UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK
-- x

DAVID FLOYD et al.,

Plaintiffs,

-against-

08 CIV. 01034 (SAS)

THE CITY OF NEW YORK et al.,

Defendants.

-- X

Report of Dennis C. Smith, Ph.D. and Robert M. Purtell, Ph.D.
in Response to the Second Supplemental Report of Jeffrey Fagan, Ph.D.

(Corrected Version)

I. OVERVIEW

This report responds to further supplemental statistical analysis that updates the scientific evidence presented in Prof. Fagan's expert report dated October 15, 2010 and in his supplemental report dated December 3, 2010, in connection with David Floyd v. The City of New York et al. Prof. Fagan's previous reports presented statistical evidence on patterns of *Terry* stops activity from January 1, 2004 — December 31, 2009. Prof. Fagan extends the analyses he presented in the previous two reports to the current period and provides statistical evidence on patterns of SQF activity from January 1, 2010 — June 30, 2012. We have previously responded in reports dated November 15, 2010 and declarations dated December 19, 2011, and February 16, 2012.

Prof. Fagan has acknowledged several of our criticisms of his earlier analysis, and changed several features of his analytic approach in response to our criticisms without acknowledging the source of the weakness identified, namely his lack of understanding and appreciation on modern, proactive policing in New York City. This lack of understanding initially informed his decision to ignore information provided by officers on the UF250 that was not pre-coded because he claimed it was unintelligible, unreliable, in fact "meaningless,"[1] and it continues to undermine his ability to interpret the rich data provided by officers on the UF250 form in their effort to comply with reporting requirements.

Among a wide range of criticisms of Prof. Fagan's use of multiple regression analysis to assess claims of racial disparity, the most serious was that he consistently

[1] Declarations of Professor Fagan in David Floyd et al v. the City of New York et al. 2-12,11, p.9.

omitted the one variable that is most critical to an analysis of racial disparity in NYPD police *Terry* stops: suspect description. While Prof. Fagan pays lip service to the idea that knowledge of crime patterns by race and ethnicity is necessary to benchmark the distribution of race and ethnicity in actual *Terry* stop patterns, he has gone to extraordinary lengths to challenge the validity, reliability and adequacy of the complaint descriptions and arrest classifications of the race/ethnicity of offenders. The majority of the shortcomings in Prof. Fagan's reports that call into question his claim of providing unbiased, "scientific" evidence are present in this latest, Second Supplemental Report. We will address those failings in as they appear in his report, and provide updates of the analyses we have presented previously to challenge his findings and interpretations of them.

A. **Qualifications**

1. Dennis C. Smith

I am an Associate Professor of Public Administration at the Robert F. Wagner Graduate School of Public Service at New York University. I have served as the Director of the Program in Public Policy and Management and Associate Dean.

Since joining the faculty of NYU in 1973, I have studied urban police policy and management, undertaking studies of police management in the Indianapolis, Indiana, Chicago, Illinois and St. Louis, Missouri metropolitan areas with Prof. Elinor Ostrom of Indiana University, recent recipient of the Nobel Prize in Economics. My dissertation was on the subject of police professionalization and performance based on a study of twenty-nine police departments in the St. Louis metropolitan area. I have done police studies with National Science Foundation and National Institute of Justice funding in the

Tampa/St. Petersburg, Florida and Rochester, New York metropolitan areas since coming to NYU.

I have been studying New York City since the late 1970s when I began an analysis of the organizational and performance effects of a twenty-five percent reduction in the size of the department in the wake of the fiscal crisis, and have studied how well the Police Academy was preparing recruits for community policing, evaluated the effects of command structure reform at the borough level on police performance, studied the introduction and impact of the CompStat (alone and with William Bratton), assessed the performance effects of Operation Impact, evaluated the management crime integrity efforts of NYPD, analyzed the relationship between crime and economic conditions at the neighborhood level, evaluated the reform of the Internal Affairs Bureau, and assessed the efficacy of *Terry* stop practices as crime prevention strategy. I also recently completed an organizational assessment of the Department of Environmental Protection Police that is charged with protecting the New York City water system.

I am currently studying the effects of the adoption of a CompStat approach to policing big cities in New York. I have also studied the adoption of evidence-based, outcome-oriented management practices in social services, non-profit organizations, and the Departments of Corrections and Parks. I have been a consultant to the NYC Office of Operations on the Mayor's Management Report, and to United Way of New York and numerous nonprofit organizations, providing expertise on the use of performance measurement and management. Additionally, I served on the Public Safety Transition Committee of Governor Andrew Cuomo.

My research on police has been published in seven books and in articles in peer reviewed journals, including the *Public Administration Review, Urban Affairs Quarterly, Journal of Criminal Justice, The Journal of Social Issues, Public Administration and Development;* and most recently, my case for evidence-based, outcome-driven performance management was an invited article in the *Journal of Public Policy Analysis and Management*. That article has now been reprinted in the Classics of Public Management issue of the *Journal of Public Policy Analysis and Management.* I co-authored the chapter on Public Safety Policy in New York State in *The Oxford Handbook of New York State Politics* (2012). I am on the editorial board of the *Journal of Comparative Policy Analysis* and of *Policy, Organization and Society*. I have a Ph.D. in Political Science from Indiana University. My curriculum vitae are presented in Appendix A.

I have been retained as an expert by defendant City of New York to render an opinion in this case. I have been compensated for this work at the rate of $250.00 per hour. My compensation is not dependent on the outcome of this case, the opinion I express or the testimony I provide.

2. Robert M. Purtell

I am an Assistant Professor of Finance at the University at Albany, Nelson A. Rockefeller College of Public Affairs & Policy. I have a B.S. in Mathematics from Manhattan College, an MBA in Finance/Economics from New York University's Stern Graduate School of Business, and a Ph.D. in Public Administration from New York University's Wagner School of Public Service. During the earlier part of my career, I held positions ranging from project manager at New York City's Human Resources

Administration to Assistant to the President at American Express to Director's level positions at several major Wall Street firms. For the ten years immediately prior to coming to Rockefeller College, I taught a broad range of finance courses at NYU. My curriculum vitae is presented in Appendix A.

My teaching and research interests focus on the application of managerial and corporate finance principals to the management and governance of nonprofit, healthcare and governmental organizations. Most recently, my research has focused on the related issues of fair-market pricing in the conversion of nonprofit organizations to for-profit status and the decision processes surrounding such decisions. I have also begun to focus on ways of applying variations on corporate-finance metrics to assess risk in nonprofits and public organizations.

I have worked with Dennis C. Smith, Ph. D. on various research projects and publications since 2006, including, but not limited to our working paper entitled *An Empirical Assessment of NYPD's "Operation Impact": A Targeted Zone Crime Reduction Strategy* (2007), and expert reports and opinions submitted in *Floyd v. City of New York*, 08 Civ. 1034 (SAS) and *Davis et al. v. City of New York et al.,* 10 Civ. 0699.

I have been retained as an expert by defendant City of New York to render an opinion in this case. I have been compensated for this work at the rate of $250.00 per hour. My compensation is not dependent on the outcome of this case, the opinion I express or the testimony I provide.

Dennis C. Smith and Robert M. Purtell hereby jointly submit this report in response to the Second Supplemental Report of Jeffrey Fagan, Ph.D.

B. Data Analyzed

We examined all of the data analyzed by Prof. Fagan,

- NYPD UF250 Stop, Question and Frisk data for the 1st through 4th Quarters of 2010 and 2011 and the 1st and 2nd Quarters of 2012. *See* Bates Nos. NYC 2 19404, 20606, 20633, 20737, 20744, 20780, 20788, 20789, 20790, 21439

- NYPD Crime Complaint Report Data for the 1st through 4th Quarters of 2010 and 2011 and thee and 2nd Quarters of 2012. *See* Bates Nos. NYC_2_20782, 20784, 20791, 20792, 21442.

We also analyzed data on suspects' race and ethnicity from the merged files of complaint and arrest reports, provided by the City also to the Plaintiff's expert, and the annual, public report of NYPD: "Complaint and Enforcement Activity in New York City (Jan. 1 –December 31, 2011," and an NYPD report, Raymond Kelly, "2011 Reasonable Suspicion Stops:Precinct Based Comparison by Stop and Suspect Description"

C. Issues Addressed and Design Modifications

1. Descriptive Statistics on Patterns of SQF Activity

The descriptive statistics are not a source of contention, but we do note some aspects of his findings which should garner more of his attention than he gives them. For example, we asked in our first report why no attention was given to the age and, especially, gender disparity on police *Terry* stops, but while that disparity remains in Prof. Fagan's demography of *Terry* stops and, even though it is unconstitutional to discriminate on the basis of gender and age, it goes without notice here. Acknowledging the fact that criminal activity is highly correlated with gender and age, as are race and ethnicity, would strengthen our argument that benchmarks used in disparity analyses should not be population based. That may explain this omission.

2. Disparate Treatment

Prof. Fagan begins his discussion of disparate treatment in the Second

Supplement Report with the statement:

> Plaintiffs claim that NYPD officers have used race and/or national origin as the factors that determine whether officers decide to stop a person, and that Black and Latino persons are the population groups most affected by the NYPD *Terry* stop practices.

Clearly, there are two separate claims here. The first focuses on individual officer decision making, and the second on the effects of an NYPD policy on Black and Hispanic/Latino persons. This is a useful distinction, one we will suggest later has applications other than assessing disparate treatment. Here it is sufficient to note that nothing in the analysis provided by Prof. Fagan supports the claim of individual officer bias influencing *Terry* stop decisions. Plaintiffs and Prof. Fagan may believe that police officer decisions to stop persons based on racial bias is real, but his efforts to date to support that belief have fallen far short. We noted in our report in the Davis case to the court a problem in Prof. Fagan's presentation of purportedly relevant social scientific findings:

> Prof. Fagan would further like us to believe that social science research supports such a notion and gestures to two academic articles to bolster his case. Neither article, however, actually supports Prof. Fagan's musings about the inner workings of police officers' psychologies. The first, Robert Sampson and Steven Raudenbush's 2004 article, "Seeing Disorder: Neighborhood Stigma and the Social Construction of 'Broken Windows'" is not, in fact, about police judgments. Rather, Sampson and Raudenbush focus on the causes and consequences of popular (and ill-informed) perceptions of neighborhoods. The second, Geoffrey P. Alpert et. al.'s 2005 "Police Suspicion and Discretionary Decision Making During Citizen Stops," explicitly contradicts Prof. Fagan's arguments regarding the role of race in officers' calculations. The detailed findings by these scholars, accordingly, merit repeating in this context. Research by Alpert and his colleagues revealed that, in contrast to their initial hypothesis and much popular belief, "minority status does *not* influence the decision to stop and

> question suspects" (emphasis added).[2] Prof. Fagan's arguments, then, can find no comfort in the scholarship he cites.[3]

In fact, common sense may be the best test of the proposition that individual NYPD officers are making stop decisions based on race and ethnicity rather than their professional assessments of observed behavior as suspicious. Prof. Fagan acknowledges that the practice on the part of NYPD of deploying police where the crime is highest results in high concentrations of officers in Black and Hispanic communities (relevant to plaintiff's claim of disparate impact) but the result of this policy decision is that in the neighborhood where most *Terry* stops occur there is very little variation on the race/ethnicity of persons police encounter on patrol. In our first report on Floyd in November, 2010, we presented calculations that showed the relative rarity of officers engaging in stops in the course of a month. On average officers make about two stops per month. In the Rand 2007 analysis of NYPD stops the officers who were "outliers" in their sample, were the ones who made more than 50 stops per year. The vast majority of officers made fewer than four stops a month, one a week.[4] In view of the tens or even hundreds of thousand of persons an officer observes in the course of each month on duty, when the overwhelming majority are Blacks and Hispanics, is it plausible to believe that the several persons actually stopped were selected based on their race?

To address the second claim "that Black and Latino persons are the population groups most affected by the NYPD Stop Question and Frisk (hereafter SQF) practices" is

---

[2] John M. MacDonald, Geoffrey P. Alpert, and Roger G. Dunham, "Police Suspicion and Discretionary Decision Making During Citizen Stops," *Criminology* 43, no. 407 (2005): 407).

[4] Greg Ridgeway, 2007.Analysis of Racial Disparities in the New York Police Department's Stop, Question, and Frisk Practices, Rand Corporation

evidence of racial bias in NYPD policing we must turn to the role of benchmarking in disparate treatment analysis.

a. Benchmarking

There are two main parts to our dispute with Prof. Fagan over benchmarking. One is the criteria for assessing benchmarks and what the social science literature says about benchmarking racial disparity in general, and in policing in particular, that calls into question the benchmark offered by Prof. Fagan. The second issue is Prof. Fagan's continuing and intensifying resistance to using in the analysis for the Floyd case (and other disparities in NYPD cases before the Court) the benchmark which he endorses in principle and in a scholarly publication, but in practice he repudiates. We will address each of these issues related to benchmarking in order.

Prof. Fagan's research design adopts a benchmark for measuring disparity that is unreliable and, as a consequence, out of step with current social science. Moreover, even for the class of benchmark Prof. Fagan employs—an adjusted census benchmark—his research design fails to incorporate specific factors called for by specialists in the field (and that Prof. Fagan himself has advocated elsewhere).

All disparate treatment claims—including the Plaintiffs' in *Floyd*—pivot on a contrast: practice X impacts members of group Y more so than similarly situated individuals who are not members of that group. And all such contrasts, in turn, invoke and depend upon a counterfactual point of comparison: what the outcomes of practice X would have been ***but for such group membership***. The injustice alleged by any disparate treatment claim is to be found in the difference between the observed outcomes of practice X in the real world and the postulated, counterfactual outcomes of practice X that would have occurred *had group membership played no role.* For disparate treatment claims that employ statistical evidence in their argument, these counterfactual points of comparison have come to be known as "benchmarks" or "baselines" in the now extensive social science literature addressing disparate treatment claims. Benchmarks help social scientists demonstrate statistically if there is an unjustifiable cause-and-effect relationship between membership in a group (say, African Americans) and a particular practice (say, *Terry* stops).

No other step in analyzing disparate treatment claims matters as much as the choice of benchmark. As should be obvious, choosing an inappropriate benchmark in an analysis will almost inevitably produce results that either misleadingly inflate or obscure disparate treatment and whatever injustice accompanies that disparity. Any analysis that starts with an inappropriate benchmark cannot be relied upon, regardless of how sophisticated subsequent aspects of that analysis might be. That the literature on benchmarks is so sizeable testifies not simply to the unique importance of benchmarking in disparate treatment claims but also to benchmarking's complexity.[5]

[5] This voluminous literature is reviewed in Greg Ridgeway and John MacDonald, "Methods for Assessing Racially Biased Policing," in *Race, Ethnicity, and Policing: New and Essential Readings*, ed. Stephen K. Rice and Michael D. White (New York: New York University Press, 2010). Notable articles include J. L. Lamberth, "Revised Statistical Analysis of the Incidence of Stops and Arrests of Black Drivers/travelers on the New Jersey Turnpike Between Exits or Interchanges 1 and 3 from the years 1988 through 1991," (Philadelphia: Temple University, Department of Psychology, 1994). M. Zingraff, Mason, H. M., Smith, W. R., Tomaskovic-Devey, D., Warren, P., McMurray, H. L., & Fenlon, C. R. et al. , "Evaluating North Carolina State Highway Patrol Data: Citations, Warnings and Searches in 1998," (North Carolina: North Carolina DEpartment of Crime Control and Public Safety and North Carolina State Highway Patrol, 2000); Sam Walker, "Searching for the Denominator: Problems with police traffic stop data and an early warning system solution," *Justice Research and Policy* 3(2001); Jeffrey Fagan, "Law, Social Science, and Racial Profiling," *Justice Research and Policy* 4, no. Fall (2002). John M. MacDonald Geoffrey P. Alpert, and Roger G. Dunham., "Police

A recent Department of Justice (DOJ) state-of-the-art "how to" guide for analyzing police data in disparate treatment highlights just how tricky the business of benchmarking can be. In situations as complex as benchmarking, the relationships between relevant factors often become more salient by constructing a simpler analogy and, accordingly, the DOJ publication asks readers to contemplate the following straightforward hypothetical. Imagine that parents of local high school students have become concerned that grading by the school's English teachers has an unfair disparate impact upon boys because the teachers—consciously or unconsciously—assign boys lower grades than they assign their female students. A disparate impact claim in this hypothetical would assume a conceptual model wherein gender (Y) has a causal impact on grades (X). To test this causal model statistically, however, we cannot construct a model that merely considers (Y) and (X). We cannot, for example, take the percentages of all A's and B's assigned to females in the class and compare them to a benchmark reflecting the percentage of the class population that is female (Such a benchmark is generally referred to in the literature as a "census" benchmark). Doing so could potentially obscure female students' English performance, creating the impression of unfair disparity in grading where none actually existed. If, for example, female students in the class preformed better in English, we would expect that teachers would assign them higher grades than they assigned to boys. In short, the existence of a gender

Suspicion and Discretionary Decision Making During Citizen Stops," *Criminology* 43, no. 407 (2005); JeffreyGrogger and Greg Ridgeway, "Testing for Racial Profiling in Traffic Stops from Behind a Veil of Darkness.," *Journal of the American Statistical Association* 101(2006); Andrew Gelman, Jeffrey Fagan, and Alex Kiss, "An Analyis of New York City's "Stop and Frisk" Policy in the Context of Claims of Racial Bias," *Journal of American Statistical Association* 102, no. 479 (2007); Ian Ayres, "Outcome Tests of Racial Disparities in Police Practices," *Justice Research and Policy* 4, no. Fall (2002).

disparity in grades alone would be insufficient to demonstrate disparate treatment in the hypothetical. Instead, a successful disparate treatment claim would require the demonstration of significant difference between the actual grades given and the grades that would have been given had gender played no role. That counterfactual, gender-meaningless situation might well still feature a disparity in the "raw" gender distribution of grades for justifiable reasons, including gender differences in English performance.

To construct an appropriate benchmark, as the Department of Justice authors argue, "our research design must 'neutralize' the impact of performance on grades." Doing so requires an external yardstick of female students' English performance that potentially-biased teachers have minimal opportunity to manipulate (say, SAT scores) so that we can control for female students' English performance in our benchmark. That control would allow us to measure more accurately the difference between the real-world grades the English teachers assigned and the counterfactual grades our benchmark suggested would have been assigned by teachers had gender played no role in their calculations.[6] To adjust this mental experiment—again, following the Department of Justice's how-to guide—so that it applies not to grades given by teachers but *Terry* stops conducted by police officers merely requires replacing each element of the hypothetical with its relevant analog for police stops. Assume that the local police jurisdiction (standing in for the local high school) has concluded that 65 percent of all *Terry* stops are of men, even though men represent only half of the population. Does this gender disparity

[6] This discussions draws heavily from Lorie A. Fridell, *By the Numbers: A Guide for Analyzing Race Data from Vehicle Stops* (Washington DC: Police Executive Research Forum, 2011), chapter 1. In *By the Numbers,* however, the author's hypothetical features grades in Math. In order to sidestep debates that Math grades are entirely objective while police actions can sometimes be subjective, I have changed the hypothetical to English. Doing so, obviously, does no harm to the hypothetical.

reveal that there are grounds for a constitutional disparate treatment challenge under the 14th Amendment? Many might pause before jumping to such a conclusion. Perhaps more men than women populate the streets during the hours when police officers are on patrol, increasing the probability that they will be stopped? Following the Department of Justice publication, this report dubs that element the *quantity factor*. Or perhaps more males than females walk in the areas where police are deployed, again increasing the probability that they will be stopped. This report refers to that element as the *location factor.* Finally, there is the possibility that males commit more crimes than do females and so are more likely to be observed by police engaging in suspicious activity that would justify a *Terry* stop. This report refers to that element of the model the *criminal participation factor*. This factor is analogous to female students' English performance in our first hypothetical (and the nature of substation here follows the recommendation of the Department of Justice how-to guide). In that mental exercise, if female students on average performed better than boys, we would expect female students to receive better grades (even absent a gender bias against boys). In much the same way, if males engage in more criminal activity, we would expect them to be subject to *Terry* stops more frequently than females (even absent a gender bias against men). A statistical benchmark that did not consider gender differences in the *criminal participation factor* when measuring gender disparities in *Terry* stops would be no more reliable than a benchmark that did not consider gender disparities in academic performance when measuring gender disparity in grades. (In point of fact, as decades of research have confirmed, crime is a young man's game, and—not surprisingly—the NYPD does conduct many more *Terry* stops of men than women. Indeed, as Prof. Fagan's report acknowledges in Table 3 of his Second Supplemental

Report more than 91% of all *Terry* stops are of males. This astounding disparity in stops, however, has aroused no disparate treatment claims—likely because of the common-sense understanding that males commit far more crimes than women.).[7]

The reader likely can predict the next step of the logical sequence: substituting race/ethnicity for gender in the above hypothetical. All three of the necessary elements for an appropriate benchmark—*location, quantity*, and *criminal participation*—remain, but now we apply these elements to racial and ethnic groups rather than to gender (and to *Terry* stops conducted by police officers rather than grades given by English teachers). Completing this final round of substitutions brings us to the situation analyzed by Prof. Fagan (and at the heart of the matter before the Court). How reliable are Prof. Fagan's conclusions regarding a pattern of unjustifiable disparity in the NYPD's practice of *Terry* stops?

Any assessment of the reliability of Fagan's conclusions must start by considering the appropriateness of his benchmark. Such a consideration, in turn, requires assessing how well Prof. Fagan's chosen benchmark addresses the three essential elements of disparity measurement in police activity that the Department of Justice how-to guide has specified (and that reflect the state of the field in measuring disparity claims).

---

[7] Again, this passage both draws heavily from and follows the logical steps of ibid., chapter one. For gender differences in crime participation, see Travis Hirschi and Michael Gottfredson., "Age and the Explanation of Crime," *The American Journal of Sociology* 89, no. 3 (1983).

Fagan's model employs an "adjusted census" benchmark that makes no effort to consider differences in "criminal participation" and, for a variety of reasons, falls short in its attempts to adjust for the "quantity" and "location" factors**.**

This report engages each of these issues in turn.

Differences in Criminal Participation

Attempts to measure racial disparity in *Terry* stops employing a benchmark that does not include differences in criminal participation by racial/ethnic groups can be relied upon, obviously, only in those contexts where no such differences exist. If, in contrast, there are significant racial/ethnic differences and the chosen benchmark does *not* address them, *any conclusion employing that faulty benchmark will itself be faulty.* (Just as any conclusions about gender disparity in grading would be faulty if the benchmark at the base of the analysis failed to consider gender differences in academic performance.)

Prof. Fagan does *not* argue in his various reports for the Court that racial/ethnic groups in New York City participate in criminal activity in rough proportion to their share of New York City's population. Instead, he argues that although using a benchmark that considered differences in the racial composition of the criminally active would be "ideal" [8], practical considerations make the available data regarding such racial differences impossible to use. That argument by Prof. Fagan raises two issues. First, since Prof. Fagan's benchmark does *not* address differences in criminal participation by race/ethnicity, does his proposed substitute benchmark adequately address the issue of racial differences in criminal participation? Merely that Prof. Fagan finds the available data untrustworthy (a contention we challenge below) does not mean that the racial

---

[8] Report of Jeffrey Fagan in David Floyd et al v, the City of New York, et al, dated October 15, 2010, p.17.

differences in crime participation cease to be critically relevant for measuring racial disparity or that his substitute benchmark is sufficient to demonstrate his claim.

For the Court to rely upon Prof. Fagan's analytic conclusions, he must demonstrate that his benchmark maps reasonably well onto the real world. If his benchmark does not address differences in criminal participation by racial/ethnic group, then we must conclude his benchmark assumes that New York City's racial/ethnic groups *do* participate in crime at rates close to their respective share of the city's population. If his benchmark assumes there are not meaningful differences in criminal participation by racial/ethnic groups, he must demonstrate the empirical accuracy of that assumption. In the absence of such a demonstration, his conclusions cannot be relied upon.

