



Zeitschrift der American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Homepage der Zeitschrift:

<https://www.tandfonline.com/loi/uasa20>

Eine Analyse der "Stop-and-Frisk"-Politik des New York City Police Department im Kontext des Vorwurfs der rassistischen Voreingenommenheit

Andrew Gelman, Jeffrey Fagan & Alex Kiss

Um diesen Artikel zu zitieren: Andrew Gelman, Jeffrey Fagan & Alex Kiss (2007) An Analysis of the New York City Police Department's "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias, Journal of the American Statistical Association, 102:479, 813-823, DOI: [10.1198/016214506000001040](https://doi.org/10.1198/016214506000001040)

Link zu diesem Artikel: <https://doi.org/10.1198/016214506000001040>



Online veröffentlicht: 01 Jan

2012.



Reichen Ihren Artikel für diese Zeitschrift ein

n



Artikelansicht: 10750



Verwandte Artikel anzeigen



Zitierte Artikel: 186 Ansicht der zitierten Artikel

Die vollständigen Zugangs- und Nutzungsbedingungen können unter
<https://www.tandfonline.com/action/journalInformation?journalCode=usa20> eingesehen werden.

Eine Analyse der "Stop-and-Frisk"-Politik des New York City Police Department im Kontext der Behauptungen über rassistische Voreingenommenheit

Andrew GELMAN, Jeffrey FAGAN, und Alex KISS

Jüngste Studien von Polizeidienststellen und Forschern bestätigen, dass die Polizei Angehörige rassistischer und ethnischer Minderheiten im Verhältnis zu ihrem Anteil an der Bevölkerung häufiger anhält als Weiße. Es wurde jedoch argumentiert, dass die Anhaltequoten die von den einzelnen ethnischen Gruppen begangenen Straftaten genauer widerspiegeln oder dass die Anhaltequoten erhöhte Quoten in bestimmten sozialen Bereichen, wie z. B. in Stadtteilen oder Bezirken, widerspiegeln. Die meisten Untersuchungen zu den Anhaltequoten und den Interaktionen zwischen Polizei und Bürgern haben sich auf Verkehrskontrollen konzentriert, und Analysen von Fußgängerkontrollen sind selten. In diesem Artikel werden die Daten von 125.000 Fußgängerkontrollen des New York Police Department über einen Zeitraum von 15 Monaten analysiert. Wir schlüsseln die Kontrollen nach Polizeibezirk auf und vergleichen die Anhalteraten nach Rasse und ethnischer Gruppe, wobei wir frühere rassenspezifische Verhaftungsraten berücksichtigen. Wir verwenden hierarchische Mehrebenenmodelle, um die Variabilität auf Bezirksebene zu berücksichtigen und gehen so direkt auf die Frage der geografischen Heterogenität ein, die sich bei der Analyse von Fußgängerkontrollen stellt. Wir stellen fest, dass Personen afrikanischer und hispanischer Abstammung häufiger angehalten wurden als Weiße, selbst nach Kontrolle der Variabilität auf Bezirksebene und rassenspezifischen Schätzungen der Kriminalitätsbeteiligung.

SCHLÜSSELWÖRTER: Kriminologie; Hierarchisches Modell; Mehrebenenmodell; Überdispersions-Poisson-Regression; Polizeikontrollen; Rassistische Voreingenommenheit.

Ohio 1968). Der Ansatz der Stadt New York

1. VOREINGENOMMENHEIT BEI POLIZEIKONTROLLEN?

In den späten 1990er Jahren wurden in den Vereinigten Staaten in der Bevölkerung, in der Rechtsprechung und in der Politik Besorgnisse über die Schikanen der Polizei gegenüber Minderheitengruppen bei ihren alltäglichen Begegnungen mit den Ordnungskräften laut. Diese Besorgnis konzentrierte sich auf das Ausmaß, in dem die Polizei Menschen auf den Autobahnen anhielt, weil sie "schwarz fahren" (siehe Weitzer 2000; Harris 2002; Lundman und Kaufman 2003). Weitere Bedenken betrafen die rassistische Voreingenommenheit bei Fußgängerkontrollen von Bürgern durch die Polizei, die auf einer "Null-Toleranz"-Politik zur Kontrolle von Straftaten gegen die Lebensqualität und auf Polizeistrategien beruhten, die sich auf Minderheitengemeinschaften konzentrierten und auf illegalen Waffenbesitz und Drogenhandel abzielten (siehe Fagan, Zimring und Kim 1998; Greene 1999; Skolnick und Caplovitz 2001; Fagan und Kim 2001).

Davies 2000, 2003; Fagan 2002; Gould und Mastrofski 2004). Diese Praktiken führten zu wütenden Reaktionen unter den ~~Bigen~~ die einer Minderheit angehören, und vergrößerten die Kluft zwischen den verschiedenen rassistischen/ethnischen Gruppen in Bezug auf ihr Vertrauen in die Polizei (Lundman und Kaufman 2003; Tyler und Huo 2003; Weitzer und Tuch 2002), was eine Legitimitätskrise mit rechtlichen, moralischen und politischen Dimensionen auslöste (siehe Wang 2001; Russell 2002; Harris 2002).

In einer Zeit sinkender Kriminalitätsraten drehen sich politische Debatten über Polizeistrategien oft um die Bewertung der Polizeistrategie von New York City in den 1990er Jahren, einer Strategie, bei der Fußgänger wegen einer Vielzahl von Straftaten in großem Umfang angehalten und durchsucht wurden (Eck und Maguire 2000; Skogan und Frydl 2004). Diese ~~Rk~~ basierte auf der rechtmäßigen Praxis, "Zivilpersonen auf der Straße vorübergehend festzuhalten, zu befragen und manchmal auch zu durchsuchen" (Spitzer 1999). Der Oberste Gerichtshof der USA hat entschieden, dass polizeiliche Anhalte- und Durchsuchungsverfahren unter bestimmten Bedingungen verfassungskonform sind (Terry v.

Andrew Gelman ist Professor am Institut für Statistik und am Institut für Politikwissenschaft der Columbia University, New York, NY 10027 (E-Mail: gelman@stat.columbia.edu). Jeffrey Fagan ist Professor an der Law School und der School of Public Health der Columbia University, New York, NY 10027 (E-Mail: jfagan@law.columbia.edu). Alex Kiss ist Biostatistiker, Abteilung für Research Design und Biostatistik, Sunnybrook and Women's College Health Sciences Center, Toronto, Ontario, Kanada. Die Autoren danken dem New York City Police Department, der New York State Division of Criminal Justice Services und dem Office of the New York State Attorney General für die Bereitstellung von Daten für diese Untersuchung. Tamara Dumanovsky und Dong Xu leisteten wichtige Beiträge zur Analyse. Joe Bafumi, Rajeev Dehejia, Jim Liebman, Dan Rabinowitz, Caroline Rosenthal Gelman und mehrere Gutachter gaben hilfreiche Kommentare ab. Diese Forschungsarbeit wurde zum Teil durch die Zuschüsse SES-9987748 und SES-0318115 der National Science Foundation unterstützt. Alle Meinungen sind die der Autoren.

Police Department (NYPD) in den 1990er Jahren wird weithin als eine der Hauptursachen für den starken Rückgang der Kriminalität in der Stadt angesehen (Zimring 2006).

Doch gegen Ende des Jahrzehnts gab es wiederholt Beschwerden über Schikanen gegenüber Minderheiten, insbesondere durch die Eliteeinheit für Straßenkriminalität (Spitzer 1999). Diese Beschwerden standen im Zusammenhang mit dem viel beachteten Überfall auf Abner Louima und den Erschießungen von Amadou Diallo und Patrick Dorismond durch die Polizei. Bürgerbeschwerden über aggressive "Stop and Frisk"-Taktiken führten schließlich zu einem Zivilprozess, in dem rassistische Voreingenommenheit in den Mustern von "Stop and Frisk" vermutet wurde, was zu einem Vergleich führte, der den Einsatz dieser Taktik regelte und umfassende Überwachungsanforderungen festlegte (Kelvin Daniels et al. v. City of New York 2004).

Wir befassen uns mit diesem Streit, indem wir das Ausmaß der rassendiskriminierenden Auswirkungen der als "New Yorker Strategie" bekannt gewordenen Maßnahme abschätzen. Wir analysieren die Raten, mit denen New Yorker unterschiedlicher ethnischer Gruppen von der Polizei auf den Straßen der Stadt angehalten wurden, um die zentrale Behauptung zu bewerten, dass rassenspezifische Anhalteraten nichts anderes widerspiegeln als rassenspezifische Kriminalitätsraten. Diese Studie basiert auf Arbeiten, die in Zusammenarbeit mit dem New York State Attorney General's Office (Spitzer

1999) durchgeführt und von der U.S. Commission on Civil Rights (2000) überprüft wurden. Die wichtigsten statistischen Aspekte sind die für den Vergleich der Raten verwendeten Basisdaten (ein Problem, das von Miller 2000, Walker 2001, Smith und Alpert 2002 erkannt wurde) und die lokalen Unterschiede in der Intensität der Polizeiarbeit, wie sie von der Street Crimes Unit durchgeführt und von Wilson und Kelling (1982) und anderen implizit empfohlen werden. Wir verwenden eine mehrstufige Modellierung (siehe Raudenbush und Bryk 2002 für einen Überblick und Sampson, Raudenbush und Earls 1997; Sampson und Raudenbush 1999; Weidner, Frase und Par-doe 2004 für Beispiele in Studien zur Kriminalität), um lokale Unterschiede beim Vergleich der Raten von Polizeikontrollen verschiedener ethnischer Gruppen in New York City zu berücksichtigen.

Wurden ethnische Minderheiten von der Polizei überproportional häufig angehalten? Wir gehen dieser Frage auf verschiedene Weise nach, indem wir Daten über Polizeikontrollen verwenden, und kommen zu dem Schluss, dass Angehörige von Minderheitengruppen häufiger angehalten wurden als Weiße, und zwar sowohl im Vergleich zu ihrer Gesamtbevölkerung als auch zu den geschätzten Raten von

© 2007 American Statistical
Association Zeitschrift der American
Statistical Association
September 2007, Vol. 102, Nr. 479, Anwendungen und Fallstudien
DOI 10.1198/016214506000001040

Verbrechen, die sie begangen haben. Wir schließen jedoch nicht unbedingt auf diskriminierende Praktiken der NYPD. Die zusammenfassenden Statistiken, die wir hier untersuchen, können nicht direkt auf Fragen der Belästigung oder Diskriminierung eingehen, sondern zeigen vielmehr statistische Muster auf, die für diese Fragen relevant sind.

Da es sich hierbei um ein kontroverses Thema handelt, das auf verschiedene Weise untersucht wurde, gehen wir in den Abschnitten 2 und 3 ausführlich auf den historischen Hintergrund und die verfügbaren Daten ein. In den Abschnitten 4 und 5 stellen wir unsere Modelle und Ergebnisse vor, und in Abschnitt 6 folgt eine Diskussion.

2. HINTERGRUND

2.1 Rasse, Nachbarschaft und Polizeikontrollen

Nahezu ein Jahrhundert rechtlicher und gesellschaftlicher Entwicklungen hat die Bühne für die aktuelle Debatte über Rasse und Polizeiarbeit bereitet. In der Vergangenheit gehörte die enge Überwachung durch die Polizei zum Alltag von Afroamerikanern und anderen Minderheitengruppen (siehe z. B. Musto 1973; Kennedy 1997). In jüngerer Zeit hat der Oberste Gerichtshof der USA in der Rechtssache *Whren et al. gegen die USA* (1996) die Verwendung der Rasse als Grundlage für eine polizeiliche Kontrolle zugelassen, sofern andere Faktoren für die Kontrolle ausschlaggebend waren. In *Brown v. Oneonta* (2000) ließ ein Bundesbezirksgericht die Verwendung der Rasse als Durchsuchungskriterium zu, wenn eine ausdrückliche rassistische Beschreibung des Verdächtigen vorlag.

