# The End of Intuition-Based High-Crime Areas

Ben Grunwald* and Jeffrey Fagan**

*In 2000, the Supreme Court held in* Illinois v. Wardlow *that a suspect's presence in a "high-crime area" is relevant in determining whether an officer has reasonable suspicion to conduct an investigative stop. Despite the importance of the decision, the Court provided no guidance about what that standard means, and over fifteen years later, we still have no idea how police officers understand and apply it in practice. This Article conducts the first empirical analysis of* Wardlow *by examining data on over two million investigative stops conducted by the New York Police Department from 2007 to 2012.*

*Our results suggest that* Wardlow *may have been wrongly decided. Specifically, we find evidence that officers often assess whether areas are high crime using a very broad geographic lens; that they call almost every block in the city high crime; that their assessments of whether an area is high crime are nearly uncorrelated with actual crime rates; that the suspect's race predicts whether an officer calls an area high crime as well as the actual crime rate; that the racial composition of the area and the identity of the officer are stronger predictors of whether an officer calls an area high crime than the crime rate itself; and that stops are less or as likely to result in the detection of contraband when an officer invokes high-crime area as a*

---

*basis of a stop. We conclude with several policy proposals for courts, police departments, and scholars to help address these problems in the doctrine.*

INTRODUCTION

Every year, police officers stop and frisk millions of pedestrians on the street.[1] One of the most common justifications they cite for these stops is that the suspect was located in a "high-crime area" (HCA).[2] In 2000, the Supreme Court gave formal approval to that practice in *Illinois v. Wardlow*, by holding that a suspect's presence in a high-crime area is relevant in determining whether an officer has reasonable suspicion to conduct a stop.[3] In doing so, the Court sanctioned a dramatic expansion in police discretion that has impacted "almost every" case challenging the constitutionality of a stop.[4]

Despite the importance of the decision, the Court provided remarkably little guidance on how to interpret and implement the high-crime area standard in practice. Indeed, the opinion said nothing at all about what "high-crime area" means, and the lower courts have made little progress filling the gap. As a result, officers haven't been told how to apply the high-crime area standard—how to think about its proper geographic scope, its relevant temporal horizon, or about the kinds of crimes that are most relevant.[5] *Wardlow* also said nothing about the relevant evidentiary standards for establishing that an area is high crime. In response, the lower courts have been remarkably lax in scrutinizing officers' claims about high-crime areas. The most common approach is to defer to the

---

1. *See* ELIZABETH DAVIS ET AL., CONTACTS BETWEEN POLICE AND THE PUBLIC, 2015, at 4 (2018), https://www.bjs.gov/content/pub/pdf/cpp15.pdf [https://perma.cc/GGS7-L3FM] (estimating that American police departments conducted 2.5 million street stops in 2015 against residents age sixteen or older); *see also* Chris Palmer, *Philly Police Decreasing Use of Stop-and-Frisk, Officials Say*, INQUIRER (May 2, 2017), http://www2.philly.com/philly/news/crime/Philly-Police-decreasing-use-of-stop-and-frisk-officials-say.html [https://perma.cc/T3T2-RQC8] (reporting that the Philadelphia Police Department conducted 140,000 stops in 2016); Jeremy Gorner & Dan Hinkel, *New Report Shows Chicago Police Street Stops Down, Minorities Still Stopped More*, CHI. TRIB. (Mar. 24, 2017), http://www.chicagotribune.com/news/local/breaking/ct-chicago-police-stop-and-frisk-report-met-20170324-story.html [https://perma.cc/M6FT-X2HM] (reporting that the Chicago Police Department conducted 54,000 stops in the first six months of 2016); Justin Fenton, *State Police Don't Analyze Stop & Frisk Data, Either*, BALT. SUN (Dec. 14, 2013), http://www.baltimoresun.com/news/maryland/sun-investigates/bs-md-sun-investigates-stop-and-frisk-20131214-story.html [https://perma.cc/VMW7-5LJT] (reporting that the Baltimore Police Department conducted 120,000 stops in 2012).

2. *See infra* Part III.

3. 528 U.S. 119, 124 (2000). A high-crime area, on its own, cannot establish reasonable suspicion. Instead, it can increase suspicion by providing an additional factor to other factors, or by enhancing the salience of other factors that, outside the context of a "high crime area," may not be sufficient to justify a stop. *See* People v. Howard, 542 N.Y.S.2d 536, 538 (N.Y. App. Div. 1989) (holding that absent additional factors, the 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).

4. Andrew Guthrie Ferguson & Damien Bernache, *The "High Crime Area" Question: Requiring Verifiable and Quantifiable Evidence for Fourth Amendment Reasonable Suspicion Analysis*, 57 AM. U. L. REV. 1587, 1590 (2008).

5. *See id.* at 1618.

expertise of the police officer, often adopting his[6] bare testimony that an area is "high crime" without additional proof.[7]

In the absence of a legal definition of high-crime areas and of meaningful judicial scrutiny, police officers enjoy wide discretion to define high-crime areas however they want. The wisdom of *Wardlow* as a constitutional doctrine thus depends heavily on how police officers exercise their discretion while implementing it in practice.

We argue that *Wardlow* depends on at least three unspoken empirical assumptions. The first assumption concerns the geographic scope of a high-crime area. The few lower courts that have confronted this question have generally agreed that high-crime areas should be analyzed through a granular geographic lens—more like a street block or intersection than a neighborhood or city.[8]

The second assumption is that officers' assessments of high-crime areas are relatively accurate. There are some good reasons to question that assumption. For one thing, officers may not always be aware of actual crime rates, which can fluctuate over time. Their assessments might also be skewed by bureaucratic pressures to increase the number of stops they conduct even if they lack constitutional justification.[9] And officers' assessments of high-crime areas might also be influenced by racial and socioeconomic biases based on the characteristics of suspects and neighborhoods in which their stops take place.[10]

---

6.    Throughout the Article, we use "he/him" pronouns, both because the vast majority officers in the NYPD—82 percent—are men and because our data do not allow us to separate male and female officers in our analysis. *See What is the Gender Breakdown of Active NYPD Officers?*, CIVILIAN COMPLAINT REVIEW BD., https://www1.nyc.gov/site/ccrb/policy/data-transparency-initiative-mos.page#gender [https://perma.cc/2PP2-55W6].

7.    *See* Ferguson & Bernache, *supra* note 4, at 1607 ("[T]he majority of jurisdictions . . . primarily have relied on an officer's testimony that an area is a 'high-crime area' without much analysis as to the basis of that conclusion."); *see also* Lenese C. Herbert, *Can't You See What I'm Saying? Making Expressive Conduct a Crime in High-Crime Areas*, 9 GEO. J. ON POVERTY L. & POL'Y 135, 135 (2002) ("As an eager young Assistant United States Attorney who 'papered' countless complaints, conducted numerous hearings, and tried a substantial number of cases, I learned how to decode police officer jargon and law enforcement terminology. One of the most commonly used—yet seldom defined—phrases was 'high crime area.' . . . [In court] judges rarely challenged the proffered label or required its definition. Judges never asked officers for data to support assertions that an area was high-crime."); United States v. Montero-Camargo, 208 F.3d 1122, 1143 (9th Cir. 2000) (Kozinski, J., concurring) ("[M]y colleagues don't even pause to ask the questions. To them, it's a high crime area, because the officers say it's a high crime area."); *see, e.g.*, State v. Morgan, 539 N.W.2d 887, 892 (Wis. 1995) ("[W]e find that an officer's perception of an area as 'high-crime' can be a factor justifying a search."); Riley v. Commonwealth, 412 S.E.2d 724, 726 (Va. Ct. App. 1992) (explaining that the officer testified that the stop took place in a "high crime area").

8.    *See infra* notes 85–87 and accompanying text.

9.    *See, e.g.*, Floyd v. New York, 959 F. Supp. 2d 540, 591–602 (S.D.N.Y. 2013) ("*Floyd I*") (documenting live testimony, depositions, roll call recordings of supervisors, internal NYPD documents, and survey results and concluding that the most plausible explanation for a 700 percent increase in stops from 2002 to 2011 was the result of "significant pressure" on police officers "to increase their stop activity").

10.    *See infra* Part II.C.

The third empirical assumption concerns predictive power. Like any other Fourth Amendment factor, a suspect's presence in a high-crime area only supports reasonable suspicion if that fact predicts, on average, whether a suspect is engaged in crime. *Wardlow* thus assumes that, controlling for other stated bases of reasonable suspicion, there is a higher probability that a suspect is engaged in a crime where the officer invokes high-crime area as a basis of a stop.

Nearly two decades have passed since the Supreme Court issued *Wardlow*, and yet we have almost no evidence about how police officers apply the high-crime area standard. We therefore don't know whether any of these empirical assumptions are satisfied in practice.

Our goal in this Article is to evaluate *Wardlow* by testing its empirical assumptions directly. To do so, we use a dataset of over two million police stops conducted by the New York Police Department (NYPD) between 2007 and 2012. The data derive from forms that officers are required to complete after every stop. The forms collect rich information on suspect demographics and the precise geographic location of each stop. The data also contain anonymized officer identifiers, which allow us to observe how the same officer behaves in different areas, and how different officers behave in the same areas. And, most important for our purposes, the forms require officers to check off a series of roughly twenty boxes, indicating the bases of suspicion that justified the stop. Fortunately, one of those boxes is for high-crime areas. We merged this dataset with crime statistics and racial and demographic information on small geographic areas in New York City.

Of course, we need to be careful about how we interpret our stop-form dataset. One possibility is that it tells us about the *ex ante*, subjective mental state of a police officer—that is, it describes the reasons an officer believed a stop was lawful moments before he carried it out. To a limited extent, we hope we can learn something about that internal mental process, but the data face significant limitations to serve that purpose. Indeed, officers fill out the form after completing their stops and they may therefore engage in post-hoc rationalization. Still, our data may offer a very rough proxy of what officers were thinking in the moment—a proxy that's better than anything else currently available.

Perhaps a more fitting interpretation of the stop-form data is that they describe the *ex post*, objective factors a police officer would use to justify a stop if he were ever asked to do so in court. This objective perspective is particularly important in the Fourth Amendment context where, under *Whren v. United States*,[11] the officer's subjective mental state is irrelevant in assessing whether a stop is unconstitutional. The data are well suited to illuminate that objective

---

11.     517 U.S. 806, 813 (1996); *see also* United States v. Willis, 431 F.3d 709, 716 (9th Cir. 2005) ("To the extent the magistrate judge made the same mistake as the dissent, by finding reasonable suspicion for a traffic stop lacking based on the officer's subjective motivations, we reverse. The parsing of police motives—as opposed to 'articulable facts'—is precisely what *Whren* tells us we may not do.") (citations omitted).

perspective. For one thing, just a few years before our data begin, the check boxes on the stop forms were created as a result of a lawsuit against the NYPD to require the department to document the bases of suspicion in every stop.[12] For another, when a stop is challenged at a suppression hearing, officers are incentivized to give testimony consistent with the contents of their stop form. Indeed, the form is typically discoverable, which means the defense can impeach an officer whose testimony deviates from it.[13] For these reasons, our data appear well suited for examining the objective factors—including the high-crime area factor—that an officer would raise to justify each stop.

Turning to our results, our empirical analyses provide significant evidence that none of *Wardlow*'s empirical assumptions are satisfied in practice. With respect to the first, our regression models suggest that officers often assess whether an area is high crime through a broad geographic lens. In many of our models, police precinct-level measures of crime (on average, four square miles) are substantially stronger predictors of whether an officer invokes HCA than measures of crime at a smaller level of geography, the census block group (.05 square miles). That's particularly true for violent- and property-crime stops. This suggests that officers frequently apply the high-crime area standard to large geographic areas such as police precincts.

Even more important, our results also provide little support for *Wardlow*'s second assumption. Officers invoke HCA in 57 percent of all stops—more often than any other basis of reasonable suspicion. And, while officers invoke HCA more often in certain parts of the city than others, they frequently do so everywhere. Indeed, in 98 percent of census block groups, officers invoked HCA in at least 30 percent of stops. In other words, officers are claiming that every block in New York City is high crime at one time or another. That claim seems implausible—particularly in the "safest big city in America."[14]

More to the point, officers' assessments of whether areas are high crime appear inaccurate. Despite our best efforts to predict HCA based on measures of crime at different levels of geography, temporal horizons, and crime types, our most predictive models produced an $R^2$ of just 0.01. In other words, actual crime

---

12.    The case that resulted in this settlement was *Daniels. v. New York*, 138 F. Supp. 2d 562 (S.D.N.Y. 2001).

13.    *See* People v. Rosario, 173 N.E.2d 881, 884 (N.Y. 1961) (holding that the defense is entitled to discovery of prior statements by a prosecution witness); JAMES E. MORRIS ET AL., VILLAGE, TOWN AND DISTRICT COURTS IN NEW YORK § 4:188 (2017) ("Police reports (insofar as they are written by an officer who will testify, or contain statements of witnesses), prior testimony, and notes relevant to a suppression issue are turned over as *Rosario* material."); PETER GERSTENZANG & ERIC H. SILLS, HANDLING THE DWI CASE IN NEW YORK § 28:3 (2017) ("Obviously, the written notes and reports of a police officer witness constitute *Rosario* material.").

14.    Pamela Engel, *Mayor Bloomberg: 'Stop and Frisk' Has Made New York City the Safest Big City in America*, BUS. INSIDER (Aug. 12, 2013), http://www.businessinsider.com/mayor-bloomberg-stop-and-frisk-has-made-new-york-city-the-safest-city-in-america-2013-8    [https://perma.cc/6DCW-XX9Z].

rates predicted only one percent of the variation in officers' assessments of whether areas are high crime.

If actual crime rates don't explain whether an officer invokes HCA as a basis of a stop, what does? One partial answer is that racial and socioeconomic biases may influence officers' determinations. When we analyze all stops together and control for local crime conditions, we find that a given officer in a given area is more likely to invoke HCA against young, Black, male suspects. When we break the data up by the type of suspected crime, we find that the higher invocation rates against Blacks is concentrated among stops for violent crime. We also find evidence that, in assessing whether an area is high crime, officers rely on neighborhood proxies, such as the racial and socioeconomic composition of residents. For example, when we analyze all stops together, across all of our models, moving a stop from an area with virtually no Black residents to an area with 100 percent Black residents is associated with a larger increase in the probability that an officer invokes HCA than moving from the single safest area in the city to the single most dangerous. This pattern appears to be concentrated in stops where the suspected crime is a violent, drug, or weapons offense.

Inter-officer disparities might also help explain when officers invoke HCA. Controlling for area of the city, roughly a quarter of officers invoke HCA in just 25 percent of stops, while another 40 percent do so over 75 percent of the time.

These results raise strong doubts as to whether the invocation of HCA has any predictive power about whether a suspect is engaged in crime. If not, the third empirical assumption of *Wardlow* does not hold. We examine this question by measuring the correlation between whether an officer invokes HCA as the basis of a stop and whether that stop results in a recorded "hit"—an arrest, the recovery of a weapon, or the recovery of other contraband. Our analysis here is necessarily limited because we can only observe the suspects that were stopped; we cannot observe suspects that officers chose not to stop (perhaps because they lacked reasonable suspicion). Still, our results are informative even if they are censored. For two of our three "hit" variables—arrest and recovery of a weapon—when we control for other observable bases of suspicion, we find that the probability of an arrest or the recovery of a weapon *decreases* when an officer invokes HCA to justify the stop. In other words, when an officer invokes HCA, the suspect is *less* likely to be engaged in a crime. This suggests that HCA may not be an indicator of guilt at all. It further suggests that officers may invoke HCA to manufacture the appearance of reasonable suspicion in their weakest stops. For our third hit variable—whether the officer recovered any contraband other than a weapon—we find that the probability of a hit remains the same when the officer invokes HCA.

Taken together, our findings provide empirical evidence that *Wardlow* may have been wrongly decided. Indeed, implementation of the high-crime area standard appears haphazard at best and discriminatory at worst. Officers call nearly every block in the city high crime at one time or another. Their

assessments of high-crime areas are only weakly correlated with actual crime rates. The suspect's race predicts whether an officer deems an area high crime as well as the actual crime rate itself. The racial composition of the area and the identity of the officer are stronger predictors of whether an officer deems an area high crime than the crime rate. And officers may even be using high-crime area as cover to bolster the appearance of constitutional validity in their weakest stops. These findings raise important questions about whether police officers can responsibly wield the discretion granted to them under *Wardlow*.

Of course, in this Article, we only evaluate the implementation of the high-crime area standard by one department during one time period.[15] Officers in other departments may be applying it with greater fidelity, and we cannot rule out this possibility with our data. But we ourselves are somewhat doubtful as the NYPD is one of the most organized, centralized, data-driven, and well-funded police departments in the country.

Short of reversing *Wardlow*, the courts have tools at their disposal to address some of the problems we have uncovered with the doctrine. Perhaps most simply, they could demand more rigorous data in suppression hearings to support an officer's claim that an area is high crime.[16] We suspect this solution would not go far enough, however, because it would only address the tiny fraction of stops that result in a criminal charge and motion to suppress.[17] Courts could go further by developing more precise definitions about the geographic scope, temporal horizon, and kinds of crimes relevant in assessing whether an area is high crime. A more aggressive judicial approach might prohibit a department from using high-crime areas to justify stops if there is evidence its officers are systematically misapplying the standard.[18] Police departments that do not faithfully implement the standard should not be able to use it to justify their stops.

We recognize that these proposals depart, at least to some extent, from how courts treat other factors under the reasonable suspicion analysis. Applying those other factors typically involves a highly discretionary, fact-bound inquiry based on the totality of circumstances and the common-sense judgments of the police

15. Our data are from 2007 to 2012, before a federal court ruled NYPD's stop and frisk program unconstitutional. See *infra* subsection I.A.

16. *See* Ferguson & Bernache, *supra* note 4, at 1593.

17. *See* Tracey L. Meares, *Programming Errors: Understanding the Constitutionality of Stop-and-Frisk as a Program, Not an Incident*, 82 U. CHI. L. REV. 159, 175 n.81 (2015) ("[O]nly a tiny fraction of stops are ever litigated because only a few result in an arrest, let alone a trial."). As others have noted, this is a general problem with the exclusionary rule as a remedy to constitutional violations. *See, e.g.*, Anthony Amsterdam, *Perspectives on the Fourth Amendment*, 58 MINN. L. REV. 349, 367–72 (1974).

18. This proposal is consistent with a recent call by Tracey Meares to review investigative stops, not as individual incidents, but instead as part of a larger organization-wide program. *See* Meares, *supra* note 17, at 174–76.

officer.[19] But, perhaps, the reason that the Fourth Amendment has taken this shape over time is that there were no other options. Historically, courts lacked access to the data needed to validate how police officers invoke Fourth Amendment factors in the field. Indeed, as the Supreme Court explained in *Wardlow* itself:

> In reviewing the propriety of an officer's conduct, courts do not have available empirical studies dealing with inferences drawn from suspicious behavior, and we cannot reasonably demand scientific certainty from judges or law enforcement officers where none exists. Thus, the determination of reasonable suspicion must be based on commonsense judgments and inferences about human behavior.[20]

As we try to show in this Article, that moment may be coming to an end.

Courts are not the only institutions that should reconsider how they handle the high-crime area standard. Police departments can promulgate regulations to guide officers.[21] Technological innovation can also help. The Philadelphia Police Department recently gave patrol officers smart phones with information on crimes occurring in the surrounding area.[22] Such devices could be used to inform officers in real time about objective crime data so that they do not need to rely on their own subjective and potentially unreliable intuitions about local crime rates. These devices could limit discretion even further by simply informing officers whether they are, at any given moment, in a high-crime area based on crime data and departmental policy.

In addition to these specific proposals, the implications of our analysis extend further—beyond the high-crime area standard—in at least two ways. First, our analysis offers a more general lesson about the response of police to different forms of judicial regulation. For example, it's perhaps unsurprising that, once courts recognized "furtive movement" as a cognizable factor in the reasonable suspicion analysis, police began to see furtive movements everywhere.[23] That concept is so vague, slippery, and contentless that any

---

19. *See, e.g.*, Navarette v. California, 572 U.S. 393, 397 (2014) ("[Reasonable suspicion] takes into account 'the totality of the circumstances—the whole picture.'"); Illinois v. Wardlow, 528 U.S. 119, 125 (2000) ("[T]he determination of reasonable suspicion must be based on commonsense judgments and inferences about human behavior.").

20. 528 U.S. at 124–25.

21. *See* Andrew Guthrie Ferguson, *Crime Mapping and the Fourth Amendment: Redrawing "High-Crime Areas,"* 63 HASTINGS L.J. 179, 219–24 (2011); Kelly K. Koss, Note, *Leveraging Predictive Policing Algorithms to Restore Fourth Amendment Protections in High-Crime Areas in a Post-Wardlow World*, 90 CHI.-KENT. L. REV. 301, 305 (2015); Hannah Rose Wisniewski, Note, *It's Time to Define High-Crime: Using Statistics in Court to Support an Officer's Subjective "High-Crime Area" Designation*, 38 NEW ENG. J. CRIM & CIV. CONFINEMENT 101, 120–22 (2012).

