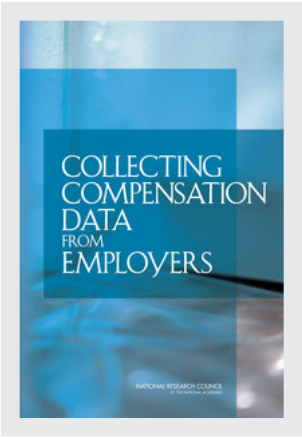

Collecting Compensation Data from Employers (2012)

DETAILS

154 pages | 6 x 9 | HARDBACK
ISBN 978-0-309-38755-2 | DOI 10.17226/13496

CONTRIBUTORS

Panel on Measuring and Collecting Pay Information from U.S. Employers by Gender, Race, and National Origin; Committee on National Statistics; Division of Behavioral and Social Sciences and Education; National Research Council

SUGGESTED CITATION

National Research Council 2012. *Collecting Compensation Data from Employers*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13496.

COLLECTING COMPENSATION DATA FROM EMPLOYERS

Panel on Measuring and Collecting Pay Information from
U.S. Employers by Gender, Race, and National Origin

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by the U.S. Equal Employment Opportunity Commission and the National Science Foundation under a grant to the National Academy of Sciences. Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (award number SES-1024012). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number-13: 978-0-309-26408-2
International Standard Book Number-10: 0-309-26408-1

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Printed in the United States of America

Suggested citation: National Research Council. (2012). *Collecting Compensation Data from Employers*. Panel on Measuring and Collecting Pay Information from U.S. Employers by Gender, Race, and National Origin. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PANEL ON MEASURING AND COLLECTING PAY INFORMATION FROM U.S. EMPLOYERS BY GENDER, RACE, AND NATIONAL ORIGIN

John M. Abowd (*Chair*), School of Industrial and Labor Relations, Cornell University
H. Juanita (Nita) Beecher, Employment Law & Litigation Group, Mercer LLC
Marc Bendick, Jr., Bendick and Egan Economic Consultants, Inc., Washington, DC
Charles C. Brown, Department of Economics, University of Michigan
Elizabeth Hirsh, Department of Sociology, University of British Columbia
Mark R. Killingsworth, Department of Economics, Rutgers University
Jonathan S. Leonard, Haas School of Business, University of California, Berkeley
Janice F. Madden, Wharton School, University of Pennsylvania
Aleksandra (Sesa) Slavkovic, Department of Statistics, Pennsylvania State University
Finis R. Welch, Welch Consulting, Bryan, TX
Valerie Rawlston Wilson, National Urban League Policy Institute, Washington, DC

Thomas J. Plewes, *Study Director*
Michael J. Siri, *Program Associate*

COMMITTEE ON NATIONAL STATISTICS
2011-2012

Contents

Preface

The U.S. Equal Employment Opportunity Commission (EEOC) collects detailed information on employment by gender and race/ethnicity by job groupings from all employers, except small employers. The agency does not collect earnings data from private employers. The only earnings data collected by EEOC are collected for employees of state and local governments, excluding school systems and educational institutions, and these earnings data are limited to major gender and race/ethnic groups for eight salary ranges. As a byproduct of the agency's enforcement programs, EEOC collects pay information during investigations of complaints and litigation, but it does not use the information collected in this manner to monitor pay trends in any structured way.

The Paycheck Fairness Act of 2009 (H.R. 12), which did not pass during the 111th Congress,[1] would have required EEOC to issue regulations to mandate data from employers to EEOC on pay by the race, gender, and national origin of employees. If the legislation had become law, EEOC would have confronted issues regarding currently available and potential data sources, methodological requirements, and appropriate statistical techniques for the measurement and collection of employer pay data.

At the suggestion of a White House Task Force, the EEOC asked the National Research Council, through its Committee on National Statistics (CNSTAT), to convene this panel to review methods for measuring and

[1]The legislation was reintroduced in both chambers in the 112th Congress. At this writing, the House version remains in committee while the Senate version failed to clear a procedural vote (to bring it up for floor consideration) on June 5, 2012.

collecting pay information by gender, race, and national origin from U.S. employers for the purpose of administering Section 709 of the Civil Rights Act of 1964, as amended. The panel was asked to consider suitable data collection instruments, procedures for reducing reporting burdens on employers, and confidentiality, disclosure, and data access issues.

In conducting this review, the panel held two workshops to gather information from data users and experts in survey methodology, wage and compensation concepts, and other methods for measuring and collecting pay information by gender, race, and national origin from U.S. employers. We particularly benefitted from papers and presentations provided by leadership and staff of EEOC, the Office of Federal Contract Compliance Programs (OFCCP) of the U.S. Department of Labor, and the U.S. Department of Justice. A paper on administrative sources of pay data was commissioned and is an appendix to this report.

The panel is grateful for the active participation of Sharon Alexander, Office of the Chair, EEOC, and Ronald Edwards, director, Program Research and Surveys Division, Office of Research, Information and Planning, EEOC, for their unhesitant cooperation with the panel during its work. Special thanks go also to Bliss Cartwright and Lucius Brown, who assisted in developing this study and in overseeing its progress on behalf of EEOC.

A large group of experts from government agencies, academia, and representing various other user organizations freely gave of their time to prepare presentations for the workshops and enter into a dialogue with the panel as it gathered information for this report.

The first workshop opened with statements by Stuart Ishimaru, commissioner, EEOC; Jocelyn Samuels, senior counselor to the assistant attorney general for civil rights, U.S. Department of Justice (DOJ); and Claudia Gordon, special assistant to the director of the OFCCP. Ron Edwards of EEOC and Pamela Coukos, senior program advisor, OFCCP, brought the panel up to date on currently available sources of equal employment opportunity and wage data. State and provincial programs that now collect earnings data by gender, race, and national origin were described by Martha Burk, formerly the senior adviser for women's issues to the governor of New Mexico, Faith Zwemke, director of the Pay Equity Office of Minnesota; and, in the second workshop, Stephanie McCleave, director of the Ontario, Canada Pay Equity Office. The general counsel of the EEOC, P. David Lopez, and three EEOC field office officials—Anna Park, regional attorney, and Rosa Viramontes, deputy regional attorney, of the Los Angeles District Office, along with Marla Stern-Knowlton, director of the San Diego Local Office—summarized the current enforcement and litigation uses of the EEO-1 data currently gathered by the agency. Bliss Cartwright of the EEOC Program Research and Surveys Division gave a presentation on national office uses of the EEO-1 data. Overviews of compensation concepts and

definitions were provided by Kevin Hallock, Cornell University, and Philip Doyle, assistant commissioner for compensation levels and trends, Bureau of Labor Statistics, Department of Labor.

In the second workshop, the panel heard from representatives of vendors who provided payroll and software products. Karen Minicozzi discussed the enterprise software offerings of Workday, Inc., Liz Balconi, consultant, and Michele Whitehead, manager of human resource services, Berkshire Associates, discussed the software that this firm uses to assist companies with understanding their equal opportunity profiles. A consultant to the panel, Nicholas Greenia, formerly of the Internal Revenue Service, gave a presentation on the availability of administrative data to yield earnings data useful for antidiscrimination purposes. A panel consisting of Ronald Edwards, EEOC; Gilberto Garcia, chief, Branch of Enforcement and Appeals, OFCCP; and Sharyn Tejani, special litigation counsel, DOJ, discussed issues of data confidentiality and data sharing.

The panel is grateful for the excellent work of the staff of CNSTAT for their support in developing and organizing the workshops and preparing this report. Tom Plewes, study director for the panel, ably supported the work of the panel. Michael Siri provided administrative support to the panel. We are especially thankful for the personal participation of Constance F. Citro, CNSTAT director, in the conduct of the workshops and in the preparation of this report.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the National Research Council. The purpose of this independent review is to provide candid and critical comments that assist the institution in making its reports as sound as possible, and to ensure that the reports meet institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

The panel thanks the following individuals for their review of the report: Frank Dobbin, Department of Sociology, Harvard University; Jon A. Geier, Employment Law Department, Paul Hastings, LLC; Kevin F. Hallock, Institute for Compensation Studies, Cornell University; Alan F. Karr, Director's Office, National Institute of Statistical Sciences; Barbara F. Reskin, Department of Sociology, University of Washington; and John H. Thompson, NORC at the University of Chicago.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of the report was overseen by Robert Michael, professor, Harris School, University of Chicago, and Michael Goodchild,

professor emeritus, University of California, Santa Barbara. Appointed by the National Research Council, they were responsible for making certain that the independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of the report rests entirely with the authoring panel and the National Research Council.

John M. Abowd, *Chair*
Panel on Measuring and Collecting Pay
Information from U.S. Employers by Gender,
Race, and National Origin

Summary

For identifying the possibility of discriminatory practices, the U.S. agencies with responsibilities for enforcing equal employment opportunity (EEO) laws have long relied, along with other sources, on detailed information that is obtained from employers on employment in job groups by gender and race/ethnicity. The U.S. Equal Employment Opportunity Commission (EEOC), the Office of Federal Contract Compliance Programs (OFCCP) of the U.S. Department of Labor, and the Civil Rights Division of the U.S. Department of Justice (DOJ) have developed processes that use these employment data as well as other sources of information to target employers for further investigation and to perform statistical analysis that is used in enforcing the antidiscrimination laws. The limited data from employers do not include (with a few exceptions) ongoing measurement of possible discrimination in compensation.

The proposed Paycheck Fairness Act of 2009 (H.R. 12) would have required EEOC to issue regulations mandating that employers provide the EEOC with information on pay by the race, gender, and national origin of employees. The legislation was not enacted. If the legislation had become law, the EEOC would have been required to confront issues regarding currently available and potential data sources, methodological requirements, and appropriate statistical techniques for the measurement and collection of employer pay data.

At the suggestion of a White House Task Force, EEOC asked the National Research Council through its Committee on National Statistics to convene a panel to review methods for measuring and collecting pay information by gender, race, and national origin from U.S. employers. The

Panel on Measuring and Collecting Pay Information from U.S. Employers by Gender, Race and National Origin considered suitable data collection instruments, procedures for reducing reporting burdens on employers, and issues of confidentiality protection and data access.

The panel concludes that the collection of earnings data would be a significant undertaking for the EEOC and that there might well be an increased reporting burden on some employers. The panel also concludes that there is, at present, no clearly articulated vision of how the data on wages could be used in the conduct of the enforcement responsibilities of the relevant agencies. The main purpose for which the wage data would be collected, as articulated to the panel by EEOC and OFCCP representatives, is for targeting employers for investigation regarding their compliance with antidiscrimination laws But beyond this general statement of purpose, the specific mechanisms by which the data would be assembled, assessed, compared, and used in a targeting operation are not well developed by either agency. An Advance Notice of Proposed Rulemaking (ANPRM), issued in August 2011 by OFCCP, posed relevant questions in seeking public comment on the development and implementation of a new compensation data collection tool. The ANPRM contained a set of 15 questions encompassing all aspects of the new tool. Questions put forth included which type of wage data to collect, appropriate job categories, the possibility of submitting data on an establishment basis, electronic data submission, etc.[1]

Nonetheless, the panel found no evidence of a clearly articulated plan for using the earnings data if they are collected. The fundamental question that would need to be answered is how the earnings data should be integrated into the compliance programs, for which the triggers for the EEOC and DOJ have primarily been a complaint process that has generated relatively few complaints about pay matters.

Furthermore, the panel concludes that existing studies of the cost-effectiveness of an instrument for collecting wage data and the resulting burden are inadequate to assess any new program. Unless the agencies have a comprehensive plan that includes the form of the data collection, it will not be possible to determine, with precision, the actual burden on employers and the probable costs and benefits of the collection. Therefore, the first recommendation is to develop such a plan.

Recommendation 1: In conjunction with the Office of Federal Contract Compliance Programs of the U.S. Department of Labor and the Civil Rights Division of the U.S. Department of Justice, the U.S. Equal Employment Opportunity Commission should prepare a comprehensive plan for use of earnings data before initiating any data collection.

[1]For the full set of questions in the ANPRM, see 76 FR 49398–49401.

The second recommendation stems from the panel's conclusion that existing evidence does not provide an adequate basis for determining the costs and benefits of the collection of wage data. Based on the data use plan, the panel recommends that a pilot study be conducted by an independent organization to provide much more reliable information about the costs and benefits of the proposed collection.

Recommendation 2: After the U.S. Equal Employment Opportunity Commission, the Office of Federal Contract Compliance Programs, and the U.S. Department of Justice complete the comprehensive plan for use of earnings data, the agencies should initiate a pilot study to test the collection instrument and the plan for the use of the data. The pilot study should be conducted by an independent contractor charged with measuring the resulting data quality, fitness for use in the comprehensive plan, cost, and respondent burden.

The panel offers two approaches to the recommended pilot study. The first pilot test—a microdata pilot approach—proposes collecting a number of core demographic variables (using the categories on the EEO-1 form) and adding an annual wage measure in order to test targeting firms for enforcement purposes. In addition, the pilot would test the collection of additional variables that are relevant to a firm's practices. For example, age and years-on-the-job variables could assist in controlling for the legitimate effect of these characteristics on wages.

The second approach—a simplified aggregated-data pilot—would develop and test an enhanced EEO-1 report that would include all the summary data required for the computation of test statistics comparing wage data within existing EEO-1 occupations. This pilot would use grouped data techniques that would produce standardized wage rates and other measures of interest. The end product would be a prototyped method for providing screening information about pay that is based on standardized information and audited test statistic formulas.

Both approaches to the pilot studies could also test various earnings definitions, such as those used in the Bureau of Labor Statistics' Occupational Employment Survey. The tests would assess the possibility of reducing employer response burden by using commercial electronic record-keeping systems in use in the larger companies. The quality of the data collected in the pilots would be independently verified by record checks or by comparison of aggregated results with administrative databases.

More needs to be done administratively to prepare the ground prior to commencing any data collection. EEOC has a small and lightly resourced data collection and analytical program that has traditionally been focused nearly exclusively on collecting employment data, developing summary

statistics, and assessing individual employer compliance through the means of rather straightforward statistical tests. If data on compensation are added to an existing form, or collected in a new instrument, it is likely that the resources for both collection and analysis in the agency would be severely strained. Thus, it is important that EEOC (and its partner antidiscrimination agencies) assess their capacity to undertake any new data collection and, when necessary, enhance their capacities to take full advantage of new opportunities for analytics and compliance, using the more sophisticated measures that will be possible.

Recommendation 3: The U.S. Equal Employment Opportunity Commission should enhance its capacity to summarize, analyze, and protect earnings data.

There are several possible means of collecting earnings information, ranging from pay bands (the clustering of pay levels method now used in the EEO-4 reports) to rates of pay. Pay band data are attractive in that they align with the way that human resource managers tend to look at compensation, but the best data are collected from payroll records, and those are most likely to be rates of pay or average earnings as computed with information on total wages and hours. Data on rates of pay have the advantage of being more likely to provide valid measures of central tendency and dispersion, thereby affording an important quality check and analytical capability. Rates of pay collection would add rigor to the collection process.

Recommendation 4: The U.S. Equal Employment Opportunity Commission should collect data on rates of pay, not actual earnings or pay bands, in a manner that permits the calculation of measures of both central tendency and dispersion.

It is important to use a definition of compensation that is measurable, collectable, and, in the end, meaningful. There are a number of definitions that are currently in use, ranging from comprehensive measures of total compensation to simple straight-time hourly pay. The panel concludes that a measure such as that used in the Occupational Employment Survey would best illuminate earnings levels. This measure has the added benefit of being generally available because earnings data by occupation are now collected with use of this definition from more than 1.2 million establishments.

Most of the firms that fall within the scope of the EEO statutes and are now required to complete an annual EEO-1 report have the ability to provide these data from their existing payroll and human resource systems. The growing penetration of highly sophisticated software-as-a-service applications into the marketplace will further enhance the ability of establishments

to provide earnings data by job group and gender, race, and national origin in the future.

Finally, the sensitivity of the data that employers provide to EEOC will be heightened if earnings data are added to EEO data records, since employee compensation data are generally considered to be highly sensitive, even proprietary information, by most employers. Therefore, it will be important for EEOC to develop more sophisticated techniques for protecting data that are provided in tabular and microdata form to the public.

Recommendation 5: In anticipation of increased user demand for microdata on pay information by demographic detail for research and analytical purposes if such data are collected by the U.S. Equal Employment Opportunity Commission, the agency should consider implementing appropriate data protection techniques, such as data perturbation and the generation of synthetic data to protect the confidentiality of the data, and it should also consider supporting research for the development of these applications.

In order to assure reporting employers that their data are indeed protected from disclosure, it will be important to establish clear and legally enforceable protections for sharing the data that employers provide in confidence. The agencies should consider whether the protections, now insured through the mechanism of interagency memoranda-of-understanding, should be incorporated in legislation.

Recommendation 6: The U.S. Equal Employment Opportunity Commission should seek legislation that would increase the ability of the agency to protect confidential data. The legislation should specifically authorize data-sharing agreements with other agencies with legislative authority to enforce antidiscrimination laws and should extend Title VII penalties to nonagency employees.

1

Background

The U.S. Equal Employment Opportunity Commission (EEOC) has a significant and active data collection program, which primarily collects information about employment status. While EEOC currently collects some pay data in its periodic reports from state and local government agencies for antidiscrimination enforcement, the agency has not collected pay data from private-sector employers, except on a case-by-case basis as necessary to support specific investigations. With that exception, the agency has no experience in collecting pay information from the private sector.[1] In its annual collection of data from private employers (EEO-1), the EEOC collects only employment classified by job category, gender, race, and national origin.

In this chapter, we briefly summarize relevant employment discrimination laws and describe the data that are currently collected in support of EEOC's enforcement program. We also describe the current roles and responsibilities of the key federal agencies that enforce those laws and that now use the EEOC data.

[1]The terms *pay*, *wages,* and *earnings* are used interchangeably in this report, depending on the context. They are taken to mean remuneration for labor or services to a worker on an hourly, daily, weekly, or annual basis or by the piece. The terms *salary* and *compensation* are also used in this report: *salary* is a fixed form of pay, wages, or earnings; compensation is the total amount of the monetary and nonmonetary pay provided to an employee by an employer in return for work performed, including money, benefits, services, and in-kind payments.

LEGISLATION, AUTHORITIES, AND RESPONSIBILITIES

Discrimination in pay on the basis of sex has been outlawed by the federal government for almost 50 years, since the Equal Pay Act of 1963. Enacted as an amendment to the Fair Labor Standards Act, the Equal Pay Act's coverage is very broad. It applies to any employer "engaging in commerce or in the production of goods for commerce" with an annual gross income of $500,000 or more (29 U.S.C. § 203(s)). Government entities and health and educational institutions are covered irrespective of size. There are narrow exceptions to coverage under the statute for certain kinds of employees (see 29 U.S.C. § 213(a)).

The Equal Pay Act requires that men and women in the same workplace be given equal pay for jobs "the performance of which requires equal skill, effort, and responsibility, and which are performed under similar working conditions" (29 U.S.C. § 206(d)(1)). Unequal pay between men and women for jobs that are substantially equal violates the act unless the employer can show that the difference in pay is attributable to a bona fide seniority, merit, or incentive system or another factor other than sex. Although the U.S. Department of Labor (DOL) was initially given authority to enforce the act, that authority was transferred to the EEOC in 1978.

Originally enacted one year after the Equal Pay Act in 1964, Title VII of the Civil Rights Act (hereafter, Title VII) prohibits a wide range of discriminatory employment practices, including discriminatory pay practices, and addresses discrimination based on sex, as well as race, color, religion, and national origin. Title VII covers private-sector employers with 15 or more employees and state and local government employers.

Under Title VII, an employee challenging pay discrimination must show that he or she is paid less than another similarly situated employee because of race, color, religion, sex, or national origin. If he or she does so, then the employer must explain the reason for the disparity. The employer may assert any of the defenses in the Equal Pay Act or a different, nondiscriminatory reason for the pay disparity. If the employer is unable to provide a satisfactory explanation for the disparity, the employer will be liable for penalties for pay discrimination. If the employer does provide a satisfactory reason for the disparity, the employee would have to show that the employer's stated reason is a pretext in order to succeed in proving pay discrimination.

Even where an employer does not intend to discriminate, a practice that is, on its face, neutral but that has the *effect* of disproportionately excluding or adversely impacting members of a protected group can violate Title VII. In such "disparate impact" cases, the individual alleging discrimination must prove—usually through statistical evidence—that the challenged practice has a substantial and significant adverse effect on a protected group.

If the individual proves this, the employer will be liable for discrimination unless it can show that the practice in question is job related and consistent with business necessity. If an employer can demonstrate that a practice is indeed justified, the individual will be given an opportunity to prove that there are other practices that would also serve the employer's purposes, but with less impact on the protected group.

Title VII's prohibitions on compensation discrimination are broader than those contained in the Equal Pay Act. For example, under Title VII, an employee can challenge not only unequal pay between men and women performing substantially equal work, but also discriminatory practices that lead to unequal compensation, such as steering women to lower paid jobs than men or maintaining "glass ceilings," artificial barriers to the advancement of women.

Title VII empowers the EEOC to accept and investigate charges of discrimination from persons who believe they have been subjected to employment discrimination and from those acting on their behalf. Title VII also allows for members of the commission itself to file charges of unlawful employment practices against employers. The EEOC is also empowered to open "directed investigations" under the Equal Pay Act, thereby allowing the EEOC to investigate the possibility of a violation of the act without having received a charge of discrimination from an aggrieved person.

Individuals must exhaust their administrative remedies through the EEOC prior to filing a lawsuit under Title VII. But under the Equal Pay Act, while aggrieved persons may file charges of discrimination with the EEOC, they are not required to do so in order to file a lawsuit under the act. Moreover, filing a charge under the act with the EEOC does not suspend the statute of limitations under the Equal Pay Act, as it does under Title VII. For this reason, and in light of the significant time it can take to exhaust administrative remedies through the EEOC, some aggrieved individuals find it preferable to file a lawsuit under the Equal Pay Act without filing a charge with the EEOC.

Under both Title VII and the Equal Pay Act, the EEOC investigates charges of discrimination and seeks to resolve them without litigation. However, the EEOC litigates a number of charges in which conciliation has failed each year. Under Title VII, the EEOC can litigate cases against private employers; charges against state and local governmental entities have to be referred to the U.S. Department of Justice (DOJ) for litigation. Under the Equal Pay Act, the EEOC may litigate against any covered employer, private or public.

In fiscal 2010, a total of 99,922 charges were filed, many for multiple allegations of discrimination (U.S. Equal Opportunity Commission, 2010). Relatively few of those multiple allegation charges included wage discrimination; far fewer were for wage discrimination only: see Table 1-1. The

TABLE 1-1 Charges Filed with U.S. Equal Employment Opportunity Commission, by Issue: October 1, 2009, to September 30, 2010

		Basis for Charge	
Issue	Total Charges	Race/National Origin	Gender
Total charges in which wage discrimination was an issue	4,478	2,314	2,164
Charges alleging only wage discrimination	638	282	356
Percent of wage discrimination charges in which wage discrimination was the only allegation	14.3%	12.2%	16.5%

SOURCE: Data from U.S. Equal Opportunity Commission.

majority of wage charges also involved other issues, most commonly terms and conditions of employment, termination, promotions, or discharges.

The Employment Litigation Section of the DOJ's Civil Rights Division is also charged with the enforcement of Title VII of the Civil Rights Act. Specifically, DOJ has jurisdiction to enforce Title VII against state and local government employers nationwide. DOJ can initiate litigation under Title VII in two ways: (1) DOJ has independent authority to bring suit against a state or local government employer when there is reason to believe that a "pattern or practice" of discrimination exists; or (2) DOJ may investigate and file suit against a state or local government employer based on an individual charge of discrimination referred by the EEOC. DOJ can initiate such a suit if the EEOC has found reasonable cause to believe that discrimination occurred, the EEOC's efforts to obtain voluntary compliance have been unsuccessful, and EEOC has referred the charge to DOJ.

The Office of Federal Contract Compliance Programs (OFCCP) in DOL is responsible for making certain federal contractors follow requirements in the Executive Order 11246 (issued in 1965) to practice equal opportunity and take affirmative action on issues of race and gender.[2] In addition, OFCCP is responsible for enforcing Section 503 of the Rehabilitation Act of 1973, covering persons with disabilities, and the Vietnam Era Veterans' Readjustment Assistance Act of 1974 (VEVRAA), covering veterans and disabled veterans. Under these laws, federal contractors must provide equal employment opportunities and take affirmative action to

[2]In addition to race and sex, Executive Order 11246 (originally implemented in 1965) addresses equal opportunity on the basis of religion, color, and national origin.

employ and advance applicants and employees; provide reasonable accommodations to disabled employees and applicants; prepare Affirmative Action Plans (AAPs); permit OFCCP access during compliance reviews; and file an annual report with the EEOC.[3]

OFCCP regulations require contractors to maintain records on employee compensation and provide them on request (41 C.F.R. § 60-1.12(a), covering records on "rates of pay or other terms of compensation"). The regulations also require contractors to "regularly" monitor their compensation systems for potential pay disparities based on race and gender, develop and implement appropriate corrections to any problem areas they identify, and report the results of their internal monitoring to management (41 C.F.R. § 60-2.17). This language apparently requires federal contractors to maintain data on earnings by demographic characteristics.

EEOC DATA COLLECTION AND REPORTS

The various laws and regulations to enforce antidiscrimination laws are accompanied by laws and regulations for the federal government to collect data that can be used in their enforcement. The EEOC uses its authority under Section 2000e-8(c) of Title VII to collect workforce data from employers. The statute requires employers to preserve "records relevant to the determinations of whether unlawful employment practices have been or are being committed," and to "make such reports therefrom as the Commission shall prescribe by regulation or order, after public hearing, as reasonable, necessary, or appropriate for the enforcement of [Title VII] or the regulations or orders thereunder."

The EEOC currently collects workforce data from private-sector employers with more than 100 employees, from federal contractors with 50 or more employees, and from all state and local government employers. Employers that meet the reporting thresholds have a legal obligation to provide the data; it is not voluntary. The data are collected through several equal employment opportunity (EEO) reports.

There are four versions of the required EEO reports, each addressed to different employer groups. Each of the versions collects employment data about gender and race/ethnicity by some type of job grouping; each provides, in essence, a snapshot of the demographics of the workplace by job category. Copies of these report forms are provided in Appendix A.

[3]The application of each of these requirements may vary on the basis of contract size and number of employees.

TABLE 1-2 EEO-1 Reports by Number of Employees Covered and Percentage Female and Minority, 2010

Size of Firm	Number of Firms	Number of Employees	Percentage Female	Percentage Minority
TOTAL	67,422	59,128,582		
5th percentile: 1–67 employees[a]	3,443	191,965	38.6	26.8
5th–25th percentile: 68–120 employees	13,511	1,312,297	41.7	29.6
25th–50th percentile: 121–194 employees	16,875	2,587,008	45.6	31.1
50th–75th percentile: 195–407 employees	16,767	4,615,048	46.6	32.6
75th–90th percentile: 408–1,118 employees	10,090	6,541,695	47.4	33.6
90th percentile and higher: more than 1,118 employees	6,736	43,880,569	50.0	34.8

[a]Includes only firms with at least 50 employees.
SOURCE: Data from U.S. Equal Employment Opportunity Commission (2010 EEO-1 Aggregate Report of U.S.).

EEO-1 Report

The EEO-1 report is required from private employers with: 100 or more employees or 50 or more employees and a federal contract. Firms must file a separate report for each facility with 50 or more employees. Approximately 67,000 firms filed more than 1.3 million EEO-1 reports (Standard Form 100) in 2009 and 2010. For 2010, the reports covered 59 million employees, which is almost one-half of the 108 million employees for all firms in the private sector. The largest 10 percent of covered firms represented about 75 percent of covered employment, and covered firms with 120 employees or less represented only about 2.5 percent of covered employment: see Table 1-2.

Employers are required to file the EEO-1 report annually to the EEO-1 Joint Reporting Committee (joint between EEOC and OFCCP) (due on September 30). The data elements that are collected include 7 race/ethnicity categories and 10 job groups, by gender.[4] Employers may use employment

[4]The race/ethnicity categories are Hispanic or Latino, and—under not Hispanic or Latino—white, black or African American, Native Hawaiian or Other Pacific Islander, Asian, American Indian or Alaska Native, and two or more races. The 10 job groups are executive/senior level officials and managers; first- or mid-level officials and managers; professionals; technicians; sales workers; administrative support workers (formerly, office and clerical workers);

figures from any pay period in July through September. Employers may submit their EEO-1 reports on paper forms, as data files[5] by electronic transfer, or by keying the data online through the EEO-1 online filing system. About 99 percent of the reports are received electronically.

There are different types of reports for single establishment employers and multiple establishment employers. Multiple establishment reports must include a consolidated form that includes all employment for the company, one for headquarters locations, and one for each location with 50 or more employees. Locations with fewer than 50 employees are required to report only the address and total number of employees at that establishment, rather than a complete matrix.

EEO-3 Report

The EEO-3 report form is used for referral unions, which are generally unions with exclusive hiring arrangements with an employer. The report is required in even-numbered years with a due date of December 31. The EEO-3 form collects data on membership and referrals by race/ethnicity and gender. In 2010, there were about 1,200 reporting unions. The reports are used for enforcement and provide basic membership and referral data for investigators. They also allow statistical analyses to examine equity in membership and referrals.

EEO-4 Report

The EEO-4 report form is used for state and local governments. It is required in odd-numbered years and is due on September 30. Approximately 6,000 jurisdictions filed EEO-4 reports in 2009. The reports that year covered 5,980,305 employees.

This is the only EEO report that now collects employment data by job group and salary ranges for race/ethnicity and gender, with separate reports by function. Data are also collected separately for part-time employees and new hires. (See Chapter 3 for discussion of the wage data that are collected in this survey.)

craft workers (formerly, craft workers, skilled); operatives (formerly, operatives, semiskilled); laborers and helpers (formerly, laborers, unskilled); and service workers.

[5]The files are sent as ASCII/text files, a simple data transfer that does not use developing techniques such as XML.

EEO-5 Report

The EEO-5 report form is used for primary and secondary public school districts. It is required in even-numbered years with a due date of November 30. For 2010, the EEOC received more than 5,800 of these reports. The data are collected from each school district with 100 or more employees by race/ethnicity and gender for relatively detailed job groups.[6] EEO-5 data are also collected for part-time employees and for new hires.[7]

WHITE HOUSE TASK FORCE REPORT AND PANEL CHARGE

Following President Obama's pledge in the 2010 State of the Union address to increase enforcement of equal pay laws, the administration established the National Equal Pay Enforcement Task Force in 2010, bringing together EEOC, DOJ, DOL, and the Office of Personnel Management (OPM). The task force identified several challenges to successful enforcement of compensation discrimination laws and made recommendations to address each challenge. Three of the five challenges identified by the task force have implications for this report:

- Three different federal agencies have distinct responsibilities to enforce the laws prohibiting pay discrimination, and the agencies do not consistently coordinate these responsibilities.
- The government's ability to understand the full scope of the wage gap and to identify and combat wage discrimination is limited by the data that are currently available. As the task force report says, "this lack of data makes identifying wage discrimination difficult and undercuts enforcement efforts. We must identify ways to collect wage data from employers that are useful to enforcement agencies but do not create unnecessary burdens on employers" (National Equal Pay Task Force, 2010, p. 5).
- Existing laws do not always provide federal officials with adequate tools to fight wage discrimination. The task force report noted the administration's strong support for the Paycheck Fairness Act, which would have required EEOC to use its data collection

[6]The job groups are officials, administrators, and managers; principals; teaching assistant principals; nonteaching assistant principals; elementary classroom teachers; secondary classroom teachers; other classroom teachers; guidance staff; psychological staff; librarians and audiovisual staff; consultants and supervisors of instruction; other professional staff; teacher aides; technicians; clerical and secretarial staff; service workers; skilled crafts; and unskilled laborers.

[7]Part-time employees are grouped by professional, instructional, and all other; new hires are grouped by officials, administrators, managers, principals and assistant principals, classroom teachers, other professional staff, and nonprofessional staff.

