Exhibit A

Representative Cases In Which Courts Determined That Officers Stopped a Defendant With Reasonable Suspicion Based Only on One or More "Conditionally Justified"[1] Circumstances Listed on Page One of NYPD UF-250

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime
1.	**People v. Jenkins, 209 A.D.2d 164 (N.Y. App. Div. 1994)** Court affirmed that defendant was stopped based on reasonable suspicion when plainclothes officers on patrol directed defendant to stop and to show his hands after the officers had made eye contact with defendant and in response defendant had turned away, began to behave nervously, reached into his waistband and removed a dark object and tossed it into a pile of trash bags.					**X**	
2.	**People v. Pegues, 208 A.D.2d 773 (N.Y. App. Div. 1994)** Court affirmed that officers had reasonable suspicion to stop and frisk defendant when defendant, who was observed driving erratically before pulling into a parking spot, was unwilling to exit the automobile when approached by officers and instead reached under the seat.					**X**	
3.	**People v. Hewitt, 247 A.D.2d 552 (N.Y. App. Div. 1998)** Court affirmed officers had reasonable suspicion to stop and frisk defendant when officers responding to a radio transmission regarding a man with a gun at the location stopped a man who did not fit the description of the suspect, but who they observed holding an open bottle in a paper bag and making furtive movements at a bulge in his waistband that was in the shape of the handle of a 9 millimeter handgun.				**X**	**X**	
4.	**People v. Bush, 171 A.D.2d 801 (N.Y. App. Div. 1991)** Court affirmed stop and frisk of defendant when officers stopped a vehicle for running a red light in which defendant was a passenger and the officer who approached the vehicle observed the defendant make hand movements toward the waistband of his pants and after directing defendant out of the vehicle observed a bulge at defendants waistline, which a frisk revealed was a gun.				**X**	**X**	

[1] The use of the term "conditionally justified" is drawn directly from Fagan's classification scheme as described in his Report and Supplemental Report wherein Fagan defined "conditionally justified" circumstances as the following: (1) carrying a suspicious object, (2) fitting a suspect description, (3) acting as a lookout, (4) wearing clothing indicative of a violent crime, (5) furtive movements or (6) suspicious bulge. See Fagan Report at 50.

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime
5.	**<u>People v. Benjamin</u>, 51 N.Y.2d 267 (1980)** Court reversed and remitted the case to the Appellate Division holding that when officers responded to a radio run advising that there were men with guns at a specified street location and upon arrival observed approximately 30 people outside, including defendant who stepped backwards while simultaneously reaching beneath his jacket with both hands to the rear of his waistband, the radio tip considered in conjunction with other supportive facts, collectively supported reasonable suspicion justifying intrusive police action, including a limited pat-down search which produced a loaded weapon on defendants person.				**X**	**X**	
6.	**<u>People v. Prochilo</u>, 41 N.Y.2d 759 (1977)** Court affirmed stop and frisk was justified when an experienced officer, on routine patrol observed that defendant, while standing and watching other officers interviewing passing pedestrians, was making continuing hand motions toward his side, and that defendant had a bulge on his right hip that the officer observed through defendant's tight outer clothing to be the complete outline of a revolver.				**X**	**X**	
7.	**<u>People v. Arps</u>, 293 A.D.2d 260 (N.Y. App. Div. 2002)** Court affirmed that an officer had reasonable suspicion to stop defendant when officer observed a bulge in defendant's waistband, as well as what appeared to be the protruding handle of a gun.				**X**		
8.	**<u>People v. Goings</u>, 41 N.Y.2d 759 (1977)** Court reversed and remanded, finding that officer's observations of defendant with a bulge in his right-hand jacket pocked which struck the officer as having the configuration and outline of a gun warranted the officer's belief that defendant was carrying a gun and ensuing frisk.				**X**		

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime
9.	**<u>United States v. Pierce</u>, 2007 U.S. Dist. LEXIS 28988 (E.D.N.Y. Apr. 19, 2007)** Court held officers had reasonable suspicion to stop defendant when officers received a specific, detailed and contemporaneous tip from a confidential informant about defendant, including where he was standing, his dress, and the fact he had a gun, in addition to other activity occurring on the street where defendant was located and officers verified each of these facts through personal observations and return calls to the confidential informant.		**X**	**X**	**X**	**X**	
10.	**<u>People v. Sharrieff</u>, 117 A.D.2d 635 (N.Y. App. Div. 1986)** Court reversed and remitted to the Supreme Court, concluding that there was a sufficient basis to stop and frisk the defendant and a second individual when the officers verified by personal observation all elements of an anonymous radio call for an auto theft in progress, including observing a second individual acting as an apparent lookout and defendant approaching the car described in the radio call and drawing away when other people drove down the street, and thereafter saw one of the men drop a shiny, metallic object and defendant drop an ice pick.		**X**	**X**		**X**	
11.	**<u>People v. Wright</u>, 8 A.D.3d 304 (N.Y. App. Div. 2004)** Court reversed and remanded, holding that officers had reasonable suspicion to believe an attempted burglary had been committed, and that it was more probable than not that the defendants, seated in a parked car directly in front of the subject residence, were participating in the crime by acting as lookouts in the getaway vehicle, when officers who responded to a radio run at 3:00 a.m. that two men were breaking into a maroon car in a residential neighborhood arrived and found defendants seated inside a vehicle which matched the description and for which they could provide no proof of ownership, and observed two other men, one of whom was wearing identical sweatshirts to defendants, attempting to break into the adjacent residence.		**X**	**X**			**X**

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime
12.	**People v Fernandez, 16 N.Y.3d 596 (2011)** Court affirmed finding that officer had reasonable suspicion to believe that defendant possessed an illegal weapon, and therefore was authorized to conduct a stop and frisk, when officer observed, in plain view, the "head" of a knife clipped to and sticking out of defendant's pocket from ten to fifteen feet away, because the officer testified that based on his experience, gravity knives are commonly carried in a person's pocket, attached with a clip, with the "head" protruding.	**X**					
13.	**People v. Lathigee, 84 A.D.2d 918 (N.Y. App. Div. 1981)** Court reversed and remanded, finding that police had reasonable suspicion that the occupants of the car had committed a burglary and acted reasonably in stopping the car and ordering the defendants to get out without conducting any preliminary inquiry when police stopped a car occupied by defendants within 30 minutes of a report of a burglary in progress and within three miles of the crime scene that matched the description of a car from which two burglars reportedly had exited, and when police knew that pry marks had been found at the crime scene and upon approaching the defendants' car police observed a "prybar" in the back seat.	**X**	**X**				
14.	**People v. Harris, 57 A.D.3d 1427 (N.Y. App. Div. 2008)** Court affirmed that the police had reasonable suspicion to stop defendant when they encountered defendant in proximity to the street where they had observed the suspects abandon their car and flee on foot, there were no other pedestrians in the area, there was minimal vehicular traffic, and defendant was dressed inappropriately for the extremely cold weather.						**X**
15.	**People v. Watkins, 40 A.D.3d 290 (N.Y. App. Div. 2007)** Court affirmed that police had reasonable suspicion justifying a stop since defendant was the only person in the area of the burglary, was wearing red, which the perpetrator had worn, attempted to walk away from an officer, and was inappropriately dressed for the weather.		**X**				**X**

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime
16.	**People v. La Daniels, 304 A.D.2d 478 (N.Y. App. Div. 2003)** Court affirmed that the police had reasonable suspicion upon which to stop the taxicab in which defendant was a passenger when defendant and the codefendant fit the general description of the perpetrators of a recent, nearby robbery, and the police observed them to be acting nervously before and after they entered the taxi, and the circumstances strongly suggested that defendant and the codefendant had switched clothing in an effort to foil identification as the codefendant was wearing ill-fitting clothes that, according to the description, should have been worn by defendant, as one man's jacket was too small while the other's was too big.		**X**			**X**	**X**

Representative Cases In Which Courts Determined That Officers Stopped a Defendant With Reasonable Suspicion Based Only on One or More "Conditionally Justified" Circumstances Listed on Page One of NYPD UF-250 and "High Crime"

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime	High Crime Area
1.	**People v. Rivera, 183 A.D.2d 674 (N.Y. App. Div. 1992)** Court affirmed stop and frisk of defendant was justified when defendant matched the radioed description of a man with a gun, was observed making furtive gestures towards his waist where the officer observed a large bulge, and could not explain to the officers what he was doing in a robbery prone location at 3:00 a.m.		**X**		**X**	**X**		**X**
2.	**United States v. Bowden, 45 Fed. Appx. 61 (2d Cir. 2002)** Court affirmed judgment of district court that the stop of defendant was justified by reasonable suspicion when defendant who was at a bar notorious for disturbances warranting a police presence, was observed by officers in an initial altercation in the parking lot and returned shortly displaying aggressive behavior and wearing different clothes, including an unseasonably heavy jacket, and attempted to flee after having been told by the police to stop and made hand movements near his waistband.					**X**	**X**	**X**
3.	**In re George G., 73 A.D.3d 624 (App. Div. 2010)** Court affirmed finding of reasonable suspicion justifying a stop and frisk when officers on patrol in a high crime area observed a bulge in defendant's waistband whose shape was consistent with the grip of a pistol and when defendant walked away and positioned his body in an effort to conceal the side where the bulge was located.				**X**	**X**		**X**

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime	High Crime Area
4.	**<u>United States v. Herring</u>, 373 F. Appx. 131 (2d Cir. 2010)** Court affirmed finding of the district court that officers stopped defendant based on reasonable suspicion when defendant was in a high crime area, in the driveway of a house known for drug activity and officers observed defendant cradling a ten- to sixteen-inch object underneath his clothing with one hand while keeping the other hand near his waistband, and defendant ignored repeated directives to stop and show his hands, instead partly turned his back to the officers and walked away.				**X**	**X**		**X**
5.	**<u>People v. Robinson</u>, 279 A.D.2d 323 (N.Y. App. Div. 2001)** Court affirmed finding of reasonable suspicion justifying a stop and frisk when defendant, who was stopped in area with high incidence of taxicab robberies, was observed by officers hailing a cab, engaging in a heated argument with the driver while reaching inside his jacket where officers observed a bulge, and the taxicab immediately drove off at a high rate of speed while defendant remained on the street.				**X**	**X**		**X**
6.	**<u>People v. Smith</u>, 267 A.D.2d 98 (N.Y. App. Div. 1999)** Court affirmed that officer's observations of defendant's furtive movements in a drug prone location in combination with a large bulge the officers believed had the configuration of a machine pistol or large semiautomatic pistol and defendant's refusal to cooperate gave rise to reasonable suspicion to stop and frisk defendant.				**X**	**X**		**X**

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime	High Crime Area
7.	**<u>United States v. Monroe</u>, 2009 U.S. Dist. LEXIS 101776 (E.D.N.Y. Nov. 2, 2009)** Court held the stop and frisk of defendant was justified by reasonable suspicion that the defendant was committing a crime when defendant was observed by officers in a high crime neighborhood, walking quickly, as if on a mission, repeatedly pulled up his pants, and engaged in a confrontation with a second group of individuals and the officers also observed the frightened reactions of bystanders.					**X**		**X**
8.	**<u>People v. Vereb</u>, 122 A.D.2d 897 (N.Y. App. Div. 1986)** Court reversed and remitted to the Supreme Court, finding that officers lawfully stopped defendant based on a reasonable suspicion that defendant had engaged in criminal activity when defendant was observed in a parking lot known to have high incidences of crimes involving automobiles and defendant was behaving in an extremely furtive manner.					**X**		**X**
9.	**<u>People v. Donello</u>, 103 A.D.2d 781 (N.Y. App. Div. 1984)** Court held that defendant's initial stop was proper because the officer had a reasonable suspicion that a crime was committed when the officer observed defendant's furtive behavior in an area known for car thefts and vandalism, but reversed defendant's conviction because defendant's responses to questions did not raise the level of suspicion to probable cause to justify the search and seizure.					**X**		**X**

	CASE	Carrying Objects in Plain View Used in Commission of Crime	Fits Description	Actions Indicative of Acting as a Lookout	Suspicious Bulge/Object	Furtive Movements	Clothes/Disguise Commonly Used in Crime	High Crime Area
10.	**<u>People v. Thurman</u>, 81 A.D.2d (N.Y. App. Div. 1981)** Court reversed the dismissal of the indictment and the suppression of certain evidence, finding that furtive behavior of defendants prior to questioning when observed by experienced officers in a neighborhood with a high rate of daytime residential burglaries gave rise to reasonable suspicion.					**X**		**X**
11.	**<u>United States v. McPhatter</u>, 2004 U.S. Dist. LEXIS 2754 (E.D.N.Y. Feb. 24, 2004)** The Court held that officer had a reasonable suspicion that defendant was committing a crime justifying a stop when defendant was in a high-crime neighborhood carrying an open bottle in a paper bag with the label and contents covered, but which the officer recognized as the bottle as a specific brand of beer.					**X**		**X**
12.	**<u>United States v. Padilla</u>, 548 F.3d 179, 189 (2d Cir. 2008)** Court affirmed the stop and frisk of defendant when he was observed in a high-crime neighborhood with another man, surreptitiously following a third man whose appearance suggested drug use, down an otherwise-deserted street and made movements indicating he was adjusting a concealed firearm.					**X**		**X**

Representative Cases In Which Courts Determined That Officers Stopped a Defendant With Reasonable Suspicion Based Only on One or More "Additional Circumstances"[1] Listed on Page Two of NYPD UF-250

		Report From Victim/ Witness	**Area has High Incidence of Reported Offense of Type Under Investigation**	**Time of Day, Day of Week, Season Corresponding to Reports of Criminal Activity**	**Suspect is Associating with Persons Known for Their Criminal Activity**	**Proximity to Crime Location**	**Evasive, False or Inconsistent Responses to Officer's Questions**	**Changing Direction at Sight of Officer/ Flight**	**Ongoing Investigations**	**Sights and Sounds of Criminal Activity**
1.	**People v. Johnson, 22 A.D.3d 371 (N.Y. App. Div. 2005)** **See also Johnson v. Artus, 2009 U.S. Dist. LEXIS 26534 (S.D.N.Y. Feb. 20, 2009) (report and recommendation of magistrate, denying habeas, adopted by Johnson v. Artus, 2009 U.S. Dist. LEXIS 44839 (SAS) (S.D.N.Y. May 28, 2009), for additional discussion of facts.** Court affirmed holding that officers had reasonable suspicion upon which to stop and frisk defendant when defendant was in a high crime area and his clothing and physical characteristics fit an armed robber's description that was sufficiently specific, given the temporal and spatial factors.	**X**	**X**	**X**		**X**			**X**	

[1] The use of the term "additional circumstances" is drawn directly from Fagan's classification scheme as described in his Report and Supplemental Report wherein Fagan defined "additional circumstances" as circumstances listed on the back of the UF-250 form:: (1) report from victim/witness, (2) area has high incidence of reported offense of type under investigation, (3) time of day, day of week, season corresponding to reports of criminal activity, (4) suspect is associating with persons known for their criminal activity, (5) proximity to crime location, (6) evasive, false or inconsistent responses to officer's questions, (7) changing direction at sight of officer/flight, (8) ongoing investigations, (9) sights and sounds of criminal activity, and (10) other. See Fagan Report at 49.

		Report From Victim/ Witness	Area has High Incidence of Reported Offense of Type Under Investigation	Time of Day, Day of Week, Season Corresponding to Reports of Criminal Activity	Suspect is Associating with Persons Known for Their Criminal Activity	Proximity to Crime Location	Evasive, False or Inconsistent Responses to Officer's Questions	Changing Direction at Sight of Officer/ Flight	Ongoing Investigations	Sights and Sounds of Criminal Activity
2.	**<u>United States v. Simmons</u>, 560 F.3d 98 (2d Cir. 2009)** Court affirmed that officers had reasonable suspicion to stop defendant when responding to an anonymous 911 call of an assault in progress, possibly involving a weapon, and the officers own observations corroborated that defendant matched the description of the suspect and was present at the specified location along with a gathering of people, late night, and in a high-crime area, and when defendants behavior – walking towards officers with his hands in his pocket and non-compliance with the first order to stop – reinforced the officers' determination that he may have been involved in criminal activity.	**X**	**X**	**X**		**X**				**X**

		Report From Victim/ Witness	Area has High Incidence of Reported Offense of Type Under Investigation	Time of Day, Day of Week, Season Corresponding to Reports of Criminal Activity	Suspect is Associating with Persons Known for Their Criminal Activity	Proximity to Crime Location	Evasive, False or Inconsistent Responses to Officer's Questions	Changing Direction at Sight of Officer/ Flight	Ongoing Investigations	Sights and Sounds of Criminal Activity
3.	**<u>United States v. Freeman</u>, 2011 U.S. Dist. LEXIS 129257 (S.D.N.Y. Nov. 8, 2011)** Court held that officers had reasonable suspicion to stop defendant when police received late night anonymous 911 calls that were sufficiently reliable – caller called twice and the physical description provided was accurate, as was the report of defendants movements – of a man with a gun in a high crime area arguing with a woman, and when the defendant was the only person in the area matching the caller's description and his evasive behavior in response to statements by the police corroborated the anonymous tip that the suspect may have a gun.	X	X	X		X		X		

		Report From Victim/ Witness	Area has High Incidence of Reported Offense of Type Under Investigation	Time of Day, Day of Week, Season Corresponding to Reports of Criminal Activity	Suspect is Associating with Persons Known for Their Criminal Activity	Proximity to Crime Location	Evasive, False or Inconsistent Responses to Officer's Questions	Changing Direction at Sight of Officer/ Flight	Ongoing Investigations	Sights and Sounds of Criminal Activity
4.	**<u>United States v. McCargo</u>, 464 F.3d 192 (2d Cir. 2006)** Court affirmed that officers had reasonable suspicion that defendant was involved in criminal activity and therefore the stop of defendant was constitutional when officers responding to a 911 call for an attempted burglary (but that did not provide a suspect description) observed defendant walking alone in a high crime area at approximately 1:00 a.m., 200 feet from the crime scene.	**X**	**X**	**X**		**X**				
5.	**<u>United States v. Muhammad</u>, 463 F.3d 115 (2d Cir. 2006)** Court held that officers had stopped defendant on the basis of reasonable suspicion and properly seized a rifle from defendant when a 911 caller provided a detailed description of the suspect including that the suspect was carrying the gun out in the open, a negligible amount of time elapsed between the call and the officers' response, no one else was in the vicinity, the neighborhood had a high incidence of crime, and the suspect attempted to flee when the officers indicated their desire to speak with him.	**X**	**X**			**X**		**X**		

		Report From Victim/ Witness	Area has High Incidence of Reported Offense of Type Under Investigation	Time of Day, Day of Week, Season Corresponding to Reports of Criminal Activity	Suspect is Associating with Persons Known for Their Criminal Activity	Proximity to Crime Location	Evasive, False or Inconsistent Responses to Officer's Questions	Changing Direction at Sight of Officer/ Flight	Ongoing Investigations	Sights and Sounds of Criminal Activity
6.	**<u>Sutton v. Duguid</u>, 2007 U.S. Dist. LEXIS 35853 (E.D.N.Y. May 16, 2007)** Court held that defendants had reasonable suspicion to stop plaintiff based on the observed narcotics activity in a high crime area, plaintiff's proximity to the individual identified as involved in the sale of narcotics, and plaintiff's effort to walk away from the commotion as soon as it broke out.		**X**			**X**		**X**		
7.	**<u>People v. Sierra</u>, 83 N.Y.2d 928 (1994)** Court affirmed that officers were justified in stopping defendant on the belief that he was committing or about to commit a drug-related crime when defendant was observed in a high crime area calling "over here, over here" to a man exiting a parked vehicle with New Jersey license plates and promptly walk away upon spotting the police, refused to approach the police vehicle and subsequently fled.		**X**					**X**		

Cases Relied on by Fagan for His Analysis of the Constitutional Sufficiency of Stops, Questions and Frisks That Have Been Either Inaccurately Interpreted or Are Subject to an Alternative Interpretation

		Fagan's Interpretation/Analysis	Inaccurate Interpretation or Alternative Interpretation
1.	**People v. Francis,** **847 N.Y.S.2d 398** **(N.Y. Sup. Ct. 2007)**	"Nevertheless, an officer cannot stop or frisk an individual simply because they possess an object that could either be contraband or be innocently possessed. *See People v. Francis*, 847 N.Y.S.2d 398, 401-02 (N.Y. Sup. Ct. 2007) (holding that an officer who observed that an object that looked like a knife, which was clipped inside a suspects [sic] pocket, did not have reasonable suspicion to believe that the knife was an illegal gravity knife and not a permissible knife)." See Fagan Report, Appendix D at B.1.	Fagan's reliance on People v. Francis for this assertion is based on an inaccurate interpretation of the court's opinion. The court **did not hold** that the officer in People v. Francis was not permitted to frisk the defendant because the officer was not 100% certain the object in defendant's pocket was an illegal knife. Rather, the court held that the officer "had a founded suspicion of criminal activity, which would have justified a common-law right of inquiry," and thus "the officer should have conducted an inquiry to determine whether his suspicions that defendant possessed an illegal knife were accurate." 847 N.Y.S.2d at 402. The court did not preclude the possibility that had the officer conducted the permitted inquiry the officer would have had reasonable suspicion sufficient to forcibly stop and frisk defendant.
2.	**People v. Saad,** **859 N.Y.S.2d 906** **(N.Y. Crim. Ct. 2008)**	"Standing alone, the fact that an individual is in possession of objects commonly used in the commission of crimes does not provide an officer with the reasonable suspicion necessary to stop or frisk that individual. *See People v. Saad*, 859 N.Y.S.2d 906 (N.Y. Crim. Ct. 2008) (holding that officers lacked reasonable suspicion to stop a man seen walking down the street, pushing a shopping cart with a tire iron protruding, and looking into parked cars)." See Fagan Report, Appendix D at B.1.	Fagan's reliance on People v. Saad for this assertion misstates the facts, and the facts set forth in People v. Saad could support an alternative assertion. First, it was not the People's assertion that defendant's possession of a tire iron alone provided the officer with reasonable suspicion to stop defendant; additional factors were defendant's presence in a high crime area and the fact he was looking into parked cars. Second, the court's decision does not preclude the possibility that, on these same facts, officers would have been justified in making a common-law right of inquiry and, depending on the answers provided, that the officer's would have had reasonable suspicion sufficient to forcibly stop and frisk defendant. See Saad, 859 N.Y.S.2d 906 ("The presence of the tire iron, the location of the encounter, the additional information gleaned, including the statement that defendant was going home, when in fact, he was traveling in a different direction, the presence of the utility knife and the open case of possession of burglar's tools, **taken together, might very tenuously support a common law right to inquire based upon a founded suspicion that criminal activity is afoot**.") (emphasis added).

		Fagan's Interpretation/Analysis	Inaccurate Interpretation or Alternative Interpretation
3.	**<u>People v. Moore</u>, 6 N.Y.3d 496 (2006);** **<u>People v. William II</u> 772 N.E.3d 1150, 1153 (2002);** **<u>Florida v. J.L.</u> 529 U.S. 266 (2000)**	Fagan asserts that "[e]ven if the anonymous information describes a specific person, this factor alone cannot justify a stop and frisk." <u>See</u> Fagan Report, Appendix D at B.2., citing <u>People v. William II</u>, 772 N.E.3d 1150, 1153 (N.Y. 2002); <u>Florida v. J.L.</u>, 529 U.S. 266 (2000). Fagan further asserts that "[a]n anonymous tip can only provide the basis for a stop if it contains predictive information 'so that the police can test the reliability of the tip.'" <u>See</u> Fagan Report, Appendix D at B.2., citing <u>People v. Moore</u>, 6 N.Y.3d 496, 499 (2006).	Fagan's interpretation of when a suspect description provided by an anonymous tipster or witness may provide the basis for a stop or frisk fails to address a significant point – that the Second Circuit has held that the officers' corroboration of anonymous information identifying a suspect that was insufficient in <u>J.L.</u>, "is entitled to **more weighty consideration in the context of an emergency 911 call...[because] a 911 call reporting an ongoing emergency is accorded a higher degree of reliability and requires a lesser showing of corroboration**." <u>See</u> <u>United States v. Simmons</u>, 560 F.3d. 98, 108 (2d Cir. 2009). Further, while Fagan asserts that an anonymous tip alone cannot justify a stop and frisk, the Second Circuit in <u>Simmons</u> declined to address that very issue because there were additional factors that supported the stop in question. <u>See</u> <u>id.</u> Accordingly, it is remains to be determined whether an anonymous 911 call that identifies the suspect and reports an ongoing emergency could, alone, justify a stop and frisk.
4.	**<u>People v. Howard</u>, 542 N.Y.S.2d 536 (N.Y. App. Div. 1989)**	Fagan asserts that "[a]bsent additional factors, the simple fact that a person is observing a location and appears to be on the lookout for something is insufficient to justify a stop and frisk." <u>See</u> Fagan Report, Appendix D at B.4, citing <u>People v. Howard</u>, 542 N.Y.S.2d 536, 538 (N.Y. App. Div. 1989)."	Fagan's reliance on <u>People v. Howard</u> for this assertion illustrates the fact that alternative interpretations can be arrived at on the same set of facts as the decision contains a lengthy dissent by Justice Smith. <u>See</u> <u>Howard</u>, 542 N.Y.S.2d at 183-185. In dissent, Justice Smith finds the conduct of the police in stopping and frisking defendant was justified.
5.	**<u>People v. Prochilo</u>, 41 N.Y.2d 759, (N.Y. 1977)**	Fagan narrowly allows that "an officer may frisk an individual if he **observes a bulge that is plainly shaped like a firearm**." <u>See</u> Fagan Report, Appendix D at B.5, citing <u>People v. Prochilo</u>, 41 N.Y.2d 759, 762 (1977) (emphasis added).	Fagan's narrow interpretation based on <u>People v. Prochilo</u> is directly contradicted by United States Supreme Court precedent, as in <u>Terry v. Ohio</u>, (392 U.S. 1 [1968]), the Court upheld the right of the police to stop and frisk a person reasonably suspected of criminal activity, notwithstanding the fact that the detective never saw any outline or bulge before he frisked the three individuals. Furthermore, in the dissent of Justice Smith in <u>Howard</u>, as he rejected Fagan's narrow reading of <u>Prochilo</u> stating it "does not stand for the proposition that no frisk can ever be made unless the police see the outline of a gun." <u>Howard</u>, 542 N.Y.S.2d at 184.

		Fagan's Interpretation/Analysis	Inaccurate Interpretation or Alternative Interpretation
6.	**People v. Hudson, 527 N.Y.S.2d 919 (N.Y. App. Div. 1988)**	Fagan asserts that "[c]arrying a suspicious object, even if sufficient to justify a stop, does not justify a frisk unless there are other indications of dangerousness." See Fagan Report, Appendix D at B.5, citing People v. Hudson, 527 N.Y.S.2d 919 (N.Y. App. Div. 1988).	Fagan's reliance on People v. Hudson for this assertion is misleading as the case is easily distinguishable. In Hudson, the officer first saw defendant carrying a three-foot object wrapped in a sheet down the street and attempted to stop defendant, but he walked away. When the officer saw defendant a second time **an hour later** he frisked him **without making an inquiry**, and while the officer testified that the frisk was for safety, the record contained no facts supporting a finding he had a reason to suspect he was in danger. By contrast, see United States v. Herring, 373 F. Appx. 131 (2d Cir. 2010), discussed on page 7, herein, in which the court affirmed a stop and a frisk conducted immediately thereafter was made with reasonable suspicion when defendant was in a high crime area, carrying a suspicious object under his clothes, ignored officer's directives to stop and walked away.
7.	**People v. Powell, 667 N.Y.S.2d 725 (N.Y. App. Div. 1998);** **United States v. McCrae, 2008 U.S. Dist. LEXIS 2314 (E.D.N.Y. Jan. 11, 2008);** **United States v. Doughty, 2008 U.S. Dist. LEXIS 74248 (S.D.N.Y. Sept. 18, 2008)**	Fagan asserts that "[w]ithout more, furtive movements potentially indicative of carrying a firearm cannot give rise to reasonable suspicion." See Fagan Report, Appendix D at B.7, citing People v. Powell, 667 N.Y.S.2d 725,727 (N.Y. App. Div. 1998); United States v. McCrae, 2008 U.S. Dist. LEXIS 2314, *9-10 (E.D.N.Y. Jan. 11, 2008); United States v. Doughty, 2008 U.S. Dist. LEXIS 74248, *18 (S.D.N.Y. Sept. 18, 2008).	Fagan's assertion is based on an unreasonably narrow reading of the UF-250 form as it excludes the possibility that the "more" that is necessary to combine with an officer's mark in the "furtive movements" circumstance to justify reasonable suspicion for a stop is included elsewhere on the face of the UF-250. As described on pages 1, 8 and 9, herein, courts have routinely upheld stops as justified and based on reasonable suspicion when officers act based on observed furtive movements in high crime areas. See, e.g., United States v. Padilla, 548 F.3d 179, 189 (2d Cir. 2008); United States v. Monroe, 2009 U.S. Dist. LEXIS 101776 (E.D.N.Y. Nov. 2, 2009); United States v. McPhatter, 2004 U.S. Dist. LEXIS 2754 (E.D.N.Y. Feb. 24, 2004); People v. Jenkins, 209 A.D.2d 164 (N.Y. App. Div. 1994); People v. Pegues, 208 A.D.2d 773 (N.Y. App. Div. 1994); People v. Vereb, 122 A.D.2d 897 (N.Y. App. Div. 1986); People v. Donello, 103 A.D.2d 781 (N.Y. App. Div. 1984); People v. Thurman, 81 A.D.2d (N.Y. App. Div. 1981).
8.	**People v. Giles, 647 N.Y.S.2d 4 (N.Y. App. Div. 1996)**	Fagan asserts that "[s]tanding alone, seasonally inappropriate attire does not justify a stop or frisk because 'wearing a long winter coat on a hot summer night…is no more than 'odd' behavior' and odd behavior alone cannot justify a stop and frisk." See Fagan Report, Appendix D at B.7, citing People v. Giles, 647 N.Y.S.2d 4, 6 (N.Y. App. Div. 1996)	Fagan's reliance on Giles is both misleading and inaccurate because the court did not find that the officer did anything to exceed the first tier of police intrusion under *DeBour* until defendant's furtive move, and therefore the court did not address whether the fact defendant was wearing seasonally inappropriate attire justified a stop and frisk. Giles 647 N.Y.S.3d at 8. In fact, when discussing the import of the clothing worn by defendant the court actually stated that the unseasonable winter coat, when taken together with the motion of adjusting an object in the rear of his waistband, assumes another possible meaning – "that the inappropriate garb is worn for the very purpose of hiding something." Id. at 6.

Exhibit B

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK
--X
DAVID FLOYD, et al.,

PLAINTIFFS,

-against- Case No.:
08 Civ. 01034

CITY OF NEW YORK et al.,

DEFENDANTS.
--X

DATE: February 9, 2011

TIME: 10:00 a.m.

EXAMINATION BEFORE TRIAL of an Expert Witness, JEFFREY A. FAGAN, Ph.D., on behalf of Plaintiffs, taken by the Defendants, pursuant to a Notice, and to the Federal Rules of Civil Procedure, held at the office of Special Federal Litigation, New York City Law Department, 100 Church Street, New York, New York 10007, before John A. Lugo, a Notary Public of the State of New York.

Q. If, at any time, you give me an answer that I'm not quite sure about, I'll ask you to clarify it. And in that respect, I'll ask your indulgence since you are the expert here and we are just lawyers. Is that agreeable to you?
A. Sure.
Q. Professor, do you have a law degree?
A. No.
Q. Do you have any formal legal training?
A. No.
Q. Have you ever taken courses at any law school?
A. No.
Q. You're a professor at Columbia Law School, correct?
A. Correct.
Q. And what courses do you receive?
A. I teach Law and Social Science, Juvenile Justice, Drug Policy, Policing, Criminal Law, Juvenile Law. I think that's about it. Seminar in Criminology.
Q. And do each of those courses have its own syllabus?
A. Yes.
Q. And could I ask you to provide your counsel for production to us a syllabus for fall 2010 and spring 2011 for the courses you teach?

A. Physical disorder was converted from the 55 PUMAs into the 75 precincts.
Q. And so your findings, then -- I just want to make sure I understand this.
The variables that you used, I take it, were precinct specific; is that correct?
A. We constructed precinct specific measures of physical disorder.
Q. Now, you controlled in your analysis for crime, obviously, correct?
A. Yes.
Q. What was the benchmark that you used to control for crime?
A. The number of crime complaints.
Q. Crime complaints per precinct?
A. Yes, in the time period, in each of time periods that we tested.
Q. Now, if you controlled for crime, why did you also control independently for factors that are known to be associated with crime, such as social and physical disorder?
MR. HELLERMAN: Object to the form of the question.
A. In our previous research we had found that these factors themselves were associated with stop patterns

before. It's an accurate rendering of crime in a particular place relative to other places and relative to other times. Whether it's an accurate count of the actual number of crimes, no.

Q. You don't agree with that?

A. No.

Q. And why do you believe that arrest data is not necessarily an accurate picture of the level of crime?

A. Because the capacity of police departments who would investigate crimes and make arrests based on probable cause is quite variable. Clearance rates range from 23 percent to 78, 80 percent.

Q. Clearance rates are rarely 100 percent, right?

A. They're never 100 percent.

Q. And since clearance rates are never 100 percent, there's always going to be some missing data there about crimes, correct?

MR. HELLERMAN: Object to the form of the question.

A. By definition that's the case.

Q. Now, in the JASA study, correct me if I'm wrong, you concluded that the arrest rate was the best available measure of the race-specific crime rate, correct?

A. We make that statement, yes.

Q. Did you agree with that statement, at the time?

A. Yeah.
Q. Did you agree with it today?
A. No. Well, I'd have to think about it. Best available where? And for which types of crime?
Q. New York City for any type of crime.
A. New York City, I think the best available data for race-specific crime, they're probably all bad. So if I said anyone was superior to another, I'd be talking about the difference between the Houston Rockets and the New Jersey Nets.
Q. Now, why do you believe that all of the available measures, race-specific measures of crime, are bad in New York City?
A. Well, arrest data as a measure of crime is confined to -- the accuracy depends on the severity of the crime for robberies, et cetera. It's probably more accurate than for larcenies and for larcenies more accurate than say for misdemeanor assaults, but these are relatively low rates.
For other measures based on suspect descriptions provided by victims, we know from the data provided in this case, for example, that that's a fairly low rate of suspect race identification.
Q. Well, would you agree that for severe crimes, violent crimes, robbery, rape, assault, first-degree

New York City in the previous year, 1997, as recorded by the Division of Criminal Justice Services (JCDS) of New York State, and categorized by ethnic group and crime type. This was deemed to be the best available measure of the local crime rates categorized by ethnicity and directly addressed concerns such as safir's, S-A-F-I-R-'S, that stop rates be related to the ethnicity of crime suspects.

Did I read the article correctly?

A. Yes.

Q. Did you qualify your use of race-specific arrest rates as a measure of race-specific crime rates?

MR. HELLERMAN: In the language you just quoted?

A. In this language?

Q. Did you in the --

MR. LARKIN: Well, before I repeat myself, please don't do that again. All right. Please don't do that again, Counsel. Really. I object to comments on the record. You know it's not appropriate. Come on, you know. Please.

Q. In the portion that I just read, did you qualify your use of race-specific arrest data as a measure of the race-specific crime rate?

A. No.

Q. Did you anywhere else in the article qualify your

descriptions, correct?
A. Yes.
Q. And there were a significant proportion of reported crimes for which there was no suspect description, right?
A. Right.
Q. If you had all the data for suspect descriptions for the reported crimes, in theory, would that have assisted you in your analysis, in this case?
A. If we had perfect data on a suspect description?
Q. Yes.
A. Sure.
Q. And in what way would it have assisted you?
A. We would have used that as an additional -- probably would have used it as a sensitivity check.
Q. And how would you have used it as a sensitively check exactly?
A. We probably would have included in a sensitively run instead of the crime complaint data, we probably would have used separate measures of the different crime complaint categories by race.
Q. Would you have used as a control for crime the racial breakdown based on suspect descriptions, assuming perfect data for suspect descriptions?
A. If it was perfect data?

Q. Yes.
A. I believe that's what I just said.
Q. Okay. I'm sorry if I missed that.
A. It's your time.
Q. Now, you state in your report at page 18 --
A. What page?
Q. Take a look at page 18, if you could.
A. Okay.
Q. You state at the top of page 18, "There is no valid basis for extrapolation of suspect race information from the small number of cases where offender case is known to the larger number of reported cases -- well, offender race is known to those cases where the suspect race is unknown, I think it's what you're saying, correct?
MR. HELLERMAN: Object to the form.
A. That's what it says.
Q. Tell us, in your own words, what that means?
MR. HELLERMAN: You just read his own words.
MR. LARKIN: All right, you know --
A. It means what it says. I'm not sure what part of it you need me to clarify.
Q. Why do you believe that there's no basis to extrapolate suspect race from the universe of cases where it's known to those where it's unknown?
A. If you know the suspect race in say one-third of

the cases, that leaves you two-thirds where you don't know. What would -- I mean, you know, I realize I'm not supposed to turn the tables, but I cannot imagine an algorithm that would allow me to accurately and confidently make any assumptions about what we know from those cases to the larger body of unknown cases, cases where suspect race is unknown.

There's no theory that would tell us that. If we did make such assumptions, the assumptions would be fraught with error. There's no way to confirm or disconfirm them. And if I tried to publish it, I would be relegated to a fifth-tier journal.

Q. Now, is there any other data that you looked at in this case that might inform the question what is the suspect description, the racial breakdown of suspect descriptions in cases where its unknown?

A. Are there any other data that we actually analyzed and reported?

Q. Any other data that you looked at in this case, that you saw?

A. I don't recall. No, I don't think so.

Q. Would arrest data, race-specific arrest data assist in that process?

A. Up to a point it could, but I'd have to analyze the data and see what the consistencies and inconsistencies

were of those data.
Q. What would you look at in the data? What would you do?
A. I'd want to look at a distribution by location, by month, by type of crime, by age, gender of the suspect, et cetera, et cetera.
Q. The city has a share of inmates in the upstate prison population, right?
A. Yes.
Q. Would the racial breakdown of the city's share of the prison population, the state prison population, inform the question of the racial breakdown of suspects in crimes?
MR. HELLERMAN: Object to the form.
A. In a minor and unreliable way.
Q. Why?
A. Minor because only a fraction of the persons arrested for crimes were sentenced to state prison. In an unreliable way because many of the variables that predict who goes to prison and who don't have nothing to do with the crime. It has to do with the quality of representation, the judge's preferences, statute.
Q. Is that true for violent crime?
A. Sure. Well, what do you mean by violence crime?
Q. Robbery, assault, murder --
A. Armed robbery or Robbery 3.

Q. Armed robbery -- let's stick to felonies. Armed robbery --
A. Robbery 3 is a felony, Counselor.
Q. Well, let's say --
MR. HELLERMAN: Let him ask his question.
Q. --- any armed robbery, a felony assault, murder, any shooting, those types of crimes, would you say that the upstate prison population, the racial breakdown of the upstate prison population is some reliable measure of the racial breakdown of crime?
A. It's more reliable than looking at the total inmate population, probably, but still it has weaknesses and is prone to error, for some of the same reasons I said before.
Q. Now, with respect to violent crimes, there is a greater proportion of reports for which we have a suspect description, right?
MR. HELLERMAN: Object to the form.
A. To the best of my recollection. There's a table actually somewhere and I --
Q. Just take a look at page 76 of your report, Table 18.
A. Right.
Q. Okay. You've divided crimes into *other* and *violent*, right?