Der rechtliche Standard für polizeiliches Verhalten bei Bürgerkontrollen geht auf das Urteil *Terry gegen Ohio* (1968) zurück, in dem es um eine Fußgängerkontrolle ging und das die Parameter für den Standard des "begründeten Verdachts" für polizeiliches Verhalten bei der Festnahme von Bürgern zum Zwecke der Durchsuchung oder Verhaftung festlegte. In jüngster Zeit haben die Gerichte das Konzept des "begründeten Verdachts" auf den Ort und das Verhalten ausgeweitet. So stellte der Oberste Gerichtshof der Vereinigten Staaten in der Rechtssache *Illinois gegen Wardlow* (2000) fest, dass die Anwesenheit einer Person in einem "Gebiet mit hoher Kriminalität" zwar nicht den Standard für einen konkreten Verdacht auf kriminelle Aktivitäten erfüllt, die Merkmale eines Ortes jedoch relevant sind, um festzustellen, ob ein Verhalten hinreichend verdächtig ist, um weitere Untersuchungen zu rechtfertigen. Da in "Gegenden mit hoher Kriminalität" oft viele Minderheiten leben (Massey und Denton 1993), birgt diese Logik die Gefahr, dass die Verdachtsmomente der Bewohner von Minderheitenvierteln erhöht werden.

Frühe Studien legen nahe, dass sowohl die rassischen Merkmale des Verdächtigen als auch die rassische Zusammensetzung der Nachbarschaft des Verdächtigen die Entscheidung der Polizei beeinflussen, einen Verdächtigen anzuhalten, zu durchsuchen oder festzunehmen (Bittner 1970; Reiss 1971). Insbesondere in städtischen Gebieten interagiert die Rasse des Verdächtigen mit den Merkmalen der Nachbarschaft, um die Verdachtsbildung der Polizeibeamten zu fördern (Thompson 1999; Smith, Makarios und Alpert 2006). Alpert, MacDonald und Dunham (2005) fanden heraus, dass die Polizei einen Bürger aus einer Minderheit mit

größerer Wahrscheinlichkeit als verdächtig einstuft, was zu einer Polizeikontrolle führt, die auf nicht-verhaltensbezogenen Hinweisen beruht, während sie sich bei weißen Bürgern häufiger auf verhaltensbezogene Hinweise stützt, um einen Verdacht zu entwickeln.

Die Polizei kann aber auch rassische Merkmale von Gemeinden durch rassische Merkmale von Einzelpersonen in ihrem kognitiven Verdachtsschema ersetzen, was zu erhöhten Anhaltequoten in Stadtvierteln mit hoher Konzentration von Minderheiten führt. So zeigte Smith (1986) in einer Studie über die Polizeiarbeit in drei Städten, dass Verdächtige in armen Stadtvierteln mit größerer Wahrscheinlichkeit festgenommen wurden, wobei das Verhalten des Verdächtigen und die Art des Verbrechens berücksichtigt wurden. Die Rasse des Verdächtigen und die rassische Zusammensetzung der Nachbarschaft des Verdächtigen waren ebenfalls signifikante Prädiktoren für die Reaktion der Polizei. Zwanghafte polizeiliche Reaktionen können mit

Gelman, Fagan und Kiss: "Stop-and-Frisk"-Politik auf die Wahrnehmung, dass arme Stadtteile nur über eine begrenzte Fähigkeit zur sozialen Kontrolle und Selbstregulierung verfügen. Diese Strategie wurde in dem einflussreichen "broken windows"-Aufsatz von Wilson und Kelling (1982) formalisiert, die argumentierten, dass polizeiliche Reaktionen auf Unordnung entscheidend seien, um Intoleranz gegenüber Verbrechen zu vermitteln und deren ansteckende Ausbreitung zu stoppen. Andere haben diese Behauptung jedoch angezweifelt (siehe Harcourt 1998, 2001; Sampson und Raudenbush 1999; Taylor 2000) und argumentiert, dass die Rasse oft als Ersatz für die Nachbarschaftsbedingungen verwendet wird, um den Verdacht der Polizei zu lenken.

Die Polizei hat rassistisch ungleiche Muster von Kontrollen mit der Begründung verteidigt, dass Minderheiten unverhältnismäßig mehr Straftaten begehen als Weiße (insbesondere die Arten von Straftaten, die die Aufmerksamkeit der Polizei auf sich ziehen), und dass die räumliche Häufung und die ungleichen Auswirkungen von Straftaten, die von und gegen Minderheiten begangen werden, eine aggressivere Durchsetzung in Minderheitengemeinschaften rechtfertigen (MacDonald 2001). Die Polizei beruft sich auf solche Unterschiede in den Verbrechensraten, um rassistische Ungleichgewichte selbst in Situationen zu rechtfertigen, in denen sie eine breite Palette möglicher Ziele hat oder in denen der Verdacht auf kriminelle Aktivitäten ansonsten eine Kontrolle oder Durchsuchung nicht rechtfertigen würde (Kennedy 1997; Harcourt 2001; Rudovsky 2001). Mit dieser Logik behauptet die Polizei, dass die höheren Anhaltequoten bei Afroamerikanern und anderen Minderheiten einfach eine vernünftige und effiziente Polizeipraxis darstellen (siehe z. B. Bratton und Knobler 1998; Goldberg 1999). Die Polizei verweist oft auf die hohen Beschlagnahmungsraten von Schmuggelware, Waffen und Flüchtigen bei solchen Kontrollen und auf eine Verringerung der Kriminalität, um diese aggressive Polizeipraxis zu rechtfertigen (Kelling und Cole 1996).

Ob rassistisch ungleiche Anhaltequoten ein Zeichen für Unverhältnismäßigkeit sind.

Der Kern der gesellschaftlichen und rechtlichen Kontroverse über Racial Profiling und Rassendiskriminierung durch die Polizei (Fagan 2002; Ayres 2002a; Harris 2002) ist die Frage, ob die Kriminalitätsrate oder die absichtliche, rassistisch motivierte Ausrichtung der Polizei auf Minderheiten über das hinausgeht, was rassistische Unterschiede in der Kriminalitätsrate rechtfertigen könnten. Diese Kontroverse stand im Mittelpunkt öffentlicher und privater Rechtsstreitigkeiten (Rudovsky 2001), politischer Mobilisierung und der Selbstkritik mehrerer Polizeidienststellen (siehe Garrett 2001; Walker 2001; Skolnick und Caplovitz 2001; Gross und Livingston 2002).

2.2 Ansätze zur Untersuchung von Daten über Polizeikontrollen

Jüngste Erkenntnisse bestätigen die Wahrnehmung von Bürgern, die einer Minderheit angehören, dass die Polizei afroamerikanische und hispanische Autofahrer überproportional häufig anhält und dass diese Bürger, wenn sie angehalten werden, mit größerer Wahrscheinlichkeit durchsucht oder verhaftet werden (Cole 1999; Veneiro und Zoubeck 1999; Harris 1999; Zingraff et al. 2000; Gross und Barnes 2002). Zwei Erhebungen mit landesweiten

Stichproben, die 1999 und 2002 abgeschlossen wurden, zeigten beispielsweise, dass Afroamerikaner weitaus häufiger als andere angehalten zu werden (Langan, Greenfeld, Smith, Durose und Levin 2001; Durose, Schmitt und Langan 2005). Beide Erhebungen zeigten, dass Fahrer, die einer Minderheit angehören, auch häufiger angaben, von der Polizei mit einem Strafzettel belegt, verhaftet, in Handschellen abgeführt oder durchsucht zu werden, und dass ihnen häufiger Gewalt angedroht wurde oder Gewalt gegen sie angewendet wurde. Diese Ungleichheiten verursachen soziale Kosten, die nach Loury (2002) kulturell bedeutsame Formen der Stigmatisierung hervorrufen, die rassistische Ungleichheiten verstärken, insbesondere in der Praxis der Strafverfolgung.

"Verdächtiges Verhalten" ist der Auslöser sowohl für Fußgänger als auch für

Verkehrskontrollen (Alpert et al. 2005). Fußgängerkontrollen sind auf der

Sie werden eingesetzt, um Drogen- und Waffengesetze durchzusetzen, flüchtige Personen oder andere Personen, gegen die ein Haftbefehl vorliegt, zu identifizieren, gemeldete Straftaten und "verdächtiges" Verhalten zu untersuchen und die Lebensqualität der Gemeinschaft zu verbessern. Für die NYPD bietet ein "Stop"-Einsatz der Polizei die Gelegenheit, mit Personen in Kontakt zu treten, die vermutlich in geringfügige Kriminalität verwickelt sind, ohne eine formelle Verhaftung vornehmen zu müssen, und zwar unter dem niedrigeren verfassungsrechtlichen Standard des "angemessenen Verdachts" (Spitzer 1999). Da Straftaten von geringer Bedeutung für die Lebensqualität und Ordnungswidrigkeiten eher in der Öffentlichkeit begangen werden, ist der Standard des "begründeten Verdachts" bei dieser Art von Straftaten leichter zu erfüllen (Rudovsky 2001).

Bei Fußgänger- und Verkehrsverstößen ist das Spektrum verdächtiger Verhaltensweisen bei der Polizeiarbeit in der Nachbarschaft jedoch so breit, dass es schwierig ist, eine geeignete Basislinie zu ermitteln, mit der sich rassenspezifische Anhalteraten vergleichen lassen (siehe Miller 2000; Smith und Alpert 2002; Gould und Mastrofski 2004). Dementsprechend ist es schwierig, Voreingenommenheit zuzuordnen; kausale Behauptungen über Diskriminierung würden weitaus mehr Informationen über solche Grundlinien erfordern, als die typischen administrativen (Beobachtungs-)Datensätze liefern können. Untersuchungen vor Ort, die sich auf die direkte Beobachtung von Polizeiverhalten stützen (z. B. Gould und Mastrofski 2004; Alpert et al. 2005), erfordern, dass die Beamten die Gründe für ihre Handlungen darlegen, eine Aufgabe, die für zahlreiche Validitätsrisiken anfällig ist. Ein zuverlässiger Nachweis ethnischer Voreingenommenheit würde stattdessen ein Versuchsdesign erfordern, bei dem andere Faktoren kontrolliert werden, um Unterschiede in den Ergebnissen, die nur auf die Rasse oder die ethnische Zugehörigkeit zurückgeführt werden können, zu isolieren. Solche Experimente werden routinemäßig bei Tests zur Diskriminierung im Wohnungs- und Beschäftigungsbereich eingesetzt (siehe z. B. Pager 2003). Beobachtungsstudien, die nicht über solche Kontrollen verfügen, sind jedoch häufig mit Verzerrungen durch ausgelassene Variablen behaftet; nur wenige Studien können alle Variablen berücksichtigen, die die Polizei bei der Entscheidung, ob sie jemanden anhält oder durchsucht, in Betracht zieht.

Ein anderer Ansatz zur Untersuchung der Rassenunterschiede umgeht die Frage, ob die Polizei beabsichtigt, aufgrund der ethnischen Zugehörigkeit oder der Rasse zu diskriminieren, und konzentriert sich stattdessen auf die unterschiedlichen Auswirkungen der polizeilichen Kontrollstrategien. Bei diesem Ansatz dienen Vergleiche der "Trefferquoten" oder der Effizienz des Anteils der Kontrollen, die zu positiven Ergebnissen führen, als Beweis für die ungleichen Auswirkungen von Polizeikontrollen. Dieser Ansatz kann zeigen, wann die Unverhältnismäßigkeit einer bestimmten Politik oder eines bestimmten Entscheidungsergebnisses nicht durch eine erhöhte institutionelle Produktivität gerechtfertigt ist. Im Zusammenhang mit der Profilerstellung gehen Ergebnistests davon aus, dass die Ex-post-Wahrscheinlichkeit, dass bei

einer polizeilichen Durchsuchung Drogen oder andere Schmuggelware entdeckt wird, vom Grad des wahrscheinlichen Grundes abhängt, den die Polizei bei der Entscheidung, einen Verdächtigen anzuhalten und zu durchsuchen, anwendet (Ayres 2002a). Die Feststellung, dass die Durchsuchung von Minderheiten weniger ergiebig ist als die Durchsuchung von Weißen, könnte ein Beweis dafür sein, dass die Polizei bei der Durchsuchung von Minderheiten eine niedrigere Schwelle für einen wahrscheinlichen Grund ansetzt. Zumindest ist dies ein Anzeichen für eine unterschiedliche Behandlung von Minderheiten, die wiederum zu unterschiedlichen Auswirkungen führt.