BOX 1-1
Statement of Task

The National Research Council through its Committee on National Statistics (CNSTAT) will convene a panel of experts to review methods for measuring and collecting pay information from U.S. employers for the purpose of administering Section 709 of the Civil Rights Act of 1964, as amended. The panel will evaluate currently available and potential data sources, methodological requirements, and appropriate statistical techniques for the measurement and collection of employer pay data. The panel will consider suitable data collection instruments, procedures for reducing reporting burdens on employers, and confidentiality, disclosure, and data access issues. It will issue a report with findings and recommendations on what data the EEOC should collect to enhance wage discrimination law enforcement efforts, which will assist the Equal Employment Opportunity Commission (EEOC) in formulating regulations at the conclusion of an 18-month study.

authority to implement a pay data collection program within 18 months of its enactment. Specifically, the bill text would require EEOC to "consider factors including the imposition of burdens on employers, the frequency of required data collection reports (including which employers should be required to prepare reports), appropriate protections for maintaining data confidentiality, and the most effective format for the data collection reports."[8] The Paycheck Fairness Act would also have amended the Equal Pay Act to prohibit employers from retaliating against employees for discussing their pay.[9]

The EEOC charge to the panel stressed that it is important for the panel to bear in mind the key considerations about the balance between enforcement utility and burdens on employers. Regardless of the fate of the Paycheck Fairness Act, the EEOC wants to ensure that any effort to collect wages takes into full account the considerations expressed in the Act regarding burden on employers, confidentiality, and appropriate format for collection. The complete statement of task is in Box 1-1.

[8]This text is from the 112th Congress version of the bill, S. 3220.

[9]The legislation passed the House of Representatives in 2009 but then failed in a cloture vote in the Senate in 2010. It has since been reintroduced in both chambers in the 112th Congress, with the Senate version (S. 3220) failing a cloture vote in June 2012.

PAY RATE INFORMATION

The employment data collected by EEOC are currently used for a variety of purposes, including enforcement, self-assessment by employers, and research. The EEOC's current collection of employment data contributes significantly to the efficiency of EEOC investigations and particularly to the development of systemic investigations. However, in a statement submitted to the panel, EEOC chair Jacqueline A. Berrien stated that the agency sees the absence of "employer-specific pay data broken down by demographic category" as a "significant barrier" to the agency's work to eradicate pay discrimination. Berrien contrasted pay discrimination, a form of discrimination she described as "largely invisible," with other forms of discrimination that are easier to detect and that EEOC can more easily confirm or refute through the use of its current data collections.

Many workplaces explicitly prohibit employees from discussing pay, and even in the absence of an explicit prohibition, employees in the United States rarely discuss their pay with one another. Because very few people know what their coworkers are paid, few people file complaints with the EEOC alleging that they are being paid in a discriminatory manner. In his testimony to the panel, EEOC commissioner Stuart Ishimaru pointed out that sex-based wage charges have made up a surprisingly small portion of the charges EEOC has received—about 2.5 percent.

Berrien contended that, in addition to strengthening the EEOC enforcement program under Title VII and the Equal Pay Act, better pay data collection would also assist employers in monitoring their compliance with federal, state, and local laws prohibiting wage discrimination. By maintaining accurate pay data, Berrien said, employers will be able "to compare and identify pay differentials that deserve closer scrutiny and to detect other patterns that may suggest departures from the standard of equal pay for equal work."

EARNINGS INFORMATION

Use by OFCCP[10]

OFCCP officials similarly argued for the collection of earnings information in a presentation to the panel. Under the authorities discussed above, federal contractors must provide equal employment opportunities, take affirmative action to employ and advance their employees, and make reasonable accommodations to employees and applicants.

[10]This section summarizes a presentation to the panel by Pamela Coukos, senior program advisor, OFCCP.

A major requirement imposed on certain covered federal contractors is to develop an AAP. To meet this requirement, contractors must maintain appropriate records by establishment or function. The AAP data requirements cover the following topics: an organizational profile; a job group analysis; and information on placement of incumbents, determining availability, and comparing incumbency to availability. The AAP should spell out placement goals and designate an individual responsible for implementation. Problem areas need to be identified and action-oriented programs specified, and the plans need to be audited periodically.

The AAP instructions call on employers to group jobs by similar pay and work content and to classify them into an appropriate EEO category based on similar duties and responsibilities, as well as similar opportunities for training, transfer, pay, and promotion, and similar jobs in lines of progression. An example of an AAP workforce analysis is shown in Table 1-3.

The OFCCP has minimum employee and contract size requirements for federal contractors[11] and different rules for construction contractors. For example, construction contractors with federal contracts or subcontracts valued at more than $10,000 in any 12-month period are covered by Executive Order 11246 at all construction worksites in the United States (Office of Federal Contract Compliance Programs, 2009).

The enforcement activities of OFCCP primarily involve full compliance reviews. These reviews begin with desk audits of information submitted by a contractor in response to a scheduling letter, and they may also include an onsite review. Contractors are identified as being subject to enforcement activities based, in part, on a system called the Federal Contractor Selection System (FCSS). This system draws information from the universe of EEO-1 reports and federal contractor databases. Using these data sources, OFCCP selects contractors based on threshold requirements, sampling procedures, and mathematical modeling.

An OFCCP compensation analysis consists of an initial review of average pay differences for job categories. The agency then performs a statistical or individual analysis as appropriate (depending on sample size and available data) and further review and analysis based on contractor pay practices and data. These data are used to assess the company's practices. The investigation is designed to answer some basic questions: Are there pay differences between employees in a protected class and otherwise similar

[11]Basically, all federal contracts and subcontracts are covered under Executive Order 11246 unless specifically exempted. Contracts and subcontracts of less than $10,000 generally are exempt, though some contracts under that amount are covered (e.g., bills of lading). Also exempt is work performed outside the United States; certain contracts with state or local governments; contracts with religious corporations, associations, and educational institutions; and contracts involving work on or near an Indian reservation. See 41 C.F.R. § 60-1.5.

TABLE 1-3 Example of an Employer's Workforce Analysis for an Affirmative Action Plan

Workforce Analysis

DEPARTMENT/WORK UNIT: Administration					MALES						FEMALES					
Job Title	Wage Rate	EEO-1 Category (EEO-1 Form or OFCCP regulations)	Job Group	Total Employees	Total	White	Black/African American	Asian/Pacific Islander	American Indian/Alaskan Native	Hispanic	Total	White	Black/African American	Asian/Pacific Islander	American Indian/Alaskan Native	Hispanic
General Manager	S-A	1	1	1	1	1										
Personnel Manager	S-D	1	1	1	1	1										
Executive Assistant	S-J	5	5	1							1	1				
Administrative Assistant	H-8	5	5	1							1		1			
File Clerk	H-11	5	5	2	1					1	1				1	
DEPARTMENT TOTAL				6	3	2				1	3	1	1		1	

SOURCE: U.S. Department of Labor, available: http://www.dol.gov/ofccp/regs/compliance/pdf/sampleaap.pdf [July 2012].

employees? Are there differences in salary/hourly rate, promotions, job assignment, and access to earning opportunities? Are there legitimate explanations for any differences?

At the time this report was being prepared, OFCCP was considering a new compensation reporting tool that would proactively allow the agency to more effectively identify potential violations of Executive Order 11246. The agency has requested public input on the kind of instrument that could be used for this purpose. This initiative is discussed in Chapter 2.

Use by DOJ[12]

According to Jocelyn Samuels, senior counselor to the assistant attorney general for civil rights of DOJ, the department uses data, including pay data, gleaned from the EEO-4 reports to fulfill its responsibilities under antidiscrimination statutes. The "pattern or practice" cases initiated based on the department's independent authority under Title VII, Samuels told the panel, "are factually and legally complex cases that seek systemic injunctive relief to alter unlawful employment practices—such as discriminatory recruitment, hiring, assignment or promotion policies—which have the purpose or the unjustified effect of denying employment or promotional opportunities to a class of individuals." DOJ may also investigate and file suit against a state or local government employer based on an individual charge of discrimination referred by the EEOC, as described above.

The department routinely consults and relies on the information included in the EEO-4 reports regarding workforce composition and new hires, in combination with other information, to determine whether or not to use its enforcement jurisdiction to investigate a specific state or local government employer. Specifically, the department relies on EEO-4 reports for data on the demographics of different job categories in an employer's workforce to assist in deciding whether to pursue investigations of allegations that may constitute a "pattern or practice" of discrimination. The EEO-4 information enables comparisons of an employer's workforce in a particular job category to an applicable benchmark—such as civilian labor force data in the relevant geographic area taken from census and survey sources—to determine whether a particular group appears to be underrepresented in that job category or in the employer's workforce as a whole. The comparison provides a basis to estimate whether there is a disparity in representation in the workforce and to make an initial assessment of the significance of the disparity, which is one factor that informs the department's

[12]Statement of Jocelyn Samuels to the Panel on Measuring and Collecting Pay Information from U.S. Employers by Gender, Race, and National Origin Workshop, May 24, 2011.

evaluation of whether to open an investigation in order to gather more detailed information from an employer.

In her presentation to the panel, Samuels stated that the demographic data collected on the EEO-4 reports are invaluable for enforcement purposes, but the wage data on the form are currently less useful. The job categories and the wage bands reported on the EEO-4 form are too broad, and the current EEO-4 form does not include any other information, such as longevity (years of service), which can be a key determinant of salary in the public sector.

In order to allow meaningful analysis, the department needs salary information in narrower job classes and information about years of service in the job class. In addition, according to Samuels, salary information should be collected in narrower bands, and should, to the extent possible, reflect the entire amount earned, not solely base pay. State reports suggest that these data are readily available in many states.[13]

In addition, DOJ has recently executed a memorandum of understanding with the EEOC in order to obtain access to EEO-1 data for private employers. DOJ anticipates that it will use these data in enforcement efforts for comparison purposes in job categories that exist in both the public and private workforce.

Use for Analysis and Research

In their presentations to the panel, the representatives of the EEOC, OFCCP, and DOJ emphasized the enforcement purposes behind the collection of data from employers and unions. However, by virtue of their depth and coverage, these data also have statistical, analytical, and research uses.

EEOC publishes annual statistical summaries of employment data from the EEO-1 and EEO-4 reports, as well as information received from federal government departments and agencies, on its website in three series: *Job Patterns for Minorities and Women in Private Industry* (EEO-1); *Job Patterns for Minorities and Women in State and Local Government* (EEO-4); and *Federal Sector Reports*. The employment data by race/ethnicity and sex are published by industry, geographic area (state and local areas), and job category.

As part of an emphasis on proactive prevention, EEOC's Office of Research, Information, and Planning has produced a series of reports based on EEO-1 data. The reports over the past decade have focused on industries and sectors (the finance industry, retail distribution centers, the media, high-end department stores, investment banking, broadcasting, and law firms)

[13]For example, see information from the Florida Bureau of State Payrolls, available at: http://www.archive.org/details/StateOfFloridaPayrollDatabase2008 [July 2012].

as well as on particular labor market topics, including *How New Business Processes Impact Minority Labor Markets*; *Women of Color: Their Employment in the Private Sector*; *Glass Ceilings: The Status of Women as Officials and Managers in the Private Sector*; and *Characteristics of Private Sector Employment Report*.

A major use of the employment data is in the context of charge-based investigations, in which the data are used to assist EEOC in identifying employers that warrant statistical comparisons, which could, in turn, trigger further investigation of their EEO practices. For example, using the EEO-1 reports of the numbers of employees in the establishment(s) in a certain job group and gender, race, and ethnic category, EEOC staff calculate a number of indicators that are designed to assess the EEO status of the firm. Those indicators include

- *Actual number*: The reported number of employees in a particular job group and gender, race, and ethnic category.
- *Expected number*: The number of employees that would be expected to exist in that certain job group and gender, race, and ethnic category according to the percentage employed by comparison establishments that have been selected based on specified geographic and industrial scope.
- *Difference*: The difference between the actual number and expected number of employees in a certain job group and gender, race, and ethnic category. If the difference is positive, the establishment is over the expected number; if it is negative, the number of employees in that category is below the expected number—a difference that is often referred to as a "shortfall."
- *Actual percent*: The percentage of employees in a certain job group and gender, race, and ethnic category.
- *Expected percent*: The percentage of employees that would be expected in that certain occupational and gender, race, and ethnic category based on that percentage in comparison establishments.
- *Two-tail probability*: A binomial statistical significance test, which is used to determine if the differences between the actual and expected numbers are statistically significant.

Administratively, EEOC primarily uses the EEO-1 data to identify potential discriminatory practices in the context of an investigation of a charge and to otherwise support investigations. The EEO-1 data are used in different ways at different stages of the investigation, and the analysis becomes more refined as the investigation progresses.

In a presentation to the panel, Bliss Cartwright of the EEOC Office of Research, Information, and Planning discussed these uses, selecting as a

hypothetical example a comparison of gender employment in one firm to employment in similar firms in the labor market. In his example, the firm had 180 female professionals of 624 total professionals, about 29 percent: in contrast, the proportion of female professionals in the labor market was 40 percent. He assumed that the labor market percentage was estimated by aggregate EEO-1 data on other firms in similar industries and locations, and he applied a one-sample binomial test of statistical significance. The main characteristics of this hypothetical example can be summarized as follows:

- Total Professionals: 624
- Female Professionals: 180
- Observed Proportion: 0.2885
- Labor Market Proportion: 0.4067
- Null Hypothesis: No Difference
- Two-Tailed Probability: <0.0000 (less than one chance out of 10,000)
- Conclusion: Substantial Evidence Against Null Hypothesis of No Difference in Proportions

Other situations may require more refined analyses. For example, sometimes a national firm has many facilities, hiring workers for the same job in different local labor markets. Alternatively, a single firm may recruit executives from a national market, midlevel managers from a regional market, and operatives from a local market. The issue is that there are multiple units of analysis, each with different employee counts and labor market estimates. In these situations, other statistical methods might be more appropriate. For example, Cartwright illustrated one approach commonly known as a pooled binomial (Gastwirth and Greenhouse, 1987), which provides an estimate of the overall shortfalls giving a single probability value. It also allows examination of homogeneity, the extent to which the units of analysis differ from each other.

The next step in an analysis is to seek additional information from an employer through a request for information (RFI) that is tailored to the potential infraction alleged in the charge. For hiring issues, for example, EEOC typically requests files with demographic information, applicant flow data, and job history records. The requested data may be extensive. The job history information typically contains the effective date of the hire or the action that distinguishes initial hires from rehires or returns by use of employee identification numbers. The requested records also include specific job titles, divisions, and salary grades. At this stage, a wide variety of statistical methods would be considered—including linear regression,

survival analysis, and stratified contingency tables—depending on the facts and issues in a particular case.

Understanding the Labor Market

Since collection of information about employment by gender, race, national origin, and job category was initiated on a regular basis in the 1970s, there has been intense interest by the academic community in using the data to understand labor markets, especially the effect of governmental programs and corporate human resource practices on employment discrimination. EEO-1 reports and enforcement data from the OFCCP have been used to examine the effect of affirmative action and other factors on the employment of minorities and women across different sectors of the economy.

Selden (2006) assessed a variety of studies that transcended disciplines,[14] pointing out that most use the EEO-1 survey data to examine the impact of affirmative action on minority and female employment shares among firms with or without federal contracts in the private sector. Selden summarized work by Leonard (1990) that concluded that affirmative action led to employment gains among women and minorities for the period 1974–1980 and rose more significantly for federal contractors than for noncontractors. Selden (2006, p. 915) concluded that "overall, studies using EEO-1 data have shown that affirmative action has significantly and positively influenced the minority employment share in the private sector, particularly in unskilled positions."

Although there have been difficulties in obtaining access to EEOC's survey data, the agency has made significant efforts to increase the access that researchers have to these data. Since 1996 the EEOC has entered into agreements with more than 35 researchers to allow access to these confidential databases. Much of this work has been published in peer-reviewed articles and books, which in many cases have raised new questions and topics for academic research. In economics, for example, Donohoe and Levitt (2001), McCrary (2007), and Miller and Segal (2011) examined the relationship between diversity and crime rates using EEO-4 data. In sociology, Dobbin, Kalev, and Kelly (2006) examined how personnel practices impact a firm's work force diversity, particularly in management. These researchers also examined the impact of OFCCP compliance reviews and Title VII lawsuits on employment profiles (Dobbin, Kalev, and Kelly, 2007; Kalev and Dobbin,

[14]Selden's assessment covered Ashenfelter and Heckman (1976); Chay (1998); Goldstein and Smith (1976); Holzer and Neumark (2000a, 2000b); Kellough (1990a, 1990b); Leonard (1984a, 1984b, 1990); Naff (2001); Naylor and Rosenbloom (2004); Rodgers and Spriggs (1996); Stephanopoulos and Edley (1995); and U.S. General Accounting Office (1991).

2009), and Kalev (2009) examined how work restructuring impacts occupational segregation based on race and gender.

A wide range of other work has also been done. Several researchers compared firm-level and sector-level changes in segregation by race, ethnicity, and sex (Stainback and Tomaskovic-Devey, 2009; Stainback, Robinson, and Tomaskovic-Devey, 2005). Huffman, Cohen, and Pearlman (2010) studied the impact of women managers on firm gender integration for the period 1975–1990. Skaggs (2008) studied how government action, including court decisions, affected female employment in food stores. Several other researchers explored the impact of various factors, including EEOC charge processing, on the employment of women and nonwhites (Hirsh, 2008, 2009; Hirsh and Kmec, 2009; Hirsh and Kornrich, 2008). Yet another group of researchers used EEO-4 data for a series of articles examining diversity in state and local governments including an examination of glass ceilings among those employers (Kerr, Reid, and Miller, 1999, 2000a, 2000b, 2002, 2003, 2004). All of this research has been done even with the difficulties of obtaining access to the data, which are discussed in Chapter 5, and in the absence of compensation data.

In the absence of employer-based earnings data by job category and demographics, however, the research community largely turned to household data to support analysis of the extent and effect of compensation discrimination in the labor market. The Current Population Survey and, more recently, the American Community Survey have emerged as powerful sources of data on earnings, industry groups, occupations, and demographics. However, these sources are limited because they do not associate the indicators of discrimination with actual employer situations and practices, nor can they be directly linked to measures of enforcement.

There is clearly a strong research and analytical interest in having an earnings dimension to establishment, occupation, and demographic data (see, e.g., Consad Research Corporation, 2009, p. 2). It is expected that there would be significant pressure on agencies that held data enriched with earnings information to make them available for analytical uses by private sector researchers. Such data could quickly become a primary source for new analytic work on equal employment and compensation issues.

Auditing the Effectiveness and Efficiency of Antidiscrimination Programs

Over the years, Congress and a number of government agencies have used data collected on EEO-1 forms to assess the effectiveness of government antidiscrimination programs. Just as the research community would benefit from the availability of earnings data, these agencies would be expected to take advantage of earnings information to sharpen their auditing reports.

The U.S. Government Accountability Office (GAO), in particular, has been at the forefront in terms of using employment data by job category and demographics. In the past two decades, GAO has published seven major studies that have been based in part on the EEOC employment data:

- *Sharing Promising Practices and Fully Implementing Strategic Human Capital Planning Can Improve Management of Growing Workload* (2008);
- *Financial Services Industry: Overall Trends in Management-Level Diversity and Diversity Initiatives* (2006);
- *Equal Employment Opportunity: The Policy Framework in the Federal Workplace and the Roles of EEOC and OPM* (2005);
- *Women's Earnings: Work Patterns Partially Explain Difference between Men's and Women's Earnings* (2003);
- *Equal Employment Opportunity: Discrimination Complaint Caseloads and Underlying Causes Require EEOC's Sustained Attention* (2000);
- *Equal Employment Opportunity: DOL Contract Compliance Reviews Could Better Target Federal Contractors* (1995); and
- *EEOC: An Overview* (1993).

Cross-Checking the Integrity of EEO Data

An additional justification for the collection of pay data is that they may help to improve the integrity of EEO employment data. Smith and Welch (1984) found some evidence that the number of minorities and women reported to be in high-level occupations by their employers on EEO-1 forms exceeded the number who reported themselves to be in those occupations in the Current Population Survey. To the extent that some employers of minority- or female-intensive occupations systematically upgrade (or misclassify) them, it would cause unusual pay compression across EEO-1 job categories and unusual pay dispersion within the higher level occupations. Being able to make such assessments by using pay data would be valuable for evaluation purposes.

2

Alternative Sources of Wage Data

The charge to this panel included a request to "evaluate currently available and potential data sources" for measuring and collecting pay information from U.S. employers for the purpose of administering Section 709 of the Civil Rights Act of 1964. We begin our response to this part of the charge with a discussion of the collection of earnings data from public-sector employers on the EEO [equal employment opportunity] form 4, or EEO-4. Indeed, the Equal Employment Opportunity Commission (EEOC) has some experience from which to draw when considering the collection of earnings data because the agency now collects wage band information on the EEO-4 form.

We also discuss other possible sources of wage information and the experiences of other agencies in collecting such information.[1] We first consider the capacity of existing federal administrative data series that include earnings information from employers to meet a requirement for wage information by gender, race, and national origin. If these administrative data, mostly from tax collections, could suffice to provide the necessary wage data for use in antidiscrimination enforcement, a new data collection process could be avoided. Unfortunately, as discussed in this chapter, the use of administrative data is not a promising path because of data incompleteness and uncertain quality.

[1]This report does not assess another data source that has appeared recently in which individual employees self-report pay by employer, occupation, and location on a variety of websites; these self-postings sometimes include pay stubs. These self-reports are not a random sample, offer little or no demographic information, have variable or in many cases no coverage of occupations, and are difficult to verify.

We then consider the experience of the Office of Federal Contract Compliance Programs (OFCCP) of the U.S. Department of Labor (DOL) with collection of earnings information on a trial basis a decade ago. The lessons learned in that experiment should be considered by EEOC as it considers collecting earnings information.

We also discuss the data collection programs of the states of New Mexico and Minnesota and the Canadian province of Ontario. These jurisdictions now gather earnings information from employers for pay equity purposes. We assess the potential of these collections to inform an EEOC decision on whether and how to collect earnings information.

Finally, we consider survey-based wage information and discuss three Bureau of Labor Statistics (BLS) surveys—the Current Employment Statistics (CES) Survey, the National Compensation Survey (NCS), and the Occupational Employment Statistics (OES) Survey. These surveys can inform the collection of wage data and provide a source of potential validation information for data series that could be collected by EEOC, but we do not judge them to be suitable sources for the wage data for EEO enforcement purposes. They do not collect data by gender, race, or national origin; they are covered by strict confidentiality provisions, which limit their use for enforcement; and they do not cover all establishments covered by EEO laws and executive orders.

DATA FROM EEO-4 REPORTS

As noted in Chapter 1, EEO-4 reports are collected in odd-numbered years from state and local governments: in 2009 approximately 6,000 jurisdictions filed EEO-4 reports that covered 3,238,769 employees. The report collects employment data by job group and salary ranges for race/ethnicity and gender, with separate reports by function (e.g., streets and highways, health, corrections). Data are also collected separately for part-time employees and new hires.

The EEO-4 report is the only one that collects any wage-related data. It collects annual salaries by job category for eight pay bands:

1. $1,000 to $15,999
2. $16,000 to $19,999
3. $20,000 to $24,999
4. $25,000 to $32,999
5. $33,000 to $42,999
6. $43,000 to $54,999
7. $55,000 to $69,999
8. $70,000 and over

The pay band data are collected for eight job categories:

1. officials and administrators
2. professionals
3. technicians
4. protective service workers
5. paraprofessionals
6. administrative support
7. skilled craft workers
8. service and maintenance workers

The wage data collected on this report have some limitations, according to EEOC commissioner Stuart Ishimiru, who addressed the panel on May 24, 2011. The form requests wage data by race, ethnic origin, and gender, but the wages are reported in broad intervals that do not allow for precise comparisons. Similarly, according to the commissioner, the job categories for which wages are reported are so broad that they are rarely if ever used to conduct wage disparity analyses. Despite these limitations, the reports are used extensively by the U.S. Department of Justice (DOJ) for administrative and enforcement purposes. Academic institutions use these reports for self-assessment purposes.

ADMINISTRATIVE DATA

The federal government and state agencies now collect a massive amount of wage data from employers and maintain them in the form of administrative records of three tax systems. Two of these systems are administered by federal agencies—the Internal Revenue Service (IRS) and the Social Security Administration (SSA)—and one by state unemployment insurance agencies under the auspices of the DOL's Employment and Training Administration (for details, see Greenia, Appendix B of this volume). The three administrative data systems are used primarily to collect taxes and determine benefits for the purposes of administering and funding the federal income tax system (by the IRS), the Social Security and Medicare programs (by SSA), and the joint state-federal unemployment insurance (UI) system.

The data are used by the programs that collect them for purposes of enforcement of their own laws and regulations. In select circumstances, federal legislation has also authorized use of these data for enforcement purposes in other programs. For example, a new hires database derived from UI filings is used by the Administration for Children and Families in the U.S. Department of Health and Human Services to facilitate finding employed parents who are not making required child support payments

under the Personal Responsibility and Work Opportunity Reconciliation Act of 1996.[2]

National compilations of statistics are produced from the three sets of data by the pertinent statistical offices of IRS and SSA, as well as the Bureau of Labor Statistics (BLS).[3] In addition, the data are used for policy analysis by the Joint Committee on Taxation of Congress, the Congressional Budget Office, and the Office of Tax Analysis in the U.S. Department of the Treasury. The data are also used for analysis by academic researchers, through the Intergovernmental Personnel Act, as well as through the U.S. Census Bureau's Research Data Centers. Table 2-1 summarizes the availability of items from each of these administrative records sources.

According to Greenia (Appendix B of this volume), the three sets of data are interrelated. For example, the three tax-based systems depend on the social security numbers (SSNs) assigned by SSA, the employer identification numbers (EINs) assigned by IRS, the reporting of employment and payroll at both the firm and individual worker level for federal and state purposes, and other information from the administrative systems, such as changes in name and address, to update the records.

The IRS has the duty to determine which workers are employees and which are contractors. "The IRS decision is obtained by the filing of a Form SS-8 for a firm or worker seeking to have IRS establish officially the employee or independent contractor status of a particular worker. This transaction then has ramifications for the other employee data collection systems that are mandated by such legislation as the State Unemployment Tax Act (SUTA) and the Federal Unemployment Tax Act (FUTA)" (Greenia, Appendix B of this volume).

Thus, although only the SSA system has data on earnings by gender, race, and national origin (items needed for enforcement purposes), it is possible, by virtue of their coverage and interrelationships, to link data from the three tax systems so that each of them could produce some data on employee earnings by gender, race and ethnicity, nativity, and age, by employer. These data could be used to inform EEOC's enforcement programs, although they most likely could not be used directly in enforcement actions.

[2]For details, see http://www.acf.hhs.gov/programs/cse/newhire/library/ndnh/background_guide.htm [July 2012].

[3]IRS data are primarily published by the Statistics of Income Division of IRS: see http://www.irs.gov/taxstats/productsandpubs/article/0,,id=125133,00.html [July 2012]. SSA data are published by the Office of Retirement and Disability Policy: see http://www.ssa.gov/policy/docs/statcomps [July 2012]. And BLS data are published in the Quarterly Census of Employment and Wages series: see http://www.bls.gov/cew/cewbultn10.htm [July 2012]. The Census Bureau also uses these data sets as input to several of its statistical programs.

TABLE 2-1 Available Items in Administrative Records Relevant to EEO

Source	Earnings at Employee Level	Identity of Employer	Employee Gender	Employee Race/Ethnicity	Employee Nativity
State Unemployment Insurance	YES	YES	NO	NO	NO
State Employment Security Agency	NO	YES	NO	NO	NO
Internal Revenue Service	YES	YES	NO	NO	YES[a]
Social Security Administration	YES	YES	YES	YES	YES

[a]Only from individual taxpayer identification number (ITIN) applications.
SOURCE: Adapted from Greenia (Appendix B of this volume).

State Unemployment Insurance Data

In addition to complying with the Federal Unemployment Tax Act, employers must also comply with the State Unemployment Tax Act by withholding and depositing tax or insurance payments from each employee's wages with state unemployment offices. These state unemployment taxes fund unemployment benefits in each state or territory (including the District of Columbia, Puerto Rico, and the Virgin Islands; see Greenia, Appendix B of this volume).

This section presents a brief summary of the UI wage records and the Quarterly Census of Employment and Wages (QCEW) Program that draws on them. It discusses how the UI data are reported, collected, and shared with the federal government, and assesses the potential usefulness of these data for EEO enforcement purposes.

UI tax rates and coverage vary by state, as do the content and format of the records a particular state collects. In general, all workers are covered by the UI system with the exception of federal employees, independent contractors, the self-employed, and some agricultural workers. A state collects detailed employment and compensation data in quarterly reports from each employer. The data include the SSN, name, and quarterly compensation for each individual employee, as well as the employer name and EIN.[4] The products of this collection are known as UI wage records.

State employment security agencies also collect aggregate monthly employment (for the pay period containing the 12th of the month) for each quarter and aggregate quarterly employee compensation from each employer in the state covered by state UI laws and for federal workers covered by the Unemployment Compensation for Federal Employees (UCFE) Program. This data collection program, the QCEW, is administered and partially funded by BLS.

Although states request data from employers at the establishment level for multiple worksites or multi-establishment employers, there is no disincentive for an employer that does not comply with the request as long as total employment is reported accurately and the appropriate amount of UI tax is paid to the states (Greenia, Appendix B of this volume).

In considering wage data for purposes of EEO enforcement, the UI data system provides the earnings data needed and at the employee level, but it also has several shortfalls:

- It is difficult, if not impossible, to disaggregate the data from multi-establishment employers to the worksite level to match with the EEO-1 reports (see Chapter 1).

[4]The coverage varies by state; see Stevens (2002) for a complete review.

- There are no gender, race and ethnicity, or nativity data collected for UI wage records, though there have been instances in which demographic data from other sources, such as driver's licenses files, have been associated with the wage records (Glover, 2011; Moore, 2011) to enable analysis of UI wage information by gender. As discussed below, it would be possible to match these records to SSA demographic data.
- In order to obtain either of the two data components provided to the states by employers—especially the detailed employee earnings—it would be necessary to obtain the data directly from employers (who would submit a copy of their UI filings to EEOC) or to enter into separate agreements with each state, and it is likely that both of these actions would require a legal action.

Internal Revenue Service Data

Since 1976, when the current simplified Combined Annual Wage Reporting (CAWR) program was established by the Tax Reform Act, employers have reported individual earnings statements and the amount of taxes withheld (including federal income tax, Social Security tax, and Medicare tax) on a single form (Form W-2 Wage and Tax Statement) for both IRS and SSA purposes. The earnings details available from the W-2 are rich: wages and salaries, deferred compensation (part of total compensation, even if not taxable currently), and certain fringe benefits are reported, in addition to capped Social Security earnings and uncapped Medicare earnings. Together, the W-2 earnings variables provide a unique and comprehensive window on earnings data at the employee level.

These individual W-2 forms are transmitted with another form (Form W-3, Transmittal of Income and Tax Statements), which cumulates the information from the W-2 forms for each reporting establishment. Because of this arrangement, it would be possible to obtain detailed annual employee compensation, quarterly and annual aggregate employee compensation, and number of employees at both the employee and employer level with links to Social Security information through an SSN and EIN crosswalk. The industry codes available at SSA, in full North American Industry Classification System (NAICS) levels, can provide a further source of rich classifier information on employers' business activities. In addition, other tax forms can provide various components of aggregate and even detailed employee compensation: for example, compensation to corporate officers. Finally, EIN and individual taxpayer identification numbers (ITINs) assignment and other transactions would enable the tracking of new business births, foreign-born workers without SSNs, and even the employee or contractor status of a worker.

An employer is required to file an annual FUTA tax return (Form 940)[5] for purposes of reporting and paying the federal unemployment taxes required by FUTA. Filing is required—at the aggregate employment level—for each nonagricultural employee earning at least $1,500 in any quarter of the year or for each employee who was employed for part or all of a day in any 20 different weeks of the year.[6] Although Form 940 does report annual total compensation, it does not report the number of employees. However, the compensation information may be useful for benchmarking compensation data reported on other federal tax forms, such as Form W-2 and Form 941, as well as the UI data.

In summary, IRS data include a wealth of earnings information for individual employees and employers. However, a limitation is that the IRS data include establishment data only when the establishment is also an enterprise (and has an EIN). Another limitation is that the tax data contain no information by gender (except, sporadically, for the IRS Statistics of Income Division individual Form 1040 tax sample), race and ethnicity, or nativity (except for ITIN applications).

Social Security Administration Data[7]

The data of most interest for examining pay equity issues are the demographic data that are available on the application for a Social Security Number (Form SS-5),[8] which can be linked to federal tax data shared by IRS. The application for an SSN captures gender, race and ethnicity, and nativity—often shortly after birth for most U.S. citizens. In addition, it captures citizenship status, which might be used as a proxy for or to supplement nativity information.

Although the Form SS-5 data are self-reported (by the individual or a parent), SSA uses supporting documentation for verification, particularly for changes, such as a marriage license (name), passport (citizenship), and birth certificate (place of birth). The Form SS-5 data, including updates, are maintained in SSA's Numerical Identification System file, referred to as the Numident file.