A. Correct.
Q. And with regards to violent crimes for the year 2005, the suspect race is missing in just over 45 percent of the cases, right?
A. Correct.
Q. So you've got a suspect description for the majority of cases, true?
A. Yes. Technically, yes.
Q. For violent crimes in 2005, right?
A. Yes.
Q. And with respect to violent crimes in 2006, again, according to your table, we have a missing suspect description, that is we don't have a race for the suspect in 46.56 percent of the cases, right?
A. Correct.
Q. And that means you've got a suspect description for the majority of those cases, right?
A. Right, 54 --
Q. Plus?
A. Percent, give or take.
Q. About. Given the fact that you have a suspect description in the majority of crimes that you've categorized as violent, would you consider suspect descriptions to be some reliable indicator of the race-specific crime rate?

A. Let me explain why not. Let's assume that there's 45, nearly 46 percent of the cases where race is unknown in 2005. Let's assume under one set of conditions that all of those cases come from neighborhoods that have let's say 50 percent or more are black population, or where the racial composition of the victims was more than 60 percent black.

Under those conditions, I might be willing to consider the possibility that a disproportionate share or a proportionate share of the unknowns would match up to the victim race. And there's a simple answer, that seems to be the way the criminological data work out.

But let's assume that all of those cases come from places that are where the local residential population is highly variable. We would have no basis rationally for making an assignment of any particular case or a collection of cases to a particular racial category for the suspect.

So, in other words, when I say to you we need to know a lot more, we need to know things like where are these cases from, what are the crimes that are charged, what are the known probabilities, et cetera, et cetera.

Q. As part of your work in this case, did you look at the racial makeup of suspect descriptions on a precinct by precinct basis?

A. I don't recall doing that.

crime.
Q. It certainly would not be unreasonable for that officer to be thinking about a possible violent crime, would it?
MR. HELLERMAN: Objection.
A. I can't answer that.
Q. Forgive me if you've answered this question, and I apologize, Professor, but, would it have assisted you in your analysis to determine the suspect descriptions in the crime complaints where you found the highest racial disparities? Do you think it would have assisted you in your --
A. I'm sorry? Could you repeat the question.
Q. Sure. You had crime complaint data broken down by precinct, right?
A. Correct.
Q. And you had suspect description information for all those crime complaints, correct?
MR. HELLERMAN: Objection.
A. Well, we had suspect crime information that was missing in a fairly large number of the cases.
Q. So, to the extent that the police department had suspect descriptions, you had that information, also, right?
MR. HELLERMAN: Object to the form.

A. We knew the percentage of cases where they had a suspect description that was not missing or unknown.

Q. Fair enough. Okay. Would it have assisted you in any way in breaking down precinct by precinct the racial or ethnic background of the suspect descriptions to the extent it was known?

A. I can't speculate if it would have assisted us because we didn't do it.

Q. Did you make a decision not to do it or was it something you just didn't think to do, or something else?

A. I think we decided, given the high rate of missing and unknown data, that it wouldn't have been useful, so we didn't do it.

Q. Now, turning to the JASA study, once again.

Turn to page 815 of the study, if you would, Professor. Now, there's a paragraph on the left-hand column that begins *however;* do you see that?

A. Yup.

Q. And at the very end you've got a clause that reads as follows: "Few studies can control for all of the variables that police consider in deciding whether to stop or search someone." Do you see that?

A. Correct.

Q. What did you mean by that?

A. Well, specifically, in this case, we don't

Q. When you say *a lot,* is it more than 50 percent or less than 50 percent?
A. Less than 50.
Q. Would it be less than 25 percent for 2004?
A. At a minimum.
Q. You would agree, even today, that few studies can control for all the variables that police consider in deciding whether to stop or search someone, true?
A. All?
Q. Yes.
A. Sure, nobody can control for all of them.
Q. Do you believe that in connection with your work in this case you controlled for all of the relevant variables?
A. Yes.
Q. You state in the JASA study on page 815, that your approach there --
A. Where are you looking?
Q. I'm sorry. Looking at page 815, the column on the right, there's a paragraph -- the second full paragraph begins *we consider hit rates briefly;* do you see that?
A. Yes.
Q. And you state, as follows: "But our main analysis attempts to resolve these supply side or omitted variable problems by controlling for race-specific rates of

the targeted behaviors in patrolled areas, assessing whether stop and search rates exceed what we would predict from knowledge of the crime rates of different racial groups." Did I read that correctly?

A. Yes.

Q. What exactly does that mean? Can you tell me in your own words?

MR. HELLERMAN: Object to the form.

A. It means what it says.

Q. Did you undertake the same analysis in your work in this case?

A. No, we did not use race-specific rates of the targeted behaviors.

Q. Why not?

A. We used the overall rates. Because, as we've already discussed, we don't have a good measure of the race-specific participation in criminal behaviors other than the arrest data.

At the time of the JASA study, the arrest data was what was available to us. At the time of this study, we made a determination to use overall crime data because that was the most comprehensive measure of crime.

Q. At any time, did you use any model in connection with your work in this case based on arrest data?

A. Did we use a model?

MR. LARKIN: Withdraw the question.
Q. At any time, did you analyze data in connection with your work in this case using race-specific arrest rates?
MR. HELLERMAN: Object to the form.
A. No, we didn't, because we knew that the -- because the clearance rates were so variable from one place to the next, that it wouldn't have been a reliable measure.
Q. Now, further down in that paragraph you state that "this approach requires estimates of the supply of individuals engaged in the targeted behaviors." Do you see that?
A. Yup.
Q. As part of your work in this case, did you try to come up with an estimate of the supply of individuals who were engaged in the targeted behaviors?
MR. HELLERMAN: Object to the form.
A. We used a measure of the percentage of the total crime complaints that were specific to the model that we were estimating of rationale for the stops or reason for the stop.
Q. So you would come up with some estimate of the supply of persons engaged in the targeted behaviors precinct by precinct; is that fair?
A. Yes.

state as follows: "Since crime and race are correlated, indexing or benchmarking to crime should account for race simultaneously. Any significant effects for the racial composition of the area suggest racial disproportionality above and beyond any disproportionality that is explained by crime." Did I read that correctly?
A. Yes.
Q. Did indexing for crime satisfy your concern that there was a need for measures of the race-specific crime rates in each precinct, as you stated in your original report?
A. It helped.
Q. Do you believe that the controls for crime in your report were sufficient to account for the racial mix of crime, as well?
A. Within each precinct?
Q. Yes.
A. Precinct by precinct?
Q. Yes.
A. It helped.
Q. Do you think that the controls for crime were sufficient for your purposes in conducting the study?
A. I think it allowed us to make the inferences that we made.
Q. Why, in this case, did you not control for the

then I could take data -- I could re-run their date with the additional variables.

Q. Now, why did you include in the JASA study parameters for racial population composition and precinct effects?

A. I think I already answered that.

Q. Can you indulge my ignorance one last time.

MR. HELLERMAN: Objection. Asked and answered.

THE WITNESS: You want me to answer?

MR. HELLERMAN: One last time.

Q. Summarize it. I didn't mean to interrupt you, Professor. Go ahead. I'm sorry.

A. Because the enterprise that we're doing here is to try and estimate the stop patterns based on the available populations to be stopped, and also the parameters of crime that would shape police behavior, what they look for, and how aggressively they would look for it.

Q. On a precinct by precinct basis, right?

A. Correct.

Q. Does the RAND analysis, as it appears in the RAND study, reflect any precinct by precinct assessment of racial disparity?

A. Well, if the RAND people said that they used our data and our models, and so I assume that they used give the written detail results, other than like these guys.

For the second three rows, we think it's unreliable because of the -- as we said on several occasions, the failure to use -- well, the fact that they use violent crime suspect data which is missing in almost 80 percent of the cases.

Q. Now, I think you said Table 7 in your report replicates JASA; is that right?

A. I believe so, yes. Or attempts to replicate JASA. There are differences between Table 7 and this report and what the JASA analysis did.

Q. And is Table 7 a replication of that equation four in the JASA study, but using data produced in this case?

A. Yes.

Q. And does Table 7 reflect racial bias, in your view?

A. Yes. Well, let me answer, it reflects a disparate racial treatment. We're fairly careful about bias, as you know.

Q. And does Table 7 reflect statistically significant disparate racial treatment, in your opinion?

A. Oh, yes.

Q. And Table 7 covers stops for the entire time

exercise of their stop authority.
Did I read that sentence correctly?
A. Yes.
Q. Were you trying to come up with some estimate of the number of people in any precinct during any period of time who would be engaged in behavior that might arouse the suspicion of a police officer and who were available to be stopped?
A. We would like to have been able to come up with -- we did come up with an estimate of the people who were engaged in the behavior. That's the measure of crime, that's the criminal activity measure.
The people who were available to the police to be potential targets who were stopped, that's people that live in the place and who were on the street and available to be stopped.
Q. So you used the crime conditions as the sole basis to come up with the number of people, that is the supply of individuals, engaged in the targeted behaviors; do I have that right?
A. That was a proxy for that, yes.
Q. Did you, at any time in your study, account for people who were engaged in behavior that did not rise to the level of a crime, but that rose to the level of reasonable suspicion?

A. There were no such data available.

Q. And does the absence of such data impair or affect, in any way, the reliability of your conclusions?

A. I don't think so. Our conclusions, with respect to the 14th Amendment, are fairly straight forward. You know, setting aside the question of suspicion for the moment, the stop rate seems to be indexed over and above the crime rate to racial composition of people in the neighborhood; over and above the crime rate. So that accounts for the supply of individuals, because we include population, and it includes the level of criminal activity.

So one would assume that there a proportion of that population is engaged in criminal activity, and the racial composition of the neighborhood.

Q. Couldn't the stop rate also reflect not only the criminal activity, but the activity of individuals in that precinct who were engaged in behavior that arouses reasonable suspicion?

A. If you go back to Table 16, that's what we did. Not 16. I'm sorry. This one. If you go back to Table 13, we took the runs from Tables 5 and added into those runs measures of reasonable suspicion.

One assumes that when the police encounter people in a neighborhood who are engaged in suspicious behavior that the police made a stop. Maybe they had looked at

MR. HELLERMAN: Objection.
A. I won't even touch that question.
Q. I mean, crime happens in a variety of circumstances, at any time of day, it depends on the specifics of any given place, any given time, and the people who are present --
A. Is there a question, Counselor?
Q. Is that true?
MR. HELLERMAN: Objection.
A. Yeah, there's a big variety of crime, sure. They make for good TV shows.
Q. And an officer contemplating whether to stop a citizen has to take into account all the circumstances and all the information that he has, at that time, right?
MR. HELLERMAN: Objection.
A. The officer should be taking into account the indicia of reasonable suspicion in deciding whether or not to stop somebody.
Q. Now, how can you control for all the individualized circumstances that might give rise to reasonable suspicion in any sociological study?
MR. HELLERMAN: Objection.
A. I don't think that's pertinent to what our endeavor was about. We simply looked at the categories of reasonable suspicion as interpreted and implied by the

officers, how they classified and what categories they fell into based on our scheme.

Q. Who, if anyone, assisted you in writing your report and supplemental report, Professor?

A. My research assistants.

Q. And how many research assistants did you have?

A. Well, there were different assistants at different points in time. One was Amanda Geller. Another was Steven Clark. I had a series of research assistants who help developed the coding scheme for suspected crime. Erin Kelly, Edith Beerdsen, Garth Davies.

Q. Was there anyone else?

A. Not that I recall.

Q. And what role did Ms. Geller play?

A. She assisted in running the models.

Q. Did she assist in developing the models?

A. She worked with me in developing the models.

Q. And you've published, co-authored articles with Ms. Geller before, correct?

A. Correct.

Q. And she's a professor?

A. She's research scientist at Columbia University.

Q. And what role did Mr. Clark play?

A. He was a research assistant.

Q. What did he do for you?

Exhibit C

SIDE 1

J = Justified
CJ = Conditionally Justified
I = Indeterminate
AC = Additional Circumstances/ Factors boxes on side 2

Justified

- **Any J box on side 1**
- **Any CJ box on side 1 with AC box on side 2**

Unjustified

- **No CJ or J box on side 1**
- **AC box only on side 2**

Indeterminate

- **Any 1 CJ or combination of CJ boxes on side 1 with no AC box on side 2**
- **"I" box only on side 1 and no AC box on side 2**

(COMPLETE ALL CAPTIONS)

STOP, QUESTION AND FRISK REPORT WORKSHEET
PD344-151A (Rev. 11-02)

Pct.Serial No.
Date | Pct. Of Occ.
Time Of Stop | Period Of Observation Prior To Stop | Radio Run/Sprint #
Address/Intersection Or Cross Streets Of Stop
☐ Inside ☐ Outside | ☐ Transit ☐ Housing | Type Of Location Describe:
Specify Which Felony/P.L. Misdemeanor Suspected | Duration Of Stop

What Were Circumstances Which Led To Stop?
(MUST CHECK AT LEAST ONE BOX)

- **CJ ("Objects")** — ☐ Carrying Objects In Plain View Used In Commission Of Crime e.g., Slim Jim/Pry Bar, etc.
- **CJ ("Description")** — ☐ Fits Description.
- **J ("Casing")** — ☐ Actions Indicative Of "Casing" Victim Or Location.
- **CJ ("Lookout")** — ☐ Actions Indicative of Acting As A Lookout.
- **CJ ("Bulge")** — ☐ Suspicious Bulge/Object (Describe)
- **I ("Other")** — ☐ Other Reasonable Suspicion Of Criminal Activity (Specify)
- ☐ Actions Indicative Of Engaging In Drug Transaction. — **J ("Drugs")**
- ☐ Furtive Movements. — **CJ ("Furtive")**
- ☐ Actions Indicative Of Engaging In Violent Crimes. — **J ("Violent")**
- ☐ Wearing Clothes/Disguises Commonly Used In Commission Of Crime. — **CJ ("Clothing")**

Name Of Person Stopped | Nickname/ Street Name | Date Of Birth
Address | Apt. No. | Tel. No.
Identification: ☐ Verbal ☐ Photo I.D. ☐ Refused ☐ Other (Specify)
Sex: ☐ Male ☐ Female | Race: ☐ White ☐ Black ☐ White Hispanic ☐ Black Hispanic ☐ Asian/Pacific Islander ☐ American Indian/Alaskan Native
Age | Height | Weight | Hair | Eyes | Build
Other (Scars, Tattoos, Etc.)
Did Officer Explain Reason For Stop ☐ Yes ☐ No | If No, Explain:
Were Other Persons Stopped/ Questioned/Frisked? ☐ Yes ☐ No | If Yes, List Pct. Serial Nos.

If Physical Force Was Used, Indicate Type:
☐ Hands On Suspect
☐ Suspect On Ground
☐ Pointing Firearm At Suspect
☐ Handcuffing Suspect
☐ Suspect Against Wall/Car
☐ Drawing Firearm
☐ Baton
☐ Pepper Spray
☐ Other (Describe)

Was Suspect Arrested? ☐ Yes ☐ No | Offense | Arrest No.
Was Summons Issued? ☐ Yes ☐ No | Offense | Summons No.
Officer In Uniform? ☐ Yes ☐ No | If No, How Identified? ☐ Shield ☐ I.D. Card ☐ Verbal

SIDE 2

Was Person Frisked? ☐ Yes ☐ No **IF YES, MUST CHECK AT LEAST ONE BOX**

☐ Inappropriate Attire - Possibly Concealing Weapon
☐ Verbal Threats Of Violence By Suspect
☐ Knowledge Of Suspects Prior Criminal Violent Behavior/Use Of Force/Use Of Weapon
☐ Other Reasonable Suspicion of Weapons (Specify)
☐ Furtive Movements
☐ Actions Indicative Of Engaging In Violent Crimes
☐ Refusal To Comply With Officer's Direction(s) Leading To Reasonable Fear For Safety
☐ Violent Crime Suspected
☐ Suspicious Bulge/Object (Describe)

Was Person Searched? ☐ Yes ☐ No **IF YES, MUST CHECK AT LEAST ONE BOX** ☐ Hard Object ☐ Admission Of Weapons Possession ☐ Outline Of Weapon ☐ Other Reasonable Suspicion of Weapons (Specify)

Was Weapon Found? ☐ Yes ☐ No If Yes, Describe: ☐ Pistol/Revolver ☐ Rifle/Shotgun ☐ Assault Weapon ☐ Knife/Cutting Instrument ☐ Machine Gun ☐ Other (Describe)

Was Other Contraband Found? ☐ Yes ☐ No If Yes, Describe Contraband And Location __________
Demeanor Of Person After Being Stopped __________
Remarks Made By Person Stopped

AC/AF — **Additional Circumstances/Factors: (Check All That Apply)**

☐ Report From Victim/Witness
☐ Area Has High Incidence Of Reported Offense Of Type Under Investigation
☐ Time Of Day, Day Of Week, Season Corresponding To Reports Of Criminal Activity
☐ Suspect Is Associating With Persons Known For Their Criminal Activity
☐ Proximity To Crime Location
☐ Other (Describe)
☐ Evasive, False Or Inconsistent Response To Officer's Questions
☐ Changing Direction At Sight Of Officer/Flight
☐ Ongoing Investigations, e.g., Robbery Pattern
☐ Sights And Sounds Of Criminal Activity, e.g., Bloodstains, Ringing Alarms

Pct. Serial No. ____ | Additional Reports Prepared: Complaint Rpt.No. ____ Juvenile Rpt. No. ____ Aided Rpt. No. ____ Other Rpt. (Specify) ____

REPORTED BY: Rank, Name (Last, First, M.I.)
Print __________ Tax# ____
Signature __________ Command ____

REVIEWED BY: Rank, Name (Last, First, M.I.)
Print __________ Tax# ____
Signature __________ Command ____

Exhibit D

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK
--------------------------------------x
DAVID FLOYD, LALIT CLARKSON, DEON DENNIS and DAVID OURLICHT, individually and on behalf of a class of all other similarly situated,

Plaintiffs, Index No.

-against- 08 CIV 01034

THE CITY OF NEW YORK, NEW YORK CITY POLICE COMMISSIONER RAYMOND KELLY, in his individual and official capacity, et al,

Defendants.
--------------------------------------x

March 22, 2010
10:04 a.m.

DEPOSITION of MARY C. CRONIN, taken by the Plaintiffs, pursuant to Notice, at the law offices of BELDOCK, LEVINE & HOFFMAN, LLP, 99 Park Avenue, New York, New York before Karen Perlman, a Shorthand Reporter and Notary Public within and for the State of New York.

GREENHOUSE REPORTING, INC.
875 Sixth Avenue - Suite 1716
New York, New York 10001
(212) 279-5108

M. Cronin

Q. And I'll read now, the top of page 6, it says: QAD shall conduct audits. At a minimum, address the following issues (A) whether and to what extent documents, i.e., UF-250s officer activity logs that have been filled out by officers to record stop, question and frisk activity have been completed in accordance with the NYPD regulations. And (B) whether and to what extent the audited stop, question and frisk activity is based upon reasonable suspicion as reflected in the UF-250 forms.

Do you see that?

A. Yes.

Q. Is it your understanding -- or I'm sorry. Does Worksheet 802 and self-inspections and audits that are based on Worksheet 802, do those address both of these issues in paragraphs A and B?

A. I believe so.

Q. Can you tell me how Worksheet 802, the self-inspection, addresses subparagraph B, whether and to what extent stop, question and frisk activity is based upon reasonable suspicion

M. Cronin

as reflected in the UF-250 forms?

A. That might be more of an explanation.

Q. Okay.

A. But we go into the commands. We take a UF-250 form that's been completed. The teams look at the form, and using the worksheet, we go down line by line, box by box, making sure things are checked off.

When you check the things off, the boxes, you have circumstances which led to the arrest. The circumstances which led to the arrest has to be based on probable cause.

Q. Well, first of all, let me back up --

A. I'm sorry. Reasonable suspicion. Forgive me.

Q. So you look at the form, and there will be one or more circumstances checked off, correct?

A. At times, yes.

Q. And so how does the reviewer from QAD or the person in the command when they're doing a self-inspection, how did they determine

M. Cronin

whether or not circumstances that are checked off in fact are accurate and were in fact the real circumstances that led to the officer to make the stop?

A. Well, we have to base it on the form that we're looking at. We're not physically there. But if the form is filled out correctly and everything is checked off, and everything that you checked off makes sense -- if you check off something, for example, the circumstances behind the stop and you put in the reason for the stop was regarding an open container, that would immediately send off -- set off bells for my people that an open container is not a reason -- reasonable suspicion to stop someone for a 250.

Q. So you're saying that one of the things that is looked for is whether or not the circumstances that are checked off correspond to the suspected crime that's listed on the --

A. Yes.

Q. -- form?

Now, other than reviewing the UF-250 form itself, and I guess also the activity logs of the officers, do the QAD evaluators or the

Exhibit E

Impact IX

Monday, July 09, 2007

BORO	Zone	Pct Zone Commander	Boundaries	Hours	RDO"s	TOTAL MOS ASSIGNED
PBMS	1	14 Capt.Berntsen	W.39 ST - W. 45 ST 7 Ave - 8 Ave	1200X2035 2000X0435	7 DAYS	72
PBMN	2	32 Capt Ehrenberg	W. 133 St. - W. 145 St Adam Clayton Powell Blvd - Lenox Ave	1730X0205	7 DAYS	48
PBMN	MN IRT	Impact Response Team Capt. Pla	26 / 30 Pct IRT "A" Broadway, W 135 St- W. 152 St W. 145 S, St. Nicolas to Broadwayt	1730X0205	7 DAYS	36
			25 Pct IRT "B" E 116 St-E. 126 St, Lexington Ave to 3 Ave	1730X0205	7 DAYS	36
PBBX	3	44 Capt. Melendez	Area bound by the following perimeter: E. 161 St to Jerome, Jerome to XBronx Expway, XBronx Expway to Grand Conc. Grand Conc to E. 169 St, E 169 St to Webster Ave, Webster to E 166 St., E 166 St to Grand Conc., Grand Conc to E 161 St	1730X0205	7 DAYS	48
PBBX	4	46 Capt.McHugh	Area bound by the following perimeter: XBronx to E 177 St, Dr MLK Jr Blvd to Jerome Ave Area bound by the following perimeter: Originating at Jerome Avenue and E Burnside Ave, to E183 St., to Grand Concourse, to E Fordham Rd., to Webster Ave, to E183 St. to Grand Concourse, to E Burnside Ave, to Jerome Ave	1730X0205	7 DAYS	48
PBBX	5	52 Capt. Corrado	E. Fordham Rd from University Ave to Decatur Ave Creston Ave to Decatur Ave, W. Kingsbridge Rd to E. Fordham Road	1730X0205	7 DAYS	48
PBBS	6	70 Capt. Mastrokostas	Area bound by the following perimeter: Originating at Ocean Ave to Clarkson Ave to Bedford Ave to Linden Blvd to Flatbush Ave to Newkirk Ave to E 21 St to Albemarle Rd to E 18 St to Church Ave, Church Ave to C.I. Ave, C.I. Ave to Caton Ave, Caton Ave E. 21 St SEE MAP	1030x1905 1800x0235	7 DAYS	70
PBBS	7	71 Capt. DiPaolo	Ocean Ave to Flatbush Ave, Parkside to Empire Blvd, with an extension on Empire Blvd from Flatbush Ave to Bedford Ave	0930x1805 1730x0205	7 DAYS	36
PBBN	MEGA	73 Capt. Tasso	Mega Zone Entire Precinct	1200X2035 1930X0405	7 DAYS	45
PBBN	MEGA	75 Zone 1- Capt.Kelly Zone 2- Capt.Farrell Zone 3 - Capt.Schweitzer	Mega Zone Entire Precinct	1200X2035 1930X0405	7 DAYS	60
PBBN	BN IRT	Impact Response Team Capt. Patti	Area bound by the following perimeter. Herkimer St to Nostrand Ave to Macon St to Throop Ave to Monroe St to Malcolm X to Patchen to MacDonugh St to Buffalo to Atlantic to Rochester to Nostrand Ave SEE MAP	1200X2035 1930X0405	7 DAYS	72
PBQS	8	103 Capt. Barrett	Jamaica Ave 153 St - Merrick Blvd Archer Ave-Jamaica Ave, Parsons Blvd-161 St	1155X2030 1430X2305	7 DAYS	48
PBQN	9	110115 Capt.Tamola	Roosevelt Ave to 37th Ave from 72 St to 104 St	2130X0605 1730X0205	7 DAYS	72

(REV. JUNE 20, 2007)

9 PSB Impact Zones
2 Mega Zones
2 Impact Response Teams
Housing PSA's 2,5,7
Transit

PSB Total	739
Housing	100
Transit	100
CW Total	939

Sgt Holohan, P.S.B.

Exhibit F

ORIGINAL

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK
---X
DAVID FLOYD, LALIT CLARKSON, DEON DENNIS, and DAVID OURLICHT, individually and on behalf of a class of all other similarly situated,

PLAINTIFFS,

-against-

Index No: 08 CIV 01034

THE CITY OF NEW YORK, NEW YORK CITY POLICE COMMISSIONER RAYMOND KELLY, in his individual and official capacity, et al,

DEFENDANTS.
---X

DATE: August 6, 2009

TIME: 10:42 a.m.

EXAMINATION BEFORE TRIAL of the Defendants, by CHIEF ROBERT GIANNELLI, taken by the Plaintiffs, pursuant to Stipulation and the Federal Rules of Civil Procedure, held at One Police Plaza, New York, New York 10038, before Scott Torrance, a Notary Public of the State of New York.

CHIEF R. GIANNELLI

same ratio of sergeant to police officer?

A. Yes, we have a designated ratio. We're very conscience because they're younger officers that we want to make sure that they're supervised properly.

Q. So then, let's move on to the impact zone officers for a second. I want to make sure I know how that works.

So, some officers, when they leave the academy, are assigned to impact zones, and you said impact zones can be certain areas of a given precinct; is that right?

A. Yes. Occasionally they may go over the boundaries and encompass two precincts, areas in two precincts.

Q. So then, each -- let me make sure I use the right terminology.

Each officer is assigned to a -- is it a squad within a specific impact zone?

A. Well, you can use that term "squad", but it would not mean the same as a squad in a precinct, because they may

Exhibit G

U.S. Census. 2000. Available: http://www.census.gov/.

Van Maanen, John. 1974. A developmental view of police behavior. In H. Jacob (Ed.), *The Potential for Reform of Criminal Justice*. Beverly Hills, CA: Sage Publications.

Vrij, Aldert, August Dragt, and Leendert Koppelaar. 1992. Interviews with ethnic interviewees: Nonverbal communication errors in impression formation. *Journal of Community and Applied Social Psychology* 2: 199–208.

Vrij, Aldert and Frans Winkel. 1994. Perceptual distortions in cross-cultural interrogations: The impact of skin color, accent, speech, style, and spoken fluency on impression formation. *Journal of Cross-Cultural Psychology* 25: 284–295.

Vrij, Aldert and Frans Winkel. 1991. Cultural patterns in Dutch and Surinam nonverbal behavior: An analysis of simulated police citizen encounters. *Journal of Nonverbal Behavior* 5: 169–184.

Walker, Samuel. 2001. Searching for the denominator: Problems with police traffic stop data and an early warning system solution. *Justice Research and Policy* 3: 63–95.

Walker, Samuel. 1993. *Taming the system: The control of discretion in Criminal Justice, 1950–1990*. New York: Oxford University Press.

Westley, William. 1953. Violence and the police. *American Journal of Sociology* 59: 34–41.

Winkel, Frans and Aldert Vrij. 1990. Interaction and impression formation in a cross-cultural dyad: Frequency and meaning of culturally determined gaze behavior in a police interview setting. *Social Behavior* 5: 335–350.

Worden, Robert. 1989. Situational and attitudinal explanations of police behavior: A theoretical reappraisal and empirical assessment. *Law and Society Review* 23: 667–771.

Worden, Robert and Robin Shepard. 1996. Demeanor, crime, and police behavior: A reexamination of the police services study data. *Criminology* 34: 83–105.

Worden, Robert Shepard, and Stephen Mastrofski. 1996. On the meaning and measurement of suspect's demeanor toward the police: A comment on "demeanor and arrest." *Journal of Research in Crime and Delinquency* 33: 324–332.

CASES CITED

State of New Jersey v. Soto et al., 734 A. 2d 350 (N.J. Super. 1996).

Wilkins v. Maryland State Police et al., Civ. No. MJG-93-468 (D.Md. 1993).

Chapter 13

Street Stops and Broken Windows Revisited

The Demography and Logic of Proactive Policing in a Safe and Changing City

Jeffrey A. Fagan, Amanda Geller, Garth Davies, and Valerie West

I. Introduction

The role of policing in New York City's crime decline has been the subject of contentious debate for well over a decade. Violent crime reached its modern peak in New York City in 1991, followed by a 10 percent decline in 1992–93 (Fagan, Zimring, and Kim, 1998). This initial crime decline was spurred by the hiring and quick deployment in 1991 of five thousand additional officers under the Safe Streets Program (McCall, 1997; Greene, 1999; Waldeck, 2000; Karmen, 2000). During this initial decline, police tactics remained largely unchanged from the preceding years. Following the mayoral election in 1993, newly appointed police commissioner William Bratton implemented a regime of "order-maintenance policing" (OMP), which—together with other management reforms and innovations—dramatically and suddenly changed both the strategy and tactics of policing across the City. The new strategy was grounded in *Broken Windows* theory (Wilson and Kelling, 1982; Kelling and Cole, 1996) and focused on the connection between physical and social disorder and violence (Greene, 1999; Livingston, 1997; Spitzer, 1999; Sampson and Raudenbush, 1999; Duneier, 1999; Waldeck, 2000; Fagan and Davies, 2000; Taylor, 2001; Harcourt, 2001).

In the new policing model, police tactics, resources, and attention were redirected toward removal of visible signs of social disorder—"broken windows"—by using police resources both for vigorous enforcement of laws on minor "quality of life" offenses, while aggressively interdicting citizens in an intensive and widespread search for weapons (Kelling and Cole, 1996; Bratton and Knobler, 1998; Silverman, 1999). Tactically, policing in this era had several faces, from frequent arrests for low-level crimes such as public drinking, graffiti, and marijuana possession (Golub, Johnson, and Dunlap, 2007; Harcourt and Ludwig, 2007; Levine and Small, 2008), to aggressive street-level interdictions and searches of citizens whose behaviors signaled their potential for any of several types of crime, but most notably carrying weapons (Harcourt, 1998; Fagan and Davies, 2000; Gelman, Fagan, and Kiss, 2007). Using aggressive "stop and frisk" tactics, this brand of OMP was designed to reduce violence and weapons possession (Spitzer, 1999; Waldeck, 2000; Fagan and Davies, 2000; Harcourt, 2001).

The origins of the tactical shift are revealed in strategy documents issued by the New York City Police Department (NYPD) in 1994. First, Police Strategy No. 5, *Reclaiming the Public Spaces of New York*, articulated a reconstructed version of *Broken Windows* theory (Wilson and Kelling, 1982) as the driving force in the development of policing policy. It stated that the NYPD would apply its enforcement efforts to "reclaim the streets" by systematically and aggressively enforcing laws against low-level *social* disorder: graffiti, aggressive panhandling, fare beating, public drunkenness, unlicensed vending, public drinking, public urination, and other misdemeanor offenses. Second, Police Strategy No. 1, *Getting Guns Off the Streets of New York*, formalized the strategic focus on the eradication of gun violence through the tactical measure of intensifying efforts to seize illegal firearms. Homicide trends in New York City since 1985 provided strong empirical support for emphasizing gun violence in enforcement policy (Davis and Matea-Gelabert, 1999). Nearly all the increases in homicides, robberies, and assaults from 1985 to 1991 were attributable to gun violence (Fagan et al., 1998). The homicide crisis was a critical theme in the mayoral election campaign of 1993, and focused the attention of the incoming Giuliani administration's crime-control policy on gun violence (Silverman, 1999).

By the end of the decade, stops and frisks of persons suspected of crimes had become a flashpoint for grievances by the City's minority communities, who came under the closest surveillance of the police and were most often stopped and frisked (Spitzer, 1999; Kocieniewski, 1999; Roane, 1999; Jackson, 2000). In a fifteen-month period from January 1998 through March 1999, non-Hispanic Black, Hispanic Black, and Hispanic White New Yorkers were three times more likely than their White counterparts to be stopped and frisked on suspicion of weapons or violent crimes relative to each group's participation in each of those two types of crimes (Gelman et al., 2007). These excess stops—stops beyond the rate that one would predict from the race-specific crime rates—could be explained neither by the crime rates in those areas in the City's poorest areas, nor by signs and manifestations of social disorder, nor by the presence of physical disorder in the form of actual "broken windows" or building or neighborhood decay. Instead, Fagan and Davies (2000) reported that policing was disproportionately concentrated in the City's poorest neighborhoods with the highest concentrations of minority citizens, even after controlling for rates of crime and physical disorder in those places (see also Gelman et al., 2007).

Despite its racial disproportionality, the harsh spotlight of a federal court order enjoining the NYPD from racially selective enforcement (*Daniels et al. v. City of New York*, 2003), and arrest rates of less than 15 percent resulting from stops (Spitzer, 1999; Gelman et al., 2007), the OMP policy continued far into the next decade (Baker, 2009). Yet New York City had changed dramatically during this period, even after rates of crime and disorder had fallen. Housing prices had soared for more than a decade in all neighborhoods, including those that had the highest violence rates in the preceding decade (Fagan and Davies, 2007), and new housing replaced abandoned lots and decaying buildings across the City (Schwartz, 1999). Welfare rolls thinned, the number of immigrants landing in the City's poorest neighborhoods rose sharply, and populations of African Americans declined by more than 10 percent (U.S. Census Bureau, 2006). With minor and random ticks up and down, crime remained nearly flat and low since 2000 (Levine and Small, 2008).

Yet, in a safe and thriving city, the number of citizen stops grew by 500 percent between 2003 and 2008 (Baker, 2008, 2009; Ridgeway, 2007), long after crime had precipitously declined to and remained at historic lows. The efficiency of these stops—that is, the rate at which crime was detected leading to an arrest—declined from about 15 percent in 1998–99 (Gelman et al., 2007) to 7.8 percent in 2003 to less than 4.1 percent in 2006 (table 13.1 infra; Ridgeway, 2007).

As we show in this chapter, street stops continue to be disproportionately concentrated in the City's poorest areas, not unlike a decade earlier. The logic of a sharp rise in street stops and a corresponding sharp decline in their efficiency, in an era of flat crime rates, demands analysis and explanation. In this chapter, we examine the exponential rise in street stops in an era of stable crime rates and look to the community contexts of these stops to identify the predictors of stops and their outcomes.

The everyday routines of New Yorkers of different ethnic and racial groups take place in vastly different local contexts, and it is in these contexts that the heterogeneity and disparate impact of policing practices are most observable. Accordingly, we identify local area characteristics of crime, disorder, and social structure that predict race-specific police stop activity. We extend the work of Fagan and Davies (2000) from 1999 to two time periods in the current decade, across an extended era of declining and then stably low crime rates. We find that the dramatic increase in stop activity in recent years is concentrated predominantly in minority neighborhoods, and that minority residents are likely to be disproportionately subjected to law enforcement contact based on the neighborhoods in which they live rather than the crime problems in those areas. Moreover, this disproportionate contact is based on more than the level of neighborhood crime and disorder; demographic makeup predicts stop activity above and beyond what local crime conditions suggest is necessary and justifiable.

We also test the efficiency of street stops to detect wrongdoing and sanction offenders, and find it to be low and declining over time: as stops have become more prevalent in recent years, they are substantially less likely to lead to arrests. These limitations are particularly pronounced in neighborhoods with high Black populations, suggesting that Black citizens are not only at an elevated risk of police contact compared to non-Hispanic Whites and Hispanics, but that the standards used to justify stops in their neighborhoods may be lower than those in neighborhoods with higher White populations. Finally, we examine and compare specific age-race-cohort impacts of policing to illustrate the extraordinary concentration of policing along racial and ethnic lines.

Our analysis begins with a brief history of the constitutional and theoretical frameworks for New York's OMP strategy, with attention to the racial dimensions of modern policing. We then discuss the data, models, and results, followed by discussion and conclusions.

II. Background

A. Race, Neighborhoods, and Police Stops

Nearly a century of legal and social trends set the stage for the current debate on race and policing. Historically, close surveillance by police has been a part of everyday life for African Americans and other minority groups (see, for example, Musto, 1973; Kennedy, 1997; Cole, 1999; Loury, 2002; Weitzer and Tuch, 2006). In recent decades, the U.S. Supreme Court has sanctioned border interdictions of persons of Mexican or Hispanic ethnicity to halt illegal immigration (*U.S. v. Martinez-Fuerte*, 1976), as well as the racial components of drug courier profiling by airlines (*U.S. v. Harvey*, 1992). In *U.S. v. Whren* (1996), the Supreme Court allowed the use of race as a basis for a police stop as long as there were other factors that motivated the stop, and in *Brown v. Oneonta* (2000), a federal district court permitted the use of race as a search criterion if there was an explicit racial description of the suspect.

The legal standard to regulate the constitutionality of police conduct in citizen stops derives from *Terry v. Ohio* (1968), which involved a pedestrian stop that established the parameters of the "reasonable suspicion" standard for police conduct in detaining citizens for purposes of search or arrest. Recently, the courts have expanded the concept of "reasonable suspicion" to include location as well as the individual's behavior. In fact, the Court has articulated and refined this "high-crime area" doctrine, in cases from *Adams v. Williams* (1972) to *Illinois v. Wardlow* (2000). This line of cases allows police to consider the character of a neighborhood as a factor justifying a standard lower than the constitutionally defined threshold in individualized "reasonable" suspicion articulated in *Terry v. Ohio* (1968) (Ferguson and Bernache, 2008). For example, in *Wardlow*, the Supreme Court noted that although an individual's presence in a "high-crime area" does not meet the standard for a particularized suspicion of criminal activity, a location's characteristics are relevant to determining whether a behavior is sufficiently suspicious to warrant further investigation. Since "high-crime areas" and social disadvantage often are conflated both perceptually and statistically with concentrations of minority citizens (Massey and Denton, 1993; Sampson and Lauritsen, 1994; Loury, 2002; Fagan, 2008; Sampson and Raudenbush, 1999, 2004; Alpert et al., 2005; Ferguson and Bernache, 2008; Massey, 2007), this logic places minority neighborhoods at risk for elevating the suspiciousness of their residents in the eyes of the police.

But in connecting race and policing, the Court was only formalizing what criminologists had known for decades. Early studies on police selection of citizens for stops suggested that both the racial characteristics of the suspect and the racial composition of the suspect's neighborhood influence police decisions to stop, search, or arrest a suspect (Reiss, 1971, Bittner, 1970). Particularly in urban areas, suspect race interacts with neighborhood characteristics to animate the formation of suspicion among police officers (Smith, 1986; Thompson, 1999; Smith et al., 2006). For example, Alpert and colleagues (2005) showed that police are more likely to view a minority citizen as suspicious—leading to a police stop—based on nonbehavioral cues while relying on behavioral cues to develop suspicion for White citizens.

Individuals—including police and political leaders—also may substitute racial characteristics of communities for racial characteristics of individuals in their cognitive schema of suspicion, and, more important, act on them. Quillan and Pager (2001) find that urban residents' perceptions of crime in their neighborhoods are significantly predicted by the prevalence of young Black men, even after crime levels and other neighborhood characteristics are controlled for. Police perceptions may be similarly skewed, resulting in elevated stop rates in neighborhoods with high concentrations of minority populations, and the pathway is through the translation of perceptions into neighborhood stigma. For example, in a study of police practices in three cities, Smith (1986) showed that suspects in poor neighborhoods were more likely to be arrested, after controlling for suspect behavior and the type of crime. Suspects' race and the racial composition of the suspect's neighborhood were also significant predictors of police response. It seems that social psychological mechanisms interact with cultural processes (patterns of behavior) and structural features of neighborhoods (poverty, concentrations of minority citizens) to produce perceptions of disorder that perpetuate urban inequality (Sampson and Raudenbush, 2004) through several forms of discrimination, including policing intensity and tactics (Fagan and Davies, 2000). Recall that Fagan and Davies showed that street stops in New York were predicted not by disorder but by race and poverty, despite policing theories that emphasized disorder as a pathway to elevated crime. Poor neighborhoods are stigmatized in this way, and people both within these areas as well as those who reside elsewhere—including those with administrative authority to withhold or allocate various services—are likely to act on their perceptions.