Knowles, Persico und Todd (2001) untersuchten diesen Ansatz der "Trefferquote" sowohl theoretisch als auch empirisch in einer Studie, in der sie feststellten, dass von den Fahrern auf der I-95 in Maryland, die von der Polizei wegen des Verdachts auf Drogenhandel angehalten wurden, Afroamerikaner genauso häufig Drogen in ihrem Auto hatten wie Weiße. Die begleitende theoretische Analyse geht von einem dynamischen Prozess aus, der die Verhaltensweisen sowohl der Polizei als auch der Bürger verschiedener Rassen berücksichtigt und ihre Entscheidungen in ein Gleichgewicht einbezieht, in dem die Polizei ihre

Gelman, Fagan und Kiss: "Stop-and-Frisk"-Politik
Verhalten an die Wahrscheinlichkeit der Entdeckung illegalen Verhaltens anpasst und Bürger verschiedener Rassengruppen ihre Neigungen anpassen, um die Wahrscheinlichkeit der Entdeckung zu erhöhen. Sie kamen zu dem Schluss, dass die Suche nach Drogen eine effiziente Zuweisung von Polizeiressourcen darstellt, trotz der ungleichen Auswirkungen dieser Kontrollen auf Bürger, die einer Minderheit angehören (Lamberth 1997; Ayres 2002a,b; Gross und Barnes 2002).

Diese Analyse lässt jedoch mehrere Faktoren außer Acht, die diese Behauptungen verfälschen könnten, wie z. B. rassische Unterschiede bei den Merkmalen, die die Polizei bei der Entscheidung berücksichtigt, welche Autofahrer sie anhält, durchsucht oder festnimmt (siehe z. B. Alpert et al. 2005; Smith et al. 2006). Darüber hinaus führen die Annahmen des Zufallsgleichgewichts im Ansatz von Persico et al. - dass sowohl die Polizei als auch potenzielle Straftäter ihr Verhalten als Reaktion auf die gemeinsamen Wahrscheinlichkeiten des Mitführens von Schmuggelware und des Anhaltens anpassen - dazu, dass sowohl bei der polizeilichen Entscheidungsfindung als auch bei der Kriminalitätsneigung der Betroffenen ein Durchschnitt über heterogene Bedingungen gebildet wird (Dharmapala und Ross 2004), und dass die Auswirkungen rassenspezifischer Empfindlichkeiten in Bezug auf Verbrechensentscheidungen unter unterschiedlichen Bedingungen des Entdeckungsrisikos bei Polizeikontrollen unberücksichtigt bleiben (Dominitz und Knowles 2005). Im Hinblick auf diese beiden Probleme haben Dharmapala und Ross (2004) verschiedene Gleichgewichte identifiziert, die zu unterschiedlichen Schlussfolgerungen in Bezug auf rassenspezifische Vorurteile bei Polizeikontrollen und -durchsuchungen führen.

Wir befassen uns kurz mit den Trefferquoten (siehe Abschnitt 5.3), aber unsere Hauptanalyse versucht, diese angebotsseitigen Probleme oder Probleme mit ausgelassenen Variablen zu lösen, indem wir die rassenspezifischen Raten der angestrebten Verhaltensweisen in den patrouillierten Gebieten kontrollieren und bewerten, ob die Anhalte- und Durchsuchungsquoten über dem liegen, was wir aufgrund der Kenntnis der Kriminalitätsraten der verschiedenen Rassengruppen vorhersagen würden. Bei diesem Ansatz wird das Anhalteverhalten an Beobachtungen über die Wahrscheinlichkeit von Verbrechen oder Schuld bei verschiedenen Rassengruppen gekoppelt. Durch die Aufschlüsselung der Daten nach Stadtteilen berücksichtigen unsere Wahrscheinlichkeitsschätzungen außerdem ausdrücklich die externen Effekte von Nachbarschaft und Rasse, die in der Vergangenheit bei der Polizeiarbeit beobachtet wurden (Skogan und Frydl 2004). Dieser Ansatz erfordert Schätzungen des Angebots an Individuen, die sich in den angestrebten Verhaltensweisen engagieren (siehe Miller 2000; Fagan und Davies 2000; Walker 2001; Smith und Alpert 2002).

Die Feststellung, dass die Polizei Minderheiten in höherem Maße anhält und durchsucht, als dies durch ihre Beteiligung an Straftaten gerechtfertigt ist, erfordert zwar nicht die Schlussfolgerung, dass die Polizei zumindest eine ungleiche Behandlung vorgenommen hat, sie liefert jedoch den Beweis dafür, dass die von der Polizei verwendeten Kriterien zu einer nicht zu rechtfertigenden ungleichen

Wirkung geführt haben.

817

3. DATEN

3.1 "Stop and Frisk" in New York City

Die NYPD führt Aufzeichnungen über Kontrollen (auf "UF-250-Formularen"). Diese Informationen wurden für alle Kontrollen (insgesamt etwa 175.000) von Januar 1998 bis März 1999 gesammelt (Spitzer 1999). Die Polizei ist nicht verpflichtet, für jede Kontrolle ein Formular auszufüllen. Vielmehr gibt es bestimmte Bedingungen, unter denen die Polizei zum Ausfüllen des Formulars verpflichtet ist. Diese "obligatorischen Kontrollen" machen 72 % der erfassten Kontrollen aus, die übrigen Berichte beziehen sich auf Kontrollen, bei denen die Berichterstattung fakultativ war. Um Bedenken hinsichtlich einer möglichen Selektionsverzerrung bei den nicht obligatorischen Kontrollen auszuräumen, haben wir unsere Hauptanalysen (siehe Abb. 2) nur für die obligatorischen Kontrollen wiederholt; die Gesamtraten der Kontrollen haben sich geändert, aber die relativen Raten für die verschiedenen ethnischen Gruppen blieben im Wesentlichen unverändert.

Das Formular UF-250 enthält ein Feld, in dem der Polizeibeamte die "Faktoren, die den Beamten dazu veranlassten, den angehaltenen Sohn zu verdächtigen (einschließlich Informationen von dritten Personen und deren Identität, falls bekannt)" eintragen kann. Wir untersuchten diese Formulare und die Gründe für die Anhaltungen für eine stadtweite Stichprobe von 5.000 Fällen sowie 10.869 weiteren Fällen, die 50 % der Fälle in einer nicht zufälligen Stichprobe von 8 der 75 Polizeibezirke repräsentieren, die so ausgewählt wurden, dass sie ein Spektrum von rassischen Bevölkerungsmerkmalen, Kriminalitätsproblemen und Anhalteraten repräsentieren und sich an den politischen Fragen der ursprünglichen Studie orientieren (Spitzer 1999, S. 158). Die folgenden Beispiele (aus Spitzer 1999) veranschaulichen die Regeln, nach denen die Polizei entscheidet, Verdächtige anzuhalten, und zeigen die sozialen und verhaltensbezogenen Faktoren, die die Polizei bei der Bildung eines angemessenen Verdachts anwendet:

- "Bei TPO [Zeit und Ort des Geschehens] war ein Mann bei einer Person, die der Beschreibung einer Person entsprach, die für GLA [schwerer Autodiebstahl] im 072 pct. log gesucht wurde. . bei der Annäherung warf der Mann eine kleine Münzrolle weg, die 5 Beutel mit al-legiertem Crack enthielt."
- "Am T/P/O beobachtete R/O [Meldebeamter] die unten genannte Person zusammen mit drei anderen, die in zahlreiche geparkte Fahrzeuge schauten. R / O hat Überwachung auf indi- viduals für ca. 20 min halten. Die Personen wurden daraufhin angehalten und befragt [sic], mit negativem Ergebnis."
- "Aufschlitzen in der Canal Street; Person passt auf Beschreibung; Person lief weg."
- "Mehrere Männer, die mehrmals in ein Fahrzeug ein- und aussteigen." "Def. Hatte einen großen Mantel mit einer Ausbuchtung in der rechten Tasche an."
- "Die angehaltene Person hörte auf zu gehen und kehrte die Richtung um, als sie die Polizei sah. Versuchte, den Laden zu betreten, als sich die Polizei näherte; wurde zur Sicherheit gefilzt."

Auf der Grundlage von Bundes- und Landesrecht sind einige dieser Gründe für die Anhaltung einer Person verfassungsgemäß, andere nicht. So haben Gerichte beispielsweise entschieden, dass eine Ausbeulung in der Hosentasche kein ausreichender Grund für die Polizei ist, eine Person ohne deren Zustimmung anzuhalten (People v. DeBour 1976; People v. Holmes 1996), und dass das Weglaufen vor der Polizei kein ausreichender Grund ist, um eine Person anzuhalten und zu filzen (Brown v. Texas 1979; siehe aber Illinois v. Wardlow 2000). Wenn die Polizei jedoch illegale Handlungen, Waffen (einschließlich "Bauchbinden"), eine Person, die auf eine Beschreibung passt, oder verdächtiges Verhalten in einem Kriminalitätsbereich beobachtet, wurden Anhalten und Durchsuchen für verfassungsgemäß erklärt (Spitzer 1999).

Das Büro des Generalstaatsanwalts des Bundesstaates New York nutzte Regeln wie diese, um die Begründungen für 61 % der Kontrollen in der Stichprobe als "begründeten Verdacht" zu charakterisieren, der eine rechtmäßige Kontrolle

rechtfertigen würde, 15 % der Kontrollen als nicht begründeten Verdacht und 24 % als unzureichende Informationen, auf die eine Entscheidung gestützt werden kann. Bei der umstrittenen Einheit für Straßenkriminalität wurde bei 23 % der Kontrollen kein hinreichender Verdacht geäußert. (Hier gab es kein ausgeprägtes Muster nach ethnischer Zugehörigkeit; der Anteil der als unangemessen bewerteten Kontrollen war für alle ethnischen Gruppen in etwa gleich). Die als "ohne hinreichenden Verdacht" eingestuft Kontrollen schienen in der Tat schwächer zu sein, da nur 1 von 29 dieser Kontrollen zu Verhaftungen führte, verglichen mit 1 von 7 der Kontrollen mit hinreichendem Verdacht.

3.2 Aggregierte Anhaltequoten für jede ethnische Gruppe

Vor diesem Hintergrund analysieren wir den gesamten Datensatz der "Stop-and-frisk"-Kontrollen, um festzustellen, in welchem Umfang verschiedene ethnische Gruppen von der Polizei angehalten wurden. Wir konzentrieren uns auf Schwarze (Afroamerikaner), Hispanoamerikaner (Latinos) und Weiße (europäische Amerikaner). Die Kategorien entsprechen den Angaben der Polizei, die die Kontrollen durchführt. Wir schließen Angehörige anderer ethnischer Gruppen (ca. 4 % der Kontrollen) wegen der Wahrscheinlichkeit von Unklarheiten bei den Klassifizierungen aus. Bei einer so geringen Häufigkeit von "anderen" kann selbst eine geringe Rate von Fehlklassifizierungen zu großen Verzerrungen bei den Schätzungen für diese Gruppe führen. Wenn beispielsweise nur 4 % der Schwarzen, Hispanoamerikaner und Weißen fälschlicherweise als "Sonstige" eingestuft würden, würde dies die Schätzungen für die Kategorie "Sonstige" fast verdoppeln, während die Auswirkungen auf die drei Hauptgruppen sehr gering wären. (Siehe Hemenway 1997 für eine ausführliche Erörterung der Probleme, die Fehleinstufungen bei Schätzungen für einen kleinen Teil der Bevölkerung verursachen können). Um einen Eindruck von den Daten zu vermitteln, zeigt Abbildung 1 die Anzahl der Anschläge für Schwarze, Hispanoamerikaner und Weiße im Laufe des 15-monatigen Zeitraums, wobei die Anschläge im Zusammenhang mit jeder der vier Arten von Straftaten ("mutmaßliche Anschuldigungen", wie im UF-250-Formular beschrieben) getrennt dargestellt werden: Gewaltverbrechen, Waffendelikte, Eigentumsdelikte und Drogendelikte.