Despite the richness of the demographic detail, the Numident file data have some limitations. They are not updated as often as tax information for such changes as name and address due to marriage or divorce (the tax

[5]The form is available at: http://www.irs.gov/pub/irs-pdf/f940.pdf [December 2011].

[6]For 2009 and 2010, agricultural employers were required to file if they paid cash wages of $20,000 or more to farm workers during any calendar quarter or if they employed 10 or more farm workers during some part of the day (whether or not at the same time) during any 20 or more different weeks in either year.

[7]Information in this section is based largely on Greenia (Appendix B of this volume).

[8]This form is available at: http://www.ssa.gov/online/ss-5.pdf [July 2012].

information at IRS may be updated before the Numident data). In addition, although nativity data classified by country might be considered relatively reliable, researchers have noted that some of the "foreign born" may be, in fact, the progeny of U.S. citizens, say, for military and other Americans stationed overseas, where birth occurs. In conjunction with citizenship status, however, the data are probably useful for indicating native versus foreign-born status.

EQUAL OPPORTUNITY SURVEY PILOT

In order to identify federal contractors with potential problems of pay discrimination that could warrant further review or evaluation by OFCCP or to support a contractor self-audit, OFCCP has long been interested in developing a screening tool to enable the agency to identify supply and service contractors whose compensation data indicate that further investigation is warranted. This interest led to initiation of a pilot survey to collect earnings data with demographic and job group information from federal government contractors. An employer survey was developed and undertaken by the OFCCP. The OFCCP experience is instructive for EEOC as it considers collecting wage information by gender, race, and national origin.

As discussed in Chapter 1, the authority for this collection rests in Executive Order 11246, as amended, which requires that federal government contractors and subcontractors "take affirmative action to ensure that applicants are employed, and that employees are treated during employment, without regard to their race, color, religion, sex, or national origin." Affirmative action under the executive order requires that contractors take affirmative steps to identify and eliminate impediments to equal employment opportunity. The affirmative steps include numerous record-keeping obligations designed, first, to assist the contractor and then OFCCP in monitoring the contractor's employment practices.[9]

In the early 2000s, the OFCCP listed three objectives for the survey (71 FR 3374):

1. to improve the deployment of scarce federal government resources toward contractors most likely to be out of compliance;
2. to increase agency efficiency by building on the tiered-review process already accomplished by OFCCP's regulatory reform efforts, thereby allowing better resource allocation; and
3. to increase compliance with equal opportunity requirements by improving contractor self-awareness and encourage self-evaluations.

[9]For full text of Executive Order 11246, as amended, see http://www.dol.gov/ofccp/regs/statutes/eo11246.htm [July 2012].

Field testing for the survey of federal contractors to collect wage information, as well as other new data items, was conducted in 1999. In 2000, OFCCP issued a requirement that nonconstruction contractor establishments designated by OFCCP prepare and file the new Equal Opportunity Survey. On a pilot basis, in April 2000, the EO Survey was sent to 7,000 contractors. One part of the survey (Part C) collected data on monetary compensation (expressed as an annual amount) and on tenure for four groups—minority females, nonminority females, minority males, and nonminority males—by the EEO-1 report categories applicable at that time: (1) officials and managers; (2) professionals; (3) technicians; (4) sales workers; (5) office and clerical workers; (6) craft workers; (7) operatives; (8) laborers; and (9) service workers. The questionnaire instructions defined annual monetary compensation as "an employee's base rate (wage or salary), plus other earnings such as cost-of-living allowance, hazard pay, or other increment paid to all employees regardless of tenure on the job, extrapolated and expressed in terms of a full year."[10] The annual monetary compensation measure was not to include the value of benefits, overtime, or one-time payments, such as relocation expenses.

The survey did obtain annual monetary compensation information—98.3 percent of respondents provided a numerical response to the compensation item. Reported median average annual compensation by gender and occupation appeared to be "broadly consistent" with other well-established data sets, such as the decennial census, the Current Population Survey, and other salary surveys (Bendick, 2000, p. 9).

After receipt of pilot survey responses, OFCCP commissioned a study to determine whether the pilot survey results could be used to predict whether a contractor would have findings of noncompliance. The study concluded, based on the first wave of survey responses, that the survey could contribute to improvements in procedures for selecting establishments for compliance evaluations (Bendick, 2000, p. i).

The OFCCP proceeded with the EO Survey that was sent to contractors beginning in December 2000 and continuing to December 2004. It included information, in summary form, about personnel activities, compensation, and tenure, as well as the contractor's affirmative action program. A total of 53,000 forms were sent.

To assess the quality and usefulness of these data, the OFCCP engaged an outside contractor to evaluate the collection to that point. The evaluation criteria were based on comparisons of survey results with the results of OFCCP compliance evaluations of a sample of supply and service contractor establishments that had completed the 2002 EO Survey. The

[10]U.S. Department of Labor form, available at: http://www.management-advantage.com/media/eosurvey.pdf [July 2012].

comparison study focused on 1,888 establishments that had completed compliance reviews and had reliable EO Survey data. Of these 1,888 cases, OFCCP found systemic discrimination in 67 cases (3.5 percent). Results of the compliance reviews and survey data were analyzed to determine whether a model could be developed that would predict which contractors in the sample were engaged in systemic discrimination based solely on the EO Survey data submitted (Abt Associates, Inc., 2005, pp. 23–37).

Based on that evaluation, OFCCP concluded that the EO Survey did not improve deployment of enforcement resources toward contractors most likely to be out of compliance and did not lead to greater self-awareness or encourage self-evaluations. OFCCP further concluded that the information in the survey largely duplicated information gathered in compliance visits, although that finding does not necessarily undercut the potential value of the survey, given that the purpose of such a survey is to obtain similar data as those gathered in compliance visits for the purpose of targeting enforcement resources.

The evaluation also found that the EO Survey imposed a burden on respondents. Each survey form was estimated to take each respondent 21 hours to complete. Based on an estimated 10,000 respondents per year, the EO Survey was estimated to cost contractor establishments 210,000 hours per year. Using data from the BLS 2004 National Compensation Survey, the total annual cost imposed on the regulated community by the survey requirements was close to $6 million. However, whether this level of burden was large or small compared with other regulatory requirements was not established, nor was the burden considered in relationship to the costs to employees of instances of wage discrimination that the survey might help uncover.

OFCCP's bottom-line conclusion was that the EO Survey had failed to provide the utility anticipated when the regulation was promulgated in 2000, and, consequently, it eliminated the survey. Reinstatement of the EO Survey, or the establishment of a similar survey, would require regulation or legislation.

Yet Bendick (2006) pointed out that the data on which both the Bendick (2000) and Abt Associates, Inc. (2005) studies drew had limitations that make it difficult to reach definitive conclusions about the value—or lack of value—of the EEO Survey for targeting enforcement resources. For example, the survey data were contaminated by the fact that compliance evaluations were conducted for many of the employers in the sample before the survey, so that employers had the opportunity to improve their practices by the time the survey was fielded.

The current OFCCP initiative is summarized by the National Equal Pay Enforcement Task Force (2010, p. 5):

> Through publication of an Advance Notice of Proposed Rulemaking (ANPRM) [in 2010],[11] OFCCP [has sought] the input of stakeholders in evaluating whether the EO Survey should be redesigned to collect different data than previously sought, and whether there are any ways to further limit the burden of data collection for employers. The implementation of a redesigned survey is expected to result in better identification of those contractors who are likely to be out of compliance, particularly with regard to compensation discrimination; a narrowing of the issues on which the resulting review will focus; and identification of contactors for corporation-wide and industry-focused reviews.[12]

U.S. STATE AND CANADIAN PROVINCIAL SURVEYS

On their own initiative, two U.S. states and a Canadian province have designed and fielded data collections to provide information on earnings by demographic characteristics for pay equity purposes. Though small and localized and based on comparable worth comparisons and thus not fully responsive to the needs of the EEOC, these initiatives provide experience in evaluating the feasibility of collecting wage information for antidiscrimination enforcement purposes. The lessons from these initiatives are summarized in this section.

Earnings Data Collection in New Mexico

As reported by Martha Burk in a presentation to the panel, in 2003 the New Mexico state legislature created a Pay Equity Task Force to study wage disparities between men and women and between minorities and nonminorities in both the public and private sectors. The task force issued a report with numerous recommendations, and, in January 2009 Governor Bill Richardson issued an executive order declaring pay equity a priority for the state.[13]

In September 2009, the Pay Equity Task Force issued a report addressing pay gaps and job segregation in the state workforce and in the workforce of state contractors. With respect to the latter, the report provided a rationale and model for requiring entities receiving state contracts to submit pay gap reports as a condition of contracting.[14] In December 2009, Governor Richardson issued a second executive order directing that

[11]The ANPRM was published as 76 FR 49399.

[12]The report further notes on p. 5 that "the EEO Survey has been rescinded, and its reinstatement, or the establishment of a similar survey, must be by regulation or legislation."

[13]For details, see *New Mexico Pay Equity Initiative*, available at: www.generalservices.state.nm.us/spd/pay_e.html [July 2012], p. 2.

[14]Available at: http://www.generalservices.state.nm.us/spd/report093009.pdf [July 2012].

a contractor reporting system be implemented and appointing a working group to design and facilitate such reporting.[15]

The new requirements were phased in gradually and, beginning July 1, 2010, all recipients of state contracts were required to submit a gender pay equity report after a contract was awarded. More recently, these reports have been required to be submitted as part of the response to a solicitation or request for proposal (RFP). The reporting requirement applies to all state agencies that let contracts and all categories of purchasing.

In her testimony to the panel, Burk stated that the early experience with collecting these data is noteworthy for the lack of resistance by employers and the absence of reports of difficulties in complying. About 3,200 firms are covered by the reporting requirement, ranging in size from Intel (New Mexico's largest employer, with more than 3,000 employees) to firms with only 10 employees. Over the first 7 months of implementation, fewer than 50 contractors contacted the state for assistance in understanding the requirement or preparing their reports.[16] Contractors seem generally to have accepted the requirement as a normal part of the contracting process and have lodged no complaints about the requirement.

As of early 2011, cataloguing of contractor reports was in progress, and systematic analysis had not begun. However, cursory examination of selected reports indicated that a majority of reporting employers have employees in only three or four of the nine EEO-1 form job categories provided for on the reporting forms, and a few (e.g., janitorial companies) have employees in only one job category besides the owner/manager. Moreover, job segregation by gender is not unusual, in which case gender pay disparities cannot be computed because of lack of wage data for both genders. When pay comparisons can be made, percentage gender pay gaps tend to range as low as 2–3 percent, with most in the 10–25 percent range. An unusual few were observed as high as 45 percent.

The report uses EEO-1 job categories for reporting because many contractors and payroll processing firms are already familiar with these categories. Using them avoided the need for a new taxonomy and also avoided the difficulties of analyzing data for job titles or groupings that were not comparable across firms.

The data each contractor is required to report consist of the number of employees by gender (including full- and part-time workers) in each EEO-1

[15]Executive Order 2009-049, available at: http://www.generalservices.state.nm.us/spd/pay_e.html [July 2012].

[16]In addition to offering "live" assistance by telephone, the state provides easy access on a website (see http://www.generalservvices.state.nm.us/spd/pay_e.html [July 2012]) to documents, including worksheets with instructions and reporting forms with instructions.

job category and the gender pay gap (stated as a percentage) in each category. Individual compensation is not reported.

Uniform reporting is enhanced by the fact that average hourly wages by gender and job category, taking into account hours worked, are computed following detailed instructions for producing these averages and entered into a worksheet. (Average hourly wages in each job category are computed by dividing the total compensation by gender by the total hours worked by that gender.)

Employers generally enter the appropriate numbers in the worksheet by use of an accounting/payroll system that is capable of classifying employees and aggregating compensation and hours worked by gender and job category. If employers do not have such a system to classify employees by job category, gender, time worked, and compensation levels, the state has provided an alternative downloadable employee data entry spreadsheet for performing the necessary calculations "from scratch." Using this spreadsheet, employers enter employee identification, job category, gender, whether full or part time, total annual compensation, and total hours worked for each employee.

Standard formulas for computing the gaps are embedded in required spreadsheets, which are provided to employers online, and the results of the computations are exported to a standard final report format. To maintain confidentiality of the wage and gender information, contractors do not turn in worksheets showing dollar amounts, but instead report only the ratios of average earnings for women to those of men in the same EEO-1 job category. Proprietary information is retained by contractors. However, they are encouraged to use this information for internal tracking of potential compensation disparities between women and men.

The Office of the State Auditor has oversight over state agencies' implementation of the reporting requirement and over the reports submitted. Procedures for auditing are still under development, and no audits have yet taken place. To date, the requirement is simply to submit a report; however, bids that fail to comply with the reporting requirement are disqualified.

Minnesota Pay Equity Survey

Since 1982 the state of Minnesota has had a pay equity law for state employees based on the concept of comparable worth. The law was extended to all city, county, school, and other public jurisdictions by the 1984 Local Government Pay Equity Act. The 1984 law requires each local government jurisdiction to use a job evaluation system to determine comparable work value and to submit a report to the state government at 3-year intervals with a comparable worth value estimate (the value of work as measured by skill, effort, responsibility, and working conditions)

and the minimum and maximum monthly salary for the job class by gender. The state uses the reports to assess how well the local jurisdictions are complying with the state law. Pay equity laws in Minnesota address only gender-based wage disparities.

Ontario Pay Equity Survey

A relatively recent pay equity survey in Ontario, Canada, collects wage information from the province's relatively large employers. The collection was enabled by a 2009 amendment to the Ontario Pay Equity Act, which provides that the Pay Equity Office may collect information for the purpose of providing reports to the Minister of Labour. The program was launched in January 2011. It involves canvassing all Ontario employers using a simple form that employers populate to provide current compensation data: see Table 2-2. The raw data are submitted to the Pay Equity Office, which assesses the data to determine if a wage gap exists.

For the first phase, Ontario workplaces with more than 500 employees were selected to enter the program. Using lists of Ontario employers developed by an external provider, employers who had not been visited by a review officer in the past 10 years and were not unionized were requested to submit current, basic wage data on the positions and incumbents in their organizations.

The information is analyzed to determine whether a wage gap appears to exist against a set of criteria developed by seasoned review officers. A committee of review officers meets monthly to review the analyses and findings and finalize the assessment. The employer of the establishment is advised whether the review officer determined there is an apparent wage gap and is provided with tools and information to allow the employer to consider whether they are pay equity compliant.

TABLE 2-2 Ontario Pay Equity Form

Job Title/ Position	Employee (may use symbol rather than name)	Held by: Male/Female	Pay as of December 31, 2010—hourly, weekly, annually	Salary Range of Position (if applicable)	Years of Service

The response rate for the first mailings was about 80 percent. This relatively high response rate was attained by significant follow-up efforts; in addition, response is encouraged because nonresponders are singled out for investigation through a proactive monitoring program for compliance.

SURVEY-BASED WAGE INFORMATION

In this section we consider the experience of surveys that collect wage information, specifically, three BLS employer surveys: the National Compensation Survey (NCS), the Current Employment Statistics (CES) survey, and the Occupational Employment Statistics (OES) survey.

We do not discuss Census Bureau data sources. The Census Bureau's business surveys have extensive establishment coverage but do not collect wage or demographic information. The decennial population census captures data on gender, race, and ethnicity, but, of course, only every 10 years and without socioeconomic establishment detail. The Current Population Survey and the American Community Survey collect wage data by gender, race/ethnicity, nativity (native/foreign born), and many other characteristics, but not by establishment. The Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) System links UI and QCEW data on employers and employees (obtained through individual agreements with states) with additional employer and employee data from censuses and surveys. The data, which include wage information, are available only to qualified researchers at one of the Census Bureau's Research Data Centers.[17]

All information collected by the federal government for statistical purposes, including the data in these three BLS surveys, is collected under a pledge of confidentiality according to the provisions of the 2002 Confidential Information Protection and Statistical Efficiency Act (CIPSEA). This means that the data cannot be shared for purposes of antidiscrimination enforcement; however, the information may be used to assist in analysis relevant to wage discrimination, and the ability of the survey to collect wage information may be instructive for EEOC.

National Compensation Survey

The NCS is an establishment-based survey that annually provides estimates of occupational earnings, employer costs for employee compensation, compensation trends, wages in one geographic area relative to other geographic areas, the incidence of employer-provided benefits among workers, and provisions of employer-provided benefit plans. The employment cost

[17]For details, see http://lehd.did.census.gov/led/ [July 2012].

index (ECI)—a principal federal economic indicator—is estimated from data collected by the NCS.[18]

The NCS samples private industry establishments with one or more workers and state and local governments across the 50 states and the District of Columbia. Each sampled establishment—over 35,000 establishments in 2010—is asked to report on selected occupations. As stated in the *BLS Handbook of Methods*, major exclusions from the survey are workers in federal and quasi-federal agencies, military personnel, agricultural workers, workers in private households, the self-employed, volunteers, unpaid workers, individuals receiving long-term disability compensation, and individuals working overseas. Currently, the NCS also excludes individuals who set their own pay (e.g., proprietors, owners, major stockholders, and partners in unincorporated firms) and family members being paid token wages; however, these exclusions are being reevaluated (Bureau of Labor Statistics, no date).

Among the products of the survey are estimated average hourly wages for over 800 occupations in approximately 80 metropolitan and selected nonmetropolitan localities, weekly and annual earnings and hours for full-time workers, and earnings by work level that permit wage comparisons across occupational groups. The survey collects no demographic detail, however, and it is therefore not directly useful for analysis that might facilitate antidiscrimination enforcement.

Current Employment Statistics Survey

The CES is an establishment payroll survey that is based on a monthly survey of approximately 141,000 businesses and government agencies representing approximately 486,000 worksites throughout the United States.[19] The primary statistics derived from the survey are monthly estimates of employment, hours, and earnings for the nation, states, and major metropolitan areas. Preliminary national estimates for a given reference month are typically released on the third Friday after the conclusion of the reference week, which is the week that includes the 12th of the month.

National estimates of average weekly hours and average hourly earnings are made for the private sector for all employees and for production and nonsupervisory employees. Detail is available for about 750 industries. Average weekly overtime hours in manufacturing are also available.

Hours and earnings are derived from reports of gross payrolls and corresponding paid hours. However, hours for salaried workers who may have

[18]For details, see http://www.bls.gov/eci/# [July 2012].

[19]Information in this section is largely reproduced from http://www.bls.gov/ces/cescope.htm [July 2012].

set compensation but volatility in their hours are often reported as standard weekly hours rather than hours actually worked and paid. The payroll for employees covered by the CES is reported before deductions of any kind, for example, for Social Security, federal and state withholding tax, union dues, or retirement plans. Included in the payroll reports is pay for overtime, vacations, holidays, and sick leave paid directly by the firm. Bonuses, commissions, and other types of nonwage cash payments are excluded unless they are earned and paid regularly (at least once a month). Employee benefits paid by the employer, as well as in-kind payments, are excluded.

Total hours during the pay period include all hours worked (including overtime hours), and hours paid for holidays, vacations, and sick leave. Total hours differ from the concept of scheduled hours worked. Average weekly hours reflect effects of numerous factors, such as unpaid absenteeism, labor turnover, part-time work, strikes, and fluctuations in work schedules for economic reasons. Overtime hours in manufacturing are collected when overtime premiums were paid and the hours were in excess of the number of straight-time hours in a workday or workweek. No information is collected by gender, race/ethnicity, or nativity.

Occupational Employment Statistics Survey

The OES survey is a semiannual mail survey designed to measure occupational employment and wage rates among full- and part-time wage and salary workers in nonfarm establishments in the United States.[20] The survey does not include the self-employed, owners and partners in unincorporated firms, household workers, or unpaid family workers.

The OES survey is a cooperative program between BLS and state workforce agencies (SWAs). BLS funds the survey and provides the procedures and technical support, while the SWAs collect most of the data.[21]

The OES is a very large survey. Its estimates are constructed from a sample of about 1.2 million establishments grouped into six semiannual panels over a 3-year period. Each year, forms are mailed to two panels of approximately 200,000 establishments, one panel in May and the other in November. Thus, for example, the May 2010 estimates were based on responses from six panels—May 2010, November 2009, May 2009, November 2008, May 2008, and November 2007.

The overall national response rate for six panels is about 78 percent based on establishments and 74 percent based on employment. The survey covers all employer size classes, and response rates are actually higher among smaller employers. The survey's coverage is extensive—approximately 63

[20]Information in this section is largely reproduced from http://www.bls.gov/oes/ [July 2012].
[21]Data for 180 large firms are collected directly by BLS.

TABLE 2-3 Occupational Employment Statistics by Number of Employees, May 2010

Number of Employees	Viable Sample Units	Respondents
1–9	437,389	380,215
10–19	195,755	155,320
20–49	202,642	148,143
50–99	107,175	71,562
100–249	84,492	55,090
250–499	35,225	22,780
500–999	13,620	8,778
More than 999	8,747	6,346

NOTE: The respondents were establishments.
SOURCE: Data from Bureau of Labor Statistics special tabulation for the panel.

percent of total national employment is represented by the unweighted employment of sampled establishments across all six semiannual panels.

The OES survey draws its sample from state UI files. The survey sample is stratified by metropolitan and nonmetropolitan area, industry, and size. To provide the most occupational coverage, larger employers are more likely to be selected than smaller employers (see Table 2-3).

The data available from the OES include cross-industry occupational employment and wage estimates for over 500 areas, including the nation, states, and the District of Columbia, metropolitan statistical areas (MSAs), metropolitan divisions (the result of MSA subdivisions) nonmetropolitan areas, and territories; national industry-specific estimates at the 2007 NAICS 3-, 4-, and selected 5-digit industry levels; and national estimates by ownership across all industries and for schools and hospitals (Bureau of Labor Statistics, 2010a). No data are collected by gender, race/ethnicity, or nativity.

The OES survey categorizes workers into nearly 800 detailed occupations based on the Office of Management and Budget's Standard Occupational Classification (SOC) system. The detailed occupations cover 22 of the 23 SOC major occupational groups. The May 2010 OES estimates mark the first set of estimates based in part on data collected using the 2010 SOC system, and after May 2012, the OES data will reflect the full set of detailed occupations in the 2010 SOC. Importantly, the 2010 SOC occupations will be capable of being cross-walked into the EEOC job categories when EEOC completes an update of the crosswalk between the EEOC job categories and the 2010 SOC.

SUMMARY

Several surveys have been developed specifically to measure pay discrimination, and there are several survey-based and administrative records-based sources of estimates of earnings. They vary widely in their approach to measurement, their coverage of employers, and their content: for example, only some of them collect demographic as well as earnings information. Only two of the data sources for establishments contain information on hours and whether the employee is on a temporary or permanent schedule, and neither of those sources includes demographic information.

It is clear that there is no current source of earnings data that incorporates the demographic, occupation, work schedule, and employer information necessary to support an antidiscrimination enforcement and analytical program. A new reporting mechanism would have to be put in place to produce earnings by gender, race, and ethnicity for establishments.

Nonetheless, the fact that earnings data are now generally reported to the taxing authorities and to federal (and state) government statistical and enforcement agencies suggests that it might be feasible to collect earnings information by gender, race, and national origin in an EEOC data collection program. It also suggests that the EEOC may be able identify other data collections that could serve as sources of benchmarks to assist in validating the information that might be collected as part of a new reporting arrangement.

3

Pay Concepts and Definitions

Pay is an important indicator of discriminatory practices. Employees with the same productivity and working conditions (including hours) in the same jobs at the same employer location could be subject to pay discrimination if they are systematically paid differently because of their membership in a particular demographic group. Employment discrimination can affect pay through a number of different channels, such as different pay rates, different noncash compensation, different hours offered, and different job assignments to otherwise similar applicants. Each of these channels poses measurement challenges.

A major challenge in considering an appropriate earnings measure to use in determining whether or not there is pay discrimination is that there is no standardized and universally accepted measure of earnings. Earnings (pay) can be represented in a variety of ways depending on the use to which the definition will be put:

- as annual, monthly, or weekly amounts;
- as totals, averages, rates, or pay bands;
- narrowly, as straight-time pay or regular salaries;
- broadly, to include straight-time pay or regular salaries and other forms of compensation, such as commissions, overtime, incentives, bonuses, shift differentials, stock options, and premium pay; and
- in relation to terms of employment, such as time worked, and the conditions of employment, such as whether the employee is full or part time and permanent or temporary.

Although disparities in any of these measures could signal discriminatory practices, the most applicable earnings measure to comprehensively identify pay discrimination may well be a broad measure that encompasses all, or nearly all, measures of compensation, such as the broadest measure noted above, as well as the terms and conditions of employment. Such a measure, however, may not be easily collectable. Furthermore, even if it were collectable, it may not be fully comparable across demographic and job groupings, as discussed below. Thus, it may be that any earnings measure selected by the U.S. Equal Employment Opportunity Commission (EEOC) may depend as much on whether the information is available and collectable than on the purpose for which it is collected and how it will be used.

In this chapter, we discuss the various components of employee compensation that can be considered when selecting the most appropriate definition of earnings for antidiscrimination purposes. We also consider trends over time in compensation practices. Finally, we assess several possible definitions from the perspectives of scope, coverage, frequency, reliability, and collectability.

ROLE OF COMPENSATION

Compensation plays many roles in the modern economy. According to Kevin Hallock, director of the Cornell University Institute of Compensation Studies, who discussed compensation issues with the panel, compensation depicts market pricing of an essential component in the production function, and, in most instances, helps to match supply and demand for a workforce and for particular skills and qualifications.[1] It can be a measure of responsiveness to offers. It can be adjusted to fit time, place, and circumstance by adjusting the pieces of compensation (wages, benefits, schedule, and other pay). Nowhere have these kinds of adjustments been more aggressive than with executive and highly paid professional compensation, for which a rich array of compensation options has emerged in recent years.

Compensation policies also play a large role in business strategy. These policies undergird and give meaning to job analysis and job evaluation processes, and they enable pay-for-performance and other productivity enhancement strategies. They facilitate internal comparisons and, when data are available, facilitate external comparisons, which are a component of competitive analysis.

[1]Various administrated pay systems (such as much of the civil service) and structures that constrain supply (e.g., licenses and apprenticeship systems) may include departures from the generalization that compensation reflects the operation of the unfettered labor market.

More and more, compensation policies are a key element in corporate strategies to improve efficiency, effectiveness, and marketplace viability. In a broad sense, they have been identified as "total rewards" strategies (WorldatWork Association, 2011). In addition to their importance as compensation in corporate business strategies, employers also seek through these policies to achieve balance in work-life considerations, performance and recognition policies, and development and career opportunities for their workforce.

There are common elements to compensation strategies across the occupational spectrum. However, one result of strategic "fine tuning" by businesses is that wages and total compensation have come to vary among occupational groups, which adds to the difficulty of making cross-occupational comparisons. Data from the National Compensation Survey (NCS)—administered by the Bureau of Labor Statistics (BLS)—indicate that wages and salaries make up a larger proportion of its definition of compensation (wages and salaries plus benefits, including supplemental pay) for management, sales, and service workers than for construction and production workers: see Figure 3-1. Total compensation may encompass much more than hourly earnings, so it is important to consider broader measures of compensation.

EARNINGS DATA AVAILABLE IN FIRMS

It is important to define earnings in a way that makes economic sense, but it is also critical to define earnings in a way that reporting employers can understand. Earnings should be capable of being reported using records readily available in the firm because they are otherwise necessary to meet the requirements of law or regulation or because they are needed for the efficient operation of the firm. Existing laws and regulations help delineate the kinds of compensation and demographic data that employers maintain.

At a minimum, all employers covered by the Fair Labor Standards Act (FLSA)[2] must keep certain records for each covered, nonexempt worker.[3] Although there is no required format for the records, the content of the records is specified: The records must include accurate information about the employee and data about the hours worked and the wages earned, to include[4]

[2]Employers covered by FLSA are those with at least two employees and an annual dollar volume of sales or business of at least $500,000. See http://www.dol.gov/whd/regs/compliance/whdfs14.pdf [December 2011].

[3]Under the FLSA, some employees are exempt from the act's overtime provisions. These employees include executive, administrative, professional, and outside sales employees who are paid on a salaried basis, some commissioned sales employees, and some seasonal employees.

[4]For details, see http://www.dol.gov/dol/topic/wages/wagesrecordkeeping.htm [July 2012].

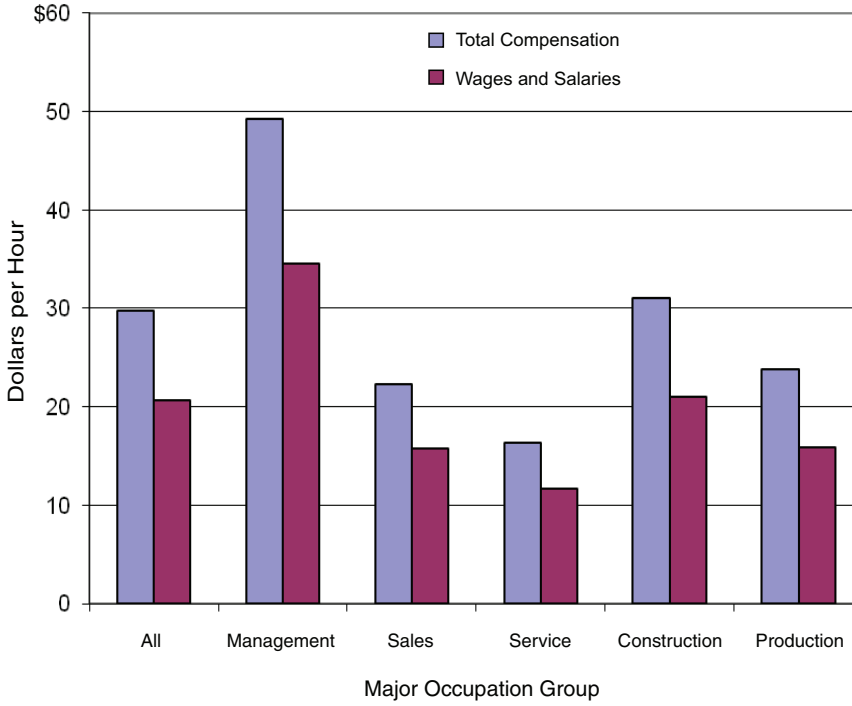

FIGURE 3-1 Hourly wage and salary and total pay by major occupational group, 2011.
SOURCE: Presentation by Kevin Hallock at panel workshop on May 24, 2011, based on data from National Compensation Survey. Reprinted with permission.

- employee's full name, as used for Social Security purposes, and on the same record, the employee's identifying symbol or number if such is used in place of name on any time, work, or payroll records;
- address, including zip code;
- birth date, if younger than 19;
- sex;
- occupation; time and day of week when employee's workweek begins; hours worked each day and total hours worked each workweek;
- basis on which employee's wages are paid;
- regular hourly pay rate;
- total daily or weekly straight-time earnings;
- total overtime earnings for the workweek;
- all additions to or deductions from the employee's wages;

- total wages paid each pay period; and
- date of payment and the pay period covered by the payment.

Given these FLSA requirements, it is safe to assume that employers covered by FLSA will maintain wage information by gender. However, wage data may not be universally available by race and national origin (data on these characteristics are required by equal employment opportunity [EEO] legislation, but not necessarily with wage data associated with them).

Other statutory and administrative requirements dictate the kind of data that employers should maintain on employee compensation. For example, those firms that have adopted employer-matching 401(k) plans called Safe Harbor plans must use the Internal Revenue Service definition of compensation, which includes all wages; salaries; other amounts received that are includible in the employee's gross income, including overtime; other items including commissions, fees for professional services, tips, bonuses, fringe benefits, and reimbursements for some other expense allowances; and foreign earned income. All of these compensation items must be accounted for: thus, for firms with this type of 401(k) plan, the compensation information is likely to be obtainable from the firm's compensation records.

Although FLSA coverage and other administrative reporting requirements tend to define the mandatory wage information that is likely to be maintained by employers that report to EEOC, the specific data that are maintained by any particular employer are defined by the particular payroll and human resource systems that support the business's operations. In many cases, these systems are developed within the company, although, increasingly, company payroll and human resource systems are developed by outside firms that specialize in providing software or "turnkey" human resources and payroll management services (see Chapter 1). Thus, a good rule of thumb would be that earnings measures for EEOC reporting would need to be compatible with data elements available from vendor systems or, at least, only require changes that could be easily implemented in vendor software.

