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

FLOYD, et al.,

Plaintiffs,

-against-

CITY OF NEW YORK, et al.,

Defendants.

--X

DECLARATION OF DENNIS C. SMITH

08 Civ. 1034 (SAS)

DENNIS C. SMITH declares, pursuant to 28 U.S.C. § 1746, under penalty of perjury, that the following is true and correct:

1. I am an Associate Professor of Public Administration at the Robert F. Wagner School of Public Service at New York University. I have been retained by the Defendants in this action as a testifying expert. I have previously submitted a report in this case, Report of Dennis C. Smith, Ph.D., dated November 15, 2010 ("11/15/10 Report"), excerpts of which are annexed hereto as Exhibit A, which also contains my curriculum vitae. I submit this declaration in support of Defendants' motion to preclude all expert reports and opinions of Jeffrey Fagan, Ph.D ("Fagan"). I have personal knowledge of the facts contained herein based on my review of documents and business records of the City of New York. I have reviewed all analyses submitted by Fagan in this case, including: the Report of Jeffrey Fagan, Ph.D., of October 15, 2010 ("Report"), the Supplemental Report of Jeffrey Fagan, Ph.D., of December 3, 2010 ("Supp. Report") (Fagan's Affidavit of September 28, 2011 ("JF Aff."), and his Declaration of

November 6, 2011 ("11/6/11 Decl.").[1] I have also reviewed the transcript of Fagan's deposition taken on February 9, 2011, as well as all articles and sources cited herein.

Legal Classification of Stops from the NYPD UF-250 Database

2. I examined Fagan's methodology and conclusions regarding his classification of all 2,805,721 stops recorded by the NYPD from 2004-2009 into three categories: legally "justified"; "unjustified"; or "indeterminate." Report (Dkt # 132) at 49. Upon receipt of Fagan's computer coding instructions on or about September 26, 2010, I attempted to replicate Fagan's findings using his computer coding instructions for the program he used to execute his classification scheme. I came upon two discrepancies between the characterization of Fagan's classification scheme in his Report (Dkt # 132) (at 50) and his coding instructions.

3. The computer coding instructions categorize UF-250s with a *single* "conditionally justified" circumstance plus an "additional circumstance" as justified; however, they fail to include those with multiple "conditionally justified" circumstances plus an "additional circumstance," which instead are categorized as indeterminate. This contradicts Fagan's description of his classification scheme in his Report, which states in part: "Stops are **justified** if the circumstances listed are conditionally justified . . . and an 'additional circumstance' is also indicated." (Report (Dkt # 132) at 50) (emphasis in original). This error deflates the count of stops that should be categorized as "justified" by 261,042.

[1] I am informed that the Reports are in the record before the Court as follows: Report and Supp. Report on plaintiffs' opposition to defendants' motion for summary judgment (filed under seal as Dkt #132), JF Aff. on plaintiffs' motion to amend/correct Order on defendants' motion for summary judgment (Dkt # 156) and JF Decl. on plaintiffs' motion for class certification (Dkt #168).

4. Fagan’s coding instructions categorize UF-250s with a single “conditionally justified” circumstance and no “additional circumstance” as unjustified. However, his Report states that “Stops are of indeterminate legality if the circumstance or circumstances listed are (all) conditionally justified, and no additional circumstances are indicated.” ((Dkt #132) at 50) (emphasis in original). This clear contradiction serves to erroneously inflate the number of stops that are categorized as “unjustified” in his analyses by 156,625. I also note that 23, 806 UF250s have no boxes checked off on the Side One of the form. Applying Fagan's methodology, under 1 percent of the stops should be classified as unjustified.

5. I also note that Fagan's analysis concludes that there are 2,805,721 stops, 697,203 Indeterminate, 179,877 Unjustified, while my analysis concludes that there are 2,811,771 stops, 698,721 Indeterminate and 180,432 Unjustified. The explanation for the difference, I assume, is that when using a large data set it is customary to "clean" the data received in a variety of ways, such as excluding cases that do not have a precinct indicated, or a precinct that for whatever reasons seems anomalous (e.g., Central Park). Typically, as in this case, if the detail of that "cleaning" are not specified by one researcher, the counts of a different researcher will not agree perfectly, and will be as we are here, slightly higher. These differences are not substantively-- or statistically--- significant. This minor difference does not impact the sum and substance of my conclusions.

