in their Figure 1A.[99]

[99] See f:\research\broken_windows2\stata\jens_descriptives_jan2105.do

Appendix Figure 1
Violent Crime Counts in New York City by Year,
Harcourt-Ludwig Dataset

f:\research\broken_windows2\stata\jens_figure1_feb1805.do

Readers with comments may address them to:

Professor Bernard Harcourt
University of Chicago Law School
1111 East 60th Street
Chicago, IL 60637
harcourt@uchicago.edu

The University of Chicago Law School
Public Law and Legal Theory Working Paper Series

1. Cass R. Sunstein and Edna Ullmann-Margalit, Second-Order Decisions (November 1999; *Ethics,* v.110, no. 1)
2. Joseph Isenbergh, Impeachment and Presidential Immunity from Judicial Process (November 1999; forthcoming *Yale Law and Policy Review* v.18 #1).
3. Cass R. Sunstein, Is the Clean Air Act Unconstitutional? (August 1999; *Michigan Law Review* #3).
4. Elizabeth Garrett, The Law and Economics of "Informed Voter" Ballot Notations (November 1999, *University of Virginia Law Review*, v. 85).
5. David A. Strauss, Do Constitutional Amendments Matter? (November 1999)
6. Cass R. Sunstein, Standing for Animals (November 1999)
7. Cass R. Sunstein, Culture and Government Money: A Guide for the Perplexed (April 2000).
8. Emily Buss, Without Peers? The Blind Spot in the Debate over How to Allocate Educational Control between Parent and State (April 2000).
9. David A. Strauss, Common Law, Common Ground, and Jefferson's Principle (June 2000).
10. Curtis A. Bradley and Jack L. Goldsmith, Treaties, Human Rights, and Conditional Consent (May 2000; *Pennsylvania Law Review* v. 149).
11. Mary Ann Case, Lessons for the Future of Affirmative Action from the Past of the Religion Clauses? (May 2001, *Supreme Court Review*, 2000)
12. Cass R. Sunstein, Social and Economic Rights? Lessons from South Africa (May, 2000).
13. Jill Elaine Hasday, Parenthood Divided: A Legal History of the Bifurcated Law of Parental Relations (June 2001)
14. Elizabeth Garrett, Institutional Lessons from the 2000 Presidential Election (May 2001).
15. Richard A. Epstein, The Allocation of the Commons: Parking and Stopping on the Commons (August 2001).
16. Jack Goldsmith, The Internet and the Legitimacy of Remote Cross-Border Searches (October 2001).
17. Adrian Vermeule, Does Commerce Clause Review Have Perverse Effects? (October 2001).
18. Cass R. Sunstein, Of Artificial Intelligence and Legal Reasoning (November 2001).
19. Elizabeth Garrett, The Future of Campaign Finance Reform Laws in the Courts and in Congress, The William J. Brennan Lecture in Constitutional Law (December 2001).
20. Julie Roin, Taxation without Coordination (March 2002).
21. Geoffrey R. Stone, Above the Law: Research Methods, Ethics, and the Law of Privilege (March 2002; forthcoming *J. Sociological Methodology* 2002).
22. Cass R. Sunstein, Is There a Constitutional Right to Clone? (March 2002).
23. Emily Buss, Parental Rights (May 2002, forthcoming Virginia Law Review).
24. David A. Strauss, Must Like Cases Be Treated Alike? (May 2002).
25. David A. Strauss, The Common Law Genius of the Warren Court (May 2002).
26. Jack Goldsmith and Ryan Goodman, U.S. Civil Litigation and International Terrorism (June 2002).
27. Jack Goldsmith and Cass R. Sunstein, Military Tribunals and Legal Culture: What a Difference Sixty Years Makes (June 2002).
28. Cass R. Sunstein and Adrian Vermeule, Interpretation and Institutions (July 2002).
29. Elizabeth Garrett, Is the Party Over? The Court and the Political Process (August 2002).
30. Cass R. Sunstein, The Rights of Animals: A Very Short Primer (August 2002).
31. Joseph Isenbergh, Activists Vote Twice (November 2002).
32. Julie Roin, Truth in Government: Beyond the Tax Expenditure Budget (November 2002).
33. Cass R. Sunstein, Hazardous Heuristics (November 2002).

