

U.S. COMMISSION ON CIVIL RIGHTS



**DISPARITY STUDIES AS
EVIDENCE OF
DISCRIMINATION IN
FEDERAL CONTRACTING**

**Briefing Report
May 2006**

U.S. COMMISSION ON CIVIL RIGHTS

The U.S. Commission on Civil Rights is an independent, bipartisan agency established by Congress in 1957. It is directed to:

- Investigate complaints alleging that citizens are being deprived of their right to vote by reason of their race, color, religion, sex, age, disability, or national origin, or by reason of fraudulent practices.
- Study and collect information relating to discrimination or a denial of equal protection of the laws under the Constitution because of race, color, religion, sex, age, disability, or national origin, or in the administration of justice.
- Appraise federal laws and policies with respect to discrimination or denial of equal protection of the laws because of race, color, religion, sex, age, disability, or national origin, or in the administration of justice.
- Serve as a national clearinghouse for information in respect to discrimination or denial of equal protection of the laws because of race, color, religion, sex, age, disability, or national origin.
- Submit reports, findings, and recommendations to the President and Congress.
- Issue public service announcements to discourage discrimination or denial of equal protection of the laws.

Members of the Commission

Gerald A. Reynolds, *Chairman*
Abigail Thernstrom, *Vice Chairman*
Jennifer C. Braceras
Peter N. Kirsanow
Arlan D. Melendez
Ashley L. Taylor
Michael Yaki

Kenneth L. Marcus, *Staff Director*

U.S. Commission on Civil Rights
624 Ninth Street, NW
Washington, DC 20425

(202) 376-8128 voice
(202) 376-8116 TTY

www.usccr.gov

This report is available on disk in ASCII Text and Microsoft Word 2003 for persons with visual impairments. Please call (202) 376-8110.

Disparity Studies as Evidence of Discrimination in Federal Contracting

A Briefing Before
The United States Commission on Civil Rights
Held in Washington, D.C., December 16, 2005

The final printed version of this report may contain additional Commissioner statements.

Briefing Report
May 2006

TABLE OF CONTENTS

OVERVIEW	1
Summary of Ian Ayres' Statement	3
Summary of Constance Citro's Statement	5
Summary of Roger Clegg's Statement	7
Summary of George LaNoue's Statement	8
Discussion	10
Is Race-Conscious Affirmative Action an Appropriate Remedy for Discrimination?	10
Government Remedies to Private Sector Discrimination	14
The Use of Race-Neutral Alternatives	14
How to Measure Discrimination	15
The Methodology of Disparity Studies	17
Individual Disparity Studies	17
National Disparity Studies	20
PREPARED REMARKS OF GEORGE R. LA NOUE, PROFESSOR OF POLITICAL SCIENCE AND DIRECTOR OF THE PROJECT ON CIVIL RIGHTS AND PUBLIC CONTRACTS, UNIVERSITY OF MARYLAND BALTIMORE COUNTY	23
Executive Summary	23
Introduction	24
State and Local Disparity Studies	26
Statistics	26
Anecdotes	30
Unrepresentative samples	30
Anonymous unverified responses	31
Affirming pre-existing assumptions	32
Federal studies	34
Urban Institute Meta-study	35
Appendix A	41
Benchmark study	43
Disparity Study Conclusions	46
Race Neutral Alternatives	49
WRITTEN STATEMENT OF CONSTANCE F. CITRO, DIRECTOR, COMMITTEE ON NATIONAL STATISTICS	53
Definitions	53
Methodological Issues for Disparity Studies	54
Extant Disparity Studies	57
Urban Institute Meta-Analysis	57
Department of Commerce "Ready, Willing, and Able" Analyses	58
Small Business Administration Study	59
Measuring Discrimination	60
Concluding Observations	60
WRITTEN STATEMENT OF IAN AYRES, WILLIAM K. TOWNSEND PROFESSOR, YALE LAW SCHOOL	63
I. There is Credible Evidence of a Compelling Government Interest	65

II. There are Persuasive Statistical Methods for Calculating Disparity Benchmarks.....	66
III. The Constitutionality of the Race-Conscious Affirmative Action in Educational Admission Strengthen the Inference that the Commerce Department’s Disparity Study Provides Credible Evidence of Narrow Tailoring. . .	68
Conclusion	69
THE USE OF DISPARITY STUDIES TO JUSTIFY RACIAL AND ETHNIC PREFERENCES IN GOVERNMENT CONTRACTING, PREPARED STATEMENT OF ROGER CLEGG, VICE PRESIDENT AND GENERAL COUNSEL, CENTER FOR EQUAL OPPORTUNITY .71	
THE COMMISSION’S FINDINGS AND RECOMMENDATIONS.....	75
States and Localities Conducting Disparity Studies	76
National Disparity Studies	79
Research on Discrimination	80
Guidance from the Department of Justice.....	81
COMMISSIONERS’ STATEMENTS [TO BE ADDED].....	83
APPENDIX A: ACRONYMS USED IN THE REPORT.....	87
APPENDIX B: PANELISTS’ BIOGRAPHIES	89

OVERVIEW

A decade ago, in the landmark case of *Adarand v. Peña*, the Supreme Court, quoting *Richmond v. Croson*, held that “[a]bsent searching judicial inquiry into the justification for ... race-based measures, [one cannot distinguish] ‘benign’ or ‘remedial’ ... classifications [from those] motivated by illegitimate notions of racial inferiority or simple racial politics.”¹ In other words, the Court held that federal programs that use racial classifications are subject to the “strictest judicial scrutiny.”² Following the Court’s decision, the Justice Department issued guidelines to govern the use of race-conscious remedies.³ The Justice Department’s guidance to federal agencies said:

[T]he mere fact that there has been generalized, historical societal discrimination in the country against minorities is an insufficient predicate for race-conscious remedial measures; the discrimination to be remedied must be identified more concretely. The federal government would have a compelling interest in taking remedial action in its procurement activities, however, if it can show with some degree of specificity just how ‘the persistence of both the practice and the lingering effects of racial discrimination’—to use Justice O’Connor’s phrase in *Adarand*—has diminished contracting opportunities for members of racial and ethnic minority groups.⁴

The Supreme Court did not decide whether the procurement program at issue in *Adarand* served a compelling interest. Thereafter, the President directed agencies to review federal affirmative action programs in light of *Adarand*. In response, the Justice Department analyzed Congressional hearings, academic social science, and local disparity studies, from which it built the case that race-based federal procurement programs indeed served a compelling interest.⁵ Next, the Department developed a method by which federal agencies could narrowly tailor programs to serve the government interest in ending discrimination. Its guidance promised that the Department of Commerce would perform industry-specific research intended to estimate “the level of minority contracting that one would reasonably expect to find in a market absent

¹ *Adarand Constructors, Inc. v. Peña*, 515 U.S. 200, 226 (1995) (hereafter cited as *Adarand*).

² *Id.* at 224. “Strict scrutiny” is the standard courts use to determine whether a government program that imposes benefits or burdens on the basis of race or national origin is constitutional. Under this standard, the government must show that the program is the least restrictive way to serve a compelling public interest and is narrowly tailored to meet that interest. Under the narrow tailoring requirement, for example, agencies must consider race-neutral means of addressing the governmental interest before resorting to race-conscious measures. *Id.*, at 224, 227, 237–38. For further discussion, see U.S. Commission on Civil Rights (USCCR), *Federal Procurement after Adarand*, September 2005, pp. 1-14. This report also presents strategies agencies may use to demonstrate that they have seriously considered race-neutral alternatives. *Ibid.*, pp. 31–66.

³ Proposed Reforms to Affirmative Action in Federal Procurement, 61 Fed. Reg. 26,042, (May 23, 1996); 48 C.F.R. § 19.201 (2005).

⁴ 61 Fed. Reg. at 26,050; *Adarand* at 237.

⁵ 61 Fed. Reg. at 26,050–63.

discrimination or its effects.”⁶ Justice asked Commerce to determine areas where the comparison between the benchmark and actual small disadvantaged business utilization did not indicate a disparity suggesting discrimination or its continuing effects. Justice prohibited federal agencies from using race-conscious mechanisms in procurements in those areas.⁷ The Justice issuance argued that the benchmarks “represent a reasonable effort to establish guidelines to limit the use of race conscious measures and to meet the requirement that such measures be narrowly tailored to accomplish the compelling interest that Congress has identified.”⁸ Federal procurement policy remains based on the Department of Justice’s analysis of a compelling interest and method of narrow tailoring.

The government relied on three seminal efforts—the 1996 appendix to Department of Justice guidance summarizing the agency’s analysis,⁹ a 1997 Urban Institute report collecting state and local research on disparities,¹⁰ and 1998 and 1999 benchmark studies from the Department of Commerce¹¹—to offer evidence that discrimination exists in federal contracting. However, some critics have expressed doubt that the studies sufficiently prove discrimination in procurement. Some have cited obsolete statistics, a lack of documentation of sources, faulty analytical methods, overly narrow application of results, failure to develop industry groupings related to federal contracting, and lack of an adequate theory of discrimination as bases for their criticisms.

The Commission conducted a briefing to gather facts so that it could better evaluate the methodological and empirical strength and quality of these seminal efforts and subsequent disparity studies, which in part form the foundation of affirmative action in federal contracting.

⁶ *Id.* at 26,045.

⁷ *Id.* at 26,046–47; 48 C.F.R. § 19.201 (b) (2005).

⁸ 61 Fed. Reg. at 26,046.

⁹ *Id.* at 26,050.

¹⁰ María Enchautequi, Michael Fix, Pamela Loprest, Sarah von der Lippe, and Douglas Wissoker, *Do Minority-Owned Businesses Get a Fair Share of Government Contracts?* (Washington, D.C.: The Urban Institute, 1997).

Many state and local governments commissioned disparity studies to demonstrate differences between the proportion of minority-owned businesses in their jurisdictions and the share of local contracting the minority enterprises received. These studies responded to the 1989 case, *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469 (1989). In *Croson*, the Supreme Court first established the strict scrutiny standard for state and local programs providing preferences to minority-owned firms—racial preferences must serve a compelling interest and be narrowly tailored to meet that need. *Id.* In 1995, *Adarand* extended the *Croson* strict scrutiny standard to apply to federal affirmative action programs. See U.S. Department of Justice, “Legal Guidance on the Implications of the Supreme Court’s Decision in *Adarand Constructors, Inc. v. Peña*,” memorandum to general counsels, June 28, 1995 <www.usdoj.gov/olc/adarand.htm> (last accessed Feb. 22, 2006), p. 1. Thus, state and local entities and the federal government face similar requirements for programs providing preferences to minority-owned businesses.

¹¹ Small Disadvantaged Business Procurement; Reform of Affirmative Action in Federal Procurement, 63 Fed. Reg. 35,714 (June 30, 1998); Small Disadvantaged Business Procurement; Reform of Affirmative Action in Federal Procurement, 64 Fed. Reg. 52,806 (Sept. 30, 1999).

Among objectives, the Commission sought to examine whether the foundation is sound and, if not, how researchers can improve the quality of evidence which studies offer.

Four experts briefed the Commission on the quality of disparity studies under use for federal procurement policy purposes: (1) Dr. Ian Ayres, William K. Townsend Professor, Yale Law School, was a consultant on the design of the Department of Commerce's 1998 and 1999 benchmark studies and has since given expert testimony on many affirmative action contracting cases; (2) Dr. Constance F. Citro, director, Committee on National Statistics, National Academy of Sciences, supervised a recently completed evaluation of disparities in federal contracting with women-owned businesses¹² and worked on a project about measuring racial discrimination;¹³ (3) Roger Clegg, President of the Center for Equal Opportunity and former senior Justice Department official, researches legal issues arising from civil rights laws and addresses them in his writing and speeches; and (4) Dr. George R. LaNoue, professor, University of Maryland, Baltimore County, has compiled a library and database of minority- and women-owned business programs containing research related to more than 160 disparity studies, which is known as the Project on Civil Rights and Public Contracts.

The panelists discussed the continuing need to conduct and improve disparity studies because the existing research relies on data which are now a decade or more old. The discussants referred to state and local disparity studies and the key national compilations of information mentioned above in articulating appropriate social science standards for new research.

Summary of Ian Ayres' Statement

Dr. Ayres asserted that quantitative methods have resulted in rigorous and persuasive evidence that race-conscious programs meet the requirements of strict scrutiny and are narrowly tailored to ensure that minorities and women receive an appropriate share of federal contracts.¹⁴

According to him, disparity studies and other research shows evidence of discrimination in some industries. The Commerce Department's research, known as the "benchmark studies"¹⁵ created a mechanism by which the government may accurately tailor the use of contracting preferences, allowing them only in industries that show evidence of underutilization of minorities. Dr. Ayres warned that the narrow tailoring requirement should not be so burdensome that agencies cannot meet it. According to Professor Ayres, disparity studies surpass the level of persuasive evidence

¹² National Research Council, *Analyzing Information on Women-Owned Small Businesses in Federal Contracting*, (Washington, DC: The National Academies Press, 2005) (hereafter cited as NRC, Women-Owned Small Businesses in Federal Contracting).

¹³ National Research Council (NRC), *Measuring Racial Discrimination*, Rebecca M. Blank, Marilyn Dabady, and Constance F. Citro, eds., (Washington, DC: The National Academies Press, 2004).

¹⁴ U.S. Commission on Civil Rights, briefing on disparity studies as evidence of discrimination in federal contracting, Washington, DC, Dec. 16, 2005, transcript, p. 26 (hereafter cited as Disparity studies briefing transcript).

¹⁵ 63 Fed. Reg. 35,714 (June 30, 1998); 64 Fed. Reg. 52,806 (Sept. 30, 1999).

the Supreme Court accepted in its recent decision on affirmative action programs at the University of Michigan Law School.¹⁶

Dr. Ayres acknowledged that, having previously reviewed their work, the panelists agreed with the objective to remedy discrimination as a compelling governmental interest. Furthermore, he argued that although the government took steps to end discrimination against minority contractors in procurement, private vendors (for example suppliers to and customers of minority contractors) continue to exhibit bias. Although individual cases cannot prove discrimination, aggregate analysis reveals patterns that justify a compelling government interest in remedying injustices.¹⁷

The disparity studies under discussion for the briefing calculate a benchmark (or minority availability percentage)—a measure of the market or share of contracts that minority firms would receive in the absence of discrimination. Benchmarks establish both whether (1) minorities face discrimination, and (2) the proposed racial preferences are sufficiently limited (that is, narrowly tailored) to remedy bias without overcorrection.¹⁸

Professor Ayres explained that over time, the methods of estimating benchmarks have become more sophisticated. Early disparity studies counted minority-owned firms as a proportion of all firms. They assumed that without discrimination minority-owned enterprises would receive a proportion of procurement dollars equal to their frequency compared to all other firms. Courts, however, rejected such methods as simplistic because many firms were not qualified to conduct business for the federal government. Thus, courts demanded approaches counting only qualified firms.¹⁹ Researchers then attempted to determine how many firms are ready, willing, and able to do business with the government, but soon recognized that even qualified businesses have substantially different capacities. For example, some firms have multinational establishments while others are sole proprietorships with a single facility. Thus, all may be qualified, but not necessarily equally available to serve the government.²⁰

To resolve this dilemma, Dr. Ayres explained, the Commerce Department computed benchmarks using a *capacity approach*. It calculated capacity of qualified firms to do business with the government using a complex formula of firms' ages and payroll (measured in thousands of dollars). With this approach, Commerce assumed that, absent discrimination, small disadvantaged businesses (SDBs) should be awarded a percentage of the industry's procurement dollars proportional to the amount of the industry's capacity they control. This approach is

¹⁶ Disparity studies briefing transcript, pp. 26–27.

¹⁷ *Ibid.*, pp. 28–29.

¹⁸ *Ibid.*, pp. 29–30.

¹⁹ *Ibid.*, pp. 30–31.

²⁰ *Ibid.*, p. 31.

conservative because it estimates benchmarks based on the prevailing state of minority businesses. For example, some argue that discrimination may have suppressed minority capacity, and the *real* disparities are even greater than these benchmarks show. Dr. Ayres supported the capacity approach as credible and rigorous for collecting evidence of discrimination and setting limits to narrowly tailor race-conscious remedies. He recommended that the Commerce Department update its benchmark analysis from 1999 before agencies abandon remedial race-conscious programs for the lack of current evidence to support their narrow tailoring.²¹

Summary of Constance Citro's Statement

Dr. Citro explained that a disparity is simply a difference between two groups on an outcome of interest. Social science theory suggests that discrimination should result in an observable disparity. However, she said, a measured disparity does not necessarily imply discrimination because any number of factors could cause the difference. Finding disparities is a first step in identifying discrimination.²²

In contracting, researchers typically measure differences using a disparity ratio. The ratio's numerator represents utilization, for example, the share of contracts or contract dollars that women or minority-owned businesses possess. The denominator includes the pool of women or minority-owned businesses that are available for contracting as a percentage of all available firms. When the numerator divided by the denominator yields a ratio of one, there is no disparity; that is, the share of contracts is commensurate with the available pool of businesses. If the ratio is less than one, the share of contracts is not as great as the proportion of available firms.²³

In contracting, the availability share for a group such as women-owned small businesses varies across industries and other characteristics of the firms. Therefore, one must use disparity ratios that are computed within meaningful categories and not merely simple counts or percentages of utilization. In addition, researchers must address many methodological issues to obtain statistically defensible, valid, and reliable measures of disparities.²⁴ Dr. Citro discussed a few such issues in her remarks and presented others in her written comments.²⁵

As an example of one weakness, Dr. Citro asserted that most work to measure disparities does not use comparable information in the numerator and denominator of ratios. Commonly, the numerator measures contract dollars awarded to a target group, such as women-owned small businesses. The typical denominator for availability measures numbers of businesses without accounting for firm size, revenues, gross sales, or similar factors, and consequently skews

²¹ Ibid., pp. 32–33.

²² Ibid., pp. 36–37.

²³ Ibid., pp. 37–38.

²⁴ Ibid., p. 38.

²⁵ Ibid., p. 38. *See her written statement herein, or NRC, Women-Owned Small Businesses in Federal Contracting.*

results, distorting disparity. Dr. Citro recommended using the same units in both numerator and denominator and using multiple measures of disparities.²⁶

The availability measure (that is, the estimate of minority businesses available to do business with the government or the denominator in the disparity ratio) is often in dispute. Dr. Citro agreed with Professor Ayres that including all businesses in the denominator as ready, willing, and able produces a too broad availability pool because many businesses are not interested or able to provide the goods or services of government contracts. On the other hand, a very narrowly tailored measure of businesses' availability probably excludes some potential bidders by falsely classifying them as unable to perform the requirements of contracts.²⁷

Dr. Citro concluded that disparity studies are a reasonable first step to identify situations in which discriminatory practices or behavior disadvantage certain types of businesses in government contracting. However, she added that observed disparities do not establish discrimination. To identify discrimination, the researcher must examine various aspects of contracting, for example, earlier causal effects such as the processes by which pools of ready, willing, and able bidders are developed.²⁸

Because disparity studies are a first step to identifying discrimination, Dr. Citro recommended that they meet high standards for validity, reliability, and reproducibility. She advised that researchers: (1) thoroughly document all data, methods, evaluations, and results; (2) use the same metric in the numerator and denominator of disparity ratios (i.e., either dollars or numbers, but not a mixture of the two); (3) represent the same time period for utilization and availability to avoid distortion from changes in the composition of the business community; (4) calculate more than one type of disparity ratio to ensure that results do not depend on a particular measure; (5) test for the sensitivity of results to variations in methods and data, effects of different groupings of industries, and the presence of outliers; and (6) develop explicit rationales for which businesses are included in the availability measure as ready, willing, and able.²⁹

Apart from these recommendations for future disparity studies, Dr. Citro commented on previous national research efforts. First, regarding the Department of Commerce's benchmark studies, Dr. Citro agreed with Dr. Ayres that a capability analysis, such as measuring payroll and years of experience, is worth exploring. However, Dr. Citro did not recommend replicating Commerce's benchmark procedure because the methodology, including the statistical equations measuring capacity, and the sensitivity of their effects on the study results, is undocumented.³⁰

²⁶ Disparity studies briefing transcript, pp. 38–39.

²⁷ *Ibid.*, p. 39.

²⁸ *Ibid.*, p. 40.

²⁹ *Ibid.*, pp. 40–41.

³⁰ *Ibid.*, p. 42.

Second, Dr. Citro briefly discussed strengths and weaknesses of the 1997 Urban Institute meta-analysis (i.e., examination of other studies). The results of this report are very out of date and apply only to state and local government contracting in specific jurisdictions. She praised the report as “a model of careful specification, sensitivity analysis, and thorough documentation.”³¹ Because the report lists the disparity studies in its analysis, other researchers can apply variations to better understand the findings. The National Research Council employed this technique to advantage in its recent review of contracting with women-owned businesses.³²

Finally, Dr. Citro explained that in 2002 the Small Business Administration (SBA) completed a preliminary analysis of disparities for women-owned businesses. By SBA’s request, a National Research Council panel, which Dr. Citro directed, reviewed the SBA effort. Based on the resulting criticisms, SBA withdrew its report for revision.³³

Dr. Citro recommended further research on the contracting process with an objective to inform government agencies of the possible sources of disparities and ways to increase access for all types of businesses. Such a research agenda would examine contracting in various industries and agencies and develop a richer and deeper understanding than emerges from current disparity studies. It could draw on case studies, administrative records, and statistical analyses.³⁴

Summary of Roger Clegg’s Statement

Mr. Clegg observed that the panelists shared certain common views. He emphasized the distinction between disparities and discrimination, agreeing that factors other than bias may explain differences. Second, Mr. Clegg noted agreement among panelists that federal agencies should not extend racial preferences unless they have found discrimination, narrowly tailored the remedy to correct the bias, and tried race-neutral measures and found them ineffective.³⁵ Mr. Clegg said that in an August 2001 brief, “the Justice Department told the Supreme Court that the federal program at issue ‘may use race-conscious remedies only as a last resort’ ‘where the effects of discrimination are stubborn, persistent, and incapable of eradication through race-neutral measures.’”³⁶

Mr. Clegg explained that, in his view, racial preferences constitute discrimination. He articulated his belief that even if a study identifies discrimination (as opposed to disparities), agencies can

³¹ Ibid., p. 42.

³² Ibid., p. 42.

³³ Ibid., p. 42.

³⁴ Ibid., p. 43.

³⁵ Ibid., pp. 43–45

³⁶ Ibid., p. 45. See Brief for the Respondents at 38–39, *Adarand Constructors v. Mineta*, No. 00-730, (U.S. Supreme Court filed in June 1, 2001) (in part citing 49 C.F.R. 26.51(a) (2005)).

eliminate such unfairness in better ways than through racial preferences. For example, he explained, if federal contracts require unrealistic or irrational bonding or are bundled together into work orders too large for small firms to fulfill, agencies should change those aspects for all companies, regardless of minority or female ownership. In instances in which the lowest bidders are denied government contracts because of race, a federal system should employ safeguards to detect discrimination, and sanctions to punish such bias. However, any safeguards and sanctions should protect all companies from discrimination, not just those that women or minorities own.³⁷

Mr. Clegg further explained that just as a disparity need not imply racial bias, the lack of under-representation does not eliminate the possibility of discrimination. He gave the example of Asian Americans, who may be over-represented in acceptance to universities relative to their proportions in the populations, but are still victims of discrimination. Measures to fight discrimination should protect everyone, he said, including individuals belonging to groups that seem over-represented.³⁸

Inevitably, a few cases of discrimination may go unremedied, Mr. Clegg suggested. However, establishing a regime to institutionalize discrimination in an opposite direction by extending preferences to minorities who are not victims may not be an appropriate, narrowly tailored, and fair means to eradicate the remaining instances. He recommended race-neutral approaches and praised a recent Commission report on federal contracting,³⁹ for collecting and discussing such techniques. He stressed that race-neutral alternatives aim to correct and end discrimination, not to achieve a particular percentage of contract awards to specific racial or ethnic groups.⁴⁰

Finally, Mr. Clegg asserted that some government entities have used race-neutral approaches very successfully. He cited the State of New Jersey, which adopted neutral alternatives after a court order struck down the use of racial preferences. Mr. Clegg supported race-neutral means even when disparity studies show persuasive evidence of discrimination.⁴¹

Summary of George LaNoue's Statement

Dr. LaNoue commented that properly done disparity studies may highlight the consequences of discrimination and provide useful information for eliminating any ill effects. Improper studies, however, create false or misleading claims of discrimination that contribute to racial polarization and suppress interest in searching for race-neutral programs that promote new opportunities. He agreed with other panelists that even at their best, disparity studies rarely identify the source of

³⁷ Disparity studies briefing transcript, pp. 45–46.

³⁸ Ibid., pp. 46–47, 119.