22. HUNCHLAB, HUNCHLAB: UNDER THE HOOD (2015), https://cdn.azavea.com/pdfs/hunchlab/HunchLab-Under-the-Hood.pdf [https://perma.cc/2DQH-NYB3].

23. *See* Jeffrey Fagan & Amanda Geller, *Following the Script: Narratives of Suspicion in* Terry *Stops in Street Policing*, 82 U. CHI. L. REV. 51, 69–81 (2015).

behavior might qualify.[24] But, in principle, the concept of a high-crime area could be operationalized in a manner that is more objective and verifiable. And yet, police officers appear able to misuse that more regulable standard too. The story of *Wardlow* thus teaches that, for the Fourth Amendment to impose a meaningful constraint on police discretion, the courts may need to develop more specific standards about what reasonable suspicion factors mean or, alternatively, to require that police do so through internal regulations.[25] Leaving the definition of those factors up to line officers on the street appears to be a dangerous proposition.

Second, our findings open the door to a largely uncharted area of empirical legal scholarship on the Fourth Amendment. Indeed, officers rely on countless factors other than high-crime areas in justifying the millions of stops they conduct each year. Officers may be applying some of those factors unfaithfully as well. Our analysis is therefore just the first step. We suggest that empirical legal scholars should begin validating other bases of reasonable suspicion on which officers regularly rely. Below, we identify several methodologies that could substantially advance this research agenda.[26]

The remainder of this paper proceeds as follows. In Part I, we briefly describe the historical development of the investigative stop and its current use in policing practice today. In Part II, we layer on the high-crime area standard, discussing *Wardlow* and how the lower courts have applied the doctrine. We also describe and justify the empirical assumptions of the high-crime area standard. Part III describes our data, and Part IV details the results of our empirical analysis. In Part V, we explore the implications of our findings for courts, police, and the Fourth Amendment more generally.

## I.
### INVESTIGATIVE STOPS

Police officers commonly invoke the high-crime area standard to justify investigative stops. In this section, we begin with a brief description of the

---

24.     The term "furtive movements" refers to a nearly infinite number of actions which an officer might find suspicious. The term often arises in cases in that an individual is suspected of carrying a firearm. Absent more bases of suspicion, furtive movements alone do not give rise to reasonable suspicion. *See, e.g.*, People v. Powell, 667 N.Y.S.2d 725, 727–28 (App. Div. 1998) (holding that officers did not have reasonable suspicion to frisk a suspect walking with his arm stiffly against his body in a high-crime area); People v. Fernandez, 928 N.Y.S.2d 293, 294–95 (N.Y. App. Div. 2011) (holding that an officer did not have reasonable suspicion to stop or frisk a suspect in a high-crime area whose hand was near his waist or in his sweatshirt pocket). Some courts have already expressed serious doubts about furtive movements. *See, e.g.*, Floyd v. New York, 959 F. Supp. 2d 668, 679 (S.D.N.Y. 2013) *("Floyd II")* ("'Furtive movements' are an insufficient basis for a stop or frisk if the officer cannot articulate anything more specific about the suspicious nature of the movement.").

25.     *See generally* John Rappaport, *Second-Order Regulation of Law Enforcement*, 103 CALIF. L. REV. 205, 213–20 (2015) (analyzing the benefits and advantages of judicial regulation aimed at line officers versus department administrators).

26.     *See infra* Part V.

historical evolution of the investigative stop and its role as a dominant policing strategy today. We then describe the basic legal framework regulating such stops and what we know about how they're typically conducted.

## A. A Brief History

Police have likely conducted investigative stops since the early days of American police departments in the mid-1800s.[27] Back then, policing strategy was largely reactive. Officers spent much of their time conducting random street patrols and responding to and investigating crimes reported by civilians.[28] Investigative stops were, therefore, primarily used to respond to and investigate crimes that had already occurred.[29]

The largely reactive character of the investigative stop began to change, however, in the early twentieth century. At the time, police experts and administrators were embracing more proactive approaches to policing, which focused not only on investigating past crimes but also preventing future ones.[30] By the 1960s, the investigative stop had already become a core crime prevention tool with regularized procedures.[31] Police officers would stop and interrogate a

---

27. For a few early cases, see Bishop v. Lucy, 50 S.W. 1029, 1029 (1899); Gisske v. Sanders, 98 P. 43, 44 (Cal. Ct. App. 1908).

28. *See* Christine N. Famega, *Proactive Policing by Post and Community Officers,* 55 CRIME & DELINQ. 78, 80 (2009).

29. For example, in one early case, after a police officer was informed that a burglary and robbery had occurred near a particular intersection in Los Angeles, he stopped and questioned a "suspicious character . . . in that vicinity." *Gisske*, 98 P. at 44. The language of the opinion implies that this tactic was relatively common at the time. Indeed, the officer's right to conduct the stop was so obvious the court saw no need to provide any citation or argument for support. *Id.* at 45 ("A police officer has a right to make inquiry in a proper manner of any one upon the public streets at a late hour as to his identity and the occasion of his presence, if the surroundings are such as to indicate to a reasonable man that the public safety demands such identification."). In another early case, two police officers in Austin were informed that a burglary had occurred at a brewing association. Soon after, they stopped and questioned a civilian two blocks away from the crime scene and arrested him for the burglary. Bishop, 50 S.W. at 1029. Once again, the opinion's language gives no indication this practice was unusual at the time.

30. *See, e.g.*, O. W. Wilson, *Police Arrest Privileges in a Free Society: A Plea for Modernization*, 51 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 395, 398 (1960) ("It is better to have an alert police force that prevents the crime than one that devotes its time to seeking to identify the assailant after the life has been taken, the daughter ravished, or the pedestrian slugged and robbed.").

31. *See* Sam B. Warner, *The Uniform Arrest Act*, 28 VA. L. REV. 315, 320 (1942) ("Every day large numbers of persons are questioned by police officers. This questioning, without immediate arrest, is essential to proper policing. . . . A man who looks round furtively, tries the door of an automobile, steps in and seems unfamiliar with its mechanism, may or may not have a right to drive the car. Under such circumstances, a passing officer ought to question the suspicious behavior."); Lawrence P. Tiffany, *Field Interrogation: Administrative, Judicial and Legislative Approaches*, 43 DENV. L.J. 389, 389 (1966) ("A common police practice, probably in all localities, is to stop and question suspects on the street when there are insufficient grounds to arrest."); *id.* at 395 ("A police officer may stop and question any person . . . whom he may have reason to suspect of unlawful design, and may demand of him his business and where he is going. . . . No law-abiding citizen will object to being questioned if it is done in a polite manner." (quoting RULES AND REGULATIONS OF THE POLICE DEP'T OF THE CITY OF PONTIAC § 230 (Jan. 1941)); Charles A. Reich, *Police Questioning of Law Abiding Citizens*, 75 YALE

"suspicious" individual on the street even if they lacked the probable cause required to conduct an arrest.[32]

Despite the expanding purpose of the investigative stop, its growth in the middle of the twentieth century was limited by at least two considerations. First, by the 1960s, the legality of the practice was unclear because the Supreme Court had not yet decided whether an investigative stop would qualify as a full-blown arrest and thus require probable cause, rather than some lower evidentiary showing.[33] Second, for much of the twentieth century, experts advocated for increased police professionalism and a reduction in police discretion.[34] As a result, police "managers began to focus all their training and resources on enforcing laws against *serious* crime," rather than on low-level, order-maintenance offenses—like drunkenness, panhandling, street prostitution, loitering, and rowdiness.[35] This new focus on serious crime likely reduced the frequency with which officers could conduct investigative stops.[36]

---

L.J. 1161, 1161 (1966) (reporting that the author had been stopped by the police nine or ten times in the last "few years" and noting that during the most recent stop the officer said "he had the right to stop anyone any place any time—and for no reason"); Meares, *supra* note 17, at 167 (describing Operation S, a police strategy in the 1950s, in which a designated unit in the San Francisco Police Department conducted thousands of stops annually).

32. *See* John A. Ronayne, *The Right to Investigate and New York's "Stop and Frisk" Law*, 33 FORDHAM L. REV. 211 (1964) (quoting from N.Y. Code Crim. Proc § 180-a, which authorizes the stopping and questioning of persons whom the police reasonably suspect are engaging in crime).

33. *See* Tiffany, *supra* note 31, at 395 (noting that many police officers in Chicago believed that field interrogations were arrests and were thus illegal absent probable cause); Loren G. Stern, *Stop and Frisk: An Historical Answer to a Modern Problem*, 58 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 532, 533 (1967) (explaining that police officers were "never quite sure whether a detention was constitutionally valid"). Some have argued that the common law provided police officers the power to conduct investigative stops based on a lower evidentiary standard than probable cause, but that view is disputed. *See* David Alan Sklansky, *The Fourth Amendment and Common Law*, 100 COLUM. L. REV. 1739, 1812 (2000).

34. *See* David Alan Sklansky, *Police and Democracy*, 103 MICH. L. REV. 1699, 1742–44 (2005) (discussing the turn to police professionalism in the 1950s and 1960s).

35. Debra Livingston, *Police Discretion and the Quality of Life in Public Places: Courts, Communities, and the New Policing*, 97 COLUM. L. REV. 551, 579 (1997) (emphasis added); *see also* DAVID B. WOLCOTT, COPS AND KIDS: POLICING JUVENILE DELINQUENCY IN URBAN AMERICA, 1890–1940, 146–47 (2005) ("Rather than focus on maintaining public order, as urban police departments had done in the late nineteenth and early twentieth centuries, police departments came to prioritize fighting crime in the 1920s and 1930s. Rather than see their primary functions as providing general services and arresting ne'er-do-wells for disorderly conduct, police came to define their purpose more narrowly as investigating and preventing crimes against persons and property, and apprehending and punishing criminals."). This trend was reinforced by a contemporary wave of judicial opinions that struck down many order-maintenance statutes as unconstitutionally vague. *See* Livingston, *supra*, at 598–600.

36. In 1960, police officers arrested 2.3 million people for drunkenness, disorderly conduct, vagrancy and other low-level crimes, which in total, accounted for roughly 52 percent of all non-traffic arrests in the country. *See* WESLEY G. SKOGAN, DISORDER AND DECLINE: CRIME AND THE SPIRAL OF DECAY IN AMERICAN NEIGHBORHOODS 89 (1990). In 1985, however, officers conducted only 1.4 million such arrests, which accounted for just 16 percent of all non-traffic arrests throughout the country. *Id.* This intense redirection of police resources—from low-level order-maintenance crimes to more serious offenses—restricted the reach of the investigative stop as a policing program.

Both this legal and philosophical limit on investigative stops eventually disappeared. In 1968, the U.S. Supreme Court resolved any lingering constitutional doubts in *Terry v. Ohio*, by holding that investigative stops do not violate the Fourth Amendment.[37] Soon enough, the philosophical aversion against order-maintenance policing also fell away. In 1982, George Kelling and James Q. Wilson published an article in *The Atlantic* called "Broken Windows."[38] The authors argued that "minor disorder in a neighborhood, if left unchecked, . . . will result in increased serious crime, and, therefore, that eliminating minor disorder . . . will have a deterrent effect on major crime."[39] According to Wilson and Kelling, focusing police resources exclusively on "serious" crime was a mistake. Instead, they argued that officers should prioritize low-level, order-maintenance enforcement and that departments should deploy more officers on foot to carry out those responsibilities.

It's hard to overstate the influence of "Broken Windows" on the use of the investigative stop.[40] Almost immediately, departments across the country began putting order-maintenance policing into action.[41] This trend expanded officers' discretion to stop a larger universe of people who were engaged in low-level offenses.[42] Indeed, police departments boasted that enforcing these low-level offenses enabled them to stop more people and thus remove more weapons and drugs from the street.[43]

In the 1990s, the NYPD—the research site of the current study—developed an approach to order-maintenance policing[44] that would once more transform the

---

37.     *See* 392 U.S. 1, 26 (1968).

38.     George L. Kelling & James Q. Wilson, *Broken Windows: The Police and Neighborhood Safety*, THE ATLANTIC, Mar. 1982.

39.     Bernard E. Harcourt, *Reflecting on the Subject: A Critique of the Social Influence Conception of Deterrence, The Broken Windows Theory, and Order-Maintenance Policing New York Style*, 97 MICH. L. REV. 291, 302 (1998).

40.     One media outlet called Broken Windows the "bible of policing;" another called order-maintenance policing the "Holy Grail of the '90s." Kevin Cullen, *The Commish*, BOSTON GLOBE, May 25, 1997, at 3; Robert A. Jones, *The Puzzle Waiting for the New Chief,* L.A. TIMES, Aug. 10, 1997, at l. Another outlet called it a "revolution in American policing." Christina Nifong, *One Man's Theory is Cutting Crime in Urban Streets*, CHRISTIAN SCI. MONITOR (Feb. 18, 1997), https://www.csmonitor.com/1997/0218/021897.us.us.4.html [https://perma.cc/PNB2-H9G9]; *see also* Harcourt, *supra* note 39, at 293–94.

41.     Livingston, *supra* note 35, at 583.

42.     *Id.* at 590.

43.     *See, e.g.*, Ruben Castaneda, *As D.C. Police Struggle on, Change Pays off in New York*, WASH. POST, Mar. 30, 1996, at Al ("In New York, laws against so-called quality-of-life violations—graffiti, aggressive panhandling, drinking in public—are enforced not only for their own sake but also because they give officers a reason to check for drugs, weapons and outstanding warrants. That has had a ripple effect, [NYPD Deputy Commissioner Jack] Maple said. 'People don't carry their guns anymore, because they know they might get stopped.'").

44.     The department issued "Police Strategy Number 5," which was entitled "Reclaiming the Public Spaces of New York." NEW YORK POLICE DEPARTMENT, POLICE STRATEGY NO. 5: RECLAIMING THE PUBLIC SPACES OF NEW YORK 5 (1994), http://marijuana-arrests.com/docs/Bratton-blueprint-1994--Reclaiming-the-public-spaces-of-NY.pdf [https://perma.cc/9CNG-YVDG]. The manual explained that order-maintenance policing would "emerge as the linchpin of efforts . . . by

investigative stop, this time, through technological innovation. Under the leadership of Commissioner William Bratton, the department created COMPSTAT, a "strategic management process that use[d] computer technology, operational strategy and managerial accountability" to increase the department's capacity to reduce crime.[45] One core feature of the program was a cutting-edge information system that could map and analyze up-to-date crime and disorder statistics on small geographic areas of the city.[46] To promote accountability, high-ranking department personnel would attend weekly meetings in which commanders presented crime data from their precincts and would explain the steps they were taking to reduce crime.[47] The investigative stop—also commonly referred to as a stop, question, and frisk—was one of the main ways commanders could show they were actively working to drive down the crime rate.[48] COMPSTAT's focus on geographic crime measures helped the department direct its resources to areas of the city with the most crime.

The NYPD's appetite for investigative stops after the adoption of COMPSTAT was nothing short of enormous. In 1998, just four years after the program began, the NYPD conducted roughly 140,000 recorded stops.[49] By 2011 that number had grown to nearly 700,000 stops per year.[50] Other cities that adopted COMPSTAT have also experienced an expansion of the practice.[51]

In the last few years, the use of investigative stops has faced legal resistance. In 2013, well after the data in our study end, Federal District Court Judge Shira Scheindlin ruled that the NYPD's stop and frisk program was systematically violating the Fourth and Fourteenth Amendments.[52] The court ordered the department to reform its program under the oversight of a court-appointed monitor.[53] Since then, the number of recorded stops in the city has

---

the . . . Department to reduce crime and fear in the city." *Id.* It further explained that by "working systematically and assertively to reduce the level of disorder in the city, the NYPD w[ould] act to undercut the ground on which more serious crimes seem possible." *Id.*

45.    William F. Walsh, *Compstat: An Analysis of an Emerging Police Managerial Paradigm*, 24 POLICING: INT'L J. POLICE STRATEGIES & MGMT. 347, 347 (2001).

46.    *Id.* at 352.

47.    *Id.*

48.    *See* ELIOT SPITZER, ATTORNEY GENERAL OF THE STATE OF NEW YORK, THE NEW YORK CITY POLICE DEPARTMENT'S 'STOP & FRISK' PRACTICES: A REPORT TO THE PEOPLE OF THE STATE OF NEW YORK FROM THE OFFICE OF THE ATTORNEY GENERAL 56–57, 59 n.48. (1999).

49.    *See id.* at 91 (citing 175,000 stops over a fifteen-month period from January 1998 through the end of March 1999.)

50.    Floyd v. New York, 959 F. Supp. 2d 540, 588 (S.D.N.Y. 2013).

51.    *See, e.g.*, Jeremy Gorner, *ACLU, Chicago Agree to Changes on Controversial Street Stops*, CHI. TRIB., Aug. 7, 2015 (noting that COMPSTAT pressured district commanders in Chicago to increase the use of stop and frisk).

52.    *Id.* at 560–63.

53.    *Id.* at 667. The *Floyd* Opinion and Order specified remedial actions that applied also to two companion stop-and-frisk cases: *Davis et al. v. City of New York*, 959 F.Supp.2d 427 (2013), and *Ligon et al. v. City of New York*, 743 F.3d 362 (2014). *See Mayor de Blasio Announces Agreement in Landmark Stop-And-Frisk Case*, NEW YORK CITY PRESS OFFICE (Jan. 30, 2014), http://www1.nyc.gov/office-of-

fallen dramatically.[54] Stop and frisk programs at a few other departments have faced public resistance as well.[55] Nonetheless, the investigative stop continues to be one of the most important tools of police departments across the country to reduce crime.

### B.   Investigative Stops Today

The legal framework governing investigative stops today is well known. The Fourth Amendment of the United States Constitution is the primary source of regulation. It requires that, before conducting a stop, an officer have "reasonable suspicion" that the suspect is committing or is about to commit a crime.[56] A "mere hunch" is insufficient;[57] officers must base their suspicions on specific and articulable facts. These facts typically relate to the suspect's behavior, clothing, and location.[58] They might include, for example, casing a store, acting as a lookout, engaging in a drug transaction, concealing potential contraband, or running away from an officer. Officers might also rely on softer signals of criminal behavior, like "furtive movements" and "bulges" in clothes or a waistband.

Less is known about how investigative stops are actually carried out in practice. Much of what we do know comes from two cities—New York and Chicago—where there has been aggressive litigation regarding stop and frisk in recent years. The data from those two jurisdictions show that investigative stops fall disproportionately on people of color. In Chicago, 71 percent of recorded stops in the first half of 2016 were against Black civilians, who accounted for a

---

the-mayor/news/726-14/mayor-de-blasio-agreement-landmark-stop-and-frisk-case#/0 [https://perma.cc/3Z6N-RT5V].

54.    NEW YORK CITY ACLU, NYC: STOP-AND-FRISK DOWN, SAFETY UP (2015), https://www.nyclu.org/sites/default/files/publications/stopfrisk_briefer_FINAL_20151210.pdf [https://perma.cc/V2RN-GMUS]; Monitor's Fifth Report: Analysis of NYPD Stops Reported, 2013–2015, at 2, Floyd et al. v. City of New York, No. 1:08-cv-01034-AT-HBP (S.D.N.Y. May 30, 2017).

55.    The rate of recorded stops conducted by the Chicago Police Department, for example, has also fallen since the department entered into a consent agree with the ACLU in 2015. Chuck Goudie, *CPD "Stop and Frisks" Down 80 Percent in 2016*, ABC CHICAGO (Feb. 1, 2016), http://abc7chicago.com/news/cpd-stop-and-frisks-down-80-percent-in-2016/1182604/ [https://perma.cc/LKB5-ECD2] (noting an 80% drop in recorded stops). The precise cause of the drop is unclear. The rate of recorded stops in Philadelphia also is declining in the wake of a consent decree in a civil rights case. Palmer, *supra* note 1 (reporting a 35 percemt decline in pedestrian stops in 2016 by the Philadelphia Police Department). *See also*, Plaintiffs' Eighth Report to Court and Monitor on Stop and Frisk Practices: Fourth Amendment Practices, Bailey et al. v City of Philadelphia, No. 10-5952, at 23 (Dec. 7, 2017), https://www.aclupa.org/download_file/view_inline/3273/198 [https://perma.cc/H66N-AJP5] (concluding that "there are still too many stops and far too many stops without reasonable suspicion").