Second, Prof. Fagan is fundamentally wrong in his rejection of the Defense's proposed benchmark for measuring disparity in *Terry* Stops—one that does address differences in criminal participation by racial/ethnic groups—for at least three reasons. Although we will address these issues in greater depth in below, they are summarized briefly here.

1. Prof. Fagan rejects the use of data on racial/ethnic differences in crime participation because, as he (accurately) notes, we have little racial data for minor property crimes from victim identification of suspects. However, neither the NYPD's interdiction efforts nor the public's fears and demands for service focus on such minor crimes, but rather on violent crimes, such as rape, robbery, murder, and assault. And for *those* crimes our knowledge of the racial/ethnic distribution of suspects is nearly complete: for violent offenses, the race of 85.4% of all suspects is known. Likewise, for trespass we know the

race of 96.8% of suspects; for disorder, 73.1% of all suspects; and for gun play, 97.9%. Only for property crimes—where victims rarely see the perpetrator—is our knowledge lacking. (Fagan, it should be noted, misrepresents the extent of our knowledge about the race of violent offenders when he claims "The rate of suspect identification for this crime type also is low compared to other crime types – 6.8%." In reality, we know the race of violent offenders in 85.4% of cases.) [9]

2. Even for the types of crimes where our knowledge about racial/ethnic participation is more limited, the share of crimes committed by each racial/ethnic group closely matches their respective participation rates in the types of crimes for which we have significant data. That racial/ethnic participation in crime is consistent, both in cases where we have a nearly complete picture and where we have only a fragmentary snapshot, suggests the accuracy of the data even with a smaller sample.

3. Fagan dismisses the data on suspect-race description from the "merge" file on a variety of technical objections. These objections, however, lack merit because they are grounded in a fundamental misunderstanding of the nature of the "merge" file. Prof. Fagan's putative objections include his doubts about reliability of arrest data and concerns over "spatial" and "temporal" matching. We will address these objections in detail below.

The Need for a Criminal Participation Benchmark

Why is including criminal participation by race/ethnicity essential to a benchmark

[9] Appendix B in the Second Supplemental Report.

in measuring racial or ethnic disparity in NYPD *Terry* stops? Recall our initial hypothetical. As explained above, if female students preformed academically, we would their English teachers to give them better grades. Only if the teachers gave better grades at a rate *over and above* what could be explained by female students' better performance compared to boys could we conclude there were grounds for a disparate treatment claim. Accordingly, in the hypothetical, we needed a benchmark that potentially-biased teachers had no control over that would provide both an independent assessment of female students' performance and could serve as a counterfactual point of comparison for the grades teachers assigned. One such independent benchmark for our hypothetical situation might be SAT test scores; since teachers have no way to manipulate such data, we can assume the scores are free of their (potential) bias. That benchmark could "neutralize" (to use the Department of Justice's argument addressed above) the impact of female students' better academic performance.

Fortunately, we have an analogous benchmark for differences in criminal participation by racial and ethnic group in New York City. The advantage of this external benchmark for measuring disparity in *Terry* stops is that it allows us to neutralize in our analysis the potential impact of differences in criminal participation. Indeed, so useful is this benchmark for measuring disparity that Prof. Fagan as recently as 2007 praised its use in disparate treatment claims (see below). This powerful benchmark is "suspect race description" data. A suspect race description is simply a victim's description of the suspect in a crime as reported to the police. Because this benchmark emerges from victims and not the police or police action, it is much less vulnerable to their manipulation and potential biases than is arrest data (much as SAT scores are independent of potentially-biased teachers).

Obviously, if particular racial or ethnic groups in New York City participate in crime at a rate disproportionate to their share of the population, we would expect officers to conduct *Terry* stops for such groups at rates higher than each groups' respective share of the City's population. The benchmark of suspect race description allows us to measure if NYPD's officers are stopping minorities at a rate *over and above* what could be explained by the racial composition of the criminally active population in New York. In contrast, Prof. Fagan's model *cannot* do so and so *cannot* be relied upon.

Prof. Fagan's Initial Praise and Later Rejection of Suspect Race Description Criminal Participation Benchmark

As noted above, as recently as 2007 in a peer-reviewed, published article on NYPD *Terry* stop practices, Prof. Fagan and his co-authors specify what they characterize as a "more relevant comparison" to use in assessing racial disparity than a census benchmark: "the number of crimes committed by each ethnic group." They cite in support of this assertion Police Commissioner's Howard Safir's claim that "the racial/ethnic distribution of the subjects of *Terry* stop reports reflect the demographics of the known violent crime suspects as reported by crime victims." Without indicating that any attempt was made to ascertain the demographics to which the Commissioner referred, the authors claimed[10]:

[10] Andrew Gelman, , Jeffrey Fagan, and Alex Kiss. "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, no. 479 (2007): 16..

> Data on actual crime are not available, of course, so as proxy we use the number of arrests from the previous year, 1997, as recorded by the Division of Criminal Justice Services of New York State, categorized by ethnic group and crime category. This was deemed to be the best available measure of local crime rates categorized by ethnicity and directly address concerns such as Safir's that stop rates be related to the ethnicity of crime suspects.

Prof. Fagan now, and in all of the litigation before this court related to racially-based disparate treatment claims arising from the NYPD's practice of *Terry* stops, insists that crime and population statistics are the two variables that, when combined, constitute the "best available measure," or benchmark for measuring disparity. His previous belief in the essential value of accounting for differences in criminal participation has mysteriously vanished.

Although Prof. Fagan and his colleagues expressed disappointment in their inability to use "the number of crimes committed by each ethic group" in their analysis in the 2007 article, he has never requested suspect data from the NYPD. Moreover, he has rejected its use throughout the litigation on *Terry* stops—even as such data has become available, with increasing detail, over time. The arrest data that Prof. Fagan used in his 2007 article was available when he wrote his *Floyd* Report, but by 2010 the weaknesses in arrest data alone was too well known for that measure to be selected again.

In his Second Supplemental Report, Prof. Fagan summarizes the basis for now rejecting what he had previously as "the best available benchmark":

> First, only files for 2010-11 are included. The analyses in this report cover a 30 month period from January 2010 — June 2012. The 2012 period is critical since there has been a decline in the number of stops since March 2012.[13] Next, the procedure for matching arrests to crime complaints relies only on time matching; there is no indication in the documentation accompanying these 'merge' files that a spatial matching procedure was used as well. And the time matching procedure itself includes anomalous findings and highly skewed distributions on the time

> concurrence.[14] Third, the validity of arrests is uncertain, given observed rates of prosecutorial declination and other reasons for non- conviction of cases. The files present arrests as if they were completed crimes. Fourth, the information gain over the use of crime complaints is trivial relative to the distribution of crimes by race. In other words, the distribution of incidents by suspect race (arrest plus victim-identified suspect) reflects the distribution of arrests by suspect race without the marginal addition of information on arrestees.

Each of Prof. Fagan's objections above deserve scrutiny.

Issue #1, the Significance of Missing Data

Prof. Fagan writes:

> First, only files for 2010-11 are included. The analyses in this report cover a 30 month period from January 2010 — June 2012. The 2012 period is critical since there has been a decline in the number of stops since March 2012.[1]

Although ideally the data on suspect description would be available for all thirty months of the study, they *are* available for twenty-four of those months. Prof. Fagan offers no explanation why having the best available measure for only 80% of the study period somehow makes the data unusable. Perhaps if the racial composition of the criminally active population in New York fluctuated meaningfully during the twenty-four months for which we have data, one could argue that the missing data mattered. Instead, however, the analysis of the months and years available show little variation in the pattern over time.

Throughout the period for which data are available, African American and Hispanics are found consistently to be disproportionately identified as perpetrators in virtually all categories of crime. Since 2008, the NYPD has issued reports on Crime and Enforcement Activity in New York City that includes analyses of the known characteristics of persons identified as suspects as well as of arrestees. Our December 2012 report for Floyd included city-wide data. Likewise, Table 1 shows remarkable continuity in the percent of Black and Hispanics suspects identified by victims, and the same consistency can be observed in the concentration of Blacks and Hispanics in the arrestee column. These two measures are incorporated in the "merged" suspect and arrest data that is now available for use as a benchmark to gauge the racial composition of the criminally active population in New York City and so represents what Fagan once called the "ideal" benchmark. Excluding the "best available data" would appear to require more than an expert's assertion that that not having all months of a study period is "critical.

Table 1. Distribution of Distribution of Victims by Race Compared to Suspects in Violent Crime Reports, by Race, 2009/ (2011)

Attributed Race of Suspect Compared to Share of Victim Population	Black (24 % of population*) **(23%)***	White (35 %) **(33%)**	Hispanic (28 %)) **(29%)**	Total number of crime victims in these categories- 2009/**(2011)**
Murder and non-negligent homicide	57.6%/**59.8%** 61.8%/**(56.3%)**	9.6%/**5.5%** 8.4%/**(5.5%)**	28.9%/**31.4%** 26.3%/ **(35.0%)**	453 **(515)**
Rape	40.5/ **52.4** 36.8/ **(48.8)**	14.7/**7.6** 17.8/ **(11.2)**	39.3/**36.6** 39.6/ **(34.5)**	1,005 **(1,412)**
Robbery	31. 0/**70.6** 31.7/ **(70.6)**	18.0/**4.3** 18.5/ **(3.9)**	38.5/ **23.8** 36.1/ **(24.0)**	18,602 **(19,764)**
Felonious Assault	46.7/**54.3** 47.8/ **56.3**	12.1/**7.6** 12.8/ **8.9**	35.5/**33.5** 33.4/ **30.1**	16,768 **(18,600)**
Grand Larceny	23.9./**62.4** 24.1/ **(63.6)**	44.7/**11.4** 43.1/ **(9.9%)**	20.0/**23.3** 20.2/ **(24.0%)**	38,877 **(38,800)**
Shooting Victims	72.8/**79.8** 74.3/ **(72.5)**	3.1/**1.4** 2.6/ **(2.5)**	23.0/**18.3** 21.7/ **(23.9)**	1,729 **(1,818)**

Source: Source: NYPD, Crime and Enforcement Activity in New York City (Jan 1-December 31, 2009/ Jan 1- December 31, 2011)

Issue #2, the Mechanics of the "Merge" File

Prof. Fagan writes:

> Next, the procedure for matching arrests to crime complaints relies only on time matching; there is no indication in the documentation accompanying these 'merge' files that a spatial matching procedure was used as well. And the time matching procedure itself includes anomalous findings and highly skewed distributions on the time concurrence.

Prof. Fagan's second basis for rejecting suspect description data appears to be based on a quite inaccurate understanding of the procedure used by NYPD to create an unduplicated tally of suspect descriptions. Prof. Fagan acknowledged in his 2007 article that both crime complaints and arrest data are sources of information regarding the racial composition of the criminally active population of New York. They are different, neither is complete, and they are interrelated. Since some "suspects" are arrested, simply adding the two would result in duplication in counts. As was described in detail for the Plaintiffs by NYPD's Crime Analysis Unit, the City's Omni Record System was used to systematically link complaints with arrests. This linking procedure must accommodate a wide range of permutations between suspect and arrest data: some crimes never result in a arrest; even when an arrest does occur, it may happen long after the reported crime; some multiple-perpetrator crimes may not result in as many arrests as there were suspects. For some crimes there are more arrests than suspects (say, a bodega robber identified by victim and a lookout or driver arrested with the robber). Rather than update potential linkages continuously and endlessly into the future, the NYPD employed a 24-hour cut off for linking arrests to specific crime, adding those previously offenders' characteristics to suspect characteristics when the former was unknown from that specific complaint. Since a relatively high proportion of the City's crime arrests take place within twenty four hours of a complaint, the NYPD's selected method *does* give an incomplete picture, but not in one that would obviously skew racial or ethnic data. Indeed, this approach is inherently conservative and likely leads to an undercounting of known suspect race. Recall, there is now and has long been a high correlation between the racial distribution of suspects in criminal complaints and in arrest data. A primary value of the merged data is to counter complaints that only a small fraction of crime perpetrators can be identified in terms of their race or ethnicity. As Table 2 shows, for virtually all categories of crime a large fraction of crime perpetrators' race and ethnicity is known.

Table 2.

PERCENT OF CRIMINAL OFFENDER WHOSE RACE/ETHNICITY IS KNOWN FROM SUSPECT DESCRIPTION OR ARREST, BY CATEGORY OF CRIME (Merged unarrested suspects and arrested suspect information for 2011)

Examining the crime complaint records and the information they contain describing Unarrested suspects as well as information on arrests associated with the complaint reveals:

All complaint reports recorded as occurring in 2011	494,551
Complaints with no unarrested suspect or arreseet records	123,907
Complaints with at least one unarrested suspect or arrestee record	370,644
% of complaints with at least one unarrested suspect or arrestee	**74.9%**
Number of unarrested suspects or arrestees associated with the 370,644 complaints	425,197

Violent felony complaints occurring in 2011	39,986
Violent felony complaints with no unarrested suspect or arrestee	161
Violent felony complaints with at least one unarrested suspect or arrestee	39,825
% of violent felony complaints with at least one unarrested suspect or arrestee	**99.6%**
Number of unarrested suspects or arrestees associated with the 39,825 complaints	57,525

(Violent felony complaints include murder, rape, robbery and felonious assault)

Robbery complaints occurring in 2011	19,725
Robbery complaints with no unarrested suspect or arrestee	57
Robbery complaints with at least one unarrested suspect or arrestee	19,668
% of robbery complaints with at least one unarrested suspect or arrestee	**99.7%**
Number of unarrested suspects or arrestees associated with the19,668 complaints	32,780

Assault 3 and Related offenses complaints occurring in 2011	50,832
Assault 3 and Related offenses complaints with no unarrested suspect or arrestee	66
Assault 3 and Related offenses complaints with at least one unarrested suspect or arrestee	50,766
% of Assault 3 and Related offenses complaints with at least one unarrested suspect or arrestee	**99.9%**
Number of unarrested suspects or arrestees associated with the 50,766 complaints	58,301

Oth. Fel.Sex Crimes & Misd. Sex Crimes complaints occurring in 2011	5,263
Oth. Fel.Sex Crimes & Misd. Sex Crimes complaints with no unarrested suspect or arrestee	839
Oth. Fel.Sex Crimes & Misd. Sex Crimes complaints with at least one unarrested suspect or arrestee	4,424
% of Oth. Fel.Sex Crimes & Misd. Sex Crimes complaints with at least one unarrested suspect or arrestee	**84.1%**
Number of unarrested suspects or arrestees associated with the 4,424 complaints	4,894

Issue #3, The Reliability of Arrest Data

Prof. Fagan writes:

> Third, the validity of arrests is uncertain, given observed rates of prosecutorial declination and other reasons for non- conviction of cases. The files present arrests as if they were completed crimes.

It is not clear what Prof. Fagan mean by “completed crimes.” It is very clear that in Prof. Fagan’s own work he has not refrained from analyzing crime data until the criminal suspects represented in the data were duly convicted; data on crimes has been sufficient for him and it sufficient for us as well. Prof. Fagan appears to insist, selectively, on a world of perfect information as suits his purposes. Such a world is far removed from the world of police, from the world prosecutors, and rather far, for that matter, from the world of Prof. Fagan’s own research. Nearly all measures available to scholars (and courts) are flawed in some degree or another. Indeed, that sea of flawed measures includes the crime statistics and census data that Prof. Fagan uses comfortably as a benchmark rather than employing victims’ own descriptions of suspects and arrest data. How do the flaws in the indices Prof. Fagan uses compare to the estimates and direction of bias in the data he challenges? For example, as we will note elsewhere in this report, census tract level data has been found to be unstable and the literature cautions against using that data for that reason.[11] In yet another example of Prof. Fagan’s selective setting of statistical thresholds, Prof. Fagan uses in his regression analysis his own, untested (and as we argue below, highly questionable) measure of patrol strength. His use of a single patrol strength

[11] Stanley K. Smith, “Tests of Forecast Accuracy and Bias for County Population Projections,” (*Journal of The American Statistical Association* 82:991-1003, 1987) and Stanley K. Smith & Mohammed Shahidullah, “An Evaluation of Population Projection Errors for Census Tracts,” (*Journal of The American Statistical Association* 90:64-71, 1995).

measure in the absence of other tests falls well short of common practice in the social sciences. Social scientists have long been advised to employ to multiple measures to counter concerns about the validity and reliability of measures. In the case of suspect description data we contend that the high correlation of suspect description data constitute multiple measures of the demography of crime perpetrators. Another multiple measure is the consistency of crime demography over time. When pollsters' survey results are challenged, analysts are reassured when subsequent surveys provide similar results.

Prof. Fagan's use of prosecutors' decisions to "decline to prosecute" (DTP) as evidence of flawed police practice and flawed police practice alone is questionable in light of the extensive criminal justice literature on prosecutorial discretion. Given the breadth of reasons prosecutors might opt to "decline to prosecute," those DTP decisions might signal faulty police work—or they might signal a half dozen other reasons that prosecutors chose DTP. Beyond the courtroom adage that an effective prosecutor "can get a grand jury to indict a ham sandwich," criminal justice scholars have long recognized that there are not only a great many reasons for DTP but that prosecutors enjoy great latitude in exercising that option. Scholars, for example, have noted that as elected officials, prosecutors have been known to employ their DTP discretion to nullify laws in response to popular sentiment.[12] This dynamic may explain the surprising variation in DTP across New York City's five district attorney's offices. Data from the New York State Division of Criminal Justice Services, for example, reveal that Robert Johnson, the Bronx's district attorney, declined to prosecute 23.4 percent of criminal cases in 2011, or nearly twice the next-highest rate of 12.1 percent,

[12] Peter Krug, "Prosecutorial Discretion and Its Limits" *The American Journal of Comparative Law* 50, (Autumn, 2002): 643-664

in Staten Island. In comparison, the District Attorney's office in Brooklyn did not prosecute 7.6 percent of cases; Queens, 5.8 percent; and Manhattan, 4.8 percent.[13] Clearly, forces in addition to faulty police practice are at work in an index that varies by more than 500% across jurisdictions in a single city with one police force. It strains credulity to assume that DTP always and only indicates failure in police practices. Prosecutors, like other public officials, have scarce budgets that can foster a form of criminal justice triage. Take, for example, an episode from Gotham's recent history. New York's Transit Police Depart under William Bratton adopted a strategy of fighting more serious crime in the subways by targeting fare beaters. Bratton assumed that criminals entering the subways to ply their trade rarely wanted to pay the fare. He was right (as research has since confirmed), but prosecutors were not initially eager to see their scant resources to prosecute theft of a subway $1.15 fare.[14] In much the same fashion, trespass arrests—which police believe are instrumental in preventing more serious crimes that require trespass—are prime candidates for rejection by prosecutors for reasons unrelated to the quality of the arrests.

Issue #4, the share of suspects for which race is known

Prof. Fagan writes:

> Fourth, the information gain over the use of crime complaints is trivial relative to the distribution of crimes by race. In other words, the distribution of incidents by suspect race (arrest plus victim-identified suspect) reflects the distribution of arrests by suspect race without the marginal addition of information on arrestees.

---

[13] Winnie Hu, "Criminal Prosecutions at a Low Rate in the Bronx, a Report Finds," (*New York Times,* 21 June, 2012)

[14] William Bratton, *Turnaround: How America's Top Cop Reversed the Crime Epidemic,* (Random House, 1998). Jack Maple, *Crime Fighter: Putting the Bad Guys Out of Business*, (Random House, 2000).

This fourth basis for rejecting the use of the merged suspect data ignores Prof. Fagan's repeated challenge to suspect data on the grounds that there are too many crimes for which the suspect is unknown. As I noted in an earlier report, Prof. Fagan has offered many and varied responses to the question of what percentage of suspects' race can be known and how from existing data—even as he consistently understates the actual percent that is known. The addition of arrest information to the victim-identified suspect information has no marginal value in terms of the measuring the racial composition of New York's criminally active population because the racial patterns are the same in both the suspect identification data and the arrest data. The addition of arrest information, however, *does* address the completeness of the evidentiary base of these assessments, something Prof. Fagan has demanded. This issue leads to Prof. Fagan's "additional" objections to merged file data:

> Two other considerations bear on the decision to use the "merge" files or the separate arrest and crime complaint files. The suggestion to use the number of violent crime suspects as the benchmark ignores the fact that violent crimes account for only a small percentage of the suspected crimes in the stops in 2010-12: as shown in Section II of this report, of the 1,624,410 stops in 2010-12, the suspected [15]crime was either major or minor violence in 23.90% of the stops. Many more stops were for either major or minor property offenses (24.95%) or weapons offenses (24.95%). In the October 2010 Report, I caution against the use of a crime benchmark with high levels of missing observations.15 Also, as shown in Appendix Table 2 *infra,* the suspect race in the "merge" file for suspected either major or minor violent offenses is known in 39.8% of all known suspects in 2010-11, and in 25.3% of all crime complaints in 2010-11. Again, the large majority of known suspects are for crimes other than violence.

[15] Dennis C. Smith, "Police," in *Setting Municipal Priorities*, (Charles Brecher and Raymond D. Horton, eds., New York: Russell Sage Foundation, 1982). This chapter cautioned that the neglect of less serious crimes may be sowing the seeds of a future crime increase, anticipating the theory of George Kelling and James Q. Wilson that order maintenance crime contributed to the effort to control more serious crime. (James Q. Wilson and George Kelling, "Broken Windows" *The Atlantic Monthly, March 1982).*

In his "other considerations," Prof. Fagan again demonstrates a lack of familiarity with the mission and strategies of the NYPD—no small matter when his task is to create a model that maps onto the world of police officers. The NYPD, like most big city police departments in America, gives priority to violent crime. That priority is part of its tradition and is built into its operating systems. Citizen calls to 911, for example, are not treated on a first come, first serve basis. Instead, violent crime goes to the front of the queue. Similarly, a study of how the NYPD managed a 25% reduction in manpower between 1975 and 1980 in the wake of the City's fiscal crisis, found the Department reduced dramatically attention to non-violent, less serious crime—a fact evident in the dramatic decrease in misdemeanor arrests during those years. As a result, during most of the period of reduced ranks, violent index crimes, especially murder, rape, robbery, and assault, did not increase dramatically. Likewise, in the proactive police management of NYPD of the past two decades, the focus on violent crime has been made quite explicit, even as the tactics have understandably changed. As Jack Maple, William Bratton's Deputy Commissioner, explains in *Crime Fighter: Putting the Bad Guys out of Business*, quality-of-life law enforcement, while worthwhile in its own right, is of greatest use as a means to an end[16]—and that end is the reduction of violent crime.

For the past decade, the dominant policing strategy of NYPD has been hot spot policing—a strategy implemented through the NYPD's Operation Impact program. NYPD has explained that the persistence and concentration of violent crime has led to a strategy of deploying focused police attention to those places where, despite dramatic

[16] Jack Maple, *Crime Fighter: Putting the Bad Guys Out of Business*, (Random House, 2000).

crime reduction in all of the City's precincts, crime remains stubbornly high. Within Operation Impact zones, designed to reduce violent crime, all levels of crime receive additional scrutiny and enforcement. A leading criminologist, Frank Zimring, has concluded that Operation Impact—analyzed in detail alongside other rival hypotheses in *The City That Became Safe*—is the best explanation for the City's unique crime decline that has been longer and deeper than elsewhere in urban America.[17] Similarly, in a recent article, David Weisburd and his colleagues demonstrate statistically that the NYPD concentrates *Terry* stop activity where violent crime is concentrated—and generally in areas no larger than a few city blocks.[18] Accordingly, Prof. Fagan's concern with the comparatively limited knowledge of the racial composition of those engaging in minor crime in the city is misplaced.