Insgesamt machten Schwarze und Hispanoamerikaner 51 % bzw. 33 % der Haltestellen aus, obwohl sie nach der Volkszählung von 1990 nur 26 % bzw. 24 % der Stadtbevölkerung ausmachten. Die Anteile ändern sich kaum, wenn wir die Bevölkerungsschätzungen von 1998 heranziehen und nur Männer im Alter von 15-30 Jahren zählen, was wohl eine bessere Ausgangsbasis darstellt. Für eine unserer ergänzenden Analysen verwenden wir auch die Bevölkerung für jede ethnische Gruppe in jedem Bezirk der Stadt. Die Bevölkerungsschätzungen für die Polizeibezirke mit geringer Wohnbevölkerung, aber hoher Tagesbevölkerung aufgrund von Handels- und Geschäftsaktivitäten wurden anhand der vom New York City Department of City Planning bereitgestellten "journey file" des U.S. Census Bureau angepasst (siehe Spitzer 1999, App. I, Tabelle 1.A.1a). Die "journey file" verwendet Algorithmen, die auf der Zeit für den Arbeitsweg und der Verteilung der Arbeitsplatzklassifizierungen basieren, um die Tages- und Nachtbevölkerung der Stadtbezirke zu schätzen. Die Trakte wurden zu den entsprechenden Polizeibezirken aggregiert, um Tag- und Nachtbevölkerungsschätzungen zu erstellen, und es wurden separate Haltestellenschätzungen für Tag- und Nachtzeitintervalle berechnet. Für diese Analysen wurden getrennte Schätzungen der Anschläge nach Tag und Nacht aggregiert, um die Gesamtanschlagraten für jedes Revier zu berechnen.

Ein vielleicht relevanterer Vergleich ist jedoch die Zahl der von Angehörigen der jeweiligen ethnischen Gruppe begangenen Straftaten. So erklärte beispielsweise der damalige New Yorker Polizeipräsident Howard Safir (Safir

1999),

819

Die rassische/ethnische Verteilung der Personen, die einer "Stop and Frisk"-Meldung unterzogen werden, spiegelt die Demografie der bekannten Verdächtigen von Gewaltverbrechen wider, wie sie von den Opfern von Straftaten angegeben wird. In ähnlicher Weise entspricht die Demografie der Festgenommenen bei Gewaltverbrechen auch der Demografie der bekannten Verdächtigen von Gewaltverbrechen.

Daten über tatsächliche Straftaten sind natürlich nicht verfügbar, daher verwenden wir als Ersatz die Anzahl der Verhaftungen in New York City im Vorjahr 1997, wie sie von der Division of Criminal Justice Services (DCJS) des Staates New York erfasst und nach ethnischer Gruppe und Art der Straftat kategorisiert wurden. Dies wurde als bestes verfügbares Maß für die nach ethnischer Zugehörigkeit aufgeschlüsselten lokalen Kriminalitätsraten angesehen und geht direkt auf die Bedenken von Safir ein, dass die Anhaltequoten mit der ethnischen Zugehörigkeit der Tatverdächtigen in Zusammenhang stehen. Wir verwenden die DCJS-Verhaftungsraten des Vorjahres zur Darstellung der Häufigkeit von

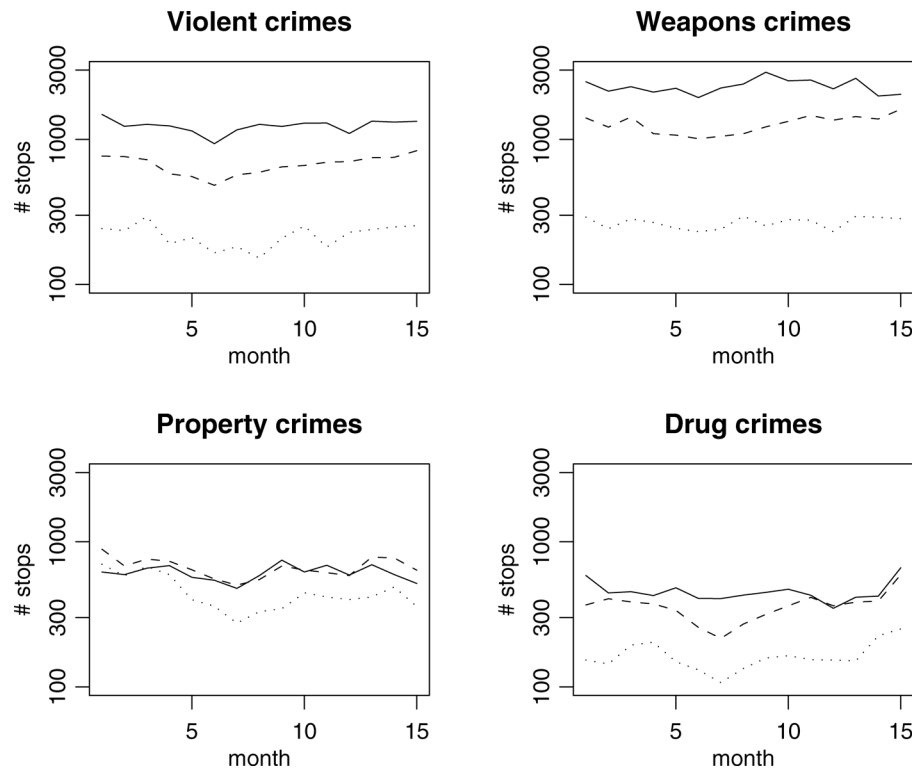


Abbildung 1. Anzahl der Polizeikontrollen in jedem der 15 Monate, charakterisiert nach Art der Straftat und ethnischer Zugehörigkeit der angehaltenen Person (—, Schwarze; -- --, Hispanoamerikaner; . . . , Weiße).

Straftaten, bei denen die Polizei vermutet, dass sie von Angehörigen der jeweiligen ethnischen Gruppe begangen wurden. Bei diesem Vergleich betrug das Verhältnis von Anschlägen zu DCJS-Festnahmen 1,24 für Weiße, 1,54 für Schwarze und 1,72 für Hispanoamerikaner; nach diesem Vergleich sind Schwarze

23 % häufiger angehalten als Weiße und Hispanics werden
angehalten
39 % häufiger als Weiße.

4. MODELLE

Die bisherigen Zusammenfassungen beschreiben Durchschnittswerte für die gesamte Stadt. Nehmen wir jedoch an, dass die Polizei in Gegenden mit hoher Kriminalität mehr Kontrollen durchführt, aber die verschiedenen ethnischen Gruppen innerhalb jeder Ortschaft gleich behandelt. Dann könnten die stadtweiten Quoten signifikante Unterschiede zwischen den ethnischen Gruppen aufweisen, selbst wenn die Kontrollen ausschließlich durch den Ort und nicht durch die ethnische Zugehörigkeit bestimmt werden. Um diese beiden Arten von Prädiktoren voneinander zu trennen, führten wir mehrstufige Analysen unter Verwendung der 75 Stadtbezirke durch. Das Zulassen von Effekten auf Bezirksebene steht im Einklang mit Theorien zur Polizeiarbeit wie "broken windows", die lokale Strategien auf Stadtteilebene betonen (Wilson und Kelling 1982; Skogan 1990). Da es möglich ist, dass sich die Muster in Stadtteilen mit unterschiedlicher ethnischer Zusammensetzung systematisch unterscheiden, haben wir die Bezirke nach ihrem Anteil an Schwarzen in drei Kategorien unterteilt: Bezirke mit einem Anteil von weniger als 10 % Schwarzen, 10-40 % Schwarzen und mehr als 40 % Schwarzen. Wir berücksichtigten auch die Unterschiede in

gression mit Indikatoren für ethnische Gruppen, ein hierarchisches Modell für Bezirke und n_{ep} , die Anzahl der DCJS-Festnahmen für diese ethnische Gruppe in diesem Bezirk (multipliziert mit 15/12, um auf einen 15-Monats-Zeitraum zu skalieren), als Basislinie oder Offset,

$$y_{ep} \sim \text{Poisson} \left(\frac{12}{e^{\mu + \alpha + \beta + E}} n_{ep} \right)$$

den Anhalteraten zwischen den Bezirken innerhalb jeder Gruppe. Jede der drei Kategorien repräsentiert etwa 1/3 der Bezirke in der Stadt, und wir haben für jede Gruppe separate Analysen durchgeführt.

4.1 Hierarchisches Poisson-Regressionsmodell

Für jede ethnische Gruppe e 1, 2, 3 und jeden Bezirk p modellierten wir die Anzahl der Haltestellen y_{ep} mit Hilfe einer überdispersen Poisson-Re

$$\beta_p \sim N(0, \sigma^2), \quad (1)$$

$$E_{ep} \sim N(0, \sigma^2),$$

wobei die Koeffizienten α_e (die sich zwangsweise zu 0 summieren) die ethnischen Gruppen kontrollieren, die β -Werte die Variation zwischen den p Bezirken (mit einer Varianz von σ_β) korrigieren und die $_{ep}E$ -Werte eine Überdispersion berücksichtigen, d. h. eine Variation in den Daten, die über die durch das Poisson-Modell erklärte hinausgeht. Wir passen das Modell mithilfe der Bayes'schen Inferenz mit einer nicht-informativen gleichmäßigen Prioritätsverteilung für die Parameter μ , α , σ_β und σ_{Ean} .

Bei der klassischen verallgemeinerten linearen Modellierung oder bei verallgemeinerten Schätzgleichungen kann die Überdispersion mithilfe einer Chi-Quadrat-Statistik geschätzt werden, wobei die Standardfehler um die Quadratwurzel der geschätzten Überdispersion erhöht werden (McCullagh und Nelder 1989). In unserer Analyse verwenden wir bereits die Bayes'sche Inferenz, um die Variation zwischen den Bezirken zu modellieren, und so stellt die Überdispersion einfach eine weitere Varianzkomponente im Modell dar; die sich daraus ergebenden Schlussfolgerungen haben in der Tat größere Standardfehler als die, die sich aus der nicht überdispersen Regression ergeben würden (was σ_{E0} entsprechen würde), und diese posterioren Standardfehler können z. B. mit einer Kreuzvalidierung der Bezirke überprüft werden.

Von größtem Interesse sind jedoch die potenzierten Koeffizienten $\exp(\alpha_e)$, die die relativen Raten der Stopps im Vergleich zu

 β E

=

Verhaftungen, nach Kontrolle für das Revier. Durch den Vergleich der Anhaltequoten mit den Verhaftungsquoten können wir auch die Anhalte im Zusammenhang mit verschiedenen Arten von Straftaten getrennt analysieren. Wir haben separate Vergleiche durchgeführt.

für Gewaltverbrechen, Waffendelikte und Eigentumsdelikte,

und Drogenkriminalität. Für jede Art von Straftat haben wir die Anzahl der Kontrollen y_{ep} nach ethnischer Gruppe e und Bezirk p modelliert, wobei wir als Basis die DCJS-Festnahmezahl n_{ep} für diese ethnische Gruppe, diesen Bezirk und diese Art von Straftat verwendet haben. (Die Unterteilung nach Verbrechenart ist in dieser Notation implizit; um die Notation einfach zu halten, haben wir keinen zusätzlichen Index für die vier Kategorien von Verbrechen eingeführt).

Wir haben daher Modell (1) für 12 separate Teilmengen der Daten geschätzt, die den vier Verbrechenarten und den drei Kategorien von Bezirken (< 10 % schwarze Bevölkerung, 10-40 % Schwarze und > 40 % Schwarze) entsprechen. Die Berechnungen wurden einfach mit der Bayes-Software BUGS (Spiegelhalter, Thomas, Best, Gilks und Lunn 1994, 2003) durchgeführt, die die Markov-Ketten-Monte-Carlo-Simulation von R (R Project 2000; Sturtz, Ligges und Gelman 2005) implementiert. Für jede Anpassung simulierten wir drei verschiedene, voneinander abhängige Markov-Ketten von unterschiedlichen Ausgangspunkten aus und hörten auf, wenn die Simulationen jeder einzelnen Kette so variabel waren wie die aller Ketten zusammen (Gelman und Rubin 1992). Wir haben dann die letzte Hälfte der simulierten Ketten gesammelt und diese zur Berechnung der Posteriorischätzungen und Standardfehler verwendet. Für die in diesem Artikel berichteten Analysen waren 10.000 Iterationen immer ausreichend, um die Sequenzen zu mischen. Wir berichten über Schlussfolgerungen unter Verwendung der posterioren Mittelwerte und Standardabweichungen, die angesichts des großen Stichprobenumfangs angemessene Zusammenfassungen sind (siehe z. B. Gelman, Carlin, Stern und Rubin 2003, Kap. 4).