6. Fagan’s review of the UF-250s in an attempt to determine the legality of the underlying stops suffers from other methodological flaws which severely undermine the reliability of his findings. He attempted to classify the stops based *only* on a review of the boxes checked on the UF-250 form. He did not review or include any handwritten notes on the UF-250

form; he also omitted information captured in other fields on the UF-250, such as details regarding the address and type of location. These omissions potentially deflate the number of Justified stops and inflate the number of Indeterminate and Unjustified stop by thousands.

7. I also observed ways in which Fagan's presentation of his findings is misleading and runs counter to generally accepted statistical practices. For example, Fagan's method of aggregating stop data for five separate years obscures the fact that, applying to Fagan's own classification scheme, the rate of stops he classifies as "justified" has increased from 2004-2009, while the proportions of "unjustified" and "indeterminate" stops have fallen. Notably, the relatively small portion of stops, 6.7 percent, which Fagan classified as "unjustified" represents an average over the five years studied. The amount of "unjustified" stops according to Fagan's scheme decreased from 9.73 percent in 2004 to one-half that amount, 4.32 percent, in 2009. This failure to distinguish temporal trends in the data limits the number and type of inferences which may be validly drawn from Fagan's findings.

8. Additionally, by combining the "unjustified" stops (6.7 percent) with the stops of "indeterminate legality" or unknown stops (24.4 percent) in his discussion of findings and concluding that "nearly 30 percent of all stops appear to be either facially unconstitutional, or lacking sufficient information to make a complete determination," (Report (Dkt #132) at 55), Fagan runs afoul of customary statistical procedures. Generally accepted statistical procedures dictate that when data is unknown or undetermined, one either sets aside the data as missing data or apportions the data to the classifications already made, in proportion to those classifications, or provides an explicit rationale for allocating the "missing" data (the cases he is unable to classify). There is no basis for associating all the missing cases with the "unjustifieds."

9. Fagan's conclusions about his findings fail to consider plausible alternative hypotheses. For example, I understand that Fagan makes the following conclusion as reflected in his Report (Dkt # 132) at 55 and Fagan Decl. of 11/6/11: "I also concluded that the fact that the legal sufficiency of 31% of all stops cannot be shown suggests that the current regime for regulating the constitutional sufficiency of the huge volume of stops is ineffective and insensitive to the actual conduct of stops." An alternative explanation is that the inability of Fagan's scheme to definitively determine the constitutionality of 24.4 % of stops more likely reflects the limitations of his classification scheme (as described above) rather than, as he suggests, documentation deficiencies or problems with the constitutional validity of the underlying stops.

10. I understand that Fagan finds that stops based at least in part on "high crime area" have become more common over time. Supp. Report (Dkt # 132) at 45. However, Fagan does not consider a likely, alternative hypothesis to explain this trend, which is that it may reflect police response to a spike in crime, conditions, or trends.

Multiple Regression Analyses

11. I also examined Fagan's methodology and conclusions regarding his various multiple regression analyses. I noted an initial discrepancy in that Fagan's classification of stops described above concluded that the vast majority of stops (70%) were "justified," i.e., based on reasonable suspicion (according to his interpretation of the caselaw as applied to the circumstances described on the UF-250 forms); yet his multiple regression analysis examined the full set of all 2,805,721 stops recorded in the UF-250 database, in spite of his separate finding that the vast majority of those same stops were "justified." An alternative way to examine the pattern of stops would have been to study only those that he classified as "indeterminate" and

"unjustified" to see if he could connect them to the race of the person stopped. Fagan's failure to do so calls into question his conclusion that "the NYPD has engaged in patterns of unconstitutional stops of City residents that are more likely to affect Black and Latino citizens." (Report (Dkt # 132) at 3).

12. I noted many other flaws in Fagan's methodology, the most significant being his use of local crime rate as a benchmark by which to measure possible evidence of bias in NYPD stop-and-frisk activity. The most logical and reliable method to assess the question of whether police are stopping individuals based on race or on constitutionally permissible grounds of suspicion of criminal activity is to use a benchmark of rates of criminal participation by race. Data on suspect race captured in crime complaints and arrest reports ("suspect description") is a reliable indicator of this measure; it is also the best available proxy for the proportions of individuals in society who are engaged in behaviors which give rise to reasonable articulable suspicion ("RAS"). Fagan drastically understates the percentage of crimes for which information on racial characteristics of offenders is known to police. *See, e.g.*, JF Dep. (HG Decl. Ex. B) at 204:3-7. Based on an aggregate of un-arrested and arrested suspect data for 2009 and 2010, suspect race is known for 85 percent of violent crimes. It is also known for 62 percent of all crimes, even though property crimes have few victim identified suspects and make up a majority of all crime. There is also no reason to believe that the proportion of unknown suspects are any different than those that are known. The charts annexed hereto as Exhibit B show the percentage of crimes for which perpetrator or suspect race is known to the police. While similar data aggregating suspect race from crime complaints and arrest reports is not yet available for the years 2004-2008, because the percentages of crimes for which suspect race is known is

consistent from 2009-2010, it is reasonable to infer that this information is known to the police for similar proportions of crimes for the years 2004-2008.