34. Cass R. Sunstein, Conformity and Dissent (November 2002).
35. Jill Elaine Hasday, The Principle and Practice of Women's "Full Citizenship": A Case Study of Sex-Segregated Public Education (December 2002).
36. Cass R. Sunstein, Why Does the American Constitution Lack Social and Economic Guarantees? (January 2003).
37. Adrian Vermeule, *Mead* in the Trenches (January 2003).
38. Cass R. Sunstein, Beyond the Precautionary Principle (January 2003).
39. Adrian Vermeule, The Constitutional Law of Congressional Procedure (February 2003).
40. Eric A. Posner and Adrian Vermeule, Transitional Justice as Ordinary Justice (March 2003).
41. Emily Buss, Children's Associational Rights? Why Less Is More (March 2003)
42. Emily Buss, The Speech Enhancing Effect of Internet Regulation (March 2003)
43. Cass R. Sunstein and Richard H. Thaler, Libertarian Paternalism Is Not an Oxymoron (May 2003)
44. Elizabeth Garrett, Legislating *Chevron* (April 2003)
45. Eric A. Posner, Transfer Regulations and Cost-Effectiveness Analysis (April 2003)
46. Mary Ann Case, Developing a Taste for Not Being Discriminated Against (May 2003)
47. Saul Levmore and Kyle Logue, Insuring against Terrorism—and Crime (June 2003)
48. Eric Posner and Adrian Vermeule, Accommodating Emergencies (September 2003)
49. Adrian Vermeule, The Judiciary Is a They, Not an It: Two Fallacies of Interpretive Theory (September 2003)
50. Cass R. Sunstein, Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation (September 2003)
51. Bernard E. Harcourt, Rethinking Racial Profiling: A Critique of the Economics, Civil Liberties, and Constitutional Literature, and of Criminal Profiling More Generally (November 2003)
52. Jenia Iontcheva, Nationalizing International Criminal Law: The International Criminal Court As a Roving Mixed Court (January 2004)
53. Lior Jacob Strahilevitz, The Right to Destroy (January 2004)
54. Adrian Vermeule, Submajority Rules (in Legislatures and Elsewhere) (January 2004)
55. Jide Nzelibe, The Credibility Imperative: The Political Dynamics of Retaliation in the World Trade Organization's Dispute Resolution Mechanism (January 2004)
56. Catharine A. MacKinnon, Directions in Sexual Harrassment Law: Afterword (January 2004)
57. Cass R. Sunstein, Black on Brown (February 2004)
58. Elizabeth F. Emens, Monogamy's Law: Compulsory Monogamy and Polyamorous Existence (February 2004)
59. Bernard E. Harcourt, You Are Entering a Gay- and Lesbian-Free Zone: On the Radical Dissents of Justice Scalia and Other (Post-) Queers (February 2004)
60. Adrian Vermeule, Selection Effects in Constitutional Law (March 2004)
61. Derek Jinks and David Sloss, Is the President Bound by the Geneva Conventions? (July 2004)
62. Derek Jinks and Ryan Goodman, How to Influence States: Socialization and International Human Rights Law (March 2004)
63. Eric A. Posner and Alan O. Sykes, Optimal War and *Jus Ad Bellum* (April 2004)
64. Derek Jinks, Protective Parity and the Law of War (April 2004)
65. Derek Jinks, The Declining Significance of POW Status (April 2004)
66. Bernard E. Harcourt, Unconstitutional Police Searches and Collective Responsibility (June 2004)
67. Bernard E. Harcourt, On Gun Registration, the NRA, Adolf Hitler, and Nazi Gun Laws: Exploding the Gun Culture Wars {A Call to Historians} (June 2004)

68. Jide Nzelibe, The Uniqueness of Foreign Affairs (July 2004)
69. Derek Jinks, Disaggregating "War" (July 2004)
70. Jill Elaine Hasday, Mitigation and the Americans with Disabilites Act (August 2004)
71. Eric A. Posner and Cass R. Sunstein, Dollars and Death (August 2004)
72. Cass R. Sunstein, Group Judgments: Deliberation, Statistical Means, and Information Markets (August 2004)
73. Adrian Vermeule, Constitutional Amendments and the Constitutional Common Law (September 2004)
74. Elizabeth Emens, The Sympathetic Discriminator: Mental Illness and the ADA (September 2004)
75. Adrian Vermeule, Three Strategies of Interpretation (October 2004)
76. Cass R. Sunstein, The Right to Marry (October 2004)
77. Jill Elaine Hasday, The Canon of Family Law (October 2004)
78. Adam M. Samaha, Litigant Sensitivity in First Amendment Law (November 2004)
79. Lior Jacob Strahilevitz, A Social Networks Theory of Privacy (December 2004)
80. Cass R. Sunstein, Minimalism at War (December 2004)
81. Eric A. Posner, The Decline of the International Court of Justice (December 2004)
82. Tim Wu, The Breach Theory of Treaty Enforcement (February 2005, revised March 2005)
83. Adrian Vermeule, Libertarian Panics (February 2005)
84. Eric A. Posner and Adrian Vermeule, Should Coercive Interrogation Be Legal? (March 2005)
85. Cass R. Sunstein and Adrian Vermeule, Is Captial Punishment Morally Required? The Relevance of Life-Life Tradeoffs (March 2005)
86. Adam B. Cox, Partisan Gerrymandering and Disaggregated Redistricting (April 2005)
87. Eric A. Posner, Political Trials in Domestic and International Law (April 2005)
88. Cass R. Sunstein, Irreversible and Catastrophic (April 2005)
89. Adam B. Cox, Partisan Fairness and Redistricting Politics (April 2005, *NYU L. Rev.* 70, #3)
90. Cass R. Sunstein, Administrative Law Goes to War (May 2005, *Harvard L. Rev., forthcoming*)
91. Cass R. Sunstein, Chevron Step Zero (May 2005)
92. Bernard E. Harcourt, Policitng L.A.'s Skid Row: Crime and Real Estate Development in Downtown Los Angeles [An Experiment in Real Time] (May 2005)
93. Bernard E. Harcourt and Jens Ludwig, Broken Windows: New Evidence from New York City and a Five-City Social Experiment (May 2005)