4.2 Alternative Modellspezifikationen

Zusätzlich zur Anpassung des Modells (1), wie zuvor beschrieben, ziehen wir zwei alternative Spezifikationen in Betracht: erstens die Anpassung desselben Modells, aber die Änderung der Stapelung der Bezirke, und zweitens die Änderung der Rolle, die die Verhaftungen des Vorjahres im Modell spielen. Wir vergleichen die Anpassungen unter diesen alternativen Modellen, um die Sensitivität für Details der Modellspezifikation zu bewerten.

Modellierung der Variabilität zwischen den Bezirken. Die Einteilung der Bezirke in drei Kategorien ist zweckmäßig und sinnvoll, da sich Stadtteile mit unterschiedlichem Anteil an Minderheitenbevölkerung in vielerlei Hinsicht unterscheiden, einschließlich der für jeden Typ angewandten Polizeistrategien (Fagan und Davies 2000). Die separate Anpassung des Modells an jede Gruppe von Bezirken ist daher eine Möglichkeit, die kontextuelle Auswirkungen. Diese Sichtweise hat jedoch einen willkürlichen Charakter. Wir untersuchen dies, indem wir die Bezirke in verschiedene Bereiche unterteilen.

wobei z_{1p} und z_{2p} den Anteil der Schwarzen und Hispanoamerikaner an der Bevölkerung im Bezirk p darstellen. Wir betrachten auch Variablen des Modells (2), einschließlich der quadratischen Terme z^2 , z^2 , und

$z_{1p} z_{2p}$, um die Empfindlichkeit gegenüber Nichtlinearität zu untersuchen.

Modellierung der Beziehung zwischen Anhalten und Festnahmen im Vorjahr. Wir betrachten auch verschiedene Möglichkeiten, die Anzahl der DCJS-Festnahmen n_{ep} im Vorjahr zu verwenden, die in Modell (1) die Rolle einer Basislinie (oder eines Offsets, in der Terminologie der verallgemeinerten linearen Modelle) spielt. Die Einbeziehung der Verhaftungsrate aus der Vergangenheit als Offset ist sinnvoll, da wir an der Rate der Festnahmen pro Straftat interessiert sind und die Verhaftungen aus der Vergangenheit als Proxy für die Kriminalitätsrate und die Erwartungen der Polizei hinsichtlich der Demografie der Täter verwenden. Eine andere Möglichkeit ist, stattdessen den Logarithmus der Anzahl der früheren Festnahmen als linearen Prädiktor zu verwenden,

$$y_{ep} \sim \text{Poisson} \left(\frac{15}{12} e^{\gamma n_{ep} + \mu + \alpha e + \beta p \log + E_{ep}} \right). \quad (3)$$

Modell (3) reduziert sich auf das Offset-Modell (1), wenn $\gamma = 1$ ist. Wir können also (3) anpassen und sehen, ob sich die Schlussfolgerungen für α_e im Vergleich zu dem früheren Modell, das γ implizit auf 1 festlegt, ändern.

Wir können diese Idee weiterführen, indem wir frühere Verhaftungen als Ersatz für die tatsächliche Kriminalitätsrate modellieren. Wir versuchen, dies auf zwei Arten zu tun, wobei jeder Ansatz die tatsächliche Kriminalitätsrate für jede ethnische Zugehörigkeit in jedem Bezirk als θ_{ep} bezeichnet, mit separaten hierarchischen Poisson-Regressionen für die diesjährigen Kontrollen und die Festnahmen des letzten Jahres (als alle Arten, einschließlich des Faktors $\frac{15}{12}$ zur Berücksichtigung unserer 15 Monate an Kontrolldaten). In der ersten Formulierung modellieren wir die Festnahmen des letzten Jahres als Poisson-verteilt mit Mittelwert θ ,

$$\begin{aligned} y_{ep} &\sim \text{Poisson} \left(\frac{15\theta}{12} e^{\mu + \alpha e + \beta p + E_{ep}} \right), \\ n_{ep} &\sim \text{Poisson}(\theta_{ep}), \\ \log \theta_{ep} &= \log N_{ep} + \tilde{\alpha}_e + \tilde{\beta} p + \tilde{E}_p. \end{aligned} \quad (4)$$

Hier verwenden wir N_{ep} , die Bevölkerung der ethnischen Gruppe e im Bezirk p , als Basis für das Modell der Kriminalitätshäufigkeit.

Die Fehlerterme der zweiten Ebene $\tilde{\beta}$ und \tilde{E} sind normale Hyper-Prioritätsverteilungen wie bei Modell (1).

Unser zweites zweistufiges Modell ähnelt (4), wobei der neue Fehlerterm \tilde{E} in das Modell für n eingefügt wird,

$$\begin{aligned} y_{ep} &\sim \text{Poisson} \left(\frac{15\theta}{12} e^{\mu + \alpha e + \beta p + E_{ep}} \right), \\ n_{ep} &\sim \text{Poisson}(\theta e^{E_{ep}}), \end{aligned} \quad (5)$$

German, Hagan, und Kiss: "Stop-and-Frisk"-Politik
Anzahl der Kategorien und sehen, wie sich die
Modellschätzungen ändern.

Ein weiterer Ansatz zur Kontrolle der systematischen Variation zwischen den Bezirken ist die Einbeziehung von Prädiktoren auf Bezirksebene, die zusammen mit den individuellen Effekten auf Bezirksebene in das Mehrebenenmodell aufgenommen werden können (siehe z. B. Raudenbush und Bryk 2002). Wie bereits erwähnt, ist die Information auf Bezirksebene, die von größtem Interesse ist und auch das größte Potenzial hat, unsere Ergebnisse zu beeinflussen, die ethnische Zusammensetzung der Bevölkerung. Daher berücksichtigen wir als Regressionsprädiktoren den Anteil der Schwarzen und Hispaniker im Bezirk und ersetzen Modell (1) durch

$$y_{ep} \sim \text{Poisson} \left(\frac{15}{12} n_{ep} \mu + \alpha_e + \zeta_1 z_1 p + \zeta_2 z_2 p + \beta p + E_{ep} \right), (2)$$

$$\log \theta_{ep} = \log N_{ep} + \alpha_e + \beta p.$$

823

Nach diesem Modell entsprechen die Verhaftungsraten n_{ep} im Durchschnitt den zugrunde liegenden Kriminalitätsraten θ_{ep} , allerdings mit einer Überdispersion im Vergleich zur Poisson-Fehlerverteilung.

5. ERGEBNISSE

5.1 Primäre Regressionsanalyse

Tabelle 1 zeigt die Schätzungen von Modell (1), die für jede der vier Kriminalitätsarten in jeder der drei Kategorien von Bezirken angepasst wurden. Die Standardabweichungen σ_β und σ der zufälligen Effekte sind beträchtlich, was darauf hindeutet

die Relevanz der hierarchischen Modellierung für diese Daten. [Es sei daran erinnert, dass diese Effekte alle auf einer logarithmischen Skala liegen, so dass ein

Tabelle 1. Schätzungen und Standardfehler für den konstanten Term μ , die Ethnizitätsparameter α und die Varianzparameter σ^2 und σE auf Bezirksebene und auf Bezirksebene nach Ethnizität für das hierarchische Poisson-Regressionsmodell (1), das getrennt für drei Kategorien des Reviers und der vier Verbrechenarten

Anteil schwarz im Revier	Parameter	Art der Straftat			
		Gewalttätig	Waffen	Eigentum	Droge
< 10%	Abfangen	-.85(.07)	.13(.07)	-.58(.21)	-1.62(.16)
	$\alpha 1$ [Schwärze]	.40(.06)	.16(.05)	-.32(.06)	-.08(.09)
	$\alpha 2$.13(.06)	.12(.04)	.32(.06)	.17(.10)
	[Hispanoamerikaner]	-.53(.06)	-.28(.05)	.00(.06)	-.08(.09)
	$\alpha 3$ [Weiße]	.33(.08)	.38(.08)	1.19(.20)	.87(.16)
	σ^2				
	σE	.30(.04)	.23(.04)	.32(.04)	.50(.07)
	Abfangen	-.97(.07)	.42(.07)	-.89(.16)	-1.87(.13)
	$\alpha 1$ [Schwärze]	.38(.04)	.24(.04)	-.16(.06)	-.05(.05)
	$\alpha 2$.08(.04)	.13(.04)	1.25(.08)	.12(.06)
10–40%	[Hispanoamerikaner]	-.46(.04)	-.36(.04)	-.08(.06)	-.07(.05)
	$\alpha 3$ [Weiße]	.49(.07)	.47(.07)		.90(.13)
	σ^2				
	σE	.24(.03)	.24(.03)	.38(.04)	.32(.04)
	Abfangen	-.47(.11)	.29(.18)	.54(.11)	-2.62(.12)
	$\alpha 1$ [Schwärze]	-1.58(.10)	.22(.11)	-1.15(.19)	.09(.07)
	$\alpha 2$.44(.08)	.30(.07)	-.08(.09)	.09(.06)
	[Hispanoamerikaner]	-.55(.08)	-.44(.08)	-.01(.07)	-.18(.09)
	$\alpha 3$ [Weiße]				
	σ^2				
> 40%	Abfangen	1.16(.06)	.14(.04)	.04(.07)	.09(.07)
	$\alpha 1$ [Schwärze]				
	$\alpha 2$				
	[Hispanoamerikaner]				
	$\alpha 3$ [Weiße]				
	σ^2				
	σE				
	Abfangen				
	$\alpha 1$ [Schwärze]				
	$\alpha 2$				

HINWEIS: Die Schätzungen von $e^{\mu + \alpha}$ sind in Abbildung 2 grafisch dargestellt, und alternative Modellspezifikationen (6a) in Tabelle 3 aufgeführt.

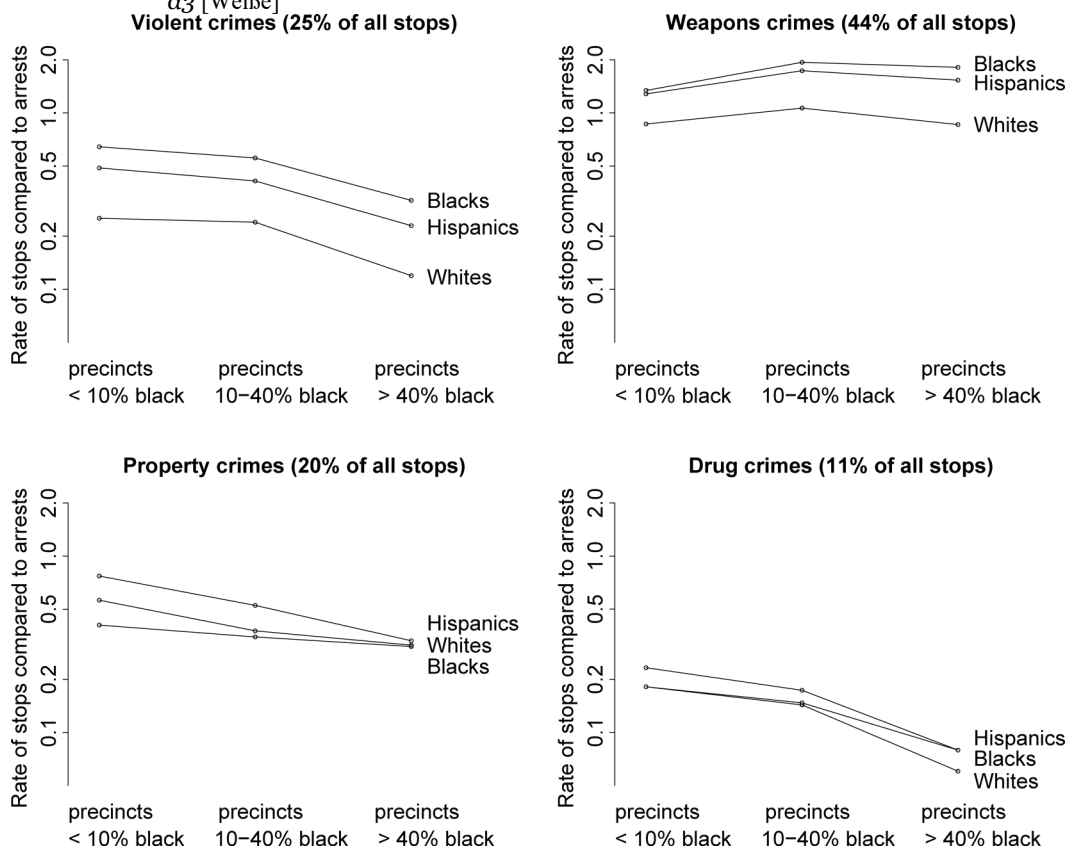


Abbildung 2. Geschätzte Raten $e^{\mu + \alpha}$, mit denen Personen verschiedener ethnischer Gruppen wegen verschiedener Kategorien von Straftaten angehalten wurden, wie aus hierarchischen Regressionen (1) geschätzt, wobei die Verhaftungen des Vorjahres als Basiswert verwendet und für Unterschiede zwischen den Bezirken kontrolliert wurden. Getrennte Analysen wurden für die Bezirke mit < 10 %, 10–40 % und > 40 % schwarzer Bevölkerung durchgeführt. Bei den häufigsten Kontrollen - Gewaltverbrechen und Waffendelikte - wurden Schwarze und Hispanoamerikaner etwa doppelt so häufig angehalten wie Weiße. Die Raten sind auf einer logarithmischen Skala aufgetragen. Die numerischen Schätzungen und Standardfehler sind in Tabelle 1 aufgeführt.