³⁹ U.S. Commission on Civil Rights, *Federal Procurement After Adarand*, September 2005.

⁴⁰ Disparity studies briefing transcript, pp. 47–48.

⁴¹ Ibid., pp. 48–49.

discrimination with any precision and, thus, require supplemental information to form the basis of an appropriate public policy.⁴²

Most disparity studies have common flaws, Dr. LaNoue said. Their statistical sections fail to measure availability according to *Croson* requirements to compare qualified, willing, and able businesses that perform similar public services. Echoing other panelists' concerns, Dr. LaNoue estimated that about 50 percent of state and local disparity studies use simple counts of businesses and are inadequate because they do not measure availability properly. He dismissed the results of another 30 percent because the studies did not factor in the capacity of firms to fulfill contracts; and 18 percent because they fail to examine disparities by industry. Together, only about 2 percent of state and local disparity studies meet requirements for availability measures and narrow tailoring.⁴³

Dr. LaNoue identified other flaws in many of the same disparity studies, some of which other panelists already had noted. He found such studies: (1) are based on obsolete or incomplete data; (2) report results in ways that exaggerate disparities; (3) fail to test for nondiscriminatory explanations for the differences; and (4) make findings of discrimination without identifying any specific instances or general sources of biased behavior.⁴⁴

Dr. LaNoue discussed the importance of anecdotal evidence for establishing discrimination and the poor quality of such information included in disparity studies. Anecdotal sections of disparity reports use samples that are not scientifically collected. Researchers base conclusions on very small percentages of the survey universe. They report allegations anonymously and fail to verify them. Furthermore, because of the lack of identifying information contained in anecdotes, government staff cannot evaluate the information.⁴⁵

The University of Maryland professor also commented on key national disparity studies. The Commerce benchmark studies are almost a decade out of date and their underlying methodology is flawed. Furthermore, the studies measured disparities for small disadvantaged businesses, but did not specify which racial and ethnic groups suffer from disparities. Current law requires that studies isolate discrimination against a particular racial or ethnic group before agencies may employ race-based strategies. Yet the outdated and flawed benchmarks provide the only statistical information to support narrow tailoring of federal race-conscious contracting programs.⁴⁶

⁴² Ibid., p. 51.

⁴³ Ibid., pp. 52–53.

⁴⁴ Ibid., pp. 53–55.

⁴⁵ Ibid., p. 55.

⁴⁶ Ibid., pp. 55–56.

The Department of Justice's 1996 Appendix A, Dr. LaNoue said, relied on secondary research, not original work; that information is now too old a source from which to draw conclusions about discrimination today, and is not credible in some points. The Urban Institute research is based on state and local disparity studies, which are overwhelmingly flawed and furthermore too old to describe today's situation. As Dr. Citro had before him, he described SBA's recent effort to create a statistical basis for justifying contracting preferences to women as flawed.⁴⁷

Neither the current Bush administration nor recent Congresses, George LaNoue reported, had made public any intentions to update disparity analyses. Nor has the Bush Justice Department appealed a decision of the Ninth Circuit, in *Western States v. Washington State Department of Transportation*, that recipients of federal transportation funding must gather local evidence of contracting discrimination against specific groups before setting any race-conscious goals. Dr. LaNoue suggested that the decision in this case amounted to requiring state and local jurisdictions to conduct new studies before they continue to offer racial preferences in federal contracting.⁴⁸

Dr. LaNoue concluded that governments need guidance about acceptable methodologies and roles for disparity studies. He further suggested that the Commission advise federal agencies to (1) provide a well-functioning complaint system to evaluate and remedy valid claims of discrimination; (2) improve communication about contracting opportunities; (3) develop technical assistance for recipients to identify and eliminate awards or processes that are discriminatory; and (4) implement race-neutral approaches to overcome problems that hinder minority as well as other new and small businesses in bidding on contracts. To develop race-neutral approaches he recommended research to examine, for example, supplier policies with differential pricing based on economies of scale or credit status, overly large contracts, requirements for bonding, difficulties in obtaining credit, and experience, or lack of it, in developing business plans to obtain financing or bonding. Any entity wishing to prove serious consideration of race-neutral alternatives, he said, should demonstrate active problem identification and creative solutions.⁴⁹

Discussion

A question and answer period followed the speakers' remarks during which time Commissioners offered insights and further probed the panelists about methodological standards and other relevant topics.

Is Race-Conscious Affirmative Action an Appropriate Remedy for Discrimination?

Panelists disagreed about whether the government should employ race-conscious affirmative action programs to remedy discrimination. Dr. Ayres opined that if affirmative action programs

⁴⁷ Ibid., p. 56.

⁴⁸ Ibid., p. 56–57.

⁴⁹ Ibid., pp. 57–58.

are narrowly tailored, race-conscious spending is normatively appropriate and constitutional.⁵⁰ He used the example of Marian Anderson, an African-American opera singer of international renown. In 1939, the Daughters of the American Revolution denied Ms. Anderson the use of the organization's facility, Constitution Hall, to give a concert because of her skin color. The Secretary of Interior intervened and arranged for her to perform an outdoor concert at the Lincoln Memorial. Dr. Ayres and Mr. Clegg disagreed about how to characterize the Secretary's act. Dr. Ayres described opening the Lincoln Memorial for the concert as constitutional, race-conscious affirmative action. Mr. Clegg, however, said the act was not discriminatory. The government did not give Marian Anderson a preference because she was black or refuse to allow any white person to sing, but merely remedied a particular case of discrimination against an individual.⁵¹ Mr. Clegg said he would object to a hypothetical government policy that would protect African Americans against discrimination, but would not protect Latinos, Asians, or whites against similar discrimination.⁵²

Commissioner Thernstrom asked for clarification on what qualifies as a race-neutral program. For example, the Institute of Justice located a clinic helping entrepreneurs comply with regulations on the south side of Chicago, where minority business-owners are the primary beneficiaries.⁵³ Mr. Clegg responded that despite its targeted location the program may well still be race-neutral. He explained that adopting a measure or criteria that disproportionately includes African Americans may still be race-neutral but that race-neutrality would require, at a minimum, that the program not turn away applicants because of race or ethnicity. The program, for example, must allow poor white people on the south side of Chicago access if they apply. However, Mr. Clegg warned against intentionally choosing neutral criteria that are related to race. For example, he explained, an employer could require janitors to have a high school diploma, which could disproportionately exclude African Americans if they are less likely to have graduated than white applicants. Choosing such hiring criteria is discriminatory, however, if they are selected with the purpose of excluding African Americans. The federal government should implement race-neutral alternatives in contracting not because they help or hurt particular racial or ethnic groups but because they eliminate discrimination and irrational contracting practices, Mr. Clegg said.⁵⁴

Dr. LaNoue added an explanation of his objections to the preferences in some federal contracting programs. Minority business programs, whether the 8(a) program (so named on the basis of its implementing statutory section) or state or local initiatives, certify firms as minority- or women-owned enterprises. To obtain such certifications, applicants must indicate the racial, ethnic, or

⁵⁰ Ibid., p. 70.

⁵¹ Ibid., pp. 70–71.

⁵² Ibid., p. 71.

⁵³ Ibid., p. 121.

⁵⁴ Ibid., pp. 121–125.

gender ownership of the business and the enterprise must meet certain size and net worth criteria. Owners need not demonstrate that they or their businesses have been victims of discrimination. The objectionable remedies, therefore, prefer firms owned by a particular racial or ethnic group and not firms of other groups such as whites (or even minorities in industries showing no disparities). He suggested that a remedy affecting any firm owned by a person of a particular group because of some statistical disparity is overly broad in the absence of proof of a pattern of discrimination against those particular firms. The preferences such remedies provide could last decades, he said, and the number of contracts obtained through them is unlimited.⁵⁵

Conversely, Dr. Ayres warned that requiring beneficiaries of race-conscious preferences to show evidence that disparate treatment harmed them in specific instances is inappropriate. He stated that so doing disaggregates information to the point of examining an individual minority contractor with respect to a specific contract. But individual cases do not help identify patterns of discrimination nor support the systemic remedies the nation needs given its ongoing history of racial discrimination.⁵⁶

Mr. Clegg, however, asserted that according to the Supreme Court, proving societal discrimination does not justify the use of racial and ethnic preferences. He argued that evidence must at a minimum prove discrimination in at least some specific cases if agencies use it to justify race conscious programs. The evidence does not need to prove that an individual did not win a contract because of race in every instance. However, Mr. Clegg added, one cannot dismiss evidence that the agency denied that individual the contract for some reason other than race. Researchers must investigate obvious explanations and creatively explore reasons that might explain disparities other than race or ethnicity, he said.⁵⁷

Mr. Clegg further opined that even though instances of discrimination continue to exist in the United States and even if the most rigorous disparity study finds evidence of it, the use of racial preferences to rid any remaining bias is unconstitutional and unjustified. Using racial preferences has enormous costs, he said. Race-based programs abandon the principle of nondiscrimination and establish policy for government to treat citizens differently on the basis of skin color or national origin. In such programs, the government does not award the contract to the best bidder, but bases decisions on color. Such awards are unfair to the nonminority firms that lose contracts, build general resentment among other owners who fear or have experienced similar losses, and stigmatize minority enterprises as not capable of winning contracts without special preferences.⁵⁸

⁵⁵ Ibid., pp. 71–72.

⁵⁶ Ibid., pp. 106–107.

⁵⁷ Ibid., p. 112.

⁵⁸ Ibid., pp. 82–84.

If a company practices racial discrimination, the remedy is to make the firm stop, said Mr. Clegg. He explained that when a victim sues a company for employment discrimination under Title VII, the court should order it to cease discriminating and hire the best qualified people regardless of their race, ethnicity, or gender. Imposing a quota on that employer to hire a certain number of Asian Americans, Latinos, and females is not the remedy, nor is such an approach nondiscriminatory.⁵⁹

Commissioner Yaki suggested that the effects of ordering a company to cease its discrimination are small. The firm's size may afford few opportunities to hire new employees. Thus, he argued, American society needs racial preferences based on disparity studies to achieve the desired goal of equality for all people and to quantify how to reach this end. The Supreme Court allows the use of racial preferences when no other way of reducing disparities proves effective, he argued. According to Mr. Yaki, Justice O'Connor clarified that this does not require the exhaustion of every conceivable race-neutral alternative.⁶⁰ Commissioner Yaki asked if Mr. Clegg could identify any condition under which a racial preference is appropriate. Mr. Clegg said that in his mind racial and ethnic preferences make no constitutional and policy sense given the alternatives available and the inevitable costs he had articulated.⁶¹

Commissioner Yaki further suggested that the poor response of the federal government in providing relief from the effects of Hurricane Katrina, and the seeming suspension of all programs to promote and encourage the use of minority firms when contracts were let for rebuilding New Orleans, provided examples of the need for racial preferences to combat continuing racial discrimination.⁶² Mr. Clegg stated that because of the enormous expenditures to repair the damage of Katrina, the federal government must award contracts to firms that can do the work at the least possible cost to avoid waste and help the thousands of people whose lives the hurricane devastated. If government awards contracts on political bases rather than to the best bidders, Congress should intervene. He suggested that white- and Asian-owned companies as well as African American-owned ones suffer if the government awards contracts politically. The remedy is to stop awarding contracts according to political connections and give them to the best-qualified companies. Racial preferences or setting aside a certain number of contracts for African Americans and another set for whites is not the solution.⁶³ Mr. Yaki suggested that, nonetheless, some filter encourages the establishment of white, male-dominated companies and feeds their proliferation while blocking minority firms. The Supreme Court has adopted, and continues to uphold, the use of race-conscious programs to combat such a mechanism.⁶⁴

⁵⁹ Ibid., pp. 84–86.

⁶⁰ Ibid., pp. 85–89.

⁶¹ Ibid., pp. 88–89.

⁶² Ibid., pp. 90–91.

⁶³ Ibid., pp. 91–92.

⁶⁴ Ibid., p. 92.

Government Remedies to Private Sector Discrimination

In the discussion, panelists, particularly Professor Ayres and Dr. LaNoue, drew distinctions between addressing government discrimination and that in the private sector. Dr. Ayres asserted that Justice O'Connor intended for local governments to use their spending power to eradicate the effects of private discrimination. Thus, Dr. Ayres suggested, local governments can use racial preferences to disrupt the effects of discrimination they cannot stop. Dr. Ayres argued that race-neutral alternatives are more appropriately directed toward remedying government discrimination. Private discrimination is less easily remedied using neutral approaches, he opined.⁶⁵

Dr. LaNoue agreed that discrimination is more likely found in the private sector than in government. He noted that state and local governments rarely prohibit discrimination in private contracting. Dr. LaNoue disagreed with Professor Ayres, however, about the effectiveness of race-neutral remedies for private sector discrimination. According to Dr. LaNoue, and Mr. Clegg agreed, one *can* identify sources of disparity in the private sector and remedy them with race-neutral means.⁶⁶

Dr. LaNoue offered two examples. First, in a St. Petersburg disparity study, minority-owned small businesses claimed notice of subcontracting opportunities was insufficient in the private sector. Although the public sector posts opportunities on a Web site accessible to everyone, the city did not require posting of local, private contracting opportunities. Dr. LaNoue suggested that in future the city should condition zoning and building permits on the willingness of the private contractor to publicly announce subcontracting opportunities in media accessible to everyone.⁶⁷

In the second example, small minority businesses encountered difficulties obtaining loans. Bank officials said that these firms lacked experience in drawing up business plans. Consequently, the Chamber of Commerce created a mentoring program for enterprises that need help in designing business plans.⁶⁸

Both examples offer race-neutral alternative methods that remove a barrier to increase opportunities. The key is to identify the obstacle, Dr. LaNoue said.⁶⁹

The Use of Race-Neutral Alternatives

Discussants suggested that contracting agencies may seldom pursue race-neutral alternatives. Dr. LaNoue commented that jurisdictions with a disparity study showing underutilization of

⁶⁵ Ibid., pp. 59–60.

⁶⁶ Ibid., pp. 60–63.

⁶⁷ Ibid., p. 61.

⁶⁸ Ibid., p. 62.

⁶⁹ Ibid., p. 62.

minorities almost always establish a program with goals for contracting with minority-owned businesses. They typically do not try to identify any source of the differences nor implement race-neutral alternatives. He elaborated that jurisdictions commonly lack complaint procedures for discrimination related to contracting that might help identify problems. Furthermore, a goals program that appears to promise contracts to minorities or women is more politically attractive to those constituencies than race-neutral alternatives.⁷⁰

How to Measure Discrimination

Dr. Citro explained that a disparity study does not demonstrate discrimination. A measured disparity does not locate discrimination in time and space, either in the contracting process or in the practices that create a pool of ready, willing, and able vendors. Disparity studies are a starting point for a research program to identify discrimination, perhaps using case studies.⁷¹

Research to find racial discrimination typically performs statistical analyses (known as “regressions”) on data from national surveys, adding in many factors that might explain disparities, Dr. Citro said. After these additions, researchers view any residual, unexplained disparities as a measure of discrimination against the designated racial or ethnic group. Dr. Citro pointed out flaws in such approaches in some of her work at the National Academies.⁷² For sound studies of discrimination, she recommended case studies examining specific kinds of businesses and the selection process—that is, the criteria and methods of decisions agencies use in awarding contracts—to lend depth to and inform a robust statistical analysis and help understand the process of federal contracting.⁷³

Furthermore, case studies are useful, she said, but researchers cannot possibly perform a case study on every business in a jurisdiction.⁷⁴ In addition to case studies, she suggested that the federal government should conduct solid research on discrimination along the lines of the field experiments, known as “paired testing,” that the Department of Housing and Urban Development has conducted with respect to housing.⁷⁵ She suggested that the National Science Foundation might provide research grants to analyze discrimination in federal contracting.⁷⁶

⁷⁰ Ibid., pp. 73–75.

⁷¹ Ibid., pp. 76–77.

⁷² Ibid., p. 77. *See also*, NRC, *Measuring Racial Discrimination*.

⁷³ Disparity studies briefing transcript, pp. 77, 113–114.

⁷⁴ Ibid., pp. 113–114.

⁷⁵ Ibid., pp. 95–96. In other work, Dr. Citro explains the audit or paired-testing methodology commonly used with respect to employment and housing.

Auditors or testers are randomly assigned to pairs (one of each race) and matched on equivalent characteristics (e.g., socioeconomic status), credentials (e.g., education), tastes and market needs. Members of each pair are typically trained to act in a similar fashion and are equipped with identical supporting documents.

Chairman Reynolds later echoed many of Dr. Citro's comments. He suggested that disparity studies do not consider many qualitative differences that determine whether or not minority-owned firms win contracts. He reiterated factors panelists mentioned, such as that minority firms are often newer and less experienced. They may lack knowledge about how to prepare a business plan to gain approval for a bank loan. To properly analyze such factors, he said, researchers must conduct expensive case studies, learning everything about a company to determine, for example, whether the price of the contract or timely product delivery determines its success. Research should capture many such factors to better measure discrimination, he said.⁷⁷

Similarly, Commissioner Thernstrom urged that research on discrimination in contracting factor in occupational choices that ethnic and national origin groups make. Groups cluster in particular occupations, she explained, not just according to standard racial and ethnic categories, but as Armenians, or Jews, for example. Cambodians completely dominate doughnut manufacturing in Los Angeles, she explained.⁷⁸

Dr. Ayres said that even the best disparity study performed in support of a broad government affirmative action program will exclude some variables that might provide alternative rationales. He agreed that a finding of disparity is not necessarily evidence of discrimination. The federal government must have additional evidence to ensure that affirmative action programs are narrowly tailored. Properly acquired anecdotal information, for example, is one type of additional evidence. He recommended using a variety of methods to measure discrimination, not just anecdotes and case studies.⁷⁹

Dr. Ayres added that studies should examine supplier industries. He recommended two methodologies aimed toward accumulating evidence of general societal discrimination. Referring to the first approach the "résumé test," Dr. Ayres explained that the researchers send out identical résumés, some with randomly assigned African American-sounding names and others with Caucasian-sounding names. Using such methodology, researchers have found that firms are much more likely to respond to Caucasian-sounding names than African American ones.

As part of the study, testers are sent sequentially to a series of relevant locations to obtain goods or services or to apply for employment, housing or college admission.... The order of arrival at the location is randomly assigned. For example, in a study of hiring, testers have identical résumés and apply for jobs, whereas in a study of rental housing, they have identical rental histories and apply for housing.

Researchers use the differences in treatment the testers experience to estimate discrimination. NRC, *Measuring Racial Discrimination*, p. 104.

⁷⁶ Disparity studies briefing transcript, pp. 77–78.

⁷⁷ Ibid., pp. 107–109.

⁷⁸ Ibid., p. 126.

⁷⁹ Ibid., pp. 115–116.

Although many studies have focused on the difficulties that minority- and women-owned firms encounter in obtaining credit, he suggested expanding this approach to examine firms that supply goods to, or are potential customers of, minority-owned firms. The results would enhance understanding about whether disparities result from discrimination.⁸⁰

Professor Ayres recommended a second study design in which the experimenter asks Internet users to sort a series of photographs of people representing different races or ethnicities according to some pertinent criterion. Even in such simple sorting tasks, Dr. Ayres reported that people engage in disparate treatment. Such results suggest an unconscious racial bias in society, Dr. Ayres said.⁸¹ However, Mr. Clegg and Commissioner Thernstrom cited work identifying problems with both methodologies that Dr. Ayres recommended.⁸²

The Methodology of Disparity Studies

Dr. Citro drew distinctions between studies that states or municipalities solicit to justify racial preferences that are narrowly tailored to local industries and situations and national research that may portray disparities or allege discrimination generally. In the former, the typical state or locality solicits a study of perhaps only 20 contracts, providing a weak basis for conclusions about disparities. To inform federal policy, however, researchers may combine the results of many such studies to obtain a clearer overall picture. The result of an amalgamation of studies, however, is not particular to a specific locality.⁸³

Individual Disparity Studies

The panelists generally agreed that existing state and local disparity studies are poor. As Commissioner Yaki pointed out, Dr. LaNoue is a critic of such studies, having supported only one—the St. Petersburg study—on which he served as consultant.⁸⁴ Dr. LaNoue explained that his negative views of disparity studies arise from the fundamental flaws in their availability measures, the unscientific manner in which researchers gather anecdotes, and the use of unverified information. He suggested that the studies were invalid on basic social science principles.⁸⁵ Dr. Citro agreed that existing disparity studies are mostly weak, but suggested that researchers could perform reasonably good studies.⁸⁶

⁸⁰ Ibid., pp. 116–117.

⁸¹ Ibid., pp. 117–118.

⁸² Ibid., pp. 118–119. See John McWhorter, *Winning the Race: Beyond the Crisis in Black America*, (New York: Manhattan Institute, forthcoming February 2006); and Amy Wax, "Some Truths About Black Disadvantage," *Wall Street Journal* Op-Ed, Feb. 7, 2005.

⁸³ Disparity studies briefing transcript, p. 98.

⁸⁴ Ibid., pp. 93–94.

⁸⁵ Ibid., p. 94.

⁸⁶ Ibid., pp. 95–97.

Dr. Citro recommended issuing guidance for state and local governments on how to conduct disparity studies with reasonably reliable and valid results. She reiterated a need for documentation, transparency (that is, allowing others to look at the data and reproduce the results), and analyzing alternative measures of disparity to ensure their robustness to different treatments.⁸⁷ She suggested experiments with variations in industry classifications, and including or excluding extremely different contracts, termed “outliers.” For example, she explained that in the current environment, very large contracts from the Iraq war could dominate results from more typical ones. Excluding the very large, atypical contracts might result in a more consistent picture of contracting. Both the Urban Institute meta-analysis and the National Research Council’s report that Dr. Citro directed engaged in such sensitivity analyses. She added that when results with different measures vary, then researchers must seek a deeper understanding of the data.⁸⁸

Dr. Citro reported that federal initiatives have improved the data available for conducting research.⁸⁹ She endorsed the Central Contractor Registration—a Department of Defense maintained Internet database of vendors seeking federal contracts—and the Census Bureau’s recently released 2002 survey of business ownership as sources for conducting credible disparity studies. These, she said, could prove useful in identifying factors contributing to disparities in federal contracting and serve as guidance for states and localities.⁹⁰ Such studies make better research possible at less cost and spread the expense among taxpayers rather than burdening individual states and localities.⁹¹

The panelists discussed different ways of examining data. Dr. Citro explained that disparity studies use inference to compare similarly situated white and black firm owners bidding on federal contracts. The statistics examine the minority-owned businesses’ share of contracts and their share of the pool of ready, willing, and able businesses. She added that some state and local studies look at these proportions within specific groups, but added that they could perform more detailed analyses. She suggested examining ratios within categories of contract amounts and business size. For example, studies usually compare small minority businesses to all other firms in the availability pool, rather than small minority businesses versus other small businesses.⁹²

Dr. LaNoue commented that researchers could look at bid success rates, for example, whether women-owned businesses are successful in 30 percent of their bids, while white male-owned firms are successful 35 or 25 percent of the time. Disparity studies do not perform such analyses,

⁸⁷ Ibid., pp. 63, 95, 97, 99.

⁸⁸ Ibid., pp. 63–64.

⁸⁹ Ibid., p. 99.

⁹⁰ Ibid., pp. 64–65.

⁹¹ Ibid., p. 78.

⁹² Ibid., p. 113.

he said. Such an analysis might reveal groups that bid frequently, but without success. It might point to reasons for their failure to win bids, for example, poor estimating skills (that training might correct), paying unnecessarily high prices for supplies or labor, or systemic racial or ethnic discrimination.⁹³

Dr. Ayres reiterated his preference for a capacity approach to developing benchmarks. He suggested that many state and local studies have not used such an approach. Some partly measure capacity by grouping qualified firms into categories by size—small, medium, and large. Dr. Ayres recognized studies using size groupings as better, but was unwilling to suggest that they adequately measure capacity or provide a strong basis for justifying benchmarks.⁹⁴

Dr. Ayres agreed that studies must be consistent in use of units in numerators and denominators of disparity ratios. He also urged that researchers pay attention to the degree of aggregation. He suggested that researchers sometimes mask statistical evidence of discrimination when analyses break the data into small subgroups.⁹⁵

Responding to a question about standards for the age of data from which to draw conclusions about discrimination, Dr. Ayres said that the Supreme Court will not allow the remedying of historical discrimination. Thus, he concluded, the data should not be so old that the government is remedying past discrimination. Contracting data more than ten years old is unreliable, he said, while urging that the federal government continue to test for discrimination with disparity studies.⁹⁶ Drs. Citro and LaNoue concurred that disparity studies must be updated.⁹⁷ Dr. LaNoue asserted that basing racial preferences on studies using data from 1996 and earlier to award contracts today is neither narrowly tailored nor fair.⁹⁸

Chairman Reynolds and others expressed concerns that a municipality may influence the result of a disparity study in the course of choosing the researcher.⁹⁹ Dr. Citro noted that perhaps only three firms have conducted most of the state and local disparity studies. She explained that the Urban Institute analysis examined whether the firm conducting the study made a difference in the results and found no effect. Dr. Citro recommended that federal analyses of local disparity studies test whether researchers prejudge the outcomes of disparity studies.¹⁰⁰ By implication,

⁹³ Ibid., pp. 114–115.