FEASIBLE DEFINITIONS OF EARNINGS

There is no single, commonly accepted definition of earnings. Table 3-1 shows the wide and rich variety of definitions embedded in the major survey and tax collection systems (discussed in Chapter 2). Because earnings data are now being collected according to various definitions, any of the definitions could be considered collectable. However, not all definitions have a history of being collectable with the addition of occupational and demographic information.

Two employer-based BLS data collections now bring together data on the establishment, compensation, occupation, and hours—the Occupational

TABLE 3-1 Comparison of Earnings Definitions and Data Availability for Key Earnings Data Sources

Data Source	Definition of Earnings	Occupational Coverage	Demographic Information
	Employer/Establishment-Based Surveys		
Occupational Employment Survey (Bureau of Labor Statistics)	Wages for the OES survey are straight-time, gross pay, exclusive of premium pay. Earnings include base rate; cost-of-living allowances; guaranteed pay; hazardous-duty pay; incentive pay, including commissions and production bonuses; and tips. Excluded are overtime pay, severance pay, shift differentials, nonproduction bonuses, employer cost for supplementary benefits, and tuition reimbursements.	The OES survey categorizes workers into nearly 800 detailed occupations based on the Office of Management and Budget's Standard Occupational Classification (SOC) system.	None
Current Employment Statistics Survey (Bureau of Labor Statistics)	Provides arithmetic averages (means) of the hourly and weekly earnings of all production and nonsupervisory jobs in the private nonfarm sector of the economy. The hours and earnings are derived from reports of gross payrolls and corresponding paid hours. Payroll is reported before deductions of any kind, e.g., for old-age and unemployment insurance, tax withholding, union dues, or retirement plans. Included in the payroll reports is pay for overtime, vacations, holidays, and sick leave paid directly by the firm. Bonuses, commissions, and other types of nonwage cash payments are excluded unless they are earned and paid regularly (at least once a month). Employee benefits paid by the employer, as well as in-kind payments, are excluded.	None	None

continued

TABLE 3-1 Continued

Data Source	Definition of Earnings	Occupational Coverage	Demographic Information
National Compensation Survey (Bureau of Labor Statistics)	Wages and salaries, or earnings, are defined as regular payments from the employer to the employee as compensation for straight-time hourly work or for any salaried work performed. Includes incentive pay, including commissions, production bonuses, and piece rates; cost-of-living allowances; hazard pay; payments of income deferred because of participation in a salary reduction plan; and deadhead pay, defined as pay given to transportation workers returning in a vehicle without freight or passengers.	Standard Occupational Classification (2010) definitions are used for initial data collection at an establishment. (The 2010 SOC system contains 840 detailed occupations, aggregated into 461 broad occupations.)	None
EEO-4 Survey (state and local governments)	Annual salary including all special increments of an employee's annual earnings that are regular and recurrent. Overtime pay is not included. Where employees are paid on an other-than-annual basis, their regular earnings in the payroll period that includes June 30 are to be expanded and expressed in terms of an annual income.	Officials and administrators; professionals; technicians; protective service workers; paraprofessionals; administrative support (including clerical and sales); skilled craft workers; service-maintenance.	White (not of Hispanic origin); Black (not of Hispanic origin); Hispanic; Asian or Pacific Islander; American Indian or Alaskan Native, by male and female
OFCCP EO Survey	Annual monetary compensation: the employee's base rate (wage or salary) plus other earnings, such as cost-of-living allowance, hazard pay, or other increment paid to employees regardless of tenure on the job. The annual monetary compensation measure was not to include the value of benefits, overtime, or one-time payments such as relocation expenses.	(1) Officials and managers; (2) professionals; (3) technicians; (4) sales workers; (5) office and clerical workers; (6) craft workers; (7) operatives; (8) laborers; and (9) service workers.	Minority females, non-minority females, minority males, non-minority males

Ontario Pay Equity Survey	Pay as of December 31 expressed in hourly, weekly, or annual amounts.	Job/position title	Male and female
Minnesota Pay Equity Survey	Minimum and maximum monthly salary.	Job class	Male and female
New Mexico Pay Equity Survey	Total annual compensation converted to average hourly wages in each job category are computed by adding the total compensation by gender divided by the total hours worked by that gender.	EEO-1 job categories	Male and female
	Administrative Records		
Employer's Quarterly Contribution and Wage Report	Total quarterly wages paid to all regular, part-time, temporary, or casual employees, without regard to age; wages paid for services performed for a partnership by the wife, husband, child, or other relative of a partner; wages paid by an individual owner to a son or daughter who is 18 or more years of age; salaries and other payments made to corporate officers for their services to the corporation (including Subchapter S corporations); tips reported by employees for Internal Revenue Service purposes by the 10th day of the month of receipt; reasonable cash value of meals, lodging, merchandise, and other types of remuneration furnished for services; commissions and bonuses paid to employees; vacation payments; dismissal pay, severance pay, or wages in lieu of notice; salary reductions pursuant to Internal Revenue Code (IRC) Section 125 (cafeteria plans) or 401(k) plans.	None	None

continued

TABLE 3-1 Continued

Data Source	Definition of Earnings	Occupational Coverage	Demographic Information
IRS W-2 Form	Wages and salaries, deferred compensation (part of total compensation, even if not taxable currently), and certain fringe benefits are reported in addition to capped Social Security earnings and uncapped Medicare earnings.	None	None
IRS 941 and 943 Forms	Total compensation; employer reported W-2 income and tips.	None	None
Social Security Master Earnings File	OASDI and Medicare taxable wages, and total wages reportable as IRS-taxable income on Form 1040, which includes wages above the OASDI taxable maximum, noncovered wages, and deferred-compensation distributions, but not deferred-compensation contributions.	None	Gender; self-reported race and ethnicity data provided on voluntary basis
Fair Labor Standards Act (FLSA)	Time and day of week when employee's workweek begins; hours worked each day and total hours worked each workweek; basis on which employee's wages are paid; regular hourly pay rate; total daily or weekly straight-time earnings; total overtime earnings for the workweek; all additions to or deductions from the employee's wages; total wages paid each pay period; date of payment and the pay period covered by the payment.	Occupation	Age; sex

Safe Harbor 401(k) Plans	All wages; salaries; other amounts received that are includible in the employee's gross income, including overtime; other items, including commissions, fees for professional services, tips, bonuses, fringe benefits, and reimbursements for some other expense allowances; and foreign earned income.	None	None

NOTE: EEO = Equal Employment Opportunity, EO = Equal Opportunity Survey, IRS = Internal Revenue Service, OASDI = Old-Age, Survivors, and Disability Insurance Program, OES = Occupational Employment Survey, OFCCP = Office of Federal Contract Compliance Programs.
SOURCES: Information from Current Employment Statistics forms (available at: http://www.bls.gov/ces/cescope.htm [July 2012]); Employer's Quarterly Contribution and Wage Report (available at: https://uitax.nvdetr.org/crppdf/nucs-4072.pdf [July 2010]); and Recordkeeping Requirements under the Fair Labor Standards Act (available at: http://www.dol.gov/whd/regs/compliance/whdfs21.htm [July 2012)].

TABLE 3-2 Occupational Employment Statistics Survey Wage Intervals, May 2010

Wage Intervals	Hourly	Annual
Range A	Under $9.25	Under $19,240
Range B	$9.25 to $11.49	$19,240 to $23,919
Range C	$11.50 to $14.49	$23,920 to $30,159
Range D	$14.50 to $18.24	$30,160 to $37,959
Range E	$18.25 to $22.74	$37,960 to $47,319
Range F	$22.75 to $28.74	$47,320 to $59,799
Range G	$28.75 to $35.99	$59,800 to $74,879
Range H	$36.00 to $45.24	$74,880 to $94,119
Range I	$45.25 to $56.99	$94,120 to $118,559
Range J	$57.00 to $71.49	$118,560 to $148,719
Range K	$71.50 to $89.99	$148,720 to $187,199
Range L	$90.00 and over	$187,200 and over

SOURCE: Bureau of Labor Statistics (2009).

Employment Statistics (OES) survey and National Compensation Survey (NCS).[5] These survey collections do not include demographic information: such information would have to be added to the compensation, occupation, and hours data collected in these two surveys to provide the information minimally needed for antidiscrimination purposes.[6] The definitions of earnings in these surveys are discussed below.

OES Wage Definition

Earnings in the OES survey are defined as straight-time gross pay, exclusive of premium pay. The definition includes a base rate of pay; cost-of-living allowances; guaranteed pay; hazardous-duty pay; incentive pay, including commissions and production bonuses; and tips. The definition excludes overtime pay, severance pay, shift differentials, nonproduction bonuses, employer costs for supplementary benefits, and tuition reimbursements.

The OES survey collects wage data from private-sector employers in 12 intervals (or bands): see Table 3-2. For each occupation, respondents are

[5]This discussion is limited to measures of compensation that can be collected from employers rather than from individuals because of the requirement to identify the possibility of pay discrimination at the point of employment, even though the most complete view of compensation and demographics can be developed from household and individual surveys (Abowd and Hallock, 2007; Zhao, 2010).

[6]As discussed in Chapter 2, data from these surveys are collected under a pledge of confidentiality and are not available for enforcement purposes. However, the data could serve a benchmarking role for EEOC surveys; moreover, the surveys indicate the feasibility of data collection by establishment on occupation, hours, and earnings.

asked to report the number of employees paid within each wage intervals. The effect of having a relatively large number of intervals in the OES survey is to narrow the bands so as to minimize the possibility of concealing pay disparities that could signal discrimination, which might occur with broad bands. The intervals are defined both as hourly rates and the computed corresponding annual rates: the annual rate for an occupation is calculated by multiplying the hourly wage rate by a typical work year of 2,080 hours.

The responding establishments are instructed to report the hourly rate for part-time workers and to report annual rates for occupations that are typically paid at an annual rate but for less than 2,080 hours per year, such as teachers, pilots, and flight attendants. Other workers, such as some entertainment workers, are paid hourly rates, but generally do not work 40 hours per week, year round. For these workers, only an hourly wage is reported.

NCS Earnings Definition[7]

In the NCS, wages and salaries, or earnings, are defined as regular payments from the employer to the employee as compensation for straight-time hourly work or for salaried work. The survey includes the following components as part of earnings:

- incentive pay, including commissions, production bonuses, and piece rates;
- cost-of-living allowances;
- hazard pay;
- payments of income deferred because of participation in a salary reduction plan; and
- deadhead pay, defined as pay given to transportation workers returning in a vehicle without freight or passengers.

The following items are not considered part of straight-time earnings, and data on them are not included in the NCS:

- uniform and tool allowances,
- free or subsidized room and board,
- payments made by third parties (e.g., tips), and
- on-call pay.

The following forms of payments are considered benefits and not part of straight-time earnings:

[7]The information in this section is largely taken from descriptions of the NCS, see http://www.bls.gov/ncs/ncswage2010.pdf [July 2012].

- payments for shift differentials, defined as extra payment for working a schedule that varies from the norm, such as night or weekend work;
- premium pay for overtime, holidays, and weekends; and
- bonuses not directly tied to production (such as Christmas and profit-sharing bonuses).

The NCS annually publishes national, Census Bureau division, and local area occupational earnings estimates of mean hourly earnings, mean and median weekly and annual earnings, and weekly and annual hours, for civilian workers (as defined by the NCS), private-industry workers, and state and local government workers. Occupational earnings data are published for some major and minor industry groups, by worker attributes (such as collective bargaining status), and by establishment characteristics (such as number of workers in the establishment). Percentile earnings by worker attributes and establishment characteristics are also published. Earnings data are presented as mean and median hourly, weekly, and annual earnings (along with hours worked weekly and annually); as percentiles; by selected worker attributes (such as full time and part time, and union and nonunion); and by establishment characteristics (such as number of employees and geographic area).

To calculate earnings for various periods (hourly, weekly, and annually), the NCS collects data on work schedules. For hourly workers, scheduled hours worked per day and per week, exclusive of overtime, are recorded, as well as the number of weeks worked annually. For salaried workers, field economists record the typical number of hours actually worked (salaried workers who are exempt from overtime provisions often work beyond the assigned work schedule).

The NCS publishes earnings estimates for occupational groups and detailed occupations; it also presents earnings estimates by work levels and combined work levels. Work levels represent a ranking of the duties and responsibilities in an occupation.

CONCLUSION

Of the two feasible wage definitions that could be used, we conclude that the definition used in the OES should be considered for use for antidiscrimination purposes because its current coverage is so widespread. Most employers who are in the industries and size classes that report employment by gender, race, and national origin to the EEOC already have experience in assembling and reporting hours and earnings together by occupation in order to complete the OES (see Chapter 2). There is strong reason to believe that the information is available and retrievable in the firms that would be called on to report earnings data to the EEOC.

4

Survey Design and Statistical Methodology

When considering the collection of earnings data by gender, race, and national origin, the U.S. Equal Employment Opportunity Commission (EEOC) confronts several key decisions in the realm of survey design and statistical methodology. The decisions involve four closely associated issues: collectability, quality (defined as fitness for use), utility for statistical analysis, and response burden.

In this chapter we discuss the pros and cons of options for collecting earnings data from employers by adding items to existing equal employment opportunity (EEO) forms or developing a new collection instrument. We consider the fitness for use of the data, which addresses the relevance of the data to users' needs. We illustrate a model-based approach to identifying the utility of the categorical variables that would also be collected if wage data are collected. We address the question of employer burden and assess various options for minimizing the burden on reporting units. The last issue is complicated by the fact that there is a differential burden faced by employers of different sizes and with different levels of sophistication in their human resource and payroll systems. In the case of collection of earnings data by gender, race, and national origin, one approach may not be appropriate for all respondents.

OPTIONS FOR DATA COLLECTION

Modify Current EEO Forms

The most direct solution to obtaining earnings information for EEOC purposes would be to add earnings items to existing EEO reports. The collection instrument that it would likely make most sense to modify for this purpose would be the EEO-1 form, for several reasons. First, it enjoys substantial coverage. As discussed in Chapter 1, the mandatory EEO-1 reports in 2010 covered about 67,000 private-sector respondents, with about 55 million employees.

Second, the form is part of the everyday operations of the antidiscrimination agencies. The EEO-1 reports are used by the EEOC and the Office of Federal Contract Compliance Programs (OFCCP) to trigger enforcement and technical assistance based on the identification of potential EEO problems, which is determined from data provided by employers on the reports.

Third, it is expected that the necessary modifications to the EEO-1 form would be quite manageable for both EEOC and the respondents. The addition of the earnings data could be accomplished in much the same way that earnings data are collected on the EEO-4 form: that is, either by adding another column to the form that requests the earnings data or adding another row for each occupation, which would collect average pay in addition to the current row that collects number of employees by race/ethnicity group. An alternate collection design would be to simply duplicate the existing EEO-1 form and have employers place in the cells of one table the number of employees, as they now do, and in the second table enter the pay corresponding to those employees.

Design a New Collection Instrument

A second option would be to design a new and, one hopes, a more streamlined collection instrument that would collect both employment and earnings information. The design of such a new instrument could be informed by the current effort by OFCCP to develop a collection instrument to replace the defunct Equal Opportunity Pilot Survey discussed in Chapter 2. As this report was being prepared, the OFCCP had issued an Advance Notice of Proposed Rulemaking (ANPRM) that solicited comments on several issues important for designing a new collection instrument. For example, OFCCP asks whether expanded information should be collected in order for OFCCP to assess whether further investigation into a contractor's compensation decisions and policies is warranted. To collect such data as average starting or initial total compensation (including paid leave, health and retirement benefits, etc.); average pay raises; average bonuses;

minimum and maximum salary; standard deviation or variance of salary; the number of workers in each gender and race/ethnicity category; average tenure; and average compensation data by job series (e.g., all engineers within a particular department or all secretaries throughout the establishment) would require a substantial redesign of the collection form.

Some of the items that might be useful in understanding the EEO environment in establishments would likely require open-ended questions, such as on topics suggested in the OFCCP ANPRM pertaining to company policies related to promotion decisions, bonuses, shift pay, and setting of initial pay. This information is difficult to collect and to process efficiently in a standardized manner.

FITNESS FOR USE

Types of Uses

Quality of information is generally defined in terms of its fitness for use. This is a multidimensional concept embracing the relevance of the information to users' needs and the accuracy, timeliness, accessibility, interpretability, and coherence that affect how the data can be used. There is a considerable literature on statistical quality and the steps that should be taken to make data useful for its intended purpose (see, e.g., Brackstone, 1999; U.S. Office of Management and Budget, 2002). The literature highlights the importance of clearly understanding the requirements for the data before collection begins. It is important in this context to consider the need of the EEOC for earnings information.

The major use of the EEO compensation data would be to aid enforcement of pay discrimination statutes in two ways: targeting enforcement actions and carrying out enforcement actions against an employer that has been targeted. Targeting is primarily a matter of selecting among the complaints the EEOC receives to identify those firms that are most likely to be found to have discriminatory practices.

There are, however, secondary uses, such as analysis of overall trends in pay discrimination and trends by industry and location, as well as research on compensation trends. If such new compensation data become available, they would be a powerful supplement to existing sources of compensation data, such as those discussed in Chapters 2 and 3.

Because the data collected by this survey would be so important to collect correctly, it is incumbent on EEOC to identify the potential uses of the data early in a design process so that the data items to be collected can be identified and issues of data quality considered. Again, the requested comments in the OFCCP ANPRM are instructive when paraphrased in EEOC terms:

- Should the data be used to conduct industrywide compensation trend analyses? If so, what type of compensation trend analyses would be appropriate to conduct on an industrywide basis?
- For each type of analysis identified, identify the categories of data that should be collected in order to compare compensation data across employers in a particular industry and the job groupings that should be used.
- Should the data be used to identify employers in specific industries for industry-focused compensation reviews?
- What specific categories of data would be most useful for identifying employers in specific industries for industry-focused compensation reviews?
- Should the data be collected by individual establishment for multiestablishment employers? What specific categories of data would be most useful for conducting compensation analyses across an employer's various establishments?

Utility of the Data Items for Statistical Analysis

In this section we consider how the EEOC could develop a statistical model for use in screening individual employers for possible violations of pay discrimination. There are several key considerations here. First, the data to be used in this model would, of course, be reported by each individual employer. In addition to the information already requested for the EEO-1 report (e.g., employment by occupation, sex, and race/ethnicity), a form would collect pay (measured as discussed in Chapter 3) and possibly other information, such as employees' years of service. Given these data, one could conduct a multiple regression analysis of pay in relation to demographic variables (e.g., the EEO-1's 14 sex and race/ethnicity groups) and other characteristics, usually called "control variables," such as occupational category and years of service. More complex models might include controls for occupation or job categories or more elaborate controls for education and labor force experience. Still more complex models might include more detailed occupational or job categories and more elaborate controls for previous experience and qualifications.

There are a large number of potential control variables that could be included in such regression models, and, especially for employers with small numbers of employees, there would be benefits from keeping the number of covariates in such models relatively small. To do that, there are a variety of statistics, including Mallows' C_p, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), that could be employed to remove control variables that were not contributing substantially to the fit of the model.

While there is substantial disagreement over the most appropriate models to use for establishing a reasonable claim of possible wage discrimination, or defending one, it is not necessary to have a definitive model to assess the potential quality of certain basic statistical tests that might be reasonably performed by EEOC. We undertake such an analysis here. We emphasize that the regression model we describe below is intended, first and foremost, as an illustrative example of a methodology for undertaking some of these basic statistical tests. For this purpose, we need to provide enough specifics to allow a clear and straightforward discussion of the general nature of the issues that would arise in such an exercise.

The regression model we use is a general linear model of the form:

$$y_i = \beta_0 + d_i\beta_1 + x_i\beta_2 + \varepsilon_i$$

Here, y_i is the logarithm of the wage measure for individual i, d_i is the vector of design variables that indicate the EEO-1 categories occupied by individual i, x_i is a vector of control variables, ε_i is the statistical error, β_0 is the intercept, β_1 is the vector of EEO-1 log wage differentials from a specified reference group (usually white, non-Hispanic males), β_2 is the vector of effects associated with the control variables, and $i = 1,...,N$, where N is the total number of employees in the analysis.[1]

For an agency such as EEOC or OFCCP, the results from this kind of regression analysis that will be of greatest concern will be the estimates of the coefficients for gender and race/ethnicity: that is, the betas, because the estimates of these coefficients indicate the extent (if any) to which women or nonwhites are paid less than men or whites who are the same in terms of the other factors (the "control variables") included in the analysis. It will be particularly important to perform a test to determine if these coefficients are statistically significantly different from zero (i.e., are unlikely to have occurred simply as a result of random or chance factors).

Assuming that design vectors d_i and x_i are statistically exogenous with respect to $\grave{o}_1$ and that $\grave{o}_1$ has a normal distribution with zero mean, constant variance, and independence over individuals, there is a well-known F-test for the null hypothesis: $\beta_1 = 0$. This statistic tests the hypothesis that all of

[1]The earliest analyses that used the logarithm of wages were Blinder (1973) and Mincer (1974). Their work discussed specifications in the logarithm and levels. Since the early 1970s the prevailing practice in economics has been to use the logarithm of the rate of pay as the dependent variable. The regression model has been selected because when analysis is expressed in logs, pay gaps can be expressed in a comparable way (i.e., as percentages) even for dates that are wide apart. This also means that estimated coefficients in log regressions can be interpreted as showing the *percentage change* in y that occurs as a result of a change in x and when x is an indicator for race or gender, it measures the percentage difference in pay between the indicated group relative to a reference group.

the EEO-1 log wage differentials are jointly zero versus the alternative that at least one of the differentials is nonzero. The usual *F*-statistic is based on the Type-III sum of squares for the model component associated with the design vector d_i: that is, the conditional model sum of squares for d_i given the other variables, x_i, in the model. This statistic is invariant to the choice of reference group.

An automated test of the hypothesis $\beta_1 = 0$ could be conducted from an enhanced EEO-1 report that included appropriate wage data. The suitability of such a test depends on how likely it is that the test would detect a departure from $\beta_1 = 0$ for realistic configurations of employer data and with appropriate controls. We approach this question by attempting to measure the power of the standard *F*-test for $\beta_1 = 0$ in scenarios that resemble best-case outcomes for such an automated procedure.

The power of a test is the probability that it will reject the null hypothesis when that hypothesis is false. In other words, the power of a test is the probability that it will actually find a sex or race/ethnicity difference when such a difference exists. In colloquial terms, one might say that the power of a test is the probability that it will detect a potentially discriminating ("guilty") party. The power depends on the magnitude of the departure from the null hypothesis (how big the differentials are) and the precision with which those differentials can be estimated. In turn, the precision of the estimate(s) depends critically on the number of data points used in forming the estimates.

In the present context, it is crucial to note that the power of the statistical model for screening employers will be sensitive to the number of data points used in its construction. It is simple common sense that, other things being equal, a poll of 1,000 people is likely to be much more precise (will have much greater power) than a poll of 100 people; similarly, regression estimates of sex or race/ethnicity pay differences that are based on many data points will have greater power than estimates based on only a few data points. Finally, note that the number of data points in an analysis of a particular employer will depend on the size of the employer's work force: the greater the number of employees, the greater the number of data points, and the greater the power of the statistical model used in screening employers. Thus, when the number of employees is small, any screening model that EEOC might develop will have very low power, and when the number of employees is large, the screening model will have high power. The important question is thus obvious: How many data points must there be—how large does the employer's work force have to be—to yield "enough" power?

For general linear models, there is standard software to assist with this power assessment. The inputs consist of estimates of the magnitude of the likely discrepancy and summary measures of the estimation precision. We next describe how we estimated those components.

We considered an employer-size power analysis that is based on the predictions and estimation precision of models fit on the March 2010 Current Population Survey (CPS) Annual Social and Economic Supplement. Essentially, we are asking: "How many employees must a respondent firm have in order for the *F*-test to have the specified power to detect log wage differentials as big as the ones in the overall economy, as measured in March 2010?" This is a "best-case" scenario for two reasons. First, the differentials in the overall economy are larger than those typically found at a single employer because the heterogeneity in job types between employers is much greater than the heterogeneity of job types for a given employer. Second, because the overall workforce is more heterogeneous than the workforce of a given employer, most effects are estimated more precisely in the March CPS than they would be in a sample drawn from a single employer.

Because the CPS data are more heterogeneous than microdata from a single employer, they permit estimation of models that strongly resemble the ones that might be used by EEOC to screen EEO-1 reports that included wage data developed according to either of the two pilots recommended in this report (see Chapter 6). And because they allow a plausibly "best-case" power analysis, it is reasonable to consider them before investing heavily in data that might permit a more precise answer.

To minimize the effects of different definitions of the wage rate, we selected previous-year wage and salary earners only. The selected individuals were full-time employees (at least 35 hours/week) for at least 50 weeks in 2009 (the reference year for the March 2010 CPS supplement) and were between the ages of 16 and 75. We coded these individuals into the appropriate gender and race/ethnicity categories corresponding to the EEO-1 form. The design of these log wage differentials has 13 degrees of freedom. We used the major occupation codes (a taxonomy of 10 occupation groups) and the detailed occupations (a taxonomy of about 500 categories).[2] The use of 10 major occupation code categories is a reasonable proxy for the EEO-1 occupations for the purposes of these power studies.

In addition to occupation categories, we also used 16 educational categories. These were entered as control variables in some analyses and used in combination with age to create a measure of time since leaving school, which is called "potential experience."

Analyses based on the public-use CPS data are necessarily between-employer estimates, rather than within employer estimates, as any analysis of EEO-1 wage data would be. We included a control for major industry (13 categories) to allow the power analyses to be closer to those that a full

[2]We chose this approach because a standardized recoding of the CPS occupational codes to EEO-1 categories would have involved about as much measurement error as the error associated with the coding to major and detailed occupations in the first place.

pilot might produce. Model 1 controls for occupation only; Model 2 controls for occupation and covariates; Model 3 controls for detailed occupations and covariates. Figure 4-1 compares the estimates of the three models.

Model 1, shown in the Table 4-1 below, estimates the EEO-1 differentials within major occupational categories. It corresponds to the test $\beta_1 = 0$ conditioning on main effects only for the major occupational group. Not surprisingly, relative to the base group of white non-Hispanic males, all of the estimated differentials are large. Jointly, the *F*-test rejects $\beta_1 = 0$ with a *P*-value of less than 0.0001, and individually all of the differentials are statistically significant at the 0.05 level or higher. The R^2 for this equation is 0.25, and the residual variance is 0.37. These two statistics are also used in the power analysis.

The first power analysis asks what the minimum employer size would be in order to detect differentials as large as those in Model 1 and with employer-specific data that had the same design and explanatory power. The line labeled "Controls EEO-1 Occupation Only" answers this question. All power analyses assume that the basic *F*-test has size 0.05 at $\beta_1 = 0$: that is, the probability of rejecting a true null hypothesis is fixed at 0.05 throughout.

A regression analysis of an employer with approximately 99 employees has power of 0.50: it is equally likely to accept or reject the null hypothesis $\beta_1 = 0$ for wage differentials on the magnitude of those in Model 1. Employment of 200 is needed to boost the power to 0.90, a value that is often used as the standard for acceptable power.[3]

Model 2, shown in Table 4-2 computes the EEO-1 log wage differentials with controls for main effects of the major occupation category as well as main effects of education, major industry, and a quartic in potential experience. The estimated log wage differentials are much smaller than in Model 1, although still quite substantial in magnitude. The *F*-test for the joint significance is 238.41 with a *P*-value less than 0.0001. The R^2 for this equation is 0.39, and the residual variance is 0.30. As can be seen in Figure 4-1, an analysis based on 155 employees delivers power of 0.50 in this case, and an analysis of an employer of size 318 is required for power of 0.90.

Model 3 is shown in Table 4-3 below. In this estimation, we control for detailed occupation in addition to the covariates that were included in Model 2. The *F*-statistic falls to 138.38 but with a *P*-value that is still less than 0.0001. Estimated differentials also fall substantially. The R^2 for

[3]All model estimation was conducted in SAS (statistical analysis software) version 9.3 using PROC GLM. All power analysis was conducted in SAS version 9.3 PROC GLMPOWER. The design matrices, estimated subgroup means, and regression summary statistics used in the power analysis were computed from the March CPS data in the statistical summaries shown in all three of our models.

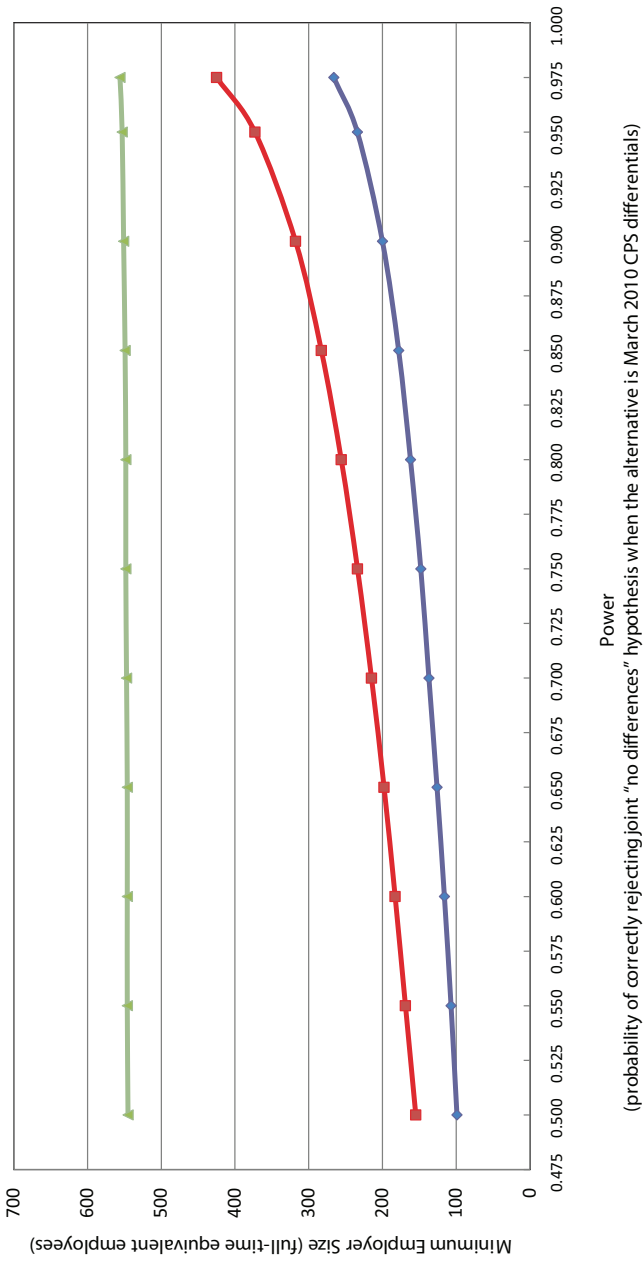

FIGURE 4-1 Comparisons of analytic power and employer size for selected EEO-1 wage reports, three models.
NOTE: See Tables 4-1, 4-2, and 4-3 and text discussion of these models.
SOURCE: Analysis by panel using Current Population Survey data.

TABLE 4-1 Base Model for Estimating EEO-1 Log Wage Differentials (Current Population Survey, March Supplement 2010)

	Model 1			
Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept [base is white (only) non-Hispanic male]	10.57427	0.010841	975.37	<.0001
Hispanic male	–0.31926	0.009651	–33.08	<.0001
Hispanic female	–0.53986	0.011632	–46.41	<.0001
White (only) non-Hispanic female	–0.35903	0.006372	–56.35	<.0001
Black or African American (only) non-Hispanic male	–0.24208	0.011809	–20.5	<.0001
Black or African American (only) non-Hispanic female	–0.46951	0.011104	–42.28	<.0001
Native Hawaiian Islander or Other Pacific Islander (only) male	–0.15631	0.072491	–2.16	0.0311
Native Hawaiian Islander or Other Pacific Islander (only) female	–0.36209	0.07278	–4.98	<.0001
Asian (only) male	–0.03405	0.015258	–2.23	0.0257
Asian (only) female	–0.23185	0.017217	–13.47	<.0001
American Indian or Alaska Native (only) male	–0.18747	0.04766	–3.93	<.0001
American Indian or Alaska Native (only) female	–0.61771	0.046857	–13.18	<.0001
Two or more races male	–0.13671	0.034945	–3.91	<.0001
Two or more races female	–0.38639	0.037565	–10.29	<.0001
	DF Model	DF Error	F Value	Pr > F
EEO-1 differentials	13	62001	410.19	<.0001

NOTE: Controls for major occupation only (10 categories).
SOURCE: Analysis by panel using Current Population Survey data.