56.    Terry v. Ohio, 392 U.S. 1, 37 (1968).

57.    Navarette v. California, 572 U.S. 393, 397 (2014) (internal quotation marks omitted).

58.    CITY OF NEW YORK POLICE DEPARTMENT, INVESTIGATIVE ENCOUNTERS REFERENCE GUIDE 14–15 (2015).

third of the city's population.[59] In New York City from 2004 to 2012, our data show that 52 percent of stops were against Black civilians, who accounted for about 23 percent of the population in 2010.[60]

In many investigative stops, the officer frisks the suspect. That's true in about half of stops in New York City[61] and roughly a third in Chicago.[62] A frisk can be highly intrusive. Seth Stoughton, a former police officer, described his technique as follows:

> I would slide my hand . . . over the area of [the] suspect's body that I was searching, moving them in small circles as I did so, so that my fingertips and palms might detect any protuberance in or under the suspect's clothing. At the same time, I would lightly clench and release my fingers . . . to shift clothing over the skin so I could ensure that I could identify items and not mistake a weapon for a seam or fold in the clothing. . . .[63]

Stoughton describes applying that technique to each area of the suspect's body, including the midsection, waistband, groin, buttocks, upper thigh, head, neckline, armpits, and chest.[64]

In a substantial number of stops, officers also use force against the suspect. Our data show that, from 2004 to 2012, NYPD officers recorded using their hands—which includes slapping and grabbing—in 19 percent of all stops. A 2015 survey of Chicago residents found that 13 percent of respondents who were stopped in the previous twelve months were pushed or shoved.[65] Officers

---

59.     RALPH B. TAYLOR & LALLEN T. JOHNSON, ANALYSIS OF CHICAGO POLICE DEPARTMENT POST-STOP OUTCOMES DURING INVESTIGATORY STOPS JANUARY THROUGH JUNE 2016: INPUT TO HON. ARLANDER KEYS' (RET.) FIRST YEAR REPORT 32 (2017), https://www.aclu-il.org/sites/default/files/appendix-b-analysis-of-cpd-post-stop-outcomes-during-investigatory-stops.pdf [https://perma.cc/9VSU-DC2B].

60.     U.S. CENSUS BUREAU, DEMOGRAPHIC PROFILE—NEW YORK CITY AND BOROUGHS 2000 AND                2010,                at                1, https://www1.nyc.gov/assets/planning/download/pdf/data-maps/nyc-population/census2010/t_sf1_dp_nyc_demo.pdf [https://perma.cc/6U37-5J8T]. From 2004 to 2012, 31 percent of stops were against Hispanic persons, who accounted for about 29 percent of the population in 2010. *Id.* But, after controlling for local crime and social conditions, stops of Hispanics were significantly more common than stops for whites. *See* Floyd v. City of New York, 959 F. Supp. 2d 540, 545 (S.D.N.Y. 2013) (citing statistical evidence that "Blacks and Hispanics are more likely than whites to be stopped within precincts and census tracts, even after controlling for other relevant variables . . . even in areas with low crime rates, racially heterogenous populations, or predominately white populations")

61.     Delores Jones-Brown & Brett G. Stoudt, *Stop, Question, and Frisk Policing Practices in New York City: A Primer (Revised)*, JOHN JAY C. CRIM. JUST. 18 (2013), www.atlanticphilanthropies.org/app/uploads/2015/09/SQF_Primer_July_2013.pdf [https://perma.cc/U2P8-P758].

62.     TAYLOR & JOHNSON, *supra* note 58, at 34.

63.     Seth W. Stoughton, Terry v. Ohio *and the (Un)Forgettable Frisk*, 15 OHIO ST. J. CRIM. L. 19, 28 (2017).

64.     *Id.* at 29.

65.     WESLEY G. SKOGAN, STOP-AND-FRISK AND TRUST IN POLICE IN CHICAGO 9 (2016), https://www.ipr.northwestern.edu/publications/docs/workingpapers/2016/WP-16-08.pdf [https://perma.cc/Y5VZ-YDRH].

sometimes use more severe forms of force. Our data show that NYPD officers reported drawing and pointing a weapon in less than 1 percent of stops. Survey data, however, suggests that officers may use weapons more frequently: respondents in the Chicago survey reported that officers drew a weapon in roughly 10 percent of stops and threatened to take out a weapon in another 6 percent.[66]

There is some evidence that the use of force during investigative stops falls disproportionately on people of color. A recent study by Roland Fryer, for example, found "large racial differences" for non-lethal uses of force.[67] However, he found no evidence of racial disparities for lethal force.[68]

The vast majority of stops do not result in any further enforcement action. In New York City, for example, only 6 percent of all stops from 2004 to 2012 resulted in an arrest.[69] A similar proportion resulted in a summons.[70] Thus, 88 percent of stops resulted in no enforcement action, underscoring how rarely officers detect contraband.[71] Indeed, officers recovered a weapon in only 1.5 percent of stops and recovered other forms of contraband in less than 2 percent of stops.[72]

These statistics suggest that the investigative stop constitutes an intrusive experience that is imposed largely on innocent civilians, particularly people of color. The Fourth Amendment is designed to balance the competing values of privacy and law enforcement by requiring officers to identify specific facts supporting reasonable suspicion. In the next section, we consider one of those facts in particular—high-crime areas.

## II.

### HIGH-CRIME AREAS

The Supreme Court held in *Wardlow* that police officers can consider whether an area is "high crime" in determining whether they have reasonable suspicion to conduct an investigative stop.[73] To be sure, this was not the first time courts had endorsed the use of crime rates in analyzing the constitutionality

---

66.    *Id.*

67.    ROLAND G. FRYER, JR., AN EMPIRICAL ANALYSIS OF RACIAL DIFFERENCES IN POLICE USE OF FORCE 3 (2016), https://law.yale.edu/system/files/area/workshop/leo/leo16_fryer.pdf [https://perma.cc/QEA8-KFYK].

68.    *Id.* at 5. *But see* Justin Feldman, *Roland Fryer is Wrong: There is Racial Bias in Shootings by Police* (July 12, 2016), https://scholar.harvard.edu/jfeldman/blog/roland-fryer-wrong-there-racial-bias-shootings-police [https://perma.cc/S92Z-N7M9t].

69.    Floyd v. New York, 959 F. Supp. 2d 540, 558 (S.D.N.Y. 2013).

70.    *Id.* A substantial fraction of all summons are for low-level violations, rather than misdemeanor or felony crimes. *Id.* at 575 n.126.

71.    *See id.* at 558–59.

72.    *Id.* at 559.

73.    *See* 528 U.S. 119 (2000).

of stops.[74] But the prior Supreme Court cases had primarily concerned illegal immigration near the border with Mexico, not typical street crime.[75] Moreover, those cases all arose in the 1970s, well before the meaning of an "investigative stop" had been transformed by the NYPD in the 1990s into the aggressive stop and frisk program that we see across the country today.[76]

As we discuss in greater detail below, *Wardlow* left many open questions about the meaning of a high-crime area and the kinds of evidence courts should consider in applying that standard in individual cases. For the most part, the lower courts have not confronted these doctrinal questions directly. Instead, the most common approach has been to defer to the expertise of police officers without engaging in meaningful scrutiny.

## A. *Illinois v. Wardlow*

In *Wardlow*, two uniformed officers working in the special operations division of the Chicago Police Department were driving towards an area of the city "known for heavy narcotics trafficking" to investigate narcotics trafficking.[77] They were traveling in a caravan with three other police cars because "they expected to find a crowd of people in the area, including lookouts and customers."[78] Sam Wardlow saw the officers' vehicle and fled in the opposite direction, clutching an opaque bag. The officers stopped and searched him and found a loaded handgun in the bag.

The Supreme Court held that the officers had reasonable suspicion to stop Wardlow. The Court acknowledged that an "individual's presence in an area of expected criminal activity, standing alone, is not enough" to establish reasonable suspicion.[79] But it also noted that "officers are not required to ignore the relevant characteristics of a location in determining whether the circumstances are sufficiently suspicious to warrant further investigation."[80] The Court thus

---

74. The New York state case that defined the standards for stops also included language suggesting that the reputation of the area as a known crime location also heightened suspicion. *See* People v. DeBour, 40 N.Y.2d 210 (1976).

75. In one other case, the Court held that an officer had the legal authority to seize a weapon in plain sight on a suspect's waistband, at least in part, due to the high levels of crime in the neighborhood. *See* Adams v. Williams, 407 U.S. 143, 147 (1972). But the Court made clear that this consideration was relevant to whether the officer, who was conducting a legal stop, had the authority to search the defendant in order to protect the officer's safety. *Id.* The other cases where the Court endorsed the use of crime levels in assessing reasonable suspicion were immigration cases. *See* United States v. Brignoni-Ponce, 422 U.S. 873, 878–84 (1975); United States v. Cortez, 449 U.S. 411, 419 (1981). In one other case, the Court limited the significance of the high-crime area factor by holding that, on its own, the factor could not establish reasonable suspicion. *See* Brown v. Texas, 443 U.S. 47, 52 (1979).

76. *See* Rachel A. Harmon & Andrew Manns, *Proactive Policing and the Legacy of* Terry, 15 OHIO ST. J. CRIM. L. 49, 55–58 (2017) (describing how *Terry* and other cases in the 1960s and 1970s predated the transformation in proactive policing without anticipating it).

77. 528 U.S. 119, 121 (2000).

78. *Id.*

79. *Id.* at 124.

80. *Id.*

concluded that the high crime level in the area, coupled with the suspect's flight, was sufficient to establish reasonable suspicion.[81] Rather than providing an empirical or analytical framework to evaluate the accuracy of officers' claims regarding high-crime areas, the Court deferred to the police and their knowledge of the area's crime rate.[82]

It is difficult to assess the precise size of *Wardlow*'s impact on subsequent Fourth Amendment case law. The courts had already considered high-crime areas in prior Fourth Amendment decisions. Moreover, it's hard to tell in any particular case whether the suspect's presence in a high-crime area was outcome determinative. Still, *Wardlow* likely had a substantial impact on focusing courts' attention on the high-crime area factor: over 4,500 federal and state decisions have cited the opinion since it was issued in 2000.[83]

### B.   *Lower-Court Interpretation of Wardlow*

*Wardlow* provided remarkably little guidance about how to apply the high-crime area standard in individual cases. It left open two kinds of questions in particular. The first are questions about the substantive definition of a high-crime area. The second concern questions about the kinds of evidence that courts should rely on in assessing whether an area is high crime. We discuss how the lower courts have approached each of these questions in turn.

### 1.   *Substantive Legal Standard*

*Wardlow* provided no substantive definition of a "high-crime area." The Court thus left unanswered at least three key questions about the meaning of the concept, and about how it should be interpreted and applied in practice.[84] First, the Court did not define the geographic scope of a high-crime area. Can it be as large as a neighborhood? Or must it be smaller, like an intersection or street block, or even a single residential or commercial building? Second, the Court did not explain how far back in time officers should look for evidence of crime. Are crimes that took place over a month ago relevant? What about six or twelve months ago? Third, the Court did not explain what kinds of crimes are relevant. Should we consider all crimes? Only serious ones? Only violent ones?

Most lower courts have entirely ignored the first question, which concerns geographic scope. But the few that have grappled with it have demanded a relatively narrow geographic area. In *United States v. Montero-Camargo*, for example, the Ninth Circuit cautioned district courts to "be particularly careful to ensure that a 'high crime' area factor is not used with respect to entire

---

81.   *Id.* at 124–25.

82.   *Id.*

83.   We obtained this number from Westlaw's citing references. For an updated number, see WESTLAW, https://1.next.westlaw.com (search for "528 U.S. 119"; then follow "Citing References" hyperlink).

84.   Ferguson & Bernache, *supra* note 4, at 1605–07.

neighborhoods . . . but is limited to specific, circumscribed locations."[85] The Sixth Circuit has similarly held that a stop was conducted with reasonable suspicion in part because the relevant high-crime area was "circumscribed to a specific intersection rather than an entire neighborhood."[86] And the First Circuit has emphasized that, in assessing whether an area is high crime, a court should consider the "limited geographic boundaries of the 'area' or 'neighborhood' being evaluated."[87]

The courts have reached little consensus, however, regarding the second and third questions about the meaning of a "high-crime area." With respect to the second question—temporal horizon—courts tend to ignore how far back in time the relevant crimes were that form the basis of an officer's judgment that an area is high crime. A few courts have demanded "temporal proximity" between the evidence of past crime and the "date of the stop" at issue.[88] But that demand, on its own, provides little guidance. And other courts appear comfortable relying on relatively old data.[89] Thus, courts have offered little clarity about how far back in time officers should look in assessing whether an area is high crime.

With respect to the third question—which concerns the kinds of crimes that are relevant—a few courts have suggested that greater weight should be given to crimes that are similar to the suspected crime that initially justified the stop. The First Circuit, for example, gives greater weight to data that establishes a "nexus between the type of crime most prevalent or common in the area and the type of crime suspected in the instant case."[90] Other courts, however, appear to reject any nexus requirement. In *United States v. Cooper*, for example, a police officer stopped an individual on the street suspected of car theft. The Sixth Circuit concluded that the area where the stop took place was high crime even though there was no evidence that "car thefts occurred in [the] area with 'unusual

---

85. United States v. Montero-Camargo, 208 F.3d 1122, 1138–39 (9th Cir. 2000). Some police departments agree. For example, at some point after our data in this study end, the NYPD amended its investigation guide to state: "[A] 'high crime area' cannot be defined too broadly, such as encompassing an entire precinct or borough." NEW YORK POLICE DEPARTMENT, INVESTIGATIVE GUIDE 15 (2015).

86. United States v. Caruthers, 458 F.3d 459, 468 (6th Cir. 2006).

87. United States v. Wright, 485 F.3d 45, 54 (1st Cir. 2007); *see also* United States v. Hill, 752 F.3d 1029, 1035 (5th Cir. 2014) ("This vague testimony about the 'overall' rise in crime in the 'fairly large county' tells us almost nothing about whether the police had reasonable suspicion to seize Hill at one single apartment complex, in one single town within the county.").

88. *See, e.g.*, *Wright*, 485 F.3d at 53–54.

89. *Montero-Camargo*, 208 F.3d at 1143 (Kozinski, J., concurring) ("One agent said he'd been involved in 15–20 stops over eight and a half years, and '[could]n't recall any . . . where we didn't have a violation of some sort.' . . . The other agent testified to 'about a dozen' stops in the same period, all but one of which led to an arrest. . . . Without hesitation, the majority treats this as a crime wave, but is it really? Does an arrest every four months or so make for a high crime area?").

90. *Wright*, 485 F.3d at 53–54. A few other courts have taken care to note that the area in which a stop took place has a high rate of the suspected crime. *See, e.g.*, *Caruthers*, 458 F.3d at 468 ("Furthermore, the crimes that frequently occur in the area are specific and related to the reason for which Caruthers was stopped.").

regularity.'"[91] Consistent with *Cooper*, courts frequently conclude that an area is high crime without any information about the frequency of the specific crime suspected by the officer.[92]

### 2. *Evidence of High-Crime Areas*

In addition to leaving open the substantive definition of a high-crime area, *Wardlow* also provided no guidance on the relevant evidentiary standards. In response, the courts have been remarkably lax, accepting a variety of different kinds of evidence to support the government's claim that an area satisfies the high-crime standard.[93] The most common approach is to accept an officer's bare testimony that an area is "high crime."[94] Indeed, courts frequently defer to such testimony without any additional information.[95] In some cases, officers may provide slightly more information about the basis of their subjective judgment. They might, for example, explain that their judgment is based on the number of arrests they have conducted in the area or on the high volume of civilian complaints received by the department.[96] When doing so, officers rarely specify

---

91.     United States v. Cooper, 431 F. App'x 399, 403 (6th Cir. 2011) (Cole, J., dissenting) (emphasis omitted).

92.     *See infra* note 95.

93.     For a more extensive review of the case law, see Ferguson & Bernache, *supra* note 4, at 1607–18.

94.     Ferguson & Bernache, *supra* note 4, at 1607 ("[T]he majority of jurisdictions . . . primarily have relied on an officer's testimony that an area is a 'high-crime area' without much analysis as to the basis of that conclusion."); *see also* Herbert, *supra* note 7, at 135 ("As an eager young Assistant United States Attorney who 'papered' countless complaints, conducted numerous hearings and tried a substantial number of cases, I learned how to decode police officer jargon and law enforcement terminology. One of the most commonly used—yet seldom defined—phrases was 'high crime area.' . . . In court . . . judges rarely challenged the label or required its definition. Judges never asked officers for data to support assertions that an area was high-crime.").

95.     Ferguson & Bernache, *supra* note 4, at 1607; Herbert, *supra* note 7, at 135; *see, e.g.*, State v. Morgan, 539 N.W.2d 887, 892 (Wis. 1995) ("[W]e find that an officer's perception of an area as 'high-crime' can be a factor justifying a search."); Riley v. Commonwealth, 412 S.E.2d 724, 726 (Va. Ct. App. 1992) (explaining that the officer testified that the stop took place in a "high crime area"); United States v. Reed, 402 F. App'x 413, 416 (11th Cir. 2010) (same); United States v. Wiley, 117 F. App'x 906, 908 (5th Cir. 2004) ("The officers testified that the restaurant was located in a high-crime area."); State v. Moyer, No. 09AP–434, 2009 WL 4936383, at *1(Ohio Ct. App. Dec. 22, 2009) ("At the hearing, the state presented Officer Harmon, who testified his encounter with defendant occurred in what is a 'known high-crime area' where narcotics and weapons arrests are common."); United States v. Bryant, No. 1:16CR060, 2017 WL 1086081, at *1 (S.D. Ohio Mar. 17, 2017) ("The location of the market itself was identified by CPD and was testified to by Officer Rogers as a high-crime area."); United States v. Singleton, No. CRIM. 3:07CR282, 2008 WL 2323487, at *2 (W.D.N.C. May 29, 2008) ("In the present case, the officers testified that they observed the Defendant walking in a high-crime area with a handgun in a holster at his side."); State v. Sanders, No. 01-0927, 2002 WL 1757659, at *2 (Iowa Ct. App. July 31, 2002) ("[T]he officer testified the two were standing behind an abandoned house in a high-crime area in Des Moines notorious for drug dealing, prostitution, loitering, and vandalism.").

96.     *See, e.g.*, United States v. Bridges, No. 14-20007, 2014 WL 1365673, at *3 (E.D. Mich. Apr. 7, 2014) ("Defendant was in a specific intersection that is a known high-crime area, where Corporal Neese has made previous arrests for firearms offenses, drug crimes, and other offenses."); Lee v. State, 868 So. 2d 577, 578 (Fla. Dist. Ct. App. 2004) ("[The officer] had made fifteen to twenty drug arrests at [the site of the stop]."); State v. Collins, 890 So. 2d 616, 619 (La. Ct. App. 2004) (noting that the police

the precise geographic area or time period from which their experience is drawn.[97]

A few courts have demanded more rigorous and objective evidence before concluding that an area is high crime.[98] In *Montero-Camargo*, for example, the Ninth Circuit instructed the district courts that calling an area high crime "requires careful examination."[99] It explained that courts "must carefully examine the testimony of police officers . . . and make a fair and forthright evaluation of the evidence they offer."[100] The court further noted that "more than mere war stories are required to establish the existence of a high-crime area . . . [and] courts should examine with care the *specific data* underlying any such assertion."[101]

Despite calling for more careful scrutiny of officers' testimony, not even the Ninth Circuit has always followed its own instructions. As the concurrence in *Montero-Camargo* pointed out, the majority opinion itself failed to scrutinize

---

department "received frequent complaints about crime in the vicinity of [the stop]."); United States v. Coates, 457 F. Supp. 2d 563, 568 (W.D. Pa. 2006) ("Detective Redpath testified that Lawrenceville . . . is a high crime area where he and his unit have made numerous drug arrests."); State v. Serna, 307 P.3d 82, 83 (Ariz. Ct. App. 2013), ("Officers described the area as 'high crime,' a 'gang neighborhood' where 'violence takes place,' and having 'numerous drug complaints.'"); United States v. Amaker, No. CRIM 2:05-00149-001, 2005 WL 3409570, at *1 (S.D.W. Va. Dec. 12, 2005) ("Patrolman Workman testified that the subject area was a high crime area and that, in his seven plus years as an officer, he had made numerous arrests in the area."); Commonwealth v. Robinson, No. 2618 EDA 2014, 2015 WL 6112184, at *4 (Pa. Super. Ct. Aug. 31, 2015) ("Officer Walsh further stated that he had personal experience with the high rate of crime in the arrest area, having previously made 'between 25 and 50 arrests' in the immediate area."); Woody v. State, 765 A.2d 1257, 1261 (Del. 2001) ("Officer Jordan testified the area was a high crime area and that the police had had numerous complaints of drug dealing and other criminal activity.").

97.    *Id.*

98.    *See, e.g.*, United States v. Wright, 485 F.3d 45, 54 (1st Cir. 2007) ("Evidence on these issues could include a mix of objective data and the testimony of police officers, describing their experiences in the area.").

99.    United States v. Montero-Camargo, 208 F.3d 1122, 1138 (9th Cir. 2000); *see also* N. Mariana Islands v. Crisostomo, No. 2013-SCC-0008-CRM, 2014 WL 7072149, at *3 (N. Mar. I. Dec. 12, 2014) ("[W]e conclude that an officer's sense of an area's criminality by itself is not enough to support a high-crime-area finding. Instead, the Commonwealth must provide objective, verifiable data showing by a preponderance of the evidence that at the time of the arrest, the disputed location had a higher crime rate than other relevant areas in a constitutionally significant manner."); United States v. Arvizu, 232 F.3d 1241, 1250 (9th Cir. 2000), *rev'd*, 534 U.S. 266 (2002)) (rejecting the district court's finding that an area was high crime because the only supporting evidence in the record was an officer's testimony that the "400 block was 'one of the most notorious areas'" for smuggling drugs and undocumented immigrants); People v. Harris, 957 N.E.2d 930, 936 (Ill. App. 1. Dist. 2011) ("A conclusory and unsubstantiated statement that a location is a 'high crime area' is insufficient to establish that consideration for purposes of a *Terry* stop.").