13. The charts annexed hereto as Exhibit C also show that, for the years 2009-2010, Blacks and Hispanics comprise a majority of violent crime suspects in all precincts except one in the City, and in most precincts are the overwhelming majority of suspects. For example, for 2009, the average of suspects of violent crime in the City who are White is 5.3%, and the average percent Black suspects is 65.9, i.e., higher by a factor of twelve. There are only four precincts in the City where more than one third of the suspects are White, including two in Staten Island (123 and 122, where 63.8% and 37.9% of suspects are White). In more than two-thirds (54) of the City's precincts, the percent of suspects who are White is *less* than 10%, and in more than two-thirds of the precincts (56 precincts) the percent of suspects who are Black exceeds 50%. There are nineteen precincts in which the percent of Black suspects is 75% or higher, and ten precincts where the Black suspects exceed 90% of all perpetrators of violent crime.

14. Data on the overall crime rate of a precinct, on which Fagan relies, is a poor baseline against which to evaluate police stop rates of individuals by race, as it captures only gross crime occurring within an area without taking into account *who* is participating in the crime. A baseline measure of overall crime would serve as an appropriate benchmark only if the question being assessed was whether total stops were proportionate to crime committed. By itself, the mere fact that a certain amount of crime exists in a precinct is irrelevant to the racial profiling inquiry; without knowing who the perpetrators are, this measure provides no information by which we can assess whether the motivation for police stops was properly based on RAS or race alone. Fagan's benchmark fails to capture the information necessary to support a

valid causal inference of racial discrimination because it fails to take into account the "offending population" at risk for police intervention.

15. By omitting data on rates of criminal participation by race, Fagan's analysis does not consider the rival hypothesis that police are acting in a way that reflects the actual patterns of criminal behavior and in terms of what is known about the racial characteristics of offenders. The tables, scatterplots and charts, annexed hereto respectively as Exhibits C, D and E, show that the proportion of criminal suspects and arrestees by race is very highly correlated with stop rates by race (r=.84) and much more so than with crime rates (r=.66) or population used by Fagan (r=. 61). Notably, the data show that even in racially heterogeneous and predominately white precincts, or suspects are disproportionately minorities – e..g. in the 6th precinct, with 8% Black and Hispanic in the population, they comprise 81.8% of suspects for violent crime, the 17th precinct has 7.8% Black and Hispanics in resident population are suspects in 75.2% of violent crime perpetrators, and in Brooklyn's 61st precinct with 11.5 resident Black and Hispanic population they are 64.9% of violent crime suspects. Fagan's statement that his multiple benchmarking strategy using both precinct-level population and crime rates adequately accounts for race in his model is inaccurate. Supp. Report (Dkt # 132) at 10. Combining population and a crime as a benchmark assumes that local residents commit crimes in equal proportion; it also ignores the fact that residents may travel to other precincts to commits crimes. Contrary to Fagan's assumptions, crime is not evenly distributed throughout or within precincts in the city.

16. The indisputably strong correlation between race and participation in crime indicates that Fagan's models are significantly biased by the omission of this variable. By omitting a variable which is so strongly correlated with race, Fagan's models attribute a

disproportionate share of explanatory power to the race variable with respect to predicting the likelihood that an individual will be stopped by the police, giving rise to a false inference of discrimination.

17. As an illustration of the omitted variable bias manifest in Fagan's model, I note that NYPD stops are not proportionally correlated with the gender of local populations. 93% of all stops in 2009 were of males while only 7% were of females, who constitute 52.5% of the population. 11/15/10 Report (annexed hereto as Exhibit A) 43-44. If an analyst were to conduct a regression analysis using Fagan's model design but including gender (rather than race) as an independent ("explanatory") variable, stop rates of men would appear disproportionately large. Without taking into account data on the radically different contributions by men and women to commission of crime, an analyst would be left to conclude erroneously that police are targeting people for stops because they are male.

18. Initially Fagan omitted unemployment data as an independent variable, as more fully explained in my 11/15/10 Report. *See* Exhibit A, at 47, 57-58. While Fagan attempted to address this omission in his supplemental report, his use of unemployment data there fails to aggregate data at a level that corresponds to the units of police action, e.g. impact zones.