TABLE 4-2 Model for Estimating EEO-1 Log Wage Differentials Controlling for Education, Major Industry, and Potential Experience (Current Population Survey, March Supplement 2010)

	Model 2			
Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept [base is white (only) non-Hispanic male]	10.76643118	0.02673094	402.77	<.0001
Hispanic male	–0.14657715	0.00912375	–16.07	<.0001
Hispanic female	–0.35794272	0.01073877	–33.33	<.0001
White (only) non-Hispanic female	–0.27917604	0.00593934	–47	<.0001
Black or African American (only) non-Hispanic male	–0.18823432	0.01062293	–17.72	<.0001
Black or African American (only) non-Hispanic female	–0.36062882	0.01018799	–35.4	<.0001
Native Hawaiian Islander or Other Pacific Islander (only) male	–0.08203944	0.06499172	–1.26	0.2068
Native Hawaiian Islander or Other Pacific Islander (only) female	–0.30479654	0.06524491	–4.67	<.0001
Asian (only) male	–0.08434892	0.01375719	–6.13	<.0001
Asian (only) female	–0.20779203	0.01551098	–13.4	<.0001
American Indian or Alaska Native (only) male	–0.12243174	0.04276004	–2.86	0.0042
American Indian or Alaska Native (only) female	–0.4567789	0.04207129	–10.86	<.0001
Two or more races male	–0.0878382	0.03133437	–2.8	0.0051
Two or more races female	–0.27586943	0.03371475	–8.18	<.0001
	DF Model	DF Error	F Value	Pr > F
EEO-1 differentials	13	61970	238.41	<.0001

NOTE: Controls for major occupation (10 categories), education (16 categories), major industry (13 categories), and potential experience (quartic).
SOURCE: Analysis by panel using Current Population Survey data.

TABLE 4-3 Model for Estimating Detailed Occupational Log Wage Differentials Controlling for Education, Major Industry, and Potential Experience (Current Population Survey, March Supplement 2010)

	Model 3			
Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept [base is white (only) non-Hispanic male]	10.86089272	0.10171189	106.78	<.0001
Hispanic male	–0.10250055	0.00874789	–11.72	<.0001
Hispanic female	–0.26943452	0.0104887	–25.69	<.0001
White (only) non-Hispanic female	–0.22408845	0.00603451	–37.13	<.0001
Black or African American (only) non-Hispanic male	–0.12759486	0.01016768	–12.55	<.0001
Black or African American (only) non-Hispanic female	–0.27720803	0.00998411	–27.76	<.0001
Native Hawaiian Islander or Other Pacific Islander (only) male	–0.06154783	0.06152922	–1	0.3172
Native Hawaiian Islander or Other Pacific Islander (only) female	–0.22666618	0.06175392	–3.67	0.0002
Asian (only) male	–0.07824855	0.01319245	–5.93	<.0001
Asian (only) female	–0.17077671	0.01501089	–11.38	<.0001
American Indian or Alaska Native (only) male	–0.10341199	0.04060558	–2.55	0.0109
American Indian or Alaska Native (only) female	–0.37084016	0.0398942	–9.3	<.0001
Two or more races male	–0.08577552	0.02967712	–2.89	0.0039
Two or more races female	–0.22016021	0.03197085	–6.89	<.0001
	DF Model	DF Error	F Value	Pr > F
EEO-1 differentials	13	61483	138.38	<.0001

NOTE: Controls for detailed occupation (497 categories), education (16 categories), major industry (13 categories), and potential experience (quartic).
SOURCE: Analysis by panel using Current Population Survey data.

this equation is 0.47, and the residual variance is 0.26. As can be seen in Figure 4-1, 545 employees are required for a power of 0.50 in this case, while about the same sample size (551 employees) yields a power of 0.90. The power curve for this model is flat because there are 496 degrees of freedom for the detailed occupation controls. Once there are adequate data to fit this model, about 50 additional observations are needed to achieve the target power for the EEO race and gender test.

MINIMIZATION OF REPORTING BURDEN

Estimation of Burden

One reason for the outcry on the part of the business community when the Paycheck Fairness Act was under consideration in Congress was the perception that the legislation would impose a significant new reporting burden on employers, particularly on small employers. The Paperwork Reduction Act of 1995 specifically requires agencies to demonstrate the practical utility of the information that they propose to collect and to balance this against the burden imposed on the public.

EEOC currently calculates the cost and burden of its data collections in its submissions of Information Collection Requests to the U.S. Office of Management and Budget (OMB). The number of respondents (including multi-establishment respondents), responses (usually at the establishment level), estimated burden hours, costs, and mode of collection for the four major EEO data collections in the most recent reports of EEOC to OMB are shown in Table 4-4.

The estimates of burden costs and hours in Table 4-4 are based on the EEOC's best estimates of the amount of time it takes for clerks to retrieve and enter the data to paper records. However, because less than one-fourth of employers who report now file paper records, the burden estimates may be overstated.

Options for Minimizing Response Burden

To the extent that the current burdens data are representative, the addition of earnings data to the existing EEOC data collection forms that do not now collect the data, in much the same manner in which earnings data are collected in the EEO-4 form, could be expected to nearly double the current burden on employers. In the case of the largest collection, the current average of 3.5 hours per EE0-1 form might increase to somewhere near the average of 6.6 hours now reported for the EE0-4 form. This is not an inconsequential increase in response burden. It would behoove EEOC to consider taking steps to reduce the increase in response burden.

TABLE 4-4 Estimated Cost and Burden of EEOC Data Collections

Form	Frequency	Respondents	Responses	Estimated Burden Hours	Estimated Cost	Percent Electronic Reported
EEO-1	Annual	45,000	170,000	599,000	$11,400,000	80
EEO-3	Biannual	1,399	1,399	2,098	85,000	79
EEO-4	Biannual	6,018	6,018	40,000	700,000	76
EEO-5	Biannual	1,135	1,135	10,000	190,000	58

SOURCE: Data from EEOC Form 83-I submissions to OMB.

Several options are available for reducing the burden on reporters. Three are discussed in this section—less frequent data collection, use of a rotating scheme for certain employer size classes, and raising the size cutoff so that fewer employers would be in the scope of the collection.

Less Frequent Collection

The EEO-1 report is now collected annually, while the other forms are collected on a biannual basis. The main issue is with the EEO-1 form. The law does not require the annual collection of EEO-1 data. The timing of collection is an administratively imposed requirement. By administratively reducing the frequency of data collection, the burden might also be reduced, though the extent to which it might be reduced is not entirely clear.

On the negative side, the less frequent availability of the reports would mean that the information that supports EEOC enforcement functions would be less current, by a year or so. This lag could be an important issue during economic turning points, when hiring or layoffs could significantly influence the employment and earnings profiles of covered firms. The time lag for EEOC's investigations of potential discrimination would increase, and the ability of the agency to be responsive to complaints in a timely manner would be negatively affected.

Rotating Sample

It might be possible to continue to collect data annually but from only a part of the current reporting population and to permit firms with certain characteristics, such as not meeting a threshold size or in a selected industry group, to report less frequently. The selection of annual versus biannual reporters could, for example, be based on an analysis by EEOC of the probability of discrimination based, in turn, on the experience of the agency with enforcement. This tailored approach to selection of those firms that could report less frequently, however, would be hard to administer and could well be difficult to implement fairly in practice.

Moreover, this nuanced approach might actually complicate matters for employers. Because so many firms automate their reporting, it is now a routine matter, and rotating the reporting requirement might actually increase the administrative burdens. Employers would need to figure out when they needed to report, and the task of developing a database to capture the reports might be much more burdensome for EEOC.

Raising the Size Cutoff

The current employment cutoff for the annual requirement to submit an EEO-1 form is 100 employees (50 employees if the firm is a federal government contractor). This cutoff limits the overall potential response burden significantly. By raising the size cutoff to, say, 200 employees (based on the statistical power analysis presented above), the number of firms that would have to report earnings would be reduced by half, but the employment coverage would be reduced by less than 10 percent (see Table 1-1, in Chapter 1). One consequence of raising the cutoff size would be a relative reduction in coverage of the earnings of females and minorities. The firms in the size classes for which the reporting requirement would be eliminated are those in which women and minorities are more heavily represented. Experiments with different cutoff sizes to better determine the tradeoffs between burden and coverage could be useful to include in the pilot study that the panel recommends (see Chapter 6).

HUMAN RESOURCE AND PAYROLL SYSTEMS

Most companies of the size covered by EEO regulations have at least somewhat automated payroll and human resource management systems. Today, larger companies are more able to comply with a potential requirement for compensation data by gender, race, and national origin because they can gather compensation information from automated payroll systems and demographic data from automated human resource systems.

The panel reviewed the state of automation of company payroll systems from the perspective of three service providers—a large payroll-providing service firm, a firm that specializes in the emerging software-as-a-service (SaaS) market, and a firm that specializes in using companies' own internal data to analyze EEO status and prepare Affirmative Action Plans for those companies. In summary, we found that automated systems were expanding rapidly among U.S. employers, but that there are differences in the extent of implementing these applications by size of firm.

Currently, larger firms are likely to have human resource and payroll management systems, and they are likely to have an easier time in complying with a new requirement to provide compensation data by demographic characteristics than would smaller firms. Over time, one would expect that the use of such systems will grow and spread among smaller firms. In the long term, these automated systems may well serve as the basis for EEOC employment and wage data collection. As discussed in Chapter 6, the panel recommends a pilot test to collect information on the extent of penetration of these human resource and automated systems: see Appendix C.

Payroll and Human Resource Providers

The industry of payroll and human resource providers is characterized by a growth in services beyond the usual provision of timekeeping and payroll functions. Most recently, the industry has expanded to include human resource management. As a result, one provider can bring together information on hours, earnings, and the demographics and work histories of the workforce. These data are captured directly from a client's data systems, often without client intervention.

The panel interviewed a large payroll-providing company to determine the influence of the growth of this sector on the reporting of earnings data to EEOC. This company lists 600,000 clients, representing, in the company's estimation, one of every six U.S. employees. The clients employ as few as 1 and as many as 1 million employees.

The company has a line of business that focuses on smaller employers—those with fewer than 100 employees—to provide a total source of payroll and human resource services. The company estimates that about 40 percent of these smaller employers use human resource services as well as payroll services. One product for the clients who use human resource services and who have an OFCCP or EEOC requirement is to produce EEO-1 reports.

Growth of Software-as-a-Service Applications

The workshop presentation by Karen Minicozzi of Workday, Inc., representing an enterprise software solution, highlighted the unified human capital management solutions offered by the enterprise software and services provider, Workday, Inc. The company is one of a growing number of firms that provide turnkey payroll and human resource management solutions to businesses under the general label of SaaS. The solutions provide a new, global core system of record to replace legacy systems that have been maintained by the establishments themselves. The approach taken by these service providers is through a multitenant architecture: that is, one version of the application with common hardware, networking, and operating systems is used for all customers ("tenants"). The applications are often supported in the "cloud," that is, through Internet connectivity. The fact that these new service approaches have so much in common allows the generation of common reports (such as EEO reports) across the system, drawing on data from both the human resource and payroll functions of the serviced companies. Most of the companies that use this service are mid-size, large, and very large companies. Workday, Inc. has 246 customers.

These SaaS providers have been enjoying remarkable growth. An annual survey of employing establishments by the consulting firm CedarCrestone,

to ascertain the penetration of human resource applications in business, found them to be widespread, and it forecast SaaS as a deployment option will likely continue that growth as organizations move from licensed on-premise solutions to the cloud. The source of this information is the *CedarCrestone 2010–2011 HR Systems Survey.* The survey is based on 1,289 responses, representing employers of over 20 million employees (CedarCrestone, 2011). The survey also found that there were measurable differences in the penetration of these administrative applications by size of firm. In the most recent survey, 94 percent of employers with 10,000 or more employees had such systems, compared with 87 percent for employers with 250 to 2,499 employees. The CedarCrestone survey found that most of the applications were still licensed software, but the subscription-based SaaS applications and outsourcing solutions were growing in use.

Analysis of Salary and Related Data for Pay Equity Purposes

In order to ensure that their firms are in compliance with the Equal Pay Act, Title VII, and Executive Order 11246 provisions, many employers use firms that perform compensation analysis and, in many cases, actually prepare automated Affirmative Action Plans. Other firms use software to support this analysis internally.

The panel heard testimony from Liz Balconi and Michele Whitehead, representatives of Berkshire Associates, a company that is very active in the compensation analysis business. This company obtains the following information from its client firms: employee identifier; job code; race; gender; date of hire; annualized base salary or hourly rate; grade, band, or classification (if applicable); time in current position, or date of last title change; date of last degree earned, or date of birth; full time or part time status; exempt or nonexempt status; title; employee location; years of relevant experience (or date of birth); factors that may legitimately impact pay in an organization, such as performance rating; education; date in grade; professional certifications; division; job group; starting salary; and annualized total compensation (including bonuses, commissions, cost of living allowances, and overtime).

The firm uses these data (which are generally available from their clients) to conduct two kinds of analyses: cohort analysis, which is a nonstatistical comparison of similarly situated incumbents within a group based on factors such as time in the company, educational background, and performance assessment; and statistical (regression) analysis to study the combined effect of factors on pay between comparator groups. Although not all of these data elements may be necessary to identify potentially discriminatory practices, prudent employers can be expected to have these types of data available and to use them to evaluate their own practices, using algorithms developed by specialty firms such as Berkshire Associates.

5

Confidentiality, Disclosure, and Data Access

In contrast to the usual situation in federal government survey data collections—in which the data are available for statistical use but are protected from being used for compliance and enforcement purposes—data on equal employment opportunity (EEO) issues are available for compliance purposes but are closely held and not often made available for research and statistical analysis purposes. This anomalous situation poses interesting challenges to the U.S. Equal Employment Opportunity Commission (EEOC) and the other federal agencies that have responsibility for the data collected from public- and private-sector employers and unions for antidiscrimination enforcement purposes.

In addition to internal EEOC compliance and analytical uses, the data collected from employers have value to other federal and state agencies for their compliance and analytical purposes, to researchers to support analysis of discrimination practices, and to those who evaluate the effectiveness and efficiency of antidiscrimination programs. These uses outside of EEOC require the agency to develop practices and procedures to protect the data that are collected from employers under a pledge of confidentiality.[1]

[1]That pledge derives from Title VII, Section 709(e) of the Civil Rights Act of 1964, which sets the requirements for confidentiality:

> It shall be unlawful for any officer or employee of the Commission to make public in any manner whatever any information obtained by the Commission pursuant to its authority under this section prior to the institution of any proceeding under this subchapter involving such information. Any officer or employee of the Commission who shall make public in any manner whatever any information in violation of this subsection shall be guilty, of a misdemeanor and upon conviction thereof, shall be fined not more than $1,000 or imprisoned not more than one year.

In this chapter we discuss current EEOC procedures for protecting confidential employer data in tabular and microdata form, evaluate the effectiveness of those measures, and suggest possible enhancements to those measures.

STATISTICAL PROTECTION OF TABULAR DATA AND MICRODATA

As discussed in Chapter 1, the EEOC now publishes a large amount of data that are derived from the collection of information from employers, both private and public. These data are generally published in aggregated form by geographic area and industry group detail in standard tabular packages that are posted on the EEOC website and otherwise made available to the public. To comply with the confidentiality provisions of Title VII that govern release of individually identifiable information from EEO-1 reports (see Chapter 1), the tables are assembled under reportedly elaborate but unpublished rules that provide for suppression of data that could identify a particular establishment or multi-establishment firm.

In releasing aggregated data of private employers collected from annual EEO-1 surveys, the EEOC uses a data suppression rule that is quite similar to the rule used by other federal government agencies for statistical data based on information collected from employers, including the Quarterly Census of Employment and Wages (QCEW) Program from the Bureau of Labor Statistics (BLS).[2] The EEOC suppression rule is triggered when it meets the two primary suppression stipulations: (1) the group has three or fewer employers, or (2) one employer makes up at least 80 percent of the group employment in the aggregate.

In applying the suppression rules to industry group or geography entity or any combination of aggregates, the EEOC withholds any group's numbers if the group (an industry or a geography entity or an industry-by-geography group, etc.) contains fewer than three firms (represented by the presence of any number of establishment(s) of an individual firm within the group) or if any one firm in the group (represented by the total numbers of all the establishment(s) of the same firm within the given group) constitutes more than 80 percent of the group totals.

Unlike some other federal agencies, EEOC does not withhold aggregated data beyond its two primary suppression rules. There are no secondary suppression rules, and the agency does not further screen the aggregated data if the data have passed the fewer-than-three rule test. But although

[2]For more information on suppression, see http://www.bls.gov/opub/hom/homch5_d.htm#Presentation [December 2011].

EEOC literature documents the above rules, as a general practice EEOC does not disclose the detailed methodology for suppression because the agency wants to prevent users from reverse-engineering the data in order to obtain the suppressed numbers.

Cell suppression is just one means of protecting tabular data. Because there is always a risk of secondary disclosure, other means have been explored in recent years by U.S. government agencies to protect data by perturbing the data in some way (see Reznek, 2006, p. 3). Two methods are discussed here: adding noise and controlled tabular adjustment.

Noise addition is accomplished by adding random "noise" to the underlying establishment-reported data before they are tabulated. In this data perturbation method, cell values that would normally meet the criteria for suppression are changed by a large amount, while cell values that are not as sensitive are changed by a smaller amount. This technique is less complicated than cell suppression, and, by adding noise, an agency can show data for all cells and for all tables, which preserves the ability to draw inferences from all cells.[3] Another effort to preserve the analytical value of protected sensitive data is being developed using a controlled tabular adjustment technique. In this technique, a sensitivity rule determines which cells are sensitive, and the technique replaces each sensitive value with a safe value that is some distance away from the sensitive value. To preserve additivity, the nonsensitive values are minimally adjusted (Reznek, 2006, p. 5).

Another increasingly popular technique that is intended to make data available for research and analytical purposes is to generate synthetic data: for generation of synthetic microdata, see Reiter (2005); for generation of synthetic tables, see Slovkovic and Lee (2010). This technique relies on sampling and simulations. Typically, a model is developed to generate synthetic or partially synthetic data that have some of the same properties as the original data by sampling from the posterior predictive distribution of the confidential data. A typical method would be to use a sequential regression imputation. In this procedure, the original value of each variable is blanked-out and replaced by a model-generated value. The technique has been used at the Census Bureau to develop a synthesized microdata file linking Social Security Administration earnings data with data from a Census Bureau demographic survey (Reznek, 2006, p. 6).

Creating publicly available data products that are statistically valid and in which confidential data are protected is a complicated process.

[3]The technique is currently being used by the Census Bureau to protect confidential microdata from the Longitudinal Employer-Household Dynamics (LEHD) Program used in the Quarterly Workforce Indicators, which use, as inputs, sensitive data from unemployment insurance wage records and Census Bureau demographic and economic information (Abowd et al., 2006).

The best procedure to use depends on the type of data and their intended purposes, as well as on the risks of disclosure. For an overview of current statistical disclosure limitation practices in the United States, see Federal Committee on Statistical Methodology (2005). Many new techniques are being developed. The most recent ones combine techniques from statistics and computer sciences and aim to account for increased disclosure risk due to the presence of more externally available information and better record linkage technologies. Recent advances in data redaction strategies and data sharing, which include, among others, virtual research data centers, remote access servers, privacy-preserving mechanisms for distributed databases, and differentially private mechanisms, are highlighted in a special 2009 issue of the *Journal of Privacy and Confidentiality* (Kinney, Karr, and Gonzalez, 2009).

PROTECTING ORIGINAL DATA

EEOC Procedures

The actual, original data collected from the forms that employers submit to EEOC are now shared with the Office of Federal Contract Compliance Programs (OFCCP) of the U.S. Department of Labor, the Civil Rights Division of the U.S. Department of Justice (DOJ), and 95 state-level fair employment practices agencies (FEPAs). There are other sharing arrangements with the U.S. Department of Education and with researchers. Often these agencies have their own procedures for assuring the confidentiality of the shared data.

The specific arrangements vary in each instance. For example, OFCCP is a statutory member of the Joint Reporting Committee with EEOC for the collection of the EEO-1 reports. This arrangement is made known in advance to companies that provide their data to the EEOC.[4] According to protocols that are in place for the Joint Reporting Committee, EEOC collects the data, edits them as needed, appends some additional identifiers to the records, and transmits a copy of the entire statistical file to OFCCP.

The DOJ Civil Rights Division is a member of a joint state and local reporting committee with EEOC for the collection of EEO-4 reports (see Chapter 1). As it does with the EEO-1 data, EEOC collects the data and at the conclusion of the survey forwards a copy of the EEO-4 statistical file

[4]The EEO-1 instruction booklet (p. 1) states:
In the interests of consistency, uniformity and economy, Standard Form 100 has been jointly developed by the Equal Employment Opportunity Commission and the Office of Federal Contract Compliance Programs of the U.S. Department of Labor, as a single form which meets the statistical needs of both programs.

to DOJ.[5] It also transmits copies of the actual individual EEO-4 reports directly to DOJ officials, allowing immediate access to reports during the reporting period, as well as access to historical data.

FEPAs are state or local authorities that investigate and resolve charges of employment discrimination filed under Title VII, the Americans with Disabilities Act (ADA), the Age Discrimination Employment Act (ADEA), and comparable state laws and local ordinances in partnership with EEOC. Over the years, EEOC has negotiated work-sharing agreements with these agencies that allow the sharing of data. EEO-1 data are shared routinely in a charge tracking system that EEOC provides, which enables the FEPAs to retrieve the reports and run statistical comparisons. Other data are shared on an ad hoc basis.

Under the auspices of a school reporting committee, the EEOC shares EEO-5 data (see Chapter 1) with DOJ and the U.S. Department of Education. Statistical files are shared with both agencies. Specific requests for EEO-5 data are also honored, most often for DOJ.

From time to time, EEOC has entered into agreements with other federal agencies to allow the sharing of survey data. Currently, the only active agreement is with DOJ to share EEO-1 data. The memorandum of understanding (MOU) agreement, discussed below, spells out strict provisions for the protection of the confidentiality of the data.

The EEOC has also historically entered into agreements with individual researchers to allow the sharing of data: see Box 5-1. This has been a practice of the EEOC since 1969, when EEOC entered into an agreement with Eleanor Brantley Schwartz of Georgia State University to study women in management. The mechanism for sharing data in a protected environment is quite detailed, complicated, and time consuming, and it relies on giving the potential data user the status of a sworn federal employee.

Office of Federal Contract Compliance Programs Procedures

OFCCP confidential data are derived from a "scheduling letter" process in which compliance reviews are initiated and certain documents and data sets are requested. The documents consist of the written Affirmative Action Plan (AAP) for the scheduled facility, certain compensation data, and information on additional personnel practices and policies to demonstrate compliance obligations.

[5]This arrangement is described in the EEO-4 booklet (p. 1):
In the interests of consistency, uniformity and economy, State and Local Government EEO-4 is being used by Federal government agencies that have responsibilities for equal employment opportunity. A joint State and Local Reporting Committee, with which this report must be filed, represents those various agencies.

BOX 5-1
Intergovernmental Personnel Act Agreements with Researchers

EEOC has used Intergovernmental Personnel Act agreements that detail outside persons to an employment arrangement to allow the sharing of survey data. These agreements give the researcher the status of a federal employee and access to the data. The researcher signs an agreement that prohibits disclosure of the data to anyone (including professors, advisers, and colleagues), except those persons directly employed by the project. It also requires the researcher to submit any work based on the EEOC information to the EEOC to (a) determine whether it contains any confidential information and (b) approve any language describing the relationship between the researcher and the EEOC. The data are to be returned to EEOC at the conclusion of the project, and all working files are to be certified as destroyed. The penalties for disclosure of confidential data in Title VII are formally transferred to the researcher.

SOURCE: Summary by panel staff of sample EEOC Intergovernmental Personnel Act agreement for external researchers, provided by EEOC staff on November 28, 2011.

Unlike EEOC, OFCCP has no formal data-sharing arrangement with federal or state agencies. Its data sharing occurs on an ad hoc or informal basis, such as when OFCCP refers cases to DOJ or EEOC to pursue enforcement. Sharing can also occur on a very limited basis under the MOU with EEOC. For data collected only by OFCCP, the past instances of data sharing have been infrequent, although additional sharing with EEOC can be foreseen.

Unlike Title VII of the Civil Rights Act of 1964 as amended, Executive Order 11264, which comprises the legal basis for OFCCP, is silent on rules and penalties for confidentiality of data from employers. However, confidentiality provisions that cover OFCCP are spelled out in the agency's regulations (see 41 C.F.R. 60-1.20(f)–(g) and 60-1.43). The regulations essentially state that the disclosure of data to the public is subject to the Freedom of Information Act and the Privacy Act and also to the procedures for preclusion of certain data due to assertion of privileges during litigation.[6]

[6]OFCCP rules were spelled out in the regulation that authorized the collection of the Equal Opportunity Survey (41 C.F.R. 60-2.18(d)). These rules state:

(d) Confidentiality. OFCCP will treat information contained in the Equal Opportunity Survey as confidential to the maximum extent the information is exempt from public disclosure under the Freedom of Information Act, 5 U.S.C. 552. It is the practice of OFCCP not to release data where the contractor is still in business and, where the contractor indicates, and through the Department of Labor review process it is determined, that the data are confidential and sensitive and that the release of data would subject the contractor to commercial harm.

The OFCCP approach to data confidentiality is evolving in the direction of greater transparency. An example is a new initiative under the umbrella of the Open Government Directive,[7] under which the Department of Labor (DOL) has developed a searchable "enforcement database" comprised of DOL enforcement agencies, including OFCCP.[8] This database is available for viewing by academic researchers, stakeholders, and the public. Users can retrieve data by state or zip code, the company name, North American Industry Classification System codes, violation, and year. The database divides OFCCP data into two categories: evaluations (compliance reviews) and investigations (complaints). In making these administrative data available for the first time, OFCCP has a policy of limiting disclosed information. For example, it provides only data specific to the facility reviewed and only summary data (yes/no) for violations found, if any. However, it should be noted that the true underlying disclosure risks with such data are not fully understood.

Department of Justice Procedures

As noted above, DOJ's Civil Rights Division obtains EEO-4 data from EEOC on a regular basis and holds it in confidence as a member of the joint state and local reporting committee. The DOJ uses the EEO-4 data to identify investigations that it believes should be launched, but it does not use the data directly in the investigation, nor are the data directly used in court cases. Instead, DOJ uses the data collected in the process of discovery to support its litigation.

The transmittal of EEO-1 data from EEOC to DOJ is covered by an MOU that was executed in May 2011.[9] The MOU calls on EEOC to provide DOJ with data for the most recent reporting period as soon as practicable after the EEOC has reconciled and finalized the statistical file. Historical EEO-1 files are also to be provided. In turn, DOJ agrees to preserve the confidentiality of the data in the same manner that EEOC employees are required by Title VII of the Civil Rights Act of 1964 as amended.

Among the steps leading to identification of a possible infringement of EEO laws, the DOJ compares the profiles of the public sector organizations

[7]White House, Memorandum on Transparency & Open Government, M-10-06. December 8, 2009. See http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf [October 2012].

[8]For details, see http://ogesdw.dol.gov [July 2012].

[9]U.S. Equal Employment Opportunity Commission, Memorandum of Understanding Between the U.S. Equal Employment Opportunity Commission and the U.S. Department of Justice–Civil Rights Division for Sharing of Employer Information Report (EEO-1) Data, May 12, 2011.

under the agency's jurisdiction with similar organizations in the private sector, using the EEO-1 data that are obtained from EEOC.

FURTHER PROTECTION OF SHARED EEO DATA

As the above discussion indicates, the EEOC shares sensitive EEO-4 and EEO-1 report data with other agencies in the federal government and with the FEPAs through rather informal arrangements, most of which are not backed by force of law. This practice is in contrast to the usual practice of federal statistical agencies that protect shared data through formal agreements backed by clear legislative authority that is enforced by stern penalties. For EEOC, even when there is an agreement, such as the one with DOJ, to share EEO-1 data, there is no indication that the data are shielded from court challenge or from requests under the Freedom of Information Act when they are shared.

In recent years, a procedure for protecting shared data has been implemented by several federal statistical agencies that might well serve as a model for protecting the EEOC employer data. The Bureau of Labor Statistics, Census Bureau, and Bureau of Economic Analysis can now share confidential data obtained from employers under provisions of the Confidential Information Protection and Statistical Efficiency Act (CIPSEA). This statute, under the umbrella of the U.S. Office of Management and Budget, prohibits disclosure or release, for nonstatistical purposes, of information collected under a pledge of confidentiality. Under this law, data may not be released to unauthorized persons. Willful and knowing disclosure of protected data to unauthorized persons is a felony punishable by up to 5 years imprisonment and up to a $250,000 fine—penalties that are significantly more stringent than those that are enumerated in the Title VII legislation.

It is certain that the sensitivity of the data that employers provide to EEOC will be heightened if earnings data were to be added to the EEO data records. Employee compensation data are generally considered to be highly sensitive; they are even considered proprietary information by many private-sector employers.

As this chapter points out, EEOC provides data to agencies that do not have the same level of confidentiality protections and are not covered by the same penalties that apply to EEOC employees and researchers under Interagency Personnel Act (IPA) agreements. Legislation patterned after the CIPSEA law could increase the protection of confidentiality of EEO data, specifically, to authorize sharing agreements between EEOC, OFCCP, DOJ, and the state and local FEPAs and extend the Title VII penalties beyond EEOC and its IPA researchers.

Such protection could be expected to increase the willingness of employers to provide detailed employment data. It could also help mitigate concerns of other federal agencies about the matching of the EEO-1 survey records to administrative data (such as those discussed in Chapter 2) if such matching was some day deemed useful to help improve the quality of the data.

6

Conclusions and Recommendations

The panel was invited to make recommendations to assist the U.S. Equal Employment Opportunity Commission (EEOC) in formulating regulations on methods for measuring and collecting pay information by gender, race, and national origin from U.S. employers for the purpose of administering Section 709 of the Civil Rights Act of 1964 as amended, if a decision is made to proceed with such a data collection. We have considered currently available and potential data sources, as well as methodologies and statistical techniques for the measurement and collection of such employer pay data. The panel's recommendations are made with an appreciation that such a new data collection would be a significant undertaking for EEOC and that it could well generate an increased reporting burden on some employers, and so any new data collection would have to be fully justified.

PURPOSE OF A NEW DATA COLLECTION

Based on the literature we reviewed and the papers and presentations made to us by the staff of EEOC, the Office of Federal Contract Compliance Programs (OFCCP) in the U.S. Department of Labor, and the U.S. Department of Justice, the panel finds that there is no clearly articulated vision of how data on wages would be used in the conduct of the enforcement responsibilities of these agencies. As discussed in Chapter 1, the use of the employment data from the EEO-1 reports for the purpose of targeting potentially noncompliant firms was highlighted by EEOC and OFCCP leadership as an objective of the collection of earnings data by gender, race, and national origin. Thus, targeting is broadly given as the objective

of collection of earnings data by both OFCCP and EEOC, but the specific mechanisms by which the data would be assembled, assessed, compared, and used in a targeting operation are not well developed by either agency. The panel found no evidence of a clearly articulated plan for using the earnings data if they are collected: the fundamental question that would need to be answered is how earnings data should be integrated into the compliance programs that have to date been triggered mainly by a complaint process, which, in their absence, includes relatively few complaints about pay matters.

With regard to existing studies of the cost-effectiveness of an instrument for collecting wage data, the panel concludes that they are inadequate to assess any new survey program. For example, unless the agencies have a comprehensive plan that includes the form of the data collection, it will not be possible to reliably determine the actual burden on employers and the costs and benefits of the collection.

As discussed in Chapter 3, it is important to clearly understand the requirement and potential uses of data as a first step in determining their fitness for use, that is, the quality of the data. Although it is assumed that, if these data are collected, they could greatly enhance the enforcement process, until EEOC and its cooperating agencies gain experience with collecting, processing, and using earnings data in field investigations and in litigation, it will not be known if the data are of sufficient reliability to support enforcement.

Other potential benefits of the possible collection of pay data remain to be fully articulated but are of interest. In addition to targeting, the collection of earnings data could well be used in research on discrimination and pay equity. Analysts would be able to associate pay differentials by type of establishment, location, job category (occupation), and demographic detail, which cannot currently be done with existing data. For such use, however, systems for maintaining, retrieving, archiving, and processing the data in a protected environment would have to be developed.

Recommendation 1: In conjunction with the Office of Federal Contract Compliance Programs of the U.S. Department of Labor and the Civil Rights Division of the U.S. Department of Justice, the U.S. Equal Employment Opportunity Commission should prepare a comprehensive plan for use of earnings data before initiating any data collection.