⁹⁴ Ibid., pp. 66–67.

⁹⁵ Ibid., pp. 67–68.

⁹⁶ Ibid., pp. 66, 68–69.

⁹⁷ Ibid., pp. 79–80.

⁹⁸ Ibid., p. 80.

⁹⁹ Ibid., p. 75.

¹⁰⁰ Ibid., p. 78.

state and local jurisdictions should heed the results of such examinations in choosing disparity researchers.

Dr. LaNoue recommended that states and localities employ universities, many of which have public policy research centers, to conduct disparity studies. So far, only a tiny fraction of academicians have done such studies, however, universities would contribute quality control to studies.¹⁰¹ Both he and Mr. Clegg offered examples of political pressure on consulting companies to produce research supporting race-based programs.¹⁰² Mr. Clegg added that such pressures are another reason he promotes race-neutral alternatives.¹⁰³ Thus, he suggests that state and local jurisdictions would not need disparity studies were their approaches to including minority- and women-owned businesses race-neutral.

National Disparity Studies

Panelists agreed that the government should conduct new studies, however, they disagreed on support for the benchmark studies or meta-analysis models. Dr. Ayres further defended the Department of Commerce's benchmark studies, taking issue with Dr. Citro's critique. Although he agreed with her quest for transparency, he pointed out that public release of raw data would violate Bureau of Census policy. He concurred, however, that Commerce could reveal documentation and statistical coefficients from past studies even in the absence of conducting more current research and called on the Department to do so. He also noted that regression coefficients and documentation appear in case law and are available through those public records.¹⁰⁴

Dr. Ayres also disagreed with Dr. LaNoue's criticism of the benchmark studies, specifying the suggestion that the studies should not include 8(a) firms as ready, willing, and able. Dr. Ayres explained that the Small Business Administration had certified the 8(a) firms as ready, willing, and able to contract with the government. Dr. Ayres dismissed concerns about whether unqualified firms slipped through this process. The capacity approach Commerce used, he said, would assign low capacity to new firms or those with little payroll, and thereby not inflate the availability percentage.¹⁰⁵

Dr. LaNoue further clarified his concern about including 8(a) firms in the availability ratio. He expressed that the benchmark study used three types of firms in its measure of availability: (1) the number of firms that bid on federal contracts; (2) firms that received source contracts; and (3) a list of 8(a) certified firms that included many that had neither bid on, nor received, contracts.

¹⁰¹ Ibid., pp. 99–100.

¹⁰² Ibid., pp. 99–100, 102–103.

¹⁰³ Ibid., p. 103.

¹⁰⁴ Ibid., pp. 100–101.

¹⁰⁵ Ibid., pp. 105–106.

Dr. LaNoue approved of measuring availability with the first two categories. Firms, he explained, do not prepare a bid unless they are willing, qualified, and have the ability to perform the work. Submitting a bid costs time and effort that firms waste if they do not win a contract or if they receive one for work they have under priced. The number of bid disqualifications is very small—one or two percent, he added.¹⁰⁶

Dr. LaNoue objected to using the third category because it included a group of firms that were potentially unqualified for, or unwilling to compete for, the available contracts. Although, as Dr. Ayres notes, the capacity measures may have reduced the effects of such unqualified firms, the Commerce study does not indicate how many such 8(a) firms the methodology added and what effects this had on the capacity ratios. Dr. LaNoue further suggested that some of the benchmark studies' capacity measures seemed implausible. For example, the study found one category of small, disadvantaged businesses was 80 percent of national capacity in that industry. Commerce did not release the benchmark data so that others could reanalyze or understand the results.¹⁰⁷

In contrast to the disagreement about Commerce's effort, Dr. Citro praised the Urban Institute's analysis of collected disparity studies. She explained that a meta-analysis discerns patterns through examination of a collection of studies. As such it gains strength from a vast amount of research and data, but at another level is no better than the quality of the individual studies it uses. Poorly conducted studies produce weak results. The Urban Institute screened studies on quality, using several criteria to discard the weaker ones and improve the reliability of results. She noted that the Urban Institute's study discarded about 40 percent of the collected studies and combined only those that met minimum standards. Dr. Citro characterized the approach as reasonable for obtaining valid, scientific results.¹⁰⁸ Citro added that after looking at the Urban Institute's work in different ways, the National Research Council's researchers concluded that no matter how they analyzed the data, women-owned businesses experienced a fairly large disparity in contracts.¹⁰⁹

The studies forming the Urban Institute's meta-analysis were carried out in the early 1990s in part using data from the 1980s. Such old data are not useful today, especially when the business world, particularly women-owned small businesses, has grown tremendously, Dr. Citro said. She advocated for new data to draw conclusions about contracting today.¹¹⁰

In concluding the briefing, discussion revolved around whether the prevailing competitive market allows or overcomes discrimination. Commissioner Thernstrom suggested that it

¹⁰⁶ Ibid., pp. 107–108, 110.

¹⁰⁷ Ibid., pp. 110–111.

¹⁰⁸ Ibid., pp. 78–79, 97.

¹⁰⁹ Ibid., pp. 98.

¹¹⁰ Ibid., pp. 79–80.

encourages businesses to hire the people or firms that will do the best job, countermanding discrimination.¹¹¹ Commissioner Yaki, on the other hand, asserted that the principles of business are not always market driven. The glass ceiling continues to bar many minorities from advancement in the competitive marketplace, he said.¹¹² However, neither Commissioner suggested that the nation is without discrimination.¹¹³ Chairman Reynolds thanked the panelists for their professionalism in presenting their different perspectives.¹¹⁴

¹¹¹ Ibid., p. 127.

¹¹² Ibid., pp. 127–128.

¹¹³ Ibid., pp. 128–129.

¹¹⁴ Ibid., p. 129.

PREPARED REMARKS OF GEORGE R. LA NOUE, PROFESSOR OF POLITICAL SCIENCE AND DIRECTOR OF THE PROJECT ON CIVIL RIGHTS AND PUBLIC CONTRACTS, UNIVERSITY OF MARYLAND BALTIMORE COUNTY

Executive Summary

Almost all public contracting programs using any form of racial classifications or preferences rely on disparity studies. Without them, it is exceedingly unlikely that a compelling interest to use race to influence contract awards could be demonstrated. The origin of this requirement is found in *City of Richmond v. Croson* where Justice O'Connor stated:

Where there is a *significant* statistical disparity between the number of *qualified* minority contractors *willing and able to perform a particular service* and the number of such contractors actually engaged by the locality or the locality's prime contractors, an inference of discriminatory exclusion *could arise*.¹

At the federal level, three documents (the 1996 Urban Institute report, the 1996 Department of Justice Appendix A and the 1998 Department of Commerce benchmark study) are all based in different ways on disparity studies methods. More than 180 disparity studies have been created by state and local governments to justify, create or maintain minority and women-owned business programs. Contracting disparity studies at all governmental levels are reviewed in this report.

In determining whether discrimination exists and, if so, what is its source and what a narrowly-tailored remedy would be, disparity studies have largely been failures. They have often been rejected by courts and have been criticized by scholars and objective government examinations.

While there are variations in the methodology and quality of disparity studies, almost all have some common flaws. In their statistical sections:

- (1) They fail to measure availability in the terms *Croson* requires of comparing "qualified" "willing and able" businesses performing "particular services";
- (2) They frequently are based on obsolete or incomplete data;
- (3) They report data in ways that exaggerate disparities;
- (4) They do not test to see if there are nondiscriminatory explanations for disparities;
- (5) They make findings of discrimination without ever identifying a single instance of discrimination or even a general source.

In their anecdotal sections:

¹ *City of Richmond v. J.A. Croson*, 488 U.S. 469, 509 (1989) (emphasis added).

- (1) They base conclusions on samples that are not gathered according to scientific methods;
- (2) They base conclusions on very small percentages of the surveyed universe.
- (3) They fail to verify any of the allegations they report;
- (4) All allegations are made anonymously, so governments receiving the reports do not have the information to judge their credibility or take action to solve the problems raised.

Despite the enormous investment of time and money in gathering anecdotal evidence in disparity studies, these efforts have made almost no difference in litigation. Judges generally have required that anecdotal testimony be given in court and subject to cross examination, if it is to have any influence in deciding the constitutionality of the program.

Despite these problems, millions of dollars are still being spent on studies which continue these flawed methodologies and the federal Justice Department (DOJ) is still relying on disparity studies whose data are now a decade old to defend race conscious procurement programs. The Commission on Civil Rights should make a major contribution to the public interest by suggesting standards that governments could use in commissioning and evaluating disparity studies.

In addition to performing disparity studies correctly in the first place, the authors should also seriously consider race neutral alternatives, if they make policy recommendations. This report concludes with some guidelines about what serious consideration should entail.

Introduction

Discrimination is a poison in the bloodstream of American life. Understanding the role of disparity studies which purport to examine discrimination in public contracting is important for several reasons. If done properly, disparity studies may be useful in highlighting the consequences of discrimination and providing the information to eliminate them. If the studies are done improperly, however, they may create claims of discrimination that are false or misleading. False claims of discrimination contribute to racial polarization and suppress interest in searching for race neutral programs that may create genuine new opportunities. Even, at their best, disparity studies can rarely identify the sources of discrimination with any precision and thus need to be supplemented with other data to create an appropriate public policy.

My background regarding this subject has been developed over several decades. From 1976 to 1982, I served as the chief trial expert for the United States Department of Labor and the United States Equal Employment Opportunity Commission regarding cases determining whether there was disparity in pay for women in universities. After the Supreme Court decided *City of Richmond v. Croson*, the landmark case in this area, in 1989, I created the Project on Civil Rights and Public Contracts to function as a library and data base on minority and women-owned business enterprise (MWBE) programs. Today, it includes about 20,000 pages of materials, including more than 160 disparity studies, the largest publicly accessible collection anywhere. I

have written seventeen law journal and other articles on disparity study related questions and wrote two editions of a guidebook for the National League of Cities on how to do disparity studies. I have worked with the cities of Albuquerque, West Palm Beach, and St. Petersburg, and multiple jurisdictions in Nashville, Tennessee and Portland Oregon as well as the states of California, Oregon, and Texas to develop or monitor disparity studies. I have been the plaintiff's expert in cases involving disparity studies in Philadelphia, Columbus, Chicago, Cincinnati, Denver, Dade County, Cook County, Atlanta Public Schools, and Jackson, MS, where MWBE programs have been declared unconstitutional. Currently, I am serving on the Maryland State Commission on Equal Pay.

This report will focus on contracting disparity studies at every level of government because the methodological procedures and problems often overlap. It will examine such federal reports as the 1996 Urban Institute report, the 1996 Justice Department's Appendix A, the 1998 Commerce Department's benchmark study, and the 2001 GAO report on evidence of discrimination in the highway construction industry, as well as the more than 150 state and local disparity studies I have examined.

The U.S. Commission on Civil Rights' interest in this subject comes at a critical time. The benchmark study's data is obsolete and its methodology flawed. Nevertheless, it is the only supporting statistical predicate to narrowly tailor federal race conscious contracting programs. The Department of Justice's 1996 Appendix A is still being introduced into cases by DOJ, but it was never based on any original research and the secondary research it relied on is now old and some of it has been found not to be credible.

Neither this Bush Administration nor recent Congresses have shown any interest in updating any disparity analysis. A recent attempt by the SBA to create a statistical basis for its 8(a) program has been heavily criticized by the National Academy of Sciences² and apparently sent back to the drawing board. Finally, a recent decision by the Ninth Circuit, *Western States Paving Co. v. Washington State Department of Transportation*, has held that recipients (highways, airports, transit systems) of federal transportation dollars must have local evidence of contracting discrimination against specific groups before, they may set race conscious goals.³ (The Bush Justice Department chose not to appeal this decision). That probably means a raft of new state and local disparity studies. In short, it is predictable that taxpayer investments in contracting disparity studies, controversy over the methods used in them, and their judicial review will go on for some time. Guidance from the Commission about acceptable methodologies and roles for disparity studies would be an important public service.

² National Research Council, *Analyzing Information on Women-Owned Small Businesses in Federal Contracting*, (Washington, DC: The National Academy Press, 2005), p. 60 (hereafter cited as NRC, *Women-Owned Small Businesses in Federal Contracting*).

³ 407 F.3d 983 (9th Cir. 2005).

State and Local Disparity Studies

Since there have been many more disparity studies completed by state and local government and much more judicial review of them and, since several federal reports depend on such studies, this report will begin by reviewing these studies. If the goal is to accurately and objectively evaluate the existence of contracting discrimination, the view is not encouraging. With few exceptions, courts have criticized the assumptions, data, and techniques of state and local disparity studies.⁴

Statistics

The validity of any disparity study is controlled by whether the creation of disparity ratios is based on appropriate statistics. In *Croson*, Justice O'Connor made a specific statement about the kind of data that should be considered and the method for comparing that data to determine a disparity. She said:

Where there is a significant statistical disparity between the number of qualified minority contractors willing and able to perform a particular service and the number of such contractors actually engaged by the locality or the locality's prime contractors, an inference of discriminatory exclusion could arise.⁵ (emphasis supplied)

The underlined parts of Justice O'Connor's test for statistical analysis in *Croson* identify several crucial requirements which must be satisfied if a public body desires to use statistical analysis to demonstrate the "extreme case" where some form of narrowly tailored remedy might be necessary to break down "patterns of deliberate exclusion."⁶ Without such specification, the probability of coming up with a false or misleading disparity ratio is very high. The key points are as follows:

⁴O'Donnell Constr. Co. v. District of Columbia, 963 F. 2d 420 (D.C. 1992); Contractors Ass'n of Eastern Pa. v. City of Philadelphia, 893 F. Supp. 438 (E.D. Pa 1995), aff'd 91 F.3d 586 (3rd Cir. 1996), cert. denied., 117 S. Ct. 953 (1997); Associated Gen. Contractors of Am. v. Columbus, 936 F. Supp. 1363 (S.D. Ohio 1996); Prior Tire v. Atlanta Public Schools, No. 1-95-CV-825-JEC (N. D. Ga. 1997); Engineering Contractors Ass'n of S. Fla. v. Metropolitan Dade County, 943 F. Supp. 1546 (S. D. Fla. 1996), aff'd 122 F. 3rd 895 (11th Cir. 1997); Phillips and Jordan v. Watts, 13 F. Supp. 2d 1308 (N. D. Fla. 1999); Webster v. Fulton Co., 51 F. Supp. 2d 1359 (N.D. Ga. 2000), aff'd 218 F.3d 1267 (11th Cir. 2000); West Tennessee Assoc. of Builders and Contractors v. Memphis City Schools, 64 F. Supp. 2d 714 (W. D. Tenn. 1999); Association for Fairness in Business, Inc., v. State of New Jersey, 82 F. Supp. 2d 353 (D. N.J. 2000); Associated Utilities Contractors v. Baltimore, 83 F. Supp. 2d 612 (D. Md. 2000); Concrete Works v. City and County of Denver, 86 F. Supp. 2d 1042 (D. Colo. 2000); Builders Ass'n of Greater Chicago v. Cook County, 123 F. Supp. 2d 1087 (N.D. Ill. 2000), aff'd 256 F.3d 642 (7th Cir. 2001); Associated General Contractors of Ohio v. Drabik, 50 F. Supp. 2d 741 (S.D. Ohio 1999), aff'd 214 F.3d 730 (6th Cir. 2000); Gill v. Metropolitan Dade County, No. 98-2300-CIV/Jordan (S.D. Fla. 2000); L. Tarango Trucking v. County of Contra Costa, 181 F. Supp. 2d 1017, 1029-31 (N.D. Ca. 2001).

⁵ City of Richmond v. J.A. Croson Co., 488 U.S. 469, 509 (1989) (emphasis added).

⁶ *Croson*, 488 U.S. at 509.

- A. In comparing participation rates, the universe of MWBE and non-MWBE contractors and subcontractors should be similarly “qualified,” and “willing and able.” Headcounts of businesses in a designated geographical area or on some list(s) are not sufficient.⁷
- B. The businesses being compared must be able to perform the same services. Companies compete for contracts only with others offering similar services. Government agencies contract for particular services. Consequently, comparisons that put companies offering distinctively different services in the same availability pool, or include businesses that do not offer services that a government purchases, are inconsistent with the *Croson* test.
- C. The data used must be recent enough to justify conclusions about the current market, if the jurisdiction is seeking to defend a current program.
- D. There must be “a significant statistical disparity.” This means more than just applying a statistical significance test to determine whether a result could have occurred by chance.⁸ Spurious statistical conclusions drawn on small or otherwise inadequate samples, or using inappropriate mathematical concepts, are not justifications for instituting governmental racial classifications. Nor are disparities that are small, or that appear only in certain measures, at certain times, sufficient⁹. To justify a race-based program, the statistics must provide evidence of “patterns of deliberate exclusion.”
- E. Statistics can provide only “an inference of discriminatory exclusion.” Without more information, they usually cannot identify the source of discrimination or provide a basis for a “narrowly tailored” remedy.¹⁰

⁷*Croson*, 488 U.S. at 509. Other *Croson* discussions of the need to measure qualifications can be found at 501 and 502. As the District Court said in *Contractors Association of Eastern Pennsylvania v. City of Philadelphia*:

“Qualified,” “willing” and “able” are the three pillars of the of the *Croson* test; a fortiori, a municipality may not enact race-based remedial measures unless it determines that qualified, willing and able minority contractors have been excluded from participating in public contracting. (893 F. Supp. 419, 432 (E. D. Pa. 1995))

See also, *Concrete Works v. City and County of Denver*, “‘Where special qualifications are necessary, the relevant statistical pool for purposes of demonstrating discriminatory exclusion must be the number of minorities *qualified to undertake the particular task*.’” (emphasis in original) 36 F.3d 1513, 1528, (10th Cir. 1994). Also, *Associated General Contractors v. Drabik* faulting a statistical comparison for not taking into account the percentage of minority- owned businesses that “were qualified, willing and able to perform state construction contracts.” 214 F.3d at 736. The need to compare qualifications in measuring disparities is well-established. See, for example in employment, *Hammon v. Barry*, 813 F.2d 412, 427, n.31 (D.C. Cir. 1987), *Long v. City of Saginaw*, 911 F.2d 1192, 1199 (6th Cir. 1992), and *Aiken v. City of Memphis*, 37 F.3d 1155, 1165 (6th Cir. 1994 en banc); and college scholarships, *Podberesky v. Kirwan*, 38 F.3d 147, 156 (4th Cir. 1994).

⁸*Associated Gen. Contractors of Am.*, 936 F. Supp. at 1363.

⁹*Engineering Contractors Ass’n of South Fla.*, 943 F. Supp. at 1560–1577.

¹⁰As Judge Bechtel said in *Contractors Association of Eastern Pennsylvania v. City of Philadelphia*:

To establish racial discrimination, the statistical disparity must in some way be linked to additional evidence. Otherwise, an inference of discriminatory exclusion simply cannot be conclusive. See *O'Donnell Construction Co. v. District of Columbia*, 963 F.2d 420, 426 (D.C. Cir., 1992) (“The idea that discrimination

Legal principles and good social science require that a disparity study articulate alternative hypotheses to the discrimination hypothesis and test them as possible explanations for the apparent disparity before reaching a conclusion about the cause of the disparity.

The key to every disparity study is the measure of availability.¹¹ If the measure of the relative availability of MWBEs is too high, the resulting disparity is an artifact of availability inflation, not discrimination. Since almost everywhere MWBEs are smaller, newer firms and in most programs are required to graduate when they are successful enough, disparity ratios based on headcounts will overstate MWBE availability. Similarly, measures that do not control for the qualifications, size or experience of firms will overstate MWBE availability. As the General Accounting Office (now the Government Accountability Office) (GAO) found in its review of disparity studies, census data “cannot adequately indicate whether a firm is truly available, that is, whether it has the qualifications, willingness, or ability to complete contracts. However, in using Census Bureau data, the studies depicted all operational firms as available for contracting.”¹² Despite this criticism, GAO found that 48.8 percent of all respondents used census data as one means of setting goals.¹³

GAO also found that directories and other listings were not a valid source of availability because they “do not contain information on firms’ qualifications, willingness or abilities.” This could result in an overstatement of how many firms are available for transportation contracting.”¹⁴

caused the low percentage [of minority participation]” is nothing more than an hypothesis....) (893 F. Supp. at 429).

¹¹ I have explored these issues at some length in George La Noue, “Who Counts: Measuring the Availability of Minority Businesses after Croson” 21 *Harvard Journal of Law and Public Policy*, Fall 1998, pp. 101-141.

¹² General Accounting Office (now the Government Accountability Office) (GAO), *Disadvantaged Business Enterprises: Critical Information is needed to Understand Program Impact*, GAO-01-586, June 2001, p. 30 (hereafter cited as GAO Report).

¹³ *Ibid.*, p. 50. See also, *Phillips & Jordan v. Watts* where the court noted the problem of overinclusiveness in using census data by stating: “The court is also unconvinced that it was appropriate to assume—as MGT did—that every firm in selected SIC codes was qualified and /or willing and able to bid on an FDOT road maintenance contract.” 13 F. Supp. 2d 1308, 1315 (N.D. Fla. 1998). This concern about overinclusiveness was echoed in *Associated General Contractors v. Drabik*, where the Court found that census data “probably overstates the percentage of MBEs qualified to provide some of the services covered by the Act. For example, firms seeking prime contracts must be able to provide performance bonds.” 50 F. Supp. 2d 741, 747 (S.D. Ohio 1999). Social scientists have come to similar conclusions:

However, there are significant problems and limitations with census data relative to the *Croson* guidelines that the availability of women and minority owned firms should reflect the number of qualified, willing and able firms. Given the number and difficulty of the required adjustments to the Census data, it is unlikely that these data will provide availability estimates that are accurate enough to allow for valid statistical tests of an inference of discriminatory exclusion. Stephen E. Celec, Dan Voich, Jr., E. Joe Nosari, and Melvin T. Stith, Sr., “Measuring Disparity in Government Procurement: Problems with Using Census Data in Estimating Availability,” *Public Administration Review*, Volume 60, No. 2, March 2002, p. 134.

¹⁴ GAO Report, p. 31.

Many directories contained obsolete information. Nevertheless, 80.5 percent of the respondents used directories as one means of setting goals.

The problem is compounded if the utilization measure is dollars and ignores contract awards. As GAO has stated:

Because MBE/WBEs are more likely to be awarded subcontracts than prime contracts, MBEs/WBEs may appear to be underutilized when the focus remains on prime contractor data. Furthermore, although some studies did include calculations based on the number of contracts, all but two based their determination of disparities on only the dollar amounts of the contracts. Because MBEs/WBEs tend to be smaller than non-MBEs/WBEs, they often are unable to perform on larger contracts. Therefore, it would appear that they were awarded a disproportionately smaller amount of contract dollars.¹⁵

Most disparity studies have incorrectly measured both availability and utilization, thus creating false disparities.

Some methods of calculating disparity ratios also create false disparities. Properly calculated, disparity ratios should compare the prime awards for each group to the total prime awards in a given chronological period. Similarly the subcontract awards for each group should be compared to the total subcontract awards. If that is not done, the results of the two different awarding processes will be conflated and thus the potential two different sources of discrimination will be masked. It is possible to find a disparity stemming from the selection process for either prime or subcontractors or both, but designing an appropriate remedy will depend on locating the source of the problem. Some disparity studies compare the percentage of MWBE subcontractor dollars awarded to the total dollars (prime and subcontractor) awarded. Since subcontracts are always a fraction of the total contract, disparity ratios comparing subcontractor awards to total dollars will always show a disparity. These studies mask that result by not calculating disparity ratios for non-MWBEs. Thus, jurisdictions may be led to believe that there is a disparity indicating discrimination against MWBEs, when in fact the data properly calculated may even show overutilization. Recently, the expert for the North Carolina Department of Transportation in a case against the state's MWBE highway program testified in deposition that he had told his disparity company of this error, but he did not know if the company had informed the governments that commissioned the studies with this particular flaw of this problem.¹⁶ There are apparently no recall procedures in the disparity study industry, but we have found these erroneous disparity ratios in 17 other studies.¹⁷

¹⁵ Ibid., p. 32.

¹⁶ Deposition of Dr. J. Vincent Eagen, in *Rowe v. North Carolina Department of Transportation*, CA No.5:03-CV-278-BO (3), August 16, 2005. Dr. Eagen testified, "I don't think you should conduct disparity ratios on that basis." (*Id.* at p. 115).