Ein Effekt von 0,3 entspricht beispielsweise einem multiplikativen Effekt von $\exp(0.3)$ 1,35 oder einer 35%igen Erhöhung der Wahrscheinlichkeit, angehalten zu werden].

Die Parameter von größtem Interesse sind die Anhaltequoten (im Vergleich zu den Festnahmen des Vorjahres) für jede ethnische Gruppe, $e^{\mu + \alpha_e}$, für e 1, 2, 3. Diese werden in Abbildung 2 grafisch dargestellt. Anhaltungen

Gewalt- und Waffendelikte waren der umstrittenste Aspekt der Stop-and-frisk-Politik (und machen mehr als zwei Drittel der Kontrollen aus), aber der Vollständigkeit halber zeigen wir hier alle vier Kategorien von Straftaten.

Abbildung 2 zeigt, dass Schwarze und Hispanoamerikaner bei den häufigsten Kategorien von Kontrollen - im Zusammenhang mit Gewalt- und Waffendelikten - in allen Kategorien von Bezirken viel häufiger angehalten wurden als Weiße. Bei Gewaltverbrechen wurden Schwarze und Hispanoamerikaner 2,5-mal bzw. 1,9-mal so häufig angehalten wie Weiße, und bei Waffendelikten wurden Schwarze und Hispanoamerikaner 1,8-mal bzw. 1,6-mal so häufig angehalten wie Weiße. Bei den weniger häufigen Kategorien von Kontrollen wurden Weiße etwas häufiger wegen Eigentumsdelikten und häufiger wegen Drogendelikten angehalten, und zwar im Verhältnis zu den Verhaftungen des Vorjahres in einem bestimmten Bezirk.

5.2 Alternative Formen des Modells

Die Anpassung der in Abschnitt 4.2 beschriebenen alternativen Modelle führte zu ähnlichen Ergebnissen wie unsere Hauptanalyse. Wir erörtern jedes alternative Modell der Reihe nach.

Abbildung 3 zeigt die geschätzten Anhaltequoten bei Gewaltverbrechen im Vergleich zu den Verhaftungen des Vorjahres für jede der drei ethnischen Gruppen für Analysen, bei denen die Bezirke in 5, 10 und 15 Kategorien unterteilt wurden, geordnet nach dem Prozentsatz der schwarzen Bevölkerung im Bezirk. Der Einfachheit halber geben wir nur die Ergebnisse für Gewaltverbrechen an; diese sind typisch für die alternativen Analysen für alle vier Verbrechenarten. Für jedes der drei Diagramme in Abbildung 3 wird das Modell für jede der drei Gruppen von Bezirken getrennt geschätzt, und diese Schätzungen werden für jede ethnische Gruppe in einer Linie verbunden. Verglichen mit der oberen linken Grafik in Abbildung 2, die die Ergebnisse der Einteilung der Bezirke in drei Kategorien zeigt, wird deutlich, dass die Einteilung in mehr Gruppen die Schätzung verrauscht, aber das Gesamtmuster der Unterschiede zwischen den Gruppen nicht verändert.

Tabelle 2 zeigt die Ergebnisse des Modells (2), das auf alle 75 Wahlbezirke angewandt wird, aber für den Anteil der Schwarzen und

Hispanoamerikaner in den Wahlbezirken. Die Schlussfolgerungen ähneln denen aus der in Abschnitt 5.1 erörterten Hauptanalyse. Einbeziehung von quadratischen Termen und Interaktionen in das Modell auf Bezirksebene (2) und die Einbeziehung der Prädiktoren auf Bezirksebene in die Modelle, die an jeden der drei Teilsätze der Daten angepasst wurden, hatten ebenfalls nur geringe Auswirkungen auf die interessierenden Parameter α_e .

Tabelle 3 zeigt Parameterschätzungen aus den Modellen, die die Verhaftungsraten n_{ep} des Vorjahres auf andere Weise berücksichtigen. Aus Gründen der Übersichtlichkeit werden die Ergebnisse nur für Gewaltverbrechen dargestellt, und der Einfachheit halber werden alle 75 Bezirke in die Modelle einbezogen. (Ähnliche Ergebnisse wurden erzielt, wenn das Modell für jede der drei Kategorien von Bezirken und für die anderen Verbrechenarten getrennt angepasst wurde). Die ersten beiden Spalten von Tabelle 3 zeigen die Ergebnisse unseres Hauptmodells (1) und des alternativen Modells (3), das $\log n_{ep}$ als Regressionsprädiktor einbezieht. Die beiden Modelle unterscheiden sich nur dadurch, dass γ im ersten Modell auf 1 beschränkt ist, aber wie wir sehen können, wird γ in der Regressionsformulierung sehr nahe an 1 geschätzt, und die Koeffizienten α_e bleiben im Wesentlichen unverändert. (Der Achsenabschnitt ändert sich ein wenig, weil $\log n_{ep}$ keinen Mittelwert von 0 hat.)

Die letzten beiden Spalten in Tabelle 3 zeigen die Schätzungen der zweistufigen Regressionsmodelle (4) und (5). Die Modelle unterscheiden sich in ihren Schätzungen der Varianzparameter σ_β und σ_E , aber die Schätzungen der Schlüsselparameter α sind im ursprünglichen Modell im Wesentlichen identisch.

Wir haben auch Analysen mit Indikatoren für den Monat der Festnahme durchgeführt. Diese Analysen brachten keine zusätzlichen Informationen für den Vergleich der ethnischen Gruppen.

5.3 Trefferquoten: Anteil der Anschläge, die zu Festnahmen führten

Eine andere Möglichkeit, ethnische Gruppen zu vergleichen, besteht darin, den Anteil der Kontrollen auf der Straße zu betrachten, die zu Festnahmen führen. Die meisten Kontrollen führen nicht zu Festnahmen, und die meisten Festnahmen gehen nicht auf Kontrollen zurück. In der zuvor beschriebenen Analyse haben wir untersucht, wie oft die Polizei Personen aus verschiedenen Gruppen angehalten hat. Jetzt schauen wir uns kurz an, was mit diesen Kontrollen geschieht.

In dem Zeitraum, für den uns Daten vorliegen, wurde 1 von 7,9 angehaltenen Weißen verhaftet, verglichen mit etwa 1 von 8,8 Hispanoamerikanern und 1 von 9,5 Schwarzen. Diese Daten stimmen mit unserer allgemeinen Schlussfolgerung überein, dass die Polizei unverhältnismäßig viele Minderheiten anhält; die Kontrollen von Weißen sind "effizienter" und führen mit größerer Wahrscheinlichkeit zu Verhaftungen, während die Kontrollen von Schwarzen und Hispanoamerikanern wahlloser sind und weniger Personen angehalten werden.

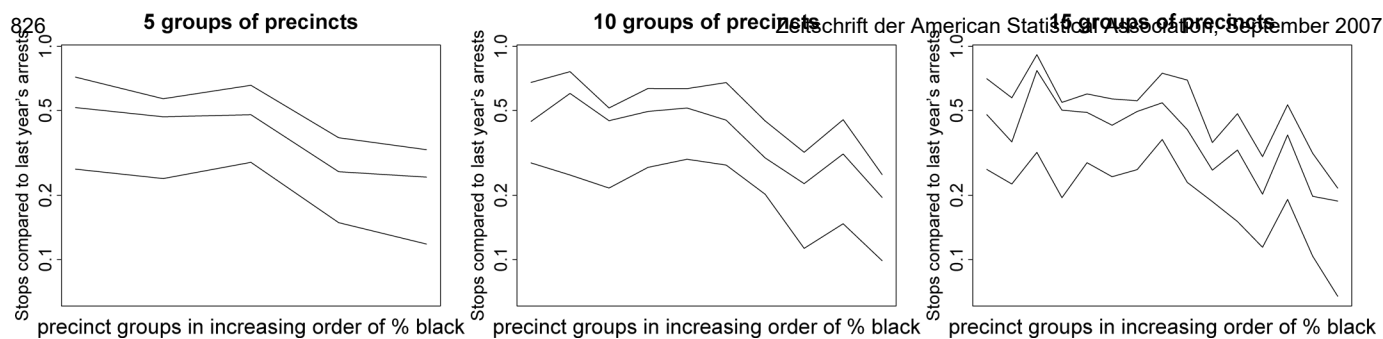


Abbildung 3. Geschätzte Raten $e\mu + \alpha e$, mit denen Personen verschiedener ethnischer Gruppen wegen Gewaltverbrechen angehalten wurden, wie aus Modellen geschätzt, die die Bezirke in 5, 10 und 15 Kategorien unterteilen. In jedem Diagramm entsprechen die obere, mittlere und untere Linie den Schwarzen, Hispanics und Weißen. Diese Diagramme zeigen die gleichen allgemeinen Muster wie das Modell mit drei Kategorien (das obere linke Diagramm in Abb. 2), jedoch mit zunehmendem Rauschen.

Tabelle 2. Schätzungen und Standardfehler für die Parameter des Modells (2), das den Anteil der Schwarzen und Hispanoamerikaner als Prädiktoren auf Bezirksebene enthält, angepasst an alle 75 Bezirke

Parameter	Art der Straftat			
	Gewalttätig	Waffen	Eigentum	Droge
Abfangen	-.66(.08)	.08(.11)	-.14(.24)	-.98(.17)
α_1 [Schwärze]	.41(.03)	.24(.03)	-.19(.04)	-.02(.04)
α_2 [Hispanoamerikaner]	.10(.03)	.12(.03)	.23(.04)	.15(.04)
α_3 [Weiße]]-.51(.03)	-.36(.03)	-.05(.04)	-.13(.04)
ζ_1 [Koeff. für Prop. Schwarz]]-1.22(.18)	.10(.19)	-1.11(.45)	-1.71(.31)
ζ_2 [Koeff. für Prop. Hispanic]	-.33(.23)	.71(.27)	-1.50(.57)	-1.89(.41)
$\sigma\beta$.40(.04)	.43(.04)	1.04(.09)	.68(.06)
σE	.25(.02)	.27(.02)	.37(.03)	.37(.03)

ANMERKUNG: Die Ergebnisse für die interessierenden Parameter α_i ähneln denen, die durch getrennte Anpassung des Basismodells an jede der drei Kategorien von Bezirken erzielt wurden, wie in Tabelle 1 und Abbildung 2 dargestellt. Wie zuvor wird das Modell separat an die Daten von vier verschiedenen Kriminalitätsarten angepasst.

dieser breiteren Rasterfahndung tatsächlich verhaftet werden.

Es ist durchaus vertretbar, dass die Polizei viele Kontrollen durchführt, die nicht zu Festnahmen führen; hier geht es um den Vergleich zwischen ethnischen Gruppen. Dies kann auch im Sinne einer einfachen ökonomischen Theorie verstanden werden (in Anlehnung an die Argumentation von Knowles,

Persico und Todd 2001 für Polizeikontrollen wegen

Drogenverdachts). Es ist vernünftig anzunehmen, dass die Rentabilität von Polizeikontrollen in dem Sinne abnimmt, dass ab einem bestimmten Punkt nur noch ein geringer Nutzen durch das Anhalten weiterer Personen erzielt wird. Wenn der Gewinn durch Verhaftungen annähernd zusammengefasst wird, dann bedeutet abnehmender Ertrag, dass die Wahrscheinlichkeit, dass eine Kontrolle zu einer Verhaftung führt - ökonomisch ausgedrückt, der Grenzgewinn aus der Kontrolle einer weiteren Person - mit der Zahl der angehaltenen Personen abnimmt. Die

Kontrollen von Schwarzen und Hispanics waren weniger "effizient" als die von Weißen, was darauf hindeutet, dass die

Polizei bei der Kontrolle von Angehörigen von Minderheitengruppen weniger strenge Maßstäbe anlegt. Zu ähnlichen Ergebnissen gelangten wir, wenn wir eine separate An-

Vergleich von Tages- und Nachtanschlügen.