100.    *Montero-Camargo*, 208 F.3d at 1138.

101.    *Id.* at 1139 n.32 (emphasis added); *see also* United States v. Thornton, No. 5:13CR522, 2014 WL 11173589, at *7 n.1 (N.D. Ohio Apr. 17, 2014) ("At the evidentiary hearing, the government provided the Court with . . . a printout of reported crimes within a one mile radius around the intersection of Copley Road and South Hawkins.").

the underlying data.[102] Other subsequent Ninth Circuit panels have done the same. One panel accepted that an area was high crime based on one officer's unsupported testimony that "in [his] experience as a patrol officer[,] . . . Harbor Boulevard in the vicinity of McFadden Boulevard is an area of heavy narcotics trafficking and other criminal activity."[103] Another panel was criticized by a dissenting judge for accepting the lower court's judgment that an area was high crime even though that judgment was based exclusively on the testimony of an officer that a particular road was "located in a high-crime area."[104] Thus, even courts that demand more rigorous scrutiny of officers' testimony often fail to conduct that scrutiny themselves.

* * *

Given the state of the case law, it's hard to draw clear generalizations about *Wardlow* and the concept of a high-crime area. The most common approach appears to be for courts to defer to the largely unsupported testimony of police officers that a particular area is high crime. As a result, courts rarely grapple with the three key questions left open by *Wardlow* concerning the geographic scope, temporal horizon, and categories of crimes most relevant to the analysis. But when the courts do scrutinize officers' testimony carefully, they appear to analyze high-crime areas through a relatively narrow geographic lens. A high-crime area appears, therefore, to be more like a street block or intersection than a neighborhood or county. It's less clear what temporal horizon or categories of crime should be considered.

## C.  *Wardlow's Unspoken Empirical Assumptions*

Based on our reading of the case law, we believe the wisdom of *Wardlow* as a constitutional doctrine depends heavily on at least three unspoken empirical assumptions about how police officers apply the high-crime area standard in practice.

The first concerns the geographic scope of a high-crime area. As we noted above, the few lower courts that have grappled with this question generally agree that officers must analyze whether an area is high crime through a granular lens:

---

102.    *Montero-Camargo*, 208 F.3d at 1143 (Kozinski, J., concurring) ("The opinion recognizes the danger in allowing the police to characterize an area as 'high-crime' to establish a basis for reasonable suspicion, but then proceeds to do just that, based on nothing more than the personal experiences of two arresting agents. As I discuss above, the agents didn't even claim this was a high crime area . . . . To [my colleagues], it's a high crime area, because the officers say it's a high crime area.").

103.    United States v. Sandoval, 131 F. App'x 614, 615–16 n.1 (9th Cir. 2005) ("While the district court may not have 'examine[d] with care the specific data underlying' this assertion, we conclude that Officer Stys was still entitled to give this factor *some* weight in forming reasonable suspicion.") (citation omitted) (quoting *Montero-Camargo*, 208 F.3d at 1139 n.32).

104.    United States v. Diaz-Juarez, 299 F.3d 1138, 1145 (9th Cir. 2002) (Ferguson, J., dissenting) (arguing that testimony claiming a certain road "was located in a high crime area" without evidence to support the officers' observation "was a far cry from the 'specific data' required to support the assertion that the stop took place in a 'high crime' area.").

something more like a street block or intersection or cluster of street blocks rather than a neighborhood or city.[105]

The second assumption is that officers' assessments of high-crime areas are relatively accurate. There are good reasons to question that assumption. For one thing, officers may not be aware of crime rates everywhere in the city, particularly if they are expected to analyze whether an area is high crime at a low geographic level, such as an individual intersection or street block.[106]

Bureaucratic incentives might also skew officers' assessments of high-crime areas. From 2002 until at least 2011, NYPD officers had significant professional incentives to increase the number of stops they conducted.[107] These incentives may have led some officers to conduct stops even when they lacked reasonable suspicion, and, to justify those stops, they may have invoked high-crime area in locations that were not high crime.[108]

An officer's intuition about whether an area is high crime may also be biased by legally suspect variables.[109] To begin with, officers might be influenced by the suspect's race. Prior empirical work has documented that, among suspects stopped by the police, people of color are more likely than whites to be arrested.[110] The process by which officers form suspicion about Black and white suspects might also differ. One particularly relevant study followed police officers in Savannah, Georgia, and observed the process by which they developed suspicion about pedestrians and motorists.[111] The authors found that, among the people the officers suspected of possible criminal behavior, suspicion of Black suspects was four times more likely than suspicion of white suspects to be based on "[n]onbehavioral" or contextual factors, such as "appearance, . . . time, and place."[112] These results suggest that officers may be

105.    *See supra* notes 85–87 and accompanying text.

106.    *See* DEREK J. PAULSEN & MATTHEW B. ROBINSON, CRIME MAPPING AND SPATIAL ASPECTS OF CRIME 38 (2d ed. 2004) (reviewing the empirical literature on police perceptions of the geographic distribution of crime and noting that "researchers have studied the spatial perceptions of police officers as they relate to crime patterns within a city and found that they . . . are incorrect").

107.    *See* Floyd v. New York, 959 F. Supp. 2d 540, 591–602 (S.D.N.Y. 2013) (documenting live testimony, depositions, roll call recordings, internal NYPD documents, and survey results, and concluding that the "most plausible explanation" for a 700 percent increase in stops from 2002 to 2011 was "significant pressure" on police officers "to increase their stop activity").

108.    *See* Fagan & Geller, *supra* note 23, at 69–81 (showing patterns in the circumstances that police officers invoke to establish reasonable suspicion in stop and frisk activities).

109.    For a recent general discussion of implicit bias related to stop and frisk, see L. Song Richardson, *Implicit Racial Bias and Racial Anxiety: Implications for Stops and Frisks*, 15 OHIO ST. J. CRIM. L. 73, 75–78 (2017).

110.    *See* Tammy Rinehart Kochel, David B. Wilson & Stephen D. Mastrofski, *Effect of Suspect Race on Officers' Arrest Decisions*, 49 CRIMINOLOGY 473, 486–90 (2011) (conducting a meta-analysis of twenty-seven independent datasets and finding that people of color are arrested more frequently than white people).

111.    Geoffrey P. Alpert, John M. MacDonald & Roger G. Dunham, *Police Suspicion and Discretionary Decision Making During Citizen Stops*, 43 CRIMINOLOGY 407, 418 (2005).

112.    *Id.* at 419 ("Behavioral criteria included specific actions by citizens that were either illegal or interpreted by the officer as suspicious. One example is observing a traffic offense. Obviously, not all

more likely to rely on softer contextual factors, including local crime rates, in forming suspicion about people of color than about white people.

Officers' high-crime assessments might also be unduly influenced by the racial, ethnic, and socioeconomic composition of the surrounding neighborhood. As the concurrence observed in *Montero-Camargo*, the concept of a high-crime area "can easily serve as a proxy for race or ethnicity."[113] Indeed, bias among officers might lead them to consciously or subconsciously believe that people of color are more likely to commit crime, and thus, to perceive that communities of color have higher crime rates.[114] Several prior empirical studies lend some support to this hypothesis. One study found a positive correlation between the proportion of young Black men in a neighborhood and residents' perceptions of neighborhood crime, even after controlling for official crime rates.[115] A second paper found that measures of neighborhood race and poverty are stronger predictors of residents' perceptions of disorder than actual disorder itself.[116] That finding is particularly important for police departments like the NYPD, which has heavily emphasized order-maintenance policing.[117] One additional study in New York City found that measures of the racial and socioeconomic composition of an area were stronger predictors of the number of police stops in that area than measures of physical and social disorder.[118] This body of empirical work suggests that the racial and socioeconomic characteristics of a neighborhood may influence a police officer's assessment of whether an area is high crime.

Furthermore, given the highly discretionary nature of the high-crime inquiry, any two given officers might disagree about whether a particular level of crime qualifies as high. Thus, an officer's high-crime area assessment might

---

police officers stop all traffic violators, but an observed traffic violation justifies an officer making a stop. Nonbehavioral criteria included officer concern about an individual's appearance, the time and place, and descriptive information provided to an officer.").

113.    United States v. Montero-Camargo, 208 F.3d 1122, 1138 (9th Cir. 2000) (Kozinski, J., concurring).

114.    *See* B. Michelle Peruche & E. Ashby Plant, *The Correlates of Law Enforcement Officers' Automatic and Controlled Race-Based Responses to Criminal Suspects*, 28 BASIC & APPLIED SOC. PSYCHOL. 193, 193–94 (2006) (discussing the theoretical and empirical literature on implicit racial bias and policing).

115.    Lincoln Quillian & Devah Pager, *Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime*, 107 AM. J. SOC. 717, 717–18 (2001); *see also* George F. Rengert & William V. Pelfrey, Jr., *Cognitive Mapping of the City Center: Comparative Perceptions of Dangerous Places*, *in* CRIME MAPPING AND CRIME PREVENTION: CRIME PREVENTION STUDIES 213–15 (D. Weisburd & T. McEwan eds., 1997) (finding a negative correlation between perceived safety and concentration of ethnic minorities).

116.    Robert J. Sampson & Stephen W. Raudenbush, *Seeing Disorder: Neighborhood Stigma and the Social Construction of "Broken Windows"*, 67 SOC. PSYCHOL. Q. 319, 319 (2004) (reporting based on data from Chicago, Seattle, and Baltimore that "[o]bserved disorder predicts perceived disorder, but racial and economic context matter [*sic*] more").

117.    *See supra* note 44–50 and accompanying text.

118.    Jeffrey Fagan & Garth Davies, *Street Stops and Broken Windows:* Terry*, Race, and Disorder in New York City*, 28 FORDHAM URB. L.J. 457, 463–64 (2000); *id.* at 484–87 (describing measures of disorder).

also be influenced by inter-officer differences in perception of what qualifies as "high-crime." We have found little empirical research on such disparities among the police, but one study surveyed residents in Baltimore in the 1980s and 1990s about their perceptions of physical and social disorder in their neighborhoods.[119] The author found that residents' perceptions of disorder varied widely, even among residents living in the same neighborhood.[120] Police officers' assessments of crime and disorder may be subject to the same disparities.

*Wardlow*'s third empirical assumption concerns predictive power. In justifying investigative stops, a police officer must be able to articulate specific facts that together form a reasonable suspicion that the suspect was engaged in or was about to engage in crime. For courts to accept "high-crime area" as a basis of reasonable suspicion, the officer's determination that an area is high crime should predict whether suspects are, on average, engaged in crime. *Wardlow* thus assumes that, controlling for other stated bases of reasonable suspicion, there is a higher probability that a suspect is engaged in a crime where the officer invokes high-crime area as a basis of the stop.

Over fifteen years have passed since the Supreme Court issued *Wardlow*, and yet we still know almost nothing about how officers apply the high-crime area standard. We therefore don't know if any of its empirical assumptions are satisfied in practice. In the remainder of this Article, we attempt to test those assumptions.

## III.

### DATA

Our dataset contains information on 2,455,030 stops conducted by the NYPD between 2007 and 2012.[121] The data derive from the Stop, Question and Frisk Report Worksheet, more commonly referred to as the UF-250. Officers must fill out the Worksheet after every stop they conduct.[122] The form collects a range of information, including the suspected crime; the suspect's race, gender,

119. RALPH B. TAYLOR, BREAKING AWAY FROM BROKEN WINDOWS: BALTIMORE NEIGHBORHOODS AND THE NATIONWIDE FIGHT AGAINST CRIME, GRIME, FEAR, AND DECLINE 164–69 (2001).

120. *Id.*

121. The original dataset also included data from 2004 to 2005. We dropped all stops from those years because we were unable to geocode a large proportion of stops due to data quality issues. We also dropped all stops classified as "radio runs" because, in those cases, the officer is searching for a specific person who matches a suspect description, usually from a recent crime in the surrounding area.

122. It is possible some officers did not always fill out a stop form for stops resulting in an arrest. However, we suspect this practice was rare given the intense bureaucratic pressure to file paperwork to demonstrate officer productivity in both stops and arrests. *See supra* note 107 and accompanying text. The New York State Attorney General conducted an analysis of arrests resulting from stops from 2009 to 2012. Of the 150,330 arrest records received, just 5 percent had no corresponding stop information. *See* ERIC T. SCHNEIDERMAN, A REPORT ON ARRESTS ARISING FROM THE NEW YORK CITY POLICE DEPARTMENT'S STOP-AND-FRISK PRACTICES 7 (2013), at https://ag.ny.gov/pdfs/OAG_REPORT_ON_SQF_PRACTICES_NOV_2013.pdf. [https://perma.cc/QA57-M2VY] (finding 142,596 records for 150,330 arrests).

and age; whether the stop resulted in the recovery of weapons, drugs, or other contraband; and whether the suspect was arrested or issued a citation. Most important for our purposes, the form also requires officers to identify the circumstances which led to the stop—that is, the bases of reasonable suspicion—by checking off a series of boxes, including a box for high-crime areas.[123]

The Worksheet also collects information on the police precinct and street address where the stop took place. We used this information to geocode the stops to each of the 2,211 census tracts and 5,722 block groups in New York City from the 2000 census.[124] We dropped 185,967 stops (5 percent) that we were unable to geocode.

We also obtained criminal-complaint data from the NYPD on each of the 4.6 million crimes reported to the department from 2004 to 2012. We geocoded the complaints to block groups[125] and then merged various lagged measures of crime with our stop data.

Finally, we obtained data on block groups in New York City from the American Community Survey 2005–2009 (ACS). Specifically, we obtained measures of the racial and socioeconomic composition of residents in each area.

Table 1 provides descriptive statistics concerning the characteristics of each of the stops in our data. Blacks are disproportionately represented. Indeed, roughly 55 percent of civilians stopped were non-Hispanic Black, while that group made up less than a quarter of the city's population.[126] Roughly 31 percent of civilians stopped were Hispanic, while that group made up about 28 percent of the city.[127] Roughly 93 percent were male. Stopped civilians were, on average, 28 years old. Stops rarely resulted in "hits." Indeed, officers recovered a weapon in 1 percent of stops and other contraband in another 2 percent. Stops resulted in an arrest just 7 percent of the time.

Table 2 provides descriptive statistics on the characteristics of the areas in which stops were conducted. We denote precinct-level measures with the abbreviation "PCT" and block group-level measures with "BG." Stops occurred

---

123. N.Y. POLICE DEP'T, STOP, QUESTION AND FRISK REPORT WORKSHEET (2011), https://assets.documentcloud.org/documents/686450/stop-question-and-frisk-report-worksheet.pdf [https://perma.cc/EA76-7C9G]. The full list of circumstances includes: (1) suspicious bulge; (2) casing a victim or location; (3) carrying objects used in commission of crime; (4) wearing clothing/disguises used in commission of crime; (5) suspect fits description; (6) visible drug transaction; (7) furtive movements; (8) acting as lookout; and (9) visible violent crime. The form also provides additional check boxes labeled as "additional" circumstances: (1) association with known criminals; (2) change of direction upon seeing police; (3) evasive actions; (4) high incidence of reported offense; (5) part of ongoing investigation; (6) proximity to crime; (7) report of victim or witness; (8) sights and sounds of crime; and (9) time of day. *Id.* Finally, the worksheet provides space for officers to identify "Other Reasonable Suspicion of Criminal Activity" in their own words. *Id.*

124. We are grateful to Matt Ruther for his exceptional work geocoding the data.

125. We were unable to code the census tract and block group of 204,923 criminal complaints (4 percent).

126. U.S. CENSUS BUREAU, *supra* note 60.

127. *Id.*

in block groups that were, on average, 28 percent white, 40 percent Black, and 36 percent Hispanic. As those averages suggest, stops are concentrated in communities of color. Roughly 40 percent of all stops occurred in block groups that were majority Black and another 30 percent took place in block groups that were majority Hispanic.

Officers also tended to stop civilians in economically disadvantaged areas of the city. On average, stops occurred in block groups in which 30 percent of households received an annual income of less than $20,000; 55 percent of households were single-headed; and 25 percent of households lived below the national poverty line.

Table 1. Stop Characteristics[128]

|  | Mean | SD | Minimum | Maximum | N |
|---|---|---|---|---|---|
| **Demographics** | | | | | |
| Black Non-Hispanic | 0.55 | 0.50 | 0 | 1 | 2,455,030 |
| Black Hispanic | 0.07 | 0.25 | 0 | 1 | 2,455,030 |
| White Non-Hispanic | 0.09 | 0.28 | 0 | 1 | 2,455,030 |
| White Hispanic | 0.24 | 0.43 | 0 | 1 | 2,455,030 |
| Other Race | 0.06 | 0.24 | 0 | 1 | 2,455,030 |
| Male | 0.93 | 0.25 | 0 | 1 | 2,455,030 |
| Age | 28.1 | 11.5 | 8 | 99 | 1,963,208 |
| **Suspected Crime** | | | | | |
| Violent | 0.20 | 0.40 | 0 | 1 | 2,455,030 |
| Property | 0.22 | 0.41 | 0 | 1 | 2,455,030 |
| Weapon | 0.29 | 0.45 | 0 | 1 | 2,455,030 |
| Drug | 0.11 | 0.31 | 0 | 1 | 2,455,030 |
| Trespass | 0.11 | 0.31 | 0 | 1 | 2,455,030 |
| Quality of Life | 0.02 | 0.13 | 0 | 1 | 2,455,030 |
| Other Crime | 0.07 | 0.25 | 0 | 1 | 2,455,030 |
| **Bases of Suspicion** | | | | | |
| Suspicious Bulge | 0.10 | 0.30 | 0 | 1 | 2,455,030 |
| Casing | 0.31 | 0.46 | 0 | 1 | 2,455,030 |
| Clothes | 0.05 | 0.21 | 0 | 1 | 2,455,030 |
| Suspect Description | 0.12 | 0.33 | 0 | 1 | 2,455,030 |
| Drug Transaction | 0.10 | 0.30 | 0 | 1 | 2,455,030 |
| Furtive Movements | 0.51 | 0.50 | 0 | 1 | 2,455,030 |
| Lookout | 0.17 | 0.38 | 0 | 1 | 2,455,030 |
| Criminal Object | 0.03 | 0.16 | 0 | 1 | 2,455,030 |
| Violent Crime | 0.09 | 0.29 | 0 | 1 | 2,455,030 |
| Criminal Associates | 0.04 | 0.20 | 0 | 1 | 2,455,030 |
| Change Direction | 0.26 | 0.44 | 0 | 1 | 2,455,030 |
| Evasive Behavior | 0.21 | 0.41 | 0 | 1 | 2,455,030 |
| High Crime | 0.59 | 0.49 | 0 | 1 | 2,455,030 |
| Ongoing Investigation | 0.13 | 0.34 | 0 | 1 | 2,455,030 |
| Proximity to Crime | 0.19 | 0.39 | 0 | 1 | 2,455,030 |
| Witness Report | 0.09 | 0.28 | 0 | 1 | 2,455,030 |
| Sights/Sounds of Crime | 0.02 | 0.14 | 0 | 1 | 2,455,030 |
| Time | 0.41 | 0.49 | 0 | 1 | 2,455,030 |
| **Stop Outcomes** | | | | | |
| Weapon Found | 0.01 | 0.11 | 0 | 1 | 2,455,030 |
| Other Contraband Found | 0.02 | 0.14 | 0 | 1 | 2,455,030 |
| Arrest Made | 0.07 | 0.25 | 0 | 1 | 2,455,030 |

---

128.    When totaled together, suspected crimes exceed 100 percent due to rounding.

To avoid the risk of overfitting, we randomly partitioned the data into a training dataset and a testing dataset.[129] The training dataset consists of one-third of all stops, and the testing dataset consists of the remaining two-thirds. All of the results we report in the following sections are based on the testing dataset.