Alternatively, these coercive police responses may relate to the perception that poor neighborhoods may have limited capacity for social control and self-regulation. This strategy was formalized in the influential "broken windows" essay of Wilson and Kelling (1982). They argued that police responses to disorder were critical to communicate intolerance for crime and to halt its contagious spread. *Broken Windows* called for the targeting of police resources to neighborhoods where public order was deteriorating, with the expectation that stopping disorderly behavior would stem the "developmental sequence" to more serious crime. In the original essay, Wilson and Kelling worried about "criminal invasion" of disorderly neighborhoods. Neighborhood disorder has explicitly been used as a criterion for allocating police resources in New York City since 1994, when commissioner William Bratton set policies to focus on minor offenses such as subway fare evasion and aggressive panhandling, in addition to felonies and other serious crime (Kelling and Cole, 1996). The policy also called for aggressive responses to social disorder that was endogenous to neighborhoods, in contrast to the "criminal invasion" concern in the theory's pristine form.

This order-maintenance approach also has been disputed, however, as critics question the causal link between disorder and more serious crime (compare Harcourt, 1998, 2001; Sampson and Raudenbush, 1999, 2004; and Taylor, 2001; with Skogan, 1990; Corman and Mocan, 2000; Rosenfeld, Fornango, and Rengifo, 2007). Moreover, these studies suggest that a focus on disorder might have a disparate impact on citizens of different races. A study of Chicago neighborhoods finds that city residents' perceptions of disorder conflate systematically observable conditions with their neighborhoods' racial and socioeconomic makeup (Sampson and Raudenbush, 2004). The association between race, poverty, and perceived disorder is significant in residents of all racial and ethnic backgrounds; race and concentrated poverty predict both residents' and outsiders' perceptions of disorder even more strongly than does systematically observed disorder. And the effect grows stronger as the concentration of poverty and minority groups increase.

So the concentration of "order maintenance" policing in poor places with high concentrations of poor residents should come as no surprise: order-maintenance policing strategies ostensibly targeted at "disorderly" neighborhoods were in fact focused on minority neighborhoods, characterized by social and economic disadvantage (Fagan and Davies, 2000). This racial bait and switch with disorder is fundamental to understanding the broad spatial and social patterns of policing in New York in the past decade. Most interesting and important is the persistence of these policies even as the objective indicia of poverty and disorder fade in what we show below is a steadily improving and safe City.

B. Approaches to Studying Police Stops

Recent empirical evidence on police stops supports perceptions among minority citizens that police disproportionately stop African American and Hispanic motorists, and that once stopped, these citizens are more likely to be searched or arrested (Cole, 1999; Veneiro and Zoubeck, 1999; Harris, 1999; Zingraff et al., 2000; Gross and Barnes, 2002; Weitzer and Tuch, 2006; Ayres, 2008). For example, two surveys with nationwide probability samples, completed in 1999 and in 2002, showed that African Americans were far more likely than other Americans to report being stopped on the highways by police (Langan et al., 2001; Durose et al., 2005). Both surveys showed that minority drivers also were more likely to report being ticketed, arrested, handcuffed, or searched by police, and that they more often were threatened with force or had force used against them. These disparities in stop rates exact high social costs that, according to Loury (2002), animate culturally meaningful forms of stigma that reinforce racial inequalities, especially in the practice of law enforcement. These stigma translate into withdrawal of minority populations from cooperation with the police and other legal authorities in the coproduction of security (Tyler and Huo, 2002; Tyler and Fagan, 2008).

Traffic violations often serve as the rationale or pretext for stops of motorists (Walker, 2001; Harris, 2002), just as "suspicious behavior" is the spark for both pedestrian and traffic stops (Alpert et al., 2005; Ayres, 2008). As with traffic violations, the range of suspicious behaviors is broad enough to challenge efforts to identify an appropriate baseline against which to compare race-specific stop rates (see Miller, 2000; Smith and Alpert, 2002; Gould and Mastrofski, 2004). Pedestrian stops are at the very core of policing, used to enforce narcotics and weapons laws, to identify fugitives or other persons for whom warrants may be outstanding, to investigate reported crimes and "suspicious" behavior, and to improve community quality of life. For the NYPD, a "stop" provides an occasion for the police to have contact with persons presumably

involved in low-level criminality without having to effect a formal arrest, and under the lower constitutional standard of "reasonable suspicion" (Spitzer, 1999). Indeed, because low-level "quality of life" and misdemeanor offenses were more likely to be committed in the open, the "reasonable suspicion" standard is more easily satisfied in these sorts of crimes (Rudovsky, 2001, 2007).

Two distinct approaches characterize recent efforts to model and understand racial disparities in police stops. Each focuses less on identifying racial bias than on understanding the role of race in explaining patterns of police behavior. Attributing bias is difficult: causal claims about discrimination would require far more information than the typical administrative (observational) data sets can supply. For example, when Officer McFadden stopped suspect Terry in the events leading to the landmark 1968 U.S. Supreme Court decision in *Terry v. Ohio*, he used his law enforcement "experience" to interpret Terry's behavior in front of the jewelry store.[1] Were McFadden's notions of "suspicious" behavior skewed by his longtime work in poor and minority neighborhoods? Was the timing of the event (shortly after the closing of the store) or the location (a deserted part of the downtown area) influential? What role did Terry's and McFadden's race play? Would Terry's actions have been interpreted differently if he were White? If McFadden were Black? If the store was in a residential neighborhood instead of downtown? In a minority neighborhood or a predominantly White one? The multiplicity of interacting factors complicated the identification of the role of race in the decision to detain Terry (Kennedy, 1997), but several analyses of the facts and jurisprudence of *Terry* suggest that the Supreme Court opinion discounted the influence of race in the opinion (Thompson, 1999; Carbado, 2002; Carbado and Gulati, 2000; Roberts, 1999; Rudovsky, 2007).

In *Terry*, it would be difficult to identify race alone, apart from the context in which race was observed, as the factor that animated McFadden's decision to stop and frisk suspect Terry. Instead, reliable evidence of ethnic or racial bias in these instances would require experimental designs that control for these competing and interacting factors—situational context, demeanor of suspect—so as to isolate differences in outcomes that could only be attributed to race or ethnicity. Such experiments are routinely used in tests of discrimination in housing and employment (see, for example, Pager, 2003, 2007; Thacher, 2008). But observational studies that lack such controls are often embarrassed by omitted variable biases: few studies can control for all the variables that police consider in deciding whether to stop or search someone, much less their several combinations or permutations. Research in situ that relies on direct observation of police behavior (e.g., Gould and Mastrofski, 2004; Alpert et al., 2005) requires officers to articulate the reasons for their actions, a task that is vulnerable to numerous validity threats. Sampling considerations, as well as the presence of the researchers in the context of the decision, also challenge the validity of observational studies.

The first approach to studying racial disparities bypasses the question of whether police intend to discriminate on the basis of ethnicity or race, and instead focuses on disparate impacts of police stop strategies. This strategy is prevalent in studies of decisions in the context of highways stops. In this approach, comparisons of "hit rates," or efficiencies in the proportion of stops that yield positive results, serve as evidence of disparate impacts of police stops. This type of analysis has been used in several studies, including Knowles, Persico, and Todd (2001); Ayres (2002a,b); Gross and Barnes (2002); and many other studies of police behaviors on highways (see, e.g., Durlauf, 2006b). This approach bypasses the supply-side question of who is stopped (and for what reason), and instead looks only at disparate impacts or outcomes for different groups.

Outcome tests are agnostic with respect to race-based motivations for stops or frisks versus a search for efficiency and deterrence (Ayres, 2002b; Dominitz and Knowles, 2006). They can show when a particular policy or decision-making outcome has a disparate impact whose racial disproportionality is not justified by heightened institutional productivity. In the context of profiling, outcome tests assume that the ex post probability that a police search will uncover drugs or other contraband is a function of the degree of probable cause that police use in deciding to stop and search a suspect (Ayres, 2002a). If searches of minorities are less productive than searches of Whites, this could be evidence that police have a lower threshold of probable cause when searching minorities. At the very least, it is a sign of differential treatment of minorities that in turn produces a disparate impact.

Knowles, Persico, and Todd (2001) consider this "hit rate" approach theoretically as well as empirically in a study finding that, of the drivers on Interstate 95 in Maryland stopped by police on suspicion of drug trafficking, African Americans were as likely as Whites to have drugs in their cars. Their theoretical analysis posits a dynamic process that considers the behaviors of police and citizens of different races, and integrates their decisions in equilibrium where police calibrate their behavior to the probabilities of detecting illegal behavior, and citizens in different racial groups adjust their propensities to accommodate the likelihood of detection. They concluded that the search for drugs was an efficient allocation of police resources, despite the disparate impacts of these stops on minority citizens (Lamberth, 1997; Ayres, 2002a; Gross and Barnes, 2002; but see Sanga, 2009, for different conclusions).

Outcome tests can be constructed as quasi experiments, with race as a treatment, to identify the role of race in the selection of citizens for searches. Ridgeway (2007) matched suspects within officers to compare the post-stop outcomes of White suspects to those of minority suspects in similar locations, stopped at similar times and for the same reasons. He reports no differences in post-stop arrests ("hit rates") despite the greater number of stops of non-Whites. But this approach seeks to explain away contextual variables, especially neighborhood context, rather than explicitly incorporate these factors in an identification strategy. Close and Mason (2007) construct a disparate outcome quasi experiment to identify the role of race in police searches by comparing the preferences of officers of different races to search motorists, controlling for the motorist's race. They use both an outcomes-based nonparametric (quasi-experimental) analysis and a standard benchmarking parametric (regression) approach, and report both personal biases and police cultural bias in their propensity to search African American and Latino drivers.

These are useful but limited strategies. The robustness of these designs is compromised by the omission of several factors—some unobservable and others usually absent from administrative data—that might bias their claims, such as racial differences in the attributes that police consider when deciding which motorists or pedestrians to stop, search, or arrest (see, for example, Alpert et al., 2005; Smith et al., 2006), or differences in police behavior in neighborhoods or other social contexts with different racial makeup (Smith, 1986; Fagan and Davies, 2000; Alpert et al., 2005). For example, Ridgeway (2007) estimated the racial proportionality of police stops of citizens based on victim reports of suspect race. This is a sound strategy, but only for the approximately 20 percent of stops based on a rationale of "fits suspect description" (see, for example, Spitzer, 1999), and only if we are confident in the accuracy of victim identification of the suspect(s) and the accompanying classification of race.[2]

The omission of neighborhood context also biases estimates of the proportionality of police stops of citizens. The randomizing equilibrium assumptions in the Persico and colleagues approach—that both police and potential offenders adjust their behavior in response to the joint probabilities of carrying contraband and being stopped—tend to average across broad heterogeneous conditions both in police decision making and offenders' propensities to crime (Dharmapala and Ross, 2004; Durlauf, 2006a, 2006b), and discount the effects of race-specific sensitivities toward crime decisions under varying conditions of detection risk via police stop (Alpert et al., 2005; Dominitz and Knowles, 2006). When these two concerns are addressed, Dharmapala and Ross (2004) identify different types of equilibria that lead to different conclusions about racial prejudice in police stops and searches.

Accordingly, the nature and extent of racial bias in the policing of motorists and pedestrians remains unsettled empirically (Persico and Todd, 2005; Antonovics and Knight, 2004; Bjerk, 2007; Donohue and Levitt, 2001; Close and Mason, 2007). Supply-side issues, both in the number and characteristics of the persons available for stops by virtue of law violation or even suspicious behavior, complicate the search game paradigm by perceptually skewing the population of stopped drivers according to the *ex ante* probabilities of criminality that police officers assign to different racial groups. Institutional or individual differences in the goals of law enforcement may also create heterogeneity both in the selection of individuals to be stopped and the decisions to engage them in searches for drugs, weapons, or other contraband. Officers may pursue one set of law enforcement goals for one group (maximizing arrests) while pursuing a different set of goals (minimizing crime) for another. Racial nepotism or antagonism may lead to differences in police stop-and-search behaviors when officers of one race face choices of whether to stop or search a driver of the same or a different racial or ethnic group (Close and Mason, 2007).

These complexities illustrate the difficulty of identifying the role of race in producing racial disparities in stops and searches, and suggest a second approach that incorporates the contexts in which individual officers consider race in their everyday interactions with citizens. Gelman and colleagues (2007) and Alpert and colleagues (2005) show how neighborhood context influences both the attribution of suspicion that animates an encounter and the outcomes of police-citizen encounters. The institutional context of policing also may influence individual officers' decisions by stigmatizing neighborhoods as "high-crime" or disorderly, skewing how officers perceive and interpret the actions of citizens. Institutional cultures also may implicitly tolerate such perceptual or cognitive schema and internalize them into policy preferences and strategic decisions, as well as internal preferences for reward, promotion, or discipline. These contextual concerns, informed by crime plus social and demographic dimensions of neighborhoods, suggest the second approach, one that explicitly incorporates either a multilevel approach that examines officer-place interactions, or shifts the focus from the actions of individual officers and individual suspects to the behaviors of cohorts of officers who collectively patrol neighborhoods with measurable attributes that incorporate race and ethnicity, and where aggregation biases from racial concentration may shape officers' preferences about crime and thresholds of suspicion.

These issues inform several features of the analyses reported in this chapter. First, to explain the distribution and predictors of street stops and then of arrests ("hit rates"), we focus on neighborhoods, not individual officers. Neighborhoods are the focal point of the underlying theories of order-maintenance policing. Place also is the unit of analysis for the allocation and deployment of police resources, and neighborhood crime rates are the metrics by which the resources of the police are managed and evaluated. Place also imparts meaning to the interpretation of routine actions and movements of citizens, whether local residents or outsiders whose appearance may evoke special attention. And the benchmark of the social composition of place, in conjunction with actual crime, is sensitive to the actual allocation of police resources as well as tactical decisions by the NYPD, and is widely used in research on selective enforcement in policing (Alpert et al., 2005; Fagan, 2002; Fridell, 2004; Skogan and Frydl, 2004).

Next we address supply-side and omitted-variable problems by controlling for the prevalence of the targeted behaviors in patrolled areas, assessing whether stop-and-search rates exceed what we would predict from knowledge of local criminal activity. This responds to the benchmark problem in research on selective enforcement. This approach requires estimates of the supply of individuals engaged in the targeted behaviors, and the extent of racial disproportionality is likely to depend on the benchmark used to measure criminal behavior (see Miller, 2000; Fagan and Davies, 2000; Walker, 2001; Smith and Alpert, 2002; Ayres, 2008; Durlauf, 2006a, 2006b; Ridgeway and MacDonald, this volume). Ideally, we would know race-specific crime rates in each social area to disaggregate benchmarks by race and ethnicity. But we observed practical problems in this approach. For example, clearance rates vary by crime type, and so the race of suspects is often unknown. Fewer than one in four stops in 2007 were based on a match between the person detained and a suspect description known to the police (Ridgeway, 2007). And suspected crimes that animate a large share of stops, such as weapons or drug possession, often do not follow from crime reports that identify the race of a suspect, so these base rates of offending are unknown.

Accordingly, we use homicide arrests as a measure of reported crime. Homicide victimization and arrests are stably measured over time, limiting measurement error. In New York, its racial distribution—both offending and victimization—is highly correlated with the demography of the neighborhood where the crime takes place (Fagan and Davies, 2004; Fagan et al., 2007). In New York City, the site of this research, homicide records are both a strong lag and lead indicator of crime, correlated

at .75 or more with reported crimes for other Part I felonies for the seventeen years from 1984 to 2000. Homicides also are the most stably and reliably measured indicator of crime over time and through police administrations, whereas other violent crimes (e.g., aggravated assault) are subject to classifications biases that vary over time and place (Zimring and Hawkins, 1997).

Following Gelman and colleagues (2007), we estimate whether the stop rate and "hit rate" within neighborhoods is predicted by local crime conditions, the physical and social composition of the neighborhood, or its racial composition. Since race is correlated with neighborhood composition and crime, we expect that race will not be a significant predictor either of stop patterns or of efficiency (the rate at which stops produce arrests), once we account for crime and other neighborhood conditions. But as we show below, race does predict stop rates and hit rates, after controlling for crime and local conditions. Is this evidence of racial animus, targeted collectively by officers in a neighborhood or through institutional and administrative levers that mark neighborhoods characterized by their racial or ethnic composition as worthy of heightened suspicion? The fact that police are stopping minorities, and others in minority neighborhoods, at a higher rate than is justified by local crime conditions does not require that we infer that police engaged in disparate treatment—but, at a minimum, it is evidence that whatever criteria the police employed produced an unjustified racially disparate impact.

III. Data and Methods

A. Data

We examine changes in OMP enforcement patterns beginning with the period examined by Spitzer (1999), Fagan and Davies (2000), and Gelman and colleagues (2007). Including that period (1998–99), we examine three distinct periods, termed the "early" (1998–1999), "middle" (2002–2004), and "recent" (2005–2006) periods. In each period, data on stop activity are based on records from the New York Police Department. The department has a policy of keeping records on stops (on "UF-250 forms") (see Spitzer, 1999; *Daniels et al. v. City of New York*, 2003); this information was collated for all stops from January 1998 through March 1999, and the 2003 and 2006 calendar years. Stops are recorded and aggregated for each precinct. Appendix A discusses the legal requirements for a stop, frisk, and arrest pursuant to a stop. Data on stops, frisks, and arrests from 2003 to 2007 were made publicly available by the NYPD following a Freedom of Information Law (FOIL) request and subsequent court order (NYCLU, 2008). Data from the "early" period were published in Spitzer (1999) and Fagan and Davies (2000).

Stop rates are analyzed in the context of citywide crime, demographic, and socioeconomic conditions. We use total stop rates (undifferentiated by suspected crime) and stop rates disaggregated by the race of person stopped. We use two measures of crime in the preceding year. First, in the figures, we use reported homicides in the police precinct in the preceding year as the measure of crime. This lagged function allows us to avoid simultaneity concerns from using contemporaneous measures of crime and police actions. Second, in the multivariate models, we use homicide arrests as the marker of crime.

We measure homicides for the "early" period using the NYPD's arrest-and-complaint file, and the city's COMPSTAT records for the "middle" and "recent" periods. In the multivariate estimates in tables 13.2 and 13.3, we use lagged homicide arrests in each neighborhood as the benchmark for estimating the proportionality of police stops and frisks. There are obvious strengths and weaknesses in this measure. Arrests are subject to police preferences for resource allocation, and also to police skills in identifying and capturing offenders. Homicide arrests also may vary by neighborhood based on externalities such as the extent of citizen cooperation with police investigations. Arrests also are vulnerable to measurement error: they often are reduced to other charges when evidence is too inconclusive to sustain a greater charge. But arrests also have strengths as a measure of crime. Reported homicides and homicide arrests are highly correlated over time across police precincts in New York: the partial correlation by month and precinct from 1989 through 2001 was .952.[3] This endogeneity of crime and policing within neighborhoods captures the preferences of police to allocate resources to particular areas in the search for offenders. Also, homicide arrests are a strong indicator of both arrests and complaints for other serious crimes.[4] To the extent that crime in the prior year is influenced *both* by crime and the policing that it attracts, the use of arrests as a measure of both the presence of police and of local crime conditions avoids omitted-variable problems when using only measures of reported crimes. Finally, arrest trends in preceding periods incorporate the priors of both individual officers and their supervisors as well as neighborhood characteristics, and in fact may capture officers' propensities to stop citizens based on the joint influence of individual and neighborhood racial markers.

We also incorporate demographic and socioeconomic variables in each area that might compete with or moderate crime as influences on stop activity: concentrated neighborhood disadvantage, residential turnover, and ethnic heterogeneity have each been associated with low levels of neighborhood collective efficacy and informal social control. These are both indicia of perceived disorder (Sampson and Raudenbush, 1999) and risk factors for crime (Fagan and Davies, 2004). More important, Fagan and Davies (2000) showed that these were salient predictors of stop activities in the "early" period, and we examine their influences over time as time-varying predictors. Areas in which these phenomena are concentrated might therefore be unable to informally regulate local residents, requiring law enforcement agencies to impose formal social control instead and leading to greater search activity.

Demographic and socioeconomic data for each period is based on the New York City Housing and Vacancy Survey (HVS), a survey completed every three years by the City's Department of Housing Preservation and Development, in cooperation with the U.S. Bureau of the Census (http://www.census.gov/hhes/www/housing/nychvs/nychvs.html). We analyze the 1999, 2002, and 2005 waves of the survey to generate baseline estimates of neighborhood social and economic status. Each wave covers approximately eighteen thousand housing units, classified into fifty-five "subboros,"

based on the Public Use Microdata Areas (PUMAs) for New York City (Community Studies of New York, 2007). We used shape files provided by the New York City Department of City Planning to reconcile the subboro boundaries with the police precincts (see Fagan and Davies, 2000). In the small number of precincts where there was overlap in the boundaries, precincts were assigned to the subboro that contained the majority of its population.

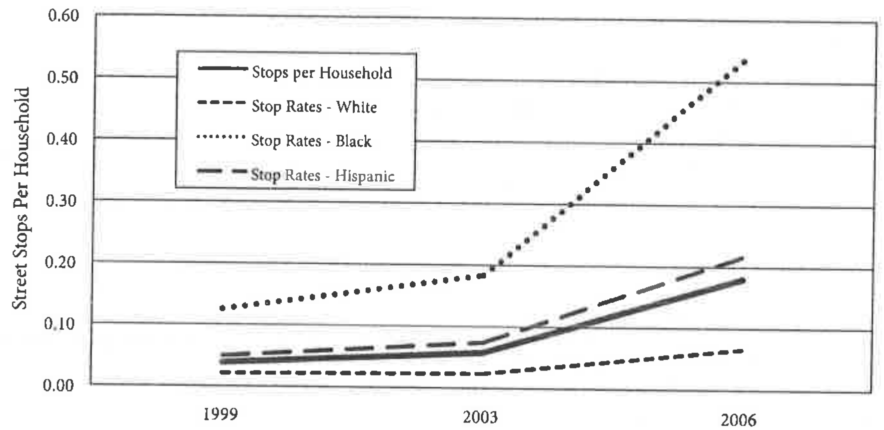
B. Base Rates and Citywide Trends

A quick look at the data on New York City neighborhoods suggests that the social and demographic makeup of the City has changed significantly since 1999. Table 13.1 shows that the city's racial and ethnic makeup has become more diverse. The bulk of the city's population growth has come from racial and ethnic minorities, plus

TABLE 13.1
Stop Activity and Neighborhood Socioeconomic Conditions

	1999 Stops per 1,000 persons		2002–2003 Stops per 1,000 persons		2005–2006 Stops per 1,000 persons		% change (99–05)
Citywide Stop Rates							
Stops per 1,000 Population							
Total Stops	12.5		19.4		60.2		381.6%
Blacks	26.6		37.7		130.8		391.7%
Whites	3.5		6.0		17.9		411.4%
Hispanics	15.1		19.5		63.9		323.2%
	Mean	SD	Mean	SD	Mean	SD	
Neighborhood Stop Activity							
Number of Stops	1813.4	1098.9	2922.5	1670.5	9208.9	6480.4	407.8%
Stops of Blacks	988.1	864.3	1411.9	1368.6	4863.0	5479.1	392.2%
Stops of Whites	187.0	145.3	320.1	273.8	972.7	860.8	420.2%
Stops of Hispanics	583.9	559.9	810.2	599.5	2688.4	2173.9	360.5%
Physical Disorder							
Exterior Walls	3.09%	0.03	2.63%	0.02	2.83%	0.02	−8.5%
Exterior Windows	3.36%	0.03	3.45%	0.03	2.36%	0.02	−29.8%
Stairways	5.25%	0.04	5.29%	0.04	4.24%	0.03	−19.3%
Floors	5.08%	0.04	4.75%	0.04	4.06%	0.03	−20.1%
Structural Characteristics							
Public Assistance	18.24%	0.13	15.17%	0.10	16.41%	0.11	−10.0%
Foreign-Born	46.19%	0.16	43.56%	0.14	49.61%	0.16	7.4%
Immigrant (different in HVS)	36.34%	0.16	43.56%	0.14	41.18%	0.16	13.3%
Entropy	89.02%	0.24	93.64%	0.25	95.48%	0.22	7.3%
Mobility (% Living < 5 years)	40.26%	0.05	35.88%	0.05	36.08%	0.05	−10.4%
Vacancy Rate	5.62%	0.03	6.87%	0.04	6.68%	0.03	18.8%
Households							
Total	52153	19305	54642	16552	55236	16803	5.9%
Black	12150	11930	13115	13382	12570	12603	3.5%
White	24112	23404	24359	22015	24191	21426	0.3%
Hispanic	11682	9155	12200	9063	12881	9206	10.3%

Sources: Socioeconomic and Household Data from New York City Housing and Vacancy Surveys, 1999, 2002, 2005, Stop data from NYPD, Population data from U.S. Census Bureau.



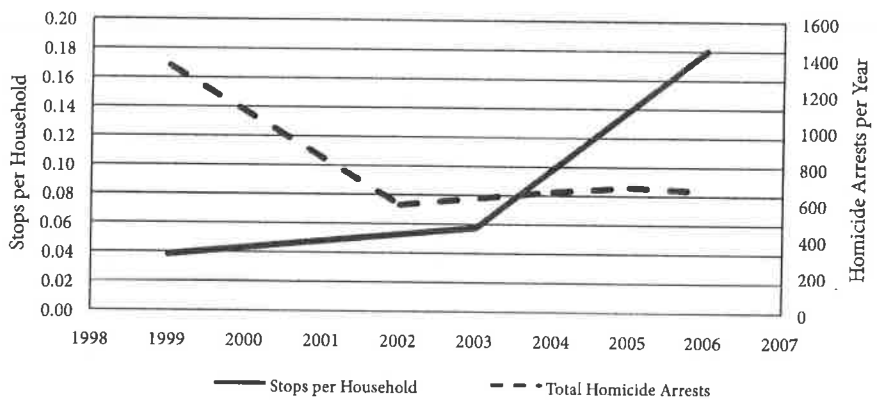


Figure 13.1 (*top*). Stops per household, New York City, 1999–2006. Sources: (Stops) NYS, Office of the Attorney General, 1999; NYC Police Department, Stop Frisks and Search Data, 2003–2007, (Households) NYC Housing and Vacancy Survey.

Figure 13.2 (*bottom*). Stops per household and total homicide arrests, New York City, 1999–2006. Sources: (Stops) NYS, Office of the Attorney General, 1999; NYC Police Department, Stop Frisks and Search Data, 2003–2007, (Households) NYC Housing and Vacancy Survey, (Arrests) NYS Division of Criminal Justice Services.

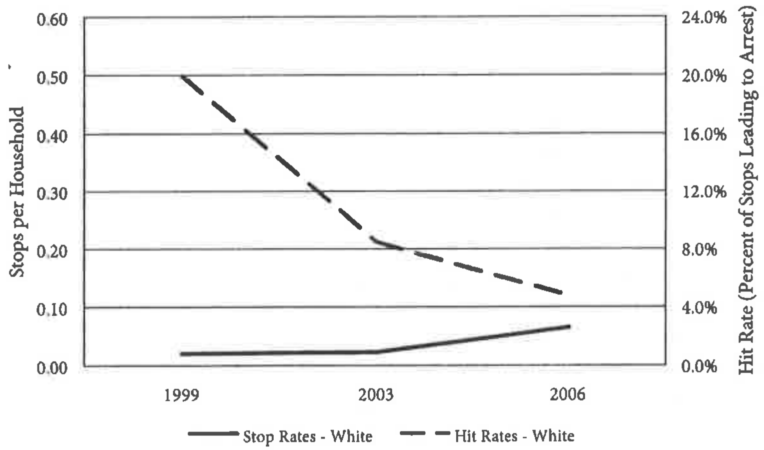
a notable increase among immigrants. Individual neighborhoods have also become more integrated, as shown by the increase in neighborhood entropy. At the same time, socioeconomic conditions have improved, with a decline in both public assistance receipt and neighborhood levels of physical disorder.

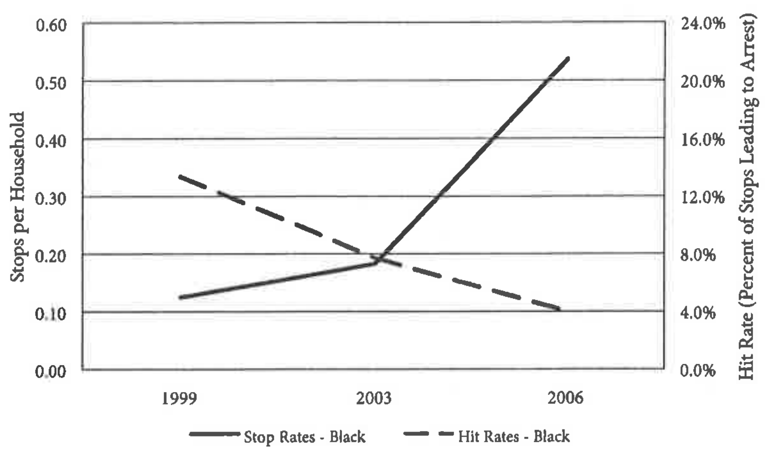
Even as the city has changed demographically and improved socioeconomically, stops and searches have become far more prevalent. Figure 13.1 shows the average neighborhood—subboro—stop rate, computed as stops per household. We use household because this is the population parameter in the HVS in each analysis period. While city residents of all races have become increasingly likely to be stopped by the police, stop rates vary dramatically by race; by 2006, Blacks were more than twice as likely to be stopped as either Whites or Hispanics. The increase in stop activity is particularly striking when considering that New York City crime rates fell dramatically between 1999 and 2006. As shown in figure 13.2, homicide arrests in the City fell by more than 50 percent between 1999 and 2002, and, albeit with a slight increase, remained low through 2006.

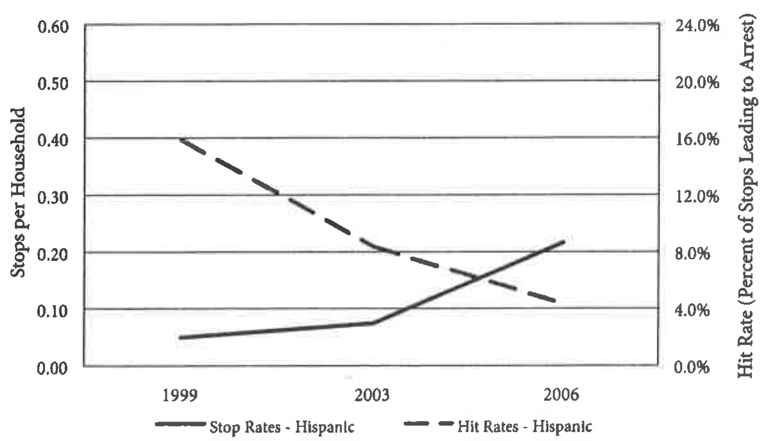
Following the examples of Knowles and colleagues (2001), Ayres (2002a,b), Gross and Barnes (2002), Gelman and colleagues (2007), and Ridgeway (2007), we measure the effectiveness of street stops by their "hit rates," the rate at which stops result in arrests. Figures 13.3a–c, like figure 13.1, present average neighborhood stop rates per household in each of the three time periods of interest, disaggregated by race, with average hit rates overlaid onto the graph. And since crime rates remained relatively stable across the period, there is no evidence that the increase in stops contributes to crime minimization. While not as pronounced as the differences in stop rates, hit rates also suggest substantial racial disparities. Figure 13.3b shows that even as stop rates have increased dramatically for Blacks from 2003 to 2006, hit rates have fallen steadily, suggesting that the increase in stop activity has added little value in maximizing efficiency via generating arrests. Stops of Whites appear more likely than stops of Blacks to lead to arrest, suggesting that Blacks are disproportionately subjected to stops, with little public safety payoff.

C. Stop Activity by Neighborhood

Stop rates have not only increased dramatically, but between-neighborhood differences in stop rates have become far more pronounced. Figure 13.4 displays one data point for each of the fifty-five HVS subboros in each period, each representing the average neighborhood stop rate per household in each year. We also show the count of homicides citywide over the same period. While earlier studies have identified neighborhoods that have the greatest racial disparity in stop-and-frisk practices, figure 13.4 shows that the dramatic growth in average stop rates from 2003 to 2006 is explained by extreme increases in a subset of neighborhoods with high rates of African American and Latino residents: Brownsville, East New York, Central Harlem, East Harlem, Bedford-Stuyvesant, and Mott Haven. Although some of this increase may be due to improved reporting, it is curious that all the improved reporting has been in neighborhoods with the highest non-White populations in the City. These neighborhoods are predominantly African American, according to the Department of City Planning.[5]







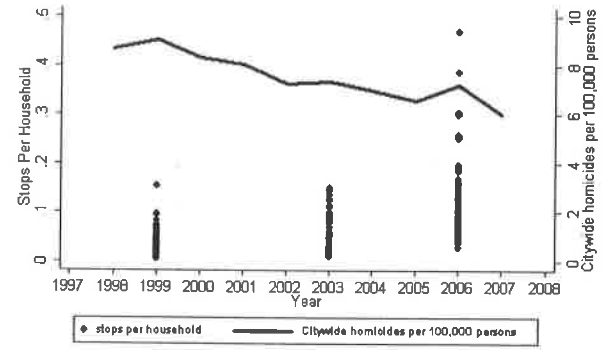


Figure 13.4. Street stops per neighborhood, selected years, 1999–2006.

Given the degree of racial segregation across New York City neighborhoods, we address this disparity below by examining neighborhood-level drivers of stop activity.

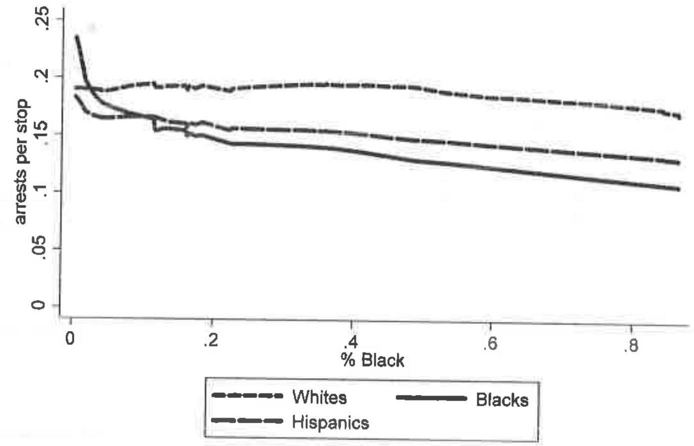
Figures 13.5a–c suggest that neighborhood racial composition explains not only stop activity but also hit rates and stop efficacy. Each figure shows, for 1999, 2003, and 2006, respectively, a LOWESS-smoothed estimate of the relationship between hit rates and the percentage of Blacks in each of the fifty-five neighborhoods for each period of time. As in figure 13.3 (a,b,c), these graphs suggest that hit rates are falling over time in stops of all racial groups. Particularly in 2006, however, the year when between-neighborhood differences are most pronounced (see figure 13.4), there is a visible difference in neighborhoods with the highest concentrations of Black households. In neighborhoods where 60 percent of households (or more) are Black, stops are not only less effective than in more mixed or White neighborhoods, but hit rates are particularly low in stops of Black and Hispanic individuals.

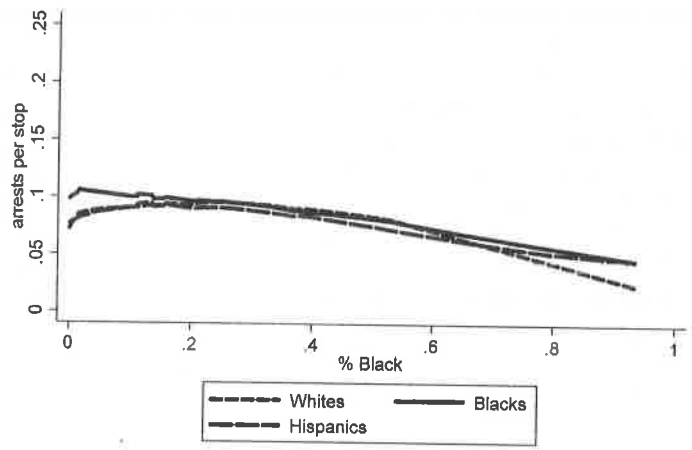
Opposite page:

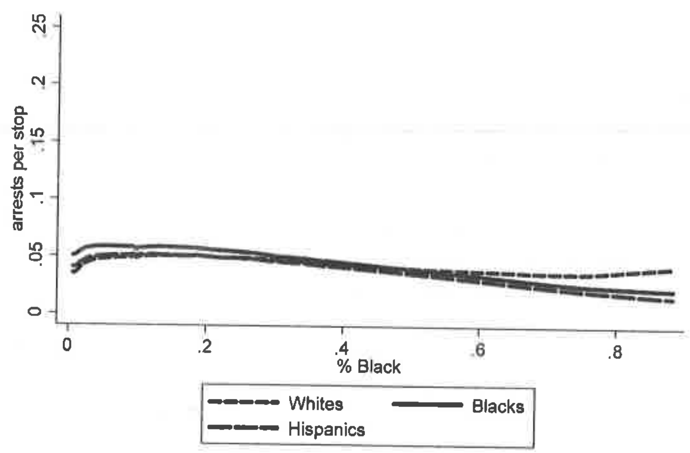
Figure 13.3a (*top*). Stops per household and arrests per stop, White suspects, New York City, 1999–2006. Source: (Stops and Arrests) New York State, Office of the Attorney General, 1999; New York City Police Department, Stop Frisks and Search Data, 2003–2007, (Households) NYC Housing and Vacancy Survey.

Figure 13.3b (*middle*). Stops per household and arrests per stop, Black suspects, New York City, 1999–2006. Source: (Stops and Arrests) New York State, Office of the Attorney General, 1999; New York City Police Department, Stop Frisks and Search Data, 2003–2007, (Households) NYC Housing and Vacancy Survey.

Figure 13.3c (*bottom*). Stops per household and arrests per stop, Hispanic suspects, New York City, 1999–2006. Source: (Stops and Arrests) New York State, Office of the Attorney General, 1999; New York City Police Department, Stop Frisks and Search Data, 2003–2007, (Households) NYC Housing and Vacancy Survey.