There is further evidence that our limited knowledge of the racial demographics of minor criminals is of narrow significance and so can not justify, as Prof. Fagan's wishes it to, a rejection of suspect-race data as a benchmark. The partial data we *do* have regarding the demography of individuals engaged in minor criminal activity closely resembles the demographics of violent criminals for which we have rather complete information. In short, whether we take a 97.9% slice of the criminal pie (as in the case of weapons offenses) or a 23.1% slice of the criminal pie (as in the case of

---

[17] Franklin E. Zimring, *The City that Became Safe: New York's Lessons for Urban Crime and Its Control (Studies in Crime and Public Policy*), (New York: Oxford University Press, 2012).

[18] David Weisburd, Cody W. Telep and Brian A. Lawton, "Could Innovations in Policing have Contributed to the New York City Crime Drop even in a Period of Declining Police Strength?: The Case of Stop, Question and Frisk as a Hot Spots Policing Strategy,"*Justice Quarterly*, (2013), doi:10.1080/07418825.2012.754920

property crimes), the resulting slices produce the same result. And even with the partial gaps in our knowledge because of the low percentage of property crime suspect for whom we have racial or ethnic data, we still know the race or ethnicity for 63.3% of all the criminally active population in New York. And let us remember why we lack some data where we do: for some crimes like burglary and auto theft, the victim rarely sees the suspect. But when the police do arrest suspects for such crimes, the racial composition of that pool of arrestees is very similar to the racial composition of offenders in New York for whom we have excellent data.

To dismiss criminal participation as a benchmark on the basis of missing information in limited areas of the suspect-race data is to ask the Court to ignore what Prof. Fagan has called the "ideal" benchmark on the off-chance that this rather unlikely scenario is true: there is a large population of unarrested white property criminals loose in Gotham who, year after year, have miraculously managed to evade the detection of both the police and their fellow citizens despite operating almost entirely in minority neighborhoods. And all the while, Prof. Fagan asks us to believe, a comparatively small number of black and Hispanic property criminals manages to get arrested relentlessly. Setting aside the absurdity of what Prof. Fagan would like us to believe, he provides no theoretical or evidentiary-based argument to support his qualm about assuming that the demography where suspects have been identified at higher rates are significantly different from those where they are not.

While the numbers change over the course of Prof. Fagan's continuing philippic against using suspect data as a benchmark, he has consistently claimed the percentage of crime perpetrators whose race or ethnicity is known as lower than what the NYPD reports and what the data we have demonstrates. See Table 2. In support of Prof. Fagan's "caution against the use of a crime benchmark with high levels of missing observations," he repeatedly misreports, more specifically, underreports, the percent of crime where the race/ethnicity of suspects is known.

Prof. Fagan's elaborate attack on the appropriateness of using suspect data as a benchmark has a precedent in his reports related to the *Floyd* case. While he did not devote an entire Appendix to his argument as he has in the Second Supplemental Report, in his February 2012, Declaration Prof. Fagan included an extensive and emphatic justification of his decision not to incorporate into his analysis a coded version of handwritten information provided by officers in the UF-250 form. As was learned in his testimony in the Daubert hearing, Prof. Fagan used a sample of 1,000 forms for which he relied on his own expert opinion to determine whether the data provided by officers in the sections of the form where "Other" could be checked, and specification supplied, met his standards of evidence as a scientific expert. What Fagan asserted next merits quoting at length:

> 9. Any attempt to code and analyze unique handwritten narrative details which officers may have entered on the UF-250's when they checked off the "other" stop circumstances on page 1 of the form could suffer from multiple sources of error. First, translation of incomplete sentences, shorthand notation with no obvious plain text meaning, and other uninterpretable scribble would introduce a level of subjectivity into the analysis that would render any meaning of these strings as unreliable.
>
> 10. Unlike the Stop Circumstances check boxes, these handwritten notations have no known inter-rater reliability. That is, the same or similar utterances ay have very different intended meanings depending on, among other things, the situation and experience of the officer. Analyses that attributed the same meaning to such similar utterances would risk errors since there is way to ascertain agreement among different officers as the meanings of these utterances. No such dilemma exists among the check

> box circumstances, where training and feedback should create a shared meaning of these established categories.
>
> 11. Second, there is no method to ascertain when and how these utterances are recorded in terms of the consistency from stop to sop. The recording of these utterances are subject to influences that may render them fragile and sensitive to the conditions in which the utterances are recorded, including the officer's level of emotional arousal or other mental state factors at varying time points after the conclusion of the encounter. This matters for the reliability of these utterances, since the meaning of an utterance during a field stop or immediately afterward may differ from its meaning if an officers completes the UF-250 form at some further point in time after the stop is concluded when emotional arousal or cognitive sharpening has faded.
>
> 12. Accuracy of recall is an enduring validity threat in the compilation of such data, with threats such as telescoping and cognitive distortions introduced when an interaction has been particularly salient, or when the officer is in a state of arousal from an encounter with a suspect. *See* Jennifer Roberts, Edward P. Mulvey, Julie Horney, John Lewis and Michael L. Arter *A test of Two Methods of Recall for Violent Events*, 21 Journal of Quantitative Criminology 175-193 (2005); Laura Campos and Maria L. Alonso-quecuty, *The Cognitive Interview: Much More than Simply "Try Again",* 5 Pscyhology, Crime & Law 47 (1999)
>
> 13. *Therefore, there is no statistically reliable way to discern reliable statistical categories from handwritten notes, or any sample of them, that would generate meaningful information about 2.8 million stops. Any attempt to do so would invite a host of potential biases and errors and would render any conclusion statistically meaningless*. (emphasis added).

These statistical truths, however, turned out not to be truths after all, according to Prof. Fagan. In his analysis of "the Constitutional Justification of Stops on the Basis of Information in the SQF Database" in his Second Supplemental Report, Prof. Fagan extensively analyzes data derived from the coded handwritten responses on the UF250 form without even addressing, much less explaining, how he has avoided the "host of potential biases and errors and [that] would render any conclusion statistically meaningless."

The "Quantity" Element of Disparity Measurements and Prof. Fagan's Report

Although Prof. Fagan's analysis does not use a simple headcount of individuals by race/ethnicity as his benchmark (a "census" benchmark) but instead employs an "adjusted census" benchmark that controls for various demographic and environmental factors, Prof. Fagan's efforts fall well short of what is considered standard practice in the now well-established field of disparity assessments.[19] To reference the essential elements of benchmarking specified by the aforementioned Department of Justice how-to guide for racial disparity measurements in police activity, Prof. Fagan's benchmarks fails to model properly the "quantity" of persons of each ethnicity that might engage in the targeted behavior and so might the subject of police exercise of their *Terry* authority. Prof Fagan's analysis assumes that the demographic make-up the geographic areas subject to police *Terry* stops remain static through the day, unchangingly reflecting the statistics to be found in the census data at the base of his calculations. But as any urbanite knows—and as academic research has consistently confirmed—the racial demographic of streets and neighborhoods frequently change by time of day. Some of this change reflects work patterns; the tidal waves of commuters flowing into the City in the morning and ebbing out in the evening mean that official census data would be a poor predictor for sidewalk demographics at lunch hour. Some of this reflects leisure patterns: video studies have shown that the share of minority residents on urban sidewalks increases significantly later

---

[19] Jeffrey Grogger and Greg Ridgeway, "Testing for Racial Profiling in Traffic Stops from Behind a Veil of Darkness," *Journal of the American Statistical Association* 101, no. 475, (September 2006): 878-887; Clayton Mosher. "Racial profiling/biased policing" *Sociological Compass* 5 (2011): 763–74; R. Tillyer, R. S. Engel, and J. Wooldredge, "The Intersection of Racial Profiling Research and the Law," *Journal of Criminal Justice* 36 (2008): 138–53.

in the evening. As a consequence, these well-understood truths of big-city life weaken any analysis that employs a population and crime rate- based benchmark. As one specialist lamented in a 2010 state-of-the field survey of racial disparity methodology, adjusted census benchmarks such as the one Fagan has adopted are "crude approximations of the population at risk for police contact" which "can lead to exaggerated estimates of racial bias."[20] In short, the reports Prof. Fagan has submitted to this Court are out of step with current science and so cannot be relied upon.

Prof. Fagan's failure to adjust his census benchmark to shifting demographics by time of day acquire become additionally damaging to the reliability of his conclusions because the hours that sidewalks and streets *least resemble* official census are the very hours when police are *most likely* to conduct Terry stops. See Appendix B. Accordingly, many real life officers on patrols will, in fact, encounter significantly more suspicious behavior than the population numbers Fagan relies upon would suggest. This phenomenon would artificially inflate the *Terry* stop rate his report claims. It is for precisely this reason that the social science literature on disparity measurement has long called for either abandoning the use of census-based benchmarks or adjusting them as required for ambient population shifts.

The mismatch between the actual demographics of the population on New York's sidewalks and the official tally from census records can be seen quite clearly in a few examples from the SQF database. Equally visible in these examples is the threat such mismatches pose to the reliability of conclusions drawn using census tract data. Take, for

---

[20] G. Ridgeway and J. MacDonald, "Methods for Assessing Racially Biased Policing." *Race, Ethnicity, and Policing: New and Essential Readings*, eds. S. K. Rice and M. D. White. (New York: New York University Press, 2010): 180 – 204, p.181.

example, the census tract identified as “6111300” in Prof. Fagan’s data. This census tract has a total black population of 36, which represents 18.4% of the tract’s residents. And yet in 2010-2011, individuals victimized in that tract identified African Americans as the suspect(s) an average of 33.8 times a month that year, representing 53.6% of the crime complaints where the race or ethnicity of the suspect was known. Likewise, the police conducted on average 146.3 *Terry* stops of African Americans per month that year in the tract. Clearly, unless all 36 of the census tract’s African Americans—regardless of their age or gender—are committing nearly a crime every month and being stopped four times a month by the police—many of the African Americans identified in the crime data as well as in the NYPD’s SQF database live elsewhere and have travelled to tract 6111300. Of course, comparing those 146.3 monthly *Terry* stops of African Americans to the official census numbers, even adjusted for the various factors Prof. Fagan does, would lead us to conclude that African Americans were being stopped by the Police disproportionately and so had grounds for a disparate treatment claim. And it is precisely such thinking that renders Prof. Fagan’s benchmark faulty, as it makes a number of clearly false assumptions. First, Prof. Fagan’s benchmark assumes that the racial demographics of a tract remains unchanged throughout the day (despite an extensive social science literature and every New Yorker’s lived experience). Second, Fagan’s benchmark assumes that the only individuals police encounter while on patrol or responding to a radio run will be those appearing as residents in the census. Neither assumption is supported by social science or common sense. Nor is census tract 6111300 unusual. 260 census tracts in New York have no African-American residents but have seen *Terry* stops of black individuals; likewise, in 250 of those 260 tracts, victims have identified African Americans as the perpetrator of a crime. Each of those hundreds of tracts would misleadingly inflate the disparity in Terry stops by race and ethnicity in Prof. Fagan's model employing an adjusted census benchmark.

Because such significant errors can creep in easily into any model using a census or even an "adjusted" census benchmark, a recent state-of-the-field article notes, "dissatisfaction with the census as a benchmark has led some researchers to develop alternate set of benchmarks."[21] Among these are "observational" benchmarks that seek to estimate the subpopulation—as shaped by both the built environment and social patterns—at risk for police contact at relevant times. Much of the research has emerged from the analogous situation of racial disparities highway stops. Numerous studies, for example, sought to measure by racial patterns in speeding—a violation that give rise to 89% of New Jersey highway stops—through the independent use of radar guns in order to craft a more accurate "at risk" population baseline for comparison. Likewise, another study trained observers to record traffic violations by race at sixteen intersections in Miami. Although these studies have their own challenges, they collectively point towards what Prof. Fagan might have done and did not in his insistence on employing a census benchmark (rather than, say, a benchmark of the criminal participation): use video to test the validity of his census benchmarking model.

In a 2002 article, in fact, Prof. Fagan called for applying precisely such methods to *Terry* cases, writing:

> Just as the refinement of observational studies on streets and highways is essential to improving estimates of the supply of persons available for

[21] G. Ridgeway and J. MacDonald, "Methods for Assessing Racially Biased Policing." *Race, Ethnicity, and Policing: New and Essential Readings*, eds. S. K. Rice and M. D. White. (New York: New York University Press, 2010): 180 – 204, p.182.

> stops, so too does improvement in measuring racial disparities in pedestrian stops depend on parallel developments in observational techniques and classification of suspicious behavior.[22]

It goes without saying that Prof. Fagan did not heed his own advice while crafting his report for the Plaintiffs. Again, Prof. Fagan's failure is *not* that he failed to use "observational" benchmarks through a logistically impossible exercise, massive exercise in videotaped urban ethnography. Rather, his failure emerges from the fact that he did not—despite clear caveats in the social science literature—test the validity of his adjusted census benchmark with "observational" benchmarks.

Applying a Criminal Participation Benchmark Standard

This report has demonstrated that any measurement of racial disparity in police action must—in order to reflect a clear consensus among social scientists—consider the racial composition of the criminally active. The logic is simple: the appropriate benchmark must "neutralize" (as the DOJ's how-to guide phrases it) differences in criminal participation that would otherwise distort disparity findings. Prof. Fagan himself, as we have noted, has called such a benchmark "ideal." We have further established that the NYPD's suspect-race data is *sufficiently* both complete and consistent to use as just such a benchmark of criminal participation. (Prof. Fagan's various objections to the use of the suspect-race data, as addressed at length in this report, lack merit). This report has also made evident that Prof. Fagan's proposed alternative to the criminal participation benchmark—an "adjusted" census benchmark—is out of step with

[22] Jeffrey Fagan, "Law, Social Science, and Racial Profiling," *Justice Research and Policy* 4 (Fall 2002): 104-129, p. 118.

current practice in the well-developed field of disparity measurement, precisely because of the errors and biases it can introduce in the type of contexts at issue in *Floyd*.

What does one find by comparing the racial composition of NYPD's *Terry* stops to the racial comparison of New York City's criminally active? Whether one makes such a comparison at the census tract level or the police precinct level; whether one makes the comparison in quarters of the city with high minority populations or low; whether one looks at violent crime or minor disorder, the racial and ethnic composition of *Terry* stops closely correlates with the racial and ethnic composition of the criminally active population of New York. In short, the NYPD does *not* stop minorities at a rate disproportionate to their share of the criminally active compared—the most relevant and appropriate benchmark. Although much of the rest of the report will address the statistical basis for these findings in detail, it might be useful to describe them briefly first in broad strokes and in plain words.

Let's consider the evidence as the police department might: at the precinct level. New York's five boroughs are divided into 76 precincts. Using all know crimes suspects, in 45 of those 76 precincts, African Americans are stopped at rates either equal to or lower than their share of the criminally active population[23]. Latinos are stopped at rates equal or lower than their share of the criminally active population in 35 of 76 precincts. Using the measure most related to police strategy, violent crime demography, the evidence that crime commission is not limited to the precinct (or census tract) in which

---

[23] New York City Police Department, Office of Management Analysis and Planning *2011 Reasonable Suspicion Stops: Precinct Based Comparison by Stop and Suspect Description*, by Michael J. Farrell, Philip McGuire, and Thomas J. Taffe (New York City, New York City Police Department). "Equal to less than" for these figures is considered the proportion of known crime suspects plus 2.5% or below.

the offender resides: In 71 of 76 NYPD precincts, African Americans are actually stopped at rates *lower* than their share of the known violent crime suspect population; and in no precinct does the percentage of African Americans stopped for violent crime exceed their share of the known violent crime suspects by more than 2.5 percentage points.[24] Latinos are stopped at rates equal to or lower than their share of known violent criminal population in 55 of 76 precincts; and in only five precincts does the differences between those rates exceed three percentage points. Let's consider the evidence as the public might see it, with violent and dangerous crimes front and center--as well as trespass crimes, which are often a necessary predicate to far more serious offenses such as burglary, rape, and homicide. Violent Crimes: With the race or ethnicity of 85.6% of all suspects known, African Americans represent 51.5% percent of known suspects and 52.2% percent of individuals stopped on suspicion of that offense. Hispanics represent 30.8% of known violent crime suspects and are 34.8% of individuals stopped for a violent offense. The race or ethnicity of 97.9% of weapons related offence suspects are known to police. 51.1% of those known suspects are African American and 60.7% of *Terry* stops for weapons related offenses are of African Americans. Hispanics are stopped and questioned for weapons related offences 30.9% of the time and represent 36.4% of known suspects. The race or ethnicity of 95.8% of suspects of trespass crimes are know. Of those suspects, 54% are identified as African American and 60.2% of stops

[24] New York City Police Department, Office of Management Analysis and Planning *2011 Reasonable Suspicion Stops: Precinct Based Comparison by Stop and Suspect Description*, by Michael J. Farrell, Philip McGuire, and Thomas J. Taffe (New York City, New York City Police Department). "Equal to less than" for these figures is considered the proportion of known crime suspects plus 2% or below.

related to trespass offenses stop African Americans. Hispanics represent 31.3% of trespass related stops and 34.7% of known trespass suspects.

Let's consider the evidence as a statistician might, by looking at a scatter plot of *Terry* stops for *all* crimes (not merely the violent crimes that most concern the public) and for census tracts (rather that the larger and more arbitrary geographical unit of precinct that the police use). With those parameters, we can compare the racial demographics of all *Terry* stops to the racial demographics of New York's criminally active population. Presented in this standard fashion, the graphs look like this:

FIGURE 1.

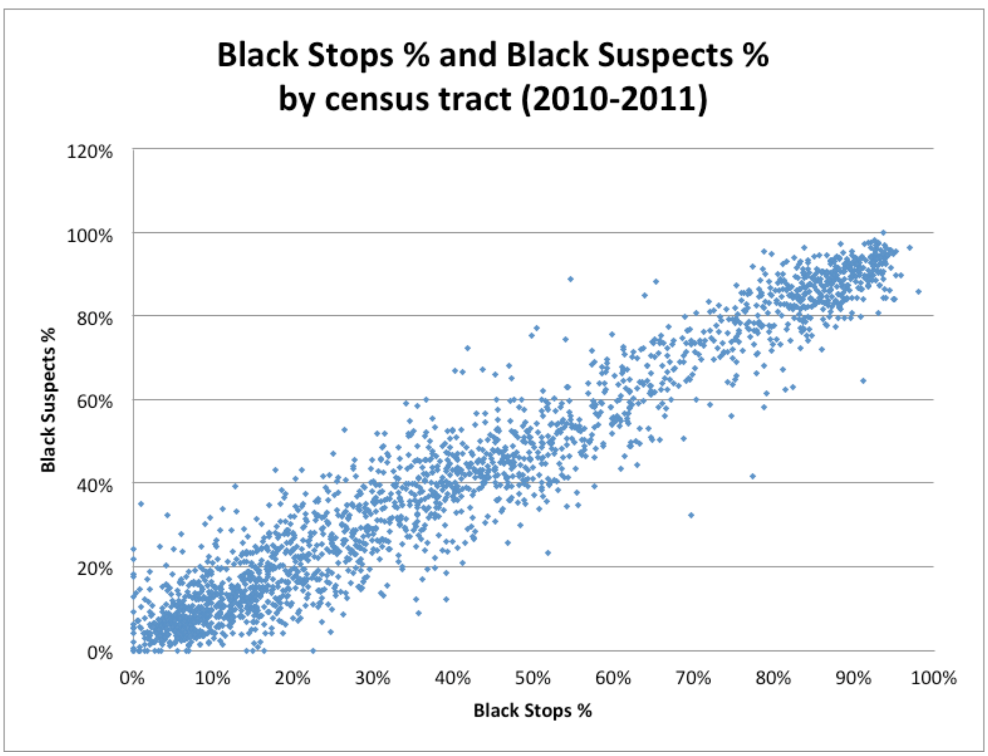
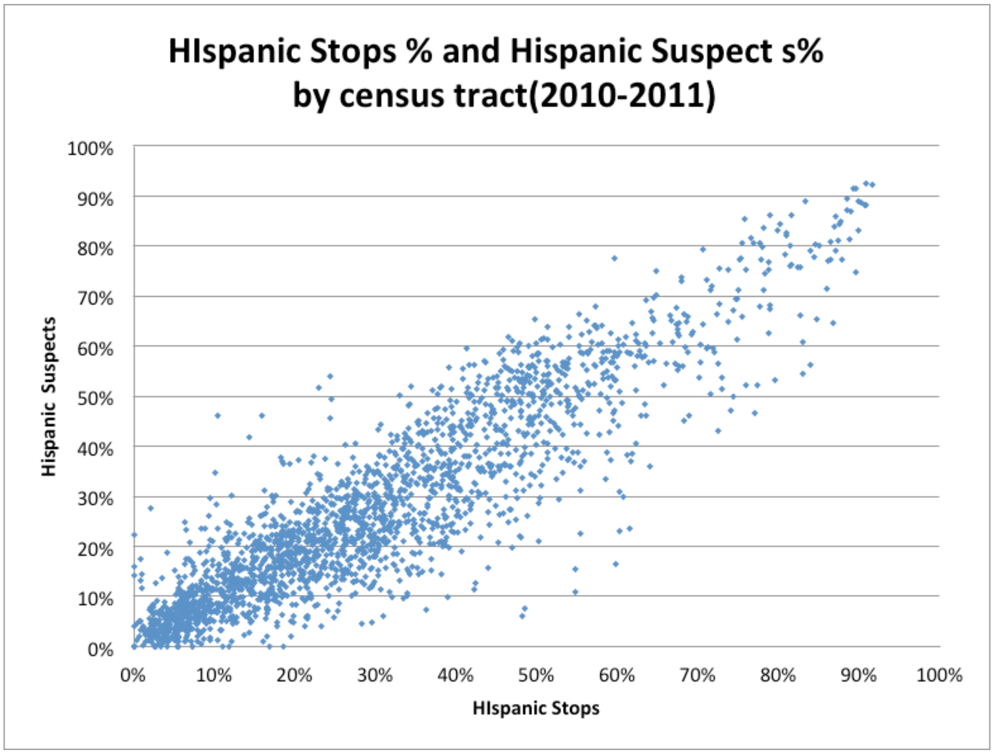


FIGURE 2.



The high correlation between the racial composition of *Terry* stops and the racial composition of the criminally active population is unavoidably apparent in these graphs. Why? Statisticians see in scatterplots the slope of relationships, the degree of dispersion around the mean, and the presence or absence of outliers in the data. Plots that show, as ours do, tight clustering around the trend line provide visible evidence of high correlation But a statistician looking over the evidence would notice more than the strength of the correlation—they would notice how much more precise the relationship is when the analyses employs a benchmark of the "criminally active" population rather than the census population. Consider, for example, the scatter plot below and then compare the pattern of dispersion around the line of central tendency.

FIGURE 3.