Table 2. Social and Crime Characteristics of Stop Areas[130]

|  | Mean | SD | Minimum | Maximum | n |
|---|---|---|---|---|---|
| **Crime** | | | | | |
| PCT Viol Index Crime 12mo | 1,801.4 | 908.2 | 204.0 | 4,468.0 | 2,455,030 |
| PCT Prop Index Crime 12mo | 2,221.8 | 1,078.9 | 471.0 | 8,172.0 | 2,455,030 |
| BG Viol Index Crime 12mo | 41.5 | 33.2 | 0.0 | 296.0 | 2,455,009 |
| BG Prop Index Crime 12mo | 52.5 | 112.9 | 0.0 | 2,365.0 | 2,455,009 |
| **Racial/Ethnic Composition** | | | | | |
| BG White | 0.28 | 0.27 | 0 | 1 | 2,408,576 |
| BG Black | 0.40 | 0.32 | 0 | 1 | 2,408,576 |
| BG Hispanic | 0.36 | 0.27 | 0 | 1 | 2,408,576 |
| **Socioeconomic Characteristics** | | | | | |
| BG No HS Graduation | 0.29 | 0.17 | 0 | 1 | 2,407,049 |
| BG Income $20k-50k | 0.30 | 0.13 | 0 | 1 | 2,402,373 |
| BG Income $50k-125k | 0.28 | 0.16 | 0 | 1 | 2,402,373 |
| BG Income >$125k | 0.09 | 0.13 | 0 | 1 | 2,402,373 |
| BG Median Income | 41,604 | 28,589 | 2,499 | 250,001 | 2,388,340 |
| BG Families in Poverty | 0.25 | 0.19 | 0.00 | 1 | 2,375,101 |
| BG Assistance | 0.08 | 0.08 | 0.00 | 0.70 | 2,402,373 |
| BG Vacant Properties | 0.09 | 0.10 | 0.00 | 1 | 2,404,982 |
| BG Single-headed Household | 0.55 | 0.25 | 0.00 | 1 | 2,375,101 |

IV.

EMPIRICAL ANALYSES

Our empirical analysis unfolds in two steps. As noted, *Wardlow* and other subsequent cases have left open many questions about the definition of a high-crime area. We therefore have little information about how officers interpret and apply the high-crime area standard—about the geographic scope and temporal horizon they use to evaluate whether an area is high crime and about the kinds of crimes they consider most relevant. In the first stage of our analysis, we seek to answer some of these questions by identifying a model of crime that best predicts when officers invoke the high-crime area standard to justify their stops. Admittedly, our answers are necessarily rough. The dataset contains stops conducted by almost 20,000 unique officers, each of whom may apply the high-

---

129.    *See* RICHARD A. BERK, STATISTICAL LEARNING FROM A REGRESSION PERSPECTIVE 31–35 (2d ed. 2016).

130.    We obtained crime data from the NYPD. Racial and socioeconomic variables derive from the American Community Survey, 2005–2009 and are constant for all years in the data. Block group-level 2009 5-year ACS data on New York City can be obtained from socialexplorer.com. PCT refers to police precinct. BG refers to census block group. Index crimes are groupings commonly used in crime reporting that were developed by the FBI. Violent Index Crimes include Murder, Armed Robbery, Aggravated Assault, and Forcible Rape. Property Index Crimes include Burglary, Larceny, Motor Vehicle Theft, and Arson. Federal Bureau of Investigation, Crime in the United States 2017, available at https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017/topic-pages/violent-crime [https://perma.cc/59UM-4A8A].

crime area standard differently. Our goal, however, is to find models of crime that best fit the data with the hopes of gaining insight about the behavior of the typical officer. In the second stage of our analysis, we then use these models to evaluate *Wardlow* and test its empirical assumptions.

### A. *Modeling Crime Rates on Officers' Invocation of HCA*

To identify a model of crime that best predicts when officers invoke the high-crime area standard, we fit a series of models regressing whether the officer invoked HCA as a basis of suspicion on a variety of different measures of crime. For ease of interpretation, we primarily report results from linear probability models.[131] We are particularly interested in each model's $R^2$, which measures how well each model fits the data on a scale from 0 to 1. We therefore report both the $R^2$ from the linear probability model and the McFadden's $R^2$ from the corresponding logit, the latter of which better fits the binary structure of the dependent variable.[132] Throughout the Article, we report cluster-robust standard errors clustered at the officer and block group-level. Our threshold for statistical significance is the 0.05 level, but our tables also report when coefficients are statistically significant at the 0.1 level.

Our first step is to examine, in broad categories, the kinds of crimes police officers appear to care about most in assessing whether an area is high crime. We begin by fitting a linear model with just one independent variable—the number of crimes in the past twelve months in the police precinct in which the stop took place. In Table 3, Model 1 shows the results when the independent variable measures total crime, Model 2 shows the results for violent crime, and Model 3 shows the results for property crime. To reduce the number of digits displayed after the decimal, we have divided each of the crime variables by 100 for all of the models presented in the remainder of the Article.

None of the models in Table 1 strongly predict whether an officer invokes HCA, but violent crime appears to be the strongest predictor. The McFadden's $R^2$ for the violent-crime model is .001, which is substantially larger than that of the total (.0001) and property-crime models (.0002). On the whole, in assessing whether an area is high crime, officers appear to focus more on violence than property offenses.

---

131. The logit models produced substantively similar results. Unfortunately, we were unable to fit logits for models with fixed effects for officers due to the computational difficulty of calculating over 20,000 fixed effects with cluster-robust standard errors.

132. *See* Paul Allison, *What's the Best R-Squared for Logistic Regression?*, STAT. HORIZONS (Feb. 13, 2013), https://statisticalhorizons.com/r2logistic [https://perma.cc/9VBV-E27V] (describing different measures of fit for logit models and endorsing McFadden's $R^2$).

Table 3. Modeling HCA on Violent and Property Crime in the Last 12 Months

|  | Mod 1 | Mod 2 | Mod 3 |
|---|---|---|---|
| **Intercept** | 0.5701** | 0.5523** | 0.6054** |
|  | [0.0127] | [0.0109] | [0.0110] |
| **PCT Total** | 0.0002 |  |  |
|  | [0.0001] |  |  |
| **PCT Violent** |  | 0.0019** |  |
|  |  | [0.0005] |  |
| **PCT Property** |  |  | -0.0008† |
|  |  |  | [0.0004] |
| **Adjusted R^2 (LPM)** | 0.0002 | 0.0013 | 0.0003 |
| **McFadden's R^2 (Logit)** | 0.0001 | 0.0009 | 0.0002 |
| **N** | 1,636,384 | 1,636,384 | 1,636,384 |

Notes: † $p<0.10$, *$p<0.05$, ** $p<0.01$

To make things more concrete, Figure 1 depicts the relationship between HCA and precinct-level violent crime. The black curve displays the best-fit line, which is nearly flat but suggests officers invoke HCA just slightly more often as violent crime increases. The gray line, which represents a local linear smoother, tells a similar story. For nearly the entire distribution—from roughly 500 to 3,300 crimes on the X-axis—the rate at which officers invoke HCA remains roughly stable, confirming that there is little correlation between HCA and crime levels. The slope of the curve only steepens above roughly 3,300 crimes. Importantly, we should be cautious in interpreting the right-hand side of the graph. The bottom of the graph contains a rug, which illustrates the number of observations located along the X-axis. As the rug shows, there are almost no observations between 3,300 and 3,900 crimes. All observations above 3,900 arise from just one precinct in Brooklyn—the 74th. Thus, the entire rise in the HCA invocation rate above 3,300 is due to just one precinct in the city.[133]

---

133.    The undue leverage of outliers can skew the estimate of a statistical relationship between variables. Richard Berk, *New Claims About Executions and General Deterrence: Déjà Vu All Over Again*, 2 J. EMPIRICAL LEGAL STUD. 303, 305–06 (2005).

Figure 1. HCA on Precinct-Level Violent Crime in Previous 12 Months
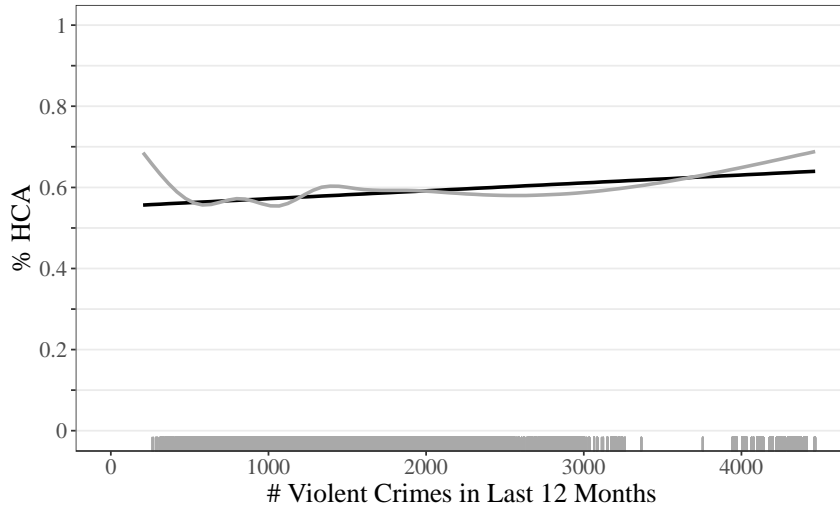


Having found that officers appear to prioritize violent crime over property crime in assessing HCA, we next turn to the question of geographic scope. Recall that police precincts cover, on average, four square miles and that census block groups cover 0.05 square miles. The first model in Table 4 replicates our most predictive model in the previous table, which regresses HCA on precinct-level (denoted by the abbreviation "PCT") violent crime in the last twelve months. Model 2 regresses HCA on block group-level (denoted by "BG") violent crime in the last twelve months, and finds a positive and statistically significant correlation. Once again, to make things more concrete, Figure 2 depicts the relationship between HCA and block group-level violent crime from Model 2. As before, the black curve displays the best-fit line, which is nearly flat but suggests officers invoke HCA just slightly more often as violent crime increases. The gray line moves up and down across the X-axis slightly, but remains relatively flat over all. The McFadden's $R^2$ for the block group-level model is substantially smaller than that of the precinct-level model, suggesting that precinct-level measures of crime are better predictors of HCA than block group-level measures.

Figure 2. HCA on Block Group-Level Violent Crime in Previous 12 Months
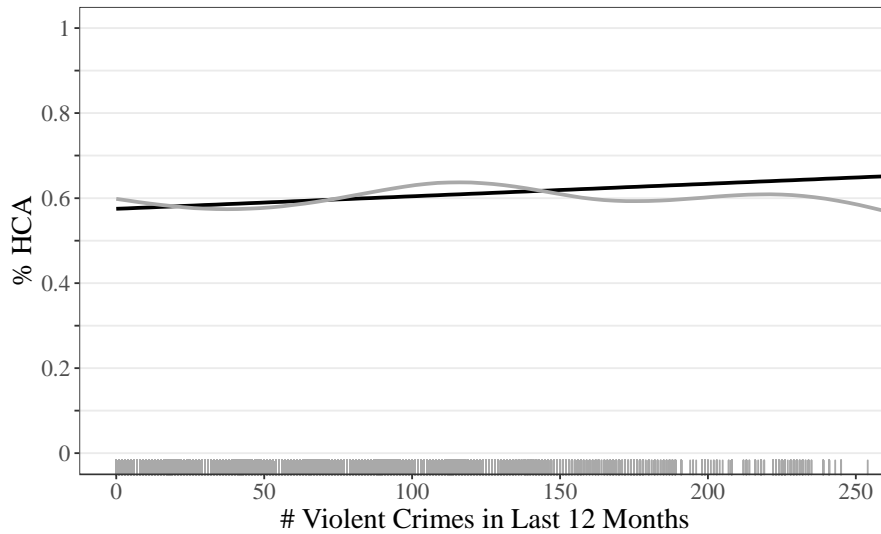


Model 3 combines both the precinct-level and block group-level variables together.[134] The coefficient on the precinct-level measure remains roughly the same and statistically significant, while the coefficient for the block group-level measure drops by about a third and is no longer statistically significant. Based on the models' respective $R^2$, precinct-level violent crime appears to be the strongest predictor of HCA. Though, the $R^2$ increases slightly when we include both the precinct and block group-level measures.

Table 4. Modeling HCA on Different Geographic Scopes

|  | Mod 1 | Mod 2 | Mod 3 |
|---|---|---|---|
| **Intercept** | 0.5523** | 0.5748** | 0.5462** |
|  | [0.0109] | [0.0059] | [0.0102] |
| **PCT Violent** | 0.0019** |  | 0.0018** |
|  | [0.0005] |  | [0.0005] |
| **BG Violent** |  | 0.0294* | 0.0206 |
|  |  | [0.0139] | [0.0139] |
| **Adjusted R^2 (LPM)** | 0.0013 | 0.0004 | 0.0014 |
| **McFadden's R^2 (Logit)** | 0.0009 | 0.0003 | 0.0011 |
| **N** | 1,636,384 | 1,636,367 | 1,636,367 |

Notes: † p<0.10, *p<0.05, ** p<0.01

---

134.     One might reasonably wonder whether modeling violent crime at both the precinct and block group-level in Model 3 introduces high levels of multi-collinearity. However, the Pearson correlation between these two variables is just 0.17 and the standard errors for each coefficient do not change. The low correlation suggests that there may be different data-generating processes at work in the precinct versus block-group decisions to invoke HCA.

We next consider temporal scope. Table 5 depicts four models regressing HCA on measures of violent crime with different time spans: 1 month, 3 months, 12 months, and 24 months. McFadden's $R^2$ shows that the variation explained by the models increases as the temporal scope expands from 1 to 12 months. However, it maxes out at 12 months, as the 24-month measures do not explain additional variation. These patterns suggest that an officer's decision to invoke HCA to justify a stop is generally based on crimes that occurred in the last 12 months.[135]

Table 5. Modeling HCA on Different Temporal Horizons

|  | Mod 1 | Mod 2 | Mod 3 | Mod 4 |
|---|---|---|---|---|
| **Intercept** | 0.5567** [0.0091] | 0.5529** [0.0095] | 0.5462** [0.0102] | 0.5460** [0.0104] |
| **PCT Violent 1mo** | 0.0180** [0.0055] |  |  |  |
| **BG Violent 1mo** | 0.1174 [0.0966] |  |  |  |
| **PCT Violent 3mo** |  | 0.0063** [0.0020] |  |  |
| **BG Violent 3mo** |  | 0.0604 [0.0463] |  |  |
| **PCT Violent 12mo** |  |  | 0.0018** [0.0005] |  |
| **BG Violent 12mo** |  |  | 0.0206 [0.0139] |  |
| **PCT Violent 24mo** |  |  |  | 0.0009** [0.0003] |
| **BG Violent 24mo** |  |  |  | 0.0099 [0.0071] |
| **Adjusted R^2 (LPM)** | 0.001 | 0.0012 | 0.0014 | 0.0014 |
| **McFadden's R^2 (Logit)** | 0.0008 | 0.0009 | 0.0011 | 0.0011 |
| **N** | 1,636,367 | 1,636,367 | 1,636,367 | 1,636,367 |

Notes: † p<0.10, *p<0.05, ** p<0.01

Having learned that violent crime in the past 12 months is the strongest predictor of how officers invoke HCA, we now return to the question of which specific crime categories matter most.

Table 6 depicts a series of models with more specific crime categories. As a baseline, Model 1 replicates Model 3 from the previous table by including measures of violent crime at the precinct- and block group-level in the last twelve months. Model 2 uses more violent crime instead: non-negligent homicide, negligent homicide, assault, and robbery. McFadden's $R^2$ shows that this model explains substantially more of the variation in HCA, but we note that nearly half of the coefficients are negative. Adding discrete categories of property, drug, and weapons crimes in Model 3 further increases the variation explained by the model, but again, we note that twelve of the twenty coefficients are negative.

---

135.    The variation explained does not substantially increase when we combine the 12-month variable with any of the other timing variables.

In Model 4, we removed all of the variables associated with negative coefficients and any additional variables that became negative thereafter. We did so because the purpose of our analysis in the next section is to evaluate how police officers invoke HCA to justify their stops. *Wardlow* does not permit officers to claim an area is high crime based on lower crime rates. We, therefore, removed variables with negative coefficients so those variables do not contribute to the $R^2$ of our models, which we use to evaluate officers' HCA assessments. For the sake of parsimony, we also removed from Model 4 any remaining variables that are not statistically significant or that add little explanatory power to the model based on McFadden's $R^2$. The remaining three variables are precinct-level non-negligent homicide, precinct-level burglary, and block group-level arson. Together, these variables explain over a third of the variation explained by the full model, Model 3.

Table 6. Modeling HCA on Specific Crime Types in the Last 12 Months

| | Mod 1 | Mod 2 | Mod 3 | Mod 4 |
|---|---|---|---|---|
| **Intercept** | 0.5462** [0.0102] | 0.5522** [0.0102] | 0.5346** [0.0123] | 0.5239** [0.0127] |
| **PCT Violent** | 0.0018** [0.0005] | | | |
| **PCT Neg Homicide** | | 0.0801 [0.3835] | -0.5705 [0.3491] | |
| **PCT Non-neg Homicide** | | 0.1965** [0.0648] | 0.1163† [0.0624] | 0.1368** [0.0506] |
| **PCT Assault** | | -0.0014 [0.0013] | 0.0095** [0.0018] | |
| **PCT Robbery** | | 0.0047 [0.0031] | -0.0057† [0.0030] | |
| **PCT Arson** | | | 0.0105 [0.0300] | |
| **PCT Burglary** | | | 0.0192** [0.0055] | 0.0124** [0.0038] |
| **PCT Motor Vehicle Theft** | | | -0.0138* [0.0062] | |
| **PCT Theft** | | | -0.0014* [0.0006] | |
| **PCT Drug** | | | -0.0065** [0.0011] | |
| **PCT Weapon** | | | -0.0079 [0.0051] | |
| **BG Violent** | 0.0206 [0.0139] | | | |
| **BG Neg Homicide** | | -3.8115† [2.0633] | -3.1249* [1.4916] | |
| **BG Non-neg Homicide** | | 0.9867** [0.3312] | 0.6913* [0.2775] | |
| **BG Assault** | | 0.0893** [0.0198] | 0.0839** [0.0244] | |
| **BG Robbery** | | -0.1859** [0.0328] | -0.0757* [0.0366] | |
| **BG Arson** | | | 1.0648** [0.2452] | 1.6380** [0.2755] |
| **BG Burglary** | | | -0.1112† [0.0624] | |
| **BG Motor Vehicle Theft** | | | -0.1719* [0.0805] | |
| **BG Theft** | | | -0.0055 [0.0056] | |
| **BG Drug** | | | 0.0878** [0.0216] | |
| **BG Weapon** | | | -0.2830** [0.0865] | |
| **Adjusted R^2 (LPM)** | 0.0014 | 0.0032 | 0.0123 | 0.0045 |
| **McFadden's R^2 (Logit)** | 0.0011 | 0.0024 | 0.0092 | 0.0034 |
| **N** | 1,636,367 | 1,636,367 | 1,636,367 | 1,636,367 |

Notes: † $p<0.10$, *$p<0.05$, ** $p<0.01$

Until now, we have assumed that police officers think about and apply the HCA standard the same way regardless of the kind of crime they believe the suspect is committing. To probe this issue further, we subsetted the data by suspected crime—violent, property, weapons, and drug crimes, which together account for roughly 80 percent of all stops. We then repeated each of our analytic steps for the full dataset on each subset. When we replicated Tables 3 through 5 for each suspected-crime type, the results were remarkably similar: the strongest variables were typically precinct-level measures of violent crime within the last 12 months. There was, however, significant variation across suspected-crime types when we replicated Table 6 (which identified the specific crimes that best predict whether an officer invokes HCA).

Table 7 thus replicates Models 3 and 4 in Table 6 for each of the four suspected crime categories. Model 1 reports the results of a regression with all of the 12-month, precinct- and block-group measures of specific crime types on the subset of stops in which the officer suspected a violent crime. Model 2 reports the ultimate model we reached after sequentially removing all variables that were negative, statistically insignificant, or that added little predictive power to the model. Model 2 shows that the strongest predictors of whether an officer invokes HCA to justify stops motivated by a suspicion of violent crime is remarkably similar to the strongest predictors for all stops (reported in Table 6): precinct-level non-negligent homicide, and burglary.

Models 4, 6, and 8 show the variables with positive coefficients that most strongly predict whether an officer invokes HCA to justify stops in which the officer suspects a property, drug, and weapons offense, respectively. For property-crime stops, the most predictive crime measures are precinct-level burglary and motor vehicle theft. For drug-crime stops, the most predictive crime measures are block group-level drugs and arson. And for weapon-crime stops, the most predictive crime measures are precinct-level non-negligent homicide and block group-level assault and arson.