¹⁷ Albany (GA) (October 1995), Baltimore (July 2000), Dougherty County (GA) (1995), Dougherty County School Board (GA) (1995), Dougherty County Water, Gas, and Light Commission (GA) (1995), Leon County (FL) (1995),

In short, there are common problems in disparity studies in measuring availability, in reporting utilization, and in calculating disparities that make them unreliable instruments for determining the existence of contracting discrimination.

Anecdotes

Anecdotal data are a part of almost every disparity study. If gathered and analyzed properly, such information could supplement statistical findings and assist in discovering discriminatory or race neutral barriers that inhibit participation. Unfortunately, despite millions of dollars of tax funds expenditures and a decade of experience, the anecdotal sections of disparity studies are rarely credible whether the mechanism for gathering the anecdotes is surveys, focus groups or interviews.¹⁸

There are several reasons:

1. The number of participants is usually very low and not representative.
2. All anecdotal information is reported anonymously and very little is ever verified.
3. The information is reported in such a way to confirm the authors pre-existing assumptions about discrimination rather than to objectively test research hypotheses.

Unrepresentative samples

In an analysis of disparity studies we have recently completed at the Project on Civil Rights and Public Contracts, we examined anecdotal sections in 154 such studies. The median rate of responses for these surveys was 19 percent. Some were as low as 2 or 3 percent. That figure is far too low to meet social science standards and, where the legal standard is strict scrutiny, it is not persuasive.¹⁹ According to these guidelines, response rates of 90 percent or more are reliable and generally can be treated as random samples of the overall population. Response rates between 75 and 90 percent usually yield reliable results, but researchers should conduct some check on the representativeness of the sample. Potential bias should receive greater scrutiny when the response rates fall below 75 percent. If the response rate drops below 50 percent, the survey should be regarded with significant caution as a basis for precise quantitative statements

Little Rock (1996), Miami Dade Community College (1995), Pam Beach County School District (1991 and 1995), Pittsburgh Public Schools (1996), Sacramento (1992), Sacramento County (1992), Sacramento Regional Transit (1992), Sacramento Housing and Redevelopment Authority (1992), South Carolina Department of Transportation (1995), and Tallahassee (FL) (1990).

¹⁸ Jeffrey Hanson, "Hanging by Yarns? Deficiencies in Anecdotal Evidence Threaten the Survival of Race Based Preference Programs for Public Contractors, 88 *Cornell Law Review* 1433 (July 2003).

¹⁹ When a very substantial non response rate exists, a crucial issue exists about whether the responders are representative of the whole universe or some subset with distinctive characteristics. According to a Federal Judicial Center publication, the former U.S. office of Statistical Standards has suggested a formula for quantifying a tolerable level of nonresponse in surveys.

about the population from which the sample was drawn.”²⁰ Moreover, proof of discrimination is required for each group preferred in an MWBE program because whether based on statistics or anecdotes the results for different groups vary widely. But the studies rarely receive proportionate responses from groups and for smaller groups there may be only a handful of responses.

Anonymous unverified responses

Almost uniformly, disparity studies report their results in aggregate data or in snippets taken from focus group or interview transcripts. The respondents have been promised anonymity, so the governments that have paid for these studies have no way of determining who is making the allegations or evaluating their validity. This is a troublesome methodology for creating any public policy.

Almost never do disparity study researchers attempt to verify the anecdotal accusations. While some “he said, she said” allegations could not be verified without enormous resources; other anecdotes claiming “I was low bidder, but did not get the contract,” can be checked. Knowing whether the anecdotes are true is absolutely essential to determining their role as a basis for supporting an MWBE program and for designing an effective remedy. Yet during litigation, when study authors and relevant government officials are asked under oath whether they know if *any* of the allegations they have relied on are true, they state they do not know.

Using tax dollars to gather anonymous allegations of mistreatment by persons of one race or gender against another, when the outcome is intended to provide contracting benefits to persons making the allegations, is a peculiar governmental activity. Anonymous allegations can not be examined or rebutted.

Political Scientist Mitchell Rice, writing in the *Public Administration Review*, states that elemental fairness requires:

Anecdotal evidence, interviews and affidavits must be from reliable and trustworthy sources and should include counter explanations and rebuttals from sources accused of bias. In other words, the gathering of evidence utilizing these approaches must be fair and deliberative.²¹

One reason for being skeptical of the role of anecdotes in disparity studies is that the authors exercise complete control over whom to interview and what to report from those anonymous interviews and may use their discretion to craft a partial or even biased view of discrimination. For example, the San Francisco United School District study states:

²⁰ As quoted in Molly Treadway Johnson, Carol Krafka, and Joe S. Cecil, “Expert Testimony in Federal Civil Trials Federal,” Federal Judicial Center, 2000, p. 245.

²¹ “Governmental Set-Asides, Minority Business Enterprises and the Supreme Court,” *Public Administration Review*, March/April, 1991, p. 485.

Since the main goal in conducting a narrative/ historical study is to gather anecdotal information on the experience of discrimination, the most appropriate methodology is a qualitative survey of a purposive sample of MBE/WBEs.... No doubt, many MBE/WBEs have had a positive experience with SFUSD but those are outside the parameters of the study.²²

When the underlying anecdotal data have been turned over to plaintiffs in litigation, it has become clear that the anecdotal sections in disparity studies often have not described the data in a balanced manner.

Affirming pre-existing assumptions

Every social scientist knows that the way questions are worded and framed can influence the results. Disparity studies often attempt to increase their dismal response rates by informing potential respondents that the studies are necessary to preserve the MWBE program. But such cues can influence who responds and what they say, since preservation of the program depends on a finding of discrimination.

While the results of surveys are often reported quantitatively, though sometimes without reporting the number of valid responses to particular questions, there are rarely any announced standards for interpreting the results. If 10 or 20 percent of the respondents allege they have suffered from discrimination of a particular type, does that evidence meet the standard *Croson* set for using the “extreme measure” of a racial preference to break up a pattern of racial exclusion or does that result suggest no such pattern exists?

Furthermore, when interviews or focus groups are the basis for generalizations about discrimination, there are no announced standards for interpreting when a credible amount of evidence has been received. If two people in focus groups say they have had a problem, the study is likely to conclude that it is a characteristic experience, even if the rest of the group has had no such experience.

For all these reasons, no race conscious program has been upheld on the basis of anecdotes. In *Coral Construction v. King County*, the Court, after noting the fifty-seven affidavits alleging discrimination by MWBE owner concluded, “While anecdotal evidence may suffice to prove individual claims of discrimination, rarely, if ever, can such evidence show a systemic pattern of discrimination necessary for the adoption of an affirmative action plan. ... [T]he MBE program cannot stand without a proper statistical foundation.”²³

²² San Francisco United School District study, p. IV-2.

²³ 941 F. 2d 910, 919, (9th Cir. 1991).

Since anonymous unverified anecdotes are at best hearsay, courts have been reluctant to rely on them for the truth of the matter asserted.²⁴ The trend is to bring anecdotal witnesses into the trial where they can be cross-examined and the court can judge the anecdotes' credibility and relevance.²⁵

The Court in *Engineering Contractors Association of South Florida v. Metropolitan Dade County* concluded:

Plaintiffs respond with several points the Court believes to be valid concerning the reliability of this anecdotal evidence. First, whether discrimination has occurred is often complex and requires a knowledge of the perspectives of both parties involved in an incident as well as knowledge about how comparably placed persons of other races, ethnicities, and genders have been treated. Persons providing anecdotes rarely have such information. Attributing an incident to discrimination when the practice is just aggressive business behavior, barriers faced by all new or small businesses, or bad communication is always a possibility.

Second, social scientists are frequently concerned about the problem of "interviewer bias" or "response bias" in any interviewing or survey situation. When the respondent is made aware of the political purpose of questions or when questions are worded in such a way as to suggest the answers the inquirer wishes to receive, "interviewer bias" can occur. If a sample is not carefully constructed, the persons providing the anecdotes may reflect a "response bias" because the persons most likely to respond are those who feel the most strongly about a problem, even though they may not be representative of the larger group.

Third, individuals who have a vested interest in preserving a benefit or entitlement may be motivated to view events in a manner that justifies the entitlement. Consequently, it is important that both sides are heard and that there are other measures of the accuracy of the claims. Attempts to investigate and verify the anecdotal evidence should be made.²⁶

Consequently, the Court found:

Without corroboration, the Court cannot distinguish between allegations that in fact represent an objective assessment of the situation, and those that are fraught with heartfelt, but erroneous, interpretations of events and circumstances. The costs associated with the

²⁴ Associated Gen. Contractors of Am. v. Columbus, 936 F. Supp. 1363, 1425–1429 (S. D. Ohio 1996) and Engineering Contractors Association of South Fla. v. Metropolitan Dade County, 943 F. Supp. 1546, 1577–1580 (S.D. Fla. 1996).

²⁵ Concrete Works v. City and County of Denver, 86 F. Supp. 1042, 1072 (D. Colo. 2000), Builders Ass'n of Greater Chicago v. Cook County, 123 F. Supp. 2d 1087 (N. D. Ill. 2000). In *Western States Paving Co., v. Washington State Department of Transportation*, the Ninth Circuit noted that WSDOT had three claims of discrimination filed. (407 F.3d 983, 1001 (9th Cir. 2005)) Two had been investigated and found unwarranted and the third was pending. This is similar to the pattern GAO found in its survey of state and transit authorities. Most (81 percent) had no complaints of discrimination in FY 1999 and 2000 and of the 29 complaints that had been investigated only four resulted in findings of discrimination. GAO Report, p. 33.

²⁶ 943 F. Supp. 1546, 1579.

imposition of race, ethnicity, and gender preferences are simply too high to sustain a patently discriminatory program on such weak evidence.²⁷

Some disparity study companies have recognized the problem. MGT of America, Inc., has written:

Hence, in today's environment, discriminatory practices sometimes exist in ways that are unrecognized by the institutions and individuals who actively continue to perpetuate these practices. They can even remain unrecognized by those who are the victims and those who are charged with the responsibility of eliminating discriminatory practices. By the same token, firms and individuals who lose contracts no doubt sometimes believe they were discriminated against even when no discrimination exists.²⁸

Given those assumptions, the research problem is obviously to distinguish between perception and reality, but disparity methodology does not permit that to happen. Confronted with the results of an MGT survey, the court in *Phillips & Jordan v. Watts* responded:

Individuals responding to FDOT's telephone survey have described their *perceptions* about barriers to FDOT's bidding procedures. But FDOT has presented no *evidence* to establish who, if anyone, in fact engaged in discriminatory acts against Black and Hispanic businesses. The record at best establishes nothing more than some ill-defined wrong caused by some unidentified wrongdoers; and under City of Richmond that is not enough! (emphasis in the original)²⁹

In short, the collection of anecdotes in disparity studies has failed to meet social science research standards and has not been helpful in documenting actual instances of discrimination or devising remedies. If sued, the anecdotal evidence courts consider will come from witnesses subject to cross examination in the courtroom.³⁰

Federal studies

Given the methodological and legal difficulties state and local disparity studies have encountered, it is surprising that the federal government has sought to rely on them. Two documents commissioned during the Clinton administration's "mend don't end" affirmative action policy illustrate the problems.

²⁷ *Id.* at 1546, 1584. *See also*, *Phillips & Jordan, Inc. v. Watts*, 13 F. Supp. 2d 1308, 1314 (N.D. Fla., 1998).

²⁸ MGT of America, Inc., 1993 North Carolina Department of Transportation Study, p. 7.

²⁹ *Phillips & Jordan, Inc.*, 13 F. Supp. 2d at 1314.

³⁰ *See for example*, *L. Tarango Trucking v. County of Contra Costa*, 181 F. Supp. 2d 1017, 1029–31 (N.D. Ca. 2001) (The plaintiffs presented several class members who claimed that they had personally suffered from discrimination because of race or sex in attempting to obtain contracts from the County. The Court finds that none of those witnesses demonstrated that the County discriminated against them because of their race or sex." (at 1029)), *Denver, Dade, Cook, and Columbus*.

Urban Institute Meta-study

In 1996, the Justice Department hired the Urban Institute at a cost of \$175,000 (supplemented with Rockefeller Foundation funding) to complete a meta-study (a study of studies) to “provide information bearing on the need for programs that assist minority-owned firms—including affirmative action in procurement.” When the Institute completed its study,³¹ it concluded, that these business received only 57 percent of their expected dollars. To reach this conclusion, the Urban Institute conducted no original research, but relied exclusively on existing state and local disparity studies, even though its authors knew many of those studies used techniques that exaggerated disparities. A meta-study relying on flawed previous work will, of course, reach inaccurate results.

In a private precursor report to the Justice Department written on January 18, 1996 or eight months before the Institute’s public report was released in October 1996, the Institute authors were much more critical of some disparity studies. The report, titled “Evaluation of Disparity Study Methodology,” rated the quantitative sections of 22 disparity studies that the Institute received from the Department as “strong,” “fully satisfactory,” “acceptable,” or “weak” and made comments about their methodologies and results.³² Ultimately the Institute rated 6 of the 22 studies as “weak” because they contained various flaws.

For example, the Institute said about the study, “City and County Public Schools, Milwaukee, Wisconsin,” that its availability data “may not reflect firms that are “ready, willing and able”; that “it was unclear whether the overall disparity ratios include prime and subcontractor data”; that “no separate analysis of prime and subcontracts is presented”; and “Additional statistical analysis is presented, but it is based on surveys with very low response rates and very small samples.”³³ This is a standard social science critique and the study rating of “weak” seems fully warranted. One would expect that rating and analysis would cause the Milwaukee study to have been excluded from the Institute’s later meta-analysis of disparity ratios. In fact, 16 disparity ratios from the “weak” Milwaukee study are included in the Institute meta-analysis.

On the other hand, the Louisiana study completed by John Lunn, then Professor of Economics at Louisiana State University, and Hugh Perry, professor of Political Science at Southern University, was rated “fully satisfactory,” but was not included in the Institute meta-study report. The Lunn/Perry study is one of the very few studies to use multiple regression analysis about which the Institute correctly says:

³¹ “María Enchautequi, Michael Fix, Pamela Loprest, Sarah von der Lippe, and Douglas Wissoker, *Do Minority-Owned Businesses Get a Fair Share of Government Contracts?* (Washington, D.C.: The Urban Institute, 1997) (hereafter cited as Meta-study).

³² The Urban Institute, *Evaluation of Disparity Study Methodology*, Jan. 18, 1996, p. 2, (hereafter cited as Analytical study).

³³ *Ibid.*, p. 29.

This method allows the authors to measure disparity after controlling for other factors that reflects a firm's ability to perform the work. These factors include size (revenues and employees) and experience of firms, whether it owned its own equipment, has necessary licenses, interest in public works jobs, bonding, line of credit, past bidding experience, education of the owner, percent of work conducted as a prime contractor. This is one of the best measures of firms ready, willing and able to perform government work.³⁴

The Lunn/Perry study is the only disparity study whose results have been published in a refereed academic journal.³⁵ This study generally did not find minority-owned business enterprises (MBEs) underutilized. One would have thought that when this much more rigorous analytical technique than the typical disparity study did not find a disparity that might have caused some doubt about the conclusions of far less sophisticated studies and that would have created some discussion of the importance of controlling for relevant variables. But the Lunn/ Perry results and significance of multiple regression are not discussed in the meta-study. Of course, neither subject is mentioned in the DOJ Appendix.

The key question in any disparity study is how availability is measured in the first place. The Institute meta-study report concluded that:

... the methods used by the disparity study authors in their quantitative analysis are remarkably consistent. Although there are differences across the studies in the sources of data used and in the definition of available firms, each study reports on the same outcome (i.e., the percentage of government contract dollars awarded to minority owned firms compared to the percentage of all available firms that are minority-owned)³⁶

Finding that studies using different data sources and different definitions of availability are remarkably consistent seems like a strange conclusion, particularly since the Urban Institute recognized that many studies used flawed availability measures. In their earlier evaluation of disparity studies, the meta-study the authors stated:

A study that uses a measure of availability that includes only firms that are ready, willing and able to perform government contract work is stronger than a study that includes all firms without trying to control for whether they are available or have the requisite ability. *This is because including all M/WBEs may overstate the true availability of these firms, and bias results towards finding a disparity even when there is none.* ... Hence, studies that employ only data from the [census] survey of Minority or Women Owned Business Enterprises (SM/SWOBEs) which enumerates minority firms to measure availability are considered comparatively weak. (emphasis added)³⁷

³⁴ Ibid., p. 15.

³⁵ "Justifying Affirmative Action: Highway Construction in Louisiana," *Industrial and Labor Relations Review*, 49(3): 464–479.

³⁶ Meta-study, p. xiii.

³⁷ Analytical study, p. 3–4.

Based on this evaluation, one would have thought that the authors would have excluded all studies that would “bias results towards finding a disparity even when there is none” or, at least, would explain the issue in the published meta-study. Neither happened.

For example, in evaluating the New York City disparity study completed by the National Economic Research Associates, the institute authors state correctly: “The study measures availability data from [census] SM/SWOBE and County Business Patterns. However, this measure may not reflect firms that are ready, willing and able to perform government contract work.”³⁸ Nevertheless, the Institute authors gave the study a “fully satisfactory” rating and included its data in the meta-analysis. Altogether about half of the studies on which the Institute based its 57 cents on the dollar conclusion are based on census data. Overall some 600 disparity ratios for construction alone are based solely on census data which the Urban Institute had found would “overstate the true availability of these firms, and bias results towards finding a disparity even when there is none.”

Ultimately in determining which studies to include in the meta-study, the team decided to use them all regardless of how they measured availability, as long as they had disparity ratios based on at least eighty contracts. The study team treated all measures of availability as equally valid. Its earlier concern that census data or other measures that were not limited to firms that were ready, willing, and able might bias the findings disappeared.

In connection with the case *Rothe v. Department of Defense*,³⁹ the lead author of both the Urban Institute reports was deposed and asked about the methodologies used.

Dr. Pam Loprest testified:

We used them all [disparity studies] in order to create the disparity measure we did.... How you interpret the final disparity ratio does depend in some sense on what the availability measure used is. So they're all valid to make the conclusion of a comparison of the share, the utilization to a share of the availability.⁴⁰

When asked whether studies were excluded which had no measure of firm qualifications in determining availability, Dr. Loprest noted there was a “huge debate” around that issue, but in the end, “We had to use the data as it was put forward.”⁴¹ Later she testified the Institute team had no definition of qualifications or capability.⁴² Dr. Loprest could not remember any studies in

³⁸ Ibid., p. 19.

³⁹ 262 F.3d 1306 (Fed. Cir. 2001).

⁴⁰ Loprest, tr. 99–100.

⁴¹ Ibid., tr.101.

⁴² Ibid., tr.103–104.

the Institute sample that considered whether a firm had bonding, licensing, or insurance to measure availability.⁴³

Another problem is that many of the disparity ratios in the studies the Institute analyzed are not statistically significant. The Institute authors knew, "A difference in these percentages can represent disparity in government contracting or it may be due to chance."⁴⁴ A statistically insignificant disparity could be caused because either the availability or utilization numbers were very small or the disparities were trivial. The Institute authors put the issue this way: "A measure of statistical significance provides information on the range of findings for which we cannot conclude a disparity exists."⁴⁵ The authors did criticize some studies, such as the Memphis/Shelby County study, for basing disparity ratios on a small number of contracts which they noted made the "figures on disparity very volatile and unreliable."⁴⁶ Again one would have thought the authors would have decided to exclude disparity ratios that were not statistically significant from their conclusions. At least, one would not expect them to include statistically insignificant results from Memphis/Shelby County in their meta-analysis. To the contrary, the authors included 66 statistically insignificant disparity ratios in construction compared to nine statistically significant disparity ratios from this study. Of the nine statistically significant disparity ratios, seven showed MWBEs were overutilized.⁴⁷ Overall in the Institute's meta-study about half of the disparity ratios for all minorities are not statistically significant. Since throughout the Institute report the authors combined the majority of results which were statistically insignificant with statistically significant results, the overall conclusion of 0.57 disparity is enormously unreliable.⁴⁸

Another methodological flaw in the meta-study is that, although it reports results as a measure of dollars, i.e. MBEs receive 57 cents of expected dollars, some of the disparity ratios the Institute used were based on shares of the number of contracts. Obviously it is improper to express a conclusion in dollar terms, if part of the data is share of contracts. Dr. Loprest testified in response to a question on that point:

Q. Isn't that kind of like comparing apples and oranges to treat data that considers the share of contracts awarded in terms of dollars versus the share of contracts awarded in shares of numbers and to somehow combine that in a figure that represents the amount of dollar award?

A. Well there are a lot of differences in every number. When you're doing a meta-analysis and when looking at these, there are differences in each of these numbers. They're in different

⁴³ Ibid., tr. 102–103.

⁴⁴ Analytical report, p. 5.

⁴⁵ Ibid.

⁴⁶ Ibid., p. 18.

⁴⁷ Meta-study, table A. 3a.

⁴⁸ Loprest tr.109.

places. There's a lot of differences in the numbers. And so, yes, you're combining a lot of things that might be called apples and oranges in that you're combining it to come [up] with an answer. And the ratio of utilization has the same thing on the top and the bottom. So then you get the percentage and then those are averaged together. I'm not sure if I am being clear.⁴⁹

The Institute's meta-study also averaged together the disparity ratios based on 100 contracts with those based on 1,000 contracts. Thus a disparity ratio based on the 2,302 construction contracts in Phoenix, AZ, is averaged into the overall total in the same way as the disparity ratio derived from the 95 construction contracts awarded by Pima County, AZ. That technique is not consistent with an outcome expressed as expected dollars received on a national basis.⁵⁰

What the result would be if the meta-study had not committed these technical errors in dealing with the disparity study data is unknown. The overwhelming difficulty with the meta-study, however, is a problem inherent in the individual disparity studies. Despite the meta-study's assurance that the "disparity study authors in their quantitative analysis are remarkably consistent," there are major and irreconcilable differences in how the studies measured the critical variable of availability. While there is substantial similarity among various studies done by the *same* consultants, there are great methodological differences among the studies completed by *different* firms. If there are differences in the availability portion of the disparity equation, then the methods are not consistent and the results being compared are also apples and oranges.

To conclude there are eight major problems in the Urban Institute's final report.

1. The meta-study does not require any control for the size of MWBE and non-MWBE firms in the disparity studies it relied on. Almost everywhere, the vast majority of public contracting dollars are awarded in a few very large contracts. The meta-study concedes that MWBEs are smaller and newer firms than non-MWBEs.⁵¹ Since no one would reasonably expect that small, new firms would receive as many dollars from government contracting as larger well-established firms, the Institute's premise for comparison in terms of "expected" dollars is false.

As the Eleventh Circuit declared:

Because they are bigger, bigger firms have a bigger chance to win bigger contracts. It follows that, all other factors being equal, and in a perfectly non-discriminatory market, one would expect the bigger (on average) non-MWBE firms to get a disproportionately higher percentage of total construction dollars awarded than smaller MWBE firms.⁵²

⁴⁹ Loprest, tr.78–79.

⁵⁰ Ibid., 86 and 89.

⁵¹ Meta-study, p. ii.

⁵² *Engineering Contractors Ass'n of South Fla. v. Metropolitan Dade County*, 122 F.3d 895, 917 (11th Cir.1997) This has been a consistent theme in judicial rulings that disparity ratios that do not control for relevant variables,

Nevertheless, the meta-study simply compared the share of MWBEs in a market with share of dollars received from government contracts without controlling for any other variables.

2. In measuring the availability of firms, the meta-study depends on headcounts of MWBEs in the disparity studies without considering the qualifications of these firms or whether they have ever bid on public contracts. Furthermore, the Institute authors knew that, “including all M/WBEs may overstate the true availability of these firms, and bias results towards finding a disparity even when there is none.”⁵³

3. In comparing minority firms to nonminority owned firms or “majority-owned” firms as the meta-study sometimes describes the latter category, the meta-study combines the receipts of firms actually owned by white men with large publicly-traded stockholder corporations which have no race or gender.

4. The meta-study is written in a legal vacuum. The authors used no legal standards in evaluating the methodologies or conclusions in the disparity studies in their sample. None of the disparity studies on which the meta-report relied has been upheld by the courts as a basis for race conscious remedies, while several have been found to be specifically invalid, and many more have been based on techniques or data sources criticized by the courts.