D. Modeling Strategy

1. PREDICTING STOP ACTIVITY

Given the between-neighborhood disparities shown in figure 13.4, we examine stop activity at the neighborhood level to identify factors that explain between-neighborhood differences both within periods and over time. Following Gelman and colleagues (2007), we estimate a series of Poisson regressions to predict the number of stops conducted in each neighborhood in each time period. The racial disparities shown in figures 13.1 and 13.3 may be driven not by race, but rather by differences in neighborhood social conditions where Blacks, Whites, and Hispanics are concentrated, or by differences in their *ex ante* crime conditions. If, for example, the police make more stops in high-crime areas, but treat individuals similarly within similarly situated localities, racial disparities in stop rates could be explained entirely by neighborhood crime conditions. Or the NYPD's focus on "broken windows" and order-maintenance policing might lead stop activity to be most prevalent in neighborhoods with disorderly conditions (Wilson and Kelling, 1982; Kelling and Cole, 1996). We therefore estimate a model where the stop count y_i in neighborhood i is distributed based on predictors X, with an expected value of:

$$E[y_i|X_i] = e^{X\beta}$$

The vector X includes a measure of neighborhood crime (homicide arrests, lagged), and several socioeconomic characteristics we expect to be correlated with crime rates and policing practices. First, we explicitly control for crime conditions in the previous year, using the number of homicide arrests in each neighborhood. To reflect the NYPD focus on disorder in the 1990s and early 2000s, we estimate and control for a single principal components factor (computed for each year) that summarizes the physical condition ("broken windows," literally) of local residences (based on the percentage of buildings whose windows, walls, floors, and stairways have problems visible to outside observers). The disorder theories animating OMP strategies considered both physical and social disorder as cues of weak informal social control and low guardianship of neighborhoods. We consider only physical disorder since some elements of social disorder—such as fighting, visible drug use—are in fact crimes and would be correlated with stop activity.[6] Also, physical disorder tends to be highly correlated with social disorder, and its component behaviors, including public intoxication, loitering, and fighting (Sampson and Raudenbush, 1999). These are targeted in OMP as a wedge to reduce crime opportunities and to identify persistent criminals. To reflect the likelihood that police activity is higher in more populated areas, we control for the logged number of households in each neighborhood.

Opposite page:

Figure 13.5a (*top*). Lowess-smoothed arrest rates by neighborhood racial composition, 1999.

Figure 13.5b (*middle*). Lowess-smoothed arrest rates by neighborhood racial composition, 2003.

Figure 13.5c (*bottom*). Lowess-smoothed arrest rates by neighborhood racial composition, 2006.

We also control for traditional and temporally stable predictors of neighborhood crime (Shaw and McKay, 1942; Sampson and Lauritsen, 1994; Land et al., 1990; Fagan and Davies, 2004, Fagan, 2008; Kirk and Laub, in press): concentrated disadvantage (measured by the percentage of households receiving public assistance), residential instability (measured by the percentage of families who have moved to the their current residence within five years, and by the residential vacancy rates), ethnic diversity (measured by the percent of residents who are Black or Hispanic, the percentage who are foreign-born, and a measure of entropy, which captures the degree of ethnic heterogeneity in the neighborhood). We expect, however, that these factors will be correlated with police activity only to the extent that they predict crime; once crime conditions are controlled for, there should be no marginal relationship between social structure and stop activity. Variables (with the exception of logged population) are standardized to a mean of zero and variance of one, and neighborhood observations are weighted based on the number of households in each.

To assess the extent to which neighborhood conditions, and their influence on policing, change over time, we first estimate three separate cross-sectional models, one for each time period of interest. We then combine the observations into a pooled cross section (model 4), and add controls for year fixed effects in Model 5. Model 6 contains year fixed effects and random intercepts with standard errors clustered by neighborhood to account for neighborhood differences.

Although the City has changed for the better over the period of analysis, and stop activity has increased dramatically over time, the crime, disorder, and socioeconomic predictors vary far more between neighborhoods than they do within neighborhoods over time, and these differences—at least in ordinal position—are stable over time (see Sampson and Morenoff, 2006). Accordingly, we rejected the option to control for neighborhood fixed effects in Model 6, preferring instead to focus on differences between neighborhoods. Controlling for neighborhood fixed effects identifies the relationship between crime and stop activity, and social structure and stop activity, solely from within-neighborhood variation. Because we acknowledge that the allocation of police resources is determined by differences between neighborhoods, model 6 is specified to reflect between-neighborhood differences, with random intercepts and standard errors clustered by neighborhood.

2. PREDICTING STOP EFFECTIVENESS

We next examine the crime and socioeconomic conditions predicting stop effectiveness, the "hit rate" at which stops lead to arrests. We expect that this rate might be tied to the same conditions of crime and disorder that predict stop activity, since "excess stops" above the crime rate are likely to be concentrated in poor neighborhoods with concentrations of minority population. Accordingly, we estimate a series of linear probability models using the predictors detailed above. As we hypothesize with stop activity, however, in the case of race-neutral policing hit rates should not be significantly related to neighborhood social structure. For these analyses, we estimate the effects of neighborhood racial composition on stop rates using both neighborhood fixed effects and, also, as above, using random intercepts.

IV. Results

A. Explaining Neighborhood Differences in Stop Rates

Table 13.2 shows the relationship between neighborhood conditions and the incidence of street stops. Models 1–3 show results for each year. As expected, stops are more frequent in neighborhoods in which crime is more prevalent for all years, but in larger neighborhoods only in 1999. Controlling for homicides, stops are more frequent in neighborhoods with higher Black populations. The effect size is fairly stable across years, even as the overall number of stops rose over time. Model 4 is a pooled cross-sectional model for all years, with no controls for time. Standard errors are clustered by neighborhoods. The effect for Black population remains significant, and population is again significant when the three time periods are pooled.

TABLE 13.2
Poisson Regressions of Stops per Neighborhood, Controlling for Social Structure and Crime, 1998–2006

	Model					
Sample Year	(1) 1999	(2) 2003	(3) 2006	(4) All Years	(5) All Years	(6) All Years
Homicide Arrests (lagged)	.202**	.163*	.182**	.172**	.183**	.027*
	[.074]	[.069]	[.067]	[.084]	[.055]	[.052]
% Receiving Public Assistance	.106	.056	.169	.257*	.159	.198*
	[.127]	[.089]	[.099]	[.131]	[.086]	[.082]
% Foreign-Born	−.011	.006	−.045	−.056	−.032	−.076
	[.079]	[.062]	[.083]	[.072]	[.060]	[.065]
Racial Entropy	.186*	.007	.091	.090	.082	.085
	[.086]	[.060]	[.064]	[.066]	[.050]	[.059]
% Black	.216*	.198**	.262**	.260**	.237**	.279***
	[.109]	[.072]	[.068]	[.068]	[.060]	[.064]
% Hispanic	.053	.002	.054	−.023	.021	.031
	[.113]	[.078]	[.083]	[.072]	[.063]	[.074]
% Moved Within 5 years	.005	−.056	−.012	−.007	−.006	.008
	[.098]	[.065]	[.098]	[.082]	[.069]	[.064]
Vacancy Rate	.038	−.074	.090	−.007	.050	.026
	[.090]	[.076]	[.076]	[.074]	[.044]	[.042]
Physical Disorder	.028	.152	−.109	−.011	−.053	−.048
	[.081]	[.074]*	[.105]	[.114]	[.071]	[.064]
Log Population	.505*	.438	.451**	.769**	.445**	.407**
	[.231]	[.230]	[.173]	[.212]	[.157]	[.065]
2003 FE					.460**	.451**
					[.060]	[.065]
2006 FE					1.590**	1.585**
					[.078]	[.083]
Constant	1.953	3.140	4.115	−.003	2.600	1.002
	[2.523]	[2.521]	[1.911]	[2.323]	[1.727]	[1.729]
Observations	55	55	55	165	165	165
Wald Chi-squared	114.76	64.32	119.12	156.3	1081.5	832.1
Neighborhood FE?	No	No	No	No	No	No[a]
Year FE?	No	No	No	No	Yes	Yes

Socioeconomic predictors are standardized to a mean of 0 and variance of 1.
Observations weighted by the number of households per neighborhood.
Robust standard errors in brackets; models 4–6 cluster standard errors by neighborhood.
[a] Model 6 includes random intercepts for neighborhoods and AR(1) covariance.
* $p<.05$; ** $p<.01$.

Model 5 includes year fixed effects, but not neighborhood fixed effects, and the standard errors are clustered by neighborhood. The results are unchanged from model 4. The year fixed effects for 2003 and 2006 are significant, reflecting the increase in the stops in the subboros in those periods relative to the 1999 rate. Physical disorder is not significant, nor are the majority of other covariates that characterize neighborhoods. But stops are more frequent in areas with higher concentrations of public assistance receipt, and with higher Black populations, after controlling for homicides and physical disorder. Since homicide rates in New York and physical disorder are correlated with Black population concentration (Fagan and Davies, 2000, 2004), we estimated models including interaction terms for percentage of Black residents and local disorder conditions (physical disorder). The relationship of Black population and the stop rate is robust to the inclusion of either interaction term (data available from authors).

Thus far, model 5 shows a strong and significant relationship between neighborhood racial composition and stop activity; police stop significantly more people in neighborhoods with more Black households. Given that all predictors are standardized, with the exception of the logged number of households, the coefficient magnitudes suggest a particularly strong relationship; racial composition is as important as local crime conditions in predicting police stop activity.

For Model 6, we also included two types of sensitivity analyses. First, we estimated the models including interactions of percent black by lagged crime and percent Hispanic by lagged crime. The results were unchanged. Next, recognizing the potential endogeneity of crime, disorder, neighborhood social and racial composition and stop rates, we estimated propensity scores for the racial composition measures and included them as predictors (results available from the authors). We estimated propensity scores to predict separately the Hispanic and Black concentrations in each neighborhood, and fixed effects for year. We then re-estimated Model 6 to include these propensity scores together with the main racial composition predictor. Following Bang and Robins (2005), we included a predictor that expressed the propensity scores for each racial composition variable in two ways:

$$(1)\ X_{i,j} = 1 / PS_{i,j}$$
$$(2)\ X_{i,j} = 1 / (1 - PS_{i,j})$$

In equations 1 and 2, X is the expression of the transformed propensity score PS_i, the estimated (predicted) racial composition for each race *i* and in neighborhood *j*. We repeated this procedure using the standardized residuals from the propensity score estimation, creating two additional propensity scores expressions. Again, the results using these estimators were unchanged (results available from the authors). Accordingly, the results in Table 13.2 are robust with respect to a variety of controls and specifications of the local crime and social conditions that might influence stop rates.

We also estimated Model 6 using both neighborhood and year fixed effects, but the model fits were unacceptably poor and the results uninterpretable. Which

modeling strategy produces the most accurate and reliable accounting for the relationship among neighborhood, crime, and stop activity? Which is a more accurate identification strategy for estimating the effects of policing on neighborhoods? We are confident in the results in models 5 and 6, and reject the unstable results for the neighborhood fixed effects model, for four reasons. First, as mentioned earlier, while there were strong within-neighborhood changes over time, the relative position of neighborhoods in terms of both crime and concentrated socioeconomic disadvantage over time was largely unchanged. In other words, the worst places still are the worst places—the places with the highest homicide rates still are the places with the highest homicide rates, the places with the highest concentration of physical disorder are still the places with the most bad housing, even as the extent of disorder in those places dissipates over time. Neighborhood fixed effects are somewhat helpful in identifying differences between places, but such differences are likely to be unimportant in this analysis. Inclusion of fixed effects for neighborhoods in this context would overdetermine the model, explaining everything and nothing at the same time.

Second, the neighborhoods are changing over time, but the rates of change are dissimilar. The social, economic, and crime conditions in poorer neighborhoods changed more than in wealthier neighborhoods (Fagan, 2008). The assumptions of stable between-unit rates of change in fixed effects models are challenged under these conditions. Third, fixed effects estimators are quite limited when the possibility exists of dynamic selection, or changes in the circumstances or preferences that would affect the assignment of the intervention—police stops, in this case—over time (Bjerk, 2008). Dynamic selection is intrinsic to the policy preferences in the allocation of police resources and tactics in the OMP model (Bratton and Knobler, 1998; Silverman, 1999). This in turn leads to our fourth concern: we think that fixed effects models in this context ask the wrong question. Our interest here is estimating the probabilities of being stopped in neighborhoods of different racial makeup and crime conditions, not with differentials by race of persons within neighborhoods. In other words, ours is a within-neighborhood design, and we seek to explain differences in stop probabilities that are quite dramatic across places and over time.

B. The Efficiency of Street Stops in Detecting Crime

Table 13.3 presents the relationship between neighborhood conditions and "hit rates," or the percent of stops that lead to arrests. As suggested earlier, by figures 13.3a–3c and 13.5a–5c, stop efficacy has declined over the period of analysis, a trend underscored by the year fixed effects in models 5 and 6. We would expect that neighborhood hit rates, driven by the likelihood of stopped residents to be engaged in illegal activity, would not be tied to neighborhood social structure; models 1–5, however, show that arrests per stop are lower in neighborhoods whose populations are predominantly Black: over time, stops in predominantly Black neighborhoods are significantly less productive in yielding arrests than in other parts of the City. Table 13.2 shows that stops are far more prevalent in these areas, to a degree beyond

TABLE 13.3
OLS Regression of Arrrests per Stop, Controlling for Social Structure and Crime, 1998–2006

	Model					
Sample Year	(1) 1999	(2) 2003	(3) 2006	(4) All Years	(5) All Years	(6) All Years
Homicide Arrests (lagged)	.010	.002	.008*	.010*	.007*	.003
	[.010]	[.005]	[.003]	[.004]	[.003]	[.007]
% Receiving Public Assistance	-.010	.000	.006	-.002	.003	-.018
	[.012]	[.007]	[.004]	[.006]	[.004]	[.012]
% Foreign-Born	.000	.002	-.003	.001	-.001	.013
	[.011]	[.005]	[.004]	[.005]	[.004]	[.016]
Racial Entropy	-.007	.000	.007	.003	.004	.011
	[.010]	[.004]	[.004]	[.004]	[.003]	[.016]
% Black	-.029*	-.009	-.012**	-.018**	-.014**	-.013
	[.011]	[.006]	[.004]	[.006]	[.004]	[.039]
% Hispanic	.007	.009	-.004	.001	.001	-.010
	[.014]	[.008]	[.005]	[.006]	[.005]	[.029]
% Moved Within 5 years	-.008	.000	.001	-.004	-.003	-.007
	[.010]	[.004]	[.003]	[.004]	[.002]	[.006]
Vacancy Rate	-.003	.011**	.001	.006	.005	.009
	[.012]	[.003]	[.004]	[.005]	[.003]	[.005]
Physical Disorder	.001	-.009	-.007	-.005	-.006	-.007
	[.011]	[.005]	[.005]	[.006]	[.004]	[.006]
Log Population	-.024	.047*	.016	-.010	.017	.080
	[.031]	[.019]	[.013]	[.014]	[.011]	[.102]
2003 FE					-.070**	-.074**
					[.009]	[.012]
2006 FE					-.108**	-.109**
					[.007]	[.011]
Constant	.412	-.433	-.131	.170	-.035	-.602
	[.339]	[.205]*	[.146]	[.156]	[.125]	[1.085]
Observations	55	55	55	165	165	165
R-squared	.280	.380	.410	.130	.690	.830
Log Likelihood (model)	92.91	119.83	143.11	278.03	363.02	412.01
BIC	-141.7	-195.6	-242.1	-499.9	-659.7	-762.8
Year FE?	No	No	No	No	Yes	Yes
Neighborhood FE?	No	No	No	No	No	Yes

Socioeconomic predictors are standardized to a mean of 0 and variance of 1.
Observations weighted by the number of stops made.
Robust standard errors in brackets; models 4–6 cluster standard errors by neighborhood.
* $p<.05$; ** $p<.01$.

what differential criminal activity would suggest; the models in Table 13.3 suggest that there is little public safety payoff. The results in model 6, however, suggest that race is no longer a significant predictor of hit rates when we treat neighborhoods as fixed effects. But when we estimate Model 6 using random intercepts and population-averaged models, we obtain the same results as in Model 5: arrest rates are significantly lower in neighborhoods with greater black population (for percent black, b=.13, s.e.=.005, p=.017). Again, we face the same issues in interpretation with respect to the neighborhood fixed effects models, and for the same reasons as discussed earlier, we reject the neighborhood fixed effects model in favor of other identification strategies that rely on clustering of standard errors by neighborhood.

Finally, to put the hit rate analysis in perspective of gains and losses, we computed the number of firearms obtained from stops. In 2003, a total of 633 firearms were

seized pursuant to stops, a rate of 3.9 firearms per 1,000 stops. More than 90 percent of the firearms seized were pistols. By 2006, following a 300 percent increase in the number of stops, the seizure rate fell to 1.4 firearms seized per 1,000 stops. The firearm seizure rates for Blacks, who were stopped more than ten times the rate per person compared to whites, were slightly higher: 4.6 firearms seized per 1,000 stops in 2003, and 1.6 seizures per stop in 2006. The seven hundred firearms seized in 2006 through stops accounted for about 10 percent of the total number of firearm seizures in New York City that were traced in the nationwide firearm trace system. On the surface, the expenditure of police resources to seize only a fraction of seizures made by other means seems inefficient, to say the least. Since removal of guns from the street was the animating goal of OMP, the low seizure rate is further evidence of the inefficiency if not futility of the strategy.

C. How Much Is Too Much? How Much Is Enough?

The burden of OMP policing in the decade since the Spitzer (1999) report has fallen disproportionately on African Americans, and, to a lesser extent, on Latinos. The strategic goal of OMP has principally been one of law enforcement—maximization of arrests and punishment. This was evident in the policy memoranda that were issued at that the outset of the OMP experiment in 1994. Crime minimization goals were path-dependent on the law enforcement goals, rooted in the putative benefits of increased stops and arrests of citizens for both minor crimes plus the detection of weapons and other contraband. Through careful allocation of police resources, the focus was on "high-crime" areas, which—in the logic of OMP—were those places with the highest concentrations of poor, non-White citizens. The high-crime area concept has proven to be elastic, though, and has expanded now to include public housing developments, despite equivocal evidence that crime in public housing is higher than in the adjacent areas (Fagan and Davies, 2000; Fagan et al., 2006). The result has been a dramatic increase in street stops since 2003, with nearly five hundred thousand New Yorkers stopped in both 2006 and 2007. In addition, tens of thousands of misdemeanor marijuana arrests (Golub et al., 2007; Levine and Small, 2008) are part of the totality of enforcement that nearly blankets some parts of the City.

Crime rates, though, have remained relatively stable in the years since 2003 as stops have increased. Figure 13.4 shows that homicide rates have remained stable after 1999, rising and falling randomly over an eight-year period. One might have expected crime rates to plunge further with the mobilization of OMP tactics, especially with the increase beginning in 2003, but that hasn't been the case. After all, a secondary benefit of maximizing punishment through street stops would be to raise the risk of detection and arrest for carrying weapons, increasing the deterrent threats of OMP tactics. But we are hard-pressed to detect such trends, given the stability of crime rates. Nor have marijuana arrests declined, despite the sharp rise in the likelihood of detection and arrest, so New Yorkers continue to use marijuana, often openly, flouting the law and discounting or ignoring the risks and consequences of arrests.

The inelasticity of crime relative to street stops raises two related questions. First, if crime minimization is the goal of OMP, rather than maximizing punishment without tangible linkages to crime reduction, how many stops are enough to maintain or lower the crime rate? Economists and criminologists have long sought algorithms that would create an optimal level of law enforcement (see Garoupa, 1997; Polinsky and Shavell, 2000, 2007; Curtin et al., 2007) or incarceration (Blumstein and Nagin, 1978) to control crime. For example, Persico and colleagues (2001) suggest that an optimal level of police searches of motorists can achieve an equilibrium across racial groups in the propensities of motorists to transport drugs or other contraband. So are five hundred thousand stops too many? Not enough to control crime? These are important questions, but we do not address them in this chapter.

The second question, though, is a first step in the process of answering the first question. Under current OMP tactics, what is the likelihood of police contact for citizens of specific racial and ethnic groups? Knowing the exposure of different population groups to detection and enforcement is a necessary antecedent to discerning whether there is leverage in these contact rates that can influence crime rates for any population group, or even for the areas where specific groups are concentrated. And if race, neighborhood, and crime are conflated to shape perceptions of "high-crime areas" that merit intensive patrol and enforcement, we would expect the exposure to be highest for non-Whites, and, as we see in figure 13.4, for African Americans in particular.

Accordingly, we estimated the probability of contact during 2006 for non-Hispanic African American males ages eighteen and nineteen, a group that has been the focus of criminal justice policy debate and research attention for nearly two decades (Fagan and Wilkinson, 1998; Cook and Laub, 1998; Loury, 2002; Feld, 1999). There were 28,945 stops of this group during 2006. The total population in 2006, according to the U.S. Bureau of the Census (U.S. Census Bureau, 2006), was 30,999. Accordingly, the point estimate for contact is .93, a figure that on its face is shocking. We reestimated this probability excluding stops made in police precincts in the City's central business districts and park areas: lower Manhattan, Midtown (including Times Square), and Central Park. With these restrictions, we reestimated the probability of contact at .92 (28,539 stops).[7] This compares to estimates of less than .20 for eighteen- and nineteen-year-old White males and .50 for Hispanic males (both Black and White Hispanics).

The stop totals are likely to include persons stopped more than once, so we reestimated these probabilities under varying assumptions about the number of persons stopped more than once and the total number of stops that were repeat stops. Table 13.4a shows that if 10 percent of the African American males ages eighteen and nineteen were stopped more than once, and these repeaters accounted for 25 percent of all stops, the probability of being stopped by the police of anyone in this age cohort is now .79. For example, if 10 percent of the population of Black men aged eighteen and nineteen (approximately 3,100 individuals) are considered "high-stop individuals," and this group makes up 25 percent of all stops within this demographic bracket, then these 3,100 people were stopped a combined 7,135 times. These men were stopped an average of 2.3 times over the course of the year, rather than the 0.92 suggested by

the raw numbers. Assuming that the remaining stops (21,404) are distributed one-per-person, the total number of people stopped over the course of the year would be 24,504. Although the raw ratio of stops to people in this demographic bracket is 0.92, the actual percentage of the population stopped by the police is lower, 0.79, shown in the upper-left cell of table 13.4a. If 25 percent of the persons were stopped more than once and they accounted for 50 percent of all stops, the probability declines to .71. Note that in table 13.4a, some cells could not be computed because the total number of stops would exceed the population in that group.[8]

We next expand the age boundaries for these estimates to include males ages eighteen to twenty-four. This age group was disproportionately involved in lethal violence throughout the 1990s in New York (Fagan and Wilkinson, 1998; Fagan et al., 1998) and elsewhere in the United States (Cook and Laub, 1998; Zimring and Hawkins, 1997). Also, desistance from crime increases substantially as persons reach their mid-twenties (Farrington, 1998). The unadjusted probability of being stopped in 2006, before accounting for repeaters, is .14 for non-Hispanic Whites, .78 for African Americans, and .39 for Hispanics.

Tables 13.4b–d show the rates accounting for different assumptions about the number of repeaters and the number of repeat stops. Given the lower stop rates of Whites and Hispanics, we rescaled the probabilities in tables 13.4c and 13.4d, hence the comparisons reflect distributions that are unique for each racial or ethnic group. Under the most likely scenarios, tables 13.4b–d show that when 10 percent of the persons account for 25 percent of the stops, the probability that an African American male is stopped (.69) is still far greater than the probability that a White or Hispanic male is stopped. Under more restrictive and conservative assumptions—that 50 percent of the persons account for 75 percent of the stops, we still estimate rates for African Americans that are twice the rate of Hispanics.

The important context in which to view these numbers is whether they are productive; by any reasonable standard, however, they are not. Figure 13.3 (a,b,c) shows two important features of hit rates: there are only negligible differences between hit rates for Whites, African Americans, and Hispanics, and the rates themselves are approximately 5 percent. Beyond the evidence of racial disparity, we are also concerned that these extraordinary stop rates of African Americans include a high volume of excess stops, stops that express unwarranted blanket suspicion that may have little marginal deterrent or law enforcement returns. But with stop rates this high and inefficiencies running at 96 percent, claims of a general deterrent effect from these stops are empirically strained by the scarcity of sanctions. So deterrence or crime control may be a secondary goal to maximization of punishment. And efficiency concerns are only one side of the social and public good of policing: equity, fairness, and distributive considerations co-occupy another dimension of policing (Moore, 2002). Even if we thought that there were crime control returns, it seems unlikely that most citizens would condone trading in the private harm of excess stops of African Americans, not to mention the stigma and internalized psychological costs, against putatively lower susceptibility to crime for the majority group. The costs of this regime lie in the harm to the 95 percent who are innocent in these excess stops.

TABLE 13.4A
Probability of Stops for African American Males, Ages 18–19, 2006

	% Repeat Stops 25%	50%	75%
% Stopped More Than Once			
10%	0.79	0.56	0.33
25%		0.71	0.48
50%			0.73

Note: Excludes stops that were made in 1st, 14th, 22d, and 18th precincts. Population: 30,999. Stops: 28,539.

TABLE 13.4B
Probability of Stops for African American Males, Ages 18–24, 2006

	% Repeat Stops 25%	50%	75%
% Stopped More Than Once			
10%	0.69	0.49	0.30
25%		0.64	0.45
50%			0.70

Note: Excludes stops that were made in 1st, 14th, 22d, and 18th precincts. Population: 104,880. Stops: 82,125.

TABLE 13.4C
Probability of Stops for Hispanic Males, Ages 18–24, 2006

	% Repeat Stops 25%	50%	75%
% Stopped More Than Once			
10%		0.29	0.20
20%			0.30
25%			0.35

Note: Excludes stops that were made in 1st, 14th, 22d, and 18th precincts. Population: 127,128. Stops: 48,968.

TABLE 13.4D
Probability of Stops for Non-Hispanic White Males, Ages 18–24, 2006

	% Repeat Stops 25%	50%	75%
% Stopped More Than Once			
2%	0.12	0.09	0.05
5%		0.12	0.08
10%			0.13

Note: Excludes stops that were made in 1st, 14th, 22d, and 18th precincts. Population: 107,936. Stops: 15,065.

V. Discussion

For nearly a decade, through a prolonged era of stably low crime rates and improving social and economic health across the City's neighborhoods, the number and rate of stops of citizens has increased by more than 500 percent while the efficiency of those stops has declined by nearly 50 percent. The burdens and benefits of these stops are disproportionately concentrated in the City's poorest neighborhoods, the places with both the highest crime rates and the highest proportions of non-White households. Our focus in this chapter is not on the race or ethnicity of individual stops of citizens, but on the rates of stops in neighborhoods with the highest concentrations of Black residents. We focus on neighborhoods because place, not individuals, has been most closely linked to the logic of policing under OMP since its inception fifteen years ago. It is place that is the focal point of the underlying theories of order-maintenance policing, place is the unit of analysis for the allocation and deployment of police resources, and the indicia of crime in places are the metrics by which the resources of the police are managed and evaluated. And the benchmark of place, in conjunction with crime, is sensitive to the actual allocation of police resources as well as tactical decisions by the NYPD, and is widely used in research on selective enforcement in policing (Alpert et al., 2005; Fagan, 2002; Fridell, 2004; Skogan and Frydl, 2004).

The effects we observe in these analyses are notable in three ways. First, stops within neighborhoods take place at rates in excess of what would be predicted from the separate and combined effects of population demography, physical and social conditions, and the crime rate. This excess seems to be concentrated in predominantly Black neighborhoods. Second, the excess stops in these neighborhoods persist over time, even as the Black population declines, crime rates remain low and effectively unchanged, the City's overall social and economic health improves, and housing and other investments increase across the City's neighborhoods, including its poorest and most segregated neighborhoods. Third, there appears to be a declining return in crime detection from marginal increases in enforcement, and this efficiency gap seems to grow over time. Like the stops that supply the arrests, the declining number of arrests that take place pursuant to stops are disproportionately concentrated in neighborhoods with higher Black populations, after controlling for crime, poverty, and disorder in those places.

The preferences for neighborhood selection for intensified stops seems to be inelastic to changes in crime rates or to the limited payoffs in arrest efficiencies from marginal increases in stops. This inelasticity is difficult to understand as either individual preferences of police officers, or as a rational tactical or management decision. As the rank and file of police in New York become more diverse and reflective of the City's demography, it is unlikely that individual preferences or subjective assessments of suspiciousness by individual officers would continue to be racially skewed over time and through changes in the social contexts of the areas they patrol.

Institutionally, the declining returns to crime control from marginal increases in stop activity is the opposite of economics. We assume, from the policy statements of police in New York, that the goal of stops is to minimize and deter crime rather than to maximize the hit rate of stops. An elastic policy sensitive to crime rates might seek to locate an optimal level of stop activity within each neighborhood or patrol area and adjust in real time. Dominitz and Knowles (2006) suggest that such a crime minimization approach works only if the priors of illegal behavior are known to vary across groups in specific ways. Perhaps the absence of assumptions or knowledge of specific variation in between-group (and by extension, between-neighborhood) crime preferences explains the persistence of these stop patterns. But we doubt that the NYPD is flying blind, since the allocation of police to neighborhoods and smaller areas is driven by real-time data about group- or area-specific crime rates.

So there is no simple explanation for the exponential growth over time in stops in the face of broad, long-term secular declines in crimes across all population groups in all places, and in the face of declining yields of legally sustainable arrests (Weiser, 2008). What then can explain the durability of a policy whose utility is weakening over time? Two possibilities come to mind. The first is that these patterns over time reflect a durable institutionalized preference to maintain these tactics even as their necessity and value is less apparent, and even as the practice's political costs mount. The practice has persisted through sharp political and legal criticism (Spitzer, 1999) and civil rights litigation against the NYPD that resulted in injunctive relief and oversight by private legal groups (*Daniels et al. v. City of New York*, 2003).

Beyond political costs, the persistence of policing tactics with disparate neighborhood impacts has salient social costs. Normative considerations—the absence of tangible returns from the policy and practice in the face of high social costs to citizens that are unevenly distributed by race and by place—suggest that the policy diminishes the social good of policing and weakens its welfarist ideology (Durlauf, 2006b), while making the job of the police harder (Skogan and Frydl, 2004; Harris, 2002). The dissipation of the social good itself has one-off costs—the withdrawal of citizens' cooperation with the police in the civic project of the coproduction of security (Tyler and Fagan, 2008; Fagan and Meares, 2007), or, in the worst case, defiance of legal and social norms (Fagan and Meares, 2007; Paternoster et al., 1997; Sherman, 1993). But such external criteria are beside the point if the preference is internalized; it need only be justified within the internal logic of the organization. Whether habit or something more, the maintenance of this policy responds to internalized incentives that remain invisible to outside observers. Its persistence requires a form of "racial blindsight" (Taslitz, 2007) to deracialize institutional recognition and acknowledgement of its consequences.

The second possibility is more mundane, and has two faces. Stops and searches of citizens are simple productivity measures for the police. Generating accurate and detailed information about stops conducted by police provides a numerical measure of police activity and outputs that is easily conveyed to citizens and oversight entities. This is especially important as crime rates decline and the traditional metrics of police productivity—arrests, crimes—no longer are sufficiently sensitive to gauge the efforts of a large and complex organization (Moore, 2002). If policing is a public good, the stop numbers provide a valuable measure of the services that produce that good.

Stops also generate a cheap form of intelligence. Intelligence was the traditional utility of the data generated in the course of stops and searches of citizens (Spitzer, 1999).[9] For years, the reports generated by stops of citizens sat in file drawers in precincts and were examined as police searched for suspects when crime patterns emerged. The information was entered into databases starting in the late 1990s, in part as a response to external investigations in reaction to political conflict following a sequence of violent, tragic, and well-publicized deaths of two citizens during encounters with the police (Spitzer, 1999). This rudimentary neural network of information was automated in the late 1990s, and has evolved into a systematic database that is one of the primary sources of information on police activity.

These institutionalized preferences, which endure in the face of persistent utility, serve the bureaucratic interests of the police hierarchy. Normative concerns over racial impacts take a backseat to the institutional interests that are indifferent to the potential for externalized costs and racial inequalities that ensue from a sustained policy with declining returns. Yet everyone has a stake in a safe society, and so security—which is primarily the province of the police—is a public good (Loader and Walker, 2007). Policing is not a discretionary service, nor is it nontrivial in the sense that it is cost-free. In New York, the cost burden of this safety—which largely accrues to White New Yorkers—is shifted to the 95 percent of African American citizens who are stopped but innocent of whatever suspected crime triggered the action. The benefits of policing—safety, calling offenders to account, conflict resolution, order, information—are social goods that are available to everyone at a low social cost, or at least at a cost that is equitably distributed. The production of this social good is not well served by the patterns we observe over the past decade of order-maintenance policing in New York.

Appendix A: Specific Police Conduct Permitted under DeBour

A. What Is a Stop?

Police stop-and-frisk procedures have been ruled constitutional under specific conditions articulated in *Terry v. Ohio* (1968). Under *Terry*, Fourth Amendment restrictions on unreasonable searches and seizures allow a police officer to stop a suspect on the street and search him or her without probable cause if the police officer has a reasonable suspicion that the person has committed, is committing, or is about to commit a crime. For their own protection, police may perform a quick surface search of the person's outer clothing for weapons if they have reasonable suspicion that the person stopped is armed. This reasonable suspicion must be based on "specific and articulable facts" and not merely on an officer's hunch.

B. Permissible Behaviors

New York law regulates police conduct more thoroughly than does *Terry*. The state law articulates a four-step analysis articulated in *People v. DeBour* (1976) and *People v. Holmes* (1996). Stops are governed by N.Y. Crim. Proc. Law § 140.50(1) (2007):

"In addition to the authority provided by this article for making an arrest without a warrant, a police officer may stop a person in a public place located within the geographical area of such officer's employment when he reasonably suspects that such person is committing, has committed or is about to commit either (a) a felony or (b) a misdemeanor defined in the penal law, and may demand of him his name, address and an explanation of his conduct."

"Stops" and "frisks" are considered separately under New York statutes. A police officer may stop a suspect but not to frisk him given the circumstances. Frisks and

TABLE 13.A1
DeBour's Four Levels of Street Encounters[a]

Predicate	Permissible Response
Level 1	Objective Credible Reason Approach to Request Information
Level 2	Founded Suspicion—Common Law Right of Inquiry
Level 3	Reasonable Suspicion Stop and (If Fear of Weapon) Frisk
Level 4	Probable Cause Arrest and Full Search Incident

[a] *People v. DeBour*, 40 N.Y. 2d 210 (1976).

TABLE 13.A2
Permissible Actions by Police Officers during Stops

Predicate	Permissible Response
Level 1	PO can ask nonthreatening questions regarding name, address, destination, and, if person carrying something unusual, police officer can ask about that. Encounter should be brief and nonthreatening. There should be an absence of harassment and intimidation. PO can: • say "STOP" (if not "forceful") • approach a stopped car • touch holster. PO cannot: • request permission to search • cause people to reasonably believe they're suspected of crime, no matter how calm and polite the tone of the questions.
Level 2	PO can ask pointed questions that would reasonably lead one to believe that he/she is suspected of a crime. Questions can be more extended and accusatory, and focus on possible criminality. PO can: • request permission to search. PO cannot: • pursue • forcibly detain.
Level 3	PO can: • forcibly detain • frisk for weapons if in fear • pull car out of traffic flow • order defendant to lie on the ground • handcuff (for good reason) • pursue.
Level 4	PO can: • arrest and search suspect.

searches are governed by N.Y. Crim. Proc. Law § 140.50(3), which requires a legitimate "stop" as a predicate to any frisk.[10] In many cases, reasonable suspicion that a person is engaging in violent or dangerous crime (such as murder, burglary, assault, etc.) will justify both a stop *and* a frisk. Table 13.A1 shows the circumstances that are necessary for a stop to escalate to a frisk and ultimately to an arrest. Table 13.A2 shows the specific police actions that are permitted at each level of a *Terry/DeBour* stop in New York.

NOTES

1. The facts of the case and its doctrinal implications have been the subject of intense interest in both constitutional criminal procedure, case law, and legal scholarship. On October 31, 1963, Cleveland police detective Martin McFadden saw two men (John W. Terry and Richard Chilton) standing on a street corner and acting suspiciously. One man would walk past a certain store window, stare in, walk on a short distance, turn back, stare in the store window again, and walk back to the other man and converse for a short period of time. The two men repeated this ritual alternately between five and six times apiece—in all, roughly a dozen trips. Each completion of the route was followed by a conference between the two on a corner, at one of which they were joined by a third man, who subsequently left swiftly. Suspecting the two men of casing the store for a robbery, McFadden followed them and saw them rejoin the third man a couple of blocks away. The officer approached the three men, identified himself as a police officer, and asked their names. When they "mumbled something" in response, McFadden patted them down for weapons and discovered that Terry and Chilton were armed. He removed their guns and arrested them for carrying concealed weapons. When the trial court denied his motion to suppress, Terry pleaded not guilty, but the Court found him guilty and sentenced him to one to three years in prison.

2. The procedure to generate a stop rationale takes place pursuant to the stop, not before, and therefore may be endogenous to the stop. Except in "radio runs," where officers are dispatched to a crime scene or location based on a citizen report or a report by another officer, and where a suspect description is provided by the dispatcher, the classification of a stop as being motivated by the match between a citizen and a "suspect description" is determined after the stop is concluded and the UF-250 form is completed. There is no method to verify the basis for the formation of suspicion for the stop. And since many stops are generated simply because the suspect "looked like a perp" (Bacon, 2009), there is considerable potential for error and theoretical misspecification. To put it less politely or scientifically, the stated rationale for the stop may in fact be either racialized, highly conditional on the conditions where the stop takes place, or simply a fiction.

3. We preferred to use both homicide arrests and homicides to test the robustness of our estimates, as well as a wider range of localized crime rates. Unfortunately, we were not privileged by the NYPD with access to its data of reported crimes that could be disaggregated to precincts, neighborhoods, and subboros. Those data were not published by the NYPD in summary form after 2001.

4. The partial correlations by year and precinct from 1984 to 2000 between homicide arrests and arrests for other Part I felony crimes was .633, and .711 for all felony crimes. For crime complaints, the partial correlation by year and precinct from 1984 to 2000 between homicide arrests and crime-specific complaints were .810 for murder, .704 for rape, .629 for robbery, and .791 for assault.

5. The stop rate and racial and ethnic distribution in these areas are:

TABLE 13.N1

Neighborhood	Stops per Household 2006	Percent African American	Percent Latino
Brownsville/Ocean Hill	.68	78	15
East New York	.65	45	38
Central Harlem	.52	71	14
East Harlem	.51	36	45
Bedford Stuyvesant	.49	72	16
Mott Haven/Hunts Point	.44	21	76

Source: New York City, Department of City Planning.

6. When arrests are made by the police upon observation of a crime, such as smoking marijuana, a stop report is completed to back-fill the case record. Accordingly, some portion of both crime complaints and stops reflect arrest-generated activity rather than independent police events.

7. In these estimates, we include Black Hispanics among Hispanics, not among African Americans.

8. Table cells are left blank in cases where the hypothesized population/stop allocations do not correspond to a "high-stop" population stopped multiple times per year. For example, in table 13.4a, the lower-left cell posits a distribution where 50 percent of the population accounts for 25 percent of the stops. If 25 percent of stops (7,135) were evenly distributed over 50 percent of the population (14,270 people), this would roughly correspond to only one-half of a stop per person. Since police stops are discrete events, an average stop rate of less than one stop per person suggests that either the "high-stop" population is overestimated, or that the portion of stops allocated to this group is underestimated. In either case, the cell is left blank, since the combination does not represent a scenario where a portion of the population is stopped repeatedly.

9. For juveniles, the parallel intelligence-gathering mechanism is the issuance of so-called YD cards to minors who are stopped by the police but not arrested. YD (for Youth Division) cards are not entered into electronic databases.

10. "When upon stopping a person under circumstances prescribed in subdivisions one and two a police officer or court officer, as the case may be, reasonably suspects that he is in danger of physical injury, he may search such person for a deadly weapon or any instrument, article or substance readily capable of causing serious physical injury and of a sort not ordinarily carried in public places by law-abiding persons. If he finds such a weapon or instrument, or any other property possession of which he reasonably believes may constitute the commission of a crime, he may take it and keep it until the completion of the questioning, at which time he shall either return it, if lawfully possessed, or arrest such person." N.Y. Crim. Proc. Law § 140.50(3).

REFERENCES

Alpert, Geoffrey, John H. MacDonald, and Roger G. Dunham (2005). "Police Suspicion and Discretionary Decision Making During Citizen Stops," *Criminology* 43(2): 407–434.