Table 7. Modeling HCA on Specific Crime Types, in the Past 12 Months, by Suspected Crime

| | Violent | | Property | | Drug | | Weapons | |
|---|---|---|---|---|---|---|---|---|
| | Mod 1 | Mod 2 | Mod 3 | Mod 4 | Mod 5 | Mod 6 | Mod 7 | Mod 8 |
| **Intercept** | 0.501** | 0.490** | 0.549** | 0.523** | 0.587** | 0.581** | 0.488** | 0.486** |
| | [0.016] | [0.014] | [0.016] | [0.015] | [0.015] | [0.007] | [0.020] | [0.011] |
| **PCT Neg Homicide** | -0.936† | | -0.171 | | -0.235 | | -0.863† | |
| | [0.555] | | [0.460] | | [0.535] | | [0.520] | |
| **PCT Non-neg Homicide** | 0.169* | 0.268** | 0.081 | | 0.106 | | 0.135 | 0.271** |
| | [0.085] | [0.070] | [0.076] | | [0.096] | | [0.085] | [0.067] |
| **PCT Assault** | 0.012** | | 0.008** | | 0.008** | | 0.014** | |
| | [0.003] | | [0.002] | | [0.003] | | [0.003] | |
| **PCT Robbery** | 0.003 | | -0.015** | | -0.006 | | -0.009* | |
| | [0.004] | | [0.004] | | [0.004] | | [0.004] | |
| **PCT Arson** | -0.096* | | -0.018 | | -0.044 | | 0.064 | |
| | [0.042] | | [0.038] | | [0.044] | | [0.041] | |
| **PCT Burglary** | 0.017* | 0.020** | 0.028** | 0.020** | 0.009 | | 0.023** | |
| | [0.008] | [0.005] | [0.006] | [0.005] | [0.008] | | [0.009] | |
| **PCT Motor Vehicle Theft** | -0.014 | | 0.007 | 0.017* | -0.029** | | -0.036** | |
| | [0.008] | | [0.007] | [0.008] | [0.009] | | [0.009] | |
| **PCT Theft** | -0.002** | | -0.001 | | 0 | | -0.002* | |
| | [0.001] | | [0.001] | | [0.001] | | [0.001] | |
| **PCT Drug** | -0.008** | | -0.005** | | -0.002 | | -0.008** | |
| | [0.002] | | [0.001] | | [0.002] | | [0.001] | |
| **PCT Weapon** | -0.005 | | -0.003 | | -0.009 | | -0.012† | |
| | [0.006] | | [0.006] | | [0.007] | | [0.007] | |
| **BG Neg Homicide** | -6.886** | | -0.898 | | -1.921 | | -3.589 | |
| | [1.821] | | [1.735] | | [2.128] | | [2.499] | |
| **BG Non-neg Homicide** | 0.184 | | 0.13 | | 0.273 | | 0.760* | |
| | [0.386] | | [0.302] | | [0.401] | | [0.364] | |
| **BG Assault** | 0.094** | | -0.009 | | 0.068** | | 0.137** | 0.079** |
| | [0.034] | | [0.024] | | [0.021] | | [0.028] | [0.024] |
| **BG Robbery** | 0.081† | | -0.102* | | -0.055 | | -0.043 | |
| | [0.042] | | [0.045] | | [0.040] | | [0.050] | |
| **BG Arson** | 0.448† | | 0.356† | | 0.729** | 0.956** | 0.932** | 1.485** |
| | [0.233] | | [0.210] | | [0.232] | [0.263] | [0.281] | [0.381] |
| **BG Burglary** | -0.099 | | 0.231** | | -0.184* | | -0.445** | |
| | [0.080] | | [0.059] | | [0.073] | | [0.091] | |
| **BG Motor Vehicle Theft** | -0.235* | | 0.133† | | -0.225† | | -0.168 | |
| | [0.096] | | [0.073] | | [0.126] | | [0.141] | |
| **BG Theft** | -0.018* | | -0.007** | | -0.025** | | -0.028* | |
| | [0.007] | | [0.003] | | [0.009] | | [0.012] | |
| **BG Drug** | -0.016 | | 0.029 | | 0.114** | 0.120** | 0.073* | |
| | [0.023] | | [0.021] | | [0.023] | [0.018] | [0.029] | |
| **BG Weapon** | -0.253* | | -0.274** | | -0.013 | | -0.113 | |
| | [0.100] | | [0.103] | | [0.071] | | [0.086] | |
| **Adjusted R^2 (LPM)** | 0.0166 | 0.0065 | 0.0143 | 0.0062 | 0.0117 | 0.007 | 0.0217 | 0.0084 |
| **McFadden's R^2 (Logit)** | 0.0124 | 0.0048 | 0.0108 | 0.0046 | 0.0091 | 0.0055 | 0.0163 | 0.0063 |
| **N** | 325,525 | 325,526 | 359,807 | 359,807 | 176,196 | 176,196 | 467,250 | 467,250 |

Notes: † $p<0.10$, *$p<0.05$, ** $p<0.01$

Taken together, our results point towards three key conclusions. First, officers' assessments of high-crime areas tend to focus most on violent crimes but the results do vary somewhat by suspected crime. For example, in stops for suspected property offenses, there is evidence officers give most weight to property crimes like burglary and motor-vehicle theft, and in stops for suspected drug crimes, there is evidence officers give most weight to drug crimes (though,

there is also evidence they give weight to arson, which is a difficult result to explain). Second, officers tend to focus on crimes that took place within the last twelve months. Third, officers tend to assess whether an area is high crime through a broad geographic lens, like precincts or neighborhoods. Though, there is also evidence they give less weight to block group-level measures of crime as well. In the following section, we apply these lessons to evaluate each of *Wardlow*'s three empirical assumptions.

## B.   Evaluating Wardlow

### 1.   Assumption 1: What is the Geographic Scope of a High-Crime Area?

We already have the data we need to assess *Wardlow*'s first empirical assumption, which concerns the geographic scope with which officers assess whether an area is high crime. Earlier in Table 4, we found that precinct-level measures of crime—which cover an average of four square miles—are substantially stronger predictors of whether an officer invokes high-crime area than block group-level measures of crime, which cover 0.05 square miles on average. This finding is consistent with officers applying the HCA standard primarily at a broad level of geography—like an entire police precinct or neighborhood.

Still, there is evidence that officers sometimes also analyze whether an area is high crime at smaller levels of geography. Indeed, as we learned from Table 7, we found that block group-level measures of crime are particularly strong predictors of whether an officer invokes HCA in drugs and weapons stops.

### 2.   Assumption 2: Are Police Officers' Assessments Accurate?

We next examine whether officers' high-crime area assessments are accurate. If they are, they should be strongly correlated with actual crime rates.

Before we describe those correlations, we first report basic descriptive data about the frequency and geographic distribution of stops in which officers invoke HCA. Officers invoke HCA quite frequently—in 59 percent of all stops—more often than any other basis of reasonable suspicion on the UF-250 form.[136] And, while they invoke HCA more often in certain parts of the city than others, they do so quite frequently everywhere. Figure 3 maps the rate at which officers invoke HCA across New York City.[137] To make them most visible, we have depicted in black those areas where officers invoke HCA *less* than 30 percent of the time. As the map shows, very few areas are black. Perhaps the most important takeaway is that officers called 98 percent of the block groups in the city high crime in at least 30 percent of stops conducted in those areas. In other words,

---

136.    *See supra* Part III tbl.1.
137.    The map depicts census tracts rather than block groups because census tracts are roughly twice as large as block groups and thus easier to depict in the figure.

officers are claiming that every block in New York City is high crime at one time or another.

Figure 3. Heat Map of Percent of Stops Officers Invoke HCA, 2007-2012
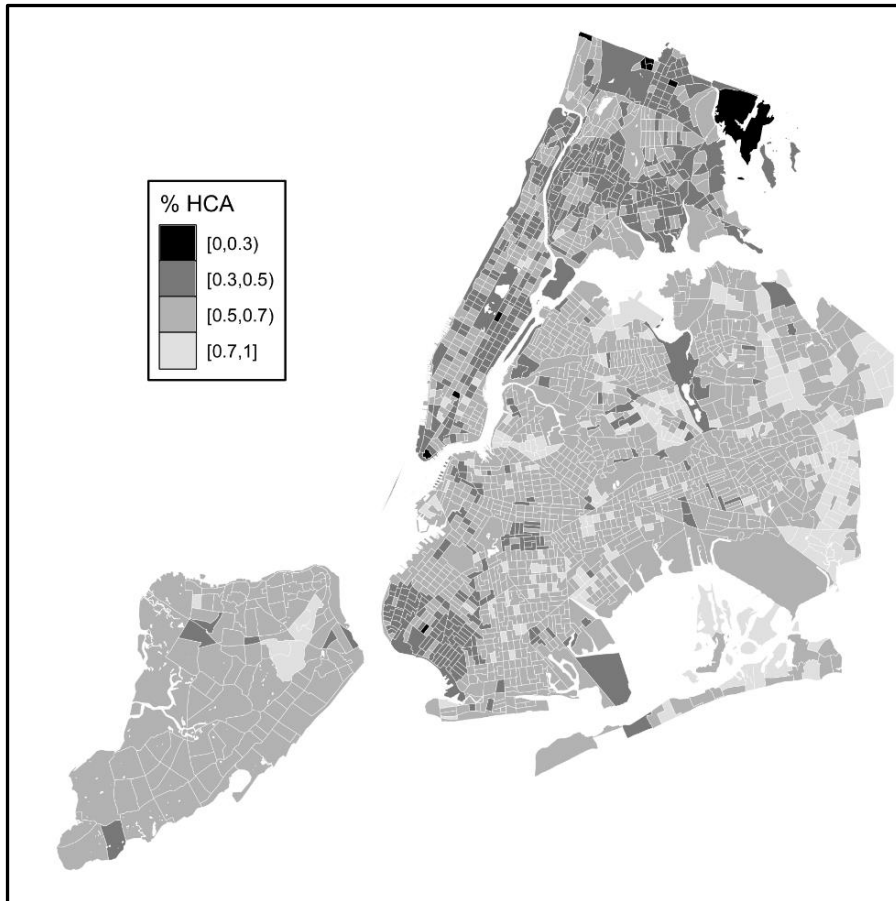


Turning now to the correlation between HCA and crime rates, once again, our earlier results shed some light. In Part IV.A, we modeled HCA on a variety of measures of crime that varied in terms of crime type, geographic scope, and temporal horizon. Across all of our models, the level of variation in HCA explained by a range of different measures of crime was remarkably close to zero. Our most predictive models had a McFadden's $R^2$ hovering around 0.01, meaning that actual crime rates explained around one percent of the variation in officers' assessments of whether areas are high crime. This pattern holds regardless of level of geography, temporal span, and categories of crime. It also held when we subsetted the data by the type of suspected crime.

If local crime rates are not the primary determinant of whether an officer invokes HCA as a basis of reasonable suspicion, what is? We next examine the

extent to which suspect demographics, the racial and socioeconomic characteristics of neighborhoods, and inter-officer disparities may help explain when and why officers invoke HCA to justify investigative stops.

### a.    Suspect Characteristics

Cognitive biases based on the demographic characteristics of the suspect might explain some of officers' inaccuracy in assessing high-crime areas. To explore this hypothesis, we fit a series of linear probability models, regressing HCA on a range of suspect characteristics, including race, gender, and age, and on local crime measures. When we analyze all stops together, we use the measures of crime in Models 1 and 4 from Table 6, which best predicted whether an officer would invoke HCA. When we break up our analysis by suspected-crime type, we use the crime measures from Models 2, 4, 6, and 8 in Table 7.

We begin by analyzing all stops together. In Table 8, Model 1 and 2 regress HCA on our two preferred sets of crime variables and on a set of dummy variables indicating the type of crime suspected by the officer. Models 3 and 4 replicate Models 1 and 2 but add variables for suspect race, ethnicity, sex, and age. For race and ethnicity, the comparison group is white non-Hispanic suspects. Both Models 3 and 4 estimate that police officers are 3 percentage points more likely to invoke HCA against a Black non-Hispanic suspect. They estimate no statistically significant difference for Black Hispanic suspects and for white Hispanic suspects. They also estimate that officers are almost 5 percentage points more likely to invoke HCA against a man than a woman, and that they are more likely to invoke HCA against younger suspects.

There are at least two possible explanations for these demographic differentials. One explanation is that areas with more young Black men have higher crime rates than our models account for. Another possible explanation is that officers who patrol areas with more young Black men have a higher propensity to invoke HCA.

To test both of these theories, Models 5 and 6 add fixed effects for block group and officer. The coefficients for Black non-Hispanic suspects and male suspects both fall by more than one half but remain statistically significant. The models also estimate that officers are 1 percentage point more likely to invoke HCA against both white Hispanic suspects and Black Hispanic suspects. The results are the same in Models 7 and 8 where we add fixed effects that interact year and block group. These results support the interpretation that officers are invoking HCA more frequently against young men of color. [138]

---

138.    We also refit the model with block-group officer fixed effects. The results were substantively similar, although there was a substantial loss in sample size because of the number of officers who conducted only one stop in a given block group.

Table 8. Modeling HCA on Suspect Characteristics and
Crime in the Last 12 Months

| | Mod 1 | Mod 2 | Mod 3 | Mod 4 | Mod 5 | Mod 6 | Mod 7 | Mod 8 |
|---|---|---|---|---|---|---|---|---|
| **Intercept** | 0.5408** | 0.5232** | 0.5059** | 0.4761** | | | | |
| | [0.0104] | [0.0125] | [0.0125] | [0.0142] | | | | |
| **Suspect Characteristics** | | | | | | | | |
| Black Non-Hispanic | | | 0.0306** | 0.0326** | 0.0155** | 0.0155** | 0.0134** | 0.0134** |
| | | | [0.0075] | [0.0075] | [0.0019] | [0.0019] | [0.0019] | [0.0015] |
| Black Hispanic | | | -0.0116 | -0.007 | 0.0122** | 0.0121** | 0.0088** | 0.0088** |
| | | | [0.0090] | [0.0090] | [0.0024] | [0.0024] | [0.0025] | [0.0020] |
| White Hispanic | | | 0.0034 | 0.0056 | 0.0113** | 0.0112** | 0.0096** | 0.0096** |
| | | | [0.0069] | [0.0070] | [0.0019] | [0.0019] | [0.0019] | [0.0015] |
| Other Race | | | -0.0252** | -0.0244** | -0.0009 | -0.001 | -0.0017 | -0.0017 |
| | | | [0.0077] | [0.0077] | [0.0025] | [0.0025] | [0.0026] | [0.0020] |
| Male | | | 0.0463** | 0.0466** | 0.0277** | 0.0277** | 0.0264** | 0.0264** |
| | | | [0.0035] | [0.0035] | [0.0017] | [0.0017] | [0.0017] | [0.0014] |
| Other Gender | | | -0.0082 | -0.0055 | 0.0169** | 0.0166** | 0.0189** | 0.0189** |
| | | | [0.0071] | [0.0070] | [0.0039] | [0.0039] | [0.0038] | [0.0031] |
| Age | | | -0.0009** | -0.0008** | -0.0001** | -0.0001** | -0.0001** | -0.0001** |
| | | | [0.0001] | [0.0001] | [0.0000] | [0.0000] | [0.0000] | [0.0000] |
| **Crime** | | | | | | | | |
| PCT Violent | 0.0017** | | 0.0014** | | -0.0030** | | -0.0010* | |
| | [0.0005] | | [0.0005] | | [0.0005] | | [0.0005] | |
| BG Violent | 0.0252* | | 0.0151 | | -0.0027 | | 0.0163† | |
| | [0.0119] | | [0.0113] | | [0.0067] | | [0.0087] | |
| PCT Non-neg Homicide | | 0.1424** | | 0.0679 | | 0.0039 | | -0.0011 |
| | | [0.0463] | | [0.0465] | | [0.0270] | | [0.0154] |
| PCT Burglary | | 0.0115** | | 0.0138** | | -0.0072** | | 0.0033* |
| | | [0.0035] | | [0.0035] | | [0.0024] | | [0.0014] |
| BG Arson | | 1.4652** | | 1.2398** | | -0.0769 | | 0.1029 |
| | | [0.2512] | | [0.2313] | | [0.0658] | | [0.0673] |
| **Suspected Crime** | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **Officer Fixed Effect** | | | | | ✓ | ✓ | ✓ | ✓ |
| **BG Fixed Effect** | | | | | ✓ | ✓ | | |
| **BG* Year Fixed Effect** | | | | | | | ✓ | ✓ |
| **Adjusted R^2 (LPM)** | 0.0098 | 0.0122 | 0.0131 | 0.0152 | 0.4068 | 0.4068 | 0.417 | 0.417 |
| **N** | 1,636,367 | 1,636,367 | 1,308,446 | 1,308,446 | 1,308,446 | 1,308,446 | 1,308,446 | 1,308,446 |

Notes: † p<0.10, *p<0.05, ** p<0.01

We next consider whether the results vary by the type of suspected crime. In Table 9, we reproduce Models 7 and 8 from Table 7 separately for stops where the officer suspected a violent, property, drug, or weapons offense. In the odd-numbered columns, we use the precinct-level and block group-level measures of violent crime we have used throughout, and in the even-numbered columns we use the specific measures of crime that best predicted HCA for each type of suspected crime (from Table 7).

As Table 9 shows, our models find that officers invoke HCA more frequently against males across all suspected-crime types. But the results for race, ethnicity, and age vary. Officers invoke HCA more often against non-Hispanic Black suspects for violent and weapons stops but not for property and drug stops. They invoke HCA more often against Black Hispanic and white Hispanic suspects for violent-crime stops but not for others. And they invoke HCA more often against younger suspects for violence, property, and weapons stops, but not for drug stops.

Taken together our models suggest several conclusions. First, controlling for the crime categories that officers appear to weigh most heavily in their assessment of high-crime areas, when we analyze all stops together, we find that officers are more likely to invoke HCA against young Black male suspects. And even when we control for officer and the narrow geographic area in which the stop took place, officers are still significantly more likely to invoke HCA against young Black men. Second, the results vary somewhat when we examine the data by suspected-crime type. We find evidence, for example, that officers invoke HCA more frequently against non-Hispanic Black suspects for violent and weapons stops but not for property and drug stops.

Table 9. Modeling HCA on Suspect Characteristics, in the Last 12 Months, By Suspected Crime

| | Violent | | Property | | Drug | | Weapons | |
|---|---|---|---|---|---|---|---|---|
| | Mod 1 | Mod 2 | Mod 3 | Mod 4 | Mod 5 | Mod 6 | Mod 7 | Mod 8 |
| **Suspect Characteristics** | | | | | | | | |
| Black Non-Hispanic | 0.0194** | 0.0194** | 0.0034 | 0.0034 | 0.0069 | 0.0068 | 0.0099* | 0.0099* |
| | [0.0055] | [0.0055] | [0.0032] | [0.0032] | [0.0055] | [0.0055] | [0.0042] | [0.0042] |
| Black Hispanic | 0.0170** | 0.0170** | 0.0032 | 0.0033 | 0.0081 | 0.0081 | 0.0008 | 0.0008 |
| | [0.0064] | [0.0064] | [0.0051] | [0.0051] | [0.0075] | [0.0075] | [0.0048] | [0.0048] |
| White Hispanic | 0.0132* | 0.0132* | 0.0057† | 0.0057† | 0.0064 | 0.0063 | 0.005 | 0.005 |
| | [0.0057] | [0.0057] | [0.0032] | [0.0032] | [0.0057] | [0.0057] | [0.0043] | [0.0043] |
| Other Race | -0.0064 | -0.0064 | -0.0018 | -0.0019 | -0.0093 | -0.0093 | -0.0055 | -0.0055 |
| | [0.0067] | [0.0067] | [0.0043] | [0.0043] | [0.0079] | [0.0079] | [0.0056] | [0.0056] |
| Male | 0.0628** | 0.0628** | 0.0171** | 0.0171** | 0.0103* | 0.0103* | 0.0261** | 0.0261** |
| | [0.0053] | [0.0053] | [0.0037] | [0.0037] | [0.0047] | [0.0047] | [0.0040] | [0.0040] |
| Other Gender | 0.0577** | 0.0577** | -0.0027 | -0.0027 | 0.0228* | 0.0227* | 0.0221** | 0.0221** |
| | [0.0097] | [0.0097] | [0.0092] | [0.0092] | [0.0116] | [0.0116] | [0.0068] | [0.0068] |
| Age | -0.0006** | -0.0006** | 0.0002* | 0.0002† | 0 | 0 | -0.0007** | -0.0007** |
| | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0001] |
| **Crime** | | | | | | | | |
| PCT Violent | -0.0003 | | -0.0004 | | -0.0023† | | -0.0004 | |
| | [0.0009] | | [0.0008] | | [0.0014] | | [0.0008] | |
| BG Violent | 0.0225 | | -0.0115 | | 0.0452† | | 0.0259* | |
| | [0.0223] | | [0.0215] | | [0.0273] | | [0.0124] | |
| PCT Non-neg Homicide | | 0.0112 | | | | | | 0.0422 |
| | | [0.0545] | | | | | | [0.0455] |
| PCT Burglary | | -0.0026 | | 0.0090† | | | | |
| | | [0.0052] | | [0.0047] | | | | |
| PCT Motor Vehicle Theft | | | | 0.0106 | | | | |
| | | | | [0.0082] | | | | |
| BG Arson | | | | | | -0.1347 | | 0.0854 |
| | | | | | | [0.2734] | | [0.1988] |
| BG Drug | | | | | | 0.0218 | | |
| | | | | | | [0.0284] | | |
| BG Assault | | | | | | | | 0.0371* |
| | | | | | | | | [0.0173] |
| **Officer Fixed Effect** | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **BG* Year Fixed Effect** | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| N | 250,347 | 250,348 | 292,411 | 292,411 | 143,047 | 143,047 | 372,471 | 372,471 |

Notes: † p<0.10, *p<0.05, ** p<0.01

### b.  *Neighborhood Characteristics*

Another possible explanation for the low correlation between HCA and crime rates is that officers may not be fully aware of local crime rates and instead use neighborhood characteristics as proxies for crime. Table 10 presents a series of models that add block group-level measures of neighborhood demographics on top of the suspect demographics we used in the previous subsection.[139] Models 1 and 2 contain variables measuring the proportion of the residential population that is Black and that is Hispanic. According to these models, moving from a block group without any Black residents to a block group with 100 percent Black residents is associated with an 8 to 9 percent increase in the probability

---

139.    We do not have such data at the precinct level.

that an officer will call the area high crime. However, the coefficient for Hispanic residents is close to zero and not statistically significant.