5. To create its finding that MBEs receive only 57 percent of expected dollars, the Institute authors in the meta-study averaged disparity results from the various local disparity studies they included in their sample. In doing so they combined statistically insignificant results with those that were statistically significant. They also combined disparities based on contracts awarded with disparities based on dollars awarded, even though the outcome was expressed in

particularly size of firms do not prove discrimination. *See also* Michigan Roadbuilders Ass’n v. Millikin, 834 F. 2d 583, 593 (6th Cir. 1987) (holding there were relatively few MBEs and those that do are generally small in size and have difficulty in competing for state contracts as a result of their size. The evidence does not prove that the State of Michigan invidiously discriminated against racial and ethnic minorities in awarding state contracts) and O’Donnell Construction Co. v. District of Columbia, 963 F.2d 420, 426 (D.C. Cir. 1992) (holding that the small proportion of District of Columbia public contracts awarded minority-owned firms did not establish discrimination because “minority firms may have not bid on ... construction contracts because they were generally small companies incapable of taking on large projects; or because they were fully occupied on other projects). Associated General Contractors of Ohio v. Drabik, 214 F.3d 730, 736 (6th Cir. 2000) (“If [minority-owned firms] comprise 10% of the total number of contracting firms in the state, but get only 3% of the dollar value of certain contracts, that does not alone show discrimination or even disparity. It does not account for the relative size of firms, either in terms of their ability to do work or in terms of the number of tasks they have the resources to complete,”) and Western States Paving Co., v. Washington State Department of Transportation, 407 F.3d 983, 1001 (9th Cir. 2005) that said comparisons of the proportion of DBE firms to the percentage of contract dollars awarded DBE firms is “oversimplified statistical evidence” “entitled to little weight, however, because it does not account for factors that affect the relative capacity of DBEs to undertake contracting work. ... DBE firms may be smaller and less experienced than non-DBE firms (especially if they are new businesses started by recent immigrants) or they are concentrated in certain areas of the state, rendering them unavailable for a disproportionate amount of work. *Id.* at 1001.

⁵³ Analytical report, pp. 3–4.

dollar terms. Courts have made clear that disparities should be calculated separately for both contracts and dollars.

6. Read carefully, rather than demonstrating a single national pattern of disparities for MBE utilization, the meta-study shows many different outcomes of underutilization, overutilization, and appropriate utilization in different industries, different regions of the country, and for different groups. The average utilization figure (57 percent) the meta-study calculates, covering many different studies, different time frames, and different types of contracts, is meaningless when there is no control for the qualifications, sizes or specializations of MBEs and non-MBEs.

7. The major disparities the study found were in the low bid prime contracting process. At the subcontracting level where there is more discretion, the meta-study found MWBEs already receive 95 percent of their “expected” construction subcontracting dollars.⁵⁴ If that number is accurate, it reflects an existing parity that does not support the assumption that prime contractors have discriminated against MBE subcontractors. Moreover, the meta-study acknowledges, “measures based on dollars actually paid out often fail to identify dollars going to minority-owned subcontractors because government payments go to prime contractors and no record is kept [by the government] of payments to subcontractors.”⁵⁵

8. The meta-study’s authors concede that they cannot determine whether the disparities they found were caused by discrimination or other factors.⁵⁶

Appendix A

When the Justice Department created a response to *Adarand* in 1996 titled, “Proposed Reforms to Affirmative Action in Federal Procurement,”⁵⁷ it attached a 23 page Appendix A titled, “The Compelling Interest for Affirmative Action in Federal Procurement: A Preliminary Survey.” This Appendix, designed to support the continuance of racial preferences in federal contracting, was developed principally by a paralegal at DOJ, who later went to law school.⁵⁸ It involved no original research and no examination of original data. It is simply a pastiche drawn from earlier Congressional reports, articles, and disparity studies. Every assertion about discrimination is accepted as valid and relevant as a support to then current federal contracting preferences. Since 1996, DOJ has done no up-date of the Appendix and at no time has Congress ever adopted it. Nevertheless, it is still introduced in court cases and some judges have found it persuasive,

⁵⁴ Meta-study, figure 3.

⁵⁵ Ibid., p. 15.

⁵⁶ Ibid., p. 63.

⁵⁷ 61 Fed. Reg. 26,042 (May 23, 1996).

⁵⁸ The background of the production of Appendix A based on deposition statements can be found in Roger Clegg and John Sullivan, “No Compelling Interest,” *NRO Online*, May 24, 2001.

particularly because it created an almost impossible problem for plaintiffs trying to critique it, given its mass of materials and the fact that DOJ never had possession of any underlying data.

In addition to relying on the Urban Institute conclusions,⁵⁹ Appendix A cites 29 disparity studies completed between 1989 and 1994. The statistical data in these studies are at least a decade old and some go back into the Seventies.⁶⁰ Most of Appendix A's references are to unverified and anonymous allegations often made by a single person in various studies. In most instances, the jurisdictions involved have regarded these studies as obsolete and/or flawed and have replaced them with newer studies that have found different results. No disparity study cited in Appendix A is evaluated or critiqued. Nor are the judicial criticisms of disparity studies which had begun to appear by 1996 mentioned.⁶¹

The use of disparities studies without examining their underlying data and compatibility with judicial standards was always a weak exercise in legal analysis. Now that a decade has passed and no up-dating of Appendix A has occurred, that document should no longer serve a predicate for racial preferences.

⁵⁹ 61 Fed. Reg. at 26,061–62 (1996).

⁶⁰ *L. Tarango Trucking v. County of Contra Costa*, 181 F. Supp. 2d 1017, 1032 (N.D. Ca.2001). “Even if the NERA report had set forth accurate availability statistics for 1992, the Court is not persuaded that those statistics accurately reflect the current availability, nearly ten years later, of women-owned and minority-owned contractors who are qualified to do business with the County. No one has attempted to do an availability study since the NERA report was issued in 1992. In the absence of current availability data, there is no data from which this Court can conclude that the County’s contracting practices have a disparate impact on women-owned and minority-owned contractors.” *Id.* at 1032.

See also the criticism on data obsolescence from Denver district court, “The 1990 Study used some data from 1977. All studies show growth in the numbers of MWBEs in the Denver MSA and what may have been true in 1977 and 1978 will not support a need for a policy in 1988, given the dynamic growth trends in the Denver MSA.” *Concrete Works v. Denver*, 86 F. Supp. 2d , 1042, 1068 (D. Colo. 2000).

⁶¹ See for example, *O’Donnell Constr. Co. v. District of Columbia*, 963 F.2d 420 (D.C. 1992), *Contractors Ass’n of Eastern Pa. v. City of Philadelphia*, 893 F. Supp. 438 (E.D. Pa. 1995), 91 F.3d 586 (3rd Cir. 1996), *Associated Gen. Contractors of Am. v. Columbus*, 936 F. Supp. 1363 (S.D. Ohio 1996). See also for example, the comments by a district court about the “Disparity Study for Memphis/Shelby County Consortium” by D.J. Miller & Associates, October 1994:

There does not seem to be any basis for the assumption that all construction firms have equal qualifications to perform construction services for the defendant. It appears that DJMA failed to determine the licensing authority or the bonding authority of MWBEs. *West Tennessee Chapter of Associated Builders and Contractors v. Memphis City Schools*, 64 F. Supp. 2d 714 (W. D. Tenn. 1999).

Or the comment by a district court about the Brimmer-Marshall Atlanta study in *Webster v. Fulton County* where the court struck down the County’s minority and female business enterprise (MFBE) program in part because, “After reviewing the Brimmer-Marshall study, the Court finds that is insufficient to establish a strong basis in evidence for the 1994 MFBE program.” *Webster v. Fulton Co.*, 51 F. Supp. 2d 1354 (N.D. Ga. 1999) *aff’d* 218 F.3d 1267 (11th Cir. 2000).

Benchmark study

Whatever the strengths and weaknesses of state and local disparity studies, they could not be used as a predicate relied on in any challenge to federal spending programs with racial preferences. On May 23, 1996, the Justice Department proposed “benchmark limits” studies for each industry which were intended to represent the “level of minority contracting that one would reasonably expect to find in a market absent discrimination or its effects.” This outcome was supposed to control the narrow tailoring decision of whether race conscious means were necessary in federal procurement related to that industry.⁶²

When the subsequent benchmark study was released by the Department of Commerce two years later, it was assumed by Congressional Judiciary committee staff that since the report was so short (twelve pages of text and tables) and left so many unanswered questions about its methodology and conclusions, that some back-up and more comprehensive document would be made available. When that did not occur because no such document exists, the staff asked for a meeting with the benchmark authors. A meeting was scheduled for July 14, 1998, but it was held at the White House and no Commerce personnel attended. The Justice and White House officials representing the Administration knew almost nothing about the benchmark’s methodology, so little useful information was exchanged. A promise was extracted that further questions could be faxed to the benchmark authors, but those inquiries were unanswered.⁶³

On its surface, the benchmark study appeared to measure disparities in the availability and capacity of small disadvantaged businesses (SDBs) and their utilization as measured in federal contract dollars during fiscal year 1996. To account for differences in separate industries, the study bases its measurements on two-digit Standard Industrial Codes (SIC) classifications created by the Census Bureau. To its credit, unlike most state and local disparity studies, the benchmark study attempted to control for capacities and specialties of SDB and non-SDB firms contracting with the federal government. The study found some sort of disparity in 40 of the 68 comparisons it made and placed an “*” by that finding which it stated justified the imposition of the 10 percent price preference in that SIC code.

Again the validity of those conclusions depends on the validity of the availability measures and other variables. Some of the study’s calculations are contradictory and implausible. The study estimates that the small disadvantaged business capacity in SIC 50 (wholesale trades and durable goods) was 33.1 percent and in SIC 58 (eating and drinking establishments) 80.5 percent. In SIC 73 (business services), the benchmark study reported that SDBs were 47 percent of the total companies and 40.2 percent of the adjusted capacity, despite the fact that the Small Business

⁶² 61 Fed. Reg. at 26,042, 26,045.

⁶³ My criticisms of the benchmark study were published in *The Public Interest*, No 138, Winter 2000, pp. 91–98, under the magazine’s title, “To the ‘Disadvantaged’ Go the Spoils? My knowledge of the benchmarks studies methodology was enhanced by the discovery and depositions taken for *Rothe v. Department of Defense*, 262 F.3d 1306 (Fed. Cir. 2001).

Administration Web site for the relevant year reported SDBs were only 2,749 of 44,258 or 6 percent of the businesses in SIC 73. If large businesses and businesses not on the SBA list were added, the percentage of SDBs in SIC 73 would decline even further. Nevertheless, according to the Commerce study, SDBs in SIC 73 received “only” 26.4 percent of the federal contracting dollars and were thus underutilized and entitled to a 10 percent price preference. That preference triggered the *Rothe* case.

When the intention to produce the benchmark study was first announced in the Federal Register in 1996, the notice stated that census data in the form of the Survey of Minority Owned Business Enterprises (SMOBE) would be used to measure the availability of minority firms for federal contracting. But, after a lot of meetings, that data base was abandoned because the census had no information on which MBEs have bonding, licensing, are certified SDBs, or willing to engage in federal contracting.

Over the two year course of the study, ten to fifteen high level meetings involving the White House, the Departments of Commerce and Justice, the Office of Management and Budget and the Council on Economic Advisors staff and several academic consultants took place to discuss the benchmark methodology. Out of these meetings decisions emerged to hire a firm called Westat to collect data on which firms bid in a large sample of federal contracts. Bidding firms can, at least, be considered to be qualified, willing and able. Using this sample, Commerce was able to determine the number of SDB and non-SDB firms that bid at least once in each SIC analyzed, though it has never reported that figure. Then Commerce took an additional step. It added to the SDB availability figure 8(a) firms which are almost all minority firms whether or not they bid. No additions were made to the non-DBE data base. Since the benchmark study's underlying data have never been released even during litigation, the full consequence of including non-bidding 8(a) firms in the SDB total is not fully knowable. Potentially it is enormous. For example, in SIC 73, business services, where the *Rothe* case focused, discovery revealed that of the 2,206 firms included in the SDB category, 1,497 were 8(a) firms. Most 8(a) firms do not bid in open competition, but compete only for set-aside contracts limited to other 8(a) firms or receive sole source contracts based on their minority status. In 1995, only 4.3 percent of all 8(a) contracts were competitively bid. Obviously, if a firm does not bid, it is not really an available competitor in the universe of competitively bid contracts the benchmark study examined to determine disparities. Therefore, evaluating the appropriate utilization of SDBs in competitive contracts could not be done because of the inappropriate addition of 8(a) firms to the availability pool.

The benchmark methodology had other flaws which inflated SDB availability. It did not consider the number of times a firm bid or the size of the contracts a firm bid on, both of which are probably highly related to the dollars a firm might be expected to be awarded. Since non-SDBs are larger businesses than SDBs, it is plausible that for contracts without racial preferences they would bid more often for larger contracts. Commerce responded that in measuring capacity by controlling for the age and payroll of firms that they might have indirectly accounted for the

differences in the number and size of contracts bid on. In fact, no one did any analysis based on the number and size of contracts bid, so no one knows what that calculation would produce.

There are other benchmark study problems. SICs are reported at the two-digit, three-digit, and four-digit level. For example, SIC 73 is called business services, SIC 737 is computers, and SIC 7378, computer maintenance and repair, but SIC 73 also includes advertising and pest control. Thus, there is an important methodological decision to be made about the level of aggregation of analysis. In its decisions about aggregating or disaggregating census categories, Commerce was not consistent. Commerce combined three separate agricultural two-digit SIC codes (01, crops, 02, livestock, and 07, agricultural services) into a single analytical category. Forestry, 08, and fishing, 09, were combined into a single category, and four different two-digit SICs for mining (10, metal, 12, coal, 13, oil and gas extraction, and 14, extraction of nonmetallic minerals) were merged into a single category. The construction industry, on the other hand, was not only broken down by building construction, SIC 15, heavy and highway construction, SIC 16, and special trades, SIC 17, but by nine geographical regions as well. There is no satisfactory explanation for combining some two-digit SIC codes and ignoring regional differences, while disaggregating others for regional analysis, except Commerce's desire to have statistically significant results given the contract sample it used.

While various circuits and the federal OFCCP rules do not consider disparities over 80 percent as statistical evidence of discrimination because such disparities are not considered significant, Commerce included nine SIC categories with a disparity ratio exceeding 80 percent as eligible for bid preferences. Indeed, for heavy construction in the east south central region and special trades contractors in New England, Commerce targeted these categories for racial preference even though utilization for SDBs reached 97 percent of calculated availability.

While a properly done disparity analysis may create an inference of discrimination,⁶⁴ it does not necessarily indicate the source of the discrimination. Identifying the source of the discrimination is necessary for a narrow-tailored remedy to be employed. While the benchmark study states that its "methodology is designed to ensure that the price adjustments authorized by reforms are narrowly tailored to remedy discrimination,"⁶⁵ it offers no explanation of how that discrimination has occurred and who committed it.

Surely if federal procurement officers are discriminating in the award of contracts, a theory for which there is not a scintilla of evidence, sanctioning or retraining these officers rather than bidding preferences would be the narrow tailored remedy. The other theory Commerce advanced is that the discrimination may occur somewhere in the private marketplace and that SDBs are weakened as competitors because of discrimination by lenders or suppliers. Aside from the fact that the study attempted no identification of that private discrimination with any particularity,

⁶⁴ *Croson*, 488 U.S. at 509.

⁶⁵ Small Disadvantaged Business Procurement; Reform of Affirmative Action in Federal Procurement, 63 Fed. Reg. 35,716 (June 30, 1998).

such a theory is not plausible for three reasons. First, the regression adjustments in the study were intended to control for variations in firm capacities, so that differences in SDB and non-SDBs capacities were already taken into account when the disparities were calculated.⁶⁶ If there were other characteristics that might have affected the relative capacity or competitiveness of SDBs and non-SDBs, the study does not specify them. If firms of equal capacity receive unequal results in federal procurement, it is more plausible to believe that is caused by disparate treatment by federal procurement officers than by anything that happened in the private sector. Second, the benchmark study found patterns of underutilization and overutilization for SDBs in various SICs that don't fit any theory of private discrimination. For example, SDBs were underutilized in SIC 34, fabricated metal products, and overutilized in SIC 33, primary metal industries. Third, if underutilization of SDBs is a sign of discrimination somewhere in the private sector, what is overutilization of SDBs (for example, a common pattern in the construction industry according to the benchmark study) a sign of? Commerce officials had no explanation for that phenomenon and clearly had given it little thought.

The benchmark study has also been criticized by the National Academy of Sciences in its review of that study in 2005. In addition to agreeing with my criticisms about "the deficiencies in documentation" and "the superficial approach to defining industries,"⁶⁷ the Academy has its own technical criticisms: "Greatly impairing the usefulness of the study results and methodology, however, is the lack of documentation for key components in the estimation, particularly for the regression equations that were used to predict capacity values."⁶⁸

The Department of Justice promised that "Benchmark limitations will be adjusted every five years, as new data regarding minority firms are made available by the Census Bureau."⁶⁹ Despite the fact that such adjustments would seem to be a necessary narrow-tailoring requirement, and that census data were abandoned as the availability source, no adjustment in the benchmarks has taken place. If the benchmark study was ever a reliable measure of the "level of minority contracting that one would reasonably expect to find in a market absent discrimination or its effects," when it was produced, it is surely an obsolete guide for creating remedies today. Indeed, Census has abandoned the SIC categories in favor of the North American Industrial Classification system which is more precise.

Disparity Study Conclusions

Although disparity studies and the methodologies employed purport to be based on social science, such work has almost never been published in peer reviewed journals. About half of the

⁶⁶ *Id.* at 35,718, fn 10.

⁶⁷ NRC, *Women-Owned Small Businesses in Federal Contracting*, p. 60.

⁶⁸ *Ibid.*, p. 2.

⁶⁹ 61 Fed. Reg. at 26,046.

disparity studies in the country have been completed by five firms. Universities or independent academic bodies have rarely been involved.

Therefore, the conclusion of the U.S. General Accounting Office (now the Government Accountability Office, GAO) about disparity studies is not surprising. In 2001 at the request of Congress, GAO reviewed 14 representative contracting disparity studies completed around the country and found all these studies flawed methodologically and therefore unreliable. GAO concluded:

the limited data used to calculate disparities, compounded by the methodological weaknesses, create uncertainties about the studies' findings.... While not all studies suffered from every problem, each suffered from enough problems to make its findings questionable. We recognize that there are difficulties inherent in conducting disparity studies and that such limitations are common to social science research; however, the disparity studies we reviewed did not sufficiently address such problems or disclose their limitations.⁷⁰

The cost to the taxpayers of various forms of post-*Croson* disparity studies has surely reached more than \$70 million dollars, perhaps the largest government expenditure for social science research ever. The Commission on Civil Rights could do a great public service by urging that these studies conform to the existing social science and legal standards and that flawed or obsolete studies no longer be relied on. If those standards are not observed, discrimination may go undetected or false claims of discrimination may pollute political rhetoric and mislead the public about real problems for which solutions exist.

The standards need not be elaborate. For the statistical sections, the *Croson* formula should be fully observed by using data bases that can actually distinguish between firms that are relatively qualified, willing, and able to perform particular services. Since census data can provide no information about individual firms and have no measure of the relative qualifications, willingness, and ability of any group of firms, they should not be used as an availability data base. Dun & Bradstreet has data on many, but not all individual firms. The information it collects, however, does not permit consistent measurement of qualifications or ability and has no measure of willingness for public work. Neither the census nor Dun & Bradstreet data can distinguish certified MWBEs from the usually larger number of uncertified firms owned by women and minorities. Vendor lists kept by jurisdictions may be a crude measure of willingness for public work, but rarely contain much about qualifications or ability. Many vendors lists are little more than mailing lists. Pre-qualification lists may be a better availability source, if the study controls for the fact that firms are pre-qualified for various amounts of work and in different specialties. If the process is open and fair, the best availability source is to examine bidders on prime contracts and firms that supply subcontractor quotes to prime contractors.⁷¹

⁷⁰ GAO Report, p. 29.

⁷¹ "The city maintains records of all firms which have submitted bids on prime contracts. This would be a ready source of information regarding the identity of the firms which are qualified to provide contracting services as prime

Those firms are willing and believe they are qualified and able to do the work. Whatever the availability source, the researcher should test the hypothesis that any disparity discovered is caused by differential treatment of similar firms and not caused by differences in the qualification, willingness, and ability of the firms themselves. Since MWBEs are generally smaller, newer firms than non-MWBEs, and indeed must graduate from the MWBE category if they are very successful, this sort of testing is necessary, if false claims that a disparity is created by discrimination is not the result.

For the anecdotal sections, the standards set out in the *Cornell Law Review* are a good guide. Anecdotal evidence should meet nine standards.⁷² (1) Anecdotal evidence should be collected from both MBE and non-MBE contractors. (2) In the collection of anecdotal evidence attempts should be made to guard against “response bias”⁷³ and “interview bias.” (3) Collected anecdotes should be appropriate in time and place. (4) Collected anecdotes should relate specifically to discrimination in public contracting. (5) Collected anecdotes should be industry specific. (6) Collected anecdotes should be group specific. (7) Collected anecdotes should contain adequate details of specific instances of discrimination. (8) Attempts should be made to corroborate anecdotes of discrimination. (9) Anonymous responses should be discouraged.

By these standards, the 1996 federal reports relying on flawed state and local disparity studies should no longer serve as predicates for racial and gender preferences in federal contracting. The Commerce benchmark study, the only disparity study based on federal data, has its own flaws. All these studies rely on data gathered in 1996 or considerably before. In any other field, a description of economic activity in 1996 would not be regarded as a basis for current policy. No person should face disadvantages in seeking public contracts based on such obsolete and flawed studies.

The multiple problems regarding disparity studies that have been identified by courts, federal agencies, and independent researchers require new standards before millions more of taxpayer dollars are spent on them and unconstitutional MWBE and DBE programs are perpetuated. In addition, there should be a requirement for transparency, so that the requests for proposal, contracts, methodology, and underlying data on which these studies rely are publicly available. In that way, an appropriate dialogue on the validity of the disparity study conclusions can take

contractors. BBC does not explain why it did not use this data. On prime contracts only the firms which submit bids are ‘available.’” *Associated Gen. Contractors of Am. v. Columbus*, 936 F. Supp 1363, 1389. (S.D. Ohio 1996).

⁷² Jeffrey Hanson, “Hanging by Yarns? Deficiencies in Anecdotal Evidence Threaten the Survival of Race Based Preference Programs for Public Contractors, 88 *Cornell Law Review* 1433 (July 2003).

⁷³ This bias can take place when respondents are told the purpose of the survey to motivate them to participate or when members of different groups are asked different questions. For example, the cover sheet on the MGT 1998 study for the North Carolina Department of Transportation informed potential respondents, after mentioning the *Croson* and *Adarand* decisions, that: “The study has been initiated to evaluate the continued need for a Disadvantaged Business program.” Most MWBEs would probably see the link between the need to allege the continuation of discrimination and the preservation of the MWBE program which benefits them. Thus the problem of response bias was created in the direction for the survey.

place within government and by the public before programs employing racial classifications are implemented.

Race Neutral Alternatives

As this Commission stated in its 2005 report, *Federal Procurement after Adarand*, an essential requirement of the Supreme Court's strict scrutiny test is that narrow tailoring requires serious consideration be given to race neutral means in overcoming discrimination and other barriers before race preferences are employed. This principal has been adopted by the federal government in the administration of federal transportation funds. U.S. Department of Transportation regulations require recipients to "meet to the maximum feasible portion of [its] overall goal by using race neutral means."⁷⁴ The Commission discussed race neutral contracting strategies in five categories: (1) enforce nondiscrimination and subcontractor compliance; (2) increase knowledge about opportunities to contract with the federal government; (3) provide education or technical assistance to improve business skills and knowledge of federal procurement and how to win contracts; (4) give financial assistance or adjustments to offset the difficulties struggling firms encounter; and (5) expand contracting opportunities and promote business development in underutilized geographic regions.⁷⁵

Since the serious consideration of race neutral means requirement stems from judicial interpretations of the Fourteenth Amendment, state and local disparity studies that make policy recommendations, as most do, ought to engage in such serious consideration. That rarely occurs. When race neutral alternatives are treated at all, it is generally in the form of simple description, with no evaluation of their effectiveness. Even worse, there is no linking of race neutral alternatives to the types of discrimination the disparity study has allegedly found. Serious consideration should always be an active process, going far beyond taking a few race neutral packaged programs off the shelf and then discovering disparities remain.

Far too often, jurisdictions believe they have a contracting disparity, often misidentified or exaggerated as explained previously, but will admit to having only the vaguest notion of what caused it and, indeed, in litigation discovery will deny under oath that they can identify any specific discrimination affecting their contracting process. The necessary beginning point of any serious consideration of race neutral programs is always to identify the problem, discriminatory or nondiscriminatory, before investing in a solution. The following list of problems and solutions is meant to be instructive, rather than exhaustive. In many respects, this list parallels the Commission's earlier recommendations listed above for federal agencies, but it is tailored for state and local governments.

⁷⁴ 49 C.F.R. § 26.51(a) (2005).

⁷⁵ U.S. Commission on Civil Rights, *Federal Procurement after Adarand*, September 2005, p. 31 ff.