Antonovics, Kate L. and Brian G. Knight (2004). "A New Look at Racial Profiling: Evidence from the Boston Police Department," NBER Working Papers 10634, National Bureau of Economic Research, Inc.

Ayres, Ian (2002a). "Outcome Tests of Racial Disparities in Police Practices," *Justice Research and Policy*, 4, 131–142.

Ayres, Ian (2002b). *Pervasive Prejudice: Unconventional Evidence of Race and Gender Discrimination.* Chicago: University of Chicago Press.

Ayres, Ian (2008). *Racial Profiling and the LAPD: A Study of Racially Disparate Outcomes in the Los Angeles Police Department.* Los Angeles: American Civil Liberties Union of Southern California; available at http://www.aclu-sc.org/documents/view/47.

Bacon, Paul C. (2009). *Bad Cop: New York's Least Likely Police Officer Tells All.* New York: Bloomsbury.

Baker, Al (2008). "Police Data Shows Increase in Street Stops," *New York Times*, May 6; available at http://www.nytimes.com/2008/05/06/nyregion/06frisk.html.

Baker, Al (2009). "Police Stops on City Streets Rose Significantly Last Year," *New York Times*, February 10, at A29.

Bittner, E. (1970). *Functions of the Police in Modern Society: A Review of Background Factors, Current Practices, and Possible Role Models.* Rockville, MD: National Institute of Mental Health.

Bjerk, David (2007). "Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate," *Journal of Public Economic Theory* 9(3): 543–567.

Bjerk, David (2008). "How Much Can We Trust Causal Interpretations of Fixed-Effects Estimators in the Context of Criminality?" Working Paper 2008–4, Robert Day School of Economics and Finance; available at http://ssrn.com/abstract=1144201.

Blumstein, Alfred and Daniel Nagin (1978). "On the Optimal Use of Incarceration for Crime Control," *Operations Research* 26(3): 381–405.

Bowling, Ben (2007). "Fair and Effective Policing Methods: Towards 'Good Enough' Policing," *Journal of Scandinavian Studies in Criminology and Crime Prevention* 8 (S1): 17–32.

Bratton, William and Peter Knobler (1998). *The Turnaround: How America's Top Cop Reversed the Crime Epidemic.* New York: Random House.

Carbado, Devon W. (2002). "(E)racing the Fourth Amendment," *Michigan Law Review* 100: 946–1044.

Carbado, Devon W. and Mitu Gulati (2000). "Working Identity," *Cornell Law Review* 85: 1259, 1280–1283.

Close, Billy R. and Patrick L. Mason (2007). "Searching for Efficient Enforcement: Officer Characteristics and Racially Biased Policing," *Review of Law and Economics* 3(2): 263–321.

Cole, David (1999). *No Equal Justice: Race and Class in the Criminal Justice System.* New York: The New Press.

Community Studies of New York (2007). "Guide to Infoshare Online." Retrieved February 14, 2007, from http://www.infoshare.org.

Cook, Philip J. and John H. Laub (1998). "The Unprecedented Epidemic of Youth Violence," *Crime and Justice: A Review of Research* 24: 27–64.

Cook, Philip J. and John H. Laub (2001). "After the Epidemic: Recent Trends in Youth Violence in the United States." Pp. 1–37 in *Crime and Justice: A Review of Research* (M. Tonry, ed.). Chicago: University of Chicago Press.

Corman, Hope and Naci Mocan (2000). "A Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City," *American Economic Review* 90(3): 584–604.

Curtin, Kevin M., Karen Hayslett-McCall, and Fang Qiu (2007). "Determining Optimal Police Patrol Areas with Maximal Covering and Backup Covering Location Models." *Network and Spatial Economics*, online only; available at http://www.gmu.edu/departments/geog/People/Curtin/NETSOnlineFirst.pdf.

Davis, Robert C. and Pedro Matea-Gelabert (1999). *Respectful and Effective Policing: Two Examples in the South Bronx*. New York: Vera Institute of Justice.

Dharmapala, D. and S. L. Ross (2004). "Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence," *Contributions to Economic Policy and Analysis* 3(1); available at http://www.bepress.com/bejeap/contributions/vol3/iss1/art12.

Dominitz, Jeffrey and John Knowles (2006). "Crime Minimization and Racial Bias: What Can We Learn from Police Search Data?" *Economic Journal* 116 (November): F368–F384.

Donohue, John and Steven Levitt (2001). "The Impact of Race on Policing, Arrest Patterns, and Crime," *Journal of Law and Economics* 44(2): 367–394.

Duneier, Mitchell (1999). *Sidewalk*. New York: Farrar, Straus and Giroux.

Durlauf, Steven N. (2006a). "Assessing Racial Profiling," *Economic Journal* 116 (November): F402–F426.

Durlauf, Steven N. (2006b). "Racial Profiling as a Public Policy Question: Efficiency, Equity, and Ambiguity," *American Economic Review* 92(2): 132–136.

Durose, Matthew. R., Erica L. Schmitt, and Patrick A. Langan (2005). "Contacts Between Police and the Public: Findings from the 2002 National Survey," NCJ 207845, Bureau of Justice Statistics, U.S. Department of Justice.

Fagan, Jeffrey (2002). "Law, Social Science, and Racial Profiling," *Justice Research and Policy* 4 (December): 104–129.

Fagan, Jeffrey (2008). "Crime and Neighborhood Change," Pp. 81–126 in *Understanding Crime Trends* (A. Goldberger and R. Rosenfeld, eds.). Washington, DC: National Academy of Sciences, National Academies Press.

Fagan, Jeffrey and Garth Davies (2000). "Street Stops and Broken Windows: *Terry*, Race, and Disorder in New York City," *Fordham Urban Law Journal* 28: 457–504.

Fagan, Jeffrey and Garth Davies (2004). "The Natural History of Neighborhood Violence," *Journal of Contemporary Criminal Justice* 20 (2): 127–147.

Fagan Jeffrey and Garth Davies (2007). "The Political Economy of the Crime Decline in New York City." Presented at the Annual Meeting of the American Association for the Advancement of Science, San Francisco, February.

Fagan, Jeffrey, Garth Davies, and Jan Holland (2006). "The Paradox of the Drug Elimination Program in New York City Public Housing," *Georgetown Journal of Poverty, Law, and Policy* 13 (September): 415–60.

Fagan, Jeffrey and Tracey L. Meares (2008). "Punishment, Deterrence, and Social Control: The Paradox of Punishment in Minority Communities," *Ohio State Journal of Criminal Law* 6: 173–229.

Fagan, Jeffrey and Deanna L. Wilkinson (1998). "Guns, Youth Violence, and Social Identity," *Crime and Justice: A Review of Research* 24: 373–456.

Fagan, Jeffrey, Deanna L. Wilkinson, and Garth Davies (2007). "Social Contagion of Violence." Pp. 688–723 in *The Cambridge Handbook of Violent Behavior* (D. Flannery, A. Vazsonyi, and I. Waldman, eds.). Cambridge: Cambridge University Press.

Fagan, Jeffrey, Franklin E. Zimring, and June Kim (1998). "Declining Homicide in New York: A Tale of Two Trends," *Journal of Criminal Law and Criminology* 88: 1277–1324.

Farrington, David P. (1998). "Predictors, Causes, and Correlates of Male Youth Violence," *Crime and Justice: A Review of Research* 24: 421–476.

Feld, Barry C. (1999). *Bad Kids: Race and the Transformation of the Juvenile Court*. New York: Oxford University Press.

Ferguson, Andrew Guthrie and Damien Bernache (2008). "The 'High Crime' Question: Requiring Verifiable and Quantifiable Evidence for Fourth Amendment Reasonable Suspicion Analysis," *American University Law Review* 57: 1587–1644.

Fridell, Lorie (2004). *By the Numbers: A Guide for Analyzing Race Data from Vehicle Stops*. Washington, DC: The Police Executive Research Forum.

Garoupa, Nuno (1997). "The Theory of Optimal Law Enforcement," *Journal of Economic Surveys* 11(3): 267–295.

Gelman, Andrew, Jeffrey Fagan, and Alex Kiss (2007). "An Analysis of the New York City Police Department's 'Stop-and-Frisk' Policy in the Context of Claims of Racial Bias." *Journal of the American Statistical Association* 102(479): 813–823.

Golub, Andrew, Bruce D. Johnson, and Eloise Dunlap (2007). "The Race/Ethnic Disparities in Marijuana Arrests in New York City," *Criminology and Public Policy* 6(1): 131–164

Gould, Jon and Stephen Mastrofski (2004). "Suspect Searches: Assess Police Behavior Under the U.S. Constitution," *Criminology and Public Policy* 3: 315–361.

Greene, Judith A. (1999). "Zero Tolerance: A Case Study of Police Policies and Practices in New York City," *Crime and Delinquency* 45(2): 171–199.

Gross, Samuel R. and Katherine Y. Barnes (2002). "Road Work: Racial Profiling and Drug Interdiction on the Highway," *Michigan Law Review* 101: 653–754.

Harcourt, Bernard E. (1998). "Reflecting on the Subject: A Critique of the Social Influence Conception of Deterrence, the Broken Windows Theory, and Order-Maintenance Policing New York Style," *Michigan Law Review* 97: 291.

Harcourt, Bernard E. (2001). *Illusion of Order: The False Promise of Broken Windows Policing*. Cambridge, MA: Harvard University Press.

Harcourt, Bernard and Jens Ludwig (2007). "Reefer Madness: Broken Windows Policing and Misdemeanor Marijuana Arrests in New York City, 1989–2000," *Criminology and Public Policy* 6(1): 165–182.

Harris, David A. (1999). "The Stories, the Statistics, and the Law: Why 'Driving While Black' Matters," *Minnesota Law Review* 84: 265–326.

Harris, David A. (2002). *Profiles in Injustice: Why Racial Profiling Cannot Work*. New York: New Press.

Jackson, David (2000). "Winning War on Crime has a Price: Giuliani Alienates Many in New York City's Black and Hispanic Communities," *Denver Post*, Aril 20, at A23.

Karmen, Andrew (2000). *New York Murder Mystery*. New York: New York University Press.

Kelling, George and Catherine Cole (1996). *Fixing Broken Windows*. New York: The Free Press.

Kennedy, Randall (1997). *Race, Crime, and Law*. Cambridge, MA: Harvard University Press.

Kirk, David and John Laub (in press). "Neighborhood Change and Crime in the Modern Metropolis," *Crime and Justice: A Review of Research*.

Knowles, John, Nicola Persico, and Petra Todd (2001). "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy* 109: 203–229.

Kocieniewski, David (1999). "Success of Elite Police Unit Exacts a Toll on the Streets," *New York Times*, February 15, at A1.

Lamberth, John (1997). "Report of John Lamberth, Ph.D.," American Civil Liberties Union; available at http://www.aclu.org/court/lamberth.html.

Land, Kenneth C., Patricia L. McCall, and Lawrence E. Cohen (1990). "Structural Covariates of Homicide Rates: Are There Any Invariances Across Time and Social Space?" *American Journal of Sociology* 95: 922–963.

Langan, Patrick, Lawrence A. Greenfeld, S. K. Smith, Matthew R. Durose, and D. J. Levin (2001). "Contacts Between Police and the Public: Findings From the 1999 National Survey," NCJ 194957, Bureau of Justice Statistics, U. S. Department of Justice.

Levine, Harry G., and Debra Peterson Small (2008). *Marijuana Arrest Crusade: Racial Bias and Police Policy in New York City, 1997–2007*. New York: New York Civil Liberties Union; available at http://www.nyclu.org/files/MARIJUANA-ARREST-CRUSADE_Final.pdf.

Livingston, Debra (1997). "Police Discretion and the Quality of Life in Public Places: Courts, Communities, and the New Policing," *Columbia Law Review* 97: 551–625.

Loader, Ian and Nigel Walker (2007). *Civilizing Security*. Cambridge: Cambridge University Press.

Loury, Glenn (2002). *The Anatomy of Racial Equality*. Cambridge, MA: Harvard University Press.

Manning, Peter K. (2001). "Theorizing Policing: The Drama and Myth of Crime Control in the NYPD," *Theoretical Criminology* 5(3): 315–344.

Massey, Douglas S. (2007). *Categorically Unequal: The American Stratification System*. New York: Russell Sage Foundation Press.

Massey, Douglas S. and Nancy Denton (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

McCall, H. Carl (1997). *Report on the Safe Streets Program of the New York City Police Department, Re: Report 96-F-40*. Albany: Office of the State Comptroller of New York State; available at http://www.osc.state.ny.us/audits/allaudits/093097/96f40.htm.

Miller, Jeffrey (2000). "Profiling Populations Available for Stops and Searches," Police Research Series Paper 121, Home Office Research, Development and Statistics Directorate, UK.

Moore, Mark H. (2002). *Recognizing Value in Policing*. Washington, DC: Police Executive Research Forum.

Musto, David (1973). *The American Disease*. New Haven, CT: Yale University Press.

New York Civil Liberties Union v. New York City Police Department (Seeking Access to NYPD Stop-and-Frisk Database Under FOIL), New York Supreme Court, Index No. 07/115154 (Direct) (2008).

Pager, Devah (2003). "The Mark of a Criminal Record," *American Journal of Sociology* 108: 937–975.

Pager, Devah (2007). *Marked: Race, Crime, and Finding Work in an Era of Mass Incarceration*. Chicago: University of Chicago Press.

Paternoster, Raymond, Robert Brame, Ronet Bachman, and Lawrence W. Sherman (1997). "Do Fair Procedures Matter? The Effect of Procedural Justice on Spouse Assault," *Law and Society Review* 31: 163–204.

Persico, Nicola and Petra Todd (2005). "Passenger Profiling, Imperfect Screening, and Airport Security," *The American Economic Association Papers and Proceedings*, 127–131.

Persico, Nicola, John Knowles, and Petra Todd (2001). "Racial Bias in Motor-Vehicle Searches: Theory and Evidence," *Journal of Political Economy* 109(1): 203–229.

Polinsky, A. Mitchell and Steven Shavell (2000). "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature, American Economic Association* 38(1): 45–76.

Polinsky, A. Mitchell and Steven Shavell (2007). "The Theory of Public Enforcement of Law." Pp. 403–454f in *Handbook of Law and Economics* (A. Mitchell Polinsky and Steven Shavell, eds.). Amsterdam: Elsevier Science Publishing.

Quillan, Lincoln and Devah Pager (2001). "Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime," *American Journal of Sociology* 107: 717–767.

Reiss, A. (1971). *The Police and the Public*. New Haven, CT: Yale University Press.

Ridgeway, Greg (2007). "Analysis of Racial Disparities in the New York Police Department's Stop, Question, and Frisk Practices," RAND Corporation, TR-534; available at http://www.rand.org/pubs/technical_reports/2007/RAND_TR534.pdf.

Roane, Kit R. (1999). "Minority Private-School Students Claim Police Harassment," *New York Times*, March 26, at B5.

Roberts, Dorothy E. (1999). "Race, Vagueness, and the Social Meaning of Order-Maintenance Policing," *Journal of Criminal Law and Criminology* 89: 775–836.

Rosenfeld, Richard, Robert Fornango, and Andres F. Rengifo (2007). "The Impact of Order-Maintenance Policing on New York City Homicide and Robbery Rates: 1988–2001," *Criminology* 45(2): 355–384.

Rudovsky, David (2001). "Law Enforcement by Stereotypes and Serendipity: Racial Profiling and Stops and Searches Without Cause," *University of Pennsylvania Journal of Constitutional Law* 3: 296–366.

Rudovsky, David (2007). "Litigating Civil Rights Cases To Reform Racially Biased Criminal Justice Practices," *Columbia Human Rights Law Review* 39(1): 97–123.

Sampson, Robert J. and Janet L. Lauritsen (1994). "Violent Victimization and Offending: Individual-, Situational-, and Community-Level Risk Factors." Pp. 1–99 in *Understanding and Preventing Violence* (Albert J. Reiss, Jr., and Jeffrey A. Roth, eds.). Washington, DC: National Academy Press.

Sampson, Robert J. and Jeffrey Morenoff (2006). "Durable Inequality: Spatial Dynamics, Social Processes, and the Persistence of Poverty in Chicago Neighborhoods." Pp. 176–203 in *Poverty Traps* (Samuel Bowles, Steve Durlauf, and Karla Hoff, eds.). Princeton, NJ: Princeton University Press.

Sampson, Robert J. and Stephen W. Raudenbush (1999). "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods," *American Journal of Sociology* 105: 622–630.

Sampson, Robert J. and Stephen W. Raudenbush (2004). "Seeing Disorder: Neighborhood Stigma and the Social Construction of 'Broken Windows,'" *Social Psychology Quarterly* 67: 319–342.

Sanga, Sarath (2009). "Reconsidering Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 16(6), in press.

Schwartz, Alex (1999). "New York City and Subsidized Housing: Impacts and Lessons of New York City's $4 Billion Capital Budget Housing Plan," *Housing Policy Debate* 10(4): 839–877.

Shaw, Clifford R. and Henry D. McKay (1942). *Juvenile Delinquency in Urban Areas*. Chicago: University of Chicago Press.

Sherman, Lawrence W. (1993). "Defiance, Deterrence, and Irrelevance: A Theory of Criminal Sanctions," *Journal of Research in Crime and Delinquency* 30(4): 4445–4473

Silverman, Eli (1999). *The NYPD Battles Crime: Innovative Ideas in Policing*. Boston: Northeastern University Press.

Skogan, Wesley G. (1990). *Disorder and Decline: Crime and the Spiral of Decay in American Cities*. New York: The Free Press.

Skogan, Wesley G. and Kathryn Frydl (eds.) (2004). *The Evidence on Policing: Fairness and Effectiveness in Law Enforcement*. Washington, DC: National Academy Press.

Smith, Douglas A. (1986). "The Neighborhood Context of Police Behavior." Pp. 313–342 in *Communities and Crime: Crime and Justice: An Annual Review of Research* (A. J. Reiss and M. Tonry, eds.). Chicago: University of Chicago Press.

Smith, Michael and Geoffrey Alpert (2002). "Searching for Direction: Courts, Social Science, and the Adjudication of Racial Profiling Claims," *Justice Quarterly* 19: 673–704.

Smith, Michael R., M. Makarios, and Geoffrey P. Alpert (2006). "Differential Suspicion: Theory Specification and Gender Effects in the Traffic Stop Context," *Justice Quarterly* 23: 271–295.

Spitzer, Elliott (1999). *The New York City Police Department's "Stop and Frisk" Practices*. Albany: Office of the New York State Attorney General; available at http://www.oag.state.ny.us/bureaus/civil_rights/pdfs/stp_frsk.pdf.

Taslitz, Andrew C. (2007). "Racial Blindsight: The Absurdity of Color-Blind Criminal Justice." *Ohio State Journal of Criminal Law* 5(1): 1–17.

Taylor, Ralph B. (2001). *Breaking Away From Broken Windows: Baltimore Neighborhoods and the Nationwide Fight Against Crime, Grime, Fear, and Decline.* Boulder, CO: Westview Press.

Thacher, David (2008). "The Rise of Criminal Background Screening in Rental Housing," *Law and Social Inquiry* 33(1): 5–30.

Thompson, Anthony (1999). "Stopping the Usual Suspects: Race and the Fourth Amendment," *New York University Law Review* 74: 956–1013.

Tyler, Tom R. and Jeffrey Fagan (2008). "Legitimacy, Compliance and Cooperation: Procedural Justice and Citizen Ties to the Law," *Ohio State Journal of Criminal Law* 6(1): 231–275.

Tyler, Tom R. and Yuen J. Huo (2002). *Trust in the Law: Encouraging Public Cooperation with the Police and Courts.* New York: Russell Sage Foundation Press.

U.S. Census Bureau (2006). "2006 American Community Survey: Overview"; available at http://www.census.gov/acs/www/Products/users_guide/ACS_2006_Overview.pdf.

Veneiro, Peter and P. H. Zoubeck (1999). "Interim Report of the State Police Review Team Regarding Allegations of Racial Profiling," Office of the New Jersey State Attorney General.

Waldeck, Sarah (2000). "Cops, Community Policing, and the Social Norms Approach to Crime Control: Should One Make Us More Comfortable with the Others?" *Georgia Law Review* 34: 1253–1297.

Walker, Samuel (2001). "Searching for the Denominator: Problems with Police Traffic Stop Data and an Early Warning System Solution," *Justice Research and Policy* 3: 63–95.

Weiser, Benjamin (2008). "Police in Gun Searches Face Disbelief in Court," *New York Times*, May 12, at B1.

Weitzer, Ronald and Steven A. Tuch (2006). *Race and Policing in America.* New York: Cambridge University Press.

Wilson, James Q. and George L. Kelling (1982). "Broken Windows: The Police and Neighborhood Safety," *Atlantic Monthly*, March, 29–38.

Zimring, Franklin E. and Gordon Hawkins (1997). *Crime Is Not the Problem: Lethal Violence in America.* New York: Oxford University Press.

Zingraff, Matthew, H. M. Mason, William R. Smith, D. Tomaskovic-Devey, Patricia Warren, H. L. McMurray, C. R. Fenlon, et al. (2000). *Evaluating North Carolina State Highway Patrol Data: Citations, Warnings, and Searches in 1998.* Raleigh: North Carolina Department of Crime Control and Public Safety.

CASES CITED

Adams v. Williams, 407 U.S. 143 (1972).
Brown v. Oneonta, 221 F.3d 329 (2000).
Daniels et al. v. City of New York, Stipulation of Settlement, 99 Civ 1695 (SAS) (2003).
Illinois v. Wardlow, 528 U.S. 119 (2000).
People v. DeBour, 40 NY 2d 210 (1976).
Terry v. Ohio, 392 U.S. 1 (1968).
U.S. v. Harvey, 961 F.2d 1361 (1992).
U.S. v. Martinez-Fuerte, 428 U.S. 543 (1976).
U.S. v. Whren, 517 U.S. 806 (1996).

Communi
Pol

Accounting
Urban Areas in the Study of Black, White, and Hispanic Searches

Karen F. Parker, Erin C. Lane, and Geoffrey P. Alpert

Police officers' decisions to conduct searches subsequent to traffic stops are based on a number of factors including, but not limited to, their own discretion.[1] Criminologists have long explored racial disparities in police behavior, ranging from arrest to incarceration.[2] More recently researchers have suggested that race plays a role in the determination to search beyond other relevant legal factors.[3] But other studies have found no significant evidence of racial disparities in searches when taking into account hit rates[4] or the constitutionality of the search.[5] The role of race in the decision making of police officers continues to elude us.

A growing body of research is interested in understanding the link between community characteristics and police behavior at the macro level.[6] Studies have identified differential treatment of suspects by police officers relative to ecological conditions.[7] According to Terrill and Mastrofski[8] and Smith,[9] neighborhood characteristics such as concentrated disadvantage, high crime rates, and racial composition increase the likelihood that police will handle suspects more coercively. Other studies have found that officers may equate neighborhood characteristics with the populations residing in them,[10] and may use the ecological characteristics of areas as cues in decision making.[11] Unfortunately, many of these studies focus on police use of force, coercive behavior, and, only recently, traffic stops.[12] Few studies examine the relationship between the ecological conditions of neighborhoods and police search rates.[13] Because this literature is limited, it remains unclear how and to what degree community characteristics contribute to search rates of distinct groups. The lack of research is increasingly troublesome in light of the growing race and ethnicity diversity of urban communities, such as the rise in Hispanic immigration and the percentage of foreign-born residents in American cities. Hispanics are a largely understudied group, and this void is particularly noticeable in the area of police searches.[14]

Exhibit H

Chapter 7

Methods for Assessing Racially Biased Policing

Greg Ridgeway and John MacDonald

Introduction

Over the past ten years there has been a proliferation of research that has attempted to estimate the level of racial bias in police behavior. Many police agencies now mandate that their officers record official contacts made with citizens during routine traffic or pedestrian stops. These administrative data sources typically include a host of information on characteristics of the stops made by police officers, including: the race/ethnicity of the driver or pedestrian; reasons for the stop; and the actions that occurred after the stop, such as searches, contraband found, and citations or arrests made. These data have been the source for the majority of studies of racially biased police behavior. Analysts have sought to apply basic social science methods to assess whether police agencies as a whole, or in some cases individual police officers, are acting in a racially biased manner. A consistent theme in this research is the search for the appropriate benchmark[1] for which one can quantitatively assess whether police behavior is conducted in a racially biased manner. Studies have linked police administrative data on stops made by officers to a variety of data sources, including: police arrest data, population estimates collected by the Census Bureau, driver's license data, motor vehicle traffic accident data, moving violations data, systematic observations of drivers, and other sources. Analysts have also attempted to estimate racial bias from assessments of post-stop outcomes and examinations of the "hit rate" (contraband found) from searches. Post-stop outcomes have also focused on matching strategies to appropriately compare minorities and whites that were similarly situated. More recently, efforts have been made to assess individual police officer bias by peer-group officer comparisons.

In the following sections we outline the various methods that have been employed in studies of racially biased policing. We provide an overview of the use of external benchmarks, internal benchmarks, and post-stop outcomes analysis for assessing racial profiling. Our discussion is not an exhaustive review of the literature. Rather, we focus on assessing the methods, their appeal, and their substantive limitations. Developing an appropriate benchmark is more complicated than is presumed in media reports. All the methods we review for assessing racially biased policing have weaknesses, but some approaches are clearly stronger than others. There is no unifying

method that can be applied to administrative data sources and definitively answer the question of whether the police are acting with racial bias. A key issue we address is the fact that the majority of approaches used do not meet the basic bedrock assumptions necessary for drawing a causal inference about the effect of race on police behavior. Yet over time the methods have improved and the policy discussions have inevitably become more nuanced and productive, leading to discussions about what the police should and should not be using as pretexts for their decisions on whom to stop and question.

External Benchmarks

There is a compulsion in media reports on racial disparities in police stops to compare the racial distribution of the stops to the racial distribution for the community's population as estimated by the U.S. Census. For example, in 2006 in New York City, 53% of stops police made of pedestrians involved black pedestrians while according to the U.S. Census they compose only 24% of the city's residential population. When the two racial distributions do not align, and they seem to do so rarely, such statistics promote the conclusion that there is evidence of racial bias in police decision making. Racial bias could be a factor in generating such disparities, but a basic introductory research methods course in the social sciences would argue that other explanations may be contributing factors. For example, differences by race in the exposure to the police or the rates of committing offenses may also contribute to racial disparities in police stop decisions. It is well documented, for example, that due to historical differences in racial segregation, housing tenure, poverty, and other sociopolitical factors, minorities in the United States are more likely to live in neighborhoods with higher rates of crime and disorder.[2] Police deployment in many cities also corresponds to differences in the demand for police services. Neighborhoods with higher volumes of calls to the police service typically have a higher presence of police.[3] Additionally, research indicates that racial minorities, and in particular blacks, are disproportionately involved in serious personal offenses as both victims and offenders.[4]

The crux of the external benchmarking analysis is to develop a benchmark that estimates the racial distribution of the individuals who would be stopped if the police were racially unbiased, and then compare that benchmark to the observed racial distribution of stopped citizens. The external benchmark can be thought of as the population at risk for official police contact. As we will see, estimating the appropriate population at risk is complicated. Crude approximations of the population at risk for police contact are poor substitutes and can hide evidence of racial bias or lead to exaggerated estimates of racial bias.

The likelihood of police stopping minority drivers involves some combination of police exposure to offending/suspicious activity, the racial distribution of the population involved in those activities, and the potential for racial bias. To provide some context, we use some hypothetical numbers and consider an unbiased officer on a foot post who makes stops only when a pedestrian matches a known-suspect description. This officer works in a precinct with forty blacks matching suspect descriptions forty whites matching suspect descriptions. If we could somehow measure such numbers we would be inclined to propose a suspect-description benchmark of 50% black and 50% white. But if the routine daily activities of whites and blacks differ, then the officer will encounter different proportions of suspects by race. Say, for example, that the majority of the forty white suspects stay inside most of the day, travel only by car, or avoid the specific areas with high police presence; then this officer will stop only a small number of white suspects, deviating substantially from the 50% benchmark. Even the less extreme situation, in which half the white suspects are exposed to the officer, results in the officer stopping blacks in 67% of all their stops decisions. The suspect benchmark in this context is only valid if the police are equally exposed to suspects from the various racial groups. Therefore, even with unbiased officers, we cannot necessarily expect what seems like a reasonable external benchmark to match the racial distribution of stops. This example effectively demonstrates that any of the external benchmarks described in this section must be viewed with caution.

The primary reason for using U.S. Census data to form the benchmark is that it is inexpensive, quick, and readily available. A number of studies attempting to assess racial bias in police behavior use population data from the census, and some rely on estimates at local-area levels like neighborhood census tracts (see Parker and colleagues in this volume). For the reasons previously listed, however, benchmarking with census data does not help us isolate the effect of racial bias from differential exposure and differential offending. Even refinements to the residential census, such as focusing on subpopulations likeliest to be involved in crime (e.g., men or driving-age young adults) are not likely to eliminate differences in the exposure of officers to criminal suspects or provide a good approximation of the population at risk for official police action. Fridell[5] summarized the problem with using the census as a benchmark with regard to offender exposure by noting that "this method does not address the alternative hypothesis that racial/ethnic groups are not equivalent in the nature and extent of their . . . *law-violating behavior*" (p. 106, emphasis in original).

Census estimates provide only the racial distribution of residents and not how these numbers vary by time of day, business attractors such as shopping centers, daily traffic patterns involving commuters, and so forth. It is quite conceivable that the residential population in many neighborhoods has little resemblance to the patterns of people on the street during the day or night. Even if refinements in the census to the neighborhood or age-prone population at risk for police involvement could give a racially unbiased estimate, the differences between the residential population and the population at different times of the day and street segments are likely to overwhelm such a determination. Commuting patterns, for example, can easily exaggerate the racial disparities in traffic stops. Imagine that 20% of traffic stops in a neighborhood that is 95% nonwhite are made of white citizens. In this context we would suggest whites are stopped four times the rate of their composition of the neighborhood population (20/5 = 4) and are subjects of racially biased police behavior. But the stop rate may be a simple reflection of the fact that daily commuters reflect 20% of drivers in this neighborhood.

Dissatisfaction with the census as a benchmark has led some researchers to develop alternate external sets of benchmarks. Some studies of traffic stops attempt to acquire more precise estimates of the racial distribution of drivers on the road to serve as the external benchmark. Under such an approach, one should be able to compare the race distribution of traffic stops made by the police to the race distributions of drivers on the same roadways. Zingraff and colleagues,[6] for example, used the race distribution of licensed drivers rather than the residential population to estimate the race distribution of drivers at risk of being stopped by the police. Although this approach accounts for racial differences in the rate at which the population holds driver's licenses, it does not account for out-of-jurisdiction drivers or for potential racial differences in travel patterns, driving behavior, or exposure to police. To address the problem with out-of-jurisdiction drivers, Farrell and colleagues[7] borrowed driving population models from the transportation literature, which use an area's ability, based on employment or retail location, to pull drivers in from outside communities or to push residents outside the area. This certainly improves on the census benchmark. But it is widely documented that minorities (and even those who possess a driver's license) are more likely to take public transit to work and vary from whites in other important ways in their daily travel patterns. Therefore, a more accurate external benchmark would be one that could reliably take into account equivalent driving patterns and behavior between race groups.

Recognizing these limitations, Alpert and colleagues[8] used data on the location of traffic accidents and the race of the not-at-fault drivers to estimate the race distribution of the at-risk population. The logic of this approach is that the race distribution of not-at-fault drivers should approximate the racial distribution of the population of drivers. Although this approach may measure the race distribution of drivers on the road, it does not account for potential racial differences in driving behavior that may be important sources for police decision making, such as the likelihood of speeding, weaving through traffic, and driving slower than usual.

Other analysts have studied the race distribution of drivers flagged by photographic stoplight enforcement cameras[9] and by aerial patrols.[10] The advantage of these benchmarks is that they are truly race-blind and measure some form of traffic violation. One can question whether they capture race differences in other aspects of stop risk, such as seatbelt usage, equipment violations, and the other cues that police use in deciding whether or not to stop a citizen.[11]

Given that the police are not likely to stop people at random, comparisons of racial distribution of stops to the residential population or the driving population on the roadways tells one very little about the race neutrality of the police. Again, it is necessary to establish a benchmark for the population at risk for official police contact. This means that one needs an accurate estimate of the subpopulation that is likely to elicit reasonable suspicion by the police.

Observation Benchmarks

Observation benchmarks are a popular approach for attempting to estimate the subpopulation at risk for police behavior. Observation benchmarks typically involve fielding teams of observers to locations to tally the racial distribution of those observed driving and violating traffic laws. More than three decades ago Albert Reiss Jr.

advocated the use of systematic social observation as a key measurement strategy for studying the police and other social phenomena.[12] By systematic, he meant that the observation of behaviors and recordings are done according to explicit standardized rules that permit replication.

This methodology was pioneered to study racial bias in police traffic stops by Lamberth[13] in his study of the New Jersey Turnpike. Observation benchmarks' greatest potential occurs in its application to racial profiling on freeways, since vehicles have essentially the same exposure to the police, and speeding is the primary violation that highway patrol focuses on. Speeding, for example, accounted for 89% of the stop reasons in a subsequent study of New Jersey Turnpike traffic stops.[14] Measuring speeding through direct observations with radar guns, for example, provides a standardized approach that is easy to replicate and less subject to measurement error than accounting for other types of traffic violations that require observers to make judgments about infractions like weaving through traffic or making illegal turns. Lang and colleagues[15] and Alpert and colleagues provide two case studies using radar guns.[16] The main wrinkle in the analysis of benchmarks based on observation of speeding is determining the appropriate speed at which drivers should be considered "at risk" for being stopped in specific sections of the highway. For example, it is conceivable that in some areas the police are more vigilant with speeding. As long as this variation is not confounded with differences in the areas that minorities and whites travel then it can provide an unbiased assessment of racial disparities in highway traffic stops.

In urban environments, however, officers stop vehicles for a variety of reasons beyond simple moving violations. Exposure to police can vary widely across different geographic segments of the city.[17] In the current volume the reader will note that a number of authors attempt to take the intra-city variation in exposure to the police into account (see, e.g., Fagan and colleagues). Eck and colleagues[18] note that in Cincinnati, Ohio, the police allocate a greater share of officers to areas with a higher volume of crime incidents, and these areas happen to be composed predominantly of black residents. Relying on direct observations of traffic violations in different segments of Cincinnati would not provide an unbiased assessment of the population at risk for police exposure, because race is confounded with the areas that police are concentrated on. One would have to develop an observation method that appropriately balanced these differences in police resource allocation.

There are few examples where investigators have attempted to take the complexity of geographic areas of a city into account in using observation methods. Alpert and colleagues[19] provide one of the few published studies where trained observers recorded traffic violations (e.g., illegal turns, running stop lights, speeding) at sixteen high-volume intersections in Miami-Dade County in areas that were classified as predominately white, black, or racially mixed. A comparison of the racial distribution of observed traffic violators to actual police traffic stops in the same areas suggested little evidence of racial bias in stop decisions. Even if observers in this study did produce an accurate benchmark for individuals at risk for exposure to the police in these areas—a challenge in its own right—several issues remain. There is no reason to believe that police stops should be representative of those simply observed committing

traffic violations in these areas. Officers target behaviors that they believe indicate drug transactions, stop individuals fitting suspect descriptions, and respond to calls for service. Once observers head down the path of trying to determine which vehicles or persons should be at risk for being stopped, the observations become more subjective and less systematic.[20] In fact, the variation between observers in such studies can exceed the estimate of the racial disparity. One observer may be more likely than others to measure some driving behavior as aggressive. Such variation in judgments in an observation study has to be taken into account, or observers have to be trained to near uniformity in judgments if one is going to produce a reliable estimate of the population at risk for police contact. Regardless, it is unclear that observational studies are relying on the same sets of markers that the police use in deciding who is suspicious and whom to stop. The courts have not consistently supported the use of observational benchmarks for this reason. In United States v. Alcaraz-Arellano[21] the court rejected the benchmark, since it was developed for a general population, not those violating the law.

Outside of traffic stop studies on speeding or moving violations on roadways, systematic observations of driving behavior are not likely to yield useful estimates for an external benchmark for an entire city. Recognizing these limitations, a number of investigators have turned to other approaches for establishing external benchmarks.

Arrest and Crime Suspect Benchmarks

Gelman, Fagan, and Kiss[22] quote then NYPD police commissioner Howard Safir: "The racial/ethnic distribution of the subjects of stop and frisk reports reflects the demographics of known violent crime suspects as reported by crime victims. Similarly, the demographics of arrestees in violent crimes also correspond with the demographics of known violent crime suspects." Safir is clearly suggesting that violent crime suspects or violent crime arrestees provide a reasonable benchmark from which the public can judge the department's racial distribution in stop percentages. This quote suggests that the arrestee population may serve as a useable benchmark for assessing racial bias in the police decision of whom to stop.

The arrestee benchmark, however, is also problematic because it is too narrow. For example, the police make stops for trespassing, vandalism, suspected drug sales, and a variety of other causes. Many stop decisions might be made for minor infractions, not serious crime incidents involving violence. The group of individuals stopped by the police in most large cities, therefore, far exceeds the group comprising the arrestee population. There are a variety of reasons that the racial distribution of individuals stopped by the police could have a racial distribution that differs greatly from that of arrestees. For one, arrests can often take place some distance away from where the crime actually occurred. Most problematic is that if officers are in fact racially biased, then we cannot use their arrests to represent what we would expect of an unbiased police force. Such a benchmark could actually hide bias. Investigators like Gelman and colleagues have attempted to control for this by using prior-year arrest decisions as an external benchmark. Again, there is no reason to expect that prior-year decisions are independent of current-year decisions—especially if, as research by Klinger[23] suggests, an established pattern of practices becomes ingrained in specific police precincts.

The criminal suspect benchmark may be a more plausible approach than the arrestee benchmark for establishing the population at risk for official police contact. It represents the public's reporting of those involved in suspicious activity and crime and would correspond more closely to racial distribution of criminals on the street.[24] Note that this benchmark is not a reasonable choice for traffic stops since police often have the intent to cite for a traffic violation without the expectation that it will lead to an arrest. Comparing the police to the public's reporting of suspicious activity at least answers the question of whether the police are finding suspicious individuals with features similar to those the public reports committing or attempting to commit crimes. Ridgeway, for example, found that in New York City black pedestrians were stopped at a rate 20 to 30% lower than their representation among the public's report of crime-suspect descriptions, and Hispanic pedestrians were stopped slightly more than their share of crime-suspect descriptions, by 5 to 10%.[25] The public may have their own racial biases, however, and they may also under- or overreport certain activities (e.g., drug market activity, suspicious individuals) depending on the area and the perceived problems that the police actively target.

Instrumental Variables

An ideal scientific method to estimate the extent of race bias in policing would be to use an experimental design and randomly assign police officers to be "race-blind" during certain periods. For example, for each officer and for each hour that an officer patrols the street, we flip a coin to determine whether that officer will be unable to perceive the race of a suspect. The difference between the percentage of stops involving minorities when the officers can perceive race and the percentage of stops involving minorities when the officers are race-blind gives us the effect of racial bias. If the officers were unbiased then the ability to perceive race should not matter in the selection of stopped individuals. If instead the officers are racially biased then we would observe more minority stops when the officers are not blinded to race.

Clearly such an experiment in the actual field is a fantasy, but instrumental variables (IV) analysis is an econometric approach that can sometimes solve such problems.[26] Instrumental variables analysis relies on the randomization that occurs in nature to replicate the classic randomized experimental design. The key hurdle is to identify an "instrument," in this case a variable that is predictive of the ability to perceive race,[27] that is not related to the actual race of suspects.[28] This is a generalization of the setup in the previous paragraph where our coin is the instrument, highly predictive of the ability to see race but unassociated with the race of potentially stopped individuals.