Models 3 and 4 add a series of block group-level socioeconomic measures. The variable for Black residents falls by almost half but remains statistically significant and the variable for Hispanic residents is now negative and statistically significant. Most of the coefficients on the socioeconomic variables are small and statistically insignificant. Two statistically significant variables are positive: the proportion of residents without a high school degree and the proportion of residents who completed high school.

Models 5 and 6 add fixed effects for officer.[140] The results for the variable measuring the proportion of Black residents remains the same, but the variable for Hispanic residents is now positive and statistically significant. Adding officer fixed effects substantially shrunk most of the socioeconomic variables. The only two statistically significant coefficients are positive: the proportion of residents without a high school degree and the percent of families in poverty.

We next compare the relative importance of these different variables in predicting whether officers invoke HCA. Starting with Model 5, to be conservative we only consider the block group-level measure of violent crime because the precinct-level measure is negative. The coefficient in the model implies that moving from the single safest block group in the city to the single most violent is associated with a 2.2 percent increase in the probability of the officer calling the area high crime. In Model 6, block group-level arson is the only positive coefficient. Moving from a block group with no arson to a block group with the most arson in the city is associated with a 2 percent increase in the probability that the officer invokes HCA.

For point of comparison, in both Models 5 and 6, substituting a white non-Hispanic suspect with a Black non-Hispanic suspect is associated with a 2.1 percent increase in the probability of the officer invoking HCA. This means that the race of the suspect predicts whether the officer will call the area high crime roughly as well as the actual crime rate in the area.

Even more striking, in both Models 5 and 6, moving from a block group with no Black residents to a block group with 100 percent Black residents is associated with a roughly 4.5 percentage point increase in the probability that the officer invokes HCA. In other words, moving from an area with no Black residents to an area with all Black residents is associated with a substantially larger increase in the probability that an officer invokes HCA than moving from the single safest neighborhood in the city to the single most dangerous.

We next consider whether the results vary by the type of suspected crime. Table 11 reproduces Columns 7 and 8 in Table 10 for violence, property, drug, and weapons stops, respectively. There is a large and statistically significant

---

140.    Unlike in prior models, we cannot include fixed effects for block group because the ACS measures are constant throughout all years of the study.

coefficient on the variables measuring the proportion of residents that are Black and that are Hispanic for three of the four crime types: violence, drugs, and weapons. At least partially because of the lower sample sizes, there are no clear patterns to the socioeconomic variables, aside from the finding that the proportion of families in poverty is positively correlated with HCA for violence, property, and weapons stops.

Table 10. Modeling HCA on Neighborhood Characteristics and Crime in the Last 12 Months

| | Mod 1 | Mod 2 | Mod 3 | Mod 4 | Mod 5 | Mod 6 |
|---|---|---|---|---|---|---|
| Intercept | 0.5526** | 0.5092** | 0.5041** | 0.4769** | | |
| | [0.0135] | [0.0155] | [0.0407] | [0.0392] | | |
| **Suspect Characteristics** | | | | | | |
| Black Non-Hispanic | 0.0105 | 0.0119† | 0.0192** | 0.0195** | 0.0214** | 0.0214** |
| | [0.0065] | [0.0066] | [0.0059] | [0.0059] | [0.0021] | [0.0021] |
| Black Hispanic | -0.0144† | -0.0121 | -0.0048 | -0.0039 | 0.0180** | 0.0179** |
| | [0.0082] | [0.0082] | [0.0077] | [0.0077] | [0.0026] | [0.0026] |
| White Hispanic | 0.0017 | 0.0015 | 0.0089 | 0.0077 | 0.0158** | 0.0158** |
| | [0.0062] | [0.0062] | [0.0058] | [0.0057] | [0.0020] | [0.0020] |
| Other Race | -0.0418** | -0.0416** | -0.0384** | -0.0388** | 0.0021 | 0.0021 |
| | [0.0073] | [0.0073] | [0.0072] | [0.0072] | [0.0026] | [0.0026] |
| Age | -0.0009** | -0.0009** | -0.0007** | -0.0007** | -0.0002** | -0.0002** |
| | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0000] | [0.0000] |
| **Area Racial/Ethnic Composition** | | | | | | |
| BG Black Residents | 0.0785** | 0.0883** | 0.0425* | 0.0552** | 0.0465** | 0.0449** |
| | [0.0160] | [0.0153] | [0.0169] | [0.0166] | [0.0048] | [0.0048] |
| BG Hispanic Residents | 0.0012 | 0.0073 | -0.0615** | -0.0537** | 0.0350** | 0.0333** |
| | [0.0160] | [0.0148] | [0.0174] | [0.0165] | [0.0051] | [0.0050] |
| **Area Socioeconomic Composition** | | | | | | |
| BG No HS Graduation | | | 0.1038** | 0.0934** | 0.0161* | 0.0179* |
| | | | [0.0283] | [0.0265] | [0.0070] | [0.0070] |
| BG HS Grad, No College | | | 0.1511** | 0.1364** | 0.0124† | 0.0114 |
| | | | [0.0270] | [0.0268] | [0.0074] | [0.0074] |
| BG Income <$20k | | | 0.0222 | 0.0118 | -0.0026 | -0.0006 |
| | | | [0.0483] | [0.0471] | [0.0168] | [0.0168] |
| BG Income $20k-50k | | | -0.027 | -0.0369 | 0.0134 | 0.0143 |
| | | | [0.0448] | [0.0429] | [0.0144] | [0.0144] |
| BG Income $50k-125k | | | 0.0456 | 0.0327 | 0.0221† | 0.0227† |
| | | | [0.0358] | [0.0345] | [0.0116] | [0.0116] |
| BG Median Income | | | 0 | 0 | 0 | 0 |
| | | | [0.0000] | [0.0000] | [0.0000] | [0.0000] |
| BG Assistance | | | -0.0113 | -0.0251 | 0.0049 | 0.0041 |
| | | | [0.0397] | [0.0370] | [0.0107] | [0.0105] |
| BG Vacant Properties | | | -0.1541** | -0.1321** | -0.0069 | -0.0066 |
| | | | [0.0369] | [0.0340] | [0.0079] | [0.0079] |
| BG Families in Poverty | | | 0.0207 | 0.0248 | 0.0197** | 0.0186** |
| | | | [0.0252] | [0.0234] | [0.0063] | [0.0063] |
| BG Single-headed Household | | | 0.0157 | 0.0164 | 0.0065 | 0.0073† |
| | | | [0.0189] | [0.0166] | [0.0041] | [0.0041] |
| **Crime** | | | | | | |
| PCT Violent | 0.0005 | | 0 | | -0.0006** | |
| | [0.0005] | | [0.0005] | | [0.0002] | |
| BG Violent | 0.0081 | | 0.0024 | | 0.0073** | |
| | [0.0114] | | [0.0112] | | [0.0028] | |
| PCT Non-neg Homicide | | -0.0936† | | -0.1431** | | -0.0156 |
| | | [0.0510] | | [0.0496] | | [0.0207] |
| PCT Burglary | | 0.0165** | | 0.0154** | | -0.0001 |
| | | [0.0034] | | [0.0035] | | [0.0014] |
| BG Arson | | 1.1207** | | 0.8356** | | 0.0877† |
| | | [0.2086] | | [0.2031] | | [0.0499] |
| **Suspected Crime** | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **Officer Fixed Effect** | | | | | ✓ | ✓ |
| **Adjusted R^2 (LPM)** | 0.0141 | 0.0163 | 0.017 | 0.0186 | 0.4031 | 0.4031 |
| **N** | 1,284,370 | 1,284,370 | 1,263,340 | 1,263,340 | 1,263,340 | 1,263,340 |

Notes: † p<0.10, *p<0.05, ** p<0.01

Table 11. Modeling HCA on Neighborhood Characteristics and Crime in the Last 12 Months, by Suspected Crime

| | Violent | | Property | | Drug | | Weapons | |
|---|---|---|---|---|---|---|---|---|
| | Mod 1 | Mod 2 | Mod 3 | Mod 4 | Mod 5 | Mod 6 | Mod 7 | Mod 8 |
| **Suspect Characteristics** | | | | | | | | |
| Black Non-Hispanic | 0.0378** | 0.0383** | 0.0046 | 0.0042 | 0.0216** | 0.0211** | 0.0283** | 0.0280** |
| | [0.0053] | [0.0053] | [0.0031] | [0.0031] | [0.0051] | [0.0051] | [0.0042] | [0.0042] |
| Black Hispanic | 0.0370** | 0.0372** | 0.0075 | 0.0071 | 0.0227** | 0.0225** | 0.0172** | 0.0169** |
| | [0.0062] | [0.0062] | [0.0047] | [0.0047] | [0.0068] | [0.0068] | [0.0048] | [0.0048] |
| White Hispanic | 0.0308** | 0.0310** | 0.0087** | 0.0083** | 0.0184** | 0.0182** | 0.0184** | 0.0182** |
| | [0.0055] | [0.0055] | [0.0029] | [0.0029] | [0.0052] | [0.0052] | [0.0043] | [0.0043] |
| Other Race | 0.0145* | 0.0148* | -0.0023 | -0.0026 | 0.0015 | 0.001 | 0.002 | 0.0017 |
| | [0.0063] | [0.0064] | [0.0040] | [0.0040] | [0.0071] | [0.0071] | [0.0054] | [0.0054] |
| Age | -0.0007** | -0.0007** | 0.0002** | 0.0002** | 0.0000 | 0.0000 | -0.0008** | -0.0008** |
| | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0001] | [0.0001] |
| **Area Racial/Ethnic Composition** | | | | | | | | |
| BG Black Residents | 0.0384** | 0.0395** | 0.0039 | 0.0049 | 0.0763** | 0.0692** | 0.0848** | 0.0815** |
| | [0.0085] | [0.0084] | [0.0077] | [0.0079] | [0.0110] | [0.0107] | [0.0078] | [0.0079] |
| BG Hispanic Residents | 0.0408** | 0.0414** | 0.002 | 0.0001 | 0.0654** | 0.0627** | 0.0656** | 0.0625** |
| | [0.0089] | [0.0088] | [0.0080] | [0.0080] | [0.0125] | [0.0123] | [0.0089] | [0.0088] |
| **Area Socioeconomic Composition** | | | | | | | | |
| BG No HS Graduation | 0.0084 | 0.0119 | 0.0038 | 0.006 | 0.0107 | 0.0063 | 0.0033 | 0.0032 |
| | [0.0132] | [0.0130] | [0.0125] | [0.0132] | [0.0174] | [0.0171] | [0.0107] | [0.0107] |
| BG HS Grad, No College | 0.0102 | 0.0112 | 0.0171 | 0.0211† | -0.0071 | -0.0097 | 0.0103 | 0.0084 |
| | [0.0133] | [0.0132] | [0.0117] | [0.0118] | [0.0174] | [0.0173] | [0.0105] | [0.0105] |
| BG Income <$20k | 0.0096 | 0.0132 | -0.0599* | -0.0599* | 0.0634 | 0.0608 | 0.0017 | 0.0016 |
| | [0.0273] | [0.0276] | [0.0257] | [0.0258] | [0.0416] | [0.0410] | [0.0224] | [0.0223] |
| BG Income $20k-50k | 0.0490* | 0.0500* | -0.0407† | -0.0387† | 0.0777* | 0.0756* | 0.0209 | 0.0202 |
| | [0.0243] | [0.0245] | [0.0224] | [0.0228] | [0.0345] | [0.0341] | [0.0194] | [0.0194] |
| BG Income $50k-125k | 0.0611** | 0.0624** | 0.0094 | 0.0094 | 0.0593* | 0.0606* | 0.0138 | 0.0135 |
| | [0.0201] | [0.0202] | [0.0178] | [0.0181] | [0.0292] | [0.0288] | [0.0178] | [0.0178] |
| BG Median Income | 0.0000* | 0.0000* | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | [0.0000] | [0.0000] | [0.0000] | [0.0000] | [0.0000] | [0.0000] | [0.0000] | [0.0000] |
| BG Assistance | -0.0350† | -0.0306 | -0.0196 | -0.0191 | 0.0242 | 0.0224 | 0.0121 | 0.0099 |
| | [0.0198] | [0.0201] | [0.0185] | [0.0188] | [0.0234] | [0.0235] | [0.0157] | [0.0155] |
| BG Vacant Properties | 0.0064 | 0.0024 | -0.0062 | -0.0054 | 0.0105 | 0.0091 | 0.0006 | 0.0016 |
| | [0.0156] | [0.0154] | [0.0185] | [0.0188] | [0.0234] | [0.0235] | [0.0157] | [0.0155] |
| BG Families in Poverty | 0.0316** | 0.0313** | 0.0362** | 0.0379** | -0.0044 | -0.0037 | 0.0176* | 0.0170* |
| | [0.0120] | [0.0119] | [0.0107] | [0.0109] | [0.0146] | [0.0143] | [0.0086] | [0.0086] |
| BG Single-headed Household | -0.0152† | -0.0137† | -0.0105 | -0.0125† | 0.0258** | 0.0247* | 0.0117† | 0.0118† |
| | [0.0081] | [0.0080] | [0.0071] | [0.0074] | [0.0099] | [0.0097] | [0.0063] | [0.0063] |
| **Crime** | | | | | | | | |
| PCT Violent | 0.0001 | | -0.0004 | | -0.0008† | | -0.0009** | |
| | [0.0004] | | [0.0004] | | [0.0004] | | [0.0003] | |
| BG Violent | 0.0194** | | -0.0004 | | 0.0066 | | 0.0095* | |
| | [0.0054] | | [0.0055] | | [0.0047] | | [0.0045] | |
| PCT Non-neg Homicide | | 0.0489 | | | | | | -0.0312 |
| | | [0.0333] | | | | | | [0.0301] |
| BG Burglary | | -0.0313 | | | | | | |
| | | [0.0321] | | | | | | |
| PCT Motor Vehicle Theft | | 1.1207** | | 0.8356** | | 0.0877† | | |
| | | [0.2086] | | [0.2031] | | [0.0499] | | |
| PCT Burglary | | | | 0.0124** | | | | |
| | | | | [0.0029] | | | | |
| BG Arson | | | | | | -0.1606 | | 0.0319 |
| | | | | | | [0.0980] | | [0.0637] |
| BG Drug | | | | | | 0.0396** | | |
| | | | | | | [0.0074] | | |
| BG Assault | | | | | | | | 0.0148* |
| | | | | | | | | [0.0059] |
| **Officer Fixed Effect** | √ | √ | √ | √ | √ | √ | √ | √ |
| **N** | 242,534 | 242,534 | 278,482 | 278,482 | 139,745 | 139,745 | 364,241 | 364,241 |

Notes: † p<0.10, *p<0.05, ** p<0.01

### c.    *Inter-officer Disparities*

We next test for evidence of inter-officer differences in the perception of whether an area is high crime. To provide a rough picture of inter-officer disparities, Figure 4 depicts a histogram of the rate at which each officer with 50 or more stops invoked HCA. The figure reveals dramatic disparities. Roughly 21 percent of officers invoked HCA in less than 25 percent of stops, while 40 percent invoked HCA over 75 percent of the time.

One conventional measure of inter-rater disparity is the mean absolute deviation—the average distance of each officer's HCA invocation rate from the absolute average for all officers.[141] To compute this number, we regress HCA on a model that contains only fixed effects for officer. We then compute the average distance of each officer's fixed effect from the average HCA invocation rate of all officers. We find that officers' HCA invocation rate is, on average, 27 percentage points away from the absolute average invocation rate of 58 percent.

Of course, at least some of this variation may be explained by the areas in which officers are assigned to patrol. We next regress HCA on fixed effects for both officer *and* census block group. Surprisingly, the absolute mean deviation of the fixed effects for each officer creeps up slightly to 29 percent. We obtain the same result when we also add suspect-level demographic variables.[142] Taken together, these empirical results provide evidence of wide inter-officer disparities in the assessment of whether an area qualifies as high crime.

---

141.    *See, e.g.*, Joel Waldfogel, *Aggregate Inter-Judge Disparity in Federal Sentencing: Evidence from Three Districts (D. Ct., S.D.N.Y., N.D. Cal.)*, 4 FED. SENT'G REP. 151, 152 (1991) (measuring inter-judge sentencing disparities by the mean absolute deviation of each judge's average punishment from the overall average punishment for all judges in a court).

142.    We cannot add ACS variables because we are using fixed effects for block groups and our block group-level ACS variables are constant throughout the study period.

Figure 4. HCA Invocation Rate Among Officer with 50 or More Stops
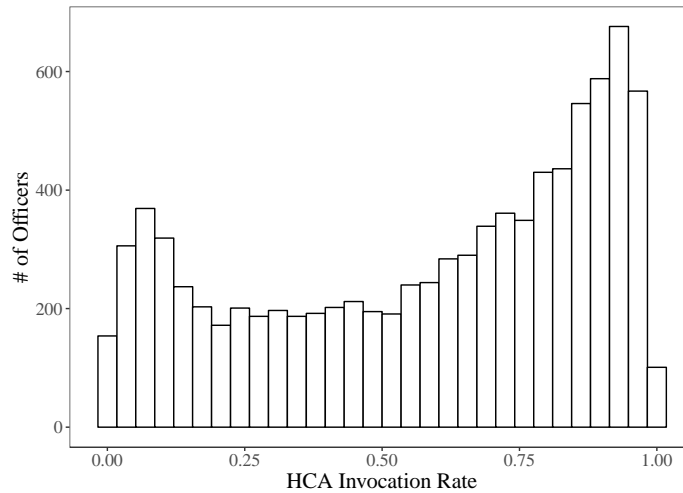


### 3.   *Assumption 3: Do High-Crime Areas Predict Hits?*

We have already found evidence that actual crime levels are poor predictors of whether an officer invokes HCA as a basis of suspicion. That finding suggests that whether an officer calls an area high crime may have little predictive power about whether the suspect is in fact engaged in a crime. We next put that hypothesis to test, to the extent we can, by fitting models to predict whether a stop results in one of three kinds of "hits": (1) arresting a suspect; (2) finding a weapon; (3) finding any other contraband.

One caveat is in order. Our analysis here is necessarily limited because we can only observe the suspects that were stopped; we cannot observe suspects that officers chose not to stop (perhaps because they lacked reasonable suspicion). Still, our results are informative even if they are censored.

Table 12 depicts a series of linear probability models predicting whether an officer arrested a suspect during a stop. Model 1 contains an independent variable indicating whether the officer invoked HCA as a basis of reasonable suspicion and variables indicating the type of suspected crime. The model estimates that when an officer invokes HCA, the officer is 1.8 percentage points less likely to arrest the suspect. That is a 27 percent relative reduction compared to the baseline arrest rate of 6.6 percent. Model 2 adds variables for all other observable bases of suspicion in our data,[143] which cut the HCA coefficient by just over half. Model 3 adds fixed effects for officers, which have little effect on the HCA

---

143.    To see the results for each of the other variables measuring bases of suspicion, see Table A.1 in the Appendix.

coefficient.[144] These results thus suggest that when an officer invokes HCA as a basis of a stop, the stop is less likely to result in an arrest.

Table 12. Predicting Arrest

|  | Mod 1 | Mod 2 | Mod 3 |
|---|---|---|---|
| Intercept | 0.077** | 0.036** |  |
|  | [0.002] | [0.002] |  |
| High Crime | -0.018** | -0.007** | -0.007** |
|  | [0.001] | [0.001] | [0.001] |
| Suspected Crime | ✓ | ✓ | ✓ |
| Other Bases of Suspicion |  | ✓ | ✓ |
| Officer Fixed Effects |  |  | ✓ |
| N | 1,636,384 | 1,636,384 | 1,636,384 |

Notes: † p<0.10, *p<0.05, ** p<0.01

In Table 13, the dependent variable is whether the officer recovered a weapon during a stop. Model 1 estimates that police officers are 0.5 percentage points less likely to recover a weapon—a 42 percent relative reduction against the baseline rate of 1.2 percent. The results are similar when, in Model 2, we add variables for all other observable bases of suspicion in our data.[145] When we add fixed effects for officers in Model 3, the coefficient remains statistically significant but drops by half. The models thus suggest that officers are less likely to recover a weapon when they invoke HCA to justify the stop.