1. Problems with the lack of a well-functioning complaint process where claims of discrimination can be evaluated and remedied if valid

Jurisdictions should establish and publicize the opportunity for all of its contracting participants to have claims of discrimination adjudicated. The process should include strong anti-retaliation sanctions and should be able to take action if any patterns or general practices of discrimination are identified. Currently, there is a wide variation among jurisdictions in whether effective complaint procedures exist and in record keeping. If jurisdictions seek to identify the discrimination they are remedying with a race conscious goals program, having an effective complaint procedure is a necessity.

2. Problems with communicating about contracting opportunities

Jurisdictions should examine their procedures in communicating prime contract opportunities by improving Web sites, publications, advertisements and listings (Dodge reports etc.), and distribution of job specifications (plan rooms etc.). Current good faith effort requirements in most jurisdictions appear to be sufficient to publicize subcontracting opportunities, except in their current form they are not race neutral, requiring outreach only to MWBEs.

3. Problems with discriminatory awards of contracts

The jurisdiction awards all prime contracts, usually through a low bid process, in which there is little opportunity for discretion. Claims that the prime contracting awards process is discriminatory are rarely made, but if such claims are valid, the first obligation of the recipient would be to clean up that process. Claims that non-MWBEs primes discriminate in subcontractor awards are more common and are made by both MWBEs and non-MWBEs. If a race neutral good faith efforts outreach system is in place, primes can be required to keep records on which firms submitted subcontractor quotes and then to justify any selection of a subcontractor who did not supply the low quote. Fraudulent behavior in the record keeping and/or in the justification part of the program should lead to serious sanctions.

As it stands now, few jurisdictions know who submitted quotes on their subcontracts or even how many non-MWBEs subcontractors have been used. The only records they keep are of MWBEs utilized.

4. Problems with suppliers

Claims are occasionally made about discriminatory pricing, but more frequently the problem is the lack of established credit for small new firms or their inability to make large purchases so they don't get the lowest unit price. The jurisdiction should investigate claims of discriminatory supplier policies. If the problem is differential pricing because of economies of scale or differences in credit status, the jurisdiction could consider screening some firms and guarantee credit or make bulk purchases for them.

5. Problems with contract size

Since MWBEs are generally smaller, newer firms, a prevalence of large contracts will almost always cause a disparity. The jurisdiction should investigate when it is financially prudent to make smaller contracting packages and adopt policies on that practice.

6. Problems with bonding and lending

Small, new firms often have trouble obtaining bonds and credit. Jurisdictions should measure the extent of the problem and, if it is believed to stem from discrimination, they should report that charge to regulatory agencies, or if they have the authority, take action against the discrimination. If the problem is race neutral, the jurisdiction may wish to consider more flexibility in its requirements or working with bonding and lending firms to secure assistance to potential contractors it has screened.

7. Problems with lack of business plans or other skills to become a viable competitor for contacts

Small, new firms may be eager to engage in public contracting, but lack the experience or knowledge to put together a business plan that may be required for financing or bonding or to complete complex forms related to public work. Jurisdictions may find a number of resources including SBA, Chambers of Commerce, and established firms that can become mentors for firms that need this help.

Serious consideration of race neutral solutions should not be an empty rhetorical ritual, but should involve active problem identification and creative solutions as the Supreme Court instructed in its *Croson* decision.⁷⁶ Consistent with its advice in *Federal Procurement after*

⁷⁶ In *Croson*, Justice O'Connor stated: "In determining whether race-conscious remedies are appropriate, we look to several factors, including the efficacy of alternative remedies." She went on to say:

Many of the barriers to minority participation in the construction industry relied upon by the city to justify a racial classification appear to be race-neutral. If MBEs disproportionately lack capital or cannot meet bonding requirements, a race-neutral program of financing for small firms would, *a fortiori*, lead to greater minority participation. (*Croson*, 488 US. at 507).

She went on:

Even in the absence of evidence of discrimination, the city has at its disposal a whole array of race-neutral devices to increase the accessibility of city contracting opportunities to small entrepreneurs of all races: simplification of bidding procedures, relaxation of bonding requirements, and training and financial aid for disadvantaged entrepreneurs of all races would open the public contracting market to all those who have suffered the effects of past societal discrimination or neglect. Many of the formal barriers to new entrants may be the product of bureaucratic inertia more than actual necessity, and may have a disproportionate effect on the opportunities open to new minority firms. Their elimination or modification would have little detrimental effect on the city's interests and would serve to increase the opportunities available to minority business without classifying individuals on the basis of race. The city may also act to prohibit discrimination in the provision of

Adarand the Commission should provide guidelines on serious consideration of race neutral alternatives at the state and local level.

credit or bonding by local suppliers and banks. Business as usual should not mean business pursuant to the unthinking exclusion of certain members of our society from its rewards. (*Id.* at 509–10).

WRITTEN STATEMENT OF CONSTANCE F. CITRO, DIRECTOR, COMMITTEE ON NATIONAL STATISTICS¹

I appreciate the opportunity to appear at the briefing organized by the U.S. Civil Rights Commission on the use of disparity studies as evidence of discrimination in federal contracting. Since May 2004, I have been the director of the Committee on National Statistics (CNSTAT) at the National Academies, where I worked as a senior study director for 20 years. I am a Fellow of the American Statistical Association and have a Ph.D. in political science from Yale University.

My remarks are based largely on my experience as study director for the CNSTAT Steering Committee for the Workshop on Women-Owned Businesses in Federal Contracting at the National Academies. The steering committee was commissioned by the U.S. Small Business Administration in summer 2003 to hold a workshop to review relevant data and methods for estimating the representation of women-owned small businesses by industry for use in designating set-aside contracting programs. The workshop was held in May 2004. The steering committee deliberated and released its report with findings and recommendations in March 2005.² In the course of its work, the committee reviewed disparity studies conducted by state and local governments that were summarized by the Urban Institute in a 1997 report, benchmark studies conducted by the Department of Commerce in the late 1990s, and a preliminary Small Business Administration (SBA) study completed in 2002.

My remarks today address the following topics from the committee's report: definitions of disparity, representation, and discrimination; methodological issues for disparity studies; the pros and cons of specific studies; and, briefly, methodological issues for discrimination studies. My remarks are also informed by my work with the CNSTAT Panel on Measuring Racial Discrimination, which issued its report in February 2004.³

Definitions

Disparity, underrepresentation, and discrimination are different concepts:

¹ Views expressed or implied herein are those of the author and not necessarily of the National Academies.

² National Research Council, *Analyzing Information on Women-Owned Small Businesses in Federal Contracting*, Steering Committee for the Workshop on Women-Owned Small Businesses in Federal Contracting, Committee on National Statistics, (Washington, DC: The National Academies Press, 2005) [available: <http://books.nap.edu/catalog/11245.html>].

³ National Research Council, *Measuring Racial Discrimination*, Panel on Methods for Assessing Discrimination, Rebecca M. Blank, Marilyn Dabady, and Constance F. Citro, editors, Committee on National Statistics, (Washington, DC: The National Academies Press, 2004) [available: <http://books.nap.edu/catalog/10887.html>].

Disparity is simply a measured difference between two groups on an outcome of interest, such as a difference in the number of contracts awarded to women-owned or minority-owned small businesses and other businesses.

Underrepresentation is a disparity in which the difference goes against a particular group—for example, a lower number of contracts awarded to women-owned or minority-owned small businesses than to other businesses. The key issue for underrepresentation is how much is “too much,” that is, the threshold that distinguishes an adverse difference that is inconsequential from an adverse difference that may suggest some kind of action, such as a preferential contracting program.

Discrimination involves differential adverse treatment of a group compared with others based on membership in the group, such as failure to award a contract to a minority-owned firm because of a prejudice against minorities. Discrimination may also involve differential treatment on the basis of other factors that results in an adverse outcome for a particular group—a hypothetical example in contracting could be a requirement that bidders have certain types of experience that are not necessary for successful performance but that operate to exclude higher proportions of minority-owned or women-owned businesses.

While one can expect discrimination to result in an observable disparity, a particular measured disparity does not necessarily imply discrimination: the disparity may be due to any number of factors, perhaps but not necessarily involving current or past discrimination. Disparities that underrepresent a group by greater than a specified value have been cited to argue for preferential contracting programs on the assumption that discrimination must be present. Typically, a disparity ratio (see below) of 0.80 or less has been used to denote “underrepresentation” for this purpose. The SBA preliminary study reviewed by the CNSTAT panel in addition used a disparity ratio of 0.50 or less to denote “substantial underrepresentation.” While the steering committee concluded that the SBA thresholds were reasonable to use to present their results, it noted that there is no scientific basis for establishing a particular threshold value for underrepresentation or substantial underrepresentation; reasoned judgment and face validity must play a role.

Methodological Issues for Disparity Studies

Because it is rarely possible to observe discriminatory behavior or treatment in action, the measurement of disparity is a starting point for analysis of the possible need for preferential programs. Such studies commonly use a measure termed the “disparity ratio,” D , which requires the calculation of values of two shares for a target group, such as women-owned or minority-owned small businesses. The two shares are a utilization share, U , and an availability share, A . The utilization share looks at an outcome of interest, such as winning government contracts, measured in number of contracts or dollars awarded. Using women as an example, the utilization share would measure contracts (or dollars) awarded to women-owned small businesses, C_w , as a share of total contracts (or dollars) awarded, C_t . The availability share looks at the available

universe or pool for contracting, measured as numbers of businesses or gross dollar receipts. The availability share in this example would measure women-owned small businesses (or their gross receipts), W , as a share of total businesses (or total gross receipts), T . Taking the two shares as a ratio gives an estimated value for the disparity ratio:

$$D = U / A, \text{ where} \tag{1}$$

U , or utilization = C_w / C_t and

A , or availability = W / T .

If D is 1.00, then there is no disparity for women-owned small businesses: their actual share, U , of contracts is the same as their expected share, A , based on their representation in the total business population. If the ratio is less than 1.00 (e.g., 0.50, which could result if women-owned small businesses had 10 percent of contracts and 20 percent of businesses), then there is an adverse disparity or underrepresentation of women-owned small businesses among successful government contractors relative to the total business population. If the ratio is more than 1.00, then women-owned small businesses are overrepresented among successful government contractors relative to the total business population.

In the contracting arena, the availability share for such a target group as women-owned small businesses can be expected to vary across industries and other characteristics of businesses and contracts. For this reason, it is critical to use disparity *ratios* to measure representation and not simple counts or percentages of utilization. For example, if industry A has 10 percent women-owned small businesses and industry B only 2 percent women-owned small businesses, a comparison of counts or percentages of contracts awarded to women versus others in each industry would be misleading because one would not expect women-owned small businesses to win as many contracts in industry B as in industry A.

Critical to appropriate calculation of disparity ratios are agreed-on definitions, measurement methods, and validation techniques for each element of the ratio—the target group, the outcome of interest, and the total population—and of any other variables to be included in the analysis, such as industry. The steering committee's report provides an extended discussion of definitional, measurement, and validation issues for these elements in Chapter 4 of its report. Some of the issues and problems that can occur include:

- Target group definition: Such target groups as “women-owned small businesses” and “minority-owned small businesses” are usually well specified in official data sources for measuring utilization, such as the federal Central Contractor Registration. However, such agency-specified definitions may not be well reflected in the available data sources for measuring availability, such as the Census Bureau's Survey of Business Owners, conducted every 5 years. There are also incentives for businesses to misrepresent their status in order to be eligible for preferential treatment.

- Outcome of interest: While typically the outcome variable used to measure utilization is the share of dollar amounts of prime contracts awarded to a target group, looking at numbers of contract actions may also be helpful, as would looking at subcontracts if data were available. Outliers, such as very large contracts awarded in special circumstances, may affect the estimates in ways that should be examined.
- Total population: This element is often the most problematic, both conceptually and with regard to measurement. Generally, courts have used a standard for measuring availability shares of “ready, willing, and able” to perform government contracting. This standard raises such issues as what to do about the very large number of very small businesses, some of which may represent occasional efforts of fully employed wage and salary workers who do not intend to “grow” their business. On the other hand, a definition that excludes all businesses with no employees or with receipts below a certain level may inappropriately omit businesses that would be able to rapidly expand by hiring independent contractors or other means to respond to a government contracting opportunity. The same kind of objection applies to excluding businesses that are not registered with a government contracting agency.
- Inconsistent measures of utilization and availability: A common problem with disparity studies is that they compare apples and oranges. For example, the utilization share may use dollars of contract awards, whereas the availability share may use numbers of businesses instead of gross receipts. Or, the utilization share may pertain to a different time period than the availability share.
- Validation and documentation: Disparity studies should, but often do not, examine several different disparity ratios, carefully evaluate the quality of the data that are used, test the sensitivity of the results to possible sources of error, and thoroughly document all assumptions, data, and methods. To the extent that carefully validated measures of different types—such as receipts-based measures, numbers-based measures, measures that vary the lower limit of receipts for inclusion in the population of businesses, and measures that exclude outliers—tell a similar story, then the justification for (or against) a preferential contracting program would be strengthened. To the extent that different measures tell very different stories (for example, if the data show a large percentage of contract actions compared with a small percentage of contract dollars awarded to a target group), then additional analysis would be required to make a case for preferring a particular measure or set of measures among all those calculated.

Extant Disparity Studies

Urban Institute Meta-Analysis

In 1997, the Urban Institute produced a meta-analysis of state and local government-commissioned disparity studies that were conducted between 1989 and 1996.⁴ In meta-analysis, the researcher produces statistics that combine and summarize the results of more than one study on an outcome variable of interest, selecting studies that meet specified criteria for content and quality.

For its review, the Urban Institute collected reports of 95 studies and excluded 37 of them as deficient in one or more respects. The remaining 58 studies most commonly measured utilization by the dollar amount of contract awards or payments; some studies also defined utilization by the numbers of contract awards. Measures of availability were expressed in numerical terms (numbers of firms) and never in monetary terms (e.g., gross receipts), and they differed in the universe definition. The six most common universe definitions were previous award winners, firms that bid on past contracts or appeared on lists to receive information about procurements, firms certified as minority owned or women owned, firms that expressed interest in government contracting work in response to a survey, all firms with paid employees, and all firms.

To combine results across studies, the Urban Institute first averaged the disparity ratios for each jurisdiction that reported more than one ratio (most did) and then took the median of the study averages in order to minimize the effects of outliers. The averaging was performed separately for groups defined by ownership status (black, Hispanic, Asian, American Indian, women-owned) and industry category (construction, construction subcontracts, goods, professional services, other services).

The Urban Institute tested the sensitivity of the results to several methodological features of the various studies, but these breakdowns did not alter the picture conveyed by the overall results. It also tested the sensitivity of the results to the universe definition for the measure of availability, specifically, whether the measure included all or most firms, or, instead, used a measure that could be construed as “ready, willing, and able” (e.g., defining the universe in terms of registered bidders). As it turned out, the median disparity ratios for women-owned businesses differed little by whether the universe of available firms was defined broadly or more narrowly. In contrast, disparity ratios for minority-owned firms were significantly higher when the universe of available firms was defined broadly than when it was defined more narrowly. This finding suggests that, in the states and localities studied, minority-owned firms were overrepresented on vendor or bidder lists compared with their share of all firms.

⁴ María Enchautegui, Michael Fix, Pamela Loprest, Sarah von der Lippe, and Douglas Wissoker, *Do Minority-Owned Businesses Get a Fair Share of Government Contracts?* (Washington, D.C.: The Urban Institute, 1997).

The Urban Institute meta-analysis is not directly useful today. The data pertain to state and local contracting for those jurisdictions that chose to conduct disparity studies, and they are out of date. The study approach is useful because it demonstrates attention to detailed documentation of data and methods and careful explication of limitations of the analysis and the sensitivity of the results to factors that could have biased them. The data for each study included in the analysis are provided in the report, so that other researchers can examine them.

Department of Commerce “Ready, Willing, and Able” Analyses

In 1996, the Department of Justice tasked the Department of Commerce to designate utilization benchmarks for determining industries in which small disadvantaged businesses could receive a price evaluation adjustment, or bid-credit, to level the playing field with larger businesses. The Justice Department intended these benchmarks to represent the “level of minority contracting that one would reasonably expect to find in a market absent discrimination or its effects.” The Commerce Department’s Office of the Chief Economist and Office of Policy Development in the Economics and Statistics Administration released the results of its first benchmark study in 1998 and updated those results in 1999.⁵ No further studies have been conducted.

The Commerce Department’s study has been strongly criticized for the lack of documentation of data sources, analytical methods, and limitations of the results; for failure to develop meaningful industry groupings for a study of federal contracting; and for the lack of a theory of discrimination underlying the study. The criticisms about the deficiencies of documentation and the superficial approach to defining industries are well founded. Articulating a theory of discrimination requires defining the universe of available firms—whether it is appropriate to use an “all firms” definition from a source such as the Census Bureau’s Survey of Business Owners (SBO), or whether a specified definition of “ready, willing, able” firms is more appropriate.

The department first sought to assemble a data set of firms that were “ready and willing” to supply the federal government. It used data for 1996 on bidders from a sample of competitive procurements over \$25,000 identified from a survey of federal contracting officers, all firms that won sole-source or other noncompetitive procurements over \$25,000, and all firms certified by the Small Business Administration as active and eligible for Section 8(a) contracts. The department matched the firms in these three data sets by taxpayer identification number and major industry grouping to eliminate duplication and added a utilization measure for each successful bidder in the integrated data set from data on contract awards.

The next step in the department’s methodology was to measure the ability or capacity of the ready and willing firms in the integrated data set by using a regression imputation method that related successful firms’ dollars of contract awards to such variables as a firm’s years of experience and size of payroll. Finally, the department used the integrated data set to measure

⁵ Small Disadvantaged Business Procurement; Reform of Affirmative Action in Federal Procurement, 63 Fed. Reg. 35,714 (June 30, 1998); Small Disadvantaged Business Procurement; Reform of Affirmative Action in Federal Procurement, 64 Fed. Reg. 52,806 (Sept. 30, 1999).

utilization and availability shares for small disadvantaged businesses compared with all firms in the data set for major industry groups. Utilization shares were calculated using the value of prime contract dollars awarded; availability shares were calculated using the capacity value assigned to firms in the data set.

Despite some strengths, such as a consistent definition of utilization and availability shares in dollar terms and an innovative attempt to measure “ready, willing, and able” firms, the usefulness of the Department of Commerce study is greatly impaired by the lack of documentation for key components in the estimation, particularly for the regression equations that were used to predict capacity values. It is not clear whether or how the department may have evaluated the robustness of the regression method or of its data sources for identifying a pool of ready and willing firms.

Small Business Administration Study

The recent (2002) SBA study of women-owned small businesses in federal contracting by industry group⁶ followed the methodology of many of the state and local disparity studies reviewed by the Urban Institute, and it refers to the Urban Institute analysis as the basis for a number of methodological decisions. Like many of the state and local studies, SBA used inconsistent definitions for the utilization and availability shares, defining utilization in monetary terms and availability in terms of numbers of businesses. The data source for measuring utilization was the Federal Procurement Data System for 1999, which contains detailed information for all federal prime contracts over \$25,000, accounting for over 90 percent of the \$200 billion spent in fiscal year 1999. The data source for measuring availability was the 1997 Survey of Business Owners, with the universe limited to business firms with paid employees. Defining underrepresentation as a disparity ratio of 0.80 or less, the SBA estimated that women-owned small businesses were underrepresented in all but five of 71 industry groups (two-digit SIC categories) for which disparity ratios were calculated. Defining substantial underrepresentation to be a disparity ratio of 0.50 or less, the study estimated that women-owned small businesses were substantially underrepresented in 56 industry groups.

A major limitation of the SBA study is the same as that of the Department of Commerce study—namely, incomplete and unclear documentation of data sources and estimation methods and the lack of any published sensitivity analysis that would indicate the robustness of the estimated disparity ratios to alternative measures of utilization and availability. The SBA document indicates that seven data sets were examined for measuring availability, including four variations of the 1997 SBO; the SBA PRO-NET database of small businesses registered to do contracting with the federal government, which carries a self-designation of women-owned status; the Central Contractor Registration; and the Department of Commerce database of firms active in federal prime contracting in fiscal year 1999, including contract winners and bidders from a survey of contracting officers.

⁶ References for the study described below are provided in the CNSTAT steering committee’s report.

The SBA document does not present disparity ratios estimated using these various sources. It presents a summary justification for selecting the 1997 SBO data for firms with paid employees, but it does not indicate whether that justification applied to all industries. Other problems with the analysis are that it uses different years for estimating utilization and availability in a period of rapid growth of women-owned small businesses and that it uses the old two-digit Standard Industrial Classification for industries instead of the newer three-digit or four-digit North American Industry Classification System.

Measuring Discrimination

Disparities are often taken to indicate the presence of discrimination, but they may be due to any number of factors. To measure discrimination, researchers must be able to answer the counterfactual question: What would have happened to a woman-owned or minority-owned small business if the business had not been woman-owned or minority-owned? Literally answering the counterfactual is impossible because one cannot clone the contracting situation, substitute an owner identical in all respects except gender or race, and rerun the decision process. In some markets, such as for housing and employment, careful field experimentation involving pairs of testers that are alike in as many ways as possible except for race or gender, has been successfully used to examine discrimination. But these kinds of experiments appear difficult to conduct for contracting. In practical terms, the question therefore becomes what are acceptable statistical methods to move from observing a statistical disparity to concluding the existence of discrimination.

The reports of the steering committee and the Panel on Methods for Measuring Discrimination describe in detail the technical issues that confront statistical models of discrimination. Suffice it to say that careful attention is needed to the quality of the input data and to address such problems as the bias from omitted variables and sample selection. Simply throwing a large number of variables into a regression equation will not likely produce results that stand up to scrutiny; there must be sufficient understanding of the process being studied to justify the necessary assumptions of the regression model. For this purpose, it may be necessary to conduct case studies of contracting procedures and venues to gain insight into the contracting process to inform the specification of a suitable statistical model.

Concluding Observations

From this brief review of issues in measuring disparities and discrimination, I conclude with six points:

- Disparity studies are a reasonable first step to identify situations in which certain types of businesses could be disadvantaged in government contracting due to current or past discriminatory practices or behaviors. Observed disparities cannot establish discrimination nor the locus of any discrimination in time or space. For example,

discrimination might—or might not—occur in one or more aspects of the contracting process, such as decisions to combine agency requirements into fewer, larger procurements or more and smaller procurements, decisions to add more work to existing contracts or let new contracts, choice of criteria and their weights in evaluating bids, and the extent of outreach programs to various types of businesses to encourage them to become qualified for contracting work. Or, discrimination might—or might not—occur much farther back in the causal chain by which pools of ready, willing, and able bidders are developed, such as in opportunities for venture and working capital for new and continuing businesses, access to technical assistance and mentoring, and the like. Measured disparities cannot speak to these or other possible causal factors, but they are a starting point for analysis.

- To be relevant and convincing, disparity studies must meet high standards for validity, reliability, and reproducibility. In general terms, all data, methods, evaluations, and results must be thoroughly documented. More specifically, in constructing a disparity ratio, the same time period and metric (dollars or numbers) should be used in measuring the utilization and availability shares. In addition, more than one type of disparity ratio should be calculated to determine if the story is the same or different depending on the measure used; the sensitivity of the results should be tested to variations in methods and data and the presence of outliers in the data; and careful evaluation should be conducted to determine the best groupings of industries to use. Finally, explicit rationales should be provided for the population or universe specification as to how narrow or broad a definition to use of businesses that are “ready, willing, and able” to perform as government contractors. Again, results that show significant underrepresentation for a target group such as women-owned or minority-owned small businesses for a range of universe definitions will be the most compelling.
- There is little point in attempting to adopt the Department of Commerce benchmark procedure given the lack of documentation of the methodology, although the notion of capability analysis along the lines of the department’s study is worth exploring.
- The specific studies used in the Urban Institute’s meta-analysis are out of date and apply only to state and local government contracting in specific jurisdictions. However, the Urban Institute study does provide a model of careful specification, sensitivity analysis, and thorough documentation.
- The SBA study should be revised as recommended in the steering committee’s report.
- A research program on the contracting process in various industries and agencies that draws on case studies, administrative records, and statistical analysis could be very useful to inform government agencies not only of the possible role of discrimination, but also of ways to improve the process for all types of businesses.