Diese Analyse der "Trefferquote" kann jedoch als unfair gegenüber der Polizei kritisiert werden, die "verdammte ist, wenn sie es tut, verdammte ist, wenn sie es nicht tut". Relativ wenige Kontrollen von Minderheiten führten zu Verhaftungen, so dass wir zu dem Schluss kommen, dass die Polizei eher bereit war, Angehörige von Minderheitengruppen mit weniger Grund anzuhalten. Wir könnten aber auch andersherum argumentieren: Da ein relativ hoher Anteil der angehaltenen Weißen verhaftet wurde, schließen wir daraus, dass die Polizei gegenüber Weißen voreingenommen ist, indem sie sie zu oft verhaftet. Analysen, die die Stichhaltigkeit von Verhaftungen nach Rasse untersuchten - das heißt,

der Anteil der Festnahmen, die zu Verurteilungen führen, würde zur Klärung dieser Frage beitragen. Leider sind solche Daten nicht ohne weiteres verfügbar. Wir glauben nicht an die letztgenannte Interpretation, aber es ist schwer, sie allein auf der Grundlage dieser Daten auszuschließen.

Aus diesem Grund betrachten wir diesen Teil der Studie lediglich als *unterstützenden* Beweis. Unsere Hauptanalyse ergab, dass Schwarze und Hispanoamerikaner überproportional häufig angehalten wurden (im Vergleich zu ihrer Bevölkerungszahl oder ihrer Kriminalitätsrate, gemessen an ihrer Quote gültiger Verhaftungen im Vorjahr), und die sekundäre Analyse der Trefferquoten oder "Verhaftungseffizienz" dieser Kontrollen stimmt mit diesem Ergebnis überein.

6. DISKUSSION UND SCHLUSSFOLGERUNGEN

In dem Zeitraum, für den uns Daten zur Verfügung standen, zeigen die Aufzeichnungen der NYPD, dass Schwarze und Hispanics häufiger angehalten wurden als Weiße, und zwar sowohl im Vergleich zu den Bevölkerungszahlen dieser Gruppen als auch zu den besten Schätzungen der von jeder Gruppe begangenen Verbrechen. Auch nach Kontrolle der Bezirke bleibt dieses Muster bestehen. Insbesondere bei Gewaltverbrechen und Waffendelikten werden Schwarze und Hispanoamerikaner etwa doppelt so häufig angehalten wie Weiße. Im Gegensatz dazu werden Weiße und Hispanoamerikaner bei den weniger häufigen Kontrollen wegen Eigentums- und Drogendelikten häufiger angehalten als Schwarze, verglichen mit der Verhaftungsrate für jede ethnische Gruppe.

Ein weiterer Beleg dafür ist, dass die Wahrscheinlichkeit, dass Schwarze angehalten und in Panik versetzt werden, geringer ist als bei Weißen, was zu einer Verhaftung führt,

Tabelle 3. Schätzungen und Standardfehler für Parameter nach Modell (1) und drei alternativen Spezifikationen für die Festnahmen des Vorjahres $_{nep}$: Behandlung von $\log(_{nep})$ als Prädiktor im Poisson-Regressionsmodell (3), und die zweistufigen Modelle (4) und (5)

Modell für die Festnahmen des Vorjahres

Achsenabschnitt-1	.08(.06))-. 94(.16))-1.07(.06))-1.13(.07)
$\alpha 1$ [Schwarz]	. 40(.03)	. 41(.03)	. 40(.03)	. 42(.08)
$\alpha 2$ [Hispanoamerikaner]	. 10(.03)	. 10(.03)	. 10(.03)	. 14(.09)
$\alpha 3$ [Weiße] -. 50(.03)) -. 51(.03)) -. 50(.03)	- . 56(.09)
γ [Koeff. für log $_{nep}$]		. 97(.03)		
$\sigma \beta$. 51(.05)	. 51(.05)	. 51(.05)	. 27(.12)	
σE . 26(.02)	. 26(.02)	. 24(.02)	. 67(.04)	

HINWEIS: Der Einfachheit halber werden die Ergebnisse nur für Gewaltverbrechen dargestellt, und zwar für das Modell, das auf alle 75 Bezirke angewendet wurde. Die drei α -Parameter sind bei allen vier Modellen nahezu identisch, wobei sich die Spezifikation nur auf den Achsenabschnitt auswirkt.

was darauf hindeutet, dass die Standards für das Anhalten von Angehörigen von Minderheitengruppen lockerer waren. Zwei verschiedene Szenarien könnten die niedrigeren "Trefferquoten" für Nichtweiße verdeutlichen, eines, das auf die gezielte Ansprache von Minderheiten hindeutet, und ein anderes, das auf die Dynamik der rassistischen Stereotypisierung und eine passivere Form der rassistischen Bevorzugung hindeutet. Im ersten Szenario nutzte die Polizei möglicherweise einen größeren Ermessensspielraum und lockere verfassungsrechtliche Standards bei der Entscheidung, Bürger aus Minderheiten anzuhalten. Diese Erklärung entspräche dem Szenario der "vorgetäuschten" Kontrollen, das in mehreren neueren Studien über Kraftfahrzeugkontrollen diskutiert wird (z. B. Lundman und Kaufman 2003), und lässt vermuten, dass die höheren Anhaltequoten absichtlich und zielgerichtet waren. Alternativ könnte die Polizei einfach häufiger die Wahrnehmung eines "Verdachts" auf der Grundlage einer breiteren Interpretation der sozialen Anhaltspunkte bilden, die die Aufmerksamkeit der Polizei erregen und offizielle Reaktionen hervorrufen (Alpert et al. 2005). Die letztgenannte Erklärung entspricht eher einem sozialpsychologischen Prozess der Rassenstereotypisierung, bei dem die Zuschreibung von Verdacht bei Minderheiten leichter mit bestimmten Verhaltensweisen und Kontexten verbunden wird als bei Weißen (Thompson 1999; Richardson und Pittinsky 2005).

Wir fanden Hinweise auf Kontrollen, die sich am besten als "rassistisch inkongruente" Kontrollen erklären lassen: hohe Raten von Minderheitskontrollen in überwiegend weißen Bezirken. In der Tat ist "fehl am Platz" einer der Auslöser für Verdächtigungen (Alpert et al. 2005; Gould und Mas- trofski 2004). Rassistisch inkongruente Kontrollen sind am auffälligsten in rassistisch homogenen Gebieten. So beobachteten wir beispielsweise hohe Anhaltequoten von Afroamerikanern im überwiegend weißen 19. Revier, ein Zeichen für eine rassenbasierte Auswahl von Bürgern für Polizeikontrollen. Auch in mehreren Bezirken der Bronx wurden hohe Anhaltequoten bei Weißen festgestellt, insbesondere bei Drogendelikten, was höchstwahrscheinlich darauf hindeutet, dass weiße Drogenkäufer in Stadtteile vordringen, in denen überwiegend Minderheiten leben und in denen Drogenmärkte auf der Straße üblich sind. Insgesamt handelte es sich jedoch um relativ seltene Ereignisse, die aufgrund der Bevölkerungsstruktur in diesen Bezirken zu irreführenden Anhaltequoten führten.

Um unsere Ergebnisse kurz zusammenzufassen: Schwarze und Hispanoamerikaner machten 51 % bzw. 33 % der Kontrollen aus, obwohl sie nur 26 % bzw. 24 % der Bevölkerung von New York City ausmachen. Verglichen mit der Anzahl der Verhaftungen jeder Gruppe im Vorjahr (als Indikator für die Häufigkeit kriminellen Verhaltens) wurden Schwarze 23 % häufiger angehalten als Weiße und Hispanics 39 % häufiger als Weiße. Die Kontrolle nach Bezirk vergrößerte diese Diskrepanzen noch, wobei Minderheiten zwischen 1,5 und

Bei den häufigsten Kategorien von Anschlägen (Gewalt- und Drogendelikte) sind ethnische Gruppen 2,5-mal so häufig angehalten worden wie Weiße (im Vergleich zu den vorherigen Verhaftungsraten der Gruppen in den Bezirken, in denen sie angehalten wurden), wobei die Unterschiede bei Eigentums- und Drogendelikten geringer sind. Die Unterschiede bei den

Anhaltequoten zwischen den ethnischen Gruppen sind real, substantiell und lassen sich nicht durch frühere Verhaftungsquoten oder Bezirke erklären. Unsere Ergebnisse bedeuten jedoch nicht unbedingt, dass die NYPD unfair oder rassistisch gehandelt hat. Es ist durchaus vertretbar, anzunehmen, dass eine wirksame Polizeiarbeit das Anhalten und Befragen vieler Personen erfordert, um Informationen über eine bestimmte Person zu sammeln.

Verbrechen.

Vor dem Hintergrund einiger schwieriger Beziehungen zwischen der Polizei und ethnischen Minderheiten in New York City ist es nützlich, ein gewisses quantitatives Verständnis für die strittigen Fragen zu haben. Angesichts der Beschwerden über die Häufigkeit, mit der die Polizei Schwarze und Hispanoamerikaner anhält, ist es wichtig zu wissen, ob es sich dabei tatsächlich um ein statistisches Muster handelt. Die NYPD

830
 hat dann die Möglichkeit, den betroffenen Gemeinden ihre Politik zu erläutern.

In den Jahren seit der Durchführung dieser Studie wurde ein umfassendes Überwachungssystem eingerichtet, mit dem zwei Ziele erreicht werden sollten. Erstens wurden Verfahren entwickelt und eingeführt, die es den Beamten ermöglichten, die Einhaltung der Vorgaben des NYPD-Streifenleitfadens für die genaue und umfassende Erfassung aller Polizeikontrollen zu überwachen. Zweitens wurden die neuen Formulare in Datenbanken eingegeben, die eine kontinuierliche Überwachung der Rassenproportionalität der Kontrollen und ihrer Ergebnisse (z. B. Durchsuchungen, Verhaftungen) ermöglichen sollten. In Verbindung mit einer genauen Berichterstattung über rassenspezifische Messgrößen für Verbrechen und Verhaftungen werden die neuen Verfahren und Überwachungsanforderungen sicherstellen, dass Untersuchungen ähnlich dieser Studie als Teil eines Rahmens von Rechenschaftsmechanismen institutionalisiert werden können.

(Eingereicht im März 2004, überarbeitet im
 Dezember 2005).