Grogger and Ridgeway[29] proposed as an instrument the natural variation in daylight and darkness that switches with the change in daylight savings. It is associated with the ability to perceive race but is not related to the race of drivers on the road. The randomization in nature that diminishes the ability of officers to view the actual

for assessing racial bias in police traffic stops. Presumably the probability of race being visible is greater in daylight. Besides the logic of the statement, there is some evidence from the literature supporting this. Lamberth described a traffic survey in which the driver's race could be identified in 95% of the vehicles, but for which nighttime observations required auxiliary lighting.[30] Greenwald canceled plans for evening surveys after his observer could identify the race of only 6% of the drivers viewed around dusk.[31]

The logic of this approach goes back to the work of Neyman[32] in the 1920s and is a special case of more general instrumental variable methods. We first have to determine the percentage of black drivers among those stopped during daylight and the percentage of black drivers stopped during darkness. Second, to account for the fact that sometimes race is not visible during the day and can be visible at night, the difference in the percentage of blacks stopped needs to be divided by the difference in the probability of race being visible in daytime and darkness. Importantly, this estimate does not require complete race blindness at night and complete visibility during the day, only a substantive diminished capacity.

One of the difficulties that Grogger and Ridgeway faced when attempting to estimate this instrumental variable is that there is no direct measure of diminished capacity due to changes in daylight, the second step of the described IV estimator. A controlled scientific experiment could be conducted to estimate visibility by daylight and darkness, but this might not reflect the types of lighting situations that officers commonly experience on the streets, especially in parts of the city that are better lit than others. As a result Grogger and Ridgeway's analysis simply assumed, logically, that the denominator is positive, such that the probability of race being visible is greater in daylight.

The validity of this instrument also depends on race being independent of daylight/darkness visibility. But the race distribution of drivers on the road and exposed to the police may be quite different between daylight hours and nighttime hours. If there were mostly black drivers on the road at night then the analysis would indicate that officers stop an excessive fraction of black drivers during the night, but this would just be because there are a larger proportion of black drivers on the road at that time. To correct this potential problem, Grogger and Ridgeway controlled for clock time and compared stops occurring near the changes to and from daylight savings time. On one Monday stops at 6 pm occur in daylight and the following Monday stops at 6 pm occur in darkness. If we can assume that the race distribution of drivers on the road at 6 pm does not change with daylight savings time and that the police do not suddenly reallocate their officers, then this provides a valid instrument.

Figure 7.1 demonstrates the idea using data from Oakland, California. The horizontal axis indicates the clock time and the vertical axis indicates hours since dark. Throughout the analysis, we omit stops carried out during the roughly thirty-minute period between sunset and the end of civil twilight, since that period is difficult to classify as either daylight or dark. The solid points indicate stops of black drivers, whereas open circles represent stops of nonblack drivers. At any time between 5:19 pm and 9:06 pm, some stops are carried out when it is ...

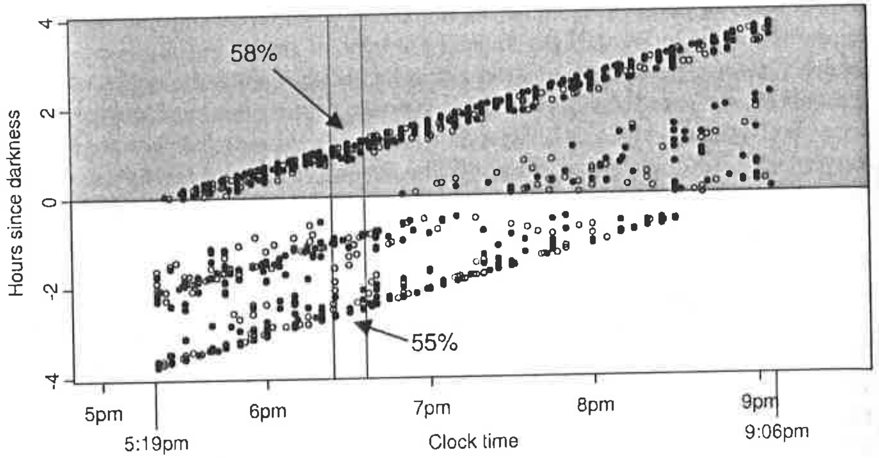


Figure 7.1

variation in daylight hours over the course of the study period. In particular, the large diagonal gap is a result of the shift from Pacific Daylight Time to Pacific Standard Time at the end of October. This shift is especially useful for our comparison since it creates extremes in visibility for fixed clock times. The vertical lines in figure 7.1 mark a period around 6:30 pm within which we can assess whether darkness influences the race of drivers stopped. During daylight hours, 55% of the stops involved black drivers, while stops after dark involved black drivers in 58% of the stops, a slight difference that, if anything, runs counter to the racial profiling hypothesis. Schell, Ridgeway, and colleagues provide a similar analysis of three years of traffic stops in Cincinnati and find similar null conclusions against racial bias in traffic stop decisions.[33]

The instrumental variables approach here, however, does have limitations. First, this method assumes that the variation in daylight/darkness gives enough of a diminished capacity to effectively remove the importance of a suspect's race in the decision of whom to stop. If the police use car profiles, such as stylistic rims or other features that are correlated with race and social class, as the primary proxy for race, then this approach will still yield an unbiased test of the race effect on police decisions but will be greatly underpowered because police will use these cues regardless of the level of daylight/darkness. Even if such proxies do not exist, the approach only measures the effect of race bias at those times of day that are sometimes light and sometimes dark. Since there is never daylight at 3 am, we cannot estimate an effect of race for stops that occur at that hour.

Internal Benchmarking

Recognizing the difficulty of assessing whether racial bias occurs on the aggregate in the decision to stop citizens has led some analysts to focus on the individual decision

making of police officers. The decision to stop a citizen is only one stage in the traffic stop process, at each stage of which police officers can introduce race bias in their decisions. Highly publicized examples of racial bias in police behavior can give an impression of systemic bias, even if the source of bias is only a few problem officers[34] (see Weitzer in this volume).[35] The Christopher Commission in its assessment of abuse of police authority among the Los Angeles Police Department (LAPD), for example, noted that 10% of officers accounted for 27.5% of complaints of excessive force and 33% of all use-of-force incidents.[36] The methods described previously, which attempt to examine bias at the departmental level, are unlikely to detect the problem if the source is a small share of individual officers, and, even if somehow there are enough biased officers to create enough statistical power to detect the problem at the department level, these previous methods do not identify potential problem officers.

Walker[37] conceptualized the internal benchmark, a framework that compares officers' stop decisions with decisions made by other officers working in similar situational contexts. This method has been applied to department data in several localities and has been adopted as a part of several "early warning systems."[38] At the LAPD, the TEAMS II Risk Management Information System places officers in one of thirty-three peer groups.[39] Officers in the same peer group presumably are expected to conduct similar policing activities. If an officer exceeds certain thresholds for their peer group, such being in the top 1% on number of complaints or number of use-of-force incidents, the system generates an "action item" for follow-up. Officer roles in LAPD, however, are certainly more diverse than thirty-three groups can capture. Similar problems are likely in other audit systems that compute a "peer-officer-based formula" to flag officers[40] but do not fully take into account the variation in environments in which officers in the same peer group work. Sometimes the peer group construction may be reasonable. For example, Decker and Rojek[41] matched each St. Louis police officer to all other officers working in the same police districts. It is unclear whether matching by district alone was sufficient to ensure validity, although they argued that officers rotated shifts sufficiently so as not to warrant concern.

While this process is useful for flagging potential problem officers, it has some drawbacks. First, if officers in the entire precinct are equally biased, the method will not flag any officers as being problematic. We must rely on other analyses to assess that issue. Second, officers whom the method flags as outliers may have legitimate explanations for the observed differences. For example, a Spanish-speaking officer may appear to make an excessive number of stops of Hispanic suspects, when, in fact, the Spanish-speaking officer gets called in to handle and document those stops. Such situations should be detectable when supervisors review cases. Otherwise, the method eliminates possible explanations based on time or place, so the range of explanations is limited.

The fundamental goal of internal benchmarking is to compare the rate of nonwhite-pedestrian stops for a particular officer with the rate of nonwhite-pedestrian stops for other officers patrolling the same area at the same time. Matching in this way assures us that the target officer and the comparison officers are exposed to the same set of offenses and offenders.

Ridgeway and MacDonald[42] developed an internal benchmark methodology to compare the racial distribution of pedestrians/drivers whom individual police officers have stopped with that of pedestrians/drivers whom other officers in the same role have stopped at the same times and places. This method has been applied in case studies in both Cincinnati[43] and New York City.[44] Utilizing an approach based on propensity score weighting, doubly robust estimation, and false discovery rates, these case studies attempt to customize the internal benchmark for each individual officer to a set of officers working in similar environments exposed to similar suspects, and to control the risk of too many officers being flagged as outliers (false positives). The

TABLE 7.1

Construction of an Internal Benchmark for a Sample Officer

Stop Characteristic	Officer A (%) (N=392)	Internal Benchmark (%) (N=3,676)
Month		
January	3	3
February	4	4
March	8	9
April	7	5
May	12	12
June	9	9
July	7	7
August	8	9
September	10	10
October	11	10
November	11	11
December	9	10
Day of the week		
Monday	13	13
Tuesday	11	10
Wednesday	14	15
Thursday	22	21
Friday	15	16
Saturday	10	11
Sunday	15	14
Time of day		
12–2 am	11	11
2–4 am	5	5
10 am–12 pm	0	1
12–2 pm	12	13
2–4 pm	13	12
4–6 pm	9	10
6–8 pm	8	8
8–10 pm	23	23
10 pm–12 am	17	17
Precinct		
A	0	0
B	98	98
C	1	1
D	1	0
Occurred inside?	4	6
Housing or transit		
Transit	0	0
Housing	0	0
Other	100	100
In uniform		
Yes	99	97
Radio run		
Yes	1	3

Note: The numbers in the table indicate the percentage of stops having that feature.

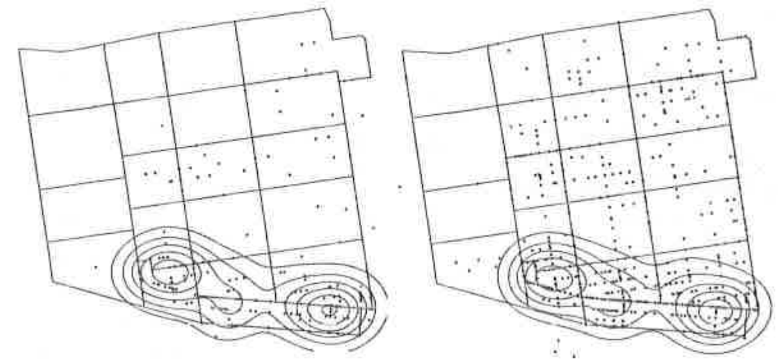


Figure 7.2

first of the three stages in this process is, for each officer, to reweight the stops made by other officers so that they have similar stop characteristics distributions.

Table 7.1 shows the results of this reweighting step for an example officer. Officer A made 392 stops. The method effectively identified 3,676 similarly situated stops made by other officers. These stops were selected as the benchmark group for Officer A because they were similar to Officer A's stops in terms of when they occurred (e.g., date, time of day), where they occurred (e.g., precinct, x-y coordinates), the assigned command of the officer making the stop, whether the officer making the stop was in uniform, and whether the stop was a result of a radio run. Figure 7.2 and table 7.1 demonstrate that this collection of 3,676 is nearly identical to the officer's stops in several respects. Furthermore, as shown in figure 7.2, the distribution of the locations of the stops can be aligned geographically so that regions of this officer's stops in 2006 can be compared to other officers making stops in the same region. An additional adjustment at this stage can improve the precision of this test. The second step of the process involves a regression model to further refine the benchmark, since some features are not perfectly matched between officers in table 7.1, such as the frequency of being in uniform and being on a radio run.

Combining propensity score analysis with a second stage regression model has recently been labeled "doubly robust estimation," since if either the propensity score weights construct a well-matched set of benchmark stops or the regression model is correctly specified, then the resulting estimate of the officer's effect on the race of those stopped can be consistently estimated.[45]

The z-statistic from these regression models is the commonly used statistical measure for assessing the magnitude of the difference between an officer's minority-stop fraction and the officer's internal benchmark group. The z-statistic scales the difference between the officer and his or her internal benchmark such that large differences based on a small number of stops are treated with greater uncertainty than large differences based on a large number of stops. Fridell[46] suggests 2.0 and Smith[47] suggests 1.645 as the appropriate z-scores to flag potentially problematic officers. But such cutoffs generate too many false positives to be useful and are one of the sources of problems for LAPD's system. In a department of one thousand officers we can expect fifty of them to have z-statistics in excess of 1.645 by chance alone.

Methods based on false discovery rates (fdr) helps address this kind of problem.[48] The fdr is the probability of no difference between the officer and the benchmark given the value of an observed test statistic, z. We should flag those officers who have values of z that suggest a low probability of being incorrectly flagged as a problem. When applied in Cincinnati this approach noted four potentially problematic officers, and in New York City fifteen potentially problematic officers.

Internal benchmark approaches provide a method for assessing individual officer bias. Again, the key to this approach is developing a reasonable peer group or comparison set of officers. This approach, however, is limited to departments with officers that make many stops. If officers make few stops (e.g., less than fifty), then chance differences from their benchmark are likely and the comparisons are underpowered. Accumulating stops across years can improve this. For departments with few officers (e.g., those with less than 100 officers), the fdr calculations become more unstable and more dependent on statistical assumptions.

Post-Stop Outcomes

The complexity of benchmarking for assessing bias in the decision to make a stop has in some cases caused analysts to abandon the endeavor in favor of assessing bias in post-stop outcomes, such as duration of the stop, decision to search, and use of force. This has its advantages, since for this analysis we have a better assessment of the race distribution of who is at risk. But substantial complexity remains.

Auditing Police-Citizen Interactions

An obstacle to understanding racial disparities in police decision making is that stopped drivers and pedestrians cannot observe how officers handle other stops, particularly those involving members of another race. They cannot answer the most pertinent question regarding racially bias policing: Would the same outcome have occurred if I had been a different race? While such counterfactual questions so far have not been answered, recordings of stops can provide some guidance to understanding the dynamics in police-citizen interactions.

Dixon and colleagues[49] used a stratified random sample of 313 vehicle-mounted video and audio recordings from Cincinnati Police Department (CPD) cars to study interactions between police and community members. The study described how the race of the driver and the race of the officer influenced the dynamics of stops, including stop features associated with "counterproductive or dissatisfying interactions," and described how typical police-motorist interactions occur as a function of race.

Among the results reported in this study is the finding that in interactions where the officer and driver are of the same race, officers are more likely to be interested in hearing the drivers' comments. The key problem that this creates in Cincinnati is that, since many more CPD officers are white, two-thirds of stops of black drivers

involve a white officer while only one-third of stops of white drivers involve a black officer. Thus the impact of degraded communication due to interracial stops will be greatest for the black drivers.

Additional research by the same research team[50] found that white officers conducted more investigative stops (e.g., asking questions about guns or drugs, asking for the IDs of passengers) while black officers were more likely to focus on the traffic infractions alone. Importantly, these differences did not depend on the race of the driver. That is, white officers also closely investigated white drivers. Such differences between white and black officers, however, can exacerbate the perception of racially biased policing. The black driver in Cincinnati who experiences one stop with a black officer and another stop with a white officer is likely to attribute the white officer's more intense investigation to race bias, even though on average this white officer treats blacks and whites with a similar level of scrutiny.

The analysis of recorded interactions is useful at identifying problem interactions, factors that can contribute to the perceptions of race, and stops that could be useful in training. But such methods do not answer the question of whether the police use race as a factor in deciding whom to stop.

Hit Rates

Hit rates, the percentage of conducted searches that turn up contraband, have been a frequently discussed outcomes test for racial equity in searches. If the hit rate for searched nonwhite suspects is less than the hit rate for searched white suspects, police might be applying a lower standard of suspicion to nonwhite suspects when deciding whether to search.

A series of papers by Persico and Todd[51] provide the theory and empirical examples of the use of hit rates with police traffic stop data. Relying on the premise of a Nash equilibrium, these authors argue that hit rates provide a race-neutral test of bias in police decision making because police decisions about which suspects to search take into account the benefits of searching different suspects, and suspects take "into account the risk of getting searched" (p. 37).[52] If officers and criminals act as rational agents, then the outcome of stops should be race neutral. Following on the logic of a Nash equilibrium that officers want to maximize their ability to find illegal contraband in traffic stops, and suspects want to reduce their likelihood of being caught, then the probability of successful "hits" should be equal once one conditions on the race of who is stopped. If, for example, police officers want to find illicit drugs and suspects want to avoid detection, the results for searches among police officers who are intentionally biased toward blacks will be offset by a higher yield of searches among whites. In the long run the differences between races in hit rates should equalize. Persico and Todd's analysis of Maryland State Police traffic stop data in several publications reports findings that the fraction of blacks stopped exceeds the fraction of black motorists on the road, but that the hit rates across racial groups are statistically equivalent.

We, however, provide an example to demonstrate that a simple comparison of hit rates can distort the true racial differences. Assume that suspects are stopped for either burglary or robbery. Further assume that there is no racial difference in the rates at which suspects carry contraband and that police are racially neutral in making stop-and-frisk decisions (essentially blind to race). Last, consider the information shown in table 7.2. Within a crime category, hit rates are equal for black and white suspects. In this example, officers detain many more white suspects on suspicion of robbery, a crime with a higher hit rate, than they do black suspects, who are more likely to be stopped for burglary. In this example, though, those large differences in the rates of stops for burglary and robbery by race are due not to officer bias but are the result of racial differences in criminal participation. As a result, the total hit rate for white suspects is 4.6% ([1+45]/1,000), and for black suspects, 1.4% ([9+5]/1,000).

One could conclude from these two numbers (4.6% vs. 1.4%) that there is racial bias in the decision to search suspects, and that whites are not searched at sufficient rates. But officers in this hypothetical example are race-neutral by design. Hit rates are equal across races for suspected burglars and robbers. This is a reminder that failing to account for an important factor—suspected crime, in this example—can distort the conclusions. In practice, the only way for the Nash equilibrium as described by Persico and Todd to work would be if black burglars and white robbers adjusted their criminal behaviors to mirror each other because they had equal probability of being stopped by the police.

This example illustrates a statistical problem that Ayres[53] termed the subgroup validity problem, in which a particular relevant feature is more prevalent for certain racial groups. Other factors may affect the hit rate as well. Officers in some precincts may be likelier to frisk, due to crime in the area, recent surges in weapon recoveries, a series of recent shootings, or more hostile attitudes displayed by suspects. An elevated frisk rate in some precincts may not meet with the community's approval, but it would be premature to attribute this variation to racial bias by police officers without examining other relevant factors. Therefore, it is critical to account for factors correlated with race that might be associated with both suspect race and the rate of contraband recovery.

In Ridgeway's analysis of hit rates in New York City, shown in table 7.3, white and Hispanic suspects stopped in situations that were similar to the collection of black suspects had hit rates of 3.2 percent and 3.8 percent, respectively, compared with a hit rate of 3.3 percent for black suspects.[54] There was no statistical evidence for a difference between these recovery rates. Furthermore, there were no differences in the rates at which officers found weapons on suspects. The unadjusted hit rates, however, suggested evidence of bias—again showing that it is important to adjust for subgroup

TABLE 7.2
Hypothetical Example of a Hit-Rate Analysis

Race	Measure	Burglary	Robbery
White	Stopped and frisked	100	900
	Had contraband (%)	1	5
	Had contraband	1	45
Black	Stopped and frisked	900	100
	Had contraband (%)	1	5
	Had contraband	9	5

TABLE 7.3
Frisked or Searched Suspects Found Having Contraband or Weapons

	Black	Hispanic	White
Any contraband	3.3	3.2	3.8
Weapon	0.7	0.7	0.8

differences in the circumstances by which different racial groups are subjected to police authority.

It is plausible that the carry rates, the percentage of stopped suspects that have contraband, differ by race. If white suspects simply carry drugs more frequently, perhaps believing that officers are unlikely to search them, then the contraband recovery rates for white suspects will be higher. Persico and Todd theorized from the logic of a Nash equilibrium that criminals will assess their risk of being searched and adjust their frequency of carrying drugs and weapons accordingly, so that an outcome test will be race-neutral. It is difficult to confirm this in practice, and, as a result, conclusions drawn from table 7.3 must allow for the possibility that carry rates are not uniform across racial groups.

Analysis of Other Stop Outcomes

Other analysts have focused on developing appropriate benchmarks for studying the stop outcomes themselves. In Cincinnati, for example, Ridgeway[55] notes that 47% of stops involving black drivers lasted less than ten minutes while 56% of stops of nonblack drivers lasted less than ten minutes. On the surface this seems to be a rather large bias. But 18% of the stopped black drivers did not have valid driver's licenses while only 5% of nonblack drivers did not have valid licenses. As a result, we cannot discern whether the disparity in stop duration is attributable to the driver's race or to the additional time required to process a stop involving an unlicensed driver.

Social scientists recognize that adjusting for confounding variables is a critical step in all proper analyses, and there are clear examples in the current book where analysts attempt to make such adjustments (see Fagan et al., and Parker and colleagues in this volume). Particular to racial profiling analyses, police may approach vehicles more cautiously and conduct pat searches for weapons in high-crime neighborhoods during peak crime times (e.g., late evening on the weekends). These decisions may occur regardless of the driver's race, but may be confounded with race due to differences in the neighborhoods in which minorities and whites live. In high-crime neighborhoods police also may be more thorough in checking for vehicle registration and driver's license records, have a longer list of recent suspect descriptions that the stopped driver may match, and may be more likely to develop probable cause. In theory and practice, all these decisions could be independent of the driver's race. As a result, the stop location and time may influence all the measured post-stop activities even in the absence of a race bias. When the race distribution of drivers differs by time and neighborhood location, one should adjust for these differences when assessing racial bias in post-stop activity. The analysis also might adjust for other features occurring after the stop, such as whether the suspect had an open warrant or a suspended driver's license.

Location and time of the stop are two among a number of factors for which post-stop activity might vary that are confounded with race of drivers or pedestrians stopped by the police. While these differences may be structurally discriminatory based on racial differences in areas that individuals live, they may not be substantively discriminatory based on police decision making.

The common practice of "adjusting for" potentially confounding factors with multivariate regression is difficult to defend in the analysis of post-stop data. The regression adjustment is only effective if there is not a strong correlation between race and the other variables in the regression model. If in the case of citizen stops the distribution of stop features of blacks differs substantially from the distribution of stop features of whites by neighborhood, type of violation, time of day, and so forth, it is uncertain whether the estimate of the race effect on police post-stop outcomes sufficiently accounts for these potentially confounding variables. Unless stops of black and white suspects occur in similar circumstances, the regression model will be sensitive to the terms in the model, such as interactions between race and other predictors (e.g., race*location). Unfortunately, this situation is often overlooked in criminological studies of racial profiling.

Earlier we showed an example in which we could reweight the stops of other officers to match the features of stops of a particular officer. In the same manner, Ridgeway[56] showed that we can construct propensity score weights to reweight the stops of, for example, nonblack drivers or pedestrians to match the characteristics of the stops of black drivers or pedestrians. Table 7.4, from a Cincinnati Police Department study of racial profiling in traffic stops described in Schell, Ridgeway, and colleagues,[57] provides a demonstration. The second column displays the percentages for the black drivers; the third column displays the percentages for the weighted nonblack drivers. The weighted percentages for the nonblack drivers are uniformly close to the percentages for the black drivers. Achieving this balance is the critical step when using propensity score techniques, and removes the problems of insufficient overlap between races and nonlinearity noted with regression models. Race, therefore, is the only factor differing between the groups by design. The fourth column in table 7.4 displays the raw percentages for the nonblack driver sample. These data indicate that very few nonblack drivers are involved in stops in Over-the-Rhine. Nonblack drivers are much more likely to be stopped on the freeways. Therefore, the weighted sample has been constructed to downweight nonblack drivers stopped on the freeways and upweight nonblack drivers stopped in Over-the-Rhine. Additionally, nonblack drivers with invalid driver's licenses are upweighted so that the rate of invalid driver's licenses in the comparison sample is closer to that of the black driver sample.

Aside from some statistical advantages, the method is also attractive because of the ease of establishing its face validity. Table 7.4 is easy to explain to a variety of policy audiences, and it is effective for arguing that the subsequent results are based on apples-to-apples comparisons.

The raw numbers indicated that black drivers were much less likely than nonblack drivers to have had a traffic stop last less than ten minutes, 47% versus 56%. After

TABLE 7.4
Comparison on a Subset of Stop Features of the Nonblack Driver Sample to Black Drivers

	% Black drivers N = 20,146	% Nonblack drivers (weighted) ESS = 5,365	% Nonblack drivers (unweighted) N = 24,383
Neighborhood			
Downtown	2.4	2.4	4.8
Over-the-Rhine	7.1	6.9	3.2
I-71	2.1	2.1	6.1
I-75	6.0	6.1	13.6
Time of day			
12–3 am	23.3	21.8	16.7
3–6 am	5.2	4.8	3.7
6–9 am	6.0	8.3	10.8
9 am–12 pm	6.8	7.8	12.7
12–3 pm	6.9	7.5	12.8
3–6 pm	16.9	17.8	15.2
6–9 pm	15.8	14.9	12.7
9 pm–12 am	19.0	17.0	15.4
Reason			
Equipment violation	24.0	22.6	12.7
Moving violation	66.1	69.7	83.4
Resident			
Cincinnati	91.8	90.8	63.2
Ohio (not Cincinnati)	3.8	4.3	18.8
Kentucky	1.9	2.6	11.7
Age			
Under 18	1.7	1.7	1.8
18–25	34.8	32.4	31.2
26–35	28.9	26.3	26.0
36–45	17.5	19.0	18.9
Invalid driver's license	18.0	13.2	5.3
Male	65.9	64.6	65.1

weighting, the nonwhite drivers stopped at similar times, places, and contexts had stops last less than ten minutes 47% of the time, the same as the black drivers. All the difference between the original numbers, 47% and 56%, can be attributable to the factors like time, place, and context.

As with the propensity score approach previously discussed, there are advantages and disadvantages to both hit rates and matching approaches. The hit rate approach has intuitive appeal, providing a clear thought experiment where all else should be equal once the police make the decision of whom to stop. The hit rates comparison assumes that selecting on whom police decide to stop equalizes the two groups so that whites and blacks should be equivalent. If blacks end up with lower hit rates than whites, then one can argue that the police are using a lower threshold in assessing suspicion for blacks. But is this reasonable? Actions transpire after the decision to stop that may be confounded with race. There is a body of research in criminology that suggests a variety of reasons for racial differences in stop outcomes. As we previously discussed, Dixon and colleagues found that black-white officer interactions in Cincinnati explained a substantial difference in the length of a stop and the decision to search a vehicle. These decisions, however, don't appear to be racially biased on the suspects but rather reflect racial differences in police officer practices. Engel and Tillyer[58] note the lengthy history of observation studies that find racial differences in suspect demeanor can affect outcomes in police-citizen interactions, such that all else but race is not equal once an officer has decided to stop a suspect.

By contrast, matching approaches try to make all the statistical adjustments available with observational data. If one has the right set of variables, then there is some confidence that a good test of the race effect in post-stop outcomes can be assessed with accuracy. White and black suspects can be compared to each other in similar situations. If the analyst does not have the right set of contextual variables, they can at least get better data and work on improving the matching strategy. There is no magic going on, no necessary thought experiment; one just wants to construct a feasible set of comparison groups.

Conclusions

The search for an appropriate method for assessing racial bias in police behavior has been a quest. Substantial improvements have been made as investigators have moved away from simple comparisons of police stop decisions to general populations estimates. The search for the appropriate benchmark, however, remains elusive. There is no clear way to establish the correct population at risk for police attention. All approaches have limitations. Clearly, the most feasible benchmarks are ones that attempt to remove as many factors that are potentially confounded with race as possible but are legally permissible on the part of the police. The key to drawing a causal inference about the importance of race is establishing a set of comparison conditions that are race-neutral. This is, however, a significant challenge because many factors are highly confounded with race. Census estimates are inappropriate benchmarks. Observations are difficult to collect in a systematic fashion, and require observers to note behaviors for which the police should consider someone suspicious. With enough training, effort, and time, observation methods can be an effective benchmark in studies that focus on traffic enforcement on highways where minorities and whites are exposed to similar circumstances, but they are less likely to be useful in highly stratified urban environments where the police focus on much more than traffic enforcement. Arrest data is too confounded with police stop decisions to be a useful benchmark. After all, arrests are often a consequence of the decision to stop and search someone. Instrumental variables offer some promise by relying on variations in nature that are independent of race, such as the switch from daylight to darkness. Here, too, instrumental variables are limited to drawing a causal inference from the conditions under which they are estimated. If, for example, the police behave systematically different toward minorities only in late night hours, variations in natural daylight won't be useful for detecting racial bias. Hit rates are attractive because of the idea that police want to maximize their ability to find contraband and make reasonable arrests, so selecting on who is stopped should provide a race-neutral test. Racial differences in the characteristics of criminal offenders, however, can make a focus on hit rates invalid. Approaches that compare like criminals will yield better hit rate assessments. Matching approaches that compare whites to minorities in similar circumstances offer promise because they attempt to make apple-to-apple comparisons.

A good matching approach, for example, could provide all relevant police fact race. Omitted variables will always be a concern. What important variables a ing can, however, be a good subject of discussion. If the police cannot arti reasonable set of missing variables that are not recorded and are associated with racial differences in who is searched, the duration of stops, and so forth, then this provides at least circumstantial evidence of race bias.

Even if police decisions on whom to stop, search, and detain are not intentionally biased, they may be structurally discriminatory. Patrolling differently in high-crime neighborhoods may place a disparate burden on minorities but may not reflect actual bias in police decision making, especially when one compares whites and minorities in similarly situated circumstances. Blacks, for example, disproportionately live in neighborhoods plagued by crime and violence, and there are few large U.S. cities where whites live in comparable circumstances. Even when one does compare whites driving or walking through predominately minority neighborhoods and finds no difference in the probability of being stopped, searched, and so forth, the reality is that these individuals likely reflect only a small fraction of police actions in minority neighborhoods. So while the decisions by the police may not be intentionally biased, they may serve to affirm perceptions of bias because the level of police activity is greater in high crime-poverty areas disproportionately settled by minorities.

Unfortunately there is no unifying method that can establish the extent to which racially biased policing occurs. All approaches have weaknesses. Social scientists should therefore be measured in their assessments.

NOTES

1. This is sometimes referred to as the denominator from the standpoint that the proportion of minority stops should be divided by the population at risk (e.g., % black stops/ % blacks at risk for being stopped) to provide an appropriate adjustment for detecting racial disparities.
2. Sampson, R. and W. J. Wilson. Toward a theory of race, crime, and urban inequality, pp. 37–54.
3. Skogan. Disorder and decline: Crime and the spiral of decay in American neighborhoods.
4. Hindelang. Variations in sex-age-race incidence rates of offending, pp. 461–475.
5. Fridell. By the numbers: A guide for analyzing race data from vehicle stops.
6. Zingraff et al. Evaluating North Carolina State Highway Patrol data.
7. Farrell et al. Rhode Island Traffic Stop Statistics Act.
8. Alpert et al. Toward a better benchmark: Assessing the utility of not-at-fault traffic crash data in racial profiling research.
9. Montgomery County Department of Police. Traffic stop data collection analysis.
10. McConnell, E. H. and A. R. Scheidegger. Race and speeding citations: Comparing speeding citations issued by air traffic officers with those issued by ground traffic officers.
11. Alpert et al. Police suspicion and discretionary decision making during citizen stops, pp. 407–434.
12. Reiss. Systematic social observation of natural social phenomena, pp. 3–33.

13. Lamberth. Revised statistical analysis of the incidence of police stops and arrests of black drivers/travelers on the New Jersey Turnpike.
14. Maxfield, R. and G. Kelling. New Jersey State Police and stop data.
15. Lange et al. Speed violation survey of the New Jersey Turnpike.
16. Alpert et al. Investigating racial profiling by the Miami-Dade Police Department, pp. 25–56.
17. Smith. The neighborhood context of police behavior, pp. 313–341.
18. Eck et al. Vehicle police stops in Cincinnati.
19. Alpert et al. Investigating racial profiling by the Miami-Dade Police Department, pp. 25–56.
20. Ibid.
21. 302 F. Supp. 2d 1217, 1229–1232, D. Kan., 2004.
22. Gelman et al. An analysis of the New York City Police Department's "stop-and-frisk" policy in the context of claims of racial bias, pp. 813–823.
23. Klinger. Negotiating order in patrol work: An ecological theory of police response to deviance, pp. 277–306.
24. For a discussion of the benefits and limitations of citizens' calls for police service data see Klinger, D. and G. Bridges. Measurement error in calls-for-service as an indicator of crime, pp. 705–726.
25. Ridgeway. Analysis of racial disparities in the New York Police Department's stop, question, and frisk practices.
26. For technical details see Angrist et al. Identification of causal effects using instrumental variables, pp. 444–455.
27. This is known as the nonzero average causal effect of the instrument on actual treatment assignment.
28. This is known as the exclusion restriction.
29. Grogger, J. and G. Ridgeway. Testing for racial profiling in traffic stops from behind a veil of darkness, pp. 878–887.
30. Lamberth. Racial profiling data analysis study: Final report for the San Antonio Police Department.
31. Greenwald. Final report: Police vehicle stops in Sacramento, California.
32. Neyman. On the application of probability theory to agricultural experiments.
33. Schell et al. Police-community relations in Cincinnati: Year three evaluation report.
34. Jefferis et al. The effect of a videotaped arrest on public perceptions of police use of force, pp. 381–395.
35. Weitzer. Incidents of police misconduct and public opinion, pp. 397–408.
36. Christopher. Report of the Independent Commission on the Los Angeles Police Department.
37. See Walker. Searching for the denominator: Problems with police traffic stop data and an early warning system solution, pp. 63–95; Walker. The citizen's guide to interpreting traffic stop data: Un-raveling the racial profiling controversy; and Walker. Internal benchmarking for traffic stop data: An early intervention system approach.
38. Walker. Early intervention systems for law enforcement agencies: A planning and management guide.
39. Birotte. Training evaluation and management system (TEAMS) II audit, phase I (fiscal year 2007/2008).
40. Walker. Early intervention systems for law enforcement agencies: A planning and management guide.

41. Decker, S. and J. Rojek. Saint Louis Metropolitan Police Department traffic stop patterns.
42. Ridgeway, G. and J. MacDonald. Doubly robust internal benchmarking and false discovery rates for detecting racial bias in police stops.
43. Ridgeway, G. et al. Police-community relations in Cincinnati: Year two evaluation report.
44. Ridgeway, G. Analysis of racial disparities in the New York Police Department's stop, question, and frisk practices.
45. For technical details see Kang, J. and J. Schafer. Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data, pp. 523–580.
46. Fridell. By the numbers: A guide for analyzing race data from vehicle stops.
47. Smith, M. R. Depoliticizing racial profiling: Suggestions for the limited use and management of race in police decision-making, pp. 219–260.
48. Benjamini, Y. and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing.
49. Dixon, T. et al. The influence of race in police-civilian interactions: A content analysis of videotaped interactions taken during Cincinnati police traffic stops, pp. 530–549.
50. Schell, T. et al. Police-community relations in Cincinnati: Year three evaluation report.
51. See Knowles, J., N. Persico and P. Todd. Racial bias in motor vehicle searches, pp. 203–229; Persico, N. and P. Todd. Generalising the hit rates test for racial bias in law enforcement, with an application to vehicle searches in Wichita, pp. F351–F367; Persico, N. and P. Todd. The hit rates test for racial bias in motor-vehicle searches, pp. 37–53.
52. Persico, N. and P. Todd. The hit rates test for racial bias in motor-vehicle searches, pp. 37–53.
53. Ayres. Outcome tests of racial disparities in police practices, pp. 131–142.
54. Ridgeway, G. Analysis of racial disparities in the New York Police Department's stop, question, and frisk practices.
55. Schell et al. Police-community relations in Cincinnati: Year three evaluation report.
56. Ridgeway. Assessing the effect of race bias in post-traffic stop outcomes using propensity scores.
57. Schell et al. Police-community relations in Cincinnati: Year three evaluation report.
58. Engel, R. S. and R. Tillyer. Searching for equilibrium: The tenuous nature of the outcome test, pp. 54–71.

REFERENCES

Alpert, Geoffrey P., John M. MacDonald, and Roger G. Dunham. 2005. Police suspicion and discretionary decision making during citizen stops. *Criminology* 43: 407–434.

Alpert, Geoffrey P., Michael R. Smith, and Roger G. Dunham. 2004. Toward a better benchmark: Assessing the utility of not-at-fault traffic crash data in racial profiling research. *Justice Research and Policy* 6: 43–70.

———. 2007. Investigating racial profiling by the Miami-Dade Police Department: A multimethod approach. *Criminology and Public Policy* 6: 25–56.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91: 444–455.

Ayres, Ian. 2002. Outcome tests of racial disparities in police practices. *Justice Research and Policy* 4: 131–142.
Benjamini, Yoav, and Yosef Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B* 57: 289–300.
Birotte, Andre. 2007. *Training evaluation and management system (TEAMS) II audit, phase I (fiscal year 2007/2008)*. Los Angeles: Office of the Inspector General, Los Angeles Police Department.
Christopher, Warren. 1991. *Report of the Independent Commission on the Los Angeles Police Department*. Los Angeles: Independent Commission on the Los Angeles Police Department.
Decker, Scott H., and Jeff Rojek. 2002. *Saint Louis Metropolitan Police Department traffic stop patterns*, St. Louis: University of Missouri.
Dixon, Travis L., Terry L. Schell, Howard Giles, and Kristin L. Drogos. 2008. The influence of race in police–civilian interactions: A content analysis of videotaped interactions taken during Cincinnati police traffic stops. *Journal of Communication* 58: 530–549.
Eck, John E., Lin Liu, and Lisa G. Bostaph. 2003. *Police vehicle stops in Cincinnati: July 1–December 31, 2001*. Cincinnati: University of Cincinnati.
Efron, Bradley. 2004. Large-scale simultaneous hypothesis testing: The choice of a null hypothesis. *Journal of the American Statistical Association* 99: 96–104.
———. 2007. Correlation and large-scale simultaneous significance testing. *Journal of the American Statistical Association* 102: 93–103.
Engel, Robin S., and Tillyer, Rob. 2008. Searching for equilibrium: The tenuous nature of the outcome test. *Justice Quarterly* 25: 54–71.
Farrell, Amy, Dean Jack McDevitt, Shea Cronin, and Erica Pierce. 2003. *Rhode Island traffic stop statistics act final report*. Boston: Northeastern University.
Fridell, Lorie A. 2004. *By the numbers: A guide for analyzing race data from vehicle stops*. Washington, D.C.: Police Executive Research Forum.
Gelman, Andrew, Jeffrey Fagan, and Alex Kiss. 2007. An analysis of the New York City Police Department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association* 102: 813–823.
Greenwald, Howard P. 2001. *Final Report: Police vehicle stops in Sacramento, California*. Sacramento: Sacramento Police Department.
Grogger, Jeffrey, and Greg Ridgeway. 2006. Testing for racial profiling in traffic stops from behind a veil of darkness. *Journal of the American Statistical Association* 101: 878–887.
Imbens, Guido. 2003. *Nonparametric estimation of average treatment effects under exogeneity: A review. Technical Working Paper 294*. Cambridge: National Bureau of Economic Research.
Hindelang, Michael. 1981 Variations in sex-age-race incidence rates of offending. *American Sociological Review* 46: 461–474.
Jefferis, Eric S., Robert J. Kaminski, Stephen Holmes, and Dena E. Hanley. 1997. The effect of a videotaped arrest on public perceptions of police use of force. *Journal of Criminal Justice* 25: 381–395.
Kang, Joseph D. Y., and Joseph L. Schafer. 2007. Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science* 22: 523–580.
Klinger, David A. 1994. Demeanor or crime? Why "hostile" citizens are more likely to be arrested. *Criminology* 32: 475–494.
———. 1997. Negotiating order in patrol work: An ecological theory of police response to deviance. *Criminology* 35: 277–306.