WRITTEN STATEMENT OF IAN AYRES, WILLIAM K. TOWNSEND PROFESSOR, YALE LAW SCHOOL

My name is Ian Ayres. I am the Townsend Professor at Yale Law School. I am a lawyer and a Ph.D. economist. A substantial part of my research has concerned empirical tests of race and gender discrimination.¹ I've tested for disparate treatment in a variety of contexts ranging from new car negotiations and bail bond setting to taxicab tipping.² I've tested for unjustified disparate impacts in kidney transplant matching rules, consumer lending and law school admissions.³

On today's issue of disparity studies, I was a consultant to the Justice Department in its effort to reform federal affirmative action programs to comply with *Adarand*'s strict scrutiny requirements.⁴ I also advised the Commerce Department in developing the statistical methodology underlying both its 1998 and 1999 disparity studies. Finally, I was the government's "narrow tailoring" expert in a number of cases challenging federal affirmative action programs, including *Rothe Development Corp. v. U.S. Dept. of Defense*⁵ and *Dynalantic Corp. v. Dept. of Defense*.⁶ While I have testified in support of certain disparity studies, I have also argued that some affirmative action programs do not have sufficient evidence of narrow tailoring:

- (1) I have testified in cases in which I have refused to support the narrow tailoring evidence in a disparity study that did not have a rigorous benchmark;⁷
- (2) I helped develop the Commerce Department methodology which eliminated race-conscious bidding credits in "red-lighted" industries;
- (3) I recommended to the Justice Department that they eliminate the "rule of 2" set asides in procurement; and

¹ Ian Ayres, *Pervasive Prejudice: Unconventional Evidence of Race and Gender Discrimination* (University of Chicago Press, 2001).

² Ian Ayres, *Fair Driving: Gender and Race Discrimination in Retail Car Negotiations*, 104 HARVARD LAW REVIEW 817 (1991); Ian Ayres & Joel Waldfogel, *A Market Test for Race Discrimination in Bail Setting*, 46 STANFORD LAW REVIEW 987 (1994); Ian Ayres, Fred Vars and Nasser Zakariya, *To Insure Prejudice: Racial Disparities in Taxicab Tipping*, 114 YALE LAW JOURNAL 1613 (2005).

³ Ian Ayres & Richard Brooks, *Does Affirmative Action Reduce the Number of Black Lawyers?*, 57 STANFORD LAW REVIEW 1807 (2005); Robert S. Gaston, Ian Ayres, et al., "Racial Equity in Renal Transplantation: The Disparate Impact of HLA-Based Allocation," *Journal of American Medical Association*, vol. 270 (1993), p. 1,352.

⁴ See *Adarand Constructors, Inc. v. Peña*, 115 S. Ct. 2097 (1995).

⁵ 49 F. Supp. 2d 937 (W.D. Texas 1999).

⁶ 115 F.3d 1012 (D.C. Cir. 1997).

⁷ *F. Buddie Contracting, Ltd. v. Cuyahoga Comm. College Dist.*, 31 F. Supp. 2d 571 (D. Ohio 1998).

- (4) I have argued (in an article which is now under submission and available on the Internet) that the Supreme Court should have demanded more narrow tailoring evidence from the University of Michigan law school.⁸

I support the requirement that race-conscious governmental programs be strictly scrutinized. Courts should demand rigorous and persuasive evidence of both a compelling governmental interest and that the race-conscious means are narrowly tailored to further that compelling interest.

My central claim here today is that quantitative methods exist and have already been used to provide this kind of evidence. The Commerce Department's disparity study is a case in point. The results of the study created a "red light/green light" system which turned off the use of bidding credits where there was not evidence of underutilization of minority contractors. But it is my opinion that the evidence of discrimination in the "green lighted" industries is both rigorous and sufficiently persuasive to make out at least a *prima facie* case of narrow tailoring.

We should guard against efforts to turn the "narrow tailoring" requirement into a burden that no government defendant could ever meet. Justice O'Connor was quite clear that the requirement of strict scrutiny was not a subterfuge to create a "fatal in fact" requirement. Indeed, the Supreme Court's willingness to accept the narrow tailoring evidence with regard to the race-conscious affirmative action policy of the University of Michigan law school is strong evidence that cutting-edge quantitative disparity studies – such as the one produced by the Commerce Department – pass constitutional muster.⁹ In the Michigan case, the Supreme Court required almost no statistical evidence that the law school used the minimum racial preference necessary to achieve its compelling interest. But the Supreme Court nonetheless was willing to sign off on the constitutionality of the affirmative action program. The best procurement disparity studies already provide much more persuasive narrow tailoring evidence and are *a fortiori* more clearly constitutional.

In the remainder of this statement, I will analyze 1) the evidence supporting the government's compelling interest in remedying discrimination; 2) the most persuasive methodologies for estimating disparity benchmarks; and 3) a comparison of the narrow tailoring evidence in procurement and educational admissions.

⁸ See Ian Ayres & Sydney Foster, "Don't Tell, Don't Ask: Narrow Tailoring After Grutter and Gratz" (unpublished working paper, 2005).

⁹ Grutter v. Bollinger, 539 U.S. 306 (2003).

I. There is Credible Evidence of a Compelling Government Interest.

No one disputes the fact that remedying discrimination is a compelling governmental interest. And there is abundant statistical evidence that discrimination not a thing of the past.¹⁰

Many commentators have argued, however, that government can only use race-conscious affirmative action to remedy its own discrimination. In fact, Justice Powell's plurality opinion in *Wygant v. Jackson Board of Education*,¹¹ noted that the Supreme Court "has insisted upon some show of prior discrimination by the governmental unit involved before allowing limited use of racial classifications in order to remedy such discrimination." Jeffrey Rosen seized upon this language to suggest that federal affirmative action was unconstitutional:

[T]he Supreme Court will only uphold federal [affirmative action] in light of convincing evidence of past discrimination by the federal government itself; but, for almost twenty years, the federal government has been discriminating in favor of minority contractors rather than against them.¹²

Because the federal government has not been discriminating against minorities for decades, Rosen argued that it has no compelling interest for affirmative action in procurement.

But the idea that affirmative action can only be used to remedy the government's own discrimination was flatly rejected by Justice O'Connor. In *City of Richmond v. J.A. Croson Co.*, Justice O'Connor in a plurality opinion joined by Chief Justice Rehnquist and Justice White, concluded that the City of Richmond "can use its spending powers to remedy private discrimination, if it identifies that discrimination with the particularity required by the Fourteenth Amendment."¹³

While government discrimination against minority contractors in procurement markets may be a thing of the past, the same cannot be said of private discrimination. Credible evidence of private discrimination by both input suppliers to, and customer of, minority contractors provides a persuasive basis for government to "use its spending powers to remedy private discrimination."¹⁴

Discrimination is predominantly practiced today in private markets. Government has a compelling interest to try to remedy it. Narrow tailoring of course requires the use of race-neutral methods – such as simply prohibiting private discrimination – whenever possible. But a great deal of private discrimination will necessarily fall below the radar screen of the law.

¹⁰ See The Compelling Interest for Affirmative Action in Federal Procurement: A Preliminary Survey, 61 Fed. Reg. 26,050–63 (May 23, 1996).

¹¹ 476 U.S. 267, 274 (1986).

¹² Jeffrey Rosen, "The Day the Quotas Died," *New Republic*, Apr. 22, 1996, p. 26.

¹³ *City of Richmond v. J. A. Croson Co.*, 488 U.S. 469, 492 (1989).

¹⁴ See Ian Ayres & Fred Vars, *When Does Private Discrimination Justify Public Affirmative Action?*, 98 COLUM. LAW REV. 1577 (1998).

Discrimination that cannot be proven in individual cases can often be identified in the aggregate. If we flip a coin once and it comes up tails, we can't tell if it is a fair coin. But if we flip it 1000 times and it keeps coming up tails, we know that something is wrong. When government through race-neutral prohibitions is unable to eliminate private discrimination itself, Justice O'Connor found that it could constitutionally use its spending power to "eradicate the effects of private discrimination."¹⁵

II. There are Persuasive Statistical Methods for Calculating Disparity Benchmarks.

The crucial and most disputed element of any disparity study is calculating the "benchmark" (sometimes referred to as the "minority availability percentage"). The benchmark attempts to measure the market share percentage of business that minority firms would receive in a world without discrimination.

The benchmark is crucial to establish (1) whether minorities face discrimination; and (2) whether the proposed racial preferences are sufficiently limited so as to only remedy the discrimination.

For example in a particular market, if a disparity study persuasively concludes that in the absence of discrimination minority contractors would have received 10 percent of the contracts, but we observe that minority firms are only receiving 4 percent of the contracts, then this shortfall in utilization is evidence of discrimination. Underutilization evidence of this kind is thus probative of the "compelling interest" prong of strict scrutiny.

But the benchmark is also crucial in testing whether an affirmative action program is narrowly tailored. Put simply, if the minority benchmark is 10 percent, an affirmative action program that induced a 13 percent minority utilization would not be narrowly tailored. Racial preferences are simply too large if they would cause minority utilization to overshoot the market share that would prevail in the absence of discrimination.

So the crucial question in disparity studies is to develop a credible methodology to estimate this benchmark share of contracts minorities would receive in the absence of discrimination. The touchstone for measuring the benchmark is to determine whether the firm is "ready, willing and able" to do business with the government. Early disparity studies attempted to calculate benchmarks on a very crude "head counting" methodology: If minorities are *X* percent of the general population, then under this theory courts should assume that absent discrimination they would be awarded *X* percent of procurement dollars.

Increasingly, however, courts have rejected mere head counting and moved toward a "qualified-firm counting" approach.¹⁶ This "qualified-firm counting" approach requires courts to identify

¹⁵ *Croson*, 488 U.S. at 491–92.

¹⁶ *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989). See also *Croson*, 488 U.S. at 501, *quoting* *Hazelwood School Dist. v. United States*, 433 U.S. 299, 307–308 (1977) ("[W]hen special qualifications are required to fill

the pool of firms which are “qualified” (in the sense of being ready, willing and able) to do business with the government. Under this approach, the minority benchmark percentage would then be calculated as the percentage of qualified firms that were minority contractors. This approach implicitly assumes that if X percent of the qualified firms are minority contractors, then absent discrimination they should be awarded X percent of the contract dollars.

While the “qualified-firm counting” approach represented a substantial advance over the “head counting approach,” it suffered from the problem that “qualified” firms may have substantially different capacities. Firms A and B may both be qualified to do *some* business with the government, but one firm may be a multinational with many plants, while the other firm may be a sole proprietorship with only a single plant. The “qualified-firm counting” approach ignores differences in capacity and deems the single-plant firm to be equally “available” to serve the government as the multiplant firm. It might assume, for example, that the manufacturers of a small micro-brewery brand and Budweiser are equally available to sell beer.

The Commerce Department’s approach for estimating the minority benchmark was far more sophisticated than either the “head counting” or the “qualified-firm counting” approaches. This methodology – which I will refer to as the “capacity” approach – calculates in dollar terms the capacity of qualified firms to do business with the government. The minority benchmark (“the SDB availability percentage”) was calculated as the SDB’s share of the industry’s total capacity. This approach more reasonably assumes that if SDBs control X percent of an industry’s capacity, then absent discrimination they would be awarded X percent of the industry’s procurement dollars.¹⁷ Unlike the “qualified-firm counting” approach, the “capacity” approach would not find that the manufacturers of a small micro-brewery brand and Budweiser are equally available, but instead would likely find that Anheuser Busch is more available in the straightforward sense that it has a larger capacity.

The details of the Commerce Department’s algorithm are discussed in the Memorandum from Jeffrey Mayer, dated June 23, 1998, re: Price Evaluation Adjustments and Benchmarking Methodology.¹⁸ But the essence of the approach is (1) to identify firms who are qualified to do business with the government, (2) to estimate the capacity of these qualified firms, and (3) to calculate what share of an SIC code’s capacity is controlled by minority contractors.

The Commerce Department’s capacity methodology is particularly conservative because it did not attempt to calculate how much greater minority capacity might have been but for discrimination. A so-called “but for” adjustment would raise the benchmark percentage by which

particular jobs, comparisons to the general population (rather than to the smaller group of individuals who possess the necessary qualifications) may have little probative value”).

¹⁷ See *Engineering Contractors Ass’n of South Fla. v. Metropolitan Dade County*, 943 F. Supp. 1546 (S.D. Fla. 1996), *aff’d*, 122 F.3d 895, 916 (11th Cir. 1997).

¹⁸ 63 Fed. Reg. 35,719 (June 30, 1998).

underutilization is judged. This methodology instead takes minority capacity as it finds it. It thus makes no attempt to remedy historical discrimination or even present discrimination by input suppliers or customers that predictably would depress the capacity of minority firms to supply government contracts.

Like Justice O'Connor, I strongly support a requirement that government justify race-conscious policies by providing persuasive evidence that the policies are narrowly tailored to promote a compelling government interest. The Commerce Department's disparity studies are rigorous and provide credible prima facie evidence of both discrimination and the potential for narrowly tailored race conscious remedies.

It is striking that the Commerce Department has not seen fit to update its benchmark analysis since 1999. I worry that the present administration is trying to achieve a backdoor sun-setting of the remedial race conscious programs by fostering the increasing desuetude of the necessary narrow tailoring evidence. Regardless of how one feels about affirmative action, we should "mend, not end" disparity studies.

III. The Constitutionality of the Race-Conscious Affirmative Action in Educational Admission Strengthen the Inference that the Commerce Department's Disparity Study Provides Credible Evidence of Narrow Tailoring.

While educational affirmative action has been the preoccupation of academics, it is plausible that affirmative action in procurement has been just as important in remedying race discrimination in the United States.

In 2003, the Supreme Court upheld the constitutionality of the University of Michigan Law School's race-conscious affirmative action policy.¹⁹ In *Grutter*, the Supreme Court required almost no statistical evidence that the law school in fact used the minimum racial preference necessary to achieve its compelling interest.

While the narrow tailoring requirement has always had multiple dimensions, a central meaning has been that the government use only the minimum racial preference necessary to achieve its compelling interest. Sometimes expressed as a requirement that plans use the "least restrictive" or "least burdensome" alternative, a core requirement is that plans should use the minimum necessary racial preference. If the government objectives could be fulfilled without use of a racial preference, then no racial preference would be allowed. If only mild racial preferences were needed to achieve the compelling government interest, then nothing more than mild preferences would be constitutionally countenanced.

But the truth is that in *Grutter* the Supreme Court required virtually no evidence that the law school used the minimum preference necessary. In part, we know that this was the case, because

¹⁹ *Grutter v. Bollinger*, 539 U.S. 306 (2003).

the Court never learned even the degree of racial preference that was given by the law school – so it could certainly not assess if this level were the minimum preference necessary to accomplish the compelling government interest.

Now that the Supreme Court has signaled its willingness to support more flexible modes of proof, it would be bizarre for it to strike down much more rigorous narrow tailoring evidence. The constitutionality of the race-conscious policy in *Grutter* suggests all the more that the evidence of narrow tailoring in the best procurement disparity studies will be found constitutional. These remedial disparity studies are more powerful statistical evidence of narrow tailoring than produced in any of the existing diversity studies. Unlike the *Grutter* narrow tailoring methodology, the Commerce Department's approach has red lights as well as green lights. It turns off preferences if preferences would lead to over utilizations.

Conclusion

It is a great honor to have the opportunity to speak to this Commission that has played such a remarkable role in this nation's struggle to secure equality for all of its citizens.

The quantum of evidence necessary to show discrimination is reasonably debatable. I support the requirement that government show persuasive evidence of discrimination and a narrowly tailored remedy. I believe that the best disparity studies currently provide prima facie evidence of narrow tailoring. But I look forward to discussing alternative views.

Thank you for giving me this opportunity to speak.

THE USE OF DISPARITY STUDIES TO JUSTIFY RACIAL AND ETHNIC PREFERENCES IN GOVERNMENT CONTRACTING, PREPARED STATEMENT OF ROGER CLEGG, VICE PRESIDENT AND GENERAL COUNSEL, CENTER FOR EQUAL OPPORTUNITY

Thank you, Mr. Chairman, for the opportunity to testify this morning before the Commission regarding the use of so-called disparity studies to justify the use of racial and ethnic preferences in government contracting. [Note: Disparity studies are also sometimes used to justify gender preferences, and the legal issues raised then are similar.]

My name is Roger Clegg, and I am vice president and general counsel of the Center for Equal Opportunity, a nonprofit research and educational organization that is based in Sterling, Virginia. Our president is Linda Chavez,¹ and our focus is on public policy issues that involve race and ethnicity, such as civil rights, bilingual education, and immigration and assimilation. I should also note that I was a deputy in the U.S. Department of Justice's Civil Rights Division for four years, from 1987 to 1991. While there, I worked on a number of cases involving contracting preferences, including an amicus brief filed by the United States in *City of Richmond v. J.A. Croson Co.*,² successfully urging the Supreme Court to strike down the municipal preferences at issue there. And I should also note that the Center for Equal Opportunity has in recent years contacted many government actors and warned them of the divisive and unfair nature, as well as the costly legal consequences, of using racial and ethnic contracting preferences and urged them, instead, to adopt race-neutral alternatives. Our efforts have been quite successful.

You will be hearing from several other witnesses today who are experts on the history and application of disparity studies. What I will focus on, since I am a lawyer, is the legal context in which disparity studies are used and the legal framework that should be used in evaluating them. My bottom line is that even a properly done disparity study is—as a constitutional and policy matter—unlikely to justify the use of racial and ethnic preferences.

The starting point of any legal analysis is: “A racial classification, regardless of purported motivation, is presumptively invalid and can be upheld only upon an extraordinary justification.”³ This “strict scrutiny” is triggered by “all racial classifications, imposed by whatever federal, state, or local governmental actor.”⁴ This is true whether there is an out-and-out set aside or quota, or whether there are race-based “discounts” made in evaluating

¹ Since the briefing, Mr. Clegg has been named president and general counsel of the organization. Mrs. Chavez remains the chair of its board of directors.

² 488 U.S. 469 (1989).

³ *Personnel Administrator v. Feeney*, 442 U.S. 256 (1979).

⁴ *Adarand Constructors, Inc. v. Peña*, 515 U.S. 200 (1995).

contracting bids (see, e.g., *Adarand*), or whether there are contracting percentage “goals”⁵—which, as a practical matter, inevitably become quotas anyway.

Once strict scrutiny is triggered, a court will look to see, first, whether the use of race is justified by some “compelling” government interest, and, second, whether the use of race is “narrowly tailored” to achieving that interest—that is, that the use of preferences is essential and that they are used no more than is necessary. Thus, in an August 2001 brief, the U.S. Department of Justice told the Supreme Court that the federal program at issue “may use race-conscious remedies only as a last resort” (citing 49 C.F.R. 26.51(a)), “where the effects of discrimination are stubborn, persistent, and incapable of eradication through race-neutral measures.”

In the government contracting context, indeed, the only compelling interest that can be put forth is the remedial one—that is, in correcting discrimination against the racial group that is awarded the preference. Thus, the government could not claim that it wanted greater intellectual or viewpoint “diversity” among its contractors: There is no black way or white way or Hispanic way or Asian way to build a guardrail or fill a pothole.

This is where disparity studies come in. The idea is that the underutilization of contractors of this or that color shows that they are suffering discrimination—a necessary precondition to the invocation of the remedial justification.

I will turn to the care that must be taken in ensuring that the evidence marshaled in a disparity study actually does demonstrate discrimination. But I want to stress at this point that, while the presence of discrimination is a *necessary* condition to the use of racial preferences, it is not *sufficient*, for the government also needs to demonstrate that the use of preferences is “narrowly tailored” to correcting whatever discrimination is found.

This means, among other things, showing that there is no other way to end the discrimination other than by using racial and ethnic preferences. And, in 2005, this simply cannot be done.

But let’s go back to the first prong of strict scrutiny—that is, demonstrating that significant racial and ethnic discrimination against minority-owned businesses in contracting exists in the first place.

There are basically two ways to do this: through anecdotal evidence, and through statistical evidence.

With regard to anecdotal evidence, that evidence should be weighed for its reliability and its recentness; considered in the context of its typicality and how many contracting decisions are made (so that there is a clear pattern of consistent discrimination—see *Croson*); and collected evenhandedly, so that discrimination against *all* racial groups (including non-Hispanic whites,

⁵ See *Lutheran Church–Missouri Synod v. FCC*, 141 F.3d 344 (D.C. Cir. 1998).

for instance) is considered. Finally, there must be a link shown between the discrimination against a particular group and the denial of contracting opportunities to members of that group.

With regard to statistical evidence, it must always be borne in mind that “disparity” means only “difference,” and it takes more than that to show *discrimination*. Thus, it is immediately obvious that simply showing a disparity between a racial group’s percentage in the general population and the percentage of contracts awarded to companies owned by that group is meaningless, because that group may simply not own companies in the same proportion as its share in the general population. We don’t expect half of all companies to be owned by women, even though half of all our people are women. For one thing, many companies are owned by hundreds or thousands of stockholders, and so are neither “male-owned” nor “female-owned.” So the figures must be adjusted accordingly.

But this is not enough either. The percentage of minority-owned companies generally may not be the same as its percentage in the particular markets where you find government contracting. For example, there may be a high percentage of Korean-owned businesses in a city, but if most of those businesses are mom-and-pop grocery stores, the percentage is misleading, since this may not be a sector in which there is much government work we expect grocery stores to bid on. So the percentage has to be adjusted again.

And this is not enough either, because it may be that the companies owned by a particular racial group in a particular market do not have the specific expertise or capability of doing the actual work sought. And, even if they are, there may not be a problem if these companies are for nondiscriminatory reasons not submitting bids in the first place.

Finally, even if it is shown that a particular group bids on contracts in a particular market and fails to receive the contracts in its proportion, this does not mean that discrimination has occurred: The explanation may be simply that the bids submitted by the minority company were not the best bids, because they were too high or failed to meet other objective qualifications.

So I am skeptical about the utility of statistical evidence in this area. And even if—through statistical evidence or anecdotal evidence—a pattern of recent discrimination is found, it does not follow that racial and ethnic preferences must be used to correct it.

At every step of the process, it is clear that there are more narrowly tailored remedies than using racial preferences. If companies are being excluded from bidding because of unrealistic or irrational bonding or bundling requirements, then those requirements should be changed for *all* companies, regardless of the skin color of the owner. If companies who could submit bids are not doing so, then the publication and other procedures used in soliciting bids should be opened up—but, again, to all potential bidders, not just some. And, finally, if it can be shown that government bids are being denied to the lowest bidder because of that bidder’s race, then there should be put in place safeguards to detect discrimination and sanctions to punish it—but, again, those safeguards and sanctions should protect all companies from racial discrimination, not just some.

Contracts are not like hiring, promoting, or even university admissions, where there is an irreducible and significant amount of subjectivity in the decisionmaking. Contracting is an area that can be made very transparent and where this transparency should make it relatively easy to detect and correct discrimination.

Even if there could still, in theory, be a few cases of discrimination that go unremedied in the absence of racial classifications, there will be many more cases of discrimination that will result from the institutionalization of racial and ethnic preferences.

The study that the Commission recently published, Mr. Chairman, did a very good job of collecting and discussing these race-neutral alternatives.⁶ My only criticism of the report is that it did not make clear that the aim of the alternatives is to correct and end discrimination—not to achieve a particular percentage of contracting by this or that racial or ethnic group.

In sum, Mr. Chairman, I think that great care must be taken in preparing a disparity study to ensure that the evidence marshaled actually demonstrates discrimination. But even where this is done, I doubt that a study can ever justify the use of racial or ethnic preferences to end the discrimination found.

Thank you again, Mr. Chairman, for the opportunity to submit this testimony today.

⁶ U.S. Commission on Civil Rights, *Federal Procurement after Adarand*, September 2005.

THE COMMISSION'S FINDINGS AND RECOMMENDATIONS

In 1996, the U.S. Department of Justice cited results from disparity studies and other sources to justify the government's continued use of racial classifications after the Supreme Court's decision in *Adarand v. Peña*.¹ As part of its mission to evaluate federal enforcement, the U.S. Commission on Civil Rights conducted a briefing on December 16, 2005 to: inform itself about the strength and quality of the social science which underlies the Justice Department's justifications; determine whether the methodologies and factual predicates that Justice utilized are sufficient to rebut the Constitution's presumption that race-based classifications are illegal; and offer recommendations for how such efforts could be improved.

At the briefing, participants acknowledged the importance of narrowly tailoring programs that redress discrimination to ensure that such practices do not result in discrimination against groups that are not afforded such benefits. Indeed, the Commission has reminded agencies of the need to seriously consider race-neutral alternatives before resorting to preferences for minority- or women-owned businesses elsewhere.² The panelists explained how identifying underlying causes

¹ *Adarand Constructors, Inc. v. Peña*, 515 U.S. 200, 226 (1995).

² See U.S. Commission on Civil Rights, *Federal Procurement after Adarand*, September 2005.

The Commission identified the following practices that agencies may use to demonstrate serious consideration of race-neutral alternatives:

“(1) Identifying and evaluating a wide range of initiatives, rather than considering only one or two alternatives.”

“(2) Documenting the underlying facts that demonstrate whether a race-neutral plan will work, rather than relying on casual observation or undocumented assumption.”