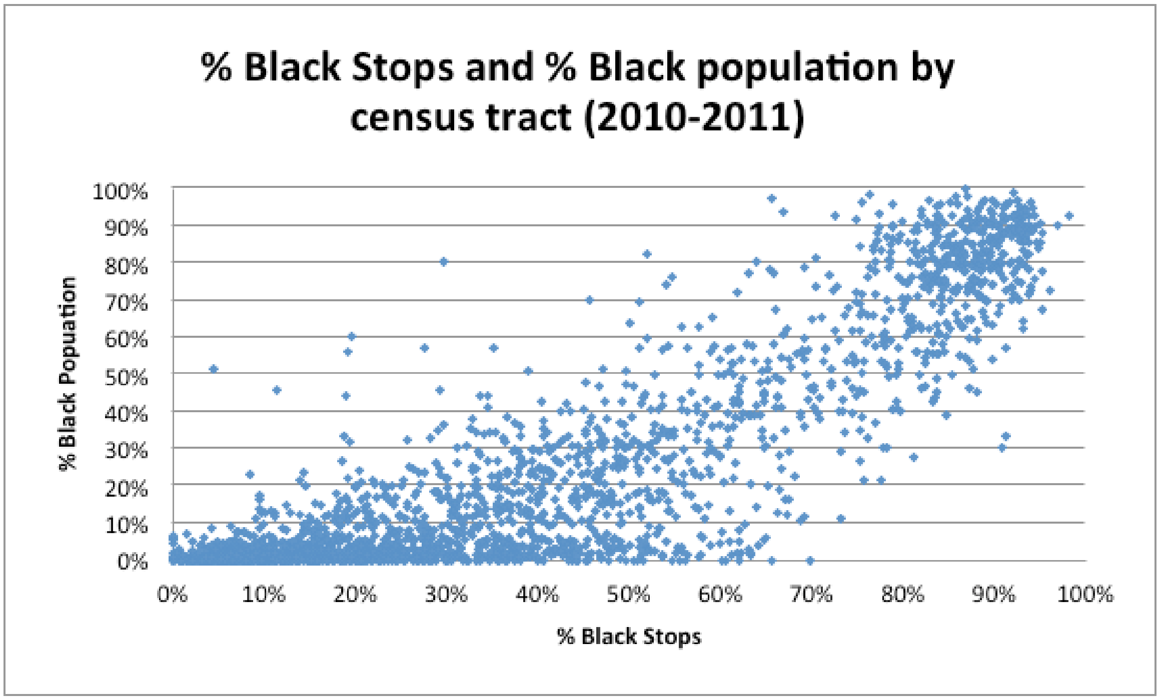
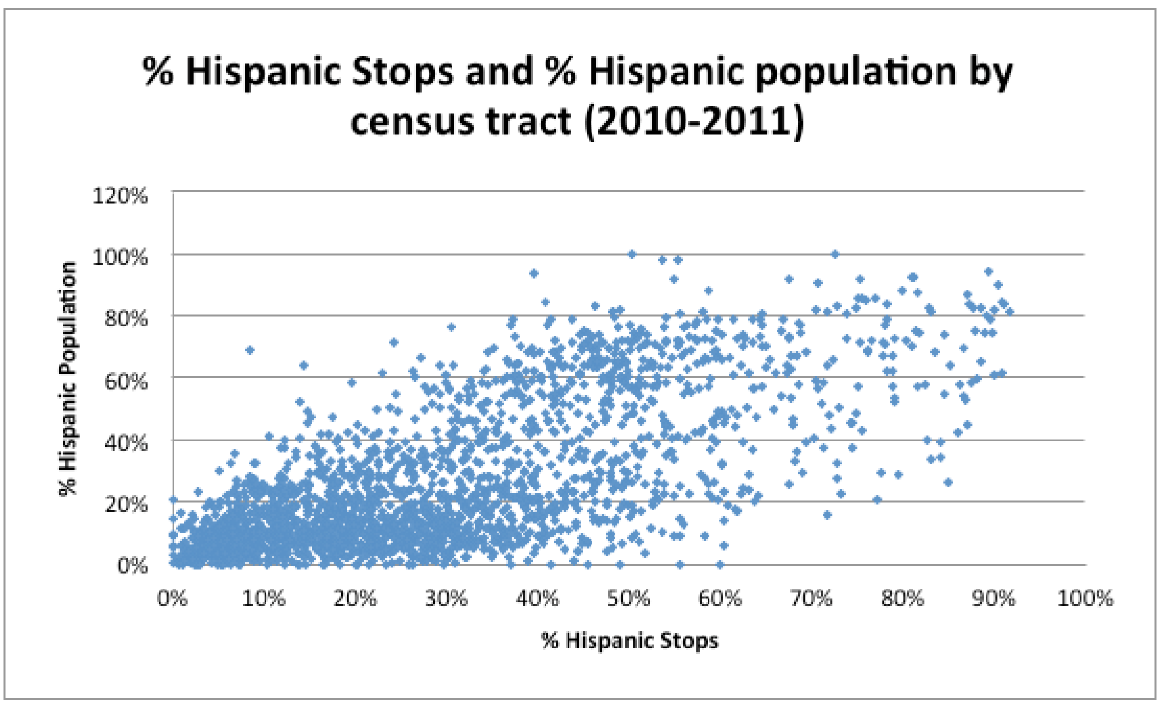


FIGURE 4.



Rather than the first figures' orderly array of plots that indicate a robust relationship, we find a graphically different scatter pattern marked by a wide dispersion —indicating a weak relationship. Why? Plots that are truly scattered—that is, have much wider and more chaotic dispersion—are taken as evidence of relatively low correlation between the variable plotted on the graph.[25] In short, Prof. Fagan's chosen benchmark is not simply inconsistent with accepted practice in the field, the relationship he claims that benchmark reveals is much weaker than what is produced when one uses the benchmark of criminal participation. At risk of belaboring the point, it is *that* benchmark the literature prefers and that we employ. We will explore each of these issues in depth below.

3. The Constitutional Justification of Stops on the Basis of Information in the SQF Database

At the heart of the dispute over the constitutionality of NYPD *Terry* stops is the information provided on UF250 forms by officers regarding their reasons for stopping a suspect, and the circumstances in which the stop occurred. In earlier reports in the case of *David Floyd et al v. the City of New York et al*, Prof. Fagan has noted the extraordinary complexity of the potential combinations ("permutations") of reasons and circumstances included in the UF250 form, and he used that overwhelming complexity to justify his choice to rely on a limited set of UF250 data in his assessment of the legal sufficiency of police stops. We criticized Prof. Fagan at the time, both for failing to acknowledge that his selective truncation of the available data potentially introduced biases into his

[25] For example in Introductory Statistics: Concepts, Models and Applications (1998), David W. Stockburger explains:

When r=0.0 the points scatter widely about the plot, the majority falling roughly in the shape of a circle. As the linear relationship increases, the circle becomes more and more elliptical in shape until the limiting case is reached (r=1.00 or r=-1.00) and all the points fall on a straight line.

analysis, and for failing to consider how officers experience that extraordinary complexity as they police the City and attempt to fit that complexity into a largely multiple choice form. We posited that even with training, officers would need time to learn to correctly master the procedure (when, for example is it required?) and content of UF250 reports. Consequently, we argued that it is appropriate to track trends over time in forming judgments about police practices based on reading the forms. Prof. Fagan consistently tracked the trend in the number of stops recorded by UF250 forms, but he paid less attention to other evidence of change over time.

In our analysis of Prof. Fagan's Second Supplemental Report presentation of his finding on the apparent legal sufficiency of stops, we encounter another example of a tendency not to provide a balanced account of what is evident in the data. We have previously noted his tendency to only selectively inquire about trends in the data he reports.[26] Prof. Fagan notes that many of his findings in the Second Supplemental Report are consistent with those reported in October, 2010, but the Report is silent in those instances where the comparison does not support his argument. Given NYPD's claims that through policy changes and training it has demonstrated its commitment to assuring the constitutionality of police practices in pursuing its strategy of crime prevention, one would expect an objective social scientist to acknowledge the full range of differences between his summary report, his analysis of the "legal sufficiency of stops" in his Tables

---

[26] Our Declaration of 2-12-11 noted: "Again, while not conceding the validity of Fagan's methodology, I observe that his presentation of the data, as ever, masks positive trends uncovered by his own methodologies. For example, applying Fagan's "corrected" coding scheme, the percentage of "unjustified" stops has decreased by nearly 50% -- from 9.7% in 2004 to 4.5% in 2009, while the percent "justified" has increased from 69.1% in 2004 to 84.4% in 2009; stops of "indeterminate" legality have also decreased by almost one-half, from 21.2% in 2004 to 11.1% in 2009. See p.3.

12 (October, 2010), and his findings presented in Table 12c in his Second Supplemental Report.

Table 3.

From Fagan Table 12c. Legal Sufficiency of Stops by Suspected Crime (% of Stops Considering "Other" Stop Factors, 2010-12, New York City [compared with data from Table 12 (2004-2009) in Prof. Fagan 10-10 Report]

		Legal Sufficiency				
Radio Runs		*Apparently Justified*		*Not Generalizable*	*Apparently Unjustified*	
Total Stops	408,573	87.51	**68.03**	7.58	4.91	**5.26**
Violent Crime Stops	109,977	94.71	**77.60**	2.66	2.63	**4.90**
Property Crime Stops	126,814	85.53	**71.36**	11.3	3.18	**3.62**
Drug Stops	31,056	90.41	**82.17**	6.38	3.21	**4.68**
Weapon Stops	67,611	89.66	**60.96**	3.59	6.76	**9.34**
Trespass Stops	33,759	68.47	**43.11**	14.32	17.21	**5.25**
QOL Stops	7,635	87.96	**64.09**	8.23	3.81	**5.46**
Other Stops	31,721	83.24	**68.99**	12.13	4.62	**4.96**

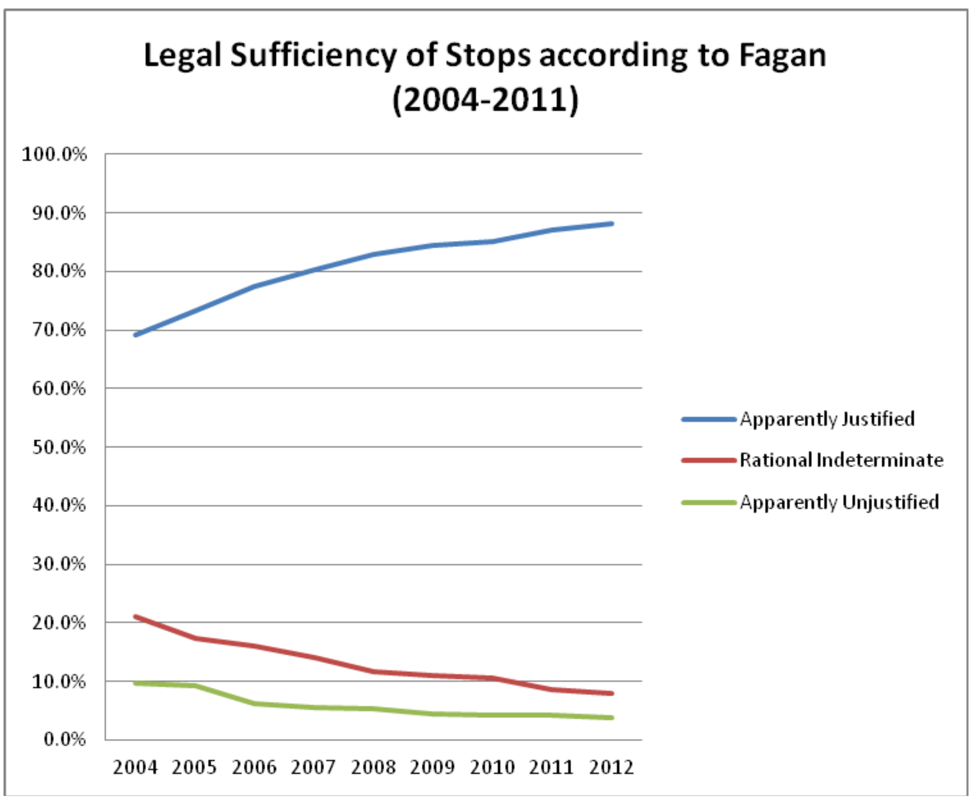
		Legal Sufficiency				
Non-Radio Runs		*Apparently Justified*		*IVot Generalizable*	*Apparently Unjustified*	
Total Stops	1,215,846	88.10	**68.92**	5.67	6.23	**6.71**
Violent Crime Stops	278,190	96.04	**82.10**	1.34	2.63	**5.06**
Property Crime Stops	278,451	88.29	**75.11**	8.42	3.29	**4.56**
Drug Stops	113,608	90.44	**85.84**	6.53	3.03	**4.16**
Weapon Stops	339,797	89.49	**63.35**	2.23	8.28	**12.30**
Trespass stops	95,583	62.33	**38.56**	14.59	23.08	**6.06**
QOL Stops	23,021	87.22	**59.77**	8.57	4.21	**7.25**
Other Stops	87,196	82.25	**69.18**	12.44	5.30	**6.0**

Even using Prof. Fagan's approach to assessing the RAS evidence in stop reports, Table 10 and the related graph below show the steady improvement in NYPD use of "*Terry* stops" in its efforts to prevent crime.

Table 4.

Legal Sufficiency of Stops according to Fagan (2004-2011)		
Apparently Justified	Rational Indeterminate	Apparently Unjustified
69.1%	21.2%	9.7%
73.2%	17.5%	9.3%
77.6%	16.2%	6.3%
80.3%	14.2%	5.5%
83.0%	11.7%	5.3%
84.4%	11.1%	4.5%
85.0%	10.7%	4.3%
87.1%	8.7%	4.2%
88.2%	7.9%	3.9%

Figure 5



The results might in fact have provided even more evidence of improved police performance if Prof. Fagan's most recent assessment strategy had not changed the assignment of so many stops he previously characterized as "indeterminate" (or ungeneralizable in the terminology directed by the Court) to "apparently unjustified." A significant number of the stops shifted to "apparently unjustified" occurred in just one category of suspected crime, trespass. We will return to the controversy surrounding police enforcement of laws against trespass, but first we must note that we are not at all certain of the status of the analytic strategy used by Prof. Fagan to produce the result he reports.

Prof. Fagan reports: "I conducted an analysis of the text strings in a stratified random sample of 3,710 cases where "other" was checked, and classified those text strings into groups or categories similar to the categories used in October 2010 Report to evaluate the apparent legal sufficiency of the stated stop factors as apparently justified, conditionally justified, or apparently unjustified." Later in the report he elaborates on his method:

> As noted in the seventh through ninth classifications above, one other combination of RAS indicia requires additional analysis before a full categorization can take place. These include those stops where (a) the only stop circumstance indicated is "Other," (b) the "Other" stop circumstance is indicated and one additional circumstance is indicated or (c) only one conditionally justified stop circumstance is indicated and the "other" stop circumstance is indicated. This group contains 156,090 stops. To better estimate the RAS indicia present in these stops,[36] an additional analysis was completed to examine the entries in the text field on the UF-250 form that accompany "other" stop circumstances. This analysis, which I previously performed for my expert report(s) in *Davis v. City of New York,* proceeded in three stages.
>
> First, I examined the text strings for a random sample of 3,710 cases where "other" was a stop circumstance indicated. After considering both the "other" stop factor and any accompanying conditionally justified Primary Stop Circumstances and Additional Circumstances, these cases were initially classified as *Not Generalizable.* The sample for this analysis was stratified based on the suspected crime categories that were the majority of stops in the larger subset where the "other" stop factor was checked.

> Specifically, I sampled 5% of trespass stops, property stops, and "other" stops, and 3% of violent crime stops, weapon stops, drug stops, and quality of life stops.

We believe that the description of the sample frame used and the rationale for it provided in the report are insufficient to enable us to judge its scientific merits. We have had previous experience with problems presented by the samples Prof. Fagan has used in the analysis he has presented to the court in this Case. Prof. Fagan used an analysis of a sample of 1,000 UF250 forms to state in both his 10-20-10 Report and his subsequent 12-20 Declaration that the data derived from handwritten responses supplied by officers after checking "Other" on the UF250 form was "meaningless" and unuseable, and that "Any attempt to do so would invite a host of potential biases and errors and would render any conclusion statistically meaningless." Data from a sample of officers' responses to "Other" on the UF250 are the basis for Prof. Fagan's refinement of the algorithm he used previously to classify the legal sufficiency of stops.

We make following observations about aspects of the sample presented in a study should have addressed but did not:

1. Prof. Fagan has not explicitly specified the process by which he randomly selected the 3,710 cases from the larger universe of 156,090 *Terry* stops. If the pool from which the sample was drawn included only cases classified initially as "indeterminate," why do we find—when we apply the coding provided by Fagan – that his sample includes 170 cases (4.58% of the sample) classified as "presumptively unjustified"?

2. Prof. Fagan describes his sample as "stratified," but offers no explanation of his stratification strategy. Prof. Fagan does not explicitly specify why he stratified his sample by suspected crime.

3. Prof. Fagan has not explicitly specified his justification for choosing to sample 5% of trespass stops, property, and "other" stops and 3% of violent crime stops, weapons stops, drug stops, and quality of life stops.

4. Prof. Fagan has not made clear from which universe the aforementioned 5 and 3 percentage samples subsets come.

5. Prof. Fagan has not specified explicitly his justification for oversampling trespass stops (as acknowledged in fn. 38).

6. Prof. Fagan has not specified the statistical power of his sample size. How and why does he believe 3,710 is sufficient?

7. Prof. Fagan has not specified what the statistical power of sample size means for the confidence intervals of his conclusions.

8. Prof. Fagan has not provided his justification for drawing his sample only from the pool of stops previously classified as indeterminate.

9. Prof. Fagan has not specified the "reliability of sampling and coding procedures that were used for 3,710 stops." (p.32) On what criteria does Prof. Fagan base his stated belief that these procedures were reliable, and to what degree does he believe they were reliable?

A careful reading of the Report will support our claim that the answers to our questions are not to be found there. Prof. Fagan does provide a limited *description* of what he did, but he does not "explain," or justify, what he did to assure that the sample is representative of the population of stops studied.

Our analyses of the sample data and coding instructions provided by the Plaintiff did raise additional questions about Prof. Fagan's methodology. When we followed the path described by Prof. Fagan, using the coding instructions from the Plaintiffs on the data used in his analysis, here is what we found that leaves the questions we posed outstanding:

Using Prof. Fagan's algorithm to reduce the 1.6 million database records to a subset from which he takes his samples (he has a long and difficult to follow command for determining which observations to use, which he claims obtains those stops that were "Not Generalizable") produces a subset from which his sample is drawn.

Mysteriously, to us at least, the subset is 84,000 observations (not the 156,090 noted in the Second Supplemental Report and reiterated in the letter from Plaintiff's attorney denying us further clarification). The subset from which Prof. Fagan drew the sample is not exclusively stops that were classified as "Not Generalizable," as the Report claims, but contains 4.9% of stops initially classified as "apparently unjustified stops." Note that once a category has been classified as apparently unjustified, there is nothing in the other text that can make it justified. See below:

Table 5 Comparison of All Stops and Subset Analyzed

Subset of observations for sample				
	Apparently Justified	**Indeterminate**	**Apparently Unjustified**	**Total**
Not Subset	1,405,958	70,423	64,038	1,540,419
Subset	0	79,919	4,081	84,000
%	0.0%	95.1%	4.9%	100.0%
Total	1,405,958	150,342	68,119	1,624,419

Source: Data provided by Plaintiff's attorney

Prof. Fagan divides these 84,000 observations into eight different databases (based on the crime suspected by officers) and takes a random sample of different sizes (3% or 5% of each of these separate databases). The merged result of these separate random samples makes up his 3,710 sample. The table below reflects the details of this process:

Table 6. Description of Stop Subset and Sample

Crime Category	Subset	Size of Sample	Observations Sampled
1: murder	61	3%	2
2: violence	5,134	3%	154
3: weapons	8,816	3%	264
4: property	26,178	5%	1,309
5: drugs	8,396	3%	252
6: trespass	20,655	5%	1,033
7: qol	2,099	3%	63
8: other	12,661	5%	633
Total	84,000		3,710

Source: Data provided by Plaintiff's Attorney

Since the goal of a careful random sample is to obtain data representative of the field of data of interest, it is standard social science practice for a researcher to explain how he or she achieves a representative sample. The main question here is the stratification of the sample and its implications for representativeness. It is not clear to us why he chose 5% samples of property and trespass crimes and 3% samples of other suspected crimes. Considering the fact that in the results of his reclassification it is mostly the trespass stops, and in lesser degree the property stops, that he moves from indeterminate to apparently unjustified, we believe it is legitimate to inquire about the impact of his sampling strategy on the findings. Taken together, while neither property nor trespass stops are the largest crime category subset, there are 534,607 stops out of 1,624,419, or 33% of all stops included in the analysis.

Typically social science researchers compare their sample characteristics to what is known of the population sampled. Prof. Fagan does not provide that analysis, but we have done a preliminary check. See below the distribution of the sample vs. the distribution of the population of "cases where other was a stop circumstance indicated" by several categories. Since Prof. Fagan stratified his sample based on suspected crime categories (newstopcat), we compared the population of stops with the sample:

Prof. Fagan reports that he stratified his sample by the suspected crime variable. However, some categories appear to be oversampled and some undersampled. Trespass and property crimes were over sampled, while weapons, drugs, violence, and quality of life were undersampled.

Table 7. Comparison Stop Subset and Sample: Crime suspected

Crime Category	Subset	% of subset	Observations Sampled	% of sample
1: murder	61	0.1%	2	0.1%
2: violence	5,134	6.1%	154	4.2%
3: weapons	8,816	10.5%	264	7.1%
4: property	26,178	31.2%	1,309	35.3%
5: drugs	8,396	10.0%	252	6.8%
6: trespass	20,655	24.6%	1,033	27.8%
7: qol	2,099	2.5%	63	1.7%
8: other	12,661	15.1%	633	17.1%
Total	84,000	100.0%	3,710	100.0%

Source: Data from Plaintiff's Attorney

The oversampled stops disproportionately include those that include the use by NYPD officers of responses supplied on the UF250 form that particularly arouse Prof. Fagan's suspicion. We turn now the role of "furtive movement" and "Area has high incidence of the offence suspected" in *Terry* stops.

The stops oversampled are ones that disproportionately include two UF250 responses by NYPD officers that particularly arouse Prof, Fagan's suspicion. We turn now the role of "furtive movement" and "Area has high incidence of the offense suspected" in *Terry* stops.

Use of the "Furtive Movements" and "High Crime Area" Stop Factors

We have noted that Prof, Fagan is sensitive to the fact that police make decisions in complex situations, and the records some of those decisions create pose challenges to the analyst. But he is particularly insensitive to the challenges confronting police officer decision making in his exploration of the use of "furtive movements" and what he calls "High Crime Areas" in his analysis of the legal sufficiency of stop decisions. The analysis in the Second Supplemental Report begins with the assertion that "furtive movements" and "high crime area"

> were the two factors most commonly checked on the UF-250 form, and are central to the classification of stops as apparently justified or apparently unjustified. As noted in the October 2010 Fagan Report,[39] both of these factors lack precise definitions or standards, and are both vulnerable to subjective and highly contextualized interpretation.[40] Both are independently legally insufficient to justify a stop.[41]

Prof. Fagan should at least acknowledge that most stops by NYPD officers occur in the precincts with the City's highest levels of crime, and that, as he observed in his work on the *Davis* case, the NYPD specifically assigns officers to small areas where the violent crime that remains in the City is concentrated. The work of Prof. Weisburd and his colleagues has clearly established that the pattern of NYPD *Terry* stops is highly correlated with violent crime patterns. (Weisburd et al, 2013). As noted in our Report for the Case of *Davis v. New York City*, in contrast to a widely reported claim by a journalist for WNYC that NYPD was not making gun stops where the guns are, Steven Romalewsky (Director of the CUNY Mapping Service at the Center for Urban Research) challenged the mapping methodology used in the WNYC report, and presented a contradictory finding:

- Quantitatively the second map reveals that most gun recoveries in 2011 were in census blocks where most of the stop and frisks took place (the opposite of WNYC's conclusion). The pink-to-hot pink blocks in the second map account for 433 recovered guns, or **56% of the total in 2011**.