Klinger, David A., and George Bridges. 1997. Measurement error in calls-for-service as an indicator of crime. *Criminology* 35: 705–726.
Knowles, John, Nicola Persico, and Petra Todd. 2001. Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy* 109: 203–229.
Lamberth, John. 1994. *Revised statistical analysis of the incidence of police stops and arrests of black drivers/travelers on the New Jersey Turnpike between exits or interchanges 1 and 3 from the years 1988 through 1991*. Philadelphia: Temple University, Department of Psychology.
Lamberth, John. 2003. *Racial profiling data analysis study: Final report for the San Antonio Police Department*. San Antonio: SAPD.
Lange, James E., Kenneth O. Blackman, and Mark B. Johnson. 2001. *Speed violation survey of the New Jersey Turnpike: Final report*. New Jersey: Office of the Attorney General of New Jersey.
Maxfield, Michael, and George L. Kelling 2005. *New Jersey State Police and stop data: What do we know, what should we know, and what should we do?* Newark: The Police Institute at Rutgers-Newark, School of Criminal Justice, Rutgers University.
McConnell, E. H., and A. R. Scheidegger. 2001. Race and speeding citations: Comparing speeding citations issued by air traffic officers with those issued by ground traffic officers. Paper presented at the annual meeting of the Academy of Criminal Justice Sciences, April 4–8, in Washington, D.C.
Montgomery County Department of Police. 2002. *Traffic stop data collection analysis, 3rd report*. Maryland: Montgomery County Department of Police.
Neyman, Jerzy 1923. On the application of probability theory to agricultural experiments: Essay on principles, Section 9. Translated in *Statistical Science* 5: 465–480, 1990.
Persico, Nicola, and Petra Todd. 2006. Generalising the hit rates test for racial bias in law enforcement, with an application to vehicle searches in Wichita. *The Economic Journal* 116: 351–367.
———. 2008. The hit rates test for racial bias in motor-vehicle searches. *Justice Quarterly* 25: 37–53.
Reiss, Albert J. 1971. Systematic observation of natural social phenomena. *Sociological Methodology* 3: 3–33.
Ridgeway, Greg. 2006. Assessing the effect of race bias in post-traffic stop outcomes using propensity scores. *Journal of Quantitative Criminology* 22: 1–26.
Ridgeway, Greg. 2007. *Analysis of racial disparities in the New York Police Department's stop, question, and frisk practices*. Santa Monica: RAND Corporation.
Ridgeway, Greg, and John M. MacDonald. 2008. Doubly robust internal benchmarking and false discovery rates for detecting racial bias in police stops. *Unpublished Manuscript*.
Ridgeway, Greg, Terry L. Schell, K. Jack Riley, Susan Turner, and Travis L. Dixon. 2006. *Police-community relations in Cincinnati: Year two evaluation report*. Santa Monica: RAND Corporation.
Sampson, Robert J., and William Julius Wilson. 1995. Toward a theory of race, crime, and urban inequality. In *Crime and Inequality*, ed. John Hagan and Ruth D. Peterson, 37–54. Stanford: Stanford University Press.
Schell, Terry L., Greg Ridgeway, Travis L. Dixon, Susan Turner, and K. Jack Riley. 2007. *Police-community relations in Cincinnati: Year three evaluation report*. Santa Monica: RAND Corporation.
Skogan, Wesley G. 1990. *Disorder and decline: Crime and the spiral of decay in American neighborhoods*. Berkeley and Los Angeles: University of California Press.
Smith, Douglas A. 1986. The neighborhood context of police behavior. In *Communities and*

Crime, ed. Albert J. Reiss and Michael H. Tonry, 313–341. Chicago: University of Chicago Press.

Smith, Michael. R. 2005. Depoliticizing racial profiling: Suggestions for the limited use and management of race in police decision-making. *George Mason University Civil Rights Law Journal* 15: 219–260.

Walker, Samuel. 2001. Searching for the denominator: Problems with police traffic stop data and an early warning system solution. *Justice Research and Policy* 3: 63–95.

———. 2002. The citizen's guide to interpreting traffic stop data: Un-raveling the racial profiling controversy. *Unpublished manuscript.*

———. 2003a. *Early intervention systems for law enforcement agencies: A planning and management guide.* Washington, D.C.: U.S. Department of Justice, Office of Community Oriented Policing Services.

———. 2003b. *Internal benchmarking for traffic stop data: An early intervention system approach.* Washington, D.C.: Police Executive Research Forum.

Weitzer, Ronald. 2002. Incidents of police misconduct and public opinion. *Journal of Criminal Justice* 30: 397–408.

Zingraff, Matthew T., Mary Mason, William Smith, Donald Tomaskovic-Devey, Patricia Warren, Harvey L. McMurray, and C. Robert Fenlon. 2000. *Evaluating North Carolina State Highway Patrol data: Citation, warnings, and searches in 1998.* North Carolina: North Carolina Department of Crime Control and Public Safety and North Carolina State Highway Patrol.

Chapter 8

Using Geographic Information Systems to Study Race, Crime, and Policing

Matt R. Nobles

Introduction

Recently, the relationships between space (in the ecological or geographical sense) and other social phenomena have benefitted from advancements of powerful technologies that put new analytical methods into the hands of researchers and practitioners alike. In particular, GIS (Geographic Information Systems) has become indispensible in the study of policing, where it is relied on to help identify patterns in offending, guide resource deployment and targeted interventions, increase awareness of police-community relations, and a host of other roles. Although many examples of the application of GIS technology to policing may be available in the field, one highly visible model is the use of CompStat, a GIS-focused approach to investigation, problem solving, resource management, and accountability in routine police patrol. CompStat represents not only an adoption of new technological tools in the fight against crime, but also a shift in strategic and tactical decision making that puts crime data and geographical information at the forefront of proactive policy. This chapter briefly acknowledges the extensive and diverse literature connecting geography, race, and policing to the study of crime before turning to a discussion of the methodological advantages of using GIS to visualize these relationships. Several case studies involving the use of GIS in the study of race, crime, and policing are presented, followed by a discussion of GIS as a less obvious tool for identifying and combating social problems.

Literature Review

Perspectives on Place, Race, and Crime

Scholars in criminology, sociology, and related fields have long embraced the idea that crime is related to geography. This concept is readily identified in some of the most influential criminological theories,[1] beginning with the Chicago School emphasizing human ecology and social disorganization,[2] and later extending to more literal interpretations and implications for urban design and crime prevention policy.[3]

Exhibit I

BY THE NUMBERS

Lorie A. Fridell

A GUIDE FOR ANALYZING RACE DATA FROM VEHICLE STOPS

Police Executive Research Forum

COPS
Community Oriented Policing Services
U.S. Department of Justice

- Address the possible intervening impact of age by breaking down the demographic profile of residents and nonresidents into two age groups: age 15 to 24 and age 25 and above;
- Match numerator and denominator. Delete from the stop data (the numerator) the stops of people who are neither residents of the target jurisdiction nor residents of the outside jurisdictions that are encompassed in the analysis; and
- Calculate a measure of racial/ethnic disparity (see Chapter 12) after developing the profile of the people stopped and the profile of the benchmark population.

Drawing Conclusions from the Results

Again we assess the strengths and weaknesses of this method in terms of the alternative hypotheses:

- Like other methods to estimate resident/nonresident driving populations, this one addresses the hypothesis that *racial/ethnic groups are not equally represented as residents in the jurisdiction*.
- By estimating the demographic profiles of nonresidents who might enter the target jurisdiction, this method addresses, in part, the possibility that *racial/ethnic groups are not equally represented as drivers on jurisdiction roads*.
- If analyses are conducted within subareas of the jurisdiction, this method addresses the hypothesis that *racial/ethnic groups are not equally represented as drivers on jurisdiction roads where stopping activity by police is high*.
- If analyses are conducted within age groups, this method takes into account the potential impact of age on driving behavior.

- This method does not address the possibility that unequal representation of racial/ethnic groups on jurisdiction roads may be attributable, in part, to differences across racial/ethnic groups in the quantity of their driving.
- This method does not address the alternative hypothesis that *racial/ethnic groups are not equivalent in the nature and extent of their traffic law-violating behavior*.

MAKING OTHER ADJUSTMENTS TO CENSUS DATA: THE RHODE ISLAND STUDY

Researchers are looking for additional ways to adjust census data to produce more valid benchmarks. For example, Amy Farrell, Jack McDevitt, Shea Cronin, and Erica Pierce of Northeastern University have recently implemented a creative adjustment model.[26] In July 2000 the Rhode Island Traffic Stop Statistics Act was passed. The Northeastern team was contracted to analyze the data collected, in response to this legislation, by the Rhode Island State Police and all municipal police departments in the state. For the municipal police departments, Farrell's team—like Novak and the Missouri team whose work is described above—adjusted census data on jurisdiction residents to account for the influx of nonresident drivers.[27] As the authors explain (Farrell et al. 2003, 29), "we created a driving population estimate based on the idea that the demographics of a target city may be better understood by weighting the population of the target city by its surrounding cities whose drivers may drive in or through the city in question." Specifically, they developed a "driving population estimate" or DPE for each municipal department based on formulas that took into account

26 See Farrell et al. (2003). The final report is available under "Reports and Publications" at www.riag.state.ri.us.

27 The team used the observation method—described in Chapter 9—to analyze the data collected by the state police.

Exhibit J

An Analysis of the New York City Police Department's "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias

Andrew GELMAN, Jeffrey FAGAN, and Alex KISS

Recent studies by police departments and researchers confirm that police stop persons of racial and ethnic minority groups more often than whites relative to their proportions in the population. However, it has been argued that stop rates more accurately reflect rates of crimes committed by each ethnic group, or that stop rates reflect elevated rates in specific social areas, such as neighborhoods or precincts. Most of the research on stop rates and police–citizen interactions has focused on traffic stops, and analyses of pedestrian stops are rare. In this article we analyze data from 125,000 pedestrian stops by the New York Police Department over a 15-month period. We disaggregate stops by police precinct and compare stop rates by racial and ethnic group, controlling for previous race-specific arrest rates. We use hierarchical multilevel models to adjust for precinct-level variability, thus directly addressing the question of geographic heterogeneity that arises in the analysis of pedestrian stops. We find that persons of African and Hispanic descent were stopped more frequently than whites, even after controlling for precinct variability and race-specific estimates of crime participation.

KEY WORDS: Criminology; Hierarchical model; Multilevel model; Overdispersed Poisson regression; Police stops; Racial bias.

1. BIAS IN POLICE STOPS?

In the late 1990s, popular, legal, and political concerns were raised across the United States about police harassment of minority groups in their everyday encounters with law enforcement. These concerns focused on the extent to which police were stopping people on the highways for "driving while black" (see Weitzer 2000; Harris 2002; Lundman and Kaufman 2003). Additional concerns were raised about racial bias in pedestrian stops of citizens by police predicated on "zero-tolerance" policies to control quality-of-life crimes and policing strategies concentrated in minority communities that targeted illegal gun possession and drug trafficking (see Fagan, Zimring, and Kim 1998; Greene 1999; Skolnick and Caplovitz 2001; Fagan and Davies 2000, 2003; Fagan 2002; Gould and Mastrofski 2004). These practices prompted angry reactions among minority citizens that widened the breach between different racial/ethnic groups in their trust in the police (Lundman and Kaufman 2003; Tyler and Huo 2003; Weitzer and Tuch 2002), provoking a crisis of legitimacy with legal, moral, and political dimensions (see Wang 2001; Russell 2002; Harris 2002).

In an era of declining crime rates, policy debates on policing strategies often pivot on the evaluation of New York City's policing strategy during the 1990s, a strategy involving aggressive stops and searches of pedestrians for a wide range of crimes (Eck and Maguire 2000; Skogan and Frydl 2004). The policy was based on the lawful practice of "temporarily detaining, questioning, and, at times, searching civilians on the street" (Spitzer 1999). The U.S. Supreme Court has ruled police stop-and-frisk procedures to be constitutional under certain restrictions (Terry v. Ohio 1968). The approach of the New York City Police Department (NYPD) during the 1990s has been widely credited as a major source of the city's sharp crime decline (Zimring 2006).

But near the end of the decade there were repeated complaints of harassment of minority communities, especially by the elite Street Crimes Unit (Spitzer 1999). These complaints came in the context of the well-publicized assault by police of Abner Louima and the shootings of Amadou Diallo and Patrick Dorismond. Citizen complaints about aggressive "stop and frisk" tactics ultimately provoked civil litigation that alleged racial bias in the patterns of "stop and frisk," leading to a settlement that regulated the use of this tactic and established extensive monitoring requirements (Kelvin Daniels et al. v. City of New York 2004).

We address this dispute by estimating the extent of racially disparate impacts of what came to be known as the "New York strategy." We analyze the rates at which New Yorkers of different ethnic groups were stopped by the police on the city streets, to assess the central claim that race-specific stop rates reflect nothing more than race-specific crime rates. This study is based on work performed with the New York State Attorney General's Office (Spitzer 1999) and reviewed by the U.S. Commission on Civil Rights (2000). Key statistical issues are the baselines used to compare rates (recognized as a problem by Miller 2000; Walker 2001; Smith and Alpert 2002) and local variation in the intensity of policing, as performed by the Street Crimes Unit and implicitly recommended by Wilson and Kelling (1982) and others. We use multilevel modeling (see Raudenbush and Bryk 2002 for an overview and Sampson, Raudenbush, and Earls 1997; Sampson and Raudenbush 1999; Weidner, Frase, and Pardoe 2004 for examples in studies of crime) to adjust for local variation in comparing the rates of police stops of different ethnic groups in New York City.

Were the police disproportionately stopping ethnic minorities? We address this question in several different ways using data on police stops and conclude that members of minority groups were stopped more often than whites, both in comparison to their overall population and to the estimated rates of

Andrew Gelman is Professor, Department of Statistics and Department of Political Science, Columbia University, New York, NY 10027 (E-mail: *gelman@stat.columbia.edu*). Jeffrey Fagan is Professor, Law School and School of Public Health, Columbia University, New York, NY 10027 (E-mail: *jfagan@law.columbia.edu*). Alex Kiss is Biostatistician, Department of Research Design and Biostatistics, Sunnybrook and Women's College Health Sciences Center, Toronto, Ontario, Canada. The authors thank the New York City Police Department, the New York State Division of Criminal Justice Services, and the Office of the New York State Attorney General for providing data for this research. Tamara Dumanovsky and Dong Xu made significant contributions to the analysis. Joe Bafumi, Rajeev Dehejia, Jim Liebman, Dan Rabinowitz, Caroline Rosenthal Gelman, and several reviewers provided helpful comments. Support for this research was provided in part by National Science Foundation grants SES-9987748 and SES-0318115. All opinions are those of the authors.

Journal of the American Statistical Association
September 2007, Vol. 102, No. 479, Applications and Case Studies
DOI 10.1198/016214506000001040

crime that they have committed. We do not necessarily conclude that the NYPD engaged in discriminatory practices, however. The summary statistics that we study here cannot directly address questions of harassment or discrimination, but rather reveal statistical patterns that are relevant to these questions.

Because this is a controversial topic that has been studied in various ways, we go into some detail in Sections 2 and 3 on the historical background and available data. We present our models and results in Sections 4 and 5, and provide some discussion in Section 6.

2. BACKGROUND

2.1 Race, Neighborhoods, and Police Stops

Nearly a century of legal and social trends has set the stage for the current debate on race and policing. Historically, close surveillance by police has been a part of everyday life for African-Americans and other minority groups (see, e.g., Musto 1973; Kennedy 1997). More recently, in Whren et al. v. U.S. (1996), the U.S. Supreme Court allowed the use of race as a basis for a police stop as long as there were other factors motivating the stop. In Brown v. Oneonta (2000), a federal district court permitted the use of race as a search criterion if there was an explicit racial description of the suspect.

The legal standard for police conduct in citizen stops derives from Terry v. Ohio (1968), which involved a pedestrian stop that set the parameters of the "reasonable suspicion" standard for police conduct in detaining citizens for search or arrest. Recently, the courts have expanded the concept of "reasonable suspicion" to include location as well as behavior. For example, the U.S. Supreme Court, in Illinois v. Wardlow (2000), noted that although a person's presence in a "high-crime area" does not meet the standard for a particularized suspicion of criminal activity, a location's characteristics are relevant to determining whether a behavior is sufficiently suspicious to warrant further investigation. Because "high-crime areas" often have high concentrations of minority citizens (Massey and Denton 1993), this logic places minority neighborhoods at risk for elevating the suspiciousness of their residents.

Early studies suggested that both the racial characteristics of the suspect and the racial composition of the suspect's neighborhood influence police decisions to stop, search, or arrest a suspect (Bittner 1970; Reiss 1971). Particularly in urban areas, suspect race interacts with neighborhood characteristics to animate the formation of suspicion among police officers (Thompson 1999; Smith, Makarios, and Alpert 2006). Alpert, MacDonald, and Dunham (2005) found that police are more likely to view a minority citizen as suspicious—leading to a police stop—based on nonbehavioral cues, while more often relying on behavioral cues to develop suspicion for white citizens.

But police also may substitute racial characteristics of communities for racial characteristics of individuals in their cognitive schema of suspicion, resulting in elevated stop rates in neighborhoods with high concentrations of minorities. For example, in a study of policing in three cities, Smith (1986) showed that suspects in poor neighborhoods were more likely to be arrested, in an analysis controlling for suspect behavior and type of crime. The suspect's race and the racial composition of the suspect's neighborhood were also significant predictors of police response. Coercive police responses may relate to the perception that poor neighborhoods may have limited capacity for social control and self-regulation. This strategy was formalized in the influential "broken windows" essay of Wilson and Kelling (1982), who argued that police responses to disorder were critical to communicate intolerance for crime and to halt its contagious spread. Others have disputed this claim, however (see Harcourt 1998, 2001; Sampson and Raudenbush 1999; Taylor 2000), arguing that race is often used as a substitute for neighborhood conditions as a marker of suspicion by police.

Police have defended racially disparate patterns of stops on the grounds that minorities commit disproportionately more crimes than whites (especially the types of crimes that capture the attention of police), and that the spatial concentration and disparate impacts of crimes committed by and against minorities justifies more aggressive enforcement in minority communities (MacDonald 2001). Police cite such differences in crime rates to justify racial imbalances even in situations where they have a wide range of possible targets or where suspicion of criminal activity would not otherwise justify a stop or search (Kennedy 1997; Harcourt 2001; Rudovsky 2001). Using this logic, police claim that the higher stop rates of African-Americans and other minorities simply represent reasonable and efficient police practice (see, e.g., Bratton and Knobler 1998; Goldberg 1999). Police often point to the high rates of seizures of contraband, weapons, and fugitives in such stops, and also to a reduction of crime, to justify such aggressive policing (Kelling and Cole 1996).

Whether racially disparate stop rates reflect disproportionate crime rates or intentional, racially biased targeting by police of minorities at rates beyond what any racial differences in crime rates might justify lies at the heart of the social and legal controversy on racial profiling and racial discrimination by police (Fagan 2002; Ayres 2002a; Harris 2002). This controversy has been the focus of public and private litigation (Rudovsky 2001), political mobilization, and self-scrutiny by several police departments (see Garrett 2001; Walker 2001; Skolnick and Caplovitz 2001; Gross and Livingston 2002).

2.2 Approaches to Studying Data on Police Stops

Recent evidence supports perceptions among minority citizens that police disproportionately stop African-American and Hispanic motorists, and that once stopped, these citizens are more likely to be searched or arrested (Cole 1999; Veneiro and Zoubeck 1999; Harris 1999; Zingraff et al. 2000; Gross and Barnes 2002). For example, two surveys with nationwide probability samples, completed in 1999 and in 2002, showed that African-Americans were far more likely than others to report being stopped on the highways by police (Langan, Greenfeld, Smith, Durose, and Levin 2001; Durose, Schmitt, and Langan 2005). Both surveys showed that minority drivers also were more likely to report being ticketed, arrested, handcuffed, or searched by police, and that they more often were threatened with force or had force used against them. These disparities exact social costs that, according to Loury (2002), animate culturally meaningful forms of stigma that reinforce racial inequalities, especially in the practice of law enforcement.

"Suspicious behavior" is the spark for both pedestrian and traffic stops (Alpert et al. 2005). Pedestrian stops are at the very core of policing, used to enforce narcotics and weapons laws, to identify fugitives or other persons for whom warrants may be outstanding, to investigate reported crimes and "suspicious" behavior, and to improve community quality of life. For the NYPD, a "stop" intervention provides an occasion for the police to have contact with persons presumably involved in low-level criminality without having to effect a formal arrest, and under the lower constitutional standard of "reasonable suspicion" (Spitzer 1999). Indeed, because low-level "quality of life" and misdemeanor offenses were more likely to be committed in the open, the "reasonable suspicion" standard is more easily satisfied in these sorts of crimes (Rudovsky 2001).

However, in pedestrian and traffic violations, the range of suspicious behaviors in neighborhood policing is sufficiently broad to challenge efforts to identify an appropriate baseline against which to compare race-specific stop rates (see Miller 2000; Smith and Alpert 2002; Gould and Mastrofski 2004). Accordingly, attributing bias is difficult; causal claims about discrimination would require far more information about such baselines than the typical administrative (observational) datasets can supply. Research in situ that relies on direct observation of police behavior (e.g., Gould and Mastrofski 2004; Alpert et al. 2005) requires officers to articulate the reasons for their actions, a task that is vulnerable to numerous validity threats. Instead, reliable evidence of ethnic bias would require experimental designs that control for other factors so as to isolate differences in outcomes that could only be attributed to race or ethnicity. Such experiments are routinely used in tests of discrimination in housing and employment (see, e.g., Pager 2003). But observational studies that lack such controls are often embarrassed by omitted variable biases; few studies can control for all of the variables that police consider in deciding whether to stop or search someone.

Another approach to studying racial disparities bypasses the question of whether police intend to discriminate on the basis of ethnicity or race and instead focuses on disparate impacts of police stop strategies. In this approach, comparisons of "hit rates," or efficiencies in the proportion of stops that yield positive results, serve as evidence of disparate impacts of police stops. This approach can show when the racial disproportionality of a particular policy or decision making outcome is not justified by heightened institutional productivity. In the context of profiling, outcome tests assume that the ex post probability that a police search will uncover drugs or other contraband is a function of the degree of probable cause that police use in deciding to stop and search a suspect (Ayres 2002a). A finding that searches of minorities are less productive than searches of whites could be evidence that police have a lower threshold of probable cause when searching minorities. At the very least, it is a sign of differential treatment of minorities that in turn produces a disparate impact.

Knowles, Persico, and Todd (2001) considered this "hit rate" approach theoretically as well as empirically in a study finding that of the drivers on I-95 in Maryland stopped by police on suspicion of drug trafficking, African-Americans were as likely as whites to have drugs in their cars. The accompanying theoretical analysis posits a dynamic process that considers the behaviors of both police and citizens of different races and integrates their decisions in an equilibrium where police calibrate their behavior to the probabilities of detecting illegal behavior and citizens in different racial groups adjust their propensities to accommodate the likelihood of detection. They concluded that the search for drugs was an efficient allocation of police resources, despite the disparate impacts of these stops on minority citizens (Lamberth 1997; Ayres 2002a,b; Gross and Barnes 2002).

However, this analysis omits several factors that might bias these claims, such as racial differences in the attributes that police consider when deciding which motorists to stop, search, or arrest (see, e.g., Alpert et al. 2005; Smith et al. 2006). Moreover, the randomizing equilibrium assumptions in the approach of Persico et al.—that both police and potential offenders adjust their behavior in response to the joint probabilities of carrying contraband and being stopped—tend to average across heterogeneous conditions both in police decision making and in offenders' propensities to crime (Dharmapala and Ross 2004), and discount the effects of race-specific sensitivities toward crime decisions under varying conditions of detection risk by police stop (Dominitz and Knowles 2005). Addressing these two concerns, Dharmapala and Ross (2004) identified different equilibria that lead to different conclusions about racial prejudice in police stops and searches.

We consider hit rates briefly (see Sec. 5.3), but our main analysis attempts to resolve these supply-side or omitted-variable problems by controlling for race-specific rates of the targeted behaviors in patrolled areas, assessing whether stop and search rates exceed what we would predict from knowledge of the crime rates of different racial groups. This approach indexes stop behavior to observables about the probability of crime or guilt among different racial groups. Moreover, by disaggregating data across neighborhoods, our probability estimates explicitly incorporate the externalities of neighborhood and race that historically have been observed in policing (Skogan and Frydl 2004). This approach requires estimates of the supply of individuals engaged in the targeted behaviors (see Miller 2000; Fagan and Davies 2000; Walker 2001; Smith and Alpert 2002).

To be sure, a finding that police are stopping and searching minorities at a higher rate than is justified by their participation in crime does not require inferring that police engaged in disparate treatment at a minimum, however, it does provide evidence that whatever criteria the police used produced an unjustified disparate impact.

3. DATA

3.1 "Stop and Frisk" in New York City

The NYPD has a policy of keeping records on stops (on "UF-250 Forms"). This information was collated for all stops (about 175,000 in total) from January 1998 through March 1999 (Spitzer 1999). The police are not required to fill out a form for every stop. Rather, there are certain conditions under which the police are required to fill out the form. These "mandated stops" represent 72% of the stops recorded, with the remaining reports being of stops for which reporting was optional. To address concerns about possible selection bias in the nonmandated stops, we repeated our main analyses (shown in Fig. 2) for the mandated stops only; the total rates of stops changed, but the relative rates for different ethnic groups remained essentially unchanged.

The UF-250 form has a place for the police officer to record the "Factors which caused officer to reasonably suspect person stopped (include information from third persons and their identity, if known)." We examined these forms and the reasons for the stops for a citywide sample of 5,000 cases, along with 10,869 others, representing 50% of the cases in a nonrandom sample of 8 of the 75 police precincts, chosen to represent a spectrum of racial population characteristics, crime problems, and stop rates, guided by the policy questions in the original study (Spitzer 1999, p. 158). The following examples (from Spitzer 1999) illustrate the rules that motivated police decisions to stop suspects and demonstrate the social and behavioral factors that police apply in the process of forming reasonable suspicion:

- "At TPO [time and place of occurrence] male was with person who fit description of person wanted for GLA [grand larceny auto] in 072 pct. log . . . upon approach male discarded small coin roller which contained 5 bags of alleged crack."
- "At T/P/O R/O [reporting officer] did observe below named person along w/3 others looking into numerous parked vehicles. R/O did maintain surveillance on individuals for approx. 20 min. Subjects subsequently stopped to questioned [sic] w/ neg results."
- "Slashing occurred at Canal street; person fit description; person was running."
- "Several men getting in and out of a vehicle several times."
- "Def. Did have on a large bubble coat with a bulge in right pocket."
- "Person stopped did stop [sic] walking and reverse direction upon seeing police. Attempted to enter store as police approached; Frisked for safety."

Based on federal and state law, some of these reasons for stopping a person are constitutional and some are not. For example, courts have ruled that a bulge in the pocket is not sufficient reason for the police to stop a person without his or her consent (People v. DeBour 1976; People v. Holmes 1996), and that walking away from the police is not a sufficient cause to stop and frisk a person (Brown v. Texas 1979; but see Illinois v. Wardlow 2000). However, when the police observe illegal activity, weapons (including "waistband bulges"), a person who fits a description, or suspicious behavior in a crime area, then stops and frisks have been ruled constitutional (Spitzer 1999).

The New York State Attorney General's office used rules such as these to characterize the rationales for 61% of the stops in the sample as articulating a "reasonable suspicion" that would justify a lawful stop, 15% of the stops as not articulating a reasonable suspicion, and 24% as providing insufficient information on which to base a decision. For the controversial Street Crimes Unit, 23% of stops were judged to not articulate a reasonable suspicion. (There was no strong pattern by ethnicity here; the rate of stops judged to be unreasonable was about the same for all ethnic groups.) The stops judged to be without "reasonable suspicion" indeed seemed to be weaker, in that only 1 in 29 of these stops led to arrests, compared with 1 in 7 of the stops with reasonable suspicion.

3.2 Aggregate Rates of Stops for Each Ethnic Group

With this as background, we analyze the entire stop-and-frisk dataset to see to what extent different ethnic groups were stopped by the police. We focus on blacks (African-Americans), Hispanics (Latinos), and whites (European-Americans). The categories are as recorded by the police making the stops. We exclude members of other ethnic groups (approximately 4% of the stops) because of the likelihood of ambiguities in classifications. With such a low frequency of "other," even a small rate of misclassification can cause large distortions in the estimates for that group. For example, if only 4% of blacks, Hispanics, and whites were mistakenly labeled as "other," this would nearly double the estimates for the "other" category while having very little effects on the three major groups. (See Hemenway 1997 for an extended discussion of the problems that misclassifications can cause in estimates of a small fraction of the population.) To give a sense of the data, Figure 1 displays the number of stops for blacks, Hispanics, and whites over the 15-month period, separately showing stops associated with each of four types of offenses ("suspected charges" as characterized on the UF-250 form): violent crimes, weapons offenses, property crimes, and drug crimes.

In total, blacks and Hispanics represented 51% and 33% of the stops, despite being only 26% and 24%, of the city population based on the 1990 Census. The proportions change little if we use 1998 population estimates and count only males age 15–30, which is arguably a better baseline. For one of our supplementary analyses, we also use the population for each ethnic group within each precinct in the city. Population estimates for the police precincts with low residential populations but high daytime populations due to commercial and business activity were adjusted using the U.S. Census Bureau "journey file," provided by the New York City Department of City Planning (see Spitzer 1999, app. I, table 1.A.1a). The journey file uses algorithms based on time traveled to work and the distribution of job classifications to estimate the day and night populations of census tracts. Tracts were aggregated to their corresponding police precinct to construct day and night population estimates, and separate stop estimates were computed for daytime and nighttime intervals. For these analyses, we aggregated separate estimates of stops by day and night to compute total stop rates for each precinct.

Perhaps a more relevant comparison, however, is to the number of crimes committed by members of each ethnic group. For example, then New York City Police Commissioner Howard Safir stated (Safir 1999),

> The racial/ethnic distribution of the subjects of "stop and frisk" reports reflects the demographics of known violent crime suspects as reported by crime victims. Similarly, the demographics of arrestees in violent crimes also correspond with the demographics of known violent crime suspects.

Data on actual crimes are not available, of course, so as a proxy we use the number of arrests within New York City in the previous year, 1997, as recorded by the Division of Criminal Justice Services (DCJS) of New York State and categorized by ethnic group and crime type. This was deemed to be the best available measure of local crime rates categorized by ethnicity and directly address concerns such as Safir's that stop rates be related to the ethnicity of crime suspects. We use the previous year's DCJS arrest rates to represent the frequency of

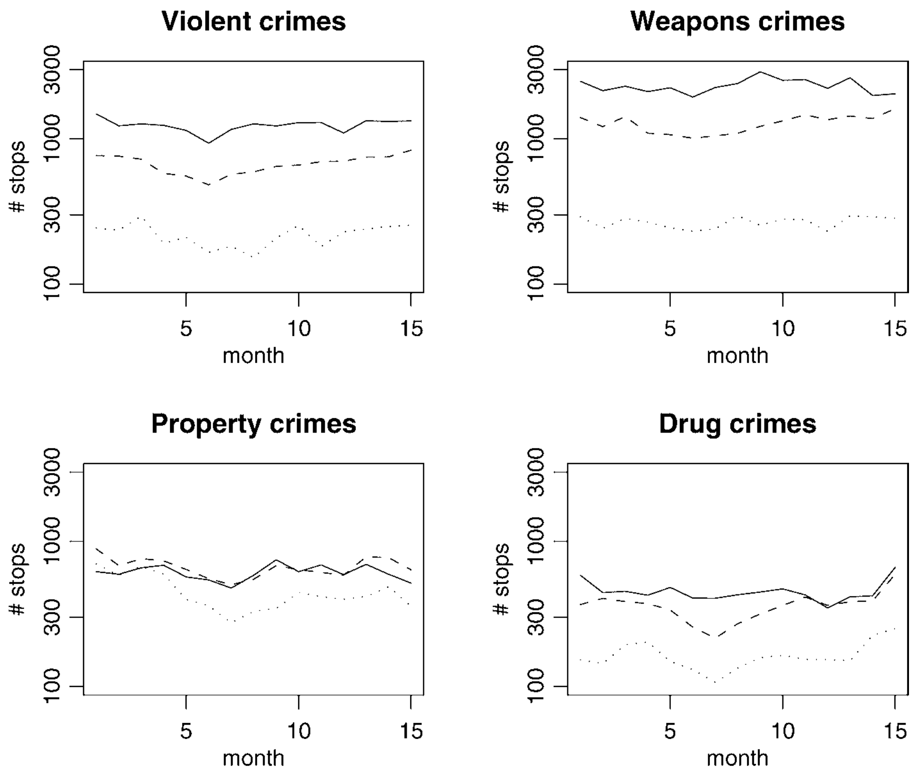


Figure 1. Number of police stops in each of 15 months, characterized by type of crime and ethnicity of person stopped (——, blacks; − − −−, Hispanics; ······, whites).

crimes that the police might suspect were committed by members of each ethnic group. When compared in that way, the ratio of stops to DCJS arrests was 1.24 for whites, 1.54 for blacks, and 1.72 for Hispanics; based on this comparison, blacks are stopped 23% more often than whites and Hispanics are stopped 39% more often than whites.

4. MODELS

The summaries given so far describe average rates for the whole city. But suppose that the police make more stops in high-crime areas but treat the different ethnic groups equally within any locality. Then the citywide ratios could show significant differences between ethnic groups even if stops were determined entirely by location rather than by ethnicity. To separate these two kinds of predictors, we performed multilevel analyses using the city's 75 precincts. Allowing precinct-level effects is consistent with theories of policing such as “broken windows” that emphasize local, neighborhood-level strategies (Wilson and Kelling 1982; Skogan 1990). Because it is possible that the patterns are systematically different in neighborhoods with different ethnic compositions, we divided the precincts into three categories in terms of their black population: precincts that were less than 10% black, 10–40% black, and more than 40% black. We also accounted for variation in stop rates between the precincts within each group. Each of the three categories represents roughly 1/3 of the precincts in the city, and we performed separate analyses for each set.

4.1 Hierarchical Poisson Regression Model

For each ethnic group $e = 1, 2, 3$ and precinct p, we modeled the number of stops, y_{ep}, using an overdispersed Poisson regression with indicators for ethnic groups, a hierarchical model for precincts, and n_{ep}, the number of DCJS arrests for that ethnic group in that precinct (multiplied by 15/12 to scale to a 15-month period), as a baseline or offset,

$$\begin{aligned} y_{ep} &\sim \text{Poisson}\left(\frac{15}{12} n_{ep} e^{\mu+\alpha_e+\beta_p+\epsilon_{ep}}\right), \\ \beta_p &\sim \text{N}(0, \sigma_\beta^2), \qquad (1) \\ \epsilon_{ep} &\sim \text{N}(0, \sigma_\epsilon^2), \end{aligned}$$

where the coefficients α_e (which we constrained to sum to 0) control for ethnic groups, the β_p's adjust for variation among precincts (with variance σ_β), and the ϵ_{ep}'s allow for overdispersion, that is, variation in the data beyond that explained by the Poisson model. We fit the model using Bayesian inference with a noninformative uniform prior distribution on the parameters μ, α, σ_β, and σ_ϵ.

In classical generalized linear modeling or generalized estimating equations, overdispersion can be estimated using a chi-squared statistic, with standard errors inflated by the square root of the estimated overdispersion (McCullagh and Nelder 1989). In our analysis, we are already using Bayesian inference to model the variation among precincts, and so the overdispersion simply represents another variance component in the model; the resulting inferences indeed have larger standard errors than would be obtained from the nonoverdispersed regression (which would correspond to $\sigma_\epsilon = 0$), and these posterior standard errors can be checked using, for example, cross-validation of precincts.

Of most interest, however, are the exponentiated coefficients $\exp(\alpha_e)$, which represent relative rates of stops compared with arrests, after controlling for precinct. By comparing stop rates to arrest rates, we can also separately analyze stops associated with different types of crimes. We conducted separate comparisons for violent crimes, weapons offenses, property crimes, and drug crimes. For each, we modeled the number of stops y_{ep} by ethnic group e and precinct p for that crime type, using as a baseline the DCJS arrest count n_{ep} for that ethnic group, precinct, and crime type. (The subsetting by crime type is implicit in this notation; to keep notation simple, we did not introduce an additional subscript for the four categories of crime.)

We thus estimated model (1) for 12 separate subsets of the data, corresponding to the four crime types and the three categories of precincts ($<$10% black population, 10–40% black, and $>$40% black). Computations were easily performed using the Bayesian software BUGS (Spiegelhalter, Thomas, Best, Gilks, and Lunn 1994, 2003), which implements Markov chain Monte Carlo simulation from R (R Project 2000; Sturtz, Ligges, and Gelman 2005). For each fit, we simulated three several independent Markov chains from different starting points, stopping when the simulations from each chain alone were as variable as those from all of the chains mixed together (Gelman and Rubin 1992). We then gathered the last half of the simulated chains and used these to compute posterior estimates and standard errors. For the analyses reported in this article, 10,000 iterations were always sufficient for mixing of the sequences. We report inferences using posterior means and standard deviations, which are reasonable summaries given the large sample size (see, e.g., Gelman, Carlin, Stern, and Rubin 2003, chap. 4).

4.2 Alternative Model Specifications

In addition to fitting model (1) as described earlier, we consider two forms of alternative specifications: first, fitting the same model but changing the batching of precincts, and second, altering the role played in the model by the previous year's arrests. We compare the fits under these alternative models to assess sensitivity to details of model specification.

Modeling Variability Across Precincts. The batching of precincts into three categories is convenient and makes sense, because neighborhoods with different levels of minority populations differ in many ways, including policing strategies applied to each type (Fagan and Davies 2000). Thus, fitting the model separately to each group of precincts is a way to include contextual effects. However, there is an arbitrariness to this division. We explore this by partitioning the precincts into different numbers of categories and seeing how the model estimates change.

Another approach to controlling for systematic variation among precincts is to include precinct-level predictors, which can be included along with the individual precinct-level effects in the multilevel model (see, e.g., Raudenbush and Bryk 2002). As discussed earlier, the precinct-level information that is of greatest interest and also has the greatest potential to affect our results, is the ethnic breakdown of the population. Thus we consider as regression predictors the proportion of black and Hispanic in the precinct, replacing model (1) by

$$y_{ep} \sim \text{Poisson}\left(\frac{15}{12} n_{ep} e^{\mu+\alpha_e+\zeta_1 z_{1p}+\zeta_2 z_{2p}+\beta_p+\epsilon_{ep}}\right), \qquad (2)$$

where z_{1p} and z_{2p} represent the proportion of the population in precinct p that are black and Hispanic. We also consider variants of model (2) including the quadratic terms, z_{1p}^2, z_{2p}^2, and $z_{1p}z_{2p}$, to examine sensitivity to nonlinearity.