PILOT STUDY

With a comprehensive plan in hand, the next logical step would be to test it. Because of the current paucity of evidence about such a data collection, the panel concludes that reliable information about the costs and

benefits of the proposed collection would best be provided by an independent pilot study. The panel's two-pronged approach to conducting a pilot study (to be done by an independent contractor) is outlined below and detailed in Appendix C.

The first approach—a microdata pilot test—would collect a number of core demographic variables—using the categories on the Equal Employment Opportunity (EEO)-1 form and adding an annual wage measure for individual employees. This approach would test targeting firms for enforcement purposes, as well as testing the collection of additional variables that could illuminate the relevant characteristics of targeted firms. For example, age and years-on-the-job variables could assist in controlling for the legitimate effect of these characteristics on wages. In developing the test, the public responses to the OFCCP ANPRM could well be instructive.

The second approach—a simplified aggregated-data pilot test—would develop and test an enhanced EEO-1 report that would include all the summary data required for the computation of test statistics comparing wage data in existing EEO-1 occupations. This pilot would use grouped data techniques that would produce standardized wage rates and other measures of interest.

Both approaches to the pilot study could test various earnings definitions. On the basis of our analysis, we conclude that the definition used in the Occupational Employment Statistics (OES) survey is the most feasible (see under heading Definition of Compensation). The tests could also assess the possibility of reducing employers' response burden through building in compatibility with the electronic record-keeping systems that are now in use in larger companies.

The quality of the data from the pilot tests would have to be assessed in light of the analysis plan that results from Recommendation 1. It would also be desirable for the quality of the data collected in the pilot to be verified by independent record checks of reporting establishments or by comparison of aggregated results with administrative databases (see Chapter 2), again using the criteria developed as part of the analysis plan in Recommendation 1.

Recommendation 2: After the U.S. Equal Employment Opportunity Commission, the Office of Federal Contract Compliance Programs, and the U.S. Department of Justice complete the comprehensive plan for use of earnings data, the agencies should initiate a pilot study to test the collection instrument and the plan for the use of the data. The pilot study should be conducted by an independent contractor charged with measuring the resulting data quality, fitness for use in the comprehensive plan, cost, and respondent burden.

AGENCY CAPACITY AND BURDEN

It is important to consider the administrative capacity for the collection, analysis, and protection of pay data. The EEOC has a small data collection and analytical program, which has traditionally been focused nearly exclusively on collecting employment data and assessing employer compliance through the means of rather straightforward statistical tests.

If EEOC undertakes a major new activity, it is not clear that it could administratively handle the work given available resources. If data on compensation is added to an existing form, or collected in a new instrument, the agency's resources for both collection and analysis are likely to be severely strained. Thus, EEOC needs to consider its capacity to undertake any new collection. To take full advantage of new opportunities for analytics and compliance using more sophisticated measures enabled by the availability of detailed earnings data will surely require an enhancement of EEOC's analytical and data processing capacity, as well as its capability to protect the confidentiality of the information.

Recommendation 3: The U.S. Equal Employment Opportunity Commission should enhance its capacity to summarize, analyze, and protect earnings data.

MEASURES FOR COLLECTION OF PAY INFORMATION

Several possible measures of pay information could be used for the possible new data collection, ranging from pay bands (the measure now used on the EEO-4 form) to rates of pay (e.g., annual salaries, hourly wages, etc.). Though pay band collection is attractive in that it aligns with the way that human resource managers tend to look at compensation, the best data are collected from payroll records, and those data are most likely to be rates of pay or average annual earnings as computed using total wage and hours information. Rates of pay as a measure have the advantage of being more likely to provide valid measures of both central tendency and dispersion, important quality checks and analytical capabilities that pay band data cannot provide. Rates of pay collection would add rigor to the collection process and subsequent analysis.

Recommendation 4: The U.S. Equal Employment Opportunity Commission should collect data on rates of pay, not actual earnings or pay bands, in a manner that permits the calculation of measures of both central tendency and dispersion.

DEFINITION OF COMPENSATION

A number of definitions of compensation are currently in use, ranging from comprehensive measures of total compensation to simple straight-time hourly pay. As noted above and in Chapter 3, we conclude that the best definition is that in the OES survey, and we urge that a test of collection of data from employers by gender, race and national origin be conducted as part of the pilot test program.

As noted in Chapter 3, earnings in the OES survey are defined as straight-time, gross pay, exclusive of premium pay. The definition includes a base rate of pay, cost-of-living allowances, guaranteed pay, hazardous-duty pay, incentive pay (including commissions and production bonuses), and tips. The definition excludes overtime pay, severance pay, shift differentials, nonproduction bonuses, employer cost for supplementary benefits, and tuition reimbursements.

Earnings data by occupation are collected in the OES survey with use of this definition from more than 1.2 million establishments in the United States with response rates of nearly 80 percent. Clearly, most of the firms that fall within the scope of the EEO statutes and are now required to complete an annual EEO-1 report have the ability to provide these data from their existing payroll and human resource systems.

With the growth of highly sophisticated electronic systems, such as those represented in software-as-a-service applications, the ability to transfer data efficiently between the payroll and human resource systems is expected to expand in the future. By monitoring these quickly changing software developments and continuing its work with reporting employers, EEOC could capitalize on advances in electronic reporting.

ACCESS TO PAY INFORMATION IN A PROTECTED ENVIRONMENT

If the pilot tests and other developmental activities recommended in this report bear fruit and if EEOC begins collection of pay data from employers, the data will comprise an important new source of information for research and analytical purposes, in addition to their intended use in enforcement. We expect that there will be great demand on the part of other federal agencies, researchers, analysts, compensation-setting bodies, and others for access to these powerful new data. EEOC would be well advised to start taking steps now to develop policies to provide access in a protected environment.

Recommendation 5: In anticipation of increased user demand for microdata on pay information by demographic detail for research and

analytical purposes if such data are collected by the U.S. Equal Employment Opportunity Commission, the agency should consider implementing appropriate data protection techniques, such as data perturbation and the generation of synthetic data, to protect the confidentiality of the data, and it should also consider supporting research for the development of these applications.

Though there have been no known breaches of the EEOC's ability to protect EEO data, the consequences of a breach in the protection of data provided in confidence are, as other federal agencies have discovered, painful and of lasting consequence. Thus, EEOC should consider providing the same protections to the organizations and individuals that become parties to data-sharing agreements as it now has with its own employees.

Recommendation 6: The U.S. Equal Employment Opportunity Commission should seek legislation that would increase the ability of the agency to protect confidential data. The legislation should specifically authorize data-sharing agreements with other agencies with legislative authority to enforce antidiscrimination laws and should extend Title VII penalties to nonagency employees.

References

Abowd, J.M., and Hallock, K.F. (2007). *Wages v. Total Compensation: Data Quality Issues and Proposals*. Prepared for the Princeton Data Improvement Initiative, February 6. Available: http://harris.princeton.edu/admin/pdfs/abowd-hallock-20070206.pdf [October 2012].

Abowd, J.M., Stephens, B., Vilhuber, L., Andersson, F., McKinney, K., Roemer, M., and Woodcock, S. (2006). *The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators*. LEHD Technical Paper No. 2006-01. Available: http://lehd.did.census.gov/library/techpapers/tp-2006-01.pdf [October 2012].

Abt Associates, Inc. (2005). An Evaluation of OFCCP's Equal Opportunity Survey. Cambridge, MA: Abt Associates, Inc.

Ashenfelter, O., and Heckman, J. (1976). Measuring the effect of an anti-discrimination program. Pp. 46–89 in *Evaluating the Labor-Market Effects of Social Programs*, O. Ashenfelter and J. Blum, eds. Princeton, NJ: Princeton University Press.

Bendick, M., Jr. (2000). *Using EEO-1 Data to Analyze Allegations of Employment Discrimination*. Presentation to the American Bar Association National Conference. Available: http://www.meetings.abanet.org/labor/lel-aba-annual/papers/2000/bendick.pdf [July 2012].

Bendick, M., Jr., Miller, J., Blumrosen, A., and Blumrosen, R. (2000). *The Equal Opportunity Survey: Analysis of a First Wave of Survey Responses*. Washington, DC: Bendick and Egan Economic Consultants.

Blinder, A. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8(4):436–455. University of Wisconsin Press. Available: http://www.jstor.org/stable/144855 [July 2012].

Brackstone, G. (1999). Managing data quality in a statistical agency. *Survey Methodology* 25:139–149.

Bureau of Labor Statistics. (2009). *Occupational Employment and Wages, May 2009*. USDL-10-0646. [News release.] Washington DC: U.S. Department of Labor. Available: http://www.bls.gov/news.release/archives/ocwage_05142010.pdf [October 2012].

Bureau of Labor Statistics. (2010a). *Occupational Employment and Wages, May 2010*. USDL 11-0722. Available: http://www.bls.gov/news.release/ocwage.htm [September 2011].

Bureau of Labor Statistics. (2010b). *Survey Methods and Reliability Statement for the May 2010 Occupational Employment Statistics Survey.* Available: http://www.bls.gov/oes/current/methods_statement.pdf [October 2010].

Bureau of Labor Statistics. (No date). *Handbook of Methods*. Washington, DC: U.S. Department of Labor. Available: http://www.bls.gov/opub/hom/homch8.htm [October 2012].

CedarCrestone. (2011). *CedarCrestone 2010–2011 HR Systems Survey: HR Technologies, Service Delivery, Approaches, and Metrics, 13th Annual Edition*. Available: http://www.cedarcrestone.com/media/whitepapers/CC_2010-2011_HRS_Survey_WP.pdf [October 2012].

Chay, K. (1998). The impact of federal civil rights policy on black economic progress: Evidence from the equal opportunity act of 1972. *Industrial Labor Relations Review* 51(4):608–632.

Consad Research Corporation. (2009). *An Analysis of Reasons for the Disparity of Wages between Men and Women*. Available: http://www.consad.com/content/reports/Gender%20Wage%20Gap%20Final%20Report.pdf [May 2012].

Dobbin, F., Kalev, A., and Kelly, E. (2006). Best practices or best guesses?: Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review* 71(August):589–617.

Donohoe, J., and Levitt, S. (2001). The impact of race on policing and arrests. *Journal of Law and Economics* 44(2):367–394.

Federal Committee on Statistical Methodology. (2005). *Report on Statistical Disclosure Methodology*. Working Paper No. 22 (2nd Version), Confidentiality and Data Access Committee, Office of Management and Budget, Executive Office of the President. Washington, DC. Available: http://www.fcsm.gov/working-papers/wp22.html [July 2012].

Follett, R.S., Ward, M.P., and Welch, F. (1993). Problems in assessing employment discrimination. *The American Economic Review* 83(2):73–78. (Papers and Proceedings of the Hundred and Fifth Annual Meeting of the American Economic Association.)

Gastwirth J.L., and Greenhouse S. (1987). Estimating a common relative risk: Application in equal employment. *Journal of the American Statistical Association* 82(397):38–45.

Gastwirth, J.L., and Haber, S. (1976). Defining the labor market for equal opportunity standards. *Monthly Labor Review* 99(March):32–36.

Glover, T. (2011). Examining the gender wage gap among new hires in Wyoming's manufacturing industry. *Wyoming Labor Force Trends* 48(8). (Wyoming Department of Workforce Services.) Available: https://doe.state.wy.us/lmi/trends/0811/toc.htm [September 2012].

Goldstein, M., and Smith, R. (1976). The estimated impact of the antidiscrimination program aimed at federal contractors. *Industrial Labor Relations Review* 29(3):524–543.

Griffin, P. (1992). The impact of affirmative action on labor demand: A test of some implications of the Le Chatelier principle. *The Review of Economics and Statistics* 74(2):251–260. Available: http://www.jstor.org/stable/2109656 [October 2011].

Hirsh, E. (2008). Settling for less? The organizational determinants of discrimination-charge outcomes. *Law and Society Review* 42(2):239–274.

Hirsh, E. (2009). The strength of weak enforcement: The impact of discrimination charges on sex and race segregation in the workplace. *American Sociological Review* 74(2):245–271.

Hirsh, E., and Kmec, J. (2009). Human resource structures: Reducing discrimination or raising rights awareness? *Industrial Relations* 48(3):512–532.

Hirsh, E., and Kornrich, S. (2008). The context of discrimination: Workplace conditions, institutional environments, and sex and race discrimination charges. *American Journal of Sociology* 113(5):1,394–1,432.

Holzer, H., and Neumark, D. (2000a). Assessing affirmative action? *Journal of Economic Literature* 38(3):483–568.

Holzer, H., and Neumark, D. (2000b). What does affirmative action do? *Industrial and Labor Relations Review* 53(2):240–272.

Huffman, M., Cohen, P., and Pearlman, J. (2010). Engendering change: Organizational dynamics and workplace gender desegregation, 1975–2005. *Administrative Science Quarterly* 55(2):255–277.

Kalev, A. (2009). Cracking the glass cages? Restructuring and ascriptive inequality at work. *American Journal of Sociology* 114(6):1,591–1,643.

Kalev, A., and Dobbin, F. (2009). Enforcement of civil rights law in private workplaces: The effects of compliance reviews and lawsuits over time. *Law & Social Inquiry* 31(4):855–903.

Kellough, J.E. (1990a). Federal agencies and affirmative action for blacks and women. *Social Science Quarterly* 71:83–92.

Kellough, J.E. (1990b). Integration in the public workplace: Determinants of minority and female employment in federal agencies. *Public Administration Review* 50(5):557–566.

Kerr, B., Reid, M., and Miller, W. (1999). A national study of gender-based occupational segregation in municipal bureaucracies: Persistence of glass walls. *Public Administration Review* 59(3):218–230.

Kerr, B., Reid, M., and Miller, W. (2000a). Occupational segregation in municipal government bureaucracies: Persistence of glass ceilings. *Women in Politics* 21(1):35–53.

Kerr, B., Reid, M., and Miller, W. (2000b). The changing face of urban bureaucracy: Is there inter-ethnic competition for municipal government jobs? *Urban Affairs Review* 35(July):770–793.

Kerr, B., Reid, M., and Miller, W. (2002). Sex-based occupational segregation in U.S. state bureaucracies, 1987–1997. *Public Administration Review* 62(July/August):412–423.

Kerr, B., Reid, M., and Miller, W. (2003). *Glass Walls and Glass Ceilings: Women's Employment Progress in State and Municipal Bureaucracies*. Westport, CT: Praeger Press.

Kerr, B., Reid, M., and Miller, W. (2004). Sex-based glass ceilings in U.S. state-level bureaucracies, 1987–1997. *Administration and Society* 36(4):377–405.

Kinney, S.K., Karr, A.F., and Gonzalez, J.F. (2009). Data confidentiality: The next five years. *Journal of Privacy and Confidentiality* 1(2):125–134.

Leonard, J.S. (1984a). Anti-discrimination or reverse discrimination: The impact of changing demographics, Title VII and affirmative action on productivity. *Journal of Human Resources* 19(2):145–174.

Leonard, J.S. (1984b). The impact of affirmative action on employment. *Journal of Labor Economics* 2(4):439–463.

Leonard, J.S. (1990). The impact of affirmative action regulation and equal employment opportunity law on black employment. *Journal of Economic Perspectives* 4(4):47–63.

McCrary, J. (2007). The effect of court-ordered hiring quotas on the composition and quality of police. *American Economic Review* 97(1):318–353.

Miller, A., and Segal, C. (2011). Does temporary affirmative action produce persistent effects? A study of black and female employment in law enforcement. *Review of Economics and Statistics* May. Available: http://www.mitpressjournals.org/doi/abs/10.1162/REST_a_00208 [July 2012].

Mincer, J. (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research. New York: Columbia University Press.

Moore, M. (2011). Wyoming new hires: Examining the wage gap. *Wyoming Labor Force Trends* 48(3):1–10.

Naff, K. (2001). *To Look Like America: Dismantling Barriers for Women and Minorities in the Federal Civil Service*. Boulder, CO: Westview Press.

National Equal Pay Enforcement Task Force. (2010). *Report of the National Equal Pay Enforcement Task Force*. Task force included Equal Employment Opportunity Commission, the U.S. Department of Justice, U.S. Department of Labor, and Office of Personnel Management. Available: http://www.whitehouse.gov/sites/default/files/rss_viewer/equal_pay_task_force.pdf [June 2012].

Naylor, L.A., and Rosenbloom, D.H. (2004). Adarand, Grutter, and Gratz: Does affirmative action in federal employment matter? *Review of Public Personnel Administration* 24(2):150–174.

Office of Federal Contract Compliance Programs. (2009). *Technical Assistance Guide for Federal Construction Contractors*. Available: https://www.dol.gov/ofccp/TAguides/consttag.pdf [May 2012].

Office of Federal Contract Compliance Programs. (2011). *Non-Discrimination in Compensation; Compensation Data Collection Tool, Advance Notice of Proposed Rulemaking*. 41 C.F.R.49398–49401. Available: http://www.gpo.gov/fdsys/pkg/FR-2011-08-10/html/2011-20299.htm [October 2012].

Office of Federal Contract Compliance Programs. (No date). *Guide for Small Businesses with Federal Contracts*. Available: http://www.dol.gov/ofccp/TAguides/sbguide.htm#Q4 [July 2012].

Reiter, J. (2005). Releasing multiply imputed, synthetic public use microdata: An illustration and empirical study. *Journal of the Royal Statistical Society. Series A* 168(1):185–205.

Reznek, A.P. (2006). *Recent Confidentiality Research Related to Access to Enterprise Microdata* Prepared for the Comparative Analysis of Enterprise Microdata Conference, September 18–19, 2006, U.S. Census Bureau. Available: http://www.oecd.org/data oecd/6/30/37503027.pdf [December 2011].

Robinson, C., Taylor, T., Tomaskovic-Devey, D., Zimmer, C., and Irvin, M. (2005). Studying race and sex segregation at the establishment-level: Methodological concerns and substantive opportunities in the use of EEO-1 data. *Work and Occupations* 32:5–38.

Rodgers, W., and Spriggs, W. (1996). The effect of federal contractor status on racial differences in establishment-level employment shares, 1979–1992. *American Economic Review* 86(2):290–293.

Selden, S.C. (2006). A solution in search of a problem? Discrimination, affirmative action, and the new public service. *Public Administration Review* 66(6):911–923.

Skaggs, S. (2008). Producing change or bagging opportunity? The effects of discrimination litigation on women in supermarket management. *American Journal of Sociology* 113:1,148–1,183.

Slavkovic, A.B., and Lee, J. (2010). Synthetic two-way contingency tables that preserve conditional frequencies. *Statistical Methodology* 7(3):225–239.

Smith, J.P., and Welch, F. (1984). Affirmative action and labor markets. *Journal of Labor Economics* 2(2):269–301. Published by The University of Chicago Press on behalf of the Society of Labor Economists and NORC at the University of Chicago. Available: http://www.jstor.org/stable/2534898 [July 2012].

Stainback, K., and Tomaskovic-Devey, D. (2009). Intersections of power and privilege: Long-term trends in managerial representation. *American Sociological Review* 74(5):800–820.

Stainback, K., Robinson, C., and Tomaskovic-Devey, D. (2005). Race and workplace integration: A politically mediated process? *American Behavioral Scientist* 48(9):1,200–1,228.

Stephanopoulos, G., and Edley, Jr., C. (1995). *Review of Federal Affirmative Action Programs*. Unpublished White House document, Washington, DC.

Stevens, D.W. (2002). *Unemployment That Is Not Covered by State Unemployment Insurance Laws, Longitudinal Employer-Household Dynamics*. Technical Paper No. TP-2002-16, U.S. Census Bureau. Available: http://lehd.did.census.gov/led/library/techpapers/tp-2002-16.pdf [December 2011].

U.S. Equal Employment Opportunity Commission. (2010). *EEOC Dramatically Slows Growth of Private Sector Charge Inventory*. Press Release, November 23. Available: http://www.eeoc.gov/eeoc/newsroom/release/11-23-10.cfm [May 2012].

U.S. General Accounting Office. (1993). *EEOC: An Overview.* GAO/T-HRD-93-30. Washington, DC: U.S. General Accounting Office.

U.S. General Accounting Office. (1995). *Equal Employment Opportunity: DOL Contract Compliance Reviews Could Better Target Federal Contractors.* GAO/HEHS-95-177. Washington, DC: U.S. General Accounting Office.

U.S. General Accounting Office. (2000). *Equal Employment Opportunity: Discrimination Complaint Caseloads and Underlying Causes Require EEOC's Sustained Attention.* GAO/T-GGD-00-104. Washington, DC: U.S. General Accounting Office.

U.S. General Accounting Office. (2003). *Women's Earnings: Work Patterns Partially Explain Difference between Men's and Women's Earnings.* GAO-04-35. Washington, DC: U.S. General Accounting Office.

U.S. Government Accountability Office. (2005). *Equal Employment Opportunity: The Policy Framework in the Federal Workplace and the Roles of EEOC and OPM.* GAO-05-195. Washington, DC: U.S. Government Accountability Office.

U.S. Government Accountability Office. (2006). *Financial Services Industry: Overall Trends in Management-Level Diversity and Diversity Initiatives, 1993–2004.* GAO-06-617. Washington, DC: U.S. Government Accountability Office.

U.S. Government Accountability Office. (2008). *Sharing Promising Practices and Fully Implementing Strategic Human Capital Planning Can Improve Management of Growing Workload.* GAO-08-589. Washington, DC: U.S. Government Accountability Office.

U.S. Office of Federal Contract Compliance Programs. (2011). *Non-Discrimination in Compensation; Compensation Data Collection Tool.* 41 C.F.R. Parts 60-1. *Federal Register* 76(154):49398.

U.S. Office of Management and Budget. (2002). *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies*. Available: http://www.whitehouse.gov/omb/fedreg/reproducible.html [May 2012].

WorldatWork Association. (2011). *What Is Total Rewards?* Available: http://www.worldatwork.org/waw/home/html/compensation_home.jsp [May 2012].

Zhao, Z. (2010). *Earnings Instability and Earnings Inequality in Urban China 1989-2006*. IZA Working Paper No. 3270. Available: http://ideas.repec.org/p/ess/wpaper/id2783.html [October 2012].

Appendix A

EEO Report Forms

This appendix reproduces the four equal employment opportunity (EEO) reports that collect data relevant to wages and employment, discussed in Chapter 1:

- EEO-1, required from private employers with 100 or more employees or 50 or more employees and a federal contract;
- EEO-3, required from referral unions, primarily unions with exclusive hiring arrangements with an employer;
- EEO-4, required from state and local governments; and
- EEO-5, required from primary and secondary public school districts.

Joint Reporting Committee
- Equal Employment Opportunity Commission
- Office of Federal Contract Compliance Programs (Labor)

EQUAL EMPLOYMENT OPPORTUNITY

EMPLOYER INFORMATION REPORT EEO—1

Standard Form 100
REV. 01/2006

O.M.B. No. 3046-0007
EXPIRES 01/2009
100-214

Section A—TYPE OF REPORT

Refer to instructions for number and types of reports to be filed.

1. Indicate by marking in the appropriate box the type of reporting unit for which this copy of the form is submitted (MARK ONLY ONE BOX).

(1) ☐ Single-establishment Employer Report

Multi-establishment Employer:
(2) ☐ Consolidated Report (Required)
(3) ☐ Headquarters Unit Report (Required)
(4) ☐ Individual Establishment Report (submit one for each establishment with 50 or more employees)
(5) ☐ Special Report

2. Total number of reports being filed by this Company (Answer on Consolidated Report only)_____

Section B—COMPANY IDENTIFICATION *(To be answered by all employers)*

OFFICE USE ONLY

1. Parent Company

a. Name of parent company (owns or controls establishment in item 2) omit if same as label — a.

Address (Number and street) — b.

City or town | State | ZIP code — c.

2. Establishment for which this report is filed. (Omit if same as label)

a. Name of establishment — d.

Address (Number and street) | City or Town | County | State | ZIP code — e.

b. Employer identification No. (IRS 9-DIGIT TAX NUMBER) — f.

c. Was an EEO–1 report filed for this establishment last year? ☐ Yes ☐ No

Section C—EMPLOYERS WHO ARE REQUIRED TO FILE *(To be answered by all employers)*

☐ Yes ☐ No 1. Does the entire company have at least 100 employees in the payroll period for which you are reporting?

☐ Yes ☐ No 2. Is your company affiliated through common ownership and/or centralized management with other entities in an enterprise with a total employment of 100 or more?

☐ Yes ☐ No 3. Does the company or any of its establishments (a) have 50 or more employees AND (b) is not exempt as provided by 41 CFR 60–1.5, AND either (1) is a prime government contractor or first-tier subcontactor, and has a contract, subcontract, or purchase order amounting to $50,000 or more, or (2) serves as a depository of Government funds in any amount or is a financial institution which is an issuing and paying agent for U.S. Savings Bonds and Savings Notes?

If the response to question C–3 is yes, please enter your Dun and Bradstreet identification number (if you have one): ☐☐☐☐☐☐☐☐☐

NOTE: If the answer is yes to questions 1, 2, or 3, complete the entire form, otherwise skip to Section G.

SF 100 – Page 2

Section D-EMPLOYMENT DATA

Employment at this establishment – Report all permanent full- and part-time employees including apprentices and on-the-job trainees unless specifically excluded as set forth in the instructions. Enter the appropriate figures on all lines and in all columns. Blank spaces will be considered as zeros.

Job Categories	Number of Employees (Report employees in only one category)														
	Race/Ethnicity														
	Hispanic or Latino		Not-Hispanic or Latino												Total Col A - N
			Male						Female						
	Male	Female	White	Black or African American	Native Hawaiian or Other Pacific Islander	Asian	Aamerican Indian or Alaska Native	Two or more races	White	Black or African American	Native Hawaiian or Other Pacific Islander	Asian	American Indian or Alaska Native	Two or more races	
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Executive/Senior Level Officials and Managers 1.1															
First/Mid-Level Officials and Managers 1.2															
Professionals 2															
Technicians 3															
Sales Workers 4															
Administrative Support Workers 5															
Craft Workers 6															
Operatives 7															
Laborers and Helpers 8															
Service Workers 9															
TOTAL 10															
PREVIOUS YEAR TOTAL 11															

1. Date(s) of payroll period used: _____ (Omit on the Consolidated Report.)

Section E - ESTABLISHMENT INFORMATION (Omit on the Consolidated Report.)

1. What is the major activity of this establishment? (Be specific, i.e., manufacturing steel castings, retail grocer, wholesale plumbing supplies, title insurance, etc. Include the specific type of product or type of service provided, as well as the principal business or industrial activity.)

Section F - REMARKS

Use this item to give any identification data appearing on the last EEO-1 report which differs from that given above, explain major changes in composition of reporting units and other pertinent information.

Section G - CERTIFICATION

Check one
1 ☐ All reports are accurate and were prepared in accordance with the instructions. (Check on Consolidated Report only.)
2 ☐ This report is accurate and was prepared in accordance with the instructions.

Name of Certifying Official	Title	Signature	Date
Name of person to contact regarding this report	Title	Address (Number and Street)	
City and State	Zip Code	Telephone No. (including Area Code and Extension)	Email Address

All reports and information obtained from individual reports will be kept confidential as required by Section 709(e) of Title VII.
WILLFULLY FALSE STATEMENTS ON THIS REPORT ARE PUNISHABLE BY LAW, U.S. CODE, TITLE 18, SECTION 1001

Union Reporting Program
Washington, DC 20507

EQUAL EMPLOYMENT OPPORTUNITY LOCAL UNION REPORT (EEO-3)

Approved by OMB
No. 3046-0006
Expires: 12/31/04

Part A. LOCAL UNION IDENTIFICATION

1. Full name of local union for which this report is filed. (Include local number, if any.)

2. Mailing address.

a. Where official mail should be sent to the union.

Number and street

City

County

State

Zip Code

b. Union office, if different from 2a.

Number and street

City

County

State Zip Code

3. Indicate type of local union report by a check in applicable box:

a. ☐ Report filed by local union in its own behalf
b. ☐ Other (explain)

4a. Are you affiliated with or chartered by a national or international union or national federation? Yes ☐ No ☐

b. If "Yes" to item 4a, give name and address of such national or international organization.

5. Are you affiliated with the AFL-CIO? Yes ☐ No ☐

Part B. LOCAL UNION PRACTICES

1. To the best of your knowledge, does your membership include any:

	Yes	No
a. Blacks (Non-Hispanic)?	☐	☐
b. Hispanics?	☐	☐
c. Women?	☐	☐

3. To the best of your knowledge, has your international union chartered a separate local within the same work and/or area jurisdiction which consists only of:

	Yes	No
a. Persons of the same race/ethnic identity	☐	☐
b. Persons of the same sex?	☐	☐

2. If "No" to any items 1a, 1b, or 1c, is this because the group or groups not represented:

	(CHECK ALL APPLICABLE BOXES) BLACK NON-HISPANIC 1 (a)	HISPANIC 1 (b)	WOMEN 1 (c)
a. Are not in the local community?			
b. Are not in the bargaining unit?			
c. Are excluded by provision in constitution or bylaws?			
d. Have not applied for membership?			
e. Have applied, but did not have a sponsor?			
f. Have applied, but did not meet qualifications other than sponsorship?			

g. Other reason(s) (Explain)

Part C. LOCAL UNIONS REQUIRED TO FILE

	Yes	No
1. Has the local union had 100 or more members at any time since December 31 of the preceding year?	☐	☐
2. Does the local union, or any unit, division, or agent of the local union, or any labor organization which performs, within a specific jurisdiction, the functions ordinarily performed by a local union, whether or not it is so designated:		
a. Operate a hiring hall or hiring office?	☐	☐
b. Have an arrangement under which one or more employers are required to consider or hire persons referred by the local union or an agent of the local union?	☐	☐
c. Have 10 percent or more of its members employed by employers which customarily and regularly look to the union, or any agent of the union, for employees to be hired on a casual or temporary basis, for a specified period of time, or for the duration of a specified job?	☐	☐

The union must complete the entire report if it answered "YES" to item 1, *AND* the answer is "YES" to any of the three questions in item 2.

The union is not required to complete the entire report if it answered "NO" to item 1, *OR* "NO" to all three questions in item 2. If that is the case, the union must complete Parts A, B, C and E and return this form to the specified address.

Part D.. REMARKS

Part E. IDENTIFICATION AND SIGNATURE

To the best of my knowledge and belief, the information contained in this report is true and complete. It is further certified that to the extent any data in Schedule I, Items 1 or 2, are based on self-identification by individuals, this information was gathered only after they were advised of its confidential nature and purposes.

1. Type or print name, title, address and telephone number for union business of designated representative

Name _____

Title _____

Work address _____

Telephone number (including area code) _____

2. Signature of designated representative | 3. Date

"Whoever, in any matter within the jurisdiction of any department or agency of the United States knowingly and willfully falsifies, conceals, or covers up by any trick, scheme, or device a material fact, or makes any false, fictitious or fraudulent statements or representations, or makes or uses any false writing or document knowing the same to contain any false, fictitious or fraudulent statement or entry, shall be fined not more than $10,000 or imprisoned not more than 5 years, or both." Title 18, Section 1001, United States Code.

SCHEDULE I—LOCAL UNION REPORT (EEO—3)

MEMBERSHIP, APPLICANT and REFERRAL INFORMATION

1. Method of identification

How was information as to race/ethnic identification and sex in Item 2 below obtained?

This information may be obtained by visual survey, from records made after employment, from personal knowledge or by self-identification. The self-identification method may be used subject to the conditions set forth in the instructions. No State law prohibiting the self-identification method applies, since the Equal Employment Opportunity Commission's regulations supersede such laws.

	Check all applicable boxes
a. Existing Record	
b. Visual Survey	
c. Tally from Personal Knowledge	
d. Self-Identification	
e. Other (Specify)	

2. Statistics

	TOTAL (COLUMNS B-K)	MALE					FEMALE				
		NON-HISPANIC ORIGIN		HISPANIC	ASIAN OR PACIFIC ISLANDER	AMERICAN INDIAN OR ALASKAN NATIVE	NON-HISPANIC ORIGIN		HISPANIC	ASIAN OR PACIFIC ISLANDER	AMERICAN INDIAN OR ALASKAN NATIVE
		WHITE	BLACK				WHITE	BLACK			
	A	B	C	D	E	F	G	H	I	J	K
a. MEMBERSHIP IN REFERRAL UNIT											
(1) MEMBERS											
(2) APPLICANTS FOR MEMBERSHIP DURING THE PAST YEAR											
b. REFERRALS DURING 2-MONTH PERIOD											
(1) NUMBER OF PERSONS REFERRED											
(2) NUMBER OF REFERRALS											
(3) APPLICANTS FOR REFERRAL											

3. Period Used For Referral Date

You should obtain the figures reported in Item 2 "Statistics" using any 2-month period between August 1 and November 30.