The following two maps show this overlap on a citywide basis, and then zoomed in on the Brooklyn-Queens border:

FIGURE 5. Modified thresholds, citywide, with gun recovery incidents

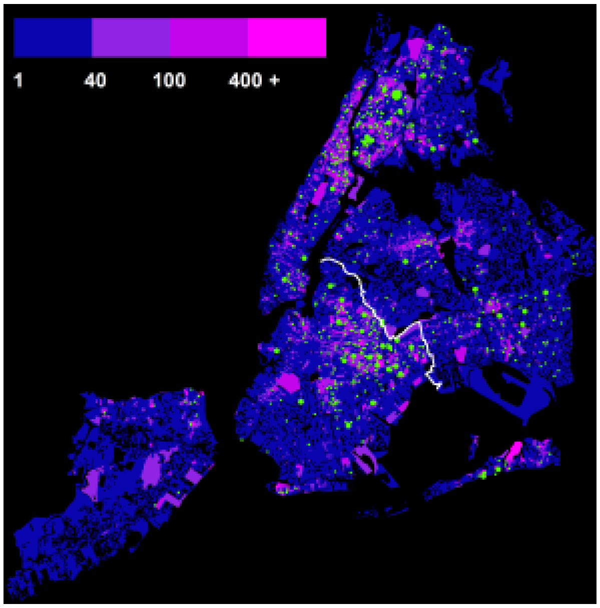
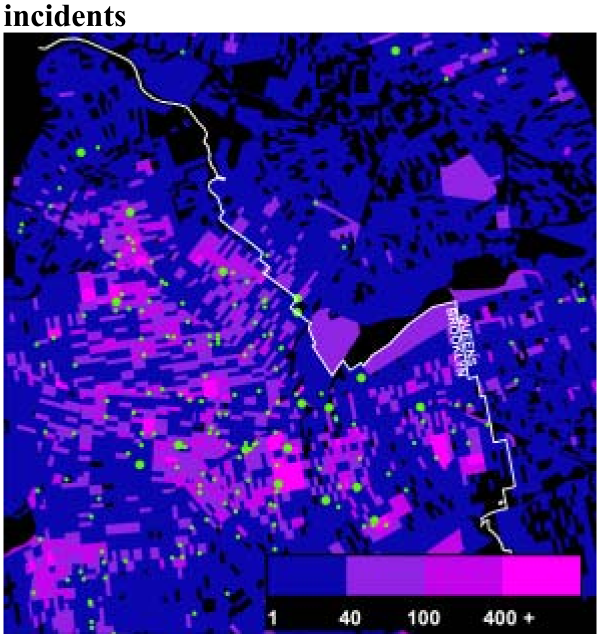


FIGURE 6. Modified thresholds, along Brooklyn-Queens border, with gun recovery incidents



In the overall case, it is therefore not surprising that on most stop reports officers indicate that it occurred in what Prof. Fagan labels a "High Crime Area." His claim that the "October 10 Fagan Report" demonstrated that both factors are poor indicia that "crime is afoot" within either the language or the jurisprudential meaning in Terry v. Ohio (or the notion of “high crime area as articulated) in *Wardlow v Illinois* is an acknowledgement that he has not answered our report’s challenge regarding the findings of his research. “Furtive movement” and “Area Has High Incidence of Reported Offenses of the Type Under Investigation” are “poor indicia” primarily if they are the only items checked. They are on the UF250 form for reasons that were apparent to the parties who agreed to the form as a tool for officers to articulate the reasons for their suspicion, and the form gives officers the opportunity to check multiple factors and write in additional ones, in recognition of the challenge of representing the complex decision leading to a *Terry* stop.

In brief, we have argued in previous reports to the Court, and maintain here, that Prof. Fagan’s now repeated assertion that “these factors are used somewhat promiscuously and indiscriminately” has not been established by his analysis. It instead results from his misunderstanding of the likelihood that furtive movement is not an uncommonly observed behavior of persons who, while planning or considering the commission of a criminal act, encounter an NYPD officer who is trained to intervene proactively to prevent crime. Prof. Fagan also asserts that both of these factors lack precise definitions or standards, and are both vulnerable to subjective and highly contextualized interpretation.   Footnote 4 refers to an article by Andrew Guthrie Ferguson and Damien Bernache, The "High Crime Area Question:" Requiring Verifiable and Quantifiable Evidence for Fourth Amendment Reasonable Suspicion Analysis, 57 *American University Law Review* 1587, 1588 (2008) and to an article by Robert J. Sampson and Stephen W. Raudenbush, Seeing Disorder: Neighborhood Stigma and the Social Construction of "Broken Windows," 67 *Social Psychology Quarterly* 319 (2004). As we have noted previously in response to Prof. Fagan’s citation of Sampson and

Raudenbush's work, their study did not include police officers but citizens; in addition, another article whose findings Prof. Fagan misrepresented concluded that even if officers developed negative views of populations they police, those views were not associated with a higher propensity to make stops based on race.[27]

Gutherie and Bernache, far from proposing that "High Crime Area" be eliminated from the set of factors used by officers in explaining their decision to make a *Terry* stop instead propose an a better operationalization of the concept that would allow courts to more systematically review its use. Here is what they propose:

- that "the claim of "High Crime Area" would require evidence of higher incidence of the particularized criminal activity than other areas of the jurisdiction,"
- that "the area must be tailored to a specific geographic location and limited to a recent temporal finding of criminal activity," and
- there would have to be a demonstrated nexus between a police officer's knowledge about a defined area and the reasonableness of that officer's observations in that area.

In an effort to ascertain legal sufficiency, Prof. Fagan cites Guthrie and Bernache's law review article because it directly addresses the use of this term, but he fails to follow the prescription they provide.

Fagan's comparison of a notation of "Area has high incidence . . ." with the previous month's crime rates in the precinct where the stop occurred[28] lacks the

[27] Geoffrey P Alpert,., John M. MacDonald, and Roger G. Dunham. "Police Suspicion And Discretionary Decision Making During Citizen Stops." *Criminology* 43 (2005)

[28] The spatial specific concern we expressed in his empirical test of the validity of the indication by an officer that the "Area Has a High Incidence . . ." is somewhat reduced by his use of census tracts, if he included factors of spatial proximity, because census tract does not have any special connection to crime patterns or police organization to respond to them. The smaller area focus does not to address the problems of using last month's crime data in general in his test of concept validity, nor does it do anything to mitigate his inability to assess officers' specific relevant knowledge of crime patterns.

geographic specificity and temporal immediacy called for by Gutherie and Bernache. In addition, Fagan did not even attempt to assess whether officers could have had up-to-date, specific information that might form the nexus between knowledge and action reflected in the decision to stop and to include "Area Has a High Incidence…" among the factors cited. This critique seems especially relevant to the failure of Prof. Fagan in this Second Supplemental Report to distinguish stops that are made by Impact Zone officers whose assignment provides many (and perhaps all) of the elements called for by Gutherie and Bernache.

Finally, imbedded in Prof. Fagan's assertion that "furtive movement" and what he calls "High Crime Area" both "lack precise definitions or standards, and are both vulnerable to subjective and highly contextualized interpretation" is a comparison with other rationales selected by officers in completing the UF250 form. The "suspicion" the *Terry* decision apparently authorized as a basis for decisions to act is inherently "subjective," contextual, and hypothetical in nature. If an officer knew, or had probable cause to believe a crime had been committed, the officer would be required to make an arrest, not conduct a *Terry* stop. While they vary, the different check off items on the UF250 form all to some degree contain an element of subjectivity; for example, if an officer checks "fits description," "actions indicative of 'Casing,'" or "actions indicative of acting as a lookout." Prof. Fagan does not question their judgments, but police officers who can choose among these may see each of them as calling for his or her professional judgment of the meaning of the behavior observed.

Modeling Disparate Treatment

The issues raised previously in this report manifest themselves in the regression analyses Prof. Fagan presented. That modeling effort contains a series of compounding errors and omissions. The flaws are a combination of: (i) Prof. Fagan's lack of understanding of policing strategy and the tactical situations officers on patrol face and (ii) his failure to apply generally accepted statistical modeling practices. These issues were raised in prior submissions to the court. However, the issues raised in previous submissions to the court, and reiterated here, have not been addressed in his analyses. Our critique is based on two fundamental principles for any rigorous analysis. First, valid statistical models must include all variables relevant to the analysis and those variables must be properly specified. Second, properly specified models must control for alternative explanations of the effect being analyzed. We believe that Prof. Fagan's proffered models fail on both counts.[29] One particular problem with interpreting regression models with a large number of observations, as is the case with the models Prof. Fagan presented, is that even very tiny and nearly meaningless effects will still be estimated to be "statistically significant". As a result, model specification, preliminary data analysis, and an explanatory discussion of the model selection process are all vitally important to meeting professional standards of rigor. Prof. Fagan fails to acknowledge these issues and offers only limited justification for his model selection and specification choices. The remainder of this report expands on the limitations of Prof. Fagan's analysis and shows the likely impact a more rigorous analysis would have on his findings. However, we were unable to address all of the issues and the interactions between and

[29] David Greenberg, Studying New York City Crime Decline: Methodological Issues, Justice Quarterly, January 13, 2013.

among them in the time allotted for our response. Even in the absence of a more thorough vetting of the data, the flaws in Prof. Fagan's most recent report are substantial, and the models he bases his conclusions on do not approach meeting the standard required for making, from a statistical and methodological vantage point, a valid claim about racial disparity in police practice.

Specific Issues with the Negative Binomial Model: We contend that Prof. Fagan has mis-specified the models he used in his analysis both in terms of the variables he failed to include in his analysis and his specification of the models. Prof. Fagan used a Negative Binomial model to investigate the issue of disproportionate enforcement. He chose this approach because stop and crime data are measured by counts which follow either a Poisson distribution (Poisson Regression) if the data are dispersed equally or a Poisson-gamma distribution (Negative Binomial Regression) if the data is over dispersed. Choosing between the two models requires careful data analysis and tests of dispersion. Prof. Fagan provided neither.

In addition, properly implementing these models requires that the analyst specify an exposure variable. Exposure variables are meant to operationalize the number of times that an event could have happened in the target area and control for entry by surrounding-area populations. .(c.f. Ridgeway & McDonald 2010) In this case, the exposure variable would have to be directly related to the probability that that a police officer might encounter an individual exhibiting a behavior that justified police action i.e. stops, arrests, etc. It is important to note that the exposure variable contains two components. First, you need an estimate of the population exhibiting the behavior which Prof. Fagan does not include. As we note below, the population estimates that Prof. Fagan does use are subject to errors and are likely to be negatively biased. (Smith, 1987, Smith & Shahidullah 1995) Second, you need an unbiased measure of police presence in the area being analyzed. As we also note below, the estimated generated by Prof. Fagan are biased and subject his analyses to issues of endogeneity. This exposure variable is the denominator in all of the rates impacted by the covariates in the model. In essence it is a measure of the number of people a police officer might encounter while on duty conditioned for the behavior they exhibit. The level of this traffic will be influenced by population density, the built environment which channels people into confined areas and, for trespass enforcement, the likelihood of people entering a building. In other work, Prof. Fagan has acknowledged the relevance of these factors but not in this Second Supplemental Report.

The Statistical Consulting Group at UCLA Academic Technology Services provides the following example to clarify this concept. (see http://www.ats.ucla.edu/stat/stata/seminars/count_presentation/count.htm)

> Count models need some sort of mechanism to deal with the fact that counts can be made over different observation periods. For example, the number of accidents are recorded for 50 different intersections. However, the number of vehicles that pass through the intersections can vary greatly. Fifteen accidents for 30,000 vehicles is very different from 15 accidents for 1,500 vehicles. Count models account for these differences by including the log of the exposure variable in the model with the coefficient constrained to one.

> The use of exposure is superior in many instances to analyzing rates as response variables because it makes use of the correct probability distributions. It should be noted that exposure is used to adjust counts on the response variable and that it is possible to use various kinds of rates, indexes or per capita measures as predictors.

The exposure variable that Prof. Fagan uses is the log of population. In other words, Prof. Fagan asserts that the number of stops is a function of the number of people who reside in a specific census tract. That assumes that all people in the tract are equally likely to be stopped for factors other than race and that an insignificant number of people come into the census tract from outside the area. It also presupposes that the population estimates used in the model are accurate and unbiased. To be correct, this measure would have to be a proxy for the probability of police contact with a person exhibiting behavior that warrants a stop. That factor would be a function of levels of traffic, the chance of police encounter and the likelihood of exhibiting a specific behavior. We argue that the exposure variable Prof. Fagan used – logged population – meets none of those requirements.

In addition, Prof. Fagan inappropriately specified the *exposure variable* he did use as the log of the population within the census tract. The use of population assumes that all people, regardless of the behavior they exhibit, are likely to be stopped by police. In addition, the use of logged population results in logging the population twice – first in the regression itself and second in the exposure variable. One can see Long and Freese (2006) for an example of the correct syntax in Stata on page 370 and 371. This has the potential effect that the true population exposure is underestimated, with the underestimation growing in a multiplicative fashion (i.e. larger population places have an even larger underestimation of the exposure than smaller places). This practice is well known for producing biased and inconsistent estimates in regression equations[30]

Moreover, Prof. Fagan uses the log of population, frozen at a fixed point in time, and the log of aggregated crime instead of modeling individual crime rates to normalize the incidence of crime for population in his model of racial disparity. This introduces two problems: 1) logs of monthly crime mask any spikes while 2) aggregation ignores crime mix. Taken together, these factors distort the relationships between crime and specific police responses to that crime. In a technical sense, they mask the relationships between policing actions and specific short-term patterns of crime, which have already been smoothed through the use of monthly data rather than a shorter time interval, leading to an additional source of bias.

In addition, the count data for stops, arrests, etc. has significant numbers of zero counts where no stops, arrests, etc. were made in a given census tract. The decision to stop or not stop—arrest or not arrest—an individual is subject to a different process than the one that drives the number of non-zero stops, etc. In cases where there are excessive numbers of zero counts driven by different processes, the zero counts must be modeled separately from the rest of the data.

---

[30] Silva, J. M. C. S., and S. Tenreyro. 2006. The log of gravity. *The Review of Economics and Statistics* 88: 641–58.

Table 7. Frequency of Zero Counts in Census Level Crime Data

Crime	0 count frequency	%
Violence	20307	30.8%
Property	10840	16.5%
Drug	37721	57.3%
Weapons	25928	39.4%
Trespass	47748	72.5%
Quality of Life	53873	81.8%
Other	37891	57.5%

There are two main approaches to doing that: zero-inflated negative binomial (or Poisson) regressions and hurdle rate negative binomial (or Poisson) regressions. The Statistical Consulting Group at UCLA Academic Technology Services (see http://www.ats.ucla.edu/stat/stata/seminars/count_presentation/count.htm) describes the models in the following way:

> Zero-inflated models attempt to account for excess zeros, i.e., there is thought to be two kinds of zeros, "true zeros" and excess zeros. Zero-inflated models estimate two equations, one for the count model and one for the excess zeros.

In a zero inflated model, the analyst specifies a variable that is hypothesized to drive the zero-generating process and uses it in the binomial portion of the zero-inflated negative binomial model. The UCLA group further observes:

> A hurdle model is a modified count model in which there are two processes, one generating the zeros and one generating the positive values. The two models are not constrained to be the same. The concept underlying the hurdle model is that a binomial probability model governs the binary outcome of whether a count variable has a zero or a positive value. If the value is positive, the "hurdle is crossed," and the conditional

distribution of the positive values is governed by a zero-truncated count model.

Hurdle models have the advantage of allowing an analyst to use the same set of predictive variables, or some subset of them, in both parts of the regression.

Given the nature of the data in this analysis, Prof. Fagan should have tested for excessive zero counts, implemented both versions of the zero-adjusting negative binomial model, selected the appropriate form and used that to generate his estimates.

The impact of zero counts can be seen by running his models with all observations with zero Terry stops in a specific census tract in a specific month removed. When we did that analysis, the coefficient for percent black which Prof. Fagan uses as a proxy for racial bias was statistically insignificant indicating that there was no bias. While this analysis is exploratory in nature, it does support the claim that the process driving the decision to stop or not stop - the zero bias – is unique and must be controlled for in any valid analysis. Prof. Fagan fails to include this crucial element in his analyses.

Table 8. Generalized Estimating Equation Regression of stops controlling for suspected crime percentages and excluding zero count observations (2010-2012)

	Total Stops	
	Coefficient	P-Value
Total complaints (logged and lagged)	1.009	(0.0208)
% violent complaints	0.936^{*}	(0.0318)
% property complaints	0.898^{**}	(0.0353)
% drug complaints	1.438^{***}	(0.102)
% weapon complaints	1.972^{***}	(0.164)
% trespass complaints	1.983^{***}	(0.197)
% qol complaints	0.886^{**}	(0.0401)
Total complaints (spatial and time lagged)	1.001	(0.00161)
% black	1.627	(0.445)
% Hispanic	2.095	(1.303)
% other race	2.014	(1.308)
SES factor	0.776	(0.122)
% foreign	0.291	(0.192)
N	44,686	

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Year fixed effects

County fixed effects

The limited number of alternative models that we were able to run in the time available generated results which either contradicted Prof. Fagan's conclusions or reduced the size of the impact he reported . For example, when we estimated an alternative model to the one Prof. Fagan used in his most recent submission to the court, we found that including lagged data on the description of suspects by race demonstrated the impact of race on *Terry* stop decisions was statistically insignificant while all of the variables that included suspect description were strongly significant.

Table 9. Panel data random effects regression of stops with suspect description 2010-2011

	Total stops per 10,000 inhabitants	
	Coefficient	P-Value
Hispanic suspects per 10,000 inh.(lagged)	0.406^{***}	(0.000)
Black suspects per 10,000 inh. (lagged)	0.876^{***}	(0.000)
White suspects per 10,000 inh. (lagged)	0.297^{***}	(0.000)
Other suspects per 10,000 inh. (lagged)	-0.363^{***}	(0.000)
Total complaints (spatial and time lagged)	0.0984	(0.684)
% black	33.62	(0.148)
% Hispanic	59.18	(0.141)
% other race	113.5^{*}	(0.028)
SES factor	-31.36^{***}	(0.000)
% foreign	-151.8^{**}	(0.002)
Constant	2.040	(0.938)
N	50,163	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Year fixed effects
Suspect description data from the merged file was available only for 2010-2011
County fixed effects

We also corrected estimated negative binomial models including suspect descriptions. In that model, race was still significant but the magnitude of the impact decreased. Our alternative analyses , then, suggest that Prof. Fagan's analyses are not robust to changes in specification and the method used to estimate the parameters. These findings provide support to the contention that police are basing their *Terry* stop decisions on empirical data rather than focusing their efforts on race.

Table 10 Generalized Estimating Equation Regression of stops controlling for suspect race ethnicity and suspected crime percentage (2010-2011)

	Total Stops	
	Coefficient	P-Value
Black suspects (logged and lagged)	0.0977***	(0.000)
White suspects (logged and lagged)	0.0384***	(0.000)
Hispanic suspects (logged and lagged)	0.0574***	(0.000)
Other suspects ((logged and lagged)	0.0270**	(0.002)
% qol complaints	-0.188***	(0.000)
% trespass complaints	0.522***	(0.000)
% weapons complaints	0.613***	(0.000)
% drug complaints	0.438***	(0.000)
% property complaints	-0.0828***	(0.000)
% violent complaints	0.109***	(0.000)
Total complaints (spatial and time lagged)	0.00642***	(0.000)
% black	0.763***	(0.000)
% Hispanic	1.102***	(0.000)
% other race	0.761***	(0.000)
SES factor	-0.0704***	(0.000)
Patrol Strength	0.0705***	(0.000)
% foreign	-0.0416	(0.672)
Constant	-1.462***	(0.000)
N	50,163	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Year fixed effects
County fixed effects
Suspect description data from the merged file was available only for 2010-2011

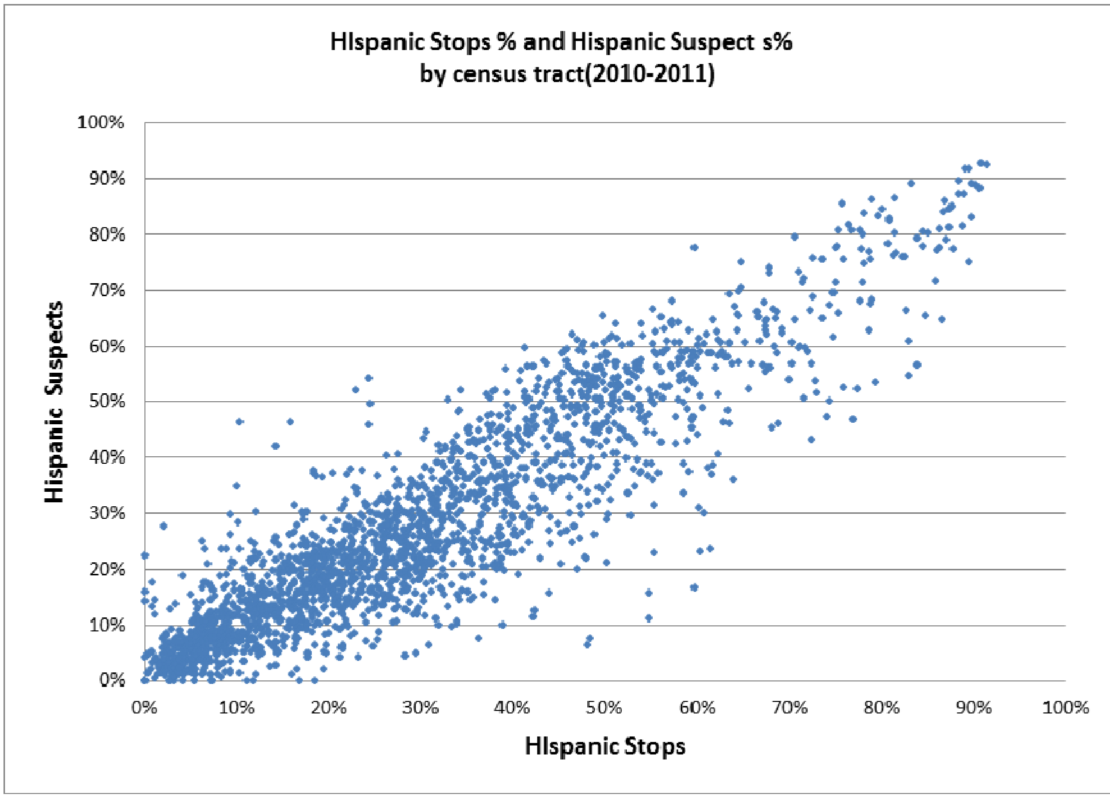
<u>Correlations between crime, area, and race</u>: There are strong correlations between race, crime, and the locations where crimes occur. To properly estimate the contribution

of race to *Terry* stop decisions, a valid analysis needs to isolate the impact of each of these factors independent of the others. To do that, the regression analyses must include interactions among and between each of these factors to separate the individual and joint effects of each. Specifically, as we showed in our response to Prof. Fagan's First Supplemental report, when descriptions of suspect by race were added to the regressions, the impact of race became insignificant and the coefficient went from positive to negative. Negative coefficients of the percentage black population in a regression suggest that Blacks were under-represented in *Terry* stop activity. Our result is consistent with the Rand Corporation report.[31], In this case, the coefficient supporting the claim that race is a determinate of stops was not statistically significant when estimated using a generalized least squares model and smaller than the one reported by Prof. Fagan. Since the parameter estimates are neither robust to changes in specification and the estimation technique used nor consistent across time, the results he presents do not support the contention that the NYPD decisions to stop an individual were influenced by race and not a combination of other factors that are correlated with race.

In the 2011-2012 data we are examining in this update of earlier reports, the scatter plots (see below) of complaints versus stops supports our previous finding. In those plots there is a very clear relationship between suspect description and the racial makeup of stops. The scatter plot of the stops by race plots are much more dispersed, showing the much weaker correlation than the relationship between racial pattern in stops and racial patterns in the suspects identified.