Modeling the Relation of Stops to Previous Year's Arrests. We also consider different ways of using the number of DCJS arrests n_{ep} in the previous year, which plays the role of a baseline (or offset, in generalized linear models terminology) in model (1). Including the past arrest rate as an offset makes sense because we are interested in the rate of stops per crime, and we are using past arrests as a proxy for crime rate and for police expectations about the demographics of perpetrators. Another option is to include the logarithm of the number of past arrests as a linear predictor instead,

$$y_{ep} \sim \text{Poisson}\left(\frac{15}{12} e^{\gamma \log n_{ep}+\mu+\alpha_e+\beta_p+\epsilon_{ep}}\right). \qquad (3)$$

Model (3) reduces to the offset model (1) if $\gamma = 1$. We thus can fit (3) and see whether the inferences for α_e change compared with the earlier model that implicitly fixes γ to 1.

We can take this idea further by modeling past arrests as a proxy of the actual crime rate. We attempt to do this in two ways, is each approach labeling the true crime rate for each ethnicity in each precinct as θ_{ep}, with separate hierarchical Poisson regressions for this year's stops and last year's arrests (as always, including the factor $\frac{15}{12}$ to account for our 15 months of stop data). In the first formulation, we model last year's arrests as Poisson distributed with mean θ,

$$\begin{aligned} y_{ep} &\sim \text{Poisson}\left(\frac{15}{12}\theta_{ep} e^{\mu+\alpha_e+\beta_p+\epsilon_{ep}}\right), \\ n_{ep} &\sim \text{Poisson}(\theta_{ep}), \qquad (4) \\ \log\theta_{ep} &= \log N_{ep} + \tilde{\alpha}_e + \tilde{\beta}_p + \tilde{\epsilon}_{ep}. \end{aligned}$$

Here we are using N_{ep}, the population of ethnic group e in precinct p, as a baseline for the model of crime frequencies. The second-level error terms $\tilde{\beta}$ and $\tilde{\epsilon}$ are given normal hyperprior distributions as for model (1).

Our second two-stage model is similar to (4) but with the new error term $\tilde{\epsilon}$ moved to the model for n_{ep},

$$\begin{aligned} y_{ep} &\sim \text{Poisson}\left(\frac{15}{12}\theta_{ep} e^{\mu+\alpha_e+\beta_p+\epsilon_{ep}}\right), \\ n_{ep} &\sim \text{Poisson}(\theta_{ep} e^{\tilde{\epsilon}_{ep}}), \qquad (5) \\ \log\theta_{ep} &= \log N_{ep} + \tilde{\alpha}_e + \tilde{\beta}_p. \end{aligned}$$

Under this model, arrest rates n_{ep} are equal to the underlying crime rates, θ_{ep}, on average, but with overdispersion compared with the Poisson error distribution.

5. RESULTS

5.1 Primary Regression Analysis

Table 1 shows the estimates from model (1) fit to each of four crime types in each of three categories of precinct. The random-effects standard deviations σ_β and σ_ϵ are substantial, indicating the relevance of hierarchical modeling for these data. [Recall that these effects are all on the logarithmic scale, so that an

Table 1. Estimates and standard errors for the constant term μ, ethnicity parameters α_e, and the precinct-level and precinct-by-ethnicity–level variance parameters σ_β and σ_ϵ, for the hierarchical Poisson regression model (1), fit separately to three categories of precinct and four crime types

Proportion black in precinct	Parameter	Crime type: Violent	Weapons	Property	Drug
<10%	Intercept	$-.85_{(.07)}$	$.13_{(.07)}$	$-.58_{(.21)}$	$-1.62_{(.16)}$
	α_1 [blacks]	$.40_{(.06)}$	$.16_{(.05)}$	$-.32_{(.06)}$	$-.08_{(.09)}$
	α_2 [Hispanics]	$.13_{(.06)}$	$.12_{(.04)}$	$.32_{(.06)}$	$.17_{(.10)}$
	α_3 [whites]	$-.53_{(.06)}$	$-.28_{(.05)}$	$.00_{(.06)}$	$-.08_{(.09)}$
	σ_β	$.33_{(.08)}$	$.38_{(.08)}$	$1.19_{(.20)}$	$.87_{(.16)}$
	σ_ϵ	$.30_{(.04)}$	$.23_{(.04)}$	$.32_{(.04)}$	$.50_{(.07)}$
10–40%	Intercept	$-.97_{(.07)}$	$.42_{(.07)}$	$-.89_{(.16)}$	$-1.87_{(.13)}$
	α_1 [blacks]	$.38_{(.04)}$	$.24_{(.04)}$	$-.16_{(.06)}$	$-.05_{(.05)}$
	α_2 [Hispanics]	$.08_{(.04)}$	$.13_{(.04)}$	$.25_{(.06)}$	$.12_{(.06)}$
	α_3 [whites]	$-.46_{(.04)}$	$-.36_{(.04)}$	$-.08_{(.06)}$	$-.07_{(.05)}$
	σ_β	$.49_{(.07)}$	$.47_{(.07)}$	$1.21_{(.17)}$	$.90_{(.13)}$
	σ_ϵ	$.24_{(.03)}$	$.24_{(.03)}$	$.38_{(.04)}$	$.32_{(.04)}$
>40%	Intercept	$-1.58_{(.10)}$	$.29_{(.11)}$	$-1.15_{(.19)}$	$-2.62_{(.12)}$
	α_1 [blacks]	$.44_{(.06)}$	$.30_{(.07)}$	$-.03_{(.07)}$	$.09_{(.06)}$
	α_2 [Hispanics]	$.11_{(.06)}$	$.14_{(.07)}$	$.04_{(.07)}$	$.09_{(.07)}$
	α_3 [whites]	$-.55_{(.08)}$	$-.44_{(.08)}$	$-.01_{(.07)}$	$-.18_{(.09)}$
	σ_β	$.48_{(.10)}$	$.47_{(.11)}$	$.96_{(.18)}$	$.54_{(.11)}$
	σ_ϵ	$.24_{(.05)}$	$.37_{(.05)}$	$.42_{(.07)}$	$.28_{(.06)}$

NOTE: The estimates of $e^{\mu+\alpha_e}$ are displayed graphically in Figure 2, and alternative model specifications are shown in Table 3.

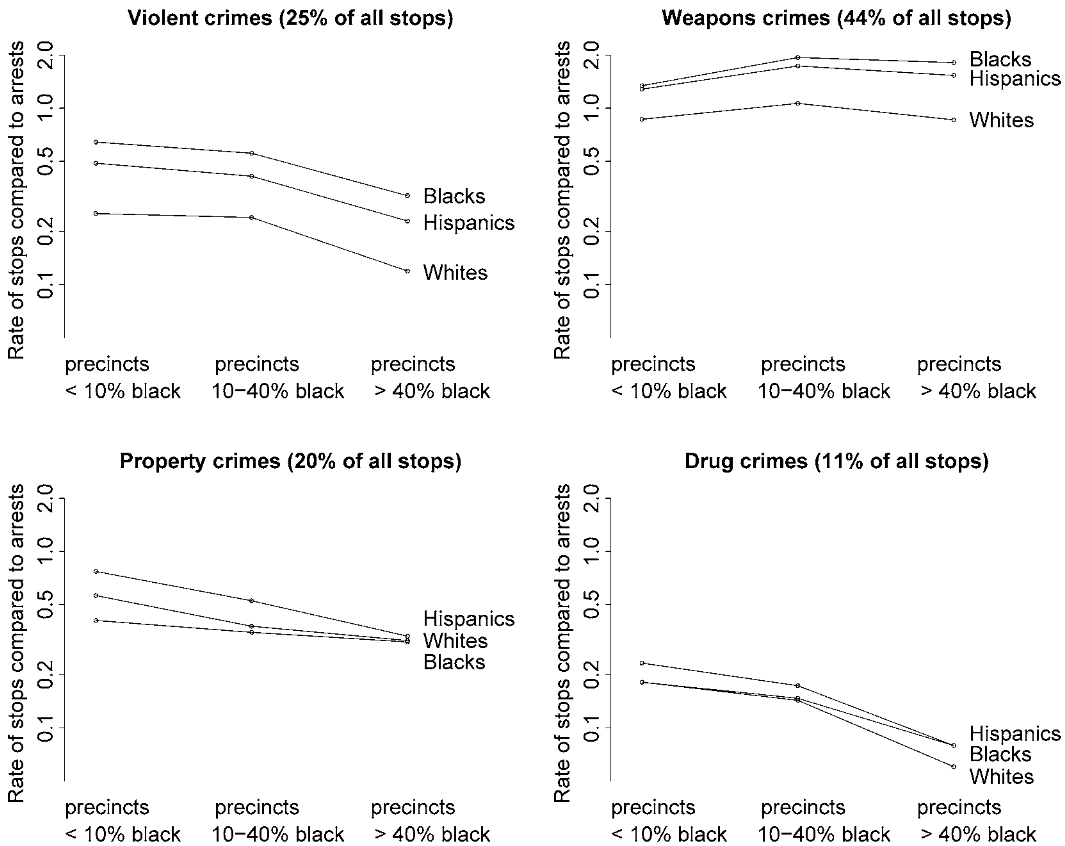


Figure 2. Estimated rates $e^{\mu+\alpha_e}$ at which people of different ethnic groups were stopped for different categories of crime, as estimated from hierarchical regressions (1) using previous year's arrests as a baseline and controlling for differences between precincts. Separate analyses were done for the precincts that had <10%, 10–40%, and >40% black population. For the most common stops—violent crimes and weapons offenses—blacks and Hispanics were stopped about twice as often as whites. Rates are plotted on a logarithmic scale. Numerical estimates and standard errors are given in Table 1.

effect of .3, for example, corresponds to a multiplicative effect of $\exp(.3) = 1.35$, or a 35% increase in the probability of being stopped.]

The parameters of most interest are the rates of stops (compared with previous year's arrests) for each ethnic group, $e^{\mu+\alpha_e}$, for $e = 1, 2, 3$. We display these graphically in Figure 2. Stops for violent crimes and weapons offenses were the most controversial aspect of the stop-and-frisk policy (and represent more than two-thirds of the stops), but for completeness we display all four categories of crime here.

Figure 2 shows that for the most frequent categories of stops—those associated with violent crimes and weapons offenses—blacks and Hispanics were much more likely to be stopped than whites, in all categories of precincts. For violent crimes, blacks and Hispanics were stopped 2.5 times and 1.9 times as often as whites, and for weapons crimes, blacks and Hispanics were stopped 1.8 times and 1.6 times as often as whites. In the less common categories of stops, whites were slightly more often stopped for property crimes and more often stopped for drug crimes in proportion to their previous year's arrests in any given precinct.

5.2 Alternative Forms of the Model

Fitting the alternative models described in Section 4.2 yielded results similar to those of our main analysis. We discuss each alternative model in turn.

Figure 3 displays the estimated rates of stops for violent crimes compared with the previous year's arrests for each of the three ethnic groups, for analyses dividing the precincts into 5, 10, and 15 categories ordered by the percentage of black population in the precinct. For simplicity, we give results only for violent crimes; these are typical of the alternative analyses for all four crime types. For each of the three graphs in Figure 3, the model is estimated separately for each of the three groups of precincts, and these estimates are connected in a line for each ethnic group. Compared with the upper-left plot in Figure 2, which shows the results from dividing the precincts into three categories, we see that dividing into more groups adds noise to the estimation but does not change the overall pattern of differences among the groups.

Table 2 shows the results from model (2), which is fit to all 75 precincts but controls for the proportions of blacks and Hispanics in precincts. The inferences are similar to those obtained from the main analysis discussed in Section 5.1. Including quadratic terms and interactions in the precinct-level model (2) and including the precinct-level predictors in the models fit to each of the three subsets of the data also had little effect on the parameters of interest, α_e.

Table 3 displays parameter estimates from the models that differently incorporate the previous year's arrest rates n_{ep}. For conciseness, results are displayed only for violent crimes, and for simplicity we include all 75 precincts in the models. (Similar results were obtained when fitting the model separately in each of three categories of precincts and for the other crime types.) The first two columns of Table 3 shows the result from our main model (1) and the alternative model (3), which includes $\log n_{ep}$ as a regression predictor. The two models differ only in that the first restricts γ to be 1, but as we can see, γ is estimated very close to 1 in the regression formulation, and the coefficients α_e remain essentially unchanged. (The intercept changes a bit because $\log n_{ep}$ does not have a mean of 0.)

The last two columns in Table 3 show the estimates from the two-stage regression models (4) and (5). The models differ in their estimates of the variance parameters σ_β and σ_ϵ, but the estimates of the key parameters α_e are essentially the same in the original model.

We also performed analyses including indicators for the month of arrest. These analyses did not add anything informative to the comparison of ethnic groups.

5.3 Hit Rates: Proportions of Stops That Led to Arrests

A different way to compare ethnic groups is to look at the fraction of stops on the street that lead to arrests. Most stops do not lead to arrests, and most arrests do not come from stops. In the analysis described earlier, we studied the rate at which the police stopped people of different groups. Now we look briefly at what happens with these stops.

In the period for which we have data, 1 in 7.9 whites stopped were arrested, compared with approximately 1 in 8.8 Hispanics and 1 in 9.5 blacks. These data are consistent with our general conclusion that the police are disproportionately stopping minorities; the stops of whites are more "efficient" and are more likely to lead to arrests, whereas those for blacks and Hispanics are more indiscriminate, and fewer of the persons stopped in

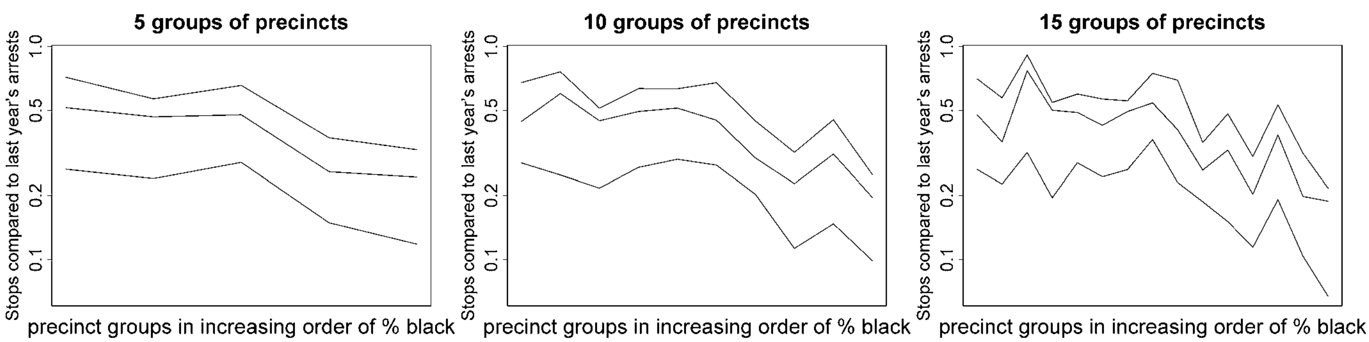


Figure 3. Estimated rates $e^{\mu+\alpha_e}$ at which people of different ethnic groups were stopped for violent crimes, as estimated from models dividing precincts into 5, 10, and 15 categories. For each graph, the top, middle, and lower lines correspond to blacks, Hispanics, and whites. These plots show the same general patterns as the model with three categories (the upper-left graph in Fig. 2) but with increasing levels of noise.

Table 2. Estimates and standard errors for the parameters of model (2) that includes proportion black and Hispanic as precinct-level predictors, fit to all 75 precincts

	Crime type			
Parameter	Violent	Weapons	Property	Drug
Intercept	$-.66_{(.08)}$	$.08_{(.11)}$	$-.14_{(.24)}$	$-.98_{(.17)}$
α_1 [blacks]	$.41_{(.03)}$	$.24_{(.03)}$	$-.19_{(.04)}$	$-.02_{(.04)}$
α_2 [Hispanics]	$.10_{(.03)}$	$.12_{(.03)}$	$.23_{(.04)}$	$.15_{(.04)}$
α_3 [whites]	$-.51_{(.03)}$	$-.36_{(.03)}$	$-.05_{(.04)}$	$-.13_{(.04)}$
ζ_1 [coeff. for prop. black]	$-1.22_{(.18)}$	$.10_{(.19)}$	$-1.11_{(.45)}$	$-1.71_{(.31)}$
ζ_2 [coeff. for prop. Hispanic]	$-.33_{(.23)}$	$.71_{(.27)}$	$-1.50_{(.57)}$	$-1.89_{(.41)}$
σ_β	$.40_{(.04)}$	$.43_{(.04)}$	$1.04_{(.09)}$	$.68_{(.06)}$
σ_ϵ	$.25_{(.02)}$	$.27_{(.02)}$	$.37_{(.03)}$	$.37_{(.03)}$

NOTE: The results for the parameters of interest, α_e, are similar to those obtained by fitting the basic model separately to each of three categories of precincts, as displayed in Table 1 and Figure 2. As before, the model is fit separately to the data from four different crime types.

these broader sweeps are actually arrested. It is perfectly reasonable for the police to make many stops that do not lead to arrests; the issue here is the comparison between ethnic groups.

This can also be understood in terms of simple economic theory (following the reasoning of Knowles, Persico, and Todd 2001 for police stops for suspected drugs). It is reasonable to suppose a diminishing return for stops in the sense that at some point, little benefit will be gained by stopping additional people. If the gain is approximately summarized by arrests, then diminishing returns mean that the probability that a stop will lead to an arrest—in economic terms, the marginal gain from stopping one more person—will decrease as the number of persons stopped increases. The stops of blacks and Hispanics were less "efficient" than those of whites, suggesting that the police have been using less rigorous standards when stopping members of minority groups. We found similar results when separately analyzing daytime and nighttime stops.

But this "hit rate" analysis can be criticized as unfair to the police, who are "damned if they do, damned if they don't." Relatively few of the stops of minorities led to arrests, and thus we conclude that police were more willing to stop minority group members with less reason. But we could also make the argument the other way around: Because a relatively high rate of whites stopped were arrested, we conclude that the police are biased against whites in the sense of arresting them too often. Analyses that examined the validity of arrests by race—that is, the proportion of arrests that lead to convictions—would help clarify this question. Unfortunately, such data are not readily available. We do not believe this latter interpretation, but it is hard to rule it out based on these data alone.

That is why we consider this part of the study to provide only *supporting* evidence. Our main analysis found that blacks and Hispanics were stopped disproportionately often (compared with their population or their crime rate, as measured by their rate of valid arrests in the previous year), and the secondary analysis of the hit rates or "arrest efficiency" of these stops is consistent with that finding.

6. DISCUSSION AND CONCLUSIONS

In the period for which we had data, the NYPD's records indicate that they were stopping blacks and Hispanics more often than whites, in comparison to both the populations of these groups and the best estimates of the rate of crimes committed by each group. After controlling for precincts, this pattern still holds. More specifically, for violent crimes and weapons offenses, blacks and Hispanics are stopped about twice as often as whites. In contrast, for the less common stops for property and drug crimes, whites and Hispanics are stopped more often than blacks, in comparison to the arrest rate for each ethnic group.

A related piece of evidence is that stops of blacks and Hispanics were less likely than those of whites to lead to arrest,

Table 3. Estimates and standard errors for parameters under model (1) and three alternative specifications for the previous year's arrests n_{ep}: treating $\log(n_{ep})$ as a predictor in the Poisson regression model (3), and the two-stage models (4) and (5)

	Model for previous year's arrests			
Parameter	Offset (1)	Regression (3)	Two-stage (5)	Two-stage (4)
Intercept	$-1.08_{(.06)}$	$-.94_{(.16)}$	$-1.07_{(.06)}$	$-1.13_{(.07)}$
α_1 [blacks]	$.40_{(.03)}$	$.41_{(.03)}$	$.40_{(.03)}$	$.42_{(.08)}$
α_2 [Hispanics]	$.10_{(.03)}$	$.10_{(.03)}$	$.10_{(.03)}$	$.14_{(.09)}$
α_3 [whites]	$-.50_{(.03)}$	$-.51_{(.03)}$	$-.50_{(.03)}$	$-.56_{(.09)}$
γ [coeff. for $\log n_{ep}$]		$.97_{(.03)}$		
σ_β	$.51_{(.05)}$	$.51_{(.05)}$	$.51_{(.05)}$	$.27_{(.12)}$
σ_ϵ	$.26_{(.02)}$	$.26_{(.02)}$	$.24_{(.02)}$	$.67_{(.04)}$

NOTE: For simplicity, results are displayed for violent crimes only, for the model fit to all 75 precincts. The three α_e parameters are nearly identical under all four models, with the specification affecting only the intercept.

suggesting that the standards were more relaxed for stopping minority group members. Two different scenarios might explain the lower "hit rates" for nonwhites, one that suggests targeting of minorities and another that suggests dynamics of racial stereotyping and a more passive form of racial preference. In the first scenario, police possibly used wider discretion and more relaxed constitutional standards in deciding to stop minority citizens. This explanation would conform to the scenario of "pretextual" stops discussed in several recent studies of motor vehicle stops (e.g., Lundman and Kaufman 2003) and suggests that the higher stop rates were intentional and purposive. Alternatively, police could simply form the perception of "suspicion" more often based on a broader interpretation of the social cues that capture police attention and evoke official reactions (Alpert et al. 2005). The latter explanation conforms more closely to a social-psychological process of racial stereotyping, where the attribution of suspicion is more readily attached to specific behaviors and contexts for minorities than it might be for whites (Thompson 1999; Richardson and Pittinsky 2005).

We did find evidence of stops that are best explained as "racial incongruity" stops: high rates of minority stops in predominantly white precincts. Indeed, being "out of place" is often a trigger for suspicion (Alpert et al. 2005; Gould and Mastrofski 2004). Racial incongruity stops are most prominent in racially homogeneous areas. For example, we observed high stop rates of African-Americans in the predominantly white 19th Precinct, a sign of race-based selection of citizens for police interdiction. We also observed high stop rates for whites in several precincts in the Bronx, especially for drug crimes, most likely evidence that white drug buyers were entering predominantly minority neighborhoods where street drug markets are common. Overall, however, these were relatively infrequent events that produced misleading stop rates due to the population skew in such precincts.

To briefly summarize our findings, blacks and Hispanics represented 51% and 33% of the stops while representing only 26% and 24% of the New York City population. Compared with the number of arrests of each group in the previous year (used as a proxy for the rate of criminal behavior), blacks were stopped 23% more often than whites and Hispanics were stopped 39% more often than whites. Controlling for precinct actually increased these discrepancies, with minorities between 1.5 and 2.5 times as often as whites (compared with the groups' previous arrest rates in the precincts where they were stopped) for the most common categories of stops (violent crimes and drug crimes), with smaller differences for property and drug crimes. The differences in stop rates among ethnic groups are real, substantial, and not explained by previous arrest rates or precincts.

Our findings do not necessarily imply that the NYPD was acting in an unfair or racist manner, however. It is quite reasonable to suppose that effective policing requires stopping and questioning many people to gather information about any given crime.

In the context of some difficult relations between the police and ethnic minority communities in New York City, it is useful to have some quantitative sense of the issues in dispute. Given that there have been complaints about the frequency with which the police have been stopping blacks and Hispanics, it is relevant to know that this is indeed a statistical pattern. The NYPD then has the opportunity to explain their policies to the affected communities.

In the years since this study was conducted, an extensive monitoring system was put into place that would accomplish two goals. First, procedures were developed and implemented that permitted monitoring of officers' compliance with the mandates of the NYPD Patrol Guide for accurate and comprehensive recording of all police stops. Second, the new forms were entered into databases that would permit continuous monitoring of the racial proportionality of stops and their outcomes (e.g., frisks, arrests). When coupled with accurate reporting on race-specific measures of crime and arrest, the new procedures and monitoring requirements will ensure that inquiries similar to this study can be institutionalized as part of a framework of accountability mechanisms.

[Received March 2004. Revised December 2005.]

REFERENCES

Alpert, G., MacDonald, J. H., and Dunham, R. G. (2005), "Police Suspicion and Discretionary Decision Making During Citizen Stops," *Criminology*, 43, 407–434.

Ayres, I. (2002a), "Outcome Tests of Racial Disparities in Police Practices," *Justice Research and Policy*, 4, 131–142.

——— (2002b), *Pervasive Prejudice: Unconventional Evidence of Race and Gender Discrimination*, Chicago: University of Chicago Press.

Bittner, E. (1970), *Functions of the Police in Modern Society: A Reviev of Background Factors, Current Practices and Possible Role Models*, Rockville, MD: National Institute of Mental Health.

Bratton, W., and Knobler, P. (1998), *Turnaround*, New York: Norton.

Brown v. Texas (1979), 443 U.S. 43, U.S. Supreme Court.

Brown v. Oneonta (2000), 221 F.3d 329, 2nd Ciruit Court.

Cole, D. (1999), *No Equal Justice: Race and Class in the Criminal Justice System*, New York: New Press.

Dharmapala, D., and Ross, S. L. (2004), "Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence," *Contributions to Economic Policy and Analysis*, 3, article 12; available at *www.bepress.com/bejeap/contributions/vol13/iss1/art12*.

Dominitz, J., and Knowles, J. (2005), "Crime Minimization and Racial Bias: What Can We Learn From Police Search Data?" PIER Working Paper Archive 05-019, University of Pennsylvania, Penn Institute for Economic Research, Dept. of Economics.

Durose, M. R., Schmitt, E. L., and Langan, P. A. (2005), "Contacts Between Police and the Public: Findings From the 2002 National Survey," NCJ 207845, Bureau of Justice Statistics, U.S. Department of Justice.

Eck, J., and Maguire, E. R. (2000), "Have Changes in Policing Reduced Violent Crime: An Assessment of the Evidence," in *The Crime Drop in America*, eds. A. Blumstein and J. Wallman, New York: Cambridge University Press, pp. 207–265.

Fagan, J. (2002), "Law, Social Science and Racial Profiling," *Justice, Research and Policy*, 4, 104–129.

Fagan, J., and Davies, G. (2000), "Street Stops and Broken Windows: Terry, Race, and Disorder in New York City," *Fordham Urban Law Journal*, 28, 457–504.

——— (2003), "Policing Guns: Order Maintenance and Crime Control in New York," in *Guns, Crime, and Punishment in America*, ed. B. Harcourt, New York: New York University Press, pp. 191–221.

Fagan, J., Zimring, F. E., and Kim, J. (1998), "Declining Homicide in New York: A Tale of Two Trends," *Journal of Criminal Law and Criminology*, 88, 1277–1324.

Fridell, L., Lunney, R., Diamond, D., Kubu, B., Scott, M., and Laing, C. (2001), "Racially Biased Policing: A Principled Response," Police Executive Research Forum; available at *www.policeforum.org/racial.html*.

Garrett, B. (2001), "Remedying Racial Profiling Through Partnership, Statistics, and Reflective Policing," *Columbia Human Rights Law Review*, 33, 41–148.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003), *Bayesian Data Analysis* (2nd ed.), London: Chapman & Hall.

Gelman, A., and Rubin, D. B. (1992), "Inference From Iterative Simulation Using Multiple Sequences" (with discussion), *Statistical Science*, 7, 457–511.

Goldberg, J. (1999), "The Color of Suspicion," *New York Times Magazine*, June 20, 51.

Gould, J., and Mastrofski, S. (2004), "Suspect Searches: Assess Police Behavior Under the U.S. Constitution," *Criminology and Public Policy*, 3, 315–361.

Greene, J. (1999), "Zero Tolerance: A Case Study of Police Policies and Practices in New York City," *Crime and Delinquency*, 45, 171–199.

Gross, S. R., and Barnes, K. Y. (2002), "Road Work: Racial Profiling and Drug Interdiction on the Highway," *Michigan Law Review*, 101, 653–754.

Gross, S. R., and Livingston, D. (2002), "Racial Profiling Under Attack," *Columbia Law Review*, 102, 1413–1438.

Harcourt, B. E. (1998), "Reflecting on the Subject: A Critique of the Social Influence Conception of Deterrence, the Broken Windows Theory, and Order-Maintenance Policing New York Style," *Michigan Law Review*, 97, 291.

——— (2001), *Illusion of Order: The False Promise of Broken Windows Policing*, Cambridge, MA: Harvard University Press.

Harris, D. A. (1999), "The Stories, the Statistics, and the Law: Why 'Driving While Black' Matters," *Minnesota Law Review*, 84, 265–326.

——— (2002), *Profiles in Injustice: Why Racial Profiling Cannot Work*, New York: New Press.

Hemenway, D. (1997), "The Myth of Millions of Annual Self-Defense Gun Uses: A Case Study of Survey Overestimates of Rare Events," *Chance*, 10, 6–10.

Illinois v. Wardlow (2000), 528 U.S. 119, U.S. Supreme Court.

Kelling, G., and Cole, C. (1996), *Fixing Broken Windows*, New York: Free Press.

Kelvin Daniels et al. v. City of New York (2004), Stipulation of Settlement, 99 Civ 1695 (SAS), ordered January 9, 2004; filed January 13, 2004.

Kennedy, R. (1997), *Race, Crime and Law*, Cambridge, MA: Harvard University Press.

Knowles, J., Persico, N., and Todd, P. (2001), "Racial Bias in Motor-Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 109, 203–229.

Lamberth, J. (1997), "Report of John Lamberth, Ph.D.," American Civil Liberties Union; available at *www.aclu.org/court/lamberth.html*.

Langan, P., Greenfeld, L. A., Smith, S. K., Durose, M. R., and Levin, D. J. (2001), "Contacts Between Police and the Public: Findings From the 1999 National Survey," NCJ 184957, Bureau of Justice Statistics, U.S. Department of Justice.

Loury, G. (2002), *The Anatomy of Racial Inequality*, Cambridge, MA: Harvard University Press.

Lundman, R. J., and Kaufman, R. L. (2003), "Driving While Black: Effects of Race, Ethnicity, and Gender on Citizen Self-Reports of Traffic Stops and Police Actions," *Criminology*, 41, 195–220.

MacDonald, H. (2001), "The Myth of Racial Profiling," *City Journal*, 11, 2–5.

Massey, D., and Denton, N. (1993), *American Apartheid: Segregation and the Making of the Underclass*, Cambridge, MA: Harvard University Press.

McCullagh, P., and Nelder, J. A. (1989), *Generalized Linear Models* (2nd ed.), New York: Chapman & Hall.

Miller, J. (2000), "Profiling Populations Available for Stops and Searches," Police Research Series Paper 121, Home Office Research, Development and Statistics Directorate, United Kingdom.

Musto, D. (1973), *The American Disease*, New Haven, CT: Yale University Press.

Pager, D. (2003), "The Mark of a Criminal Record," *American Journal of Sociology*, 108, 937–975.

People v. DeBour (1976), 40 N.Y.2d 210, 386 N.Y.S.2d 375.

People v. Holmes (1996), 89 N.Y.2d 838, 652 N.Y.S.2d 725.

R Project (2000), "The R Project for Statistical Computing," available at *www.r-project.org*.

Raudenbush, S. W., and Bryk, A. S. (2002), *Hierarchical Linear Models* (2nd ed.), Thousand Oaks, CA: Sage.

Reiss, A. (1971), *The Police and the Public*, New Haven, CT: Yale University Press.

Richardson, M., and Pittinsky, T. L. (2005), "The Mistaken Assumption of Intentionality in Equal Protection Law: Psychological Science and the Interpretation of the Fourteenth Amendment," KSG Working Paper RWP05-011; available at *ssrn.com/abstract=722647*.

Rudovsky, D. (2001), "Law Enforcement by Stereotypes and Serendipity: Racial Profiling and Stops and Searches Without Cause," *University of Pennsylvania Journal of Constitutional Law*, 3, 296–366.

Russell, K. K. (2002), "Racial Profiling: A Status Report of the Legal, Legislative, and Empirical Literature," *Rutgers Race and the Law Review*, 3, 61–81.

Safir, H. (1999), "Statement Before the New York City Council Public Safety Commission," April 19; cited in Spitzer (1999).

Sampson, R. J., and Raudenbush, S. W. (1999), "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods," *American Journal of Sociology*, 105, 622–630.

Sampson, R. J., Raudenbush, S. W., and Earls, F. (1997), "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy," *Science*, 277, 918–924.

Skogan, W. (1990), *Disorder and Decline: Crime and the Spiral of Decay in American Cities*, Berkeley, CA: University of California Press.

Skogan, W., and Frydl, K. (eds.) (2004), *The Evidence on Policing: Fairness and Efffectiveness in Law Enforcement*, Washington, DC: National Academy Press.

Skolnick, J., and Capolovitz, A. (2001), "Guns, Drugs and Profiling: Ways to Target Guns and Minimize Racial Profiling," *Arizona Law Review*, 43, 413–448.

Smith, D. A. (1986), "The Neighborhood Context of Police Behavior," in *Communities and Crime: Crime and Justice: An Annual Review of Research*, eds. A. J. Reiss and M. Tonry, Chicago: University of Chicago Press, pp. 313–342.

Smith, M., and Alpert, G. (2002), "Searching for Direction: Courts, Social Science, and the Adjudication of Racial Profiling Claims," *Justice Quarterly*, 19, 673–704.

Smith, M. R., Makarios, M., and Alpert, G. P. (2006), "Differential Suspicion: Theory Specification and Gender Effects in the Traffic Stop Context," *Justice Quarterly*, 23, 271–295.

Spiegelhalter, D., Thomas, A., Best, N., Gilks, W., and Lunn, D. (1994, 2003), "BUGS: Bayesian Inference Using Gibbs Sampling," MRC Biostatistics Unit, Cambridge, U.K.; available at *www.mrc-bsu.cam.ac.uk/bugs/*.

Spitzer, E. (1999), "The New York City Police Department's "Stop and Frisk" Practices," Office of the New York State Attorney General; available at *www.oag.state.ny.us/press/reports/stop_frisk/stop_frisk.html*.

Sturtz, S., Ligges, U., and Gelman, A. (2005), "R2WinBUGS: A Package for Running WinBUGS From R," *Journal of Statistical Software*, 12 (3), 1–16.

Sykes, R. E., and Clark, J. P. (1976), "A Theory of Deference Exchange in Police-Civilian Encounters," *American Journal of Sociology*, 81, 587–600.

Taylor, R. B. (2000), *Breaking Away From Broken Windows: Baltimore Neighborhoods and the Nationwide Fight Against Crime, Grime, Fear, and Decline*, Boulder, CO: Westview Press.

Terry v. Ohio (1968), 392 U.S. 1, U.S. Supreme Court.

Thompson, A. (1999), "Stopping the Usual Suspects: Race and the Fourth Amendment," *New York University Law Review*, 74, 956–1013.

Tyler, T. R., and Huo, Y. J. (2003), *Trust in the Law*, New York: Russell Sage Foundation.

U.S. Commission on Civil Rights (2000), "Police Practices and Civil Rights in New York City," available at *www.uscrr.gov/pubs/nypolice/main.htm*.

U.S. v. Martinez-Fuerte (1976), 428 U.S. 543. U.S. Supreme Court.

Veneiro, P., and Zoubeck, P. H. (1999), "Interim Report of the State Police Review Team Regarding Allegations of Racial Profiling," Office of the New Jersey State Attorney General.

Walker, S. (2001), "Searching for the Denominator: Problems With Police Traffic Stop Data and an Early Warning System Solution," *Justice Research and Policy*, 3, 63–95.

Wang, A. (2001), "Illinois v. Wardlow and the Crisis of Legitimacy: An Argument for a "Real Cost" Balancing Test," *Law and Inequality*, 19, 1–30.

Weidner, R. R., Frase, R., and Pardoe, I. (2004), "Explaining Sentence Severity in Large Urban Counties: A Multilevel Analysis of Contextual and Case-Level Factors," *Prison Journal*, 84, 184–207.

Weitzer, R. (2000), "Racialized Policing: Residents' Perceptions in Three Neighborhoods," *Law and Society Review*, 34, 129–155.

Weitzer, R., and Tuch, S. A. (2002), "Perceptions of Racial Profiling: Race, Class and Personal Experience," *Criminology*, 50, 435–456.

Whren et al. v. U.S. (1996), 517 U.S. 806, U.S. Supreme Court.

Wilson, J. Q., and Kelling, G. L. (1982), "The Police and Neighborhood Safety: Broken Windows," *Atlantic Monthly*, March, 29–38.

Zimring, F. E. (2006), *The Great American Crime Decline*, Oxford, U.K.: Oxford University Press.

Zingraff, M., Mason, H. M., Smith, W. R., Tomaskovic-Devey, D., Warren, P., McMurray, H. L., and Fenlon, C. R., et al. (2000), "Evaluating North Carolina State Highway Patrol Data: Citations, Warnings and Searches in 1998," available at *www.nccrimecontrol.org/shp/ncshpreport.htm*.

Exhibit K

Reference Guide on Multiple Regression

DANIEL L. RUBINFELD

Daniel L. Rubinfeld, Ph.D., is Robert L. Bridges Professor of Law and Professor of Economics Emeritus, University of California, Berkeley, and Visiting Professor of Law at New York University Law School.

CONTENTS

Not all possible variables that might influence the dependent variable can be included if the analysis is to be successful; some cannot be measured, and others may make little difference.[30] If a preliminary analysis shows the unexplained portion of the multiple regression to be unacceptably high, the expert may seek to discover whether some previously undetected variable is missing from the analysis.[31]

Failure to include a major explanatory variable that is correlated with the variable of interest in a regression model may cause an included variable to be credited with an effect that actually is caused by the excluded variable.[32] In general, omitted variables that are correlated with the dependent variable reduce the probative value of the regression analysis. The importance of omitting a relevant variable depends on the strength of the relationship between the omitted variable and the dependent variable and the strength of the correlation between the omitted variable and the explanatory variables of interest. Other things being equal, the greater the correlation between the omitted variable and the variable of interest, the greater the bias caused by the omission. As a result, the omission of an important variable may lead to inferences made from regression analyses that do not assist the trier of fact.[33]

discrimination), *cert. denied*, 504 U.S. 913 (1992). Whether a particular variable reflects "legitimate" considerations or itself reflects or incorporates illegitimate biases is a recurring theme in discrimination cases. *See, e.g.,* Smith v. Virginia Commonwealth Univ., 84 F.3d 672, 677 (4th Cir. 1996) (en banc) (suggesting that whether "performance factors" should have been included in a regression analysis was a question of material fact); *id.* at 681–82 (Luttig, J., concurring in part) (suggesting that the failure of the regression analysis to include "performance factors" rendered it so incomplete as to be inadmissible); *id.* at 690–91 (Michael, J., dissenting) (suggesting that the regression analysis properly excluded "performance factors"); *see also* Diehl v. Xerox Corp., 933 F. Supp. 1157, 1168 (W.D.N.Y. 1996).

30. The summary effect of the excluded variables shows up as a random error term in the regression model, as does any modeling error. *See* Appendix, *infra,* for details. *But see* David W. Peterson, *Reference Guide on Multiple Regression*, 36 Jurimetrics J. 213, 214 n.2 (1996) (review essay) (asserting that "the presumption that the combined effect of the explanatory variables omitted from the model are uncorrelated with the included explanatory variables" is "a knife-edge condition . . . not likely to occur").

31. A very low R-squared (R^2) is one indication of an unexplained portion of the multiple regression model that is unacceptably high. However, the inference that one makes from a particular value of R^2 will depend, of necessity, on the context of the particular issues under study and the particular dataset that is being analyzed. For reasons discussed in the Appendix, a low R^2 does not necessarily imply a poor model (and vice versa).

32. Technically, the omission of explanatory variables that are correlated with the variable of interest can cause biased estimates of regression parameters.

33. *See* Bazemore v. Friday, 751 F.2d 662, 671–72 (4th Cir. 1984) (upholding the district court's refusal to accept a multiple regression analysis as proof of discrimination by a preponderance of the evidence, the court of appeals stated that, although the regression used four variable factors (race, education, tenure, and job title), the failure to use other factors, including pay increases that varied by county, precluded their introduction into evidence), *aff'd in part*, *vacated in part*, 478 U.S. 385 (1986).

Note, however, that in *Sobel v. Yeshiva University,* 839 F.2d 18, 33, 34 (2d Cir. 1988), *cert. denied*, 490 U.S. 1105 (1989), the court made clear that "a [Title VII] defendant challenging the validity of