Dates of 2-month Period _____

EEOC FORM 274, JUN 88 PREVIOUS EDITIONS ARE OBSOLETE

EEOC ORIGINAL

Page 2

EQUAL EMPLOYMENT OPPORTUNITY COMMISSION

STATE AND LOCAL GOVERNMENT INFORMATION (EEO-4)

EXCLUDE SCHOOL SYSTEMS AND EDUCATIONAL INSTITUTIONS
(Read attached instructions prior to completing this form)

APPROVED BY OMB 3046-0008
EXPIRES 12/31/2005

DO NOT ALTER INFORMATION PRINTED IN THIS BOX

MAIL COMPLETED FORM TO:
EEO-4 Reporting Center
PO Box 1898
Chicago, IL 60690-1898

A. TYPE OF GOVERNMENT (Check one box only)

☐ 1. State ☐ 2. County ☐ 3. City ☐ 4. Township ☐ 5. Special District

☐ 6. Other (Specify)_____

B. IDENTIFICATION

1. NAME OF POLITICAL JURISDICTION (If same as label, skip to Item C)

2. Address--Number and Street	CITY/TOWN	COUNTY	STATE/ZIP	EEOC USE ONLY A B

C. FUNCTION

(Check one box to indicate the function(s) for which this form is being submitted. Data should be reported for all departments and agencies in your government covered by the function(s) indicated. If you cannot supply the data for every agency within the function(s) attach a list showing name and address of agencies whose data are not included.)

	1.Financial Administration. Tax billing and collection, budgeting, purchasing, central accounting and similar financial administration carried on by a treasurer's, auditor's or comptroller's office and GENERAL CONTROL. Duties usually performed by boards of supervisors or commissioners, central administration offices and agencies, central personnel or planning agencies, all judicial offices and employees (judges, magistrates, bailiffs, etc.)		8. HEALTH. Provision of public health services, out-patient clinics, visiting nurses, food and sanitary inspections, mental health, alcohol rehabilitation service, etc. 9. HOUSING. Code enforcement, low rent public housing, fair housing ordinance enforcement, housing for elderly, housing rehabilitation, rent control.
	2. STREETS AND HIGHWAYS. Maintenance, repair, construction and administration of streets, alleys, sidewalks, roads, highways and bridges.		10. COMMUNITY DEVELOPMENT. Planning, zoning, land development, open space, beautification, preservation.
	3. PUBLIC WELFARE. Maintenance of homes and other institutions for the needy; administration of public assistance. (Hospitals and sanatoriums should be reported as item7.)		11. CORRECTIONS. Jails, reformatories, detention homes, half-way houses, prisons, parole and probation activities
	4. POLICE PROTECTION. Duties of a police department sheriff's, constable's, coroner's office, etc., including technical and clerical employees engaged in police activities.		12. UTILITIES AND TRANSPORTATION. Includes water supply, electric power, transit, gas, airports, water transportation and terminals.
	5. FIRE PROTECTION. Duties of the uniformed fire force and clerical employees. (Report any forest fire protection activities as item 6.)		13. SANITATION AND SEWAGE. Street cleaning, garbage and refuse collection and disposal. Provision, maintenance and operation of sanitary and storm sewer systems and sewage disposal plants.
	6. NATURAL RESOURCES. Agriculture, forestry, forest fire protection, irrigation drainage, flood control, etc., and PARKS AND RECREATION. Provision, maintenance and operation of parks, playgrounds, swimming pools, auditoriums, museums, marinas, zoos, etc.		14. EMPLOYMENT SECURITY STATE GOVERNMENTS ONLY
	7. HOSPITALS AND SANATORIUMS. Operation and maintenance of institutions for inpatient medical care.		15. OTHER (Specify on Page Four)

EEOC FORM 164 FEB 97 (Previous Editions are Obsolete) PAGE 1

D. EMPLOYMENT DATA AS OF JUNE 30

(Do not include elected/appointed officials. Blanks will be counted as zero)

1. FULL-TIME EMPLOYEES (Temporary employees are not included)

JOB CATEGORIES	ANNUAL SALARY (In thousands 000)	MALE						FEMALE				
		TOTAL (COLUMNS B-K)	NON-HISPANIC ORIGIN		HISPANIC	ASIAN OR PACIFIC ISLANDER	AMERICAN INDIAN OR ALASKAN NATIVE	NON-HISPANIC ORIGIN		HISPANIC	ASIAN OR PACIFIC ISLANDER	AMERICAN INDIAN OR ALASKAN NATIVE
			WHITE	Black				White	Black			
		A	B	C	D	E	F	G	H	I	J	K
OFFICIALS ADMINISTRATORS	1. $0.1-15.9											
	2. 16.0-19.9											
	3. 20.0-24.9											
	4. 25.0-32.9											
	5. 33.0-42.9											
	6. 43.0-54.9											
	7. 55.0-69.9											
	8. 70.0 PLUS											
PROFESSIONALS	9. $0.1-15.9											
	10. 16.0-19.9											
	11. 20.0-24.9											
	12. 25.0-32.9											
	13. 33.0-42.9											
	14. 43.0-54.9											
	15. 55.0-69.9											
	16. 70.0 PLUS											
TECHNICIANS	17. $0.1-15.9											
	18. 16.0-19.9											
	19. 20.0-24.9											
	20. 25.0-32.9											
	21. 33.0-42.9											
	22. 43.0-54.9											
	23. 55.0-69.9											
	24. 70.0 PLUS											
PROTECTIVE SERVICE	25. $0.1-15.9											
	26. 16.0-19.9											
	27. 20.0-24.9											
	28. 25.0-32.9											
	29. 33.0-42.9											
	30. 43.0-54.9											
	31. 55.0-69.9											
	32. 70.0 PLUS											
PARA- PROFESSIONALS	33. $0.1-15.9											
	34. 16.0-19.9											
	35. 20.0-24.9											
	36. 25.0-32.9											
	37. 33.0-42.9											
	38. 43.0-54.9											
	39. 55.0-69.9											
	40. 70.0 PLUS											
ADMINISTRATIVE SUPPORT	41. $0.1-15.9											
	42. 16.0-19.9											
	43. 20.0-24.9											
	44. 25.0-32.9											
	45. 33.0-42.9											
	46. 43.0-54.9											
	47. 55.0-69.9											
	48. 70.0 PLUS											

EEOC FORM 164, FEB 97 (Previous Editions are Obsolete) PAGE 2

D. EMPLOYMENT DATA AS OF JUNE 30 (Cont.)

(Do not include elected/appointed officials. Blanks will be counted as zero)

1. FULL-TIME EMPLOYEES (Temporary employees are not included)

JOB CATEGORIES	ANNUAL SALARY (In thousands 000)	MALE						FEMALE				
		TOTAL (COLUMNS B-K)	NON-HISPANIC ORIGIN		HISPANIC	ASIAN OR PACIFIC ISLANDER	AMERICAN INDIAN OR ALASKAN NATIVE	NON-HISPANIC ORIGIN		HISPANIC	ASIAN OR PACIFIC ISLANDER	AMERICAN INDIAN OR ALASKAN NATIVE
			WHITE	Black				White	Black			
		A	B	C	D	E	F	G	H	I	J	K
SKILLEDCRAFT	49. $0.1-15.9											
	50. 16.0-19.9											
	51. 20.0-24.9											
	52. 25.0-32.9											
	53. 33.0-42.9											
	54. 43.0-54.9											
	55. 55.0-69.9											
	56. 70.0 PLUS											
SERVICE MAINTENANCE	57. $0.1-15.9											
	58. 16.0-19.9											
	59. 20.0-24.9											
	60. 25.0-32.9											
	61. 33.0-42.9											
	62. 43.0-54.9											
	63. 55.0-69.9											
	64. 70.0 PLUS											
65. TOTAL FULL TIME (LINES 1 – 64)												
2. OTHER THAN FULL-TIME EMPLOYEES (Including temporary employees												
66. OFFICIALS/ADMIN												
67. PROFESSIONALS												
68. TECHNICIANS												
69. PROTECTIVE SERVICE												
70. PARA-PROFESSIONAL												
71. ADMIN. SUPPORT												
72. SKILLED CRAFT												
73. SERVICE/MAINTENANCE												
74. TOTAL OTHER THAN FULL TIME (LINES 66 – 73)												
3. NEW HIRES DURING FISCAL YEAR - Permanent full time only JULY 1 – JUNE 30												
75. OFFICIALS/ADMIN												
76. PROFESSIONALS												
77. TECHNICIANS												
78. PROTECTIVE SERVICE												
79. PARA-PROFESSIONAL												
80. ADMIN. SUPPORT												
81. SKILLED CRAFT												
82. SERVICE/MAINTENANCE												
83. TOTAL NEW HIRES (LINES 75 – 82)												

EEOC FORM 164, FEB 97 (Previous Editions are Obsolete) PAGE 3

REMARKS (List National Crime Information Center (NCIC) number assigned to any Criminal Justice Agencies whose data are included in this report)
LIST AGENCIES INCLUDED ON THIS FORM
CERTIFICATION. I certify that the information given in this report is correct and true to the best of my knowledge and was reported in accordance with accompanying instructions. (Willfully false statements on this report are punishable by law, US Code, Title 18, Section 1001.)

NAME OF PERSON TO CONTACT REGARDING THIS FORM	TITLE
ADDRESS (Number and Street, City, State, Zip Code)	TELEPHONE NUMBER extension: FAX NUMBER

DATE E-MAIL	TYPED NAME/TITLE OF AUTHORIZED OFFICIAL	SIGNATURE

EEOC FORM 164, FEB 97 (Previous Editions Obsolete) PAGE 4

EQUAL EMPLOYMENT OPPORTUNITY COMMISSION ELEMENTARY-SECONDARY STAFF INFORMATION (EEO-5) Public school systems	FORM APPROVED BY OMB NO. 3046-0003 APPROVAL EXPIRES This is a joint requirement of the EEOC and the Office for Civil Rights, U.S. Department of Education and the U.S. Department of Justice.

DO NOT ALTER INFORMATION PRINTED IN THIS BOX.

NOTE: ALL EMPLOYEES IN YOUR SCHOOL DISTRICT MUST BE INCLUDED ON THIS FORM. Additional Copies of this form may be obtained from the address below. Send your full report to:

PART I. IDENTIFICATION

A. TYPE OF AGENCY WHICH OPERATES THE REPORTING SCHOOL SYSTEM

☐ Local Public School ☐ Special Regional Agency ☐ State Education Agency

☐ Other (Specify) _____

B. SCHOOLS SYSTEMS IDENTIFICATION (OMIT IS SAME AS LABEL)

NAME

STREET AND NO. OR POST OFFICE BOX	CITY/TOWN	COUNTY	STATE	ZIP

C. GENERAL STATISTICS

NUMBER OF SCHOOLS OPERATED	NUMBER OF ANNEXES OPERATED	OCTOBER 1ST ENROLLMENT

D. REMARKS

EEOC FORM 168A Page 1

PART II. Staff Statistics of (DATE) _____ DO NOT INCLUDE ELECTED/APPOINTED OFFICIALS (SEE DEFINITION IN APPENDIX)

DISTRICT NAME: _____

Activity Assignment Classification	**A. FULL-TIME STAFF**														
	Race/Ethnicity														
	Hispanic or Latino		Non -Hispanic or Latino												Total Col A - N
			Male						Female						
	Male	Female	White	Black or African American	Asian	Native Hawaiian or Other Pacific Islander	American Indian or Alaska Native	Two or more races	White	Black or African American	Asian	Native Hawaiian or Other Pacific Islander	American Indian or Alaska Native	Two or more races	
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1. Officials, Administrators, Managers															
2. Principals															
3. Assistant Principals, Teaching															
4. Assistant Principals, Non-teaching															
5. Elementary Classroom Teachers															
6. Secondary Classroom Teachers															
7. Other Classroom Teachers															
8. Guidance															
9. Psychological															
10. Librarians/ Audiovisual Staff															
11. Consultants & Supervisors of Instruction															
12. Other Professional Staff															
13. Teachers Aides															
14. Technicians															
15. Administrative Support Workers															
16. Service Workers															
17. Skilled Crafts															
18. Laborers and Helpers															
19. TOTALS (1-18)															

Activity Assignment / Classification	B. PART-TIME STAFF															
	Race/Ethnicity															
	Hispanic or Latino		Non -Hispanic or Latino													
			Male						Female							
	Male	Female	White	Black or African American	Asian	Native Hawaiian or Other Pacific Islander	American Indian or Alaska Native	Two or more races	White	Black or African American	Asian	Native Hawaiian or Other Pacific Islander	American Indian or Alaska Native	Two or more races		
20. Professional Instructional																
21. All Other																
22. TOTALS (20-21)																
C. NEW HIRES FULL-TIME (JULY THRU SEPT. OF THE SURVEY YEAR)																
23. Officials, Administrators, Managers																
24. Principals/Assistant Principals																
25. Classroom Teachers																
26. Other Professional Staff																
27. Nonprofessional Staff																
28. TOTALS (23-27)																

CERTIFICATION: I certify that the information given in this report is correct and true to the best of my knowledge and was prepared in accordance with accompanying instructions. Willfully false statements on this report are punishable by law, U.S. Code, Title 18, Section 1001.

Date	Phone: Fax: Email:	Typed Name/Title of Person Responsible for Report	Signature

EEOC FORM 168A PREVIOUS EDITIONS ARE OBSOLETE

Appendix B

Study of Employment Earnings for the Equal Employment Opportunity Program: A Possible Role for Administrative Data from Three Tax Systems

Nicholas Greenia

INTRODUCTION

The proposed Paycheck Fairness Act of 2009 (H.R. 12 in the 112th Congress) would have required the U.S. Equal Employment Opportunity Commission (EEOC) to issue regulations mandating the provision of earnings data from employers to the EEOC classified by the race, gender, and national origin of their employees. According to the proposed legislation, these pay or earnings data are needed to bolster the related employment and other data already collected through the equal employment opportunity (EEO) reports, particularly the EEO-1 reports, for purposes of enforcing compliance with statutory nondiscrimination employment practices. The new data were argued to be critical in continuing to administer Section 709 of the Civil Rights Act of 1964, as amended.

This paper explores the feasibility of using existing data from the administrative records of three tax systems for accomplishing the EEO-1 stated goals for new data collection. It discusses the data collected from and the interrelationships among three tax systems: two administered by federal agencies, the Internal Revenue Service (IRS) and the Social Security Administration (SSA), and one by the state agencies, the unemployment insurance (UI) offices that operate as federal-state partnerships under the Employment and Training Administration (ETA) of the U.S. Department of Labor. It continues by discussing how the interrelationships of the three tax systems benefit data quality, including timeliness, for EEO-1 purposes. It also provides an overview of the sources, including the forms that could

provide the needed data. The paper concludes by presenting major concerns on confidentiality.

These systems hold particular promise for a number of reasons. One is the coverage of the taxes reported and collected: federal income taxes for funding many federal programs that benefit all U.S. residents, taxes that help fund the Social Security and Medicare programs for retirees and other qualified recipients, and unemployment insurance taxes that fund the unemployment benefits of workers who are laid off during difficult economic times, particularly for extended periods such as during the recent deep recession. Another is data quality: the data records tend to have, in general, high levels of compliance because of the importance of these programs—highlighted by the penalties for noncompliance—for the nation's safety net and in funding congressionally mandated expenditures. A third is the potential for triangulation of firm and worker levels of reporting by the use of all three systems. Although there are some issues with response rates in each system, such as the tax gap for federal income taxes, partial participation is likely to result in detection by one of the three systems.

Although each administrative record data set holds promise for supplementing EEO-1 data, there are also challenges associated with the use of these administrative data. Like any data system, these three administrative record systems are imperfect in terms of response rates, accuracy, and all levels of granularity, such as multiemployer member reporting in the UI system. In addition, each also has constraints, including purposes and access.

How the EEOC decides to approach the enhancement of its data, including any redesign of its own EEO-1 collection system, may be key to determining not only the most useful plan, but also the most viable for purposes of obtaining earnings data classified by gender, race/ethnicity, and nativity.

A BRIEF OVERVIEW OF THE THREE TAX SYSTEMS

This section presents an overview of the purpose, coverage, data availability, national importance, and interrelationships of the three systems. These administrative earnings data are captured by multiple administrative forms, reported in various components, and available across multiple years from the three tax systems. The classifier variables for gender, race/ethnicity, nativity, and even age, also exist at the employee record level although they are not universally captured in the databases. All of these data could be linked to a specific employer for an employee, including for multiple employers.

Purpose

The three data systems are used primarily to collect taxes for administering and funding vital mandated programs: the federal income tax system by IRS, the Social Security and Medicare programs by SSA, and the state UI systems, which are operated by state Employment Security Agencies (ESAs) under a federal-state partnership. Related national statistics are produced from all three sets of data by the statistical offices of SSA and IRS, as well as the Department of Labor's Bureau of Labor Statistics and the U.S. Census Bureau. In addition, they are used for policy analysis in a wide range of offices, including the Joint (Senate-House) Committee on Taxation, the Congressional Budget Office, and the Office of Tax Analysis at the U.S. Department of the Treasury, and for analytical research by top academics through the Intergovernmental Personnel Act as well as the Census Bureau's Research Data Centers.

Such robust—and visible—uses of the data have beneficial consequences for the EEO-1 Program because weaknesses, limitations, and inaccuracies in the data systems tend to become known and corrective measures taken in order to ensure the utility and consistency of the data over time. In addition, because the U.S. statistical system is decentralized, it is more difficult for any one system's data anomalies to go unnoticed, given the cross-checks implicitly or explicitly built in across these quasi independent systems—particularly for financial data, including employment earnings.

Coverage

Across the three systems, as well as the U.S. Census Bureau, establishments and workers needed for EEO-1 purposes would be covered. The data are reported on IRS income tax returns (for individuals and businesses), employment tax returns (for both the Federal Income Contributions Act [FICA] and the Federal Unemployment Tax Act [FUTA]), information returns (including for tax-exempt nonprofit organizations), applications for Social Security Numbers (SSNs), and on UI-related forms. Several federal agencies play major roles in either funding or helping process the data and payments for these programs: the Department of Labor (DOL)—particularly its Bureau of Labor Statistics (BLS) and ETA—SSA and IRS. In addition, the states play a major role in administering the State Unemployment Tax Authority (SUTA) Program as well as the employment and training administration system funded in large part by ETA.

Data Availability

Table B-1 summarizes availability of the EEO-1 items needed by source, including the Census Bureau.

National Importance

The data are critical for funding many federal programs that benefit all U.S. residents, taxes that help fund the Social Security and Medicare programs for retirees and other qualified recipients, and unemployment insurance taxes that fund the unemployment benefits of workers who are laid off during difficult economic times.

Inter-Relationships

The three sets of data are interrelated, albeit sometimes in subtle ways. For example, all three systems depend on the SSNs assigned by SSA, the employer identification numbers (EINs) assigned by IRS, the reporting of employment and payroll at both the firm and individual worker level for federal and state purposes, and related information to update them, such as changes in name or address. Similarly, the IRS determination of which workers are employees and which are contractors has an impact on the other systems. The IRS decision is obtained by the filing of a Form SS-8 for a firm or worker seeking to have IRS establish officially the employee or independent contractor status of a particular worker. This transaction then has ramifications for the other employee data collection systems, such as SUTA and FUTA, and could also be used to inform and supplement the EEO-1 reports.

EEO-1 Utility

Because of the coverage, availability, and interrelationships, the three tax systems hold considerable promise for providing the employee earnings data needed by gender, race/ethnicity, nativity, and even age, by employer. In addition, these systems could be useful also because of the other data they contain, in addition to employee earnings, for supplementing the EEO-1 report data currently collected, including across time both retrospectively and prospectively.

STATE UNEMPLOYMENT INSURANCE DATA

This section presents a brief summary of why and how UI and Quarterly Census of Employment and Wages (QCEW) data are reported, collected,

TABLE B-1 Summary of Available Population/Universe Data by Agency

Source	Earnings at Employee Level	Earnings at Employer Level	Employee Gender	Employee Race/Eth.	Employee Nativity
State UI[a]	YES	YES	NO	NO	NO
State ESA[b]	NO	YES	NO	NO	NO
IRS	YES	YES	NO	NO	YES[c]
SSA	YES	YES	YES	YES	YES
Census[d]	YES	YES	YES	YES	YES

[a]Based on household surveys (samples), BLS does publish employment and compensation data by gender, race/ethnicity, and nativity. These could be useful, at a minimum, as benchmark estimates for purposes of EEO-1 expansion. However, BLS does not presently have access to the states' detailed employee earnings records.

[b]ESA = Employment Security Agency.

[c]Only from applications for individual taxpayer identification numbers (ITINs).

[d]Earnings data exist annually in extracts of tax data from IRS. The decennial census captures gender, race, and ethnicity, but, only every 10 years. However, the Census Bureau's American Community Survey (ACS) captures gender, race/ethnicity, and nativity (native/foreign born). Although it is a sample, the ACS is sufficiently large—approximately 2 million respondent households annually—that, at a minimum, its data could be useful for benchmark estimates and perhaps for matching to the EEO-1 report population on a statistically significant basis.

and shared with the federal sector, and the significance for the EEO-1 Program.

Purpose

In addition to complying with FUTA, employers must also comply with SUTA by withholding and depositing tax or insurance payments from each employee's wages with the state unemployment offices. Although federal unemployment taxes serve several purposes, state unemployment taxes are used only to fund unemployment benefits in a particular state or territory (including the District of Columbia, Puerto Rico, and the Virgin Islands).

Coverage

Tax rates and coverage vary by state, as do the content and format of the records a particular state collects. In general, workers not covered by this system include federal employees, independent contractors, the self-employed, and some agricultural workers.

Content

A state collects the employment and compensation data in two parts. The first part is detailed earnings data[1] collected as part of the UI system. The state UI agency collects reports from each employer that include the SSN, name, and quarterly compensation for each individual employee (as well as the employer name and EIN).[2] This collection of detailed employee earnings, often called UI wage records, provides the most frequent and granular information about employee earnings across the three tax systems.

For the second part, the state ESA collects aggregate monthly employment (for the pay period containing the 12th of the month[3]) for each quarter and the aggregate quarterly employee compensation from each employer in the state covered by state UI laws and federal workers covered by the Unemployment Compensation for Federal Employees (UCFE) Program.[4] This program, administered by the BLS, also includes the collection of monthly

[1]See, for example, http://detr.state.nv.us/uicont/forms/NUCS-4072.PDF [July 2012].

[2]The coverage varies by state. For a complete review, see Stevens (2002), available at: http://lehd.did.census.gov/led/library/techpapers/tp-2007-04.pdf [July 2012].

[3]The 12th of the month is the same date used for reporting of employment on the IRS quarterly employment FICA tax returns (Form 941 series) that is, March 12, June 12, etc.

[4]This quarterly reporting of aggregate compensation provides more commonality with the IRS Form 941 series, which also reports quarterly aggregate employee compensation: see, for example, http://www.bls.gov/cew/forms/mwr_nm.pdf [July 2012], also see http://www.bls.gov/cew/cewover.htm [July 2012].

employment data and provides the most frequent aggregate employment data across the three tax systems.

The second part data collection is partly funded by BLS, and after a state edits the data, it transmits electronic summaries to BLS for its statistical needs. Although data are also requested for multiple worksite or multi-establishment employers, there is no disincentive for an employer that does not comply with the multisite request as long as total employment is reported accurately and the appropriate amount of UI taxes is paid to the states.

EEO-1 Utility

For purposes of expanding the EEO-1 Program, the UI data system provides the earnings data needed and at the employee level, but it also presents three problems. First, because of the lack of a disincentive for nonreporting of multisite employer detail, there may be a disconnect in matching to multi-establishment employer data at the worksite level—but not the enterprise level—from the EEO-1 reports. It would be up to the EEOC to determine how big a problem this represents for its enforcement needs. Second, gender, race/ethnicity, and nativity data are not collected for either of the two parts described above. However, if the detailed employee earnings data could be matched to SSA Numident (Numerical Identification System) data, this problem could be reduced if not resolved. Third, and perhaps most daunting, in order to obtain either of the two data parts provided to the states—especially the detailed employee earnings—it would be necessary to obtain separate agreements with each state as was done so laboriously for the Longitudinal Employer-Household Dynamics (LEHD) Program at the Census Bureau starting in the 1990s.[5]

INTERNAL REVENUE SERVICE DATA

This section presents a summary of several tax and information forms, especially Form W-2, Form 941, and Form 940, and why they might be of interest to expand the EEO-1 reports on employment and earnings data. In addition, it discusses the close relationship IRS and SSA have in terms of the first two forms, particularly for validating and reconciling amounts withheld for income, Social Security, and Medicare taxes.

[5]The LEHD program is briefly described in Chapter 2.

Purpose

In 1976,[6] the current simplified Combined Annual Wage Reporting (CAWR) Program was established by law to ensure that employers pay and report the correct amount of tax, including federal income tax withholding and that they file timely all necessary forms with SSA. That same year, Form W-2 (Wage and Tax Statement) was redesigned to include Social Security information, and Form W-3 (Transmittal of Income and Tax Statements), was amended to include cumulative totals of each money field appearing on the associated Form W-2s.

Content

Detailed annual employee compensation, quarterly and annual aggregate employee compensation, and number of employees are provided at both the employee and employer level and are linkable by the SSN/EIN crosswalk also provided. In addition, other tax forms provide various components of aggregate and even detailed employee compensation, such as compensation to corporate officers. Finally, EIN and ITIN assignment and other transactions enable the tracking of new business births, foreign born workers without SSNs, and even the employee or contractor status of a worker.

For purposes of expanding EEO-1 reports, three forms in particular figure prominently in the CAWR process: Form W-2, Form 941, and Form 940.[7]

Form 940, Employer's Annual Federal Unemployment Tax Act (FUTA) Tax Return[8]

Purpose

Form 940 is required to be filed annually by an employer for purposes of reporting and paying the federal unemployment taxes required by FUTA. These taxes are used to fund state workforce agencies, pay half the cost of extended unemployment benefits in severe economic downturns, and also for loans to states to help them pay unemployment benefits, including extended unemployment benefits.

[6]The Tax Reform Act of 1976 (TRA76) also established the present confidentiality statute in the tax code, namely, Section 6103.

[7]Schedule H, filed with Form 1040 to report household employees, is omitted from this discussion.

[8]See http://www.irs.gov/pub/irs-pdf/f940.pdf [July 2012].

Coverage

Filing is required—at the aggregate employment level—for each nonagricultural employee earning at least $1,500 in any quarter of the year or for each employee who was employed for part/all of a day in any 20 different weeks of the year.[9]

EEO-1 Utility

Although Form 940 does report annual total compensation, it does not report the number of employees. However, for purposes of this analysis, the compensation information may be useful for benchmarking compensation data reported on other federal tax forms, say, Form W-2, and Form 941, as well as the UI data.

Form W-2, Wage and Tax Statement[10]

Purpose

Form W-2 is required to be filed by both employees, with their individual tax returns (Form 1040), and employers, transmitted under the summary Form W-3. The form's major tax purpose is threefold: reporting of federal income tax, Social Security tax, and Medicare tax withheld from employees' compensation. The W-2 is also required to be filed if these taxes were not withheld but should have been.

Coverage

Withholding of federal income tax is not required for an employee who had no federal tax liability in the previous year and is expected to have none in the current year. However, because Social Security and Medicare taxes must be withheld, a Form W-2 must be filed for such an employee. Thus, this is an extremely potent building block for employment and wage data—at the employee level, but cross-referenced to the employer level by the cross-walk of SSN/EIN—even for low-wage employees. In addition, because a different W-2 must be filed by each employer of an employee, these data can provide multiple employer information for an employee with multiple jobs.

[9]For 2009 and 2010, agricultural employers were required to file if they paid cash wages of $20,000 or more to farm workers during any calendar quarter or if they employed 10 or more farm workers during some part of the day (whether or not at the same time) during any 20 or more different weeks in either year.

[10]See http://www.irs.gov/pub/irs-pdf/fw2.pdf [July 2012].

EEO-1 Utility

The industry codes available at SSA (at the full 6-digit level of the North American Industry Classification System) can provide a further source of rich classifier information on employers' business activities. Earnings detail is also rich: wages and salaries, deferred compensation (part of total compensation, even if not taxable currently), and certain fringe benefits are reported, in addition to capped Social Security earnings and uncapped Medicare earnings. Together, the W-2 earnings variables provide a unique and comprehensive window on earnings data at the employee level.

Form 941, Employer's Quarterly Tax Return[11]

Purpose

Form 941 is required to be filed quarterly by an employer in order to report and pay federal income tax withheld for employees, and both the employer's and employees' share of Social Security and Medicare Taxes. Similarly, Form 943, Employer's Annual Federal Tax Return for Agricultural Employees,[12] is required to be filed annually for the same reasons, but for agricultural employees.

Coverage

In general, coverage of FICA employees by the Form 941 series is very similar to that of FUTA employees by the Form 940.

EEO-1 Utility

Both the Form 941 series and Form 943 contain a number of useful fields, especially the total number of employees and their total compensation—quarterly for the Form 941, annually for Form 943. In addition, the forms report taxable Social Security wages (which are capped at the SSA ceiling) and taxable Medicare wages (which are not capped and, thus, equivalent to total wages).

Data Quality: IRS and SSA Reconciliation

IRS and SSA use a reconciliation process involving the filings of both Form W-2 and Form 941 in order to determine discrepancies and possible

[11]See http://www.irs.gov/pub/irs-pdf/f941.pdf [July 2012].
[12]See http://www.irs.gov/pub/irs-pdf/f943.pdf [July 2012].

tax delinquencies. Specifically, they compare taxable SSA wages, taxable SSA tips, taxable Medicare wages, and federal income tax withheld. Discrepancies result in the direct contact of employers, and consequences for noncompliance—and even nonresponse—can be serious. For example, in addition to monetary penalties that may result, so-called bad boy employers have been required to file Form 941 on a monthly, instead of a quarterly, basis.

IRS uses a similar cross-check system involving more tax forms, such as the Form 1040[13] series of individual tax returns, to ensure that an individual's total reported income jibes with other reports of the income source; for example, the Form W-2 for earnings and other compensation and Form 1099R for income such as interest, dividends, and pension distributions.

The consequences of being noncompliant with the federal income tax system are well known and potentially include not only prison, monetary penalties and interest, but also damage to one's credit ratings for both individuals and firms. For a firm, such damage can extend to its reputation in the business community, for example, for partnering and other collaborative efforts, and adversely affect attempts to raise capital publicly, say, with an initial public offering, and privately.

Because of the adverse consequences of tax noncompliance, firms are generally highly incentivized to comply and provide accurate and timely information to both IRS and SSA. If they are not, IRS enlists an array of tools for enforcing compliance that include a discriminant function (DIF) system scoring[14] of individual and some business tax returns and numerous auditors and agents to ensure that tax laws are obeyed and corrective measures taken when they are not.

Data Quality: IRS and State UI Reconciliation

Although a similar relationship exists between IRS and the state workforce agencies[15] for purposes of ensuring the timely and accurate payment of both state and federal unemployment taxes, Form 940 earnings data—annual employment compensation by employer—may be less useful for

[13]See http://www.irs.gov/pub/irs-pdf/f1040.pdf [July 2012].

[14]Under this system, IRS computer programs assign each return a numeric DIF score rating the potential for necessary changes to the return, based on past IRS experience with similar returns. The unreported income DIF score is used to rate the return for the potential of unreported income. IRS staff screen the highest-scoring returns, selecting some for audit and identifying the items on these returns that are most likely to need review.

[15]Under Section 6103 of the tax code (and reciprocating state and municipal laws), IRS, state, and even municipal tax authorities have long shared data for mutual benefit involving tax administration. For states, such sharing has included data to administer both income taxes and employment or payroll taxes.

purposes of expanding EEO-1 reports than the more detailed information on the Form W-2 and Form 941. However, the information sharing between IRS and the state workforce agencies also helps ensure the accuracy of the data reported to the states at both the firm and employee level for purposes of both federal and state unemployment taxes. The importance of the interagency relationship for ensuring that these taxes are paid correctly and timely is a major reason these data from all three tax systems may hold such promise for expanding the earnings data on the EEO-1 reports.