---

[31] Greg Ridgway, Analysis of Racial Disparities in the New York Police Department's Stop, Question, and Frisk Practices, Rand Corporation, 2007.



f

FIGURE 7. Hispanic Stops % and Hispanic Suspects % by Census Tract (2010-2011)

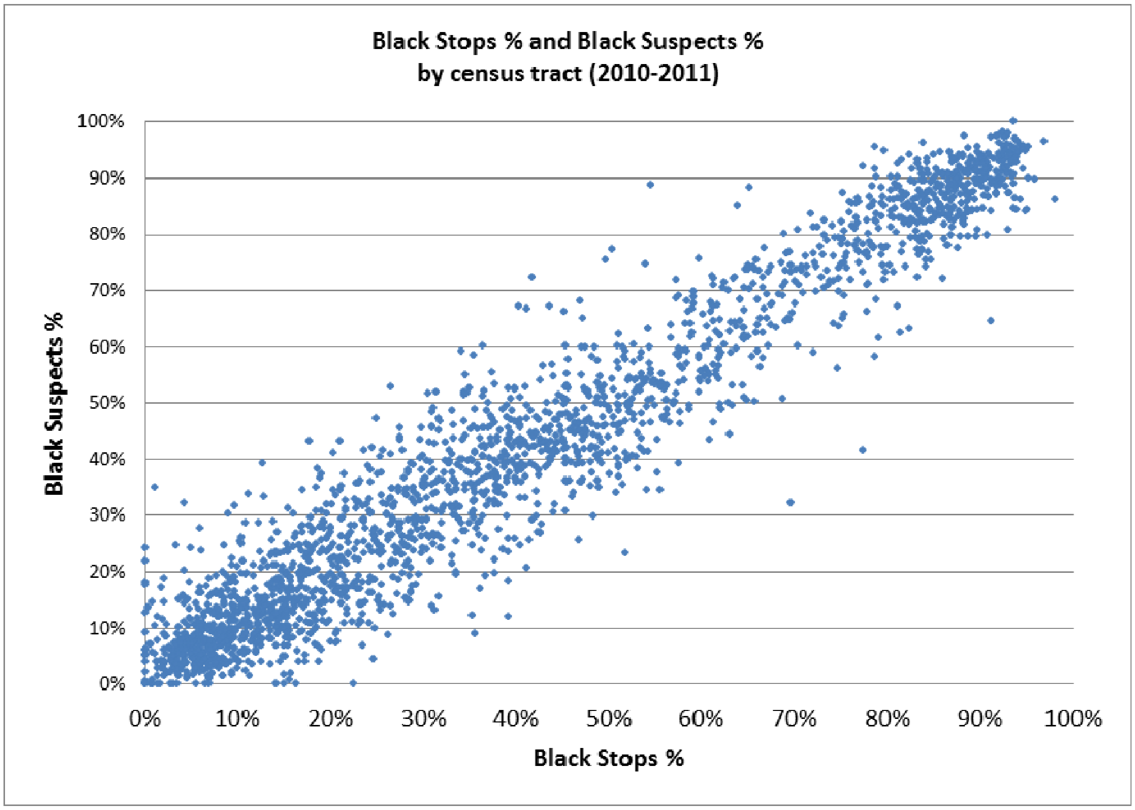


FIGURE 8. Black Stops % and Black Suspect % by Census Tract (2010-2011)

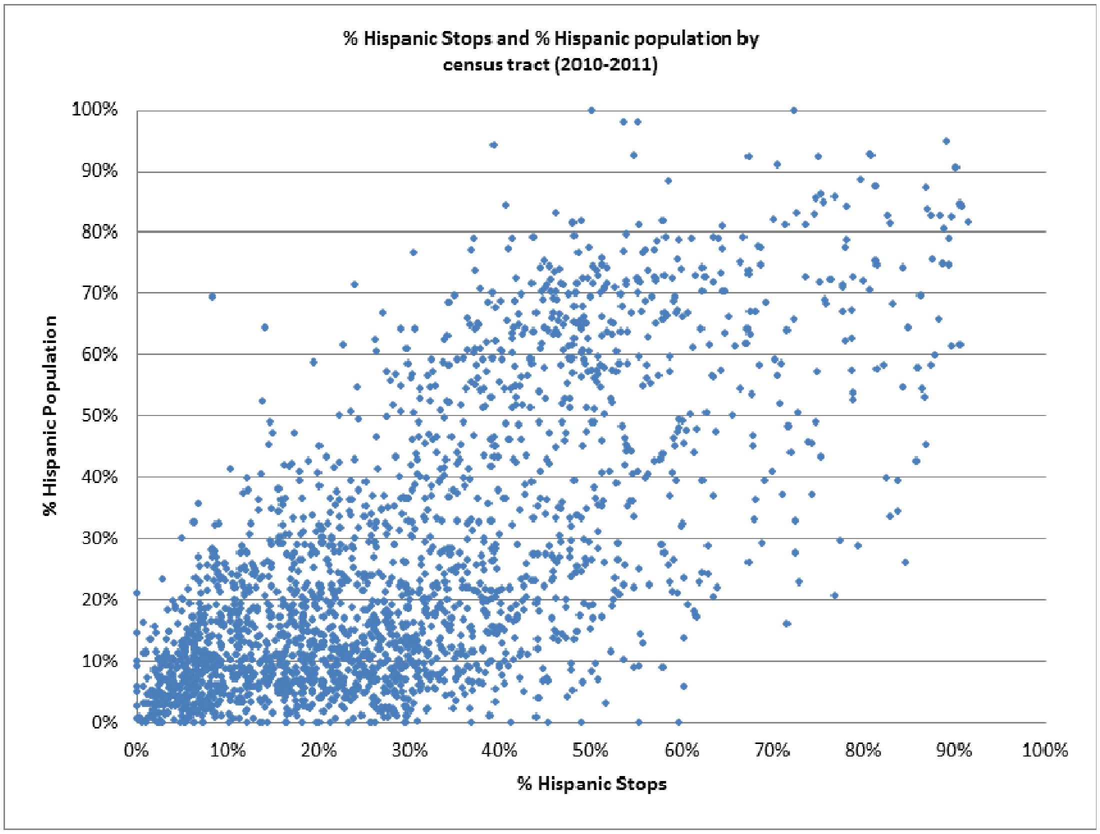


FIGURE 9 % Hispanic Stops and % Hispanic population by census tract (2010-2011)

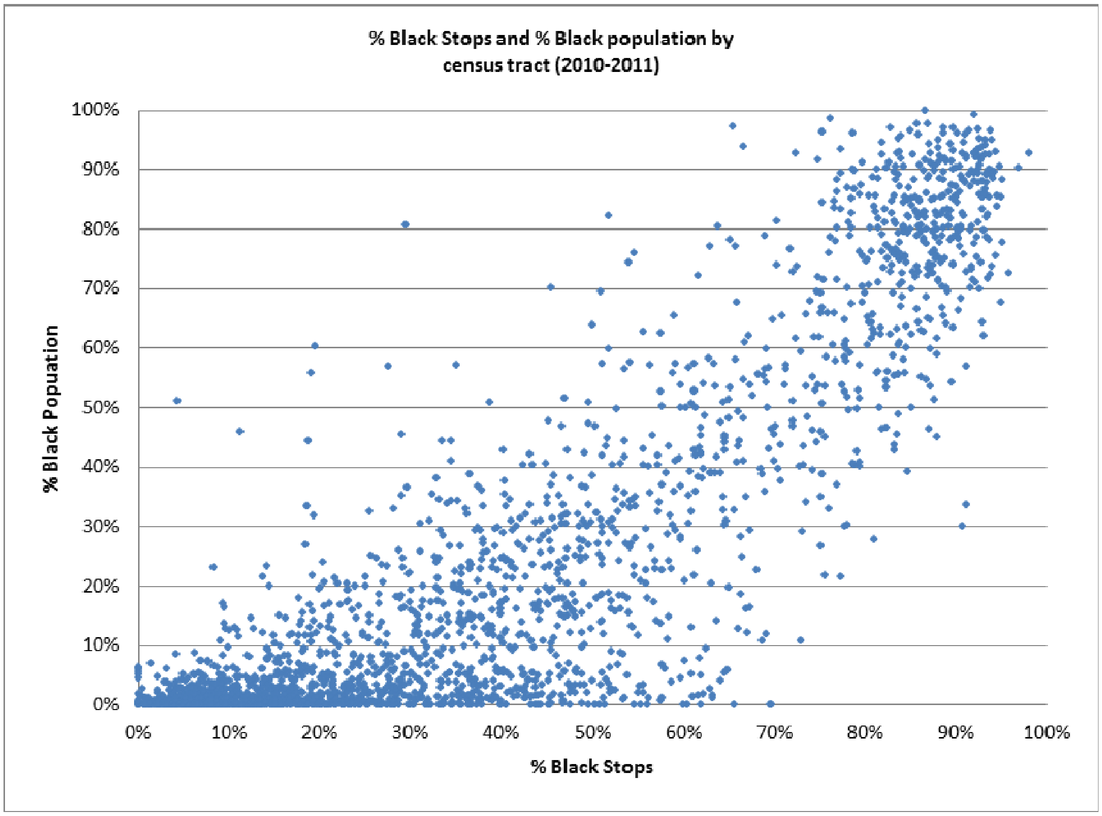


FIGURE 10 % Black Stops and % Hispanic population by census tract (2010-2011)

In these graphs and the ones included in our reply to early reports by Prof. Fagan, the contrast in the dispersion of dots is dramatic evidence of the weaker correlation of race alone and stops, and race of those actively engaged in crime.

A good example of how stop patterns are related to crime patterns rather than population is in census tract with Census ID 6111300. This census tract has a total black population of 36 (18.4% of the total census tract population) but averages 146.3 monthly black stops for the period 2010-2011. This census tract averages 33.8 complaints of blacks committing crimes for the same period. The black stops are 53% of the total stops where the race is known, while the suspect identified represent 53.6% of the crime complaints with known race.

In his models, Prof. Fagan fails to include a suspect description by race for each crime, which has proven to be a better predictor of the racial makeup of those stopped than individual measures of crime and race taken separately. This is an example of a missing variable bias. Prof. Fagan's results are evidence of correlations, not causation. The strong correlations between the data – crime, area populations, etc. – that police use for both deployment and *Terry* stop decisions at the patrol level, and race compound the problem of attributing cause based on a statistical correlation. Moreover, our analysis indicates an appropriate accounting of the relevant variables indicates race is not a significant factor in *Terry* stop decisions.

<u>Failure to Include Time Series Dynamics</u>: Prof. Fagan does not acknowledge the effect that changes in police practice might have had over time. Nor does he allow for the impact of time specific crime patterns on police responses. Rather, he only allows for annual fixed differences in the relationships among crime, race and location across all precincts in past analyses, census tracts in his Second Supplemental Report. Using the coding scheme developed by Prof. Fagan, with which we take issue, the graph below shows that in at least the dimension of legal sufficiency there has been a downward trend in unjustifiable trespass stops. Analogous time-dependent changes exist in crime counts, population size, demographic factors and stops. Failure to control properly for time leaves the issue of serial autocorrelation within each unit of observation unresolved, which can thereby result in biased estimates. To control for this and determine whether there were period specific differences in the stop trends, Prof. Fagan needed to introduce a time variable and interact it with the key factors in his model. He failed to do so.

This is compounded by the issue of changing crime patterns and changes in what the police would define as "high crime" areas over time, especially at the census tract level, both in terms of location and the time of the occurrence of the circumstances that might warrant a stop. The dynamic nature of crime patterns and police deployment decisions may well be a contributing factor to the incidence of census tracts reporting zero stops in a given time period. Fagan's analysis implicitly assumes a stagnant environment but patrol officers face one that is constantly changing. Without including some measure of dynamics in his models, Prof. Fagan fails to capture any of these factors that might influence *Terry* stop decisions. As Table 18 shows changes overtime are the rule for trespass stop, like other crimes suspected patterns.

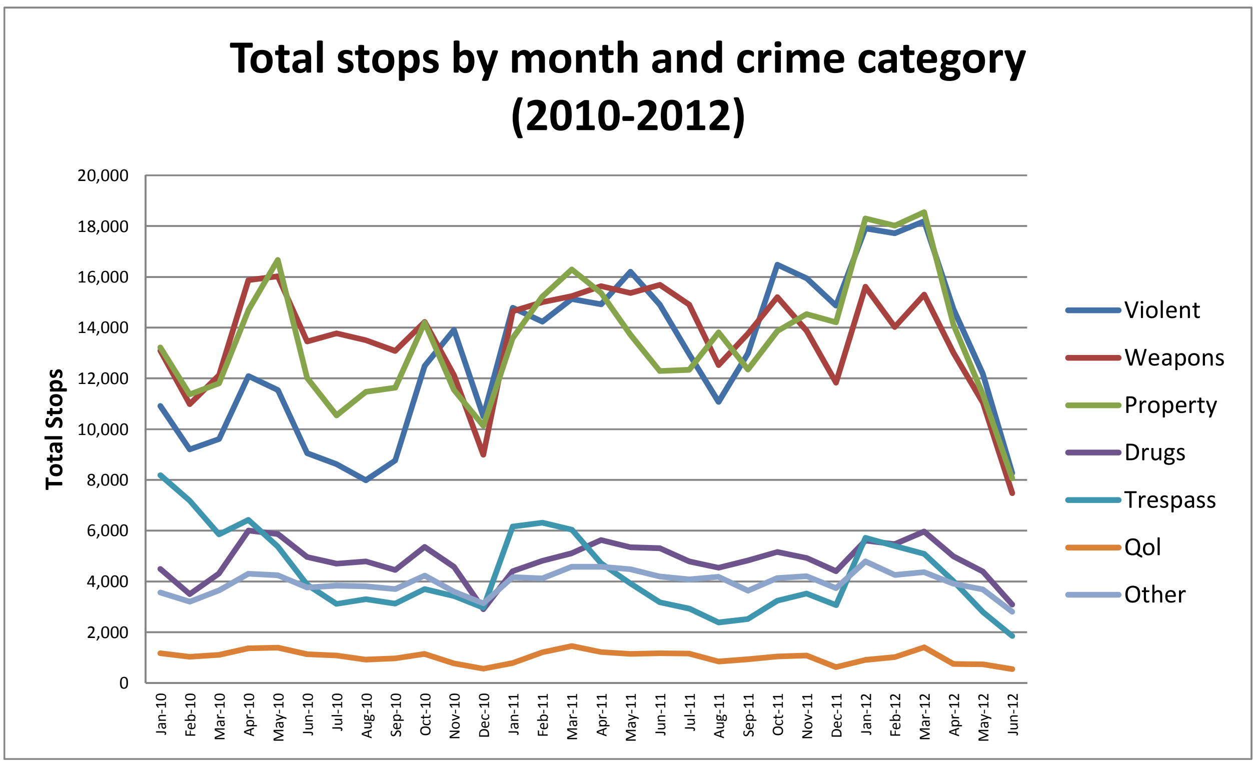


FIGURE 11. .

As we illustrated in the body of the report, the results for other stops follow similar patterns.

Failure to Control for Changing Crime Mix: By modeling aggregated crime rather than looking at the actual mix of individual crimes in a geographic area, Prof. Fagan implicitly assumes that all crimes are equal with respect to likely police responses. To properly control for the bias this can introduce, he needed to control for intra-crime variation – for example a domestic homicide versus a gang shooting, which are likely to illicit dramatically different police responses on the street – and crime patterns that cut across crime categories -- for example an increase in drug related crime versus larceny. He assumes that crime patterns are constant in both dimensions and that the police respond to aggregate crime data rather than patterns of specific crimes and that policing/*Terry* stop responses are the same regardless of the changing patterns. This compounds both the time-series and the missing data issues discussed above.

To correct for these problems, Prof. Fagan needs to include measures of individual crimes in a model that accounts for the time series dynamics . Differing responses based on patterns of crime and other available data are consistent with modern policing practice but missing in Prof. Fagan's models. At the very least, he should use factor analysis to generate a crime index (if the crimes are highly collinear) or include individual crime counts if they are not. In the time available, we were unable to complete that alternative analysis.

Modeling for Change in NYPD Training Procedures: Prof. Fagan chose to model the most recent data used in his Second Supplemental report as if it were independent of prior results, was not impacted by time trends and was independent of the NYPD change in training practices. A more generally accepted approach to including the most recent data in his analysis would have been to acknowledge the time series components in the data as we have argued above, to include the most recent period in the model and specify the period of time after the NYPD's change in training procedures in an interrupted time series design with the post training periods identified using dummy variables and those dummies interacting with time. That would have allowed him to control the findings for the dynamics of the situation without losing the ability to distinguish between older and more recent, post-training results.

Police Practice: Prof. Fagan bases his analysis on, and draws inferences from, the more restricted set of information than the police use to make both macro patrol deployment decisions at the command level, and micro *Terry* stop decisions at the patrol level. The department uses information about special events like the St. Patty's parade, the Caribbean festival in Brooklyn, major sports events, political events like rallies and conventions, terrorism alerts, and the NYC Marathon, in addition to crime trend reports, to alter patrol deployment decisions.

At the precinct level, individual officers have roll call briefings, many now have experience in Impact Zones, some have relatively stable assignments and are familiar with many of the people in their patrol areas. In addition, there are many different specialized patrol assignments in NYPD, some of which are formally organized around crime issues and patterns that intersect with --but are not based on-- race, ethnicity, and national backgrounds. Additionally, individual officers have local knowledge about their specific patrol sectors. The NYPD makes all of this information available to patrol officers for use in their day-to-day, minute-to-minute decisions. In addition, the information available to command and patrol level officers differs by type of unit. Specialized units like Housing, Transit, Narcotics, Organized Crime, all have different sources of information and priorities. Prof. Fagan fails to even acknowledge the rich data set used in modern policing, much less account for the available data in his analysis.

Prof. Fagan measures patrol strength based on the number of officers who make one or more stops in a specific census tract. He ignores the presence of partners. He also fails to account for the low incidence of stops per officer – on the order of 2-3 per month per officer on average. Further, he ignores the mobility of police in patrol areas. As the table below shows, it is not uncommon for officers to make stops in multiple census tracts in a given month and may not be uncommon for them to do so in a given day.

Table 11. Distribution by Census Tracts of Officer Stops

# of census tracts officers made stops during a month	
# of census tracts	**%**
1 census tract	33.4%
2 census tracts	21.9%
3 census tracts	14.6%
4 census tracts	10.0%
5 census tracts	6.5%
More than 5 census tracts	13.8%

In addition, Prof. Fagan ignores differences in patrol strength by time of day and between types of officer assignments. For example, a routine patrolman will make stops when time permits and the need presents itself, but an officer with a special assignment (e.g. drug enforcement) will be more likely to devote more time to making stops. These special assignment officers are also more likely to be assigned to high crime areas, suggesting the measurement of the number of unique officers in a census tract is not a reliable measure of patrol strength. To this general observation must be added the fact that Prof. Fagan's analysis in this Second Supplement Report, unlike his Report in the Davis case, ignores the widespread use of "special assignment" officers by NYPD during the time of his study. The Second Supplemental Report ignores completely the City's reliance on Operation Impact in its effort to prevent crime. Impact Zone officers are not mobile and are more likely than officers on other assignments to make stops, again a fact acknowledged by Prof. Fagan in his Davis Report. In Prof. Fagan's new measure of patrol strength, after the officer's first stop in an Impact Zone deployment, that stop results in one officer being added to patrol strength for the census tract for that month, the same as an officer who makes one stop in that tract while en route to another assignment.

The impact of this approach to estimating patrol strength leads to significant measurement issues. According to the patrol-strength data generated by Prof. Fagan, there were 9,051 observations – census tract months – that have no patrol presence at all (13.74% of the observations). There were 9747 (14.7%) census areas that had patrol strength of 1 despite the fact that officers typically work with partners. At the other extreme, there were census tracts that have patrol strength of 144.

Consider that one officer (reptaxnm=108355) made stops in 32 different census tracts in the same month. It is clear that estimating patrol strength by counting officers who make a single stop is subject to errors and removes any interpretable variation between "patrol strength," as now measured by Prof. Fagan, and stops. The way Prof. Fagan has constructed his measure, stops and patrol strength are endogenous. Stops are used to determine patrol strength but patrol strength is used as a factor in the analysis of stops in Prof. Fagan's regression analyses. From a purely statistical perspective, Prof. Fagan's measure of patrol strength is a serious flaw in his analysis, that casts his conclusions as dubious.

If, as Prof. Fagan asserts, race is a causal determinant of *Terry* stop decisions, then any model that supports that claim must include as much of the relevant data that may have a significant impact as possible. Excluding the available and explicit decision making bases from the model potentially masks other factors that are heavily correlated with a variable such as race that might make race appear a factor when it is not. To the extent that relevant data are excluded from the model or incorrectly measured, the parameter estimates, as well as the error terms used to estimate statistical significance of those estimates, are biased, and therefore valid inferences cannot be drawn from the results of the model.

<u>Population:</u> The second significant data issue in Prof. Fagan's analysis is his use of population. There are several problems with his usage. First, he uses a static measure of population rather than trending population estimates across time, the latter of which is the most common practice in criminal-justice modeling. Second, he does not control for the

well-known level of sampling error in population estimates for very small areas. As a geographic area gets smaller, the accuracy of population estimates derived from samples falls. In addition, he does not adjust for ghost population – illegal immigrants not accurately represented in either the census or subsequent population samples, and tenants who are not on leases or variations in population within a census tract by time of day.[32] (Smith, 1987, Smith & Shahidullah 1995)

In addition, errors between census and available population (people actually in the area) grow more disparate for smaller areas. For instance, Manhattan's population swells by 20% for commuters during the daytime (Ridgeway 2007 and Miller 2000). A comparison of observational counts of residents found the census under-counted available minority populations by as much as 300%. Such a significant underestimate casts into doubt the comparisons Prof. Fagan makes to minority population figures in a census tract.

Potentially of even greater concern is the use of census tracts as a measure of population. Using population within a census tract implicitly assumes that people are not mobile. The proper population measure would need to control for stops of people from other census tracts who were stopped in the census tract where the stop was reported but live in other areas. This might be particularly important in census tracts with significant numbers of stores and offices, those that draw tourists, areas with significant numbers of commuters and areas where drug trafficking is prevalent. The only way census tracts could be an accurate reflection of the population in an area at a given time of day is if population flows followed a purely random pattern such that it did not significantly alter

[32] Smith, 1987, Smith & Shahidullah 1995

the demographic, particularly the racial, composition of an area. One extreme example of areas with shifting populations is Central Park which Prof. Fagan includes in his analysis but which has no permanent population. The implicit assumption that daily population flows have no impact on demographic composition is a highly dubious assumption but one of several unstated assumptions Prof. Fagan seems to make.

Interpreting the Practical Significance of the Findings. The results obtained from statistical analyses can be both statistically significant and simultaneously inconsequential. We argue that is the case with the evidence of racial bias offered by Prof. Fagan. To demonstrate that, we have converted the coefficients for the percent black presented in Table 5 on page 18 of his Second Supplemental Report to the court to show the impact of increasing the percent black in a given area has on the likelihood of those persons being stopped. As we argued earlier, the relationship between lagged reported crimes – used as an independent explanatory variable – and the suspected crime shown on the Form 250 as the reason for the stop is tenuous. NYPD reacts to a much richer set of information including observed behavior when deciding whether to conduct a Terry stop. It would not be surprising to find weak relationships between reported crime counts and specific reasons for a stop. We now show, in Table 20 below we show the impact of population changes for all of the stop reasons reported by Prof. Fagan.

The raw percent-black coefficients in Professor Fagan's regression results in Table 5 show the impact of a one-percent change in the black population on what is called the log-odds ratio of a person being stopped for the reason shown on the Form 250. Since log odds ratios are not obvious to non-statisticians, we have converted them to odds ratios which are easier to interpret. We have done that for each of the coefficients reported in Table 5.