19. I also note that the percentage of all crimes for which suspect race is known to the police based on aggregates of crime complaints and arrest data represents a significantly higher rate of availability than data Fagan himself has used for benchmarks in previous studies, including the 2007 JASA study in which Fagan used race-specific arrest rate as a benchmark (*see* Gelman, Andrew, Jeffrey Fagan, and Alex Kiss, "An Analysis of the NYPD's Stop-and-Frisk Policy in the Context of Claims of Racial Bias," 102 *Journal of the American*

Statistical Association 813 (2007)) (hereinafter, "Gelman, Fagan & Kiss"). *See* Declaration of Heidi Grossman, December 19, 2011 ("HG Decl."), at Exhibit J.

20. Fagan's observation that "fewer than one in four stops in 2009 were based on a… suspect description known to the police," (Report (Dkt # 132) at 17) is of no consequence. Even if the majority of police stops are not based on a search for a particular individual, the valid benchmark by which to assess the possibility of police bias is the share of the population by race engaged in "targeted behaviors", for which the best proxy based on the best available data is criminal participation by race as measured by suspect descriptions. Since 1994 NYPD has increased its focus on crime prevention. Focusing on "known suspects" implies that police are primarily interested in intervening after a crime has been committed, the reactive policing model of the past.

21. Fagan's multiple regression models suffer from additional methodological deficiencies which render his conclusions unreliable. His selection of independent variables creates a multicolinearity problem, meaning that two or more of his independent, aka "explanatory," variables (e.g., crime rate, poverty, social and physical disorder) are highly correlated with race – his explanatory variable of interest. Multicolinearity produces skewed and unreliable results, it impairs the analyst's ability to distinguish among the competing possible explanations for the observed outcome (in Fagan's model, likelihood of being stopped by the police).

22. As explained more fully in my 11/15/10 Report (Exhibit A, at 53-54, 58-59), Fagan's use of precinct-level data in his analysis is a flaw which renders the following of his control variables inaccurate and unreliable: crime rates, population characteristics, patrol strength, and police response to crime conditions. Fagan's assumption that "the regulation and

oversight of stop and frisk policy and activities takes place at the precinct level" (Report (Dkt # 132) at 30) is incorrect, since, dating back prior to 2004, NYPD policing strategy, crime data analysis and deployment decisions, has focused on crime "hot spots," which may in a few cases comprise entire precincts but for the most part consist of very small, local areas *within* precincts. *See* Smith and Purtell, "An Empirical Assessment of Operation Impact" (annexed hereto as Exhibit F); Al Baker, " City Is Doubling Police Program to Reduce Crime," NEW YORK TIMES, Dec. 27, 2007 (annexed hereto as Exhibit G). Analyzing data at the precinct level is also misleading because crime, and accordingly police deployment and police stops, are not evenly distributed throughout precincts; aggregating data at too general of a level masks such variations. As such, the only proper approach for the questions addressed in Fagan's reports is to analyze the data according to "hotspots" or NYPD-designated "impact zones," which are operationally and functionally distinct from precincts.

23. Fagan likewise employed inappropriate time lags in his attempt to control for police response to crime. *See* my 11/15/10 Report (Exhibit A), at 18-19, 62-63. His first Report used a lag of one calendar quarter between crime complaints and police stops recorded (Report (Dkt # 132) at 10), while the Supp. Report shortened this lag to one month (Supp. Report (Dkt # 132) at 2). However, neither model accounts for the fact that NYPD allocation decisions are made in response to crime information on an immediate basis. Therefore, both lags utilized by Fagan present an inaccurate and misleading picture of the effect of crime conditions on stop rates.

24. Fagan's attempt to control for crime as a predictor of stop rates in his model is further flawed due to his improper aggregation of crime statistics across crime categories, a practice which the FBI has deemed to yield misleading and unreliable data. *See*

article annexed hereto as Exhibit H. Fagan's practice of logging crime complaint data also distorts his measure of crime. Logging the data "smoothes it out" – that is, it fails to capture distinctions over time such as such as spikes and valleys. Logging is an inappropriate technique in an inquiry of this nature, which purports to account for police response to crime trends. Fagan states that logging is necessary because: "All models control for the one-calendar-quarter lag of logged crime complaints. The log transformation of the actual number of crime complaints is used. Log transformation is necessary to adjust when distributions are skewed and non-linear." (Report (Dkt # 132) at 31, n.52). However, Fagan chose not to log the stop-and-frisk numbers, which show a very similar distribution to that of the crime complaints; while using logs in data where instability in trends is a focus of interest is in my view in appropriate, applying Fagan's logic above, he should have logged the stops as well as the crime complaints. This is another reason that Fagan's regression models do not accurately control for changes in police stop rates in response to crime, and thus cannot reliably estimate the influence of crime trends on stop patterns.