REFERENZEN

- Alpert, G., MacDonald, J. H., und Dunham, R. G. (2005), "Police Suspicion and Discretionary Decision Making During Citizen Stops", *Criminology*, 43, 407-434.
- Ayres, I. (2002a), "Outcome Tests of Racial Disparities in Police Practices" (Ergebnistests zu rassischen Unterschieden in der Polizeipraxis). *Justizforschung und -politik*, 4, 131-142.
- (2002b), *Pervasive Prejudice: Unconventional Evidence of Race and Gender Discrimination*, Chicago: University of Chicago Press.
- Bittner, E. (1970), *Functions of the Police in Modern Society: A Review of Background Factors, Current Practices and Possible Role Models*, Rockville, MD: National Institute of Mental Health.
- Bratton, W., und Knobler, P. (1998), *Turnaround*, New York: Norton.
- Brown v. Texas (1979), 443 U.S. 43, U.S. Supreme Court.
- Brown gegen Oncont (2000), 221 F.3d 329, 2nd Circuit Court.
- Cole, D. (1999), *No Equal Justice: Race and Class in the Criminal Justice System*, New York: New Press.
- Dharmapala, D., und Ross, S. L. (2004), "Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence", *Contributions to Economic Policy and Analysis*, 3, Artikel 12; verfügbar unter www.bepress.com/bejeap/contributions/vol13/iss1/art12.
- Dominitz, J., und Knowles, J. (2005), "Crime Minimization and Racial Bias: What Can We Learn From Police Search Data?" PIER Working Paper Archive 05-019, University of Pennsylvania, Penn Institute for Economic Research, Dept. of Economics.
- Durose, M. R., Schmitt, E. L., und Langan, P. A. (2005), "Contacts Between Police and the Public: Findings From the 2002 National Survey," NCJ 207845, Bureau of Justice Statistics, U.S. Department of Justice.
- Eck, J., und Maguire, E. R. (2000), "Have Changes in Policing Reduced Violent Crime: An Assessment of the Evidence," in *The Crime Drop in America*, eds. A. Blumstein und J. Wallman, New York: Cambridge University Press, S. 207-265.
- Fagan, J. (2002), "Law, Social Science and Racial Profiling", *Justice, Research and Policy*, 4, 104-129.
- Fagan, J., und Davies, G. (2000), "Street Stops and Broken Windows: Terry, Race, and Disorder in New York City", *Fordham Urban Law Journal*, 28, 457-504.
- (2003), "Policing Guns: Order Maintenance and Crime Control in New York", in *Guns, Crime, and Punishment in America*, ed. B. Harcourt, New York: New York University Press, S. 191-221.
- Fagan, J., Zimring, F. E., und Kim, J. (1998), "Declining Homicide in New York: A Tale of Two Trends," *Journal of Criminal Law and Criminology*, 88, 1277-1324.
- Fridell, L., Lunney, R., Diamond, D., Kubu, B., Scott, M., und Laing, C. (2001), "Racially Biased Policing: A Principled Response", Police Executive Research Forum; verfügbar unter www.policeforum.org/racial.html.
- Garrett, B. (2001), "Remedying Racial Profiling Through Partnership, Statistics, and Reflective Policing", *Columbia Human Rights Law Review*, 33, 41-148.
- Gelman, A., Carlin, J. B., Stern, H. S., und Rubin, D. B. (2003), *Bayesian Data Analysis* (2nd ed.), London: Chapman & Hall.
- Gelman, A., und Rubin, D. B. (1992), "Inference From Iterative Simulation Using Multiple Sequences" (mit Diskussion), *Statistical Science*, 7, 457-511.

Zeitschrift der American Statistical Association, September 2007
 Goldbeitz, 19, 20, 51.

Gould, J., und Mastrofski, S. (2004), "Suspect Searches: Assess Police Behavior Under the U.S. Constitution", *Kriminologie und öffentliche Ordnung*, 3, 315-361.

- Greene, J. (1999), "Zero Tolerance: A Case Study of Police Policies and Practices in New York City", *Crime and Delinquency*, 45, 171-199.
- Gross, S. R., und Barnes, K. Y. (2002), "Road Work: Racial Profiling and Drug Interdiction on the Highway", *Michigan Law Review*, 101, 653-754.
- Gross, S. R., und Livingston, D. (2002), "Racial Profiling Under Attack", *Columbia Law Review*, 102, 1413-1438.
- Harcourt, B. E. (1998), "Reflecting on the Subject: A Critique of the Social Influence Conception of Deterrence, the Broken Windows Theory, and Order-Maintenance Policing New York Style", *Michigan Law Review*, 97, 291.
- _____. (2001), *Illusion of Order: The False Promise of Broken Windows Policing*, Cambridge, MA: Harvard University Press.
- Harris, D. A. (1999), "The Stories, the Statistics, and the Law: Why 'Driving While Black' Matters", *Minnesota Law Review*, 84, 265-326.
- _____. (2002), *Profiles in Injustice: Why Racial Profiling Cannot Work*, New York: New Press.
- Hemenway, D. (1997), "The Myth of Millions of Annual Self-Defense Gun Uses: A Case Study of Survey Overestimates of Rare Events", *Chance*, 10, 6-10.
- Illinois gegen Wardlow (2000), 528 U.S. 119, U.S. Supreme Court.
- Kelling, G., und Cole, C. (1996), *Fixing Broken Windows*, New York: Free Press.
- Kelvin Daniels et al. gegen die Stadt New York (2004), Vergleichsvereinbarung, 99 Civ 1695 (SAS), angeordnet am 9. Januar 2004; eingereicht am 13. Januar 2004.
- Kennedy, R. (1997), *Race, Crime and Law*, Cambridge, MA: Harvard University Press.
- Knowles, J., Persico, N., und Todd, P. (2001), "Racial Bias in Motor-Vehicle Searches: Theory and Evidence", *Journal of Political Economy*, 109, 203-229.
- Lamberth, J. (1997), "Report of John Lamberth, Ph.D.", American Civil Liberties Union; verfügbar unter www.aclu.org/court/lamberth.html.
- Langan, P., Greenfeld, L. A., Smith, S. K., Durose, M. R., und Levin, D. J. (2001), "Contacts Between Police and the Public: Findings From the 1999 National Survey", NCJ 184957, Bureau of Justice Statistics, U.S. Department of Justice.
- Loury, G. (2002), *The Anatomy of Racial Inequality*, Cambridge, MA: Harvard University Press.
- Lundman, R. J., und Kaufman, R. L. (2003), "Driving While Black: Effects of Race, Ethnicity, and Gender on Citizen Self-Reports of Traffic Stops and Police Actions", *Criminology*, 41, 195-220.
- MacDonald, H. (2001), "The Myth of Racial Profiling", *City Journal*, 11, 2-5.
- Massey, D., und Denton, N. (1993), *American Apartheid: Segregation and the Making of the Underclass*, Cambridge, MA: Harvard University Press.
- McCullagh, P., und Nelder, J. A. (1989), *Generalized Linear Models* (2. Aufl.), New York: Chapman & Hall.
- Miller, J. (2000), "Profiling Populations Available for Stops and Searches", Police Research Series Paper 121, Home Office Research, Development and Statistics Directorate, Vereinigtes Königreich.
- Musto, D. (1973), *The American Disease*, New Haven, CT: Yale University Press.
- Pager, D. (2003), "The Mark of a Criminal Record", *American Journal of Sociology*, 108, 937-975.
- Volk v. DeBour (1976), 40 N.Y.2d 210, 386 N.Y.S.2d 375.
- Volk v. Holmes (1996), 89 N.Y.2d 838, 652 N.Y.S.2d 725.
- R Project (2000), "The R Project for Statistical Computing", verfügbar unter www.r-project.org.
- Raudenbush, S. W., und Bryk, A. S. (2002), *Hierarchische lineare Modelle* (2. Aufl.), Thousand Oaks, CA: Sage.
- Reiss, A. (1971), *The Police and the Public*, New Haven, CT: Yale University Press.
- Richardson, M., und Pittinsky, T. L. (2005), "The Mistaken Assumption of Intentionality in Equal Protection Law: Psychological Science and the Interpretation of the Fourteenth Amendment" (Psychologische Wissenschaft und die Auslegung des vierzehnten Verfassungszusatzes), KSG Working Paper RWP05-011; verfügbar unter ssrn.com/abstract=722647.
- Rudovsky, D. (2001), "Law Enforcement by Stereotypes and Serendipity: Racial Profiling and Stops and Searches Without Cause", *University of Pennsylvania Journal of Constitutional Law*, 3, 296-366.
- Russell, K. K. (2002), "Racial Profiling: A Status Report of the Legal, Legislative, and Empirical Literature", *Rutgers Race and the Law Review*, 3, 61-81.
- Safir, H. (1999), "Statement Before the New York City Council Public Safety Kommission", 19. April; zitiert in Spitzer (1999).
- Sampson, R. J., und Raudenbush, S. W. (1999), "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods", *American Journal of Sociology*, 105, 622-630.
- Sampson, R. J., Raudenbush, S. W., und Earls, F. (1997), "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy", *Wissenschaft*, 277, 918-924.
- Skogan, W. (1990), *Disorder and Decline: Crime and the Spiral of Decay in American Cities*, Berkeley, CA: University of California Press.
- Skogan, W., und Frydl, K. (Hrsg.) (2004), *The Evidence on Policing: Fairness and Effectiveness in Law Enforcement*, Washington, DC: National Academy Press.
- Skolnick, J., und Capolovitz, A. (2001), "Guns, Drugs and Profiling: Ways to Target Guns and Minimize Racial Profiling", *Arizona Law Review*, 43, 413-448.
- Smith, D. A. (1986), "The Neighborhood Context of Police Behavior", in *Communities and Crime: Crime and Justice: An Annual Review of Research*, eds. A. J. Reiss und M. Tonry, Chicago: University of Chicago Press, S. 313-342.
- Smith, M., und Alpert, G. (2002), "Searching for Direction: Courts, Social Science, and the Adjudication of Racial Profiling Claims", *Justice Quarterly*, 19, 673-704.
- Smith, M. R., Makarios, M., und Alpert, G. P. (2006), "Differential Suspicion: Theory Specification and Gender Effects in the Traffic Stop Context", *Justice Quarterly*, 23, 271-295.
- Spiegelhalter, D., Thomas, A., Best, N., Gilks, W., und Lunn, D. (1994, 2003), "BUGS: Bayesian Inference Using Gibbs Sampling", MRC Biostatistics Unit, Cambridge, U.K.; verfügbar unter www.mrc-bsu.cam.ac.uk/bugs/.
- Spitzer, E. (1999), "The New York City Police Department's 'Stop and Frisk' Practices", Office of the New York State Attorney General; verfügbar unter www.oag.state.ny.us/press/reports/stop_frisk/stop_frisk.html.
- Sturtz, S., Ligges, U., und Gelman, A. (2005), "R2WinBUGS: A Package for Running WinBUGS From R", *Journal of Statistical Software*, 12 (3), 1-16.
- Sykes, R. E., und Clark, J. P. (1976), "A Theory of Deference Exchange in Police-Civilian Encounters", *American Journal of Sociology*, 81, 587-600.
- Taylor, R. B. (2000), *Breaking Away From Broken Windows: Baltimore Neighborhoods and the Nationwide Fight Against Crime, Grime, Fear, and Decline*, Boulder, CO: Westview Press.
- Terry gegen Ohio (1968), 392 U.S. 1, U.S. Supreme Court.
- Thompson, A. (1999), "Stopping the Usual Suspects: Race and the Fourth Amendment", *New York University Law Review*, 74, 956-1013.
- Tyler, T. R., und Huo, Y. J. (2003), *Trust in the Law*, New York: Russell Sage Foundation.
- U.S. Commission on Civil Rights (2000), "Police Practices and Civil Rights in New York City", verfügbar unter www.uscrr.gov/pubs/nypolice/main.htm.
- U.S. gegen Martinez-Fuerte (1976), 428 U.S. 543, U.S. Supreme Court.
- Veneiro, P., und Zoubeck, P. H. (1999), "Interim Report of the State Police Review Team Regarding Allegations of Racial Profiling", Office of the New Jersey State Attorney General.
- Walker, S. (2001), "Searching for the Denominator: Problems With Police Traffic Stop Data and an Early Warning System Solution", *Justice Research and Policy*, 3, 63-95.
- Wang, A. (2001), "Illinois v. Wardlow and the Crisis of Legitimacy: An Argument for a 'Real Cost' Balancing Test", *Law and Inequality*, 19, 1-30.
- Weidner, R. R., Frase, R., und Pardoe, I. (2004), "Explaining Sentence Severity in Large Urban Counties: A Multilevel Analysis of Contextual and Case-Level Factors", *Prison Journal*, 84, 184-207.
- Weitzer, R. (2000), "Racialized Policing: Residents' Perceptions in Three Neighborhoods", *Law and Society Review*, 34, 129-155.
- Weitzer, R., und Tuch, S. A. (2002), "Perceptions of Racial Profiling: Race, Class and Personal Experience", *Kriminologie*, 50, 435-456.
- Whren et al. v. U.S. (1996), 517 U.S. 806, Oberster Gerichtshof der USA.
- Wilson, J. Q., und Kelling, G. L. (1982), "The Police and Neighborhood Safety: Broken Windows", *Atlantic Monthly*, März, 29-38.
- Zimring, F. E. (2006), *The Great American Crime Decline*, Oxford, U.K.: Oxford University Press.
- Zingraff, M., Mason, H. M., Smith, W. R., Tomaskovic-Devey, D., Warren, P., McMurray, H. L., und Fenlon, C. R., et al. (2000), "Evaluating North Carolina State Highway Patrol Data: Citations, Warnings and Searches in 1998", verfügbar unter www.nccrimecontrol.org/shp/neshpreport.htm.