To illustrate the concept of an odds ratio, let's start with a coin tossing example. If one tosses a fair coin the chances that it will come up heads or tails is 50%. To test to see if the coin is fair we can either look at the odds of each outcome to see if they are the same or create a ratio of the two probabilities of occurrence. In this case of a coin toss, that would be 50% divided by 50% or 1. When things are equally likely to happen, the odds ratio is one.

The process for converting logs odds ratios for a 1% increase or decrease in the proportion of blacks in an area to odds ratios for the same changes in proportion is as follows: We multiply the coefficient for the percent-black variable in each column of Table 5 by 1% to find the impact of a one-percent change in the black population on the log odds that they might be the subject of a Terry stop. For the "total stops" column the calculation is .883 x .01 or .0083. That is, the impact of a one-percent change in percent black on the log odds of a black person being involved in a Terry stop is .0083. Taking the antilog of .0083 results in an odds ratio of roughly 1.009.

Since this transformation does not alter the meaning of the findings, we can interpret this as meaning that a 1% change in the black population results in a slight increase in the probability of being stopped. Most analysts interpret an odds ratio impact of approximately one as having *no practical significance* even if the result is statistically significant. (Of course, with data sets the size of those used in this analysis, infinitesimal differences will be *statistically significant.* As Table 20 below shows, only the results for drugs, weapons, trespass and QOL/Disorder Terry stops had odds ratio impacts of .01 or more, and even these results are minimal. In short, by using this odds ratio analysis of Professor Fagan's findings we demonstrate that the "statistically significance" of race in his equations has no practical significance in terms of a black person's likelihood of being stopped. If, as is undeniably the case, Blacks and Hispanics are stopped at a higher rate than their share of the population, we must look to other factors for an explanation. As we have shown when suspect description is entered in to our regression, even statistical significance of race disappears.

TABLE 12. Illustrating the Impact of the Effect Size

Presented in Table 5 of Prof. Fagan's 2nd Supplemental Report

Crime shown on the 250	TotalStops	Violent Crime	Property	Drugs	Weapons	Trespass	Other Crimes	QOL Didorder
Percent Black Coefficient	0.883	0.938	0.281	1.042	2.078	1.121	-0.518	-1.389
1% log odds ratio impact	0.00883	0.00938	0.00281	0.01042	0.02078	0.01121	-0.00518	-0.01389
1% odds ratio impact	1.0089	1.0094	1.0028	1.0105	1.0210	1.0113	0.9948	0.9862

Taken together, these modeling, structural and data issues, along with the other problems described above, have the potential to generate biased parameter estimates and make it impossible to draw valid inferences from the models presented by Prof. Fagan as evidence of racial bias in *Terry* stop activity.

This report is signed this 1st day of February, 2013 in Rensselaerville, New York

Dennis C. Smith, Ph.D.

This report is signed this 1st day of February, 2013 in Rensselaerville,New York

Robert M. Purtell, Ph.D.

Appendix A

CURRICULUM VITAE

DENNIS CHARLES SMITH
Associate Professor of Public Policy

NEW YORK UNIVERSITY
ROBERT F. WAGNER GRADUATE SCHOOL OF PUBLIC SERVICE

Professor in Residence, New York State Assembly

The Puck Building 295 Lafayette, Room 3073
New York, NY 10012
telephone: 212- 998-7425
telefax: 212- 995-4162
e-mail: dennis.smithde@nyu.edu

BORN: August 12, 1945 - Chicago, IL

DEGREES:
B.A. University of New Mexico - Political Science 1967
M.A. Indiana University - Political Science 1969
Ph.D. Indiana University - Political Science 1976

AWARDS AND HONORS:
B.A. *magna cum laude* in Political Science
magna cum laude in General Honors

Phi Beta Kappa
Woodrow Wilson fellowship

National Defense Education Act Fellowship, Title IV, Department of Political Science, Indiana University

TEACHING:
Associate Professor (Fall 1978 - present)
Professor in Residence New York State Assembly(Spring,2006-present)
Assistant Professor (1973 - 78)
Robert F. Wagner Graduate School of Public Service
New York University
Lecturer in Political Science, Indiana University Southeast, Jefferson, Division of Social Sciences (1969-71)
Visiting Professor Boconni University, Milan, Italy. Fall, 2005.
Visiting Professor, American University in Paris, Spring 2007 to present.

ADMINISTRATION at Robert F. Wagner Graduate School of Public Service
New York University:

Program Manager, Wagner/Accenture Leading Large Scale Change Executive Briefing Series, July 2007- present.
Director, Wagner International Initiative, 1998-2002.
Director, Public Policy Specialization, 1992-1997
Director, Public Administration Program, 1982-90
Associate Dean, 1986-88

RESEARCH/CONSULTANT:

Expert witness for NYC Department of Law in Floyd v City of New York and Davis v, City of New York (cases involving evaluation of NYPD Stop. Question and Frisk policy and practice)

Evaluation consultant, Assessment of Public Involvement Strategies of the NY Metropolitan Transportation Authority, Federal Transportation Administration funded project,. Rudin Center for Transportation Policy and Management,

Consultant: Office of the Commissioner, NYC Department of Environmental Protection, Study of the organization and management of the DEP Police, May, 2007 to April, 2008.

Consultant., Office of the Commissioner, New York City Police Department, Assessment of Operation Impact: Strategies to reduce crime hotspots in New York City. November, 2005- June 2007.

Consultant, Office of the Commissioner, New York City Police Department, Assessment of the process of maintaining the integrity of crime reports, 2005.

Consultant, NYC Human Resources Administration, Assessing the implementation of the WeCare Initiative, 2005-present.

Consultant, Office of the Commissioner, New York City Police Department, Assessment of Borough Command Structure, 2003-2004.

Consultant, Office of the NYC Deputy Mayor for Operations, Project on Performance Based Contracting 2002-2004.

Consultant, Charles Hayden Foundation, Evaluation of General Support Initiative, 1996-1998

Consultant, Dewitt-Wallace-Reader's Digest Fund, Evaluation of the "Management Initiative" A Program to Develop the Management Capacity of Youth Serving Organization (1995).
Consultant, New York City Police Foundation, Study of the Recruit Training Program of the New York Police Department Academy (1994)

Co-Director, study of ambulance service in New York City, with James R. Knickman, Health Research Program, with support from the Commonwealth Fund (1989 - 1991)

Director, study of the New York City Mayor's Management Planning and Reporting System, in collaboration with Barbara Gunn, Director of the Mayor's Office of Operations, with support from the Fund for the City of New York (1988-90).

Co-Director, study of the impact of retrenchment on the New York City Police Department (1980-81), under the auspices of the NYPD Research Advisory Committee.

Principal Investigator of "A Two-Wave Panel Study of the Impact of Education on Police Attitudes and Performance," a study funded by the Office of Criminal Justice Education and Training, Law Enforcement Assistance Administration (1977-80).

Director of Survey of Police Officials in the Police Services Study, Phase II, a study of police performance in three metropolitan areas under a grant from the National Science Foundation, Research Applied to National Needs Division (full time, Summers 1976 and 1977).

Design and execution of a study, "Institutional Arrangements and the Police: St. Louis Metropolitan Area," 1971-73, a grant to Dr. Elinor Ostrom, Associate Professor of Political Science, Indiana University, from the Center for Studies of Metropolitan Problems, National Institute of Mental Health.

Research Consultant, studies on citizen evaluations of the police in Indianapolis and Chicago metropolitan areas, 1969-71, under a grant to Dr. Elinor Ostrom from the National Science Foundation.

Research Associate, "The Organization of Government Response to Civil Disorders in Indiana," under a grant to Dr. Philip S. Kronenberg, Assistant Professor of Political Science, Indiana University, from the Law Enforcement Assistance Administration, Summer 1969.

INTERNATIONAL EXPERIENCE

Director, Wagner International Initiative, 1998-2002.

Taught International Capstone course at NYU Wagner, directing teams of students in year long capacity building projects for UN agencies and international non governmental organizations that have worked in Mozambique, Bosnia, South Africa, Cote D'Ivoire, Indonesia, Peru, Nicaragua, Tanzania, Egypt, Brazil, Mongolia, Uganda, and India. Current team project in Jordan, developing and leading a public management training program for young government officials for the Education for Employment Foundation, at the request of the Royal Court., 1995 to present.

Designed and lead a new Masters degree program at NYU Wagner in Management of International Public Organizations, 1998-2002.

Directed NYU Wagner Partnership in Democratic Decision-making with the Ukrainian Institute of Public Administration, jointly with the European Institute of Public Administration
(funded by US Department of State and the European Union)

Directed University Partnership with Eduardo Mondlane University, Maputo Mozambique, to develop Public Policy Analysis Curriculum and faculty development (with Department of State Funding)

Directed an evaluation of family heath initiative in Southern Africa operated by Planned Parenthood of New York with USAID funding.

Taught summer course on The Use of Policy Analysis in Europe at the Ecole Nationale de Travaux Publics de l' Etat, 1992-1999.

Taught televideo course on European Union/American Union with the Institute of European Studies of the Universite Libre de Bruxelles, 1996- 2001, with European Union and Department of State funding.

Participant in TransAtlantic Dialogue of the American Society of Public Administration and the European

PROFESSIONAL ACTIVITIES:

Member, Task Force on Performance Reporting, NYC Mayor's Office of Operations, Spring, 2012

Keynote speaker, New York State Leadership and Accountability Conference, Albany, May, ` 2008.

Senior Consultant on Performance Management, SEEDCO/N-PAC, 1996 to present.

Member of Board, Institute of Public Administration, 1998-present.

Member, New Progressive Scholars Network of the Progressive Policy Institute, 1996-present.

Consultant, Innovation in Government Award Program, Harvard University, 1990 – 1998.

UNIVERSITY AND COMMUNITY SERVICE:

Faculty Advisory Committee, King Juan Carlos I Center at NYU, 1998- present.

Member, Faculty Advisory Committee, European Union Center at New York University, 1998-present.

Member, NYU Graduate Commission, 1996-present.

Chairman, Subcommittee on Graduate and Professional Education, New York University,
Chancellor's Task Force on Internal and External Communication., 1983

Chairman, New York University Faculty Council, 1982-83

Vice-Chairman, New York University Faculty Council, 1980-82

Member, Editorial Board, NYU Press, 1980-83.

Member, Presidential Search Internal Advisory Commission, 1980-81

RESEARCH AND TEACHING INTERESTS:

Management of international public service organizations

Performance measurement in the public and nonprofit sector

Program evaluation and public policy impact analysis

Urban public service delivery systems

EDITORIAL BOARD MEMBERSHIP:

The Journal of Comparative Policy Analysis
Policy, Organization, and Society

ARTICLES AND PUBLICATIONS:

"A Multi-Strata, Similar Design for Measuring Police Performance," with Elinor Ostrom and Roger B. Parks; paper presented at the Annual Meeting of the Midwest Political Science Association (Chicago, 1973).

"The Effects of Training and Education on Police Attitudes and Performance: A Preliminary Analysis," with Elinor Ostrom, in Herbert Jacob, ed., **The Potential for Reform of Criminal Justice** (Volume III, **Sage Criminal Justice Systems Annuals**, 1974).

"On the Fate of Lilliputs in Metropolitan Policing," with Elinor Ostrom, **Public Administration Review** (March-April, 1976). Earlier version presented at the American Society for Public Administration meetings in Chicago, April 2-5, 1975. Excerpts from this paper comprised the main article in the **Criminal Justice Newsletter: A Bi-Weekly Report on Significant Developments for Leaders in Criminal Justice Administration**, Vol. 6, No. 11, May 26, 1975. Edited version in D. Hagman, **Public Planning and Control of Urban and Land Development** (West, 1980).

Police Professionalization and Performance: An Analysis of Public Policy from the Perspective of Police as Producers and Citizens as Consumers of a Public Service (unpublished Ph.D. dissertation, Indiana University, 1976).

"Dangers of Police Professionalization: An Empirical Analysis," **Journal of Criminal Justice**, Vol. 6, Fall 1978. Earlier version presented to the American Society for Public Administration, Annual Meetings, Washington DC, 1976).

"Police Attitudes and Performance: The Impact of Residency," **Urban Affairs Quarterly**, Vol. 15, No. 3 (March, 1980).

The Effects of Higher Education on Police Performance: A Critical Review of Findings, a consultant report for the National Advisory Commission on Higher Education for Police Officers, Washington DC: The Police Foundation, 1978.

"Racial Context as a Factor in Changing Police Organizations," with Diane Baillargeon, presented at the Annual Meeting of the American Society for Public Administration, Phoenix AZ, 1978.

"Value Biases in Performance Assessment," presented at the Annual Meeting of the American Political Science Association, New York, 1978. Accepted for publication in **Evaluation Review**.

"Reforming the Police: Organizational Strategies for the Urban Crisis," in Joseph Hawes, ed. **Law and Order in American History**; Port Washington NY: Kennikat Press, 1979.

Educating the Police: An Interim Assessment, with Diane Baillargeon, the final report of "A Two-Wave Panel Study on Police Attitudes and Performance" to the Office of Criminal Justice Education and Training (LEAA Grant 78-CD-AX-00027, August 1979).

Booking the Police: Police Education Re-examined, with Diane Baillargeon, the final report of "A Two-Wave Panel Study..." (LEAA *op. cit.*) An earlier version was presented at the annual meeting of the American Society for Criminology, 1979.

"In Pursuit of Safety: Alternative Patterns of Police Production in Three Metropolitan Areas," with Diane Baillargeon, in **Journal of Social Issues**, Vol. 30, No. 4 (1980).

"Police," in **Setting Municipal Priorities, 1982**, Charles Brecher and Raymond D. Horton, eds., New York: Russell Sage Foundation, 1982. Reprinted in **Setting Municipal Priorities: American Cities and the New York Experience**, C. Brecher and R.D. Horton, eds., NYU Press, 1984.

John Mathiason and Dennis Smith, "The Diagnostic of Reform: The Evolving Tasks and Functions of the United Nations," **Public Administration and Development** (Vol. 7, No.2, 1987).

Performance Management in New York City: A Review of the Mayor's Management Plan and Reporting System (Preliminary Report, October 1990).

Improving Ambulance Use in New York City: A Final Report (with James R. Knickman and Carolyn Berry) New York University Health Research Program report to the Commonwealth Fund, March 1991.

"Managing the Demand for Emergency Service: The New York City EMS" (with James R. Knickman and Carolyn Berry); a paper presented at the 13th Annual Research Conference of the Association of Public Policy and Management, Denver, Colorado, October 1992.

"HRA Adrift: Social Spending without Direction" (with William Grinker) in **City Journal**, September 1993.

"Performance Management in New York City: The Mayor's Management Plan and Report System in the Koch Administration, a paper presented at the 15th Annual Research Conference of the Association of Public Policy and Management, Washington, D.C., October, 1993.

"Managing Organizational Transformations: The Case of Problem-solving Community Policing in New York City," a paper presented at the 16th Annual Research Conference of the Association of Public Policy and Management, October, 1994.

“Implementing UN CIVPOL: The Challenges of International Public Management, presented at the International Studies Association Toronto Convention, March 19, 1997

“What can public managers learn from police reform in New York? COMSTAT and the promise of performance management,” presented at the 19th Annual Research Conference of the Association of Public Policy and Management (APPAM) in Washington, D.C., Nov. 6-8,1997.

“Using Technology to Create International Educational Partnerships,” a paper presented at parallel plenary sessions at the 50th Anniversary Conference of the Council on International Education Exchange in Barcelona, Spain, November 18-20, 1997.

"Making Management Count: Toward Theory-Based Performance Management," (with R. Barnes) 20th Annual Research Conference of the Association of Public Policy and Management (APPAM) in New York, NY., November 2-4,1998. Revised version submitted for final review to **Nonprofit Leadership and Management**.

"Performance Management in New York City: COMPSTAT and the Revolution in Police Management,"(with William Bratton) in **Quicker, Better, Cheaper ?: Managing Performance in American Government,** edited Dall Forsythe, SUNY Press Albany, 2001.

"Electronic government, transparency, and performance management in the governance of cities," a paper presented at the United Nations/Metropolitan Seoul Conference on E-Governance, Seoul, Korea, August, 2001.

"Old Wine, New Bottles? The Distinctive Challenges of Managing International Public Service Organizations," A paper presented at the 23rd Annual Research Conference of the Association for Public Policy Analysis and Management (APPAM) in Washington DC, November 1-3, 2001.

"Managing UNCIVPOL: The potential of performance management in international public services," in Dijkzeul, D., Beigbeder, Y (eds.) **Rethinking International Organizations: Pathologies and Promise**, Berghahn Books, Oxford/New York, 2003.

“The Promise and Pitfalls of Performance Base Contracting.” A paper presented at the 25th Annual Research Conference of the Association for Public Policy Analysis and Management (APPAM) in Washington DC, November 6-7, 2004.

“ An Empirical Assessment of Seven Years of SATCOM: The NYPD Command Structure in Brooklyn North” (with Joseph Benning) A paper presented at the 26th Annual Research Conference of the Association for Public Policy Analysis and Management (APPAM) in Atlanta, Georgia November 3-5, 2005.

“The Transformation of Social Services Management in New York City: “CompStating” Welfare” (with William Grinker) A paper presented at the 26th Annual Research Conference of the Association for Public Policy Analysis and Management (APPAM) in Atlanta, Georgia November 3-5, 2005.

“Partners in Performance: Effectiveness and integrity in the public sector,” with Frank Anechiarico, paper presented at the ASPA conference “Ethics and Integrity in Governance: A Trans-Atlantic Dialogue, in Leuven, Belgium, June 1-3, 2005.

"Practice, practice, practice: The education and training of policy analysts at NYU/ Wagner" in Iris Geva-May ed., **Thinking Like a Policy Analyst: A Clinical Approach to Policy Analysis,** Palgrave, 2005.

“Putting it all together: E-government, Transparency and Performance Management.” Presented at the APEC/Korean Independent Commission Against Corruption Seminar on E-government, Transparency and Governance, Seoul, Korea, September 1-2, 2005.

“Managing for Performance and Integrity: Administrative Reform in New York City Government” (with Frank Anechiarico). Presented at the Annual Meetings of the American Society for Public Administration, April 4, 2006, Denver, Colorado.

“Performance as Integrity, Integrity as Performance: A New Paradigm for Public Administration” (with Frank Anechiarico). Presented at the ASPA conference “Public Sector Performance: A Trans-Atlantic Dialogue, in Leuven, Belgium, June 1-3, 2006. Also presented at City University of Hong Kong, June 9, 2006.

“Crime Reduction and Economic Development in New York City: The Re-distributional Effects of Improving Public Safety “ (with Robert Purtell) A paper presented at the 27th Annual Research Conference of the Association for Public Policy Analysis and Management (APPAM) in Madison, Wisconsin, November 3-5, 2006.

“An Empirical Assessment of NYPD’s ‘Operation Impact’: A Targeted Zone Crime-Reduction Strategy” (with Robert Purtell), a paper presented at the APPAM Annual Research Conference, Washington DC, November, 2007.

“Can New York CompStat State Government Performance?” an invited paper presented in Workshop on Performance Measurement in Multi-level Governments at the 4th TransAtlantic Public Administration Dialogue in Milan, Italy, June, 2008.

“Does Stop and Frisk Stop Crime” (with Robert Purtell) A paper presented at the 29[th] Annual Research Conference of the Association for Public Policy Analysis and Management (APPAM) in Los Angeles, California, November 6-9, 2008.

“Evaluation of the New York Integrity System” in Local Integrity Systems: World Cities Fighting Corruption and Safeguarding Integrity, edited by Leo Huberts,et al.,BJU Legal Publishers, 2008.

“Right from the Start: The Managerial Advantages of Combining Effectiveness and Integrity in Policy Design,” (with Frank Anechiarico) paper presented and annual research conference of the Association of Public Policy and Management, Washington DC, November 5-7, 2009.

“Implementing Police Management Reform: the diffusion of Compstat in the cities of New York State” With Robert Purtell, paper presented and annual research conference of the Association of Public Policy and Management, Washington DC, November 5-7, 2009.

“Making Management Count: A case for theory and evidence based public management,” **Journal of Policy Analysis and Management**, Summer 2009.

“Are New York State’s Public Authorities Performing Well? Who knows?” **NYSBA Government, Law and Policy Journal** | Fall 2009 | Vol. 11 | No. 2 63

Perspectives on Knowledge Converted into Action, **Public Administration Review,** Volume 70 Issue 2, Pages 317 – 319, February, 2010.

“Public Safety Policy in New York State” (with Martin Horn) in Gerald Benjamin, **Public Policy in New York State**, Oxford University Press, 2012.

“The Joint Pursuit of Government Integrity and High Performance” with Frank Anechiarico, in **NYSBA Government, Law and Policy Journal** | Winter 2011 | Vol. 13 | No. 2

“Policy Analysis in the Budgetary Process: A Case Study of the New York Legislature” with Robert Purtell, paper presented and annual research conference of the Association of Public Policy and Management, Washington DC, November 5-7, 2011.

“Combining Effectiveness and Integrity in Public Management: Outcome oriented e-governance and transparency in the fight against corruption” (with Frank Anechiarico) paper presented and annual research conference of the Association of Public Policy and Management, Washington DC, November 5-7, 2011.

“Twinning Effectiveness and Integrity: Lessons from Management Reforms in the Fight Against Police Corruption in New York” (with Frank Anechiarico) *Journal of Police Studies(forthcoming)*

Curriculum Vitae

Robert M. Purtell, Ph. D.

Phone 518-442-5276

Email: rpurtell@albany.edu

Research and Teaching Interests:

Managerial Finance, Corporate Finance and Governance of Nonprofit and Governmental Organizations, Evidence-Based Management, Risk Management, Program Evaluation and Strategy

Published Papers:

Managing Crime Counts: An Assessment of the Quality Control of NYPD Crime Data, (2006) with Dennis Smith, NYU Wagner Graduate School, The Center for Law and Criminal Justice, New York University Law School (October, 2006)

Beyond Budgeting: Public-Service Financial Education for the 21st Century (2009) with James Fossett, Rockefeller College, University at Albany, Journal of Public Affairs Education, Vol. 16(1) p. 95-110

Hey You Never Know: Selling State Lotteries in America (2010) with James Fossett, Rockefeller College, University at Albany, Journal of Public Budgeting, Accounting and Financial Management, Vol. 22(3) p. 376-406

Published Cases:

Problems and Cases in Financial Management for Public, Health, and Not-for-Profit Organizations by Steven A. Finkler, Prentice Hall, Upper Saddle River, NJ, 3rd Edition, (2009)

Books:

Financial Management for Public, Health, and Not-for-Profit Organizations by Finkler, Calabreze, Purtell & Smith, Prentice Hall, Upper Saddle River, NJ, 4th Edition, under revision for publication in 2012

Papers Under Review:

Public Assets at Risk: Evaluating Nonprofit Hospital Conversions to For-Profit Ownership, Revise and Resubmit, Journal of Public Budgeting, Accounting and Financial Management

Lessons in Policing Crime Reporting, with Dennis Smith, NYU, Wagner Graduate School

To Lease or Not to Lease: Is That the Question?, with James Fossett and Niyousha Hosseinichimeh, Rockefeller College, University at Albany

Working Papers:

Estimating Cost: Theoretical and Practical Considerations in Regression-Based Cost Estimation

Price and Size: Are All Municipal Bond Trades Fairly Priced? with Louis Stewart, Howard University

Thinking at the Margins: The Use of Marginal-Contribution Analysis in Hospital Planning with Denise Tahara, New York Medical College

An Empirical Assessment of NYPD's "Operation Impact": A Targeted Zone Crime Reduction Strategy, with Dennis Smith, NYU Wagner Graduate School.