25. Fagan's attempt to control for differential police deployment ("patrol strength") as a possible variable to explain police stop patterns is undermined by his use of unreliable data. Fagan explained that this was a key consideration because "Police deployment patterns frequently involve the saturation of police patrols in crime-prone areas, which often leads to more encounters with minority citizens as compared to Whites. This differential exposure of citizens to police may result in differential enforcement patterns across racial/ethnic groups, especially under conditions where there are differences in the racial makeup and concentrations of neighborhoods or police precincts." (Report (Dkt #132) at 9-10). However, Fagan testified that he did not have data on the numbers of officers in each precinct who were

actually out on patrol making stops. (JF Dep. at 78:12-81:21, 79:14-20 (*see* HG Decl. at Ex. B)). Instead, he used raw data provided by the NYPD which just showed the overall number of officers assigned to each command, without indicating the breakdown of how many officers were engaged in enforcement duties and potentially responsible for citizen stops, and how many were involved in administrative or other types of duties. (Report (Dkt # 132) at 9-10, Appendix E). Therefore Fagan's estimate of "patrol strength" does not reliably or accurately control for the differential exposure of citizens to police he described above. Furthermore, Fagan's regression models in his Supp. Report do not make any attempt to account for this variable.

26. Because Fagan's control variables are fundamentally flawed, including the fact that key explanatory variables were omitted from this analysis, his model does not accurately account for the influences of race-neutral factors on police stops. Thus it misrepresents and artificially inflates the potential influence of individual race in predicting the likelihood of stops.

Fagan's Declaration of November 6, 2011

27. With regard to Fagan's 11/6/11 Decl. (Dkt # 168), I offer the following observations. His analysis of stops in deciles based on crime/stop ratios appears to establish that Blacks and Hispanic are disproportionately stopped in precincts that differ in levels of crime. By ignoring the fact that crime suspects are disproportionately Black and Hispanic across almost all precincts, his finding lead to a misleading conclusion about police practice. In decile #1 which recorded 4% of the reported stops, and one of the lowest percent Black in the resident population, in all but one precinct Blacks constituted a majority of the suspects (and in the one that is not a majority Black suspects: 74.4% are Hispanic). In five of the eight precincts included

in this decile, the percents of Black suspects exceeds 60%. The highest White percent of suspects is 15.8 %, and two are less than 3% White suspects.

28. Among the seven precincts in decile #10, with 24% of the total stops, the percent Black among the violent crime suspects is higher than 65% in all included precincts; four are higher than 75%. In all seven precincts, the highest percent of suspects who are White is 7.5 and three are less than 1.0%.

29. Considered in the light of this data showing disparities in crime patterns by race of suspect, the observed racial disparities in NYPD stop patterns are, in my opinion, consistent with the expected practice of a police department committed to preventing crime through alertness to suspicious behavior.

Alternative Regression Analysis

30. In order to demonstrate the effect of the omitted key variable of suspect race on the outcome of Fagan's multiple regression analysis, I conducted an alternative analysis using Fagan's regression model but adding data on suspect race aggregated from crime complaints and arrest reports (annexed hereto as Exhibit I); I performed this analysis for 2009-2010, the years for which the aggregated data was available. I used the number of crimes by suspect race as an independent variable instead of the logged total crime complaints which Fagan used. My analysis demonstrates that when corresponding suspect race is added to the crime complaints data, the percent Black and the percent Hispanic coefficients are no longer significant. Additionally, the percent Black coefficient changes sign. If this co-efficient were statistically significant, which it is not, it would indicate that stops are inversely related to percent Black in the population, which is the opposite of Fagan's finding.

31. These two results are clear evidence of an omitted variable bias. In respect to the new variables introduced, the Black suspects and Hispanic suspects are significant at the .001 level with positive coefficients. This shows that the total stops in a precinct in a month, are explained by the number of total Black and Hispanic suspects rather than by the percentage Black or Hispanic population, demonstrating that Fagan's model missed variables that are central to the analysis, and contradict his central claim that race per se explains stops. This results table demonstrates how the regression results change dramatically, and evidence of racial discrimination disappears, when the proper variables are taken into account.

I declare under penalty of perjury that the foregoing is true and correct. Executed in Westerlo, New York, on December 19, 2010

Dated: Westerlo, New York
December 19, 2011

Dennis C. Smith