Table 13. Predicting Recovery of a Weapon

|  | Mod 1 | Mod 2 | Mod 3 |
|---|---|---|---|
| Intercept | 0.015** | 0.009** |  |
|  | [0.001] | [0.001] |  |
| High Crime | -0.005** | -0.004** | -0.002** |
|  | [0.001] | [0.001] | [0.000] |
| Suspected Crime | ✓ | ✓ | ✓ |
| Other Bases of Suspicion |  | ✓ | ✓ |
| Officer Fixed Effect |  |  | ✓ |
| N | 1,636,384 | 1,636,384 | 1,636,384 |

Notes: † p<0.10, *p<0.05, ** p<0.01

In Table 10, the dependent variable is whether the officer recovered contraband other than a weapon. Model 1 estimates that the correlation between HCA and the recovery of other contraband is near-zero and statistically

---

144. One other potential "hit" variable is whether the officer issued a summons, which is typically applicable to only very low-level offenses. When we substituted the arrest variable for the summons variable, the results were similar except that for Model 3, the p-value for the HCA variable was just above the 0.10 statistical significance threshold, at 0.13.

145. For the results for the variables measuring other bases of suspicion, see Table A.2 in the Appendix.

insignificant. That result doesn't change when we add variables for all observable bases of suspicion in Model 2,[146] or when we add fixed effects for officer in Model 3. Thus, the models suggest that when an officer invokes HCA to justify the stop, the stop is no more likely to result in the recovery of contraband.

Table 14. Predicting Recovery of Other Contraband

|  | Mod 1 | Mod 2 | Mod 3 |
|---|---|---|---|
| **Intercept** | 0.02** | -0.001 |  |
|  | [0.001] | [0.001] |  |
| **High Crime** | 0 | 0.001 | 0 |
|  | [0.001] | [0.001] | [0.000] |
| **Suspected Crime** | ✓ | ✓ | ✓ |
| **Other Bases of Suspicion** |  | ✓ | ✓ |
| **Officer Fixed Effect** |  |  | ✓ |
| **N** | 1,636,384 | 1,636,384 | 1,636,384 |

Notes: † p<0.10, *p<0.05, ** p<0.01

Taken together, these results have at least two important implications. First, the absence of a positive correlation between HCA and all three hit variables implies that suspects are not more likely to be engaged in a crime when officers invoke HCA as a basis of suspicion to justify a stop. Second, the fact that the correlation between hits and HCA is *negative* for arrests and weapons suggests that officers may be invoking HCA as a basis of suspicion to manufacture the appearance of reasonable suspicion in some of their weakest stops.

## V.

### IMPLICATIONS

Our empirical investigation raises serious questions about whether *Wardlow*'s empirical assumptions are satisfied in practice. Indeed, at least based on administrative data from the NYPD during an era of intensive use of stop and frisk policing, implementation of the high-crime area standard appears haphazard at best, and discriminatory at worst. Officers call nearly every block in the city high crime. Their assessments of high-crime areas are only weakly correlated with actual crime rates. The suspect's race predicts whether an officer deems an area high crime as well as the actual crime rate itself. The racial composition of the area and the identity of the officer are stronger predictors of whether an officer deems an area high crime than the crime rate. And officers may even be using high-crime area as cover to bolster the appearance of constitutional validity in their weakest stops.

---

146.    For the results for the variables measuring other bases of suspicion, see Table A.3 in the Appendix.

Short of reversing *Wardlow*, the courts have a number of tools at their disposal to address some of these problems with the doctrine. Perhaps most simply, they could develop more precise definitions about the geographic scope, temporal horizon, and kinds of crimes relevant in assessing whether an area is high crime. They could also demand more rigorous data in suppression hearings to support an officer's claim that an area is high crime.[147] But we suspect this solution would not go far enough because it would only address the tiny fraction of stops that result in a criminal charge and motion to suppress.[148] A more aggressive judicial approach might prohibit a department from using high-crime areas to justify stops if there is evidence its officers are systematically misapplying the standard.[149] Arguably, police departments that do not faithfully implement the high-crime area standard should not be able to use it to justify stops.

Police departments also have several options to regulate officers' assessments of high-crime areas. As others have noted, one possible solution is for departments to promulgate guidelines officially designating certain areas as "high crime."[150] Under this system, officers could only invoke HCA to justify stops occurring within officially designated zones. Another option is for police supervisors to conduct routine audits of stop forms to provide feedback to officers about how they should apply the high-crime area standard. A department could also release public data to allow for external review.

Police departments can also look to new technological innovations in the field of predictive policing for help. Companies like PredPol and HunchLab[151] have recently developed systems to deliver data to officers' smart phones about crime occurring in the surrounding area in real time.[152] We can see at least two potential benefits. First, departments could use this technology to improve the quality of officers' information about local crime rates and relieve officers of the need to rely on their own subjective and potentially unreliable intuitions. Second, going a step further, this technology could also reduce officers' discretion in deciding whether an area is high crime. HunchLab, for example, uses machine

---

147.    *See* Ferguson & Bernache, *supra* note 4, at 1593.

148.    *See supra* note 18, and accompanying text.

149.    Most simply, departments could provide data on the rate of HCA invocation by crime quantile. But departments could also provide some or all of the validation procedures we have applied in this paper.

150.    *See supra* note 21.

151.    For a more detailed discussion of these companies' technologies and services, see ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING 62–67, 69–72 (2017).

152.    The Philadelphia Police Department has purchased technology from HunchLab. HunchLab, HUNCHLAB: UNDER THE HOOD (2015), https://cdn.azavea.com/pdfs/hunchlab/HunchLab-Under-the-Hood.pdf [https://perma.cc/9L98-SUQ2]; HunchLab, *10-minute Overview*, YOUTUBE (JAN. 13, 2017), https://www.youtube.com/watch?v=WRdcWkH7g0E [https://perma.cc/U2KT-S2NZ]. The Los Angeles Police Department has purchased software from Pred Pol. *See* Nick O'Malley, *To Predict and to Serve: The Future of Law Enforcement,* SYDNEY MORNING HERALD (March 31, 2013), http://www.smh.com.au/world/to-predict-and-to-serve-the-future-of-law-enforcement-20130330-2h0rb.html [https://perma.cc/BMP2-5EBJ].

learning algorithms to map "high-crime areas"—typically no larger than a few city blocks—and then sends that information to officers on patrol.[153] While HunchLab's algorithms currently define high-crime areas based on departmental goals of crime reduction, police departments could develop new algorithms based on the Fourth Amendment's definition of high-crime areas.

In defining high-crime areas empirically, departments must choose their data carefully. First, they should take care to minimize the influence of racial and socioeconomic biases on the construction or definition of high-crime areas. As many others have noted, the output of machine learning algorithms can be affected by biases in the data on which they are trained.[154] For example, if a police department assigns a disproportionate number of officers to patrol communities of color, those communities will contain a disproportionate number of arrests. Defining high-crime areas based on arrest data would then make communities of color appear more dangerous than they are and might also create a kind of high-crime feedback loop.[155] Data on criminal complaints filed by citizens are perhaps a better—albeit imperfect—measure of crime that are less likely to incorporate biases or generate feedback loops.[156]

In addition to the specific proposals we have offered here, the implications of our analysis extend beyond *Wardlow* and the high-crime area standard. Indeed, officers rely on countless other factors in justifying the hundreds of thousands of stops they conduct each year, and officers may very likely be applying some of those factors unfaithfully as well. That's particularly true for softer factors, like suspicious bulges and furtive movements, which officers frequently cite as bases for stops.[157] Yet, we are unaware of any other empirical

---

153.    HunchLab, *supra* note 152.

154.    *See* Aziz Z. Huq, *Racial Equity in Algorithmic Criminal Justice*, 68 DUKE L.J. (forthcoming 2019).

155.    *See, e.g.*, Danielle Ensign et al., *Runaway Feedback Loops in Predictive Policing*, 81 PROCEEDINGS OF MACHINE LEARNING RES. 1 (2018), https://arxiv.org/abs/1706.09847 [https://perma.cc/8HD4-4246] (finding crime data systems can "be susceptible to runaway feedback loops, where police are repeatedly sent back to the same neighborhoods based on prior deployments regardless of the true crime rate").

156.    Still, complaint data is not without its own set of biases. Residents of certain neighborhoods—particularly those that are wealthier and have healthier relationships with local police—may report crimes more frequently, thereby giving the appearance that they have disproportionately more crime relative to neighborhoods with lower reporting rates.

157.    Some courts have already expressed serious doubts about some softer factors, like furtive movements. *See, e.g.*, Floyd v. New York, 959 F. Supp. 2d 668, 679 (S.D.N.Y. 2013) ("'Furtive movements' are an insufficient basis for a stop or frisk if the officer cannot articulate anything more specific about the suspicious nature of the movement."). Other subjective factors may also include suspicious bulges, sights or sounds of crime, or evasive actions. These factors are likely vulnerable to cognitive distortion and bias, especially in the context of race or threatening situations. *See* Jennifer Eberhardt et al., *Seeing Black: Race, Crime, and Visual Processing*, 87 J. PERSONALITY & SOC. PSYCHOL. 876, 880 (2004) (finding that subjects were more likely to perceive a weapon after seeing an image of a person with a darker skin shade); Andrew R. Todd, Kelsey C. Thiem & Rebecca Neel, *Does Seeing Faces of Young Black Boys Facilitate the Identification of Threatening Stimuli?*, 27 PSYCHOL. SCI. 384, 384 (2016) (finding that "participants had less difficulty . . . identifying threatening stimuli and more difficulty identifying nonthreatening stimuli after seeing [images of] Black faces than after seeing

studies that attempt to validate the reliability and predictive validity of the factors that police use to form reasonable suspicion. Our analysis of high-crime areas is therefore just the first step. Empirical legal scholars should begin validating other common bases of reasonable suspicion.

Due to challenges in data availability, that project will be more difficult than our efforts to evaluate the high-crime area standard here. Indeed, unlike crime data, which is collected and published by police departments across the country, data on other common Fourth Amendment factors are not readily available. To gather that data, empirical legal scholars would need to directly observe officers' conduct and the conduct of suspects on the street.

At least two research methodologies can help advance this research agenda. First, the popularization of high-definition body cameras offers a new window to collect data on the process by which officers form reasonable suspicion. Researchers today can systematically code body-camera footage to evaluate the accuracy of police officers' claims that, for example, suspects were engaged in suspicious or evasive movements, were casing a commercial establishment, were concealing contraband, or were engaged in any other actions indicative of crime.

Second, empirical researchers can apply systematic social observation (SSO), a traditional method of data collection in the policing literature.[158] In SSO studies, a neutral observer accompanies an officer on patrol, recording what the officer does and says based on predetermined rules and protocols. SSO is useful because it allows participant observers to directly witness how officers form suspicion and decide whether to stop suspects. That officers would be aware they are being observed, of course, could bias their behavior,[159] but these concerns can be diminished by careful training and monitoring of researcher-police interactions.[160]

At least one study by Geoffrey Alpert, John MacDonald, and Roger Dunham successfully used this approach to study how officers in the Savannah Police Department—an agency with approximately 400 officers—develop

---

White faces"); *see also* Richard R. Johnson & Mark A. Morgan, *Suspicion Formation Among Police Officers: An International Literature Review*, 26 CRIM. JUST. STUD. 99, 100, 107–09 (2013) (discussing how officers use racial characteristics and non-verbal cues in developing suspicion about suspects on the street).

158.    *See* STEPHEN MASTROFSKI ET AL., SYSTEMATIC OBSERVATION OF PUBLIC POLICE: APPLYING FIELD RESEARCH METHODS TO POLICY ISSUES, vii (1998) (describing the methodology in depth); Albert J. Reiss, Jr., *Systematic Observation of Natural Social Phenomena*, 3 SOC. METHODOLOGY 3, 4 (1971) (detailing methods of observation and recording of social interactions *in situ*).

159.    E.M. Hoeben, W. Steenbeek & L.J.R. Pauwels, *Measuring Disorder: Observer Bias in Systematic Social Observation at Streets and Neighborhoods*, 34 J. QUANTITATIVE CRIMINOLOGY 221, 224–27 (2018), https://doi.org/10.1007/s10940-016-9333-6 [https://perma.cc/XUM8-9REY] (noting sources of bias including inter- and intra-observer variation, prior experience with police, and reactivity of the officers under observation).

160.    Reiss, Jr., *supra* note 158, at 3–5, 27–30 (discussing potential challenges of estimating the effects of observation in the measurement of social phenomena).

reasonable suspicion.[161] Trained observers accompanied randomly selected officers on 132 8-hour shifts during a period of 8 months. The observers were trained to "document the police officer's actions and reactions as well as any interactions that occurred with citizens" and "record the sequence in which the events unfolded."[162] They were also trained to "remind and prompt officers to 'think out loud'" when something or someone raised officers' suspicion.[163] And they were also trained to record "when officers seemed to take notice of something and whether they acted on it, and to question the officer about his or her thoughts and feelings about the observation."[164] The most important data that the observers recorded were the bases of suspicion identified by the officers, but they also recorded other relevant variables, including suspect race, the racial composition of the area, and the type of action in which the suspect was engaged.[165] In total, the study observed officers forming suspicion 174 times, which demonstrates the feasibility of using SSO to collect data on the process by which officers form reasonable suspicion.

In future empirical work, legal scholars can use body camera footage or SSO to validate the most common factors invoked by officers to establish reasonable suspicion. They can do so by assessing the factual accuracy of police officers' claims about the presence of these factors and by examining whether they are in fact predictive of criminal behavior. Armed with that information, scholars and courts will have a clearer picture of what Fourth Amendment factors provide a reliable basis for reasonable suspicion and meaningful protection against unreasonable intrusions on personal privacy.

---

161. Alpert, MacDonald & Dunham, *supra* note 111, at 417–19.
162. *Id.* at 418.
163. *Id.*
164. *Id.*
165. *Id.* at 419 (noting that observers reported "behavioral criteria"—which the authors defined as "specific actions by citizens that were either illegal or interpreted . . . as suspicious"—and nonbehavioral criteria—such as the suspect's "appearance, the time and place" and any suspect descriptions provided to the officer).

APPENDIX

Table A.1. Modeling Arrest, Full Results

|  | Mod 1 | Mod 2 | Mod 3 |
|---|---|---|---|
| **Intercept** | 0.077** | 0.036** |  |
|  | [0.002] | [0.002] |  |
| **Circumstances Leading to Stop** |  |  |  |
| High Crime | -0.018** | -0.007** | -0.007** |
|  | [0.001] | [0.001] | [0.001] |
| Criminal Object |  | 0.129** | 0.123** |
|  |  | [0.006] | [0.005] |
| Suspect Description |  | 0.034** | 0.036** |
|  |  | [0.002] | [0.002] |
| Suspicious Bulge |  | 0.004* | 0.015** |
|  |  | [0.002] | [0.001] |
| Casing |  | -0.016** | -0.018** |
|  |  | [0.001] | [0.001] |
| Lookout |  | -0.019** | -0.018** |
|  |  | [0.001] | [0.001] |
| Clothes |  | -0.019** | -0.015** |
|  |  | [0.002] | [0.002] |
| Drug Transaction |  | 0.024** | 0.012** |
|  |  | [0.004] | [0.002] |
| Furtive Movements |  | -0.001 | 0.001 |
|  |  | [0.001] | [0.001] |
| Violent Crime |  | 0.003† | 0.005** |
|  |  | [0.002] | [0.002] |
| Other |  | 0.026** | 0.033** |
|  |  | [0.002] | [0.002] |
| **Additional Circumstances** |  |  |  |
| Witness Report |  | 0.084** | 0.085** |
|  |  | [0.003] | [0.002] |
| Proximity to Crime |  | 0.003* | 0.014** |
|  |  | [0.001] | [0.001] |
| Criminal Associates |  | 0.004* | 0.002 |
|  |  | [0.002] | [0.002] |
| Sights/Sounds of Crime |  | 0.042** | 0.052** |
|  |  | [0.005] | [0.004] |
| Change Direction |  | -0.003* | -0.001 |
|  |  | [0.001] | [0.001] |
| Evasive Behavior |  | 0.027** | 0.029** |
|  |  | [0.002] | [0.001] |
| Ongoing Investigation |  | 0.009** | 0.004** |
|  |  | [0.002] | [0.001] |
| Time |  | -0.005** | -0.002* |
|  |  | [0.001] | [0.001] |
| Other |  | 0.035** | 0.046** |
|  |  | [0.004] | [0.003] |
| **Suspected Crime** | ✓ | ✓ | ✓ |
| **Officer Fixed Effect** |  |  | ✓ |
| **N** | 1,636,384 | 1,636,384 | 1,636,384 |

Table A.2. Modeling Recovery of a Weapon, Full Results

|  | Mod 1 | Mod 2 | Mod 3 |
|---|---|---|---|
| **Intercept** | 0.015** | 0.009** | |
|  | [0.001] | [0.001] | |
| **Circumstances Leading to Stop** | | | |
| High Crime | -0.005** | -0.004** | -0.002** |
|  | [0.001] | [0.001] | [0.000] |
| Criminal Object | | 0.087** | 0.085** |
|  | | [0.004] | [0.004] |
| Suspect Description | | -0.002** | -0.001* |
|  | | [0.001] | [0.000] |
| Suspicious Bulge | | 0.016** | 0.020** |
|  | | [0.002] | [0.002] |
| Casing | | 0 | -0.006** |
|  | | [0.000] | [0.001] |
| Lookout | | -0.001** | -0.001* |
|  | | [0.000] | [0.000] |
| Clothes | | -0.006** | -0.004** |
|  | | [0.001] | [0.001] |
| Drug Transaction | | 0 | 0 |
|  | | [0.000] | [0.000] |
| Furtive Movements | | -0.006** | -0.005** |
|  | | [0.001] | [0.001] |
| Violent Crime | | 0.001* | 0.001† |
|  | | [0.001] | [0.001] |
| Other | | 0.008** | 0.011** |
|  | | [0.001] | [0.001] |
| **Additional Circumstances** | | | |
| Witness Report | | 0.005** | 0.006** |
|  | | [0.001] | [0.001] |
| Proximity to Crime | | -0.002** | 0.001** |
|  | | [0.001] | [0.000] |
| Criminal Associates | | -0.001* | -0.001 |
|  | | [0.001] | [0.001] |
| Sights/Sounds of Crime | | 0.009** | 0.010** |
|  | | [0.002] | [0.002] |
| Change Direction | | -0.001** | -0.001** |
|  | | [0.000] | [0.000] |
| Evasive Behavior | | 0.001† | 0 |
|  | | [0.000] | [0.000] |
| Ongoing Investigation | | 0.003** | 0.003** |
|  | | [0.001] | [0.000] |
| Time | | 0.001 | 0 |
|  | | [0.001] | [0.000] |
| Other | | 0.014** | 0.012** |
|  | | [0.002] | [0.001] |
| **Suspected Crime** | ✓ | ✓ | ✓ |
| **Officer Fixed Effect** | | | ✓ |
| **N** | 1,636,384 | 1,636,384 | 1,636,384 |

Table A.3. Modeling Recovery of Other Contraband, Full Results

| | Mod 1 | Mod 2 | Mod 3 |
|---|---|---|---|
| **Intercept** | 0.02** | -0.001 | |
| | [0.001] | [0.001] | |
| **Circumstances Leading to Stop** | | | |
| High Crime | 0 | 0.001 | 0 |
| | [0.001] | [0.001] | [0.000] |
| Criminal Object | | 0.058** | 0.054** |
| | | [0.003] | [0.003] |
| Suspect Description | | 0.001* | 0.002* |
| | | [0.001] | [0.001] |
| Suspicious Bulge | | 0.003** | 0.007** |
| | | [0.001] | [0.001] |
| Casing | | -0.001 | 0 |
| | | [0.000] | [0.000] |
| Lookout | | -0.006** | -0.005** |
| | | [0.001] | [0.000] |
| Clothes | | -0.006** | -0.004** |
| | | [0.001] | [0.001] |
| Drug Transaction | | 0.020** | 0.013** |
| | | [0.003] | [0.002] |
| Furtive Movements | | 0.003** | 0.005** |
| | | [0.001] | [0.001] |
| Violent Crime | | -0.002** | -0.002* |
| | | [0.001] | [0.001] |
| Other | | 0.011** | 0.012** |
| | | [0.001] | [0.001] |
| **Additional Circumstances** | | | |
| Witness Report | | 0.007** | 0.006** |
| | | [0.001] | [0.001] |
| Proximity to Crime | | 0 | 0.003** |
| | | [0.001] | [0.000] |
| Criminal Associates | | 0.003* | 0 |
| | | [0.001] | [0.001] |
| Sights/Sounds of Crime | | 0.018** | 0.021** |
| | | [0.002] | [0.002] |
| Change Direction | | 0.001† | 0.001† |
| | | [0.001] | [0.000] |
| Evasive Behavior | | 0.012** | 0.012** |
| | | [0.001] | [0.001] |
| Ongoing Investigation | | 0.006** | 0.002** |
| | | [0.001] | [0.001] |
| Time | | -0.001* | 0 |
| | | [0.001] | [0.000] |
| Other | | 0.014** | 0.016** |
| | | [0.002] | [0.001] |
| **Suspected Crime** | ✓ | ✓ | ✓ |
| **Officer Fixed Effect** | | | ✓ |
| **N** | 1,636,384 | 1,636,384 | 1,636,384 |