Crime Reduction and Economic Development in New York City: The Redistribution Effects of Public Safety with Dennis Smith, NYU Wagner Graduate School

Does Stop and Frisk Stop Crime? with Dennis Smith, NYU Wagner Graduate School

Evaluation Studies:

Managing Crime Counts: An Assessment of the Quality Control of NYPD Data, conducted for the New York City Police Department, with Dennis Smith, October 2006

An Empirical Assessment of NYPD's "Operation Impact": A Targeted Zone Crime Reduction Strategy, conducted for the New York City Police Department, with Dennis Smith, June, 2007

Assessing the Impact of Stop-and-Frisk Activities on Major Crimes in New York City, with Dennis Smith, NYU Wagner Graduate School, October 2008

Dissertation:

Hospital Conversions: Have They Been Done at Fair-Market Value? Committee Chair: Steven Finkler, Committee Members: Roy Sparrow, David Yermack.

Conference Presentations and Seminars:

To Lease or Not to Lease: Is That the Question?, APPAM 2010 Research Conference, November 2010

Managing the Integrity of Performance Data: An Empirical Approach to Fighting Fudge in Performance Reporting, APPAM 2010 Research Conference, November 2010

To Lease or Not to Lease: Is that the (Infrastructure) Question?, 2010 Budgeting & Financial Management Meeting, New York University, August 2010

Implementing Management Reform in State and Local Government, APPAM 2009 Research Conference, November 2009

Causes of the Financial Crisis, 2009 NASPAA Annual Conference, October 2009
Hey You Never Know: Selling State Lotteries in America, ASPA Research Conference 2008
Assessing the Impact of Stop-and-Frisk Activities on Major Crimes in New York City, APPAM 2008 Research Conference, November 2008

An Empirical Assessment of NYPD's "Operation Impact": A Targeted Zone Crime Reduction Strategy, APPAM 2007 Research Conference, November 2007

Crime Reduction and Economic Development in New York City: The Redistribution Effects of Public Safety, APPAM 2006 Research Conference presented with Dennis Smith, November 2006

Beyond Budgeting: Public-Service Financial Education for the 21st Century, 2006 NASPAA Annual Conference: The Future of the Public Sector, October 2006

Do the Right Thing: Conflicts of Interest., Healthcare Financial Management Association Conference, "Corporate Accountability: Lessons for the Healthcare Industry", Panelist, March 2003, New York.

Comparative Governance Mechanisms in Not-For-Profit and For-Profit Organizations. Seminar for the students and faculty of the Wagner Graduate School at New York University, December 2002, New York University

Introducing Financial Management to Not-for-Profit, Healthcare and Public Sector Students. Seminar for Faculty from the Ukrainian Academy of Public Administration at the Robert F. Wagner Graduate School of Public Service, June 1999, New York University

Healthcare Mergers and Acquisitions: Valuations, Value and Price in a Changing Payer Environment. Seminar at the Leonard N. Stern School of Business Healthcare Executive-MBA Symposium, May 1998, New York University

Issues in Healthcare Finance and Governance. Panelist at the Leonard N. Stern School of Business Healthcare Executive-MBA Symposium, May 1998, New York University

Fundamental Considerations in Mergers and Acquisitions Involving Nonprofit Organizations at the Strategic Alliance Fund's Conference on "Making Strategic Alliances: A Look at Nonprofit Collaborations", December 1996, New York

Press Citations:

Gardiner, Sean (2010), City Violent Crimes Spike Even as Overall Rate Falls, The Wall Street Journal, **http://online.wsj.com/article/SB10001424052748703567304575628580735588388.html**, November 22

McDonald, H. 2010, "Compstat and Its Enemies: A Study Questioning the NYPD's Statistics is Irredeemably Flawed." The City Journal, http://www.city-journal.org/2010/eon0217hm.html, November 11

Rivera, R. and Baker, A., 2010, "Data Elusive on Low-Level Crime in New York City", The New York Times, http://www.nytimes.com/2010/11/02/nyregion/02secrecy.html?r=1&emc=eta1, November 1

Hope, B., 2010, "City Police Crime Statistics Trustworthy, Researchers Say" The New York Sun, http://www.nysun.com/new-york/city-police-crime-statistics-trustworthy/42235/, October 25.

Milchman, K, 2010, "It's the Size That Counts", New York Press, http://www.nypress.com/blog-1155-it%5Cs-the-size-that-counts.html, June 28

Staff, 2010, "Study Lauds Police Effort; Commissioner Criticizes Low Starting Pay" The New York Times, http://query.nytimes.com/gst/fullpage.html?res=9403E0DB163EF93BA15755C0A9619C8B63&sec=&spon=&pagewanted=all, May 17

Manioscalo, J. 2010, "Rosy Crime Stats Spark Questions", Brooklyn Heights Courier, http://www.yournabe.com/articles/2010/02/09/brooklyn_heights_courier/news/courier-yn_brooklyn_heights-compstatchallenge.txt, February 9.

Kelly, R.W., 2010, "Why Crime Stata Can Be Trusted: Commissioner Raymond Kelly Defends the Numbers", The New York Daily News, http://www.nydailynews.com/opinions/2010/02/14/2010-02-14_why_crime_stats_can_be_trusted.html, February 14

Walker, K. 2008, "Lottery Lease Proposals Prompt Warnings", Baptist Press http://www.baptiststandard.com/index.php?option=com_content&task=view&id=8726&Itemid=53, October 30

Walker K., 2008, "Warnings raised Over Lottery Leases", Baptist Press, http://www.bpnews.net/bpnews.asp?ID=29184, October 23

Walstin, B., 2008, "Is Vermont's Lottery-leasing Scheme Too Risky? A Report from SUNY-Albany Runs the Numbers" Seven Days: Local Matters, **http://www.7dvt.com/2008/vermont-s-lottery-leasing-scheme-too-risky-report-suny-albany-runs-numbers**, February 27

Blau, R., 2007, "Find NYPD 'Impact' Has One on Crime; Attacks Trouble Spots" The Chief, July 6, 2007 http://www.thechief-leader.com/news/2007-07-06/news/003.html, July 6

Lueck, T., 2007, "Study Lauds Police Effort; Commissioner Criticizes Low Starting Pay", The New York Times, June 28, http://sba.beforeid.com/news/nyt062807.html

Gendar, A., 2007, "Recruiting Woes May Cut Impact" Daily News, http://www.nydailynews.com/ news/2007/06/28/2007-06-28_recruiting_woes_may _cut_impact.html, June 28

2007, "NY Deployment of Rookie Cops Lowered Crime, Study Says", The Crime Reporter, http://www.thecrimereport.org/archive/ny-deployment-of-rookie-cops-lowered-crime-study-says/, June 28

2007, "Policing Strategy Called Success", The New York Sun, http://www.nysun.com/new-york/policing-strategy-called-success/57489/

Hope, B. 2006, "City Police Crime Statistics Trustworthy, Researchers Say" The Sun **http://www.nysun.com/new-york/city-police-crime-statistics-trustworthy/42235/**, October 25

Blau, R., 2006, "Skeptics Question Findings: Call NYPD Crime Stats True-Blue" The Chief, http://www.thechief-leader.com/news/2006-11-10/News_of_the_week/Skeptics_Question_Findings_Call_NYPD_Crime_Stats_T.html, November 10

Fraser, J.A., 1995, "Forecasting: Staying on Target" Inc. Magazine, http://www.inc.com/magazine/19951101/2492.html, November 1

Fraser, J.A., 1995, "Checklist – A Must for Business Survival", Inc. Magazine, http://www.inc.com/magazine/19950501/2286.html, May 1

1984, "Commodities for the Masses" Inc. Magazine , http://www.inc.com/magazine/19840201/9784.html, February 1

Curriculum and Course Development:

Healthcare Financial Management: I co-designed the course and prepared the standard teaching notes and class overheads for the course. I also wrote integrated problem sets to correspond to the three course modules. The first teaches acuity-based budgeting and variance-analysis. The second shows how costs and payer-mix impact pricing-decisions. The third focuses on healthcare corporate-financing decisions.

Financial Management for Public, Nonprofit and Health Organizations: (Core Course) I prepared the first version teaching notes that accompany *Financial Management for Public, Health, and Not-for-Profit Organizations* by Steven A. Finkler

Cost Management for Not-For-Profit and Governmental Organizations. I designed the course, wrote a complete set of lecture notes and three cases. The first, illustrates the use of multiple regression and interrupted time series to isolate fixed and variable cost components. The second uses break-even analysis to address a compensation decision. The third is a comprehensive case illustrating the use of activity-based costing in a contract negotiation.

Infrastructure Planning, Management and Finance: I designed this three-part course for urban-planning and public-management students. The first module introduces the topic broadly and places it in historical and policy perspective. The second part of the course deals with political, managerial, accounting, analytical and forecasting problems endemic to decisions involving public expenditures. The final segment combines public and corporate finance to investigate alternative methods that governments can use to finance and pay for infrastructure investments.

Curriculum Development: Designed the finance core-course and a four-course sequence for a healthcare-finance concentration for Kean University's MPA program.

Education:

New York University**, Wagner Graduate School of Public Service*, ***M. Phil., Ph.D.

***New York University**, Stern Graduate School of Business* - ***MBA,** Finance/Economics*
Elected to Omicron Delta Epsilon: International Economics Honor Society

***Manhattan College**, BS, Mathematics*

Affiliations:

Association for Public Policy Analysis and Management, American Society for Public Administration and Association for Public Budgeting and Finance

Teaching Experience:

The University at Albany, Nelson A. Rockefeller College of Public Affairs & Policy

Assistant Professor of Finance, 2005 to present.

Director of the MPA Program, 2009 to present. Courses taught:

- Economic and Financial Theory (Doctoral Seminar - 2005)
- Public Budgeting (2006-07)
- Financial Management for Government and Nonprofit Organizations (2006-11)
- Cost Management for Government and Nonprofit Organizations (2006-11)
- Capital Markets, Risk & Governments (2008-11)
- Healthcare Financial Management (2011 – co-instructor)
- Nonprofit Financial Management (2011)

NYU, Wagner Graduate School of Public Service,

Adjunct Assistant Professor of Public Administration, 2001 to 2005

Visiting Lecturer on Public Administration, 1999 to 2000

Adjunct Lecturer on Public Administration, 1995 to 1999, and 2000-01

Member of the Finance and Health Faculties. Courses taught:

- Financial Management for Public, Nonprofit and Health Organizations (1995-05)
- Healthcare Financial Management (1996-05, Co-Designed the Course)
- Healthcare Cost Accounting (1997-05)
- Cost Management for Nonprofit and Governmental Organizations (2002-04)

Kean University, College of Business & Public Administration

Assistant Professor, Health Services Administration, 2000-01. Courses taught:

- Financial Condition Analysis (Undergraduate, Designed the Course)

- Financial Management for Health and Social-Welfare Organizations (Graduate)
- Advanced Topics in Healthcare Finance (Graduate, Designed the Course)

NYU, School of Continuing and Professional Studies

Adjunct Instructor on Finance, Law and Taxation, 1993 to 1995. Courses taught:

- Introduction to Investments and the Markets (1993-4)
- Mergers and Acquisitions (1994-5, Designed the Course)
- Corporate Finance (1995, Designed the Course)

Teaching Evaluations :

My overall teaching evaluations have consistently ranged from 4.6 to 5.0 with a weighted-average of approximately 4.75 on a scale where one is the lowest rating and five the highest.

Service and Awards:

- 1999 – NYU/Wagner Doctoral-Student Teaching Award
- 2000 – NYU/Wagner ACE Award for Health-Management-Program service.
- 2001 – Dissertation nominated for the NASPAA and APPAM awards
- 2007-11 – Member Empire State Capital Area ASPA Awards Committee
- 2006-10 – Member of Economics and Management Search Committees
- 2009 to present – Member of the Rockefeller College Admissions Committee
- 2009 to present – Member of the Committee on Academic Standing and Retention
- 2009 to present – Member of Public Administration Department Executive Committee
- 2009 to present – reviewer for the Journal of Public Budgeting, Accounting & Financial Management
- 2010 to present – Member of NASPAA Program Guidelines Committee for Managerial Finance
- 2010 – Member of Rockefeller College Presidential Management Fellows Candidate-Selection Committee
- 2011 – Member UAlbany Ad-Hoc-Revenue Committee

Professional Experience:

Boardroom Capital Partners, New York, NY, 1985 to 2005, *Principal*: Advised clients in the areas of corporate-finance and the design of structured-finance instruments.

Merrill Lynch, New York, NY, 1983-5, *Director - Merrill Lynch Futures Management:* Designed and marketed managed futures funds.

Coopers & Lybrand, New York, NY, 1982-3, *Director - Financial-Services Management-Consulting Practice:* Managed corporate-finance, planning, and financial-product-design engagements.

A.G. Becker, New York, NY, 1978-82, *Director & CFO, Correspondent Services*, a $70 million business that provided clearing and corporate-finance services to broker/dealers in the United States, Canada and Asia.

American Express, New York, NY, 1972-8, *Assistant to the President 1975-8 and Internal-Consulting Manager 1972-5:* Managed corporate and managerial finance, merger and acquisition, fixed-income and venture-capital, portfolio-management, planning and reporting, and information-systems projects.

Human Resources Administration, City of New York, 1973-4, *Project Manager:* On temporary assignment from American Express, I directed financial-control and systems projects.

Eastern Airlines, New York, NY, 1969-72, *Senior Financial Analyst:* Developed integrated capital-budgeting and simulation models used to analyze over $2 billion of capital projects.

Sikorsky Aircraft, Stratford, CT, 1968-9, *Mathematical Analyst:* Developed optimization algorithms and mathematical models of dynamic engineering and manufacturing systems.

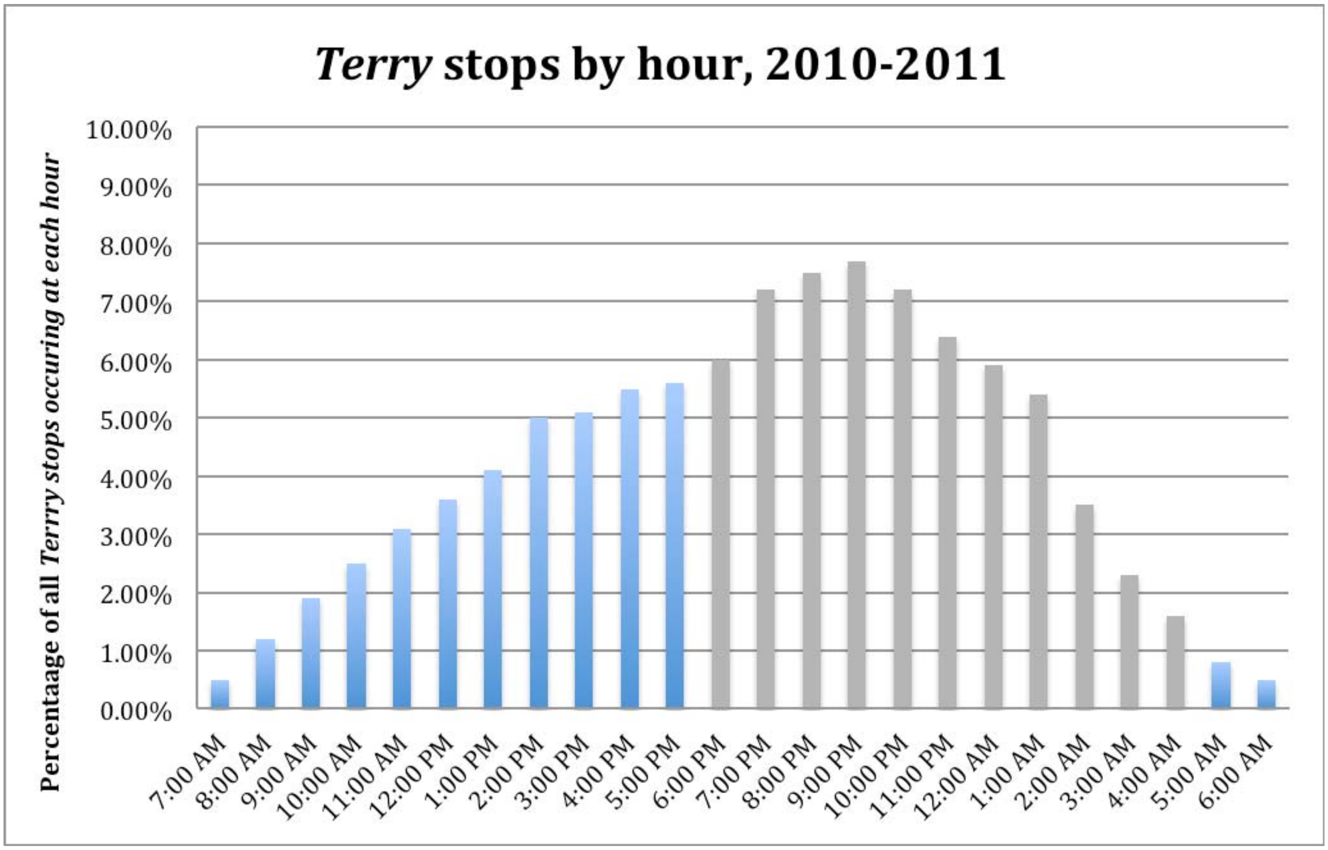
Representative Structured-Finance Designs:

Guaranteed-Principal Venture-Capital Fund. This structure guarantees investors the full return of their principal without any loss of upside potential. It has been used by the US Small Business Administration to fund narrowly-focused venture-capital funds.

Dynamically-Optimized Equipment Trust. This securitization structure gives borrowers optimal longitudinal advance-rates while lowering their all-in cost of funds without materially increasing the risk of default.

Appendix B

Appendix B: *Terry* Stops by Hour



Appendix C

Documents and Materials Considered

I have reviewed the following documents and materials in the course of forming my opinion. I continue to review materials and documents in this case and reserve the right to revise my opinion as my work continues.

1. Data produced by NYPD
    a. Merged NYPD arrest and crime complaint data 2010
    b. Merged NYPD arrest and crime complaint data 2011
    c. Arrest Report and Complaint Report Data for 2010
    d. Arrest Report and Complaint Report Data for the 1st and 2nd Quarters of 2011
    e. Arrest Report and Complaint Report Data for the 3rd Quarter of 2011
    f. Arrest Report and Complaint Report Data for the 4th Quarter of 2011
    g. Arrest Report and Complaint Report Data for 1st 6 Months of 2012
    h. UF 250 Database with fictionalized Tax IDs 1st Quarter 2010
    i. UF 250 Database with fictionalized Tax IDs 2nd Quarter 2010
    j. UF 250 Database with fictionalized Tax IDs 3rd Quarter 2010
    k. UF 250 Database with fictionalized Tax IDs 4th Quarter 2010
    l. UF 250 Database with fictionalized Tax IDs 1st Quarter 2011
    m. UF 250 Database with fictionalized Tax IDs 2nd Quarter 2011
    n. UF 250 Database with fictionalized Tax IDs 3rd Quarter 2011 (corrected)
    o. UF 250 Database with fictionalized Tax IDs 4th Quarter 2011
    p. UF 250 Database with fictionalized Tax IDs 3rd Quarter 2011
    q. UF 250 Database with fictionalized Tax IDs 1th Quarter 2012
    r. UF 250 Database with fictionalized Tax IDs 2rd Quarter 2012
2. Data Produced by Plaintiffs:
    a. 1 DVD (PL-EX 933) containing data produced by plaintiffs on December 17, 2012
    b. CD (PL-EX 934) containing corrected data produced by plaintiffs on January 10, 2013

3. Academic Publications:

    a. Alpert, Geoffrey P., John M. MacDonald, and Roger G. Dunham. "Police Suspicion And Discretionary Decision Making During Citizen Stops." *Criminology* 43 (2005): 2.
    b. Ayres, Ian. "Outcome tests of racial disparities in police practices." *Justice Research and Policy* 4, no. -1 (2002): 131-142
    c. Bratton, William, and Peter Knobler. *The turnaround: How America's top cop reversed the crime epidemic*. Random House, 2009.
    d. Fagan, Jeffrey. "Law, social science, and racial profiling." *Justice Research and Policy* 4, no. -1 (2002): 103-130.

e. Fridell, Lorie A. "By the numbers: A guide for analyzing race data from vehicle stops." Washington, DC: Police Executive Research Forum, 2004.
f. Gelman, Andrew, Jeffrey Fagan, and Alex Kiss. "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, no. 479 (2007): 813-823.
g. Grogger, Jeffrey, and Greg Ridgeway. "Testing for racial profiling in traffic stops from behind a veil of darkness." *Journal of the American Statistical Association*101, no. 475 (2006): 878-887.
h. Hirschi, Travis, and Michael Gottfredson. "Age and the explanation of crime."*American Journal of Sociology* (1983): 552-584.
i. Krug, Peter. "Prosecutorial discretion and its limits." *The American Journal of Comparative Law* (2002): 643-664.
j. Lamberth, John. "Revised statistical analysis of the incidence of police stops and arrests of black drivers/travelers on the New Jersey Turnpike between exits or interchanges 1 and 3 from the years 1988 through 1991." *State v. Pedro Soto* 734 (1994).
k. Maple, Jack, and Chris Mitchell. *The crime fighter: Putting the bad guys out of business*. Broadway, 2000.
l. Mosher, Clayton. "Racial Profiling/Biased Policing." *Sociology Compass* 5, no. 9 (2011): 763-774.
m. Ridgeway, Greg. *Analysis of racial disparities in the New York Police Department's stop, question, and frisk practices*. Rand Corporation, 2007.
n. Ridgeway, Greg, and John MacDonald. "Methods for assessing racially biased policing." *Race, ethnicity, and policing* (2010): 180-205.
o. Smith, Dennis C. “Police.” In *Setting Municipal Priorities,* edited by Charles Brecher and Raymond D Horton. New York: Russell Sage Foundation, 1981.
p. Smith, Stanley K., and Mohammed Shahidullah. "An evaluation of population projection errors for census tracts." *Journal of the American Statistical Association* 90, no. 429 (1995): 64-71.
q. Stockburger, David W. *Introductory statistics: Concepts, models, and applications*. David Stockburger, 1996.
r. Tillyer, Rob, Robin S. Engel, and John Wooldredge. "The intersection of racial profiling research and the law." *Journal of Criminal Justice* 36, no. 2 (2008): 138-153.
s. Walker, Samuel. "Searching for the denominator: Problems with police traffic stop data and an early warning system solution." *Justice Research and Policy*3, no. 1 (2001): 63-96.

t. Weisburd, David, Cody W. Telep, and Brian A. Lawton. "Could Innovations in Policing have Contributed to the New York City Crime Drop even in a Period of Declining Police Strength?: The Case of Stop, Question and Frisk as a Hot Spots Policing Strategy." *Justice Quarterly* ahead-of-print (2013): 1-25.
u. Zimring, Franklin E. *The City that Became Safe: New York's Lessons for Urban Crime and Its Control: New York's Lessons for Urban Crime and Its Control*. Oxford University Press, USA, 2011.
v. Zingraff, Matthew T., H. Marcinda Mason, William R. Smith, Donald Tomaskovic-Devey, Patricia Warren, Harvey L. McMurray, and C. Robert Fenlon. "Evaluating North Carolina state highway patrol data: citations, warnings and searches in 1998." *report submitted to North Carolina Department of Crime Control and Public Safety and North Carolina State Highway Patrol*(2000).

4. NYPD Reports

a. New York City Police Department, Office of Management Analysis and Planning *2011 Reasonable Suspicion Stops: Precinct Based Comparison by Stop and Suspect Description*, by Michael J. Farrell, Philip McGuire, and Thomas J. Taffe (New York City, New York City Police Department).