Form SS-4, Application for Employer Identification Number[16]

In addition to starting the process for assigning an EIN for an entity (usually, but not always, a business), the Form SS-4 establishes an employer's account on the IRS Business Master File (similar to the business registers at BLS and the Census Bureau, but for tax administration), including filing requirements for income tax returns (Form 1120 series, Form 1065 series, Form 990 series) and employment tax returns (Form 940 and Form 941 series). It also provides the SSN-EIN crosswalk for a sole proprietorship converting from nonemployer to employer status, important information in order to link the Schedule C posting to the Individual Master File on SSN with the accompanying Form 1040, to the sole proprietorship's employment tax returns posting on EIN to the Business Master File. IRS also provides SS-4 population data to SSA (and the Census Bureau), which uses the detailed alpha information on business activity to assign full 6-digit industry codes,[17] which should be useful industry classification for EEO-1 reports. In summary, this short form initiates actions in several systems—both statistical and administrative—which begin the cross-tracking of many events for a central use of the form, the identification of new businesses.

Form W-4, Employee's Withholding Allowance Certificate[18]

Form W-4 identifies a new employee's withholding status for purposes of the required Form W-2 that is later filed with an employee's Form 1040 individual income tax return. Although the W-4 is not required to be filed with IRS, it is required to be filed by federal and state agencies for

[16]See http://www.irs.gov/pub/irs-pdf/fss4.pdf [July 2012].

[17]The 6-digit North American Industry Classification System (NAICS) codes replaced the 4-digit Standard Industrial Classification (SIC) codes in 1997, but continuity mappings (from SIC to NAICS) exist at many federal agencies using these codes. IRS uses NAICS-based codes for its tax returns, but only what can fit on the allotted one page of the form instructions. These vary by business entity according to the business activity distribution; for example, Form 1065 codes differ from those for Form 1120.

[18]See http://www.irs.gov/pub/irs-pdf/fw4.pdf [July 2012].

employers, as part of the National Directory of New Hires at the U.S. Department of Health and Human Services (HHS) (see related discussion below under Confidentiality). One use of this form, in addition to its potential for identifying increases in national employment on practically a real-time basis, is that it individually identifies new employees, something that may be of interest for EEO-1 reports to track employment by employers.

Form W-7, Application for IRS Individual Taxpayer Identification Number (ITIN)[19]

Form W-7 is filed for foreign workers, regardless of immigration status,[20] in place of an application for SSN. The ITIN is important not only because of the foreign nativity information it contains, but also because it helps complete identification of the worker universe information, supplementing and complementing the SSNs reported for more permanent status workers. Thus, it indirectly helps provide detailed worker information on the forms filed with the states as well as a more complete picture of the employee/employer relationships revealed by Form W-2 filings. In addition, most immigrants—even in illegal status—have incentives to have an ITIN so that they and their employers can file required tax returns. The worker may have the additional incentive of obtaining a tax refund later.

Form SS-8, Determination of Worker Status for Purposes of Federal Employment Taxes and Income Tax Withholding[21]

Although Form SS-8 is not required, it may be filed by either a worker or firm to determine whether a worker should be considered an employee or independent contractor. The resulting determination may have ramifications for not only IRS forms, such as the W-2 and employment tax returns, but also for UI and related record filings with the states for SUTA and their employment training administration programs. One purpose of a related return, Form 1099 Miscellaneous,[22] is to report payments to contractor workers. Thus, this information, in conjunction with compensation information reported for employees, can help provide a complete worker compensation picture by employer.

In addition to helping capture information for contractors required to complete EEO-1 reports, such information might also be helpful for EEOC

[19]See http://www.irs.gov/pub/irs-pdf/fw7.pdf [July 2012].

[20]From a general policy perspective, IRS has not cared about an immigrant's legal or illegal status, only that the employee and employer file required returns and withhold and pay all required taxes.

[21]See http://www.irs.gov/pub/irs-pdf/fss8.pdf [July 2012].

[22]See http://www.irs.gov/pub/irs-pdf/f1099msc.pdf [July 2012].

in determining which employers might be avoiding compliance with EEOC requirements and which are evading compliance. To paraphrase IRS compliance parlance, avoidance would be considered legal, but not evasion.

Additional Income Tax Returns

Finally, several returns report earnings at both the individual and firm levels. For the former, Form 1040 and the related Schedule C (for sole proprietorships) report individual and self-employment earnings. Moreover, when the Schedule C's filer is also an employer, the Schedule C will contain compensation information for the firm's workers; for example, Cost of Labor. At the firm level, aggregate employment compensation—salaries and wages, cost of labor—can be found on the Form 1120 series,[23] in addition to an item of possible interest for expanding EEO-1 reports, namely, compensation to officers of the corporation. Aggregate employment compensation is also reported on pass-through forms, such as the Form 1065 series[24] for partnerships and Form 1120-S[25] for subchapter S investors. Income and taxes are reported for the individual partner or investor on Schedule K-1[26] and the respective Form 1040 (although partners and investors may be businesses, not individuals).

An additional sector of employers may also be of interest for the EEOC, namely, nonprofit or tax-exempt organizations that have to file Form 990, Return of Organization Exempt from Income Tax,[27] (or the related Form 990-T,[28] Exempt Organization Business Income Tax Return). Both forms, especially the former, report a number of earnings items of potential interest, including aggregate cost of labor and compensation to officers, as well as detailed individual compensation to officers, directors, trustees, and highly compensated employees.

Limitations of IRS Data

Although IRS data include a wealth of earnings data by individual employee and employer, they include establishment data only when an establishment is also an enterprise (and has an EIN). Another limitation is that they contain no data by gender (except, sporadically, for the Statistics of

[23] See http://www.irs.gov/pub/irs-pdf/f1120.pdf [July 2012].
[24] See http://www.irs.gov/pub/irs-pdf/f1065.pdf [July 2012].
[25] See http://www.irs.gov/pub/irs-pdf/f1120s.pdf [July 2012].
[26] See http://www.irs.gov/pub/irs-pdf/f1065sk1.pdf [July 2012] and http://www.irs.gov/pub/irs-pdf/f1120ssk.pdf [July 2012].
[27] See http://www.irs.gov/pub/irs-pdf/f990.pdf [July 2012].
[28] See http://www.irs.gov/pub/irs-pdf/f990t.pdf [July 2012].

Income [SOI] individual Form 1040 tax sample), race/ethnicity, or nativity (except for ITIN applications).

SOCIAL SECURITY DATA

Although a massive amount of data exist at SSA, the data of most interest for expanded EEO-1 reports are captured from the application for an SSN and the linkable federal tax data shared by IRS. Thus, only these data are discussed below.

Purpose

The data at SSA are used for administering the Social Security and Medicare programs mandated by law. Nevertheless, a related purpose is the statistical analysis necessary for such administration, conducted by not only the Office of the Actuary, but also the Office of Research, Evaluation, and Statistics (ORES). The latter would most likely be the office with which the EEOC would need to discuss any future work involving EEO-1 report data.

Content

Form SS-5,[29] Application for Social Security Number, is administered by SSA and captures gender, race/ethnicity, and nativity—often shortly after birth for most U.S. citizens. In addition, it captures citizenship status, which might be used as a proxy for or to supplement nativity information. Although the Form SS-5 data are self-reported, SSA uses supporting documentation for verification—particularly for changes, such as a marriage license (name), passport (citizenship), and birth certificate (place of birth). The Form SS-5 data, including updates, are maintained on SSAs Numident file. Because many people, such as nonretirees, have more incentives to update their tax information changes, say, name and address due to marriage or divorce, the tax information at IRS may be updated before the Numident data. However, because of the Form W-2/941 reconciliation process partnered by SSA and IRS on withholding for income, Social Security, and Medicare taxes, SSA has these data as an additional source for updating changes to the Numident, and can also query the individuals and firms in case of doubt.

[29]See http://www.ssa.gov/online/ss-5.pdf [July 2012].

Quality

Because of the supporting documentation, the SSA-IRS relationship (as well as the SSA-Census Bureau relationship), and penalties for noncompliance, filers should have incentives to provide accurate and timely data, although some limitations may be inherent. For example, although nativity data classified by country might be considered relatively reliable, researchers have noted that some of the "foreign born" may be, in fact, the progeny of American citizens, say, for military and other Americans stationed overseas, where birth occurs. In conjunction with the citizenship status, however, the data are probably useful for indicating native vs. foreign born status—the same nativity classifications published by BLS for its household surveys.

Like IRS and UI administrative data, SSA data are imperfect, but the interrelationship of these seemingly disparate data sets is usually a strength. For example, incorrect decedent data[30] might be passed on from SSA to the IRS (for individuals on its Individual Master File), but ultimately, a tax return, say, the Form W-2, could help rectify the mistake, even if it had not already been corrected through other means.

Utility for EEO-1

Together with the detailed earnings data obtained from the IRS Form W-2 and Form 941, SSA classification data of gender, race/ethnicity, and nativity should be considered a potent source of information at the individual employee level, with the essential crosswalk of SSN (or ITIN) and EIN enabling linkage to the respective employers.

U.S. CENSUS BUREAU

Although the U.S. Census Bureau is a federal statistical office and does not collect administrative data directly or participate in the administration of the related programs, it may play a unique role in the utilization of administrative data for purposes of expanded EEO-1 reports for several reasons.

First, the Census Bureau is an established and long-time user of administrative data for statistical purposes, and it has developed institutional expertise in the integration of these data with its own statistical survey and census data. For example, the bureau has long considered IRS tax data to be the "lifeblood" of its business register and related business programs, and

[30]For example, see "Social Security Wrongly Declares 14,000 People Dead Each Year," CNN, August 17, 2011, available at: http://money.cnn.com/2011/08/17/pf/social_security_deaths_mistakes/index.htm [July 2012].

uses these tax data both for sampling frame purposes and to supplement and improve the quality of its own data programs. Moreover, although IRS data are not necessarily reported at the establishment level, the Census Bureau integrates them in the business data system it maintains at the establishment level. Thus, the bureau could be an important resource in the establishment as unit of measurement.

Second, the Census Bureau has partnered with other statistical agencies, such as ORES and BLS, on mutually beneficial programs. For example, SSA and the Census Bureau collaborate on work matching individual record data from the Current Population Survey (CPS) and the Survey of Income and Program Participation (SIPP) to Numident data and Form W-2 tax data at SSA. In addition, the Census Bureau and BLS have collaborated on work involving the record to record comparison of their respective business registers, including IRS tax data on the bureau's register.

Third, the Census Bureau is authorized to access earnings data from federal tax data that most other federal agencies, including SSA, are not authorized to access, for example, from income tax returns for individuals and employers.

Fourth, the Census Bureau has a unique vehicle, the special sworn status process, which enables outsiders—from other federal agencies and even outside the federal sector—to access non-anonymized or confidential data. Such access might be granted to experts to help ensure that the work being done by the bureau for an outside sponsor is conducted according to the sponsor's ultimate needs.

Finally, the Census Bureau has not only considerable resources as an advantage over most other statistical agencies—conferring significant economies of scale—but also a wealth of data in its own Title 13 programs, such as individual gender and race/ethnicity data from the decennial census.

For purposes of the EEO-1 reports, these factors suggest that the Census Bureau might play an important role in providing data access to EEOC staff (or its contractors, with special sworn status), under mutually suitable terms, and also in partnering with other agencies, such as SSA, in order to use data that the Census Bureau may not possess on a regular basis for the population of employees.

CONFIDENTIALITY AND DATA ACCESS

IRS Data

In general, individually identifiable tax data at IRS, frequently referred to as federal tax information, are considered confidential and nondisclosable unless such disclosure or access is authorized by statute, meaning that both bodies of Congress have passed a bill signed into law by the president.

De-identification (removing identifier information such as name, SSN, address, etc.) of the data is considered a necessary but insufficient measure for anonymizing tax data, and statute (e.g., sections 6108(c) and 6103(j)(4) of Title 26 of the U.S. Code) requires that publicly released data be protected from both direct and indirect re-identification (e.g., in conjunction with other publicly released data by other agencies). Also, in general, tax data must be used for purposes of tax administration, which includes both a statistical and research component. Thus, the IRS SOI and the IRS Research Office are statutorily authorized users of identifiable tax data within the tax agency. Non-tax administration uses are discouraged, and IRS policy—and sometimes statute, e.g., for the Census Bureau—has been that even authorized tax data access should be to the minimal extent necessary to accomplish an authorized purpose. Over the years the overriding concern manifested by IRS has been the protection of the voluntary tax system, and confidentiality protection is seen as a cornerstone of such protection. Thus, any usage or proposed usage of tax data, even by authorized users, that is perceived by the public or IRS (or third-party scrutiny—Congress, the media, privacy advocates, etc.) to threaten taxpayer confidentiality is unlikely to be entertained and likely to be rejected.

However, non-tax administration accesses exist statutorily and certain exceptions have been made—rarely—to statutory access. Thus, HHS is authorized to access limited federal tax information for purposes of its child support program. The Census Bureau and a few other agencies (such as the Bureau of Economic Analysis) are statutorily authorized to access selected tax data for another non-tax administration purpose, namely, statistical purposes. Sometimes, a statute requires that the Treasury Department's regulations specify exactly which tax items may be accessed; e.g., for Census, but not for the Congressional Budget Office and SSA. These latter uses must be argued on a case-by-case basis, a less arduous process than changing statute and the regulations, but not an easy one, as IRS must be convinced[31] with a compelling business case that does not compromise even the perception of confidentiality protection and the voluntary compliance tax system.

BLS, which is not statutorily authorized to access federal tax information, has accessed identifiable tax data, both through special sworn status at Census/IRS-approved facilities, say, for the Census-BLS business register comparison project mentioned above, but also as BLS employees. The latter access involves only "minor" tax data access, such as EINs, for a long-standing procedure enabling the two agencies to compare industry codes in order to better harmonize some national statistics by that classification.

[31]Furthermore, Treasury's Assistant Secretary for Tax Policy must officially approve an amendment to the regulations.

Another way to access tax data is to obtain the taxpayers' waivers of their confidentiality rights for specific tax data through an established, formal IRS documentation process. This method has been employed successfully for a limited amount of tax data for the Health and Retirement Study (HRS). However, as one might imagine, obtaining these waivers for a sample size similar to that of HRS is probably more plausible than for the EEO-1 universe.

SSA Data

As indicated above, SSA's classification data—gender, race/ethnicity, nativity—on the Numident have been linked to tax data as well as data from important surveys, such as the CPS and SIPP, in a long-standing collaborative arrangement with the Census Bureau. However, this work has been for exclusively statistical purposes, a factor that EEOC would need to consider in formulating a data request.

Nevertheless, statistical needs might be joined to the administrative needs of both SSA and the EEO-1 Program, perhaps under a mutually beneficial agreement involving data from the National Directory of New Hires (NDNH) database at HHS (maintained at the SSA National Computing Center). SSA's National Computing Center stores and maintains the NDNH, which also includes other data that may be of interest to the EEO-1 Program, such as from Form W-4, Employee's Withholding Allowance Certificate,[32] completed for most employees on the first day of employment. The W-4 can thus be used to identify not only new employees, but also new employment in the aggregate, and on a virtually real-time basis. It may be worth considering whether such an arrangement, perhaps most viable as a statistical research program under the child support statute underpinning the NDNH, might make sense, especially if HHS might derive some benefit as well. One proposal might involve a detailed analysis of EEO-1 employees who had experienced job discrimination and any relationship of such discrimination to the status of such employees as absent parents. This is only a simple example for illustration; surely, more sophisticated research proposals could be devised that might meet some of the needs of all three agencies.

State Data

Although the collection of UI data is largely funded by ETA, the quarterly detailed employment and compensation data are retained by states and shared outside the states only rarely. Two important exceptions include the NDNH and the Census Bureau's LEHD Program.

[32]See http://www.irs.gov/pub/irs-pdf/fw4.pdf [July 2012].

HHS's Federal Office of Child Support Enforcement (OCSE) operates the NDNH, a database established pursuant to the Personal Responsibility and Work Opportunity Reconciliation Act of 1996. The primary purpose of the NDNH is to assist state child support agencies in locating parents and enforcing child support orders[33]; however, Congress has authorized a limited number of other state and federal agencies to receive information from the NDNH for authorized purposes, including for statistical research related to the child support mandate.

Beginning in the 1990s, the Census Bureau began negotiating individual agreements with the states that resulted in their providing these detailed data to Census for the LEHD Program. Any use of these data by the EEOC should probably include some benefits for the states—as occurred with the Census Bureau collaboration.

The above examples are only illustrative of what is exceptional: in general, the data are not shared, even with a statistical agency such as the Bureau of Labor Statistics.

[33]This program is sometimes referred to as the "deadbeat dads" program, although either parent may be in its scope.

Appendix C

Proposed Pilot Tests of Compensation Data Collection

In Chapter 6, the panel recommends that the Equal Employment Opportunity Commission (EEOC) sponsor a pilot study to be conducted by an independent contractor that would provide much more reliable information about the costs and benefits of the proposed collection than the panel can provide based on existing evidence. In this Appendix, we outline two possible approaches to conducting an independent pilot study.

MICRODATA PILOT

This approach would seek to gain an understanding of the availability, sensitivity, and reliability of the employer earnings data for their employees. The pilot would be conducted by an independent contractor. This recommendation that EEOC should use an independent contractor to conduct the pilot is based on the belief that an independent contractor could gain access to information that might not be possible to collect if the data were to be used for enforcement purposes. The process of contracting with an independent contractor would also give EEOC a strong incentive to make the plan for data use sufficiently comprehensive so that the potential contractors for the pilot would develop competitive study designs and plans for analysis of the results.

The first priority of the microdata pilot test would be to assess the availability and retrievability of data items of interest for individual employees. This would likely require conducting an employer records check of employers in different industries and size classes. The records check would focus on questions of interest and would parallel the questions that have

been posed in the Office of Federal Contract Compliance Programs (OFCCP) Advanced Notice of Proposed Rulemaking (ANPRM) (OFCCP, 2011). In particular, at the establishment level, how do employers record and maintain compensation data? What internal actions would employers need to take to assemble earnings and earnings-related data? What categories of compensation-related data are maintained in computer-accessible personnel and payroll systems? How easy is it for employers to cross-haul data between the payroll record systems (earnings and hours measures) and their human resource management systems (employee characteristics and work histories)? How do these systems vary by type and size of company? How costly is it on the margin to retrieve and report these data? How much time is spent in retrieving and reporting these data (for purposes of quantifying the response burden to be reported to the Office of Management and Budget under the Paperwork Reduction Act)?

Second, it would be important to validate the availability and reliability of variables that would help to gain an understanding of the role of earnings in affirmative action and antidiscrimination enforcement. In addition to annual and hourly wages, the pilot would collect a number of core demographic variables by the present EEO-1 categories using an annual wage measure in order to test targeting firms for enforcement purposes. To the extent it is possible, the pilot should also collect additional variables that could help to explain the equal opportunity environment in the establishment and the possible influence of these variables on wage differences that may be observed. These variables might include birth date, entry level wage, and hire date. Such variables could assist in refining indicators that could help identify the possibility of discrimination based on age or seniority. These new earnings-related variables should be audited, on a random basis, to ensure that they appropriately reflect the data that reside in the employer's data systems.

The pilot could also test various earnings definitions, such as that used in the Bureau of Labor Statistics Occupational Employment Survey. In developing the test, the public responses to the OFCCP ANPRM could well be instructive.

With the results of the pilot data collection in hand, the contractor could turn to assessing the quality of the data. The contractor should compute earnings means with measures of dispersion for the estimates. With these measures in hand, the contractor should analyze the data as if the data were used by EEOC to select cases for further investigation (e.g., does earnings data assist in identifying potential discrimination cases and does it support a reasonable "threshold analysis" to determine whether or not an employer should be subject to further review and evaluation), and also as if the data were the subject of litigation (e.g., how well do the earnings data meet tests of reliability and appropriateness). The power of tests (as

illustrated in Chapter 4 of this report) to detect discrimination can be calculated to shed light on the size of the groups required to have a good chance of detecting discrimination.

Another useful test would be to use data from the Social Security Administration, the Internal Revenue Service, the states, and the Census Bureau (as discussed in Chapter 2)—or even from other data systems such as state driver's license files, passports, and visas—to explore the development of an ongoing data quality assessment tool through which the EEOC earnings data would be benchmarked against other sources, at fine levels of granularity, to assess the closeness of their match to the EEOC earnings data. There is precedent for gaining such access to otherwise confidential tax and administrative data sources. For example, the Internal Revenue Service (IRS) Statistics of Income statistical sample files have been provided to compliance offices in the IRS for statistical analysis related to their compliance (audit) needs in a manner that does not adversely compromise the statistical integrity of those sample files (see discussion in Appendix B).

The results of the pilot test could be anonymized to permit a peer review of the independent findings.

SIMPLIFIED AGGREGATED-DATA PILOT

An alternative, simplified pilot could be performed by an independent contractor using a revised EEO-1 report format. The design would be a simplified version of the pilot described below. The contractor would prepare prototype EEO-1 reports that contained sufficient wage information to permit the EEOC to calculate grouped-data test statistics for differences in the mean standardized earnings across race/ethnicity and gender groups.

The standardized wage rates (full-year earnings, not actual earnings) would have to be integrated with the other data used to produce the EEO-1 report. Audited formulas for computing the average and the relevant variances would be developed for the data within EEO-1 occupation and race/ethnicity or gender group. Audited formulas for computing all relevant test statistics would also be developed. These could be based on existing statistical analysis software, or simply vetted using such software.

In prototyping a report that permitted statistical screening using grouped data techniques, the contractor would also be directed to experiment with tabulations that controlled for birth and hire date. Once again, the goal would be to produce standardized enhancements to the EEO-1 report that properly integrated the relevant data on standardized wage rates, birth dates, and hire dates with the other data used to compute the EEO-1 report. Once the data have been integrated, the report would generate validated sufficient statistics for the grouped data comparison of conditional means, given birth date and hire date. Audited formulas for computing

all relevant test statistics based on the conditional means would also be developed. Once again, these could be computed using existing statistical analysis software or simply vetted with such software.

Because of the complexity of these calculations, and the difficulty of interpreting the raw report data, the contractor would be used to develop an electronic reporting format that the agency could then use for preliminary screening of the EEO-1 reports. The electronic reporting format, encompassing the audited formulas, could then be implemented by integration into payroll and human resource management software reporting systems, just as the option to produce the current EEO-1 report has been incorporated into such products.

Appendix D

Biographical Sketches of Panel Members and Staff

JOHN M. ABOWD (*Chair*) is Edmund Ezra Day professor of economics and professor of information science at Cornell University, research associate at the National Bureau of Economic Research, research affiliate at the Centre de Recherche en Economie et Statistique (Paris, France), and research fellow at the Institute for Labor Economics (Bonn, Germany). He is a fellow of the American Statistical Association (2009) and of the Society of Labor Economists (2006). He was also the distinguished senior research fellow at the U.S. Census Bureau (1998–2009). He served as director of the Cornell Institute for Social and Economic Research from 1999 to 2007. His current research focuses on the creation, dissemination, privacy protection, and use of linked, longitudinal data on employees and employers. In his work at the Census Bureau, he provided scientific leadership for the Longitudinal Employer-Household Dynamics Program, which produces research and public-use data integrating censuses, demographic surveys, economic surveys, and administrative data. His research on integrated labor market data is done in collaboration with the Institut National de la Statistique et des Etudes Economiques, the French national statistical institute. He is currently the principal investigator or co-principal investigator for multiyear grants and contracts from the National Science Foundation and the U.S. Census Bureau. He has published articles in the *American Economic Review*, *Econometrica*, the *Review of Economics and Statistics*, the *Quarterly Journal of Economics*, the *Journal of the American Statistical Association*, and other major economics and statistics journals. Dr. Abowd served on the faculty at Princeton University, the University of Chicago, and the Massachusetts Institute of Technology before coming to Cornell. His National

Academies' service includes membership on the Committee on National Statistics (CNSTAT) Panel on Measuring Business Formation, Dynamics, and Performance and the Panel on Access to Research Data: Balancing Risks and Opportunities, and he is currently a member of CNSTAT. He has an M.A. and a Ph.D. in economics from the University of Chicago.

H. JUANITA (NITA) BEECHER is chair of the Employment Law & Litigation Group (ELLG), a network of senior corporate in-house labor and employment lawyers for Mercer LLC, and the compliance chair for the company's U.S. diversity networks. She has more than 25 years experience in labor and employment law, particularly with class investigations by the U.S. Equal Employment Opportunity Commission (EEOC) and the Office of Federal Contract Compliance Programs (OFCCP). After serving as an in-house counsel in major corporations for more than 20 years, Ms. Beecher joined Mercer (then ORC) in 2000, and became the chair of the ELLG in 2003. In her role as compliance chair, she works with the OFCCP and the EEOC on matters of interest to Workforce Opportunity Network and ELLG members. Her corporate experience includes positions at E.I. du Pont de Nemours & Company, the Consolidation Coal Company, Arch Mineral (now Arch Coal), and McDonnell Douglas/The Boeing Company. She is a graduate of the University of North Carolina School of Law.

MARC BENDICK, JR., is an economist specializing in public and private initiatives to enhance mainstream economic opportunities for traditionally excluded individuals, families, businesses, and communities. He is the author of more than 125 books, articles in refereed journals, and testimony before congressional committees. He has also served as an expert witness in more than 175 employment discrimination cases representing both plaintiffs and defendants, including many of the nation's largest class actions. He has been a consultant on discrimination and workforce diversity management to the U.S. Equal Employment Opportunity Commission, the Office of Federal Contract Compliance Programs, the U.S. Department of Justice, and some of the nation's largest employers. Since 1984, he has been a principal in Bendick and Egan Economic Consultants, Inc. He holds a Ph.D. in economics from the University of Wisconsin and a B.A. in economics and social psychology from the University of California, Berkeley.

CHARLES C. BROWN is professor of economics and a research professor at the Survey Research Center, Institute for Social Research (ISR), University of Michigan. His past research has focused on topics such as compensating differentials, effects of minimum wage laws and of equal employment opportunity policies, the determinants of enlistment and reenlistment in the military, and the relationship between employer size and

labor market outcomes. Current work focuses on measurement error in survey data, early-retirement windows, and consequences of the relatively equal opportunity in the military for children of black soldiers. He has been involved in the design and updating of the labor market status sections of the Heath and Retirement Study (HRS), and is currently analyzing data on early out windows offered to HRS respondents. Other current projects include an analysis of the relationship between age of firm and wages, and an exploratory study on children from military families. In addition to his research responsibilities for ISR's Michigan Retirement Research Center, he is assisting the director in an advisory capacity. He holds a Ph.D. in economics from Harvard University.

ELIZABETH HIRSH is assistant professor in the Department of Sociology, University of British Columbia. Her research interests include examining gender and race inequality, organizations, and the law. Much of her research in these areas focuses on employment discrimination and the consequences of legal prohibitions and organizational policies on labor market inequality. Current research includes a project examining the market, political, and organizational conditions under which employment discrimination lawsuits filed under U.S. equal employment opportunity laws bring about change in sex and race inequality in the workplace and a study of the impact of human resources practices on discrimination disputes. Other projects include an analysis of how status characteristics, workplace conditions, and neighborhood contexts influence workers' self reports of race discrimination; an analysis of corporate adoption of gender identity and expression non-discrimination policies, and a study of the extent of occupational segregation by sex, race, ethnicity, and Hispanic origin in the U.S. labor force. She holds a Ph.D. in sociology from the University of Washington.

MARK R. KILLINGSWORTH is a professor of economics at Rutgers University in New Brunswick, New Jersey. He was previously on the faculty of Barnard College and Fisk University. His research focuses on labor economics. He is the author of *Labor Supply* and *The Economics of Comparable Worth*, and has written on comparable worth and pay equity issues. He has testified on immigration reform and comparable worth before committees of the U.S. Congress, and has been a consultant to the Canadian Department of Justice, and the U.S. Departments of Justice and Labor. He was an undergraduate at the University of Michigan, and received M.Phil. and D.Phil. degrees from the University of Oxford, where he was a Rhodes Scholar. His recent work has been concerned with family members' labor force participation decisions, labor market influences on fertility, and the effect of childhood religious instruction on adult earnings.

JONATHAN S. LEONARD is George Quist chair in business ethics in the Economics Analysis and Policy Group at the Haas School of Business, University of California, Berkeley. He has served as a senior economist for the President's Council of Economic Advisors and a fellow of the National Bureau of Economic Research. He holds a Ph.D. in economics from Harvard University. His research focuses on affirmative action, workplace regulation, job creation, and employee incentives.

JANICE F. MADDEN is professor of regional science, sociology, and real estate at the University of Pennsylvania where she has been vice provost for graduate education. She is also a research associate at the University of Pennsylvania's Population Studies Center and has previously served as director of the Alice Paul Research Center and the Women's Studies Program at the university. She has been a founder and has served on the board of directors of, and a consultant with, Econsult Corporation of Philadelphia. She has written on the economics of sex discrimination, changes in income and inequality within U.S. metropolitan areas, and wages and poverty. She has previously served on the National Research Council's Committee on Vocational Education and Economic Development in Depressed Areas and chaired the Committee to Assess the Portfolio of the Division of Science Resources Studies of the National Science Foundation. She holds a Ph.D. in economics from Duke University and a B.A. in economics from the University of Denver.

THOMAS J. PLEWES (*Study Director*) is a senior program officer for the Committee on National Statistics (CNSTAT) and interim director of the Committee on Population. He was director for previous National Research Council studies on data collection sponsored by the U.S. Department of Education—a study of the estimation of English language learners—and the U.S. Department of Agriculture—a review of the Agricultural Resource Management Survey. He has also directed several studies for CNSTAT sponsored by the National Science Foundation on topics such as research and development statistics and survey methodology. Previously, he was associate commissioner for employment and unemployment statistics of the Bureau of Labor Statistics. He was a member of the Federal Committee on Statistical Methodology and is a fellow of the American Statistical Association. He has a B.A. in economics from Hope College and an M.A. in economics from the George Washington University.

ALEKSANDRA (SESA) SLAVKOVIC is associate professor of statistics with appointments in the Department of Statistics and the Institute for CyberScience at the Pennsylvania State University, University Park, and in

the Department of Public Health Sciences, Pennsylvania State College of Medicine, Hershey. She is currently serving as an associate editor of the *Annals of Applied Statistics* and *Journal of Privacy and Confidentiality*. Her primary research interest is in the area of data privacy and confidentiality. Other related past and current research interests include statistical analysis of usability evaluation methods and human performance in virtual environments, statistical data mining, application of statistics to social sciences, algebraic statistics, and causal inference. She served as a consultant to the National Academy of Sciences/National Research Council Committee to Review the Scientific Evidence on the Polygraph in 2001 and part of 2002. She holds a Ph.D. in statistics from Carnegie Mellon University.

FINIS R. WELCH is president of Welch Consulting. He testifies frequently on statistical and economic issues involving a variety of issues from allegations of employment discrimination to underwriting criteria for insurance companies. He is also distinguished professor emeritus of economics at Texas A&M University and professor emeritus of economics at the University of California, Los Angeles. He earned his Ph.D. at the University of Chicago and taught microeconomic theory, econometrics, and labor economics to graduate students for 39 years. He has testified before Congress on various issues relating to public policy; his publications on the economics of income, education, and employment have been frequently cited in the professional literature. He is an elected member of the American Academy for the Advancement of Science and a fellow of the Econometric Society. He is past vice president of the American Economic Association and past president and vice president of the Society of Labor Economists.

VALERIE RAWLSTON WILSON is an economist and vice president of research at the National Urban League Policy Institute in Washington, DC, where she chairs the National Urban League's Research Council and is responsible for planning and directing the Policy Institute's Research Agenda. She is also a member of the National Urban League President's Council of Economic Advisors, which assists the League in shaping national economic policy. Under her direction, the Policy Institute recently launched *State of Urban Jobs,* a component of Iamempowered.com, that features the Institute's research and policy analysis and serves as a vehicle for communicating the latest information related to African American and urban employment issues. Dr. Wilson has served as managing editor, associate editor, and contributing author for the National Urban League's annual *The State of Black America* report and oversees production of the National Urban League's annual Equality Index™. In 2001, a report she co-wrote with William E. Spriggs—formerly executive director of the National Urban

League's Institute for Opportunity and Equality (IOE)—earned the IOE the Winn Newman Award from the National Committee on Pay Equity. Dr. Wilson earned a Ph.D. from the Department of Economics at the University of North Carolina at Chapel Hill. Her fields of specialization include labor economics, racial and economic inequality, and economics of higher education.

COMMITTEE ON NATIONAL STATISTICS

The Committee on National Statistics (CNSTAT) was established in 1972 at the National Academies to improve the statistical methods and information on which public policy decisions are based. The committee carries out studies, workshops, and other activities to foster better measures and fuller understanding of the economy, the environment, public health, crime, education, immigration, poverty, welfare, and other public policy issues. It also evaluates ongoing statistical programs and tracks the statistical policy and coordinating activities of the federal government, serving a unique role at the intersection of statistics and public policy. The committee's work is supported by a consortium of federal agencies through a National Science Foundation grant.