“(3) Demonstrating an empirical basis for determining whether race-neutral plans will be effective, rather than relying on speculation.”

“(4) Ensuring data to demonstrate assessments are current, competent, and comprehensive.”

“(5) Periodically revising race-conscious plans, perhaps annually or biennially, to determine the need for continuing them and the viability of implementing race-neutral plans.” (Ibid., p. 24.) Such reviews also ensure that race-conscious programs are narrowly tailored.

“(6) Analyzing data to establish causal relationships, rather than relying on broad assumptions, particularly before concluding that a race-neutral plan is ineffective.” (Ibid., p. 24.)

Because the Commission found that agencies were not systematically engaging in such activities, it offered the following framework, representing actions agencies could undertake to meet this obligation. Agencies should (1) develop standards and policies based upon sound benchmark data, that is, the social science research used to demonstrate that minority groups have not realized the program benefits that they would expect in the absence of discrimination; (2) identify and develop a wide range of race-neutral alternatives to strategies that are based upon race; (3) routinely evaluate the effects of race-neutral and -conscious strategies; and communicate and coordinate with other agencies to share information about successful race-neutral programs and best practices. Ibid., pp. 24–30.

of disparities can lead to race-neutral strategies that will make federal contracts more accessible to minority- and women-owned businesses as well as nonminority firms that may be similarly disadvantaged.

Additionally, the briefing identified three types of research relevant to the Commission's examination, specifically: (1) state and local disparity studies which formed part of the basis for the Justice Department's rebuttal; (2) national disparity studies; and (3) other research helpful to identifying the extent and nature of discrimination and explaining causes of disparities correctable with race-neutral approaches.

Panelists and Commissioners discussed the applicability of government contracting programs to private sector discrimination. Although many such problems could not be resolved in a single briefing, facts gathered for the forum clarify and instruct entities who serve various roles in policy formation and implementation. Federal contract personnel, state and local program administrators (who face similar strict scrutiny requirements through *Croson*), legal analysts, researchers and social scientists, policy makers, and civil rights enforcement officers are among the myriad entities who will find the Commission's conclusions useful. To these corps the Commission offers the following findings and recommendations.

States and Localities Conducting Disparity Studies

Finding 1: State and local jurisdictions must have local evidence of discrimination against specific groups in contracting to narrowly tailor any race-conscious goals they set.

Finding 2: Most current disparity studies are not only outdated, but have common flaws. They fail to measure availability according to requirements to compare qualified, willing, and able businesses that perform similar services. They use simple counts of businesses without taking capacity into account. The researchers (1) use obsolete or incomplete data; (2) report results in ways that exaggerate disparities; (3) fail to test for nondiscriminatory explanations for the differences; (4) find purported discrimination without identifying instances of bias or general sources; (5) rely on anecdotal information that they have not collected scientifically or verified; (6) do not examine disparities by industry; and (7) neglect to identify which racial and ethnic groups suffer from the disparities.

Recommendation 1: States and localities must discard disparity studies conducted using data that is more than five years old. The results are too outdated to justify preferential awards given today. Researchers must use current data.

Recommendation 2: Officials that operate affirmative action programs must ensure that the methodology of any disparity study justifying racial or ethnic preferences adheres to generally accepted social science research standards. Disparity studies must demonstrate validity, reliability, and reproducibility of results. Researchers must thoroughly document all data,

methods, evaluations, and results. Reports must provide sufficient information that others can examine the data and generate the same findings.

Recommendation 3: Researchers must develop an explicit rationale for including businesses in the availability measure as qualified, willing, and able to carry out contract work. Their work should compare only businesses that are able to perform the same services. Analysts should remove from the pool of available businesses any companies offering services that a government does not purchase or that are distinctively different.

Recommendation 4: Researchers, and the federal, state, or local jurisdictions relying on their work, must use multiple measures of disparities, estimating the available pool of firms through both broad and narrow definitions, to ensure that results are not unique to one particular definition.

Recommendation 5: Researchers should perform detailed analyses. They should calculate disparity ratios within meaningful categories, such as specific industries, racial and ethnic groups, or contract amounts, so that government can use the research results to appropriately narrowly tailor preferences.

Recommendation 6: Analysts should use measures of available firms that account for the businesses' capacity to perform the work. At a minimum, they should examine disparity ratios by size of business. For example, instead of contrasting small minority businesses with all other firms, researchers should compare them to other small businesses. Yet, categorizing businesses as small, medium, and large is only a weak measure of capacity. The research should attempt to include additional and more fine-tuned measures of capacity, such as revenue, number of employees, or the firm's payroll.

Recommendation 7: Studies must calculate ratios using numerators and denominators with comparable units, for example, either dollars or numbers, but not a mixture of the two.

Recommendation 8: Similarly, utilization and availability measures, that is, numerators and denominators, must represent the same time period to avoid any distortion from changes in the composition of the business community.

Recommendation 9: Researchers should draw comparisons between minority- and majority-owned firms. They should exclude large, publicly-traded stock corporations from the analysis.

Recommendation 10: Researchers should test for the sensitivity of results to variations in methods and data, effects of different groupings of industries, and the presence of extreme contracts, called "outliers," that inordinately affect results. When results with different measures vary, the researchers must seek a deeper understanding of the data to explain underlying causes.

Recommendation 11: Social scientists should conduct separate analyses for prime and subcontractor data. Because minority owned enterprises more likely receive subcontracts,

analysts should compile disparity ratios comparing only prime awards for each group to the total prime awards. Similarly, they should compare subcontract awards for each group to the total subcontract awards.³

Recommendation 12: In analyzing disparities, social scientists should consider new ways of examining data. They could present bid success rates for the racial, ethnic, and gender groups, to determine disparities in frequency of competing for contracts and success in winning them.

Recommendation 13: Before concluding that disparities result from bias, researchers must test for nondiscriminatory explanations for the differences and identify specific instances or possible sources of group differences. They might factor in the experience of firms; whether the company owns its own equipment, has necessary licenses, bonding, a line of credit, and past bidding experience; education of the owner; and percent of work conducted as a prime contractor.

Recommendation 14: State and local jurisdictions must use a proper statistical foundation to show a pattern of discrimination or justify the use of affirmative action programs. They may rely on anecdotal evidence only to provide a deeper understanding of how disparities relate to discrimination.

Recommendation 15: Researchers must apply scientific standards to their use of anecdotes to bolster claims of discrimination. They must systematically collect anecdotes from a scientifically drawn sample, document the size of the survey universe that responded, and make an effort to investigate and verify content. Recognizing that reports represent anecdotes anonymously, authors should include sufficient detail and other information to lend credibility to the allegation.

To the extent possible, study directors must ensure that any collected anecdotal evidence meets the following rigorous standards: It is (1) collected from both minority and nonminority contractors; (2) not subject to response bias (i.e., arising from self selected study participants rather than a broadly representative sample) nor interviewer bias (i.e., responses the researcher encouraged through the wording of questions or informing the respondent of the political purpose of the inquiry); (3) appropriate in time and place; (4) directly related to discrimination in public contracting; (5) industry specific; (6) group specific; (7) highly detailed about specific instances of discrimination; (8) corroborated with other evidence; and (9) not anonymous (even though names are withheld from the report).

In verifying anecdotes, researchers should pose the following questions: Does the claimant have a vested interest in preserving a benefit or entitlement? Is the person providing the anecdote representative of a larger group or merely someone who feels most strongly about a problem? Investigators should distinguish between perception and reality; and obtain knowledge of the

³ See National Research Council, *Analyzing Information on Women-Owned Small Businesses in Federal Contracting*, Steering Committee for the Workshop on Women-Owned Small Businesses in Federal Contracting, Committee on National Statistics, (Washington, DC: The National Academies Press, 2005), pp. 7, 88–89 (hereafter cited as NRC, *Women-Owned Small Businesses in Federal Contracting*).

perspectives of both parties involved in an incident as well as information about the treatment of comparably placed persons of other races, ethnicities, and genders.

Recommendation 16: To ensure the highest quality research, states and localities should seek universities, many of which have public policy research centers, to conduct their disparity studies.

National Disparity Studies

Finding 3: The three national studies of disparities—the Department of Justice’s 1996 appendix to its guidance, the Urban Institute’s meta-analysis, and the Department of Commerce’s benchmark studies—are outdated and inappropriate now to serve as a basis for federal policy or agency action.

Recommendation 17: The government should conduct new research on disparities nationwide to support both its policy on discrimination in contracting and federal, state, and local jurisdictions’ efforts to narrowly tailor race-conscious programs. Federal officials must discard disparity studies conducted using data that is more than five years old. To justify the award of contracts, the government must use current data, not outdated results.

Recommendation 18: When conducting national research on disparities, study directors must apply all of the above recommendations offered for state and local efforts. For example, they should ensure that methodology follows generally accepted social science research standards; develop a rationale for including businesses as ready, willing, and able to carry out contract work; perform detailed analyses; and test for nondiscriminatory explanations of group differences.

Recommendation 19: When updating national disparity studies, the federal government should model its methodology on the best aspects of the Urban Institute and Department of Commerce studies, but not replicate problematic features of them such as failing to explain the effects of a capacity measure on results. It should, for example, follow the Urban Institute report in carefully specifying the methodology, providing thorough documentation, and conducting sensitivity analyses.

Recommendation 20: Researchers must carefully develop estimates of availability, including only firms that bid on or receive sole source contracts. Furthermore, the research must recognize that many firms certified for the 8(a) program are newly formed businesses still in a developmental stage. If analysts include all 8(a) certified businesses in the availability measure of a national study, they must present detailed characteristics of the firms. For example, what proportion of 8(a) businesses previously received contracts? What proportion previously bid on contracts? What proportion neither received nor bid on contracts? How did inclusion of 8(a) businesses that had no experience in bidding on or executing contracts affect disparities?

Recommendation 21: Any federal study of disparities should control for capacity in its availability measure. A report, widely accessible to interested parties, should provide detailed information on the capacity measure, including the resulting statistical equations (that is, its regression coefficients) and the sensitivity of the study results to its inclusion.

Recommendation 22: Any meta-analysis of state and local disparity studies must develop standards of scientific quality, use them to systematically weigh the merits of each local study, and discard (or relegate to separate tiers) work that fails to meet the criteria. The report should explain the standards it applies and indicate which studies surpass them, forming the basis of the meta-analysis.

Recommendation 23: As always, when combining results from many different studies in a meta-analysis, researchers should calculate disparities separately for the number of contract awards and dollar amounts.

Recommendation 24: Researchers performing meta-analyses must compute multidimensional results to demonstrate outcomes in different industries, regions of the country, minority (or other disadvantaged) groups, time frames, and types of contracts. Additionally, they should address firm qualifications, sizes, or specializations of minority- versus nonminority-owned enterprises.

Recommendation 25: National disparity studies also must meet high standards of validity, reliability, and reproducibility of results. Researchers must thoroughly document all data sources, methods, evaluations, and findings.

Research on Discrimination

Finding 4: Disparities are not necessarily evidence of discrimination. Yet, the federal government uses disparity studies to justify programs that provide preferences to minority- and women-owned businesses. Further research could identify either the existence of discrimination, and its nature, or nondiscriminatory causes of disparities that might lead to initiatives that could improve access to contracting for all enterprises.

Recommendation 26: To better understand discrimination and develop appropriate policy, Congress should direct the National Science Foundation to establish a research program to identify sources of disparities and provide grants that analyze discrimination.

The National Science Foundation should develop a research program that draws on case studies, administrative records, and statistical analyses. The analyses should examine earlier causal effects, for example, the process of developing pools of ready, willing, and able bidders. It should examine the occupational choices that ethnic and national origin groups make and analyze supplier industries to determine how they affect minority-owned businesses' competitive bids. In addition, the research agenda should look at the many qualitative differences that determine whether or not minority-owned firms win contracts, such as, their years in existence and owners' experience, knowledge of how to prepare business plans, and acumen regarding

procedures for obtaining bank loans. Research should also examine the criteria and selection process for contracts, including, for example, what weight decision makers give to price versus timely product delivery.

The research program should also pursue evidence of discrimination through other methods, such as the paired testing commonly used to measure housing discrimination or scientifically acquired anecdotal information. The Foundation may collect information from other studies that purport to measure unconscious racial bias, such as using a “resume test” to find employment discrimination or asking subjects to sort photographs of diverse people. However, any such research should articulate direct relevance to federal contracting and recognize known weaknesses of the approach.

Recommendation 27: The Small Business Administration should develop a research agenda on the contracting experience of small and disadvantaged businesses, including both prime and subcontractors. Such research should help identify the sources of discrimination in various stages of the contracting process and, to the extent feasible, in the process of business formation and development. To better carry out this research agenda, the agency should initiate data collection of information on subcontractors either through surveys or administrative records.⁴

Recommendation 28: Federal agencies should use any sources of disparities that research identifies to develop strategies to increase access to contracting opportunities for all types of businesses.

Guidance from the Department of Justice

Finding 5: The Department of Justice’s guidance on affirmative action in federal contracting, including its appendix surveying research that establishes a compelling interest for race-conscious programs is now a decade old. The business world experienced rapid growth and change since then. Requirements for narrowly tailoring race-conscious programs dictate that federal officials must justify preferences using current information, not evidence of discrimination that is now outdated.

Recommendation 29: The Department of Justice should offer agencies guidance on the design of federal programs to ensure that all firms have access to contracting opportunities. It should emphasize all of the available tools, including race-conscious programs, race-neutral approaches, and enforcement, for example, through complaints processing. It should warn agencies of the problems in using outdated disparity studies to narrowly tailor their race-conscious programs and clarify the nature of neutral alternatives that require no such justification.

⁴ See NRC, *Women-Owned Small Businesses in Federal Contracting*, pp. 7–8.

Finding 6: Two panelists agreed that today the private sector may discriminate more than government. One panelist suggested that the Supreme Court intends for the federal government to offer racial preferences to combat the impact of private sector discrimination.

Recommendation 30: The Department of Justice should advise federal, state, and local agencies on whether and how government can use racial preferences to disrupt the effects of discrimination in the private sector.

Finding 7: Social science is expensive. Even the best studies will exclude some factors that could provide nondiscriminatory explanations for differences between groups.

Recommendation 31: The legal system should not make the requirement for narrow tailoring so burdensome that agencies, whether federal, state, or local jurisdictions, cannot meet it.

COMMISSIONERS' STATEMENTS [TO BE ADDED] 

APPENDIX A: ACRONYMS USED IN THE REPORT

BBC	BBC Research and Consulting, a company that has conducted state and local disparity studies
CNSTAT	Committee on National Statistics, an arm of the National Research Council
DBE	Disadvantaged business enterprise
DOJ	U.S. Department of Justice
FDOT	State of Florida Department of Transportation
FY	Fiscal year
GAO	General Accounting Office, known since July 7, 2004, as the Government Accountability Office
MBE	Minority-owned business enterprise
MBE/WBE; MWBE; M/WBE; MFBE	Minority- or woman-owned business enterprise; minority- or female-owned business enterprise
MGT	MGT of America, Inc., a vendor responsible for conducting numerous state and local disparity studies
MSA	Metropolitan statistical area, a census data grouping for cities of at least 50,000 people or urbanized areas of at least 100,000 people, and the counties that include these areas
NAICS	North American Industry Classification System, an economic classification scheme adopted in 1997 to replace the Standard Industrial Codes (SICs)
NERA	National Economic Research Associates, a consulting firm that provides studies of disadvantaged business enterprises and helps state and local governments and other public agencies develop affirmative action programs for procuring goods and services from businesses owned by minorities and women
NRC	The National Research Council, part of the National Academies
OFCCP	U.S. Department of Labor's Office of Federal Contract Compliance Programs
Ph.D.	Doctor of Philosophy
SBA	U.S. Small Business Administration
SBO	Bureau of the Census' survey of business owners
SDB	Small disadvantaged business
SFUSD	San Francisco United School District
SIC	Standard Industrial Code, the industry classification scheme in use until 1997

SM/SWOBE	Survey of Minority-Owned Business Enterprises and Survey of Women-Owned Business Enterprises
SMOBE	Survey of Minority-Owned Business Enterprises
USDOT	U.S. Department of Transportation
WSDOT	Washington State Department of Transportation

APPENDIX B: PANELISTS' BIOGRAPHIES

IAN AYRES is the William K. Townsend Professor at Yale Law School. He was born and raised in Kansas City, Missouri, received his B.A. (majoring in Russian studies and economics) and J.D. from Yale and his Ph.D in economics from M.I.T. Professor Ayres clerked for the Honorable James K. Logan of the Tenth Circuit Court of Appeals. He has previously taught at Illinois, Northwestern, Stanford and Virginia law schools and has been a research fellow of the American Bar Foundation.

He is a regular commentator on public radio's *Marketplace* and a columnist for FORBES magazine and regularly writes opeds for the NEW YORK TIMES.

Professor Ayres has published 8 books and over 100 articles on a wide range of topics. In the spring 2005, he published three books, *OPTIONAL LAW: REAL OPTIONS IN THE STRUCTURE OF LEGAL ENTITLEMENTS* (University of Chicago Press 2005); *INSINCERE PROMISES: THE LAW OF MISREPRESENTED INTENT* (Yale University Press 2005) (with Gregory Klass); *STRAIGHTFORWARD: HOW TO MOBILIZE HETEROSEXUAL SUPPORT FOR GAY RIGHTS* (Princeton University Press 2005) (with Jennifer Gerarda Brown). He is also the author of *WHY NOT?: HOW TO USE EVERYDAY INGENUITY TO SOLVE PROBLEMS BIG AND SMALL* (2003) (with Barry Nalebuff); *VOTING WITH DOLLARS* (2002) (with Bruce Ackerman) and *PERVASIVE PREJUDICE?* (2001).

He is also the author of several empirical studies: *Does Affirmative Action Reduce the Number of Black Lawyers?*, *STANFORD LAW REVIEW* (forthcoming 2005) (with Richard Brooks); *To Insure Prejudice: Racial Disparities in Taxicab Tipping*, *YALE LAW JOURNAL* (forthcoming 2005) (with Fred Vars and Nasser Zakariya); *Shooting Down the More Guns, Less Crime Hypothesis*, 55 *STANFORD LAW REVIEW* 1193 (2003) (with John J. Donohue III); *Measuring the Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack*, 113 *QUARTERLY JOURNAL OF ECONOMICS* 43 (1998) (with Steven D. Levitt); *Pursuing Deficit Reduction Through Diversity: How Affirmative Action at the FCC Increased Auction Competition*, 48 *STANFORD LAW REVIEW* 761 (1996) (with Peter Cramton); *A Market Test for Race Discrimination in Bail Setting* (with Joel Waldfogel); *Racial Equity in Renal Transplantation: The Disparate Impact of HLA-Based Allocation*, 270 *JOURNAL OF AMERICAN MEDICAL ASSOCIATION* 1352 (1993) (with Robert Gaston, Laura Dooley and Arnold Diethelm); and *Fair Driving: Gender and Race Discrimination in Retail Car Negotiations*, 104 *HARVARD LAW REVIEW* 817 (1991).

Professor Ayres has been ranked as one of the most prolific and most-cited law professors of my generation.¹ Excerpts of his publications can be found on the internet at: www.law.yale.edu/ayres/.

CONSTANCE F. CITRO was named director of the Committee on National Statistics in May 2004. She is a former vice president and deputy director of Mathematica Policy Research, Inc., and was an American Statistical Association/National Science Foundation research fellow at the U.S. Census Bureau. From 1984-2004, she served as study director for numerous projects of the committee, including the Panel to Review the 2000 Census, the Panel on Estimates of Poverty for Small Geographic Areas, the Panel on Poverty and Family Assistance, the Panel to Evaluate the Survey of Income and Program Participation, the Panel to Evaluate Microsimulation Models for Social Welfare Programs, and the Panel on Decennial Census Methodology. Her research has focused on the quality and accessibility of large, complex microdata files, as well as analysis related to income and poverty measurement. She is a fellow of the American Statistical Association. She received a B.A. degree from the University of Rochester and M.A. and Ph.D. degrees in political science from Yale University.

ROGER CLEGG is President of the Center for Equal Opportunity. Mr. Clegg writes, speaks, and conducts research on legal issues raised by the civil rights laws. The Center for Equal Opportunity is a conservative research and educational organization based in Sterling, Virginia, that specializes in civil rights, immigration, and bilingual education issues. Mr. Clegg also is a contributing editor at *National Review Online*, and writes frequently for *USA Today*, *The Weekly Standard*, *The Legal Times*, *The Chronicle of Higher Education*, and other popular periodicals and law journals.

From 1982 to 1993, Mr. Clegg held a number of positions at the U.S. Department of Justice, including Assistant to the Solicitor General, where he argued three cases before the United States Supreme Court, and the number-two official in the Civil Rights Division and Environment Division. From 1993 to 1997, Mr. Clegg was vice president and general counsel of the National Legal Center for the Public Interest, where he wrote and edited a variety of publications on legal issues of interest to business. He is a graduate of Rice University and Yale Law School.

Mr. Clegg lives in Fairfax, Virginia, with his wife and son.

GEORGE R. LANOUE is Professor of Political Science in the Public Policy Graduate Program at the University of Maryland Graduate School, Baltimore. For eighteen years (1969-1997), he served as Program Director. Public Policy is one of the largest programs in the graduate school enrolling 230 Masters and Ph.D. students. Prior to coming to the University of Maryland, he was Director of the Teacher's College - Columbia University Graduate Program in Politics and

¹ See James Lindgren & Daniel Seltzer, *The Most Prolific Law Professors and Faculties*, 71 CHI.-KENT L. REV. 781 (1996); Fred R. Shapiro, *The Most-Cited Legal Scholars*, 29 J. LEGAL STUD. 409 (2000).

Education. He has taught at American University, the University of Chicago, the University of Miami, and the University of Strasbourg (France).

He graduated *magna cum laude* from Hanover College in 1959 and received his M.A. in 1961 and Ph.D. in 1966 in Political Science from Yale University. He has been awarded three national fellowships - the Woodrow Wilson, Danforth, and Public Administration.

Dr. LaNoue's research has focused on education law, civil rights law, and public policy. He has published four books and numerous articles. His book, *Academics in Court: the Consequences of Academic Discrimination Litigation*, University of Michigan Press, 1987, was funded by the Carnegie Corporation of New York. He completed a guidebook for local officials on disparity studies after *City of Richmond v. Croson* for the National League of Cities in 1991 and revised it in 1994. He also wrote a comparison of federal and state minority business programs for the *Annals of the American Academy of Political and Social Sciences*. He published pieces analyzing minority business programs after *Croson* in the *American Behavioral Scientist*, *The Public Interest*, *The Public Administration Review*, *The Journal of Policy History*, *Journal of Urban Affairs*, *The Public Historian*, *Population and Environment*, and the *American Behavioral Scientist*. He has published five law review articles on post-*Croson* law in the *Albany Law Review*, the *Columbia Human Rights Law Review*, the *Harvard Journal of Law and Policy*, *Santa Clara Law Review*, and *The Urban Lawyer*. He is the author of three encyclopedia entries: "School Decentralization," *Encyclopedia of Education*, "Political Science," *Encyclopedia of Educational Research*, and "Affirmative Action," *World Book Encyclopedia*.

A frequent witness in Congressional testimony, Dr. LaNoue is also a well-seasoned trial expert on civil rights cases in federal courts. He has been an Assistant to the Executive Director of the U.S. Equal Opportunity Commission and been the U.S. Department of Labor's principal trial expert in academic equal pay litigation. He has also served as consultant on a wide variety of educational and legal problems to the American Civil Liberties Union, the American Council of Trustees and Alumni, the Association of Governing Boards, the National Council of Churches, the U.S. Department of Education Office for Civil Rights, the United States Commission On Civil Rights, and several state governments and universities. He has been a consultant for the State of Texas and California, as well as the cities of Albuquerque, Nashville, Portland (Or.), St. Petersburg and West Palm Beach on post-*Croson* disparity studies. He has been the plaintiff's expert in cases involving disparity studies in Philadelphia, Columbus, Chicago, Cincinnati, Denver, Dade County, Cook County, Atlanta Public Schools, and Jackson, Miss. where MBE programs have been declared unconstitutional.

Dr. LaNoue has been listed in *Who's Who in American Law*, *Who's Who in the East*, *International Who's Who in Education*, and *American Men and Women of Science*. In 1992, Dr. LaNoue was the American representative to the International Experts Conference on "Legal Measures Against discrimination on Nationality, Ethnic, and Racial Grounds," hosted by the Commission on Foreign Affairs of the Senate of Berlin, Federal Republic of Germany. Sponsored by the U.S. Information Agency, the Swedish government, the German government,

the Fredrich Ebert Foundation, and others, Dr. LaNoue has had the opportunity to do research and lecture in fifteen countries.

U.S. Commission on Civil Rights
Washington, DC 20425

Official Business
Penalty for Private Use \$300

Visit us on the Web:
www.usccr.gov