

Evaluation of Compensation Data Collected Through the EEO-1 Form

The Utility of Pay Data Collected for National Comparisons and Initial Investigations

The Equal Employment Opportunity Commission (EEOC) was created in 1964 to enforce Title VII of the Civil Rights Act. Later, its authority expanded to enforce the Equal Pay Act of 1963, the Age Discrimination in Employment Act of 1967, the Americans with Disabilities Act of 1990, and the Civil Rights Act of 1991.

Starting in 1966, EEOC initiated the EEO-1: a mandatory survey of private employers with 100 or more employees and federal contractors who meet certain criteria. The survey collects data about employee race/ethnicity, sex, and job type. For 2017 and 2018, EEOC added new questions to the survey, known as Component 2, that asked about employees' pay rates (using 12 categories) and the number of hours worked. In 2020, EEOC requested the National Academies of Science, Engineering, and Medicine to evaluate the data collected by these new additions to the survey.

This brief focuses on the utility of Component 2 data for 1) national comparisons and 2) initial investigations by EEOC staff. The panel analyzed the survey data in two ways. To assess the usefulness of Component 2 data for national comparisons, the panel compared Component 2 data to national wage data collected by the U.S. Census Bureau through the American Community Survey (ACS). Second, to determine the value of the data for assisting EEOC staffs' initial investigations of individual employers, the panel examined a specific industry (the tech sector) in a particular geographic area (Silicon

Valley), using as an exemplar a previously published EEOC report.¹

Other briefs discuss other aspects of the data, such as the overall quality of the data, the measurement of pay, the usefulness of the data for employers' self-assessment, and the usefulness of the data for evaluation of pay among groups protected under EEOC's authority. The panel's findings, conclusions, and recommendations across all of these topics are summarized in the report's highlights, and presented in its full text.

FINDINGS

Exploring National Patterns of Pay Disparities The EEO-1 pay data provide a unique opportunity to examine pay disparities at the employer level. Although other federal surveys collect pay data at the individual level, the EEO-1 pay data instrument is the only survey to collect pay, occupation, sex, and race/ethnicity data of employees at the employer and establishment level. Since investigation and resolution of workplace discrimination charges require information regarding specific employers and their establishments, collecting and analyzing such data has considerable value.

After addressing data quality issues, the panel found that EEO-1 pay data produced raw pay gap results by sex and

¹ U.S. EEOC. (2021). *Diversity in High Tech*. Available: <https://www.eeoc.gov/special-report/diversity-high-tech>.

race/ethnicity similar to those produced from the ACS. This supported the potential utility of EEO-1 pay data.

However, although the EEO-1 survey collected establishment data not available in other surveys, it also omitted data that other surveys have shown to be valuable. In order to produce comparable data, the ACS data were modified by categorizing the exact pay amounts into pay bands; by lessening the detail contained in the ACS occupation groups, and by dropping some of the explanatory variables (such as age and education) that have been shown to be important in describing pay differences.

Assisting Initial Investigations of Individual Employers

One of EEOC's planned uses of the data is as a preliminary tool for charge investigations. Such analyses would provide an initial indication of the extent, if any, of raw sex and race/ethnicity pay gaps at a particular establishment and within the particular occupation. Comparing the sorting of employees across occupations and pay bands could indicate the degree to which internal inequalities are produced by segregation processes. Such information could be used to guide the next steps of the investigation, such as assigning priority and resources to the case, requesting more information from the employer so that analysts could adjust raw pay gaps by individual factors known to affect pay (such as education and job tenure), or closing the investigation.

The panel found that the data could be suitable for calculating raw sex-race/ethnicity pay gaps by occupation, establishing local labor-market/industry pay gaps and comparing targeted establishments with local averages, and for employer self-assessment of raw pay gaps.

Through its exemplar case study profiling four individual establishments, the panel identified some situations in which a charge could warrant further investigation, and other situations in which pay gaps were so small that further investigation was unlikely to be productive.

Data Quality Issues The panel's analysis of the data was limited by several data quality issues that substantially

restrict the utility of the data. Some issues were unavoidable: for example, if an employer is small with relatively little variation in employee characteristics, then there may not be sufficient data to support statistical analysis. Other data quality issues relate to the measures used, the data collected, and how the data were processed. For instance, the aggregation of hours worked data, without information on part-time or part-year status, made it difficult to convert reported wages into hourly wage rates.

If pay data were collected in greater detail, rather than in 12 broad pay bands, the data would better support analysis at the level in which pay disparities often occur. If part-time, full-time, and part-year status were collected, comparisons of annual pay and hourly wage rates would be more robust. If the EEO-1's coverage of private firms were more complete, then an employer under investigation would be more likely to appear in the EEO-1 pay data file, and pay comparisons to peer employers would be more robust. If additional data were collected on other factors that affect pay levels, such as education and experience, EEOC would be in a better position to determine which pay disparities were legitimate. If sex and race/ethnicity were collected in a way that permitted more detailed analysis, then comparisons of protected groups to similarly situated employees would be more robust. If EEOC conducted better survey design and data cleaning operations, then the data would be less subject to extreme outliers that appear to be errors rather than legitimate data.

RECOMMENDATIONS

Following are recommendations to improve the usefulness of the data.

Before 2017-2018 Component 2 data are used to assist initial investigations of charges, for employer self-assessment, or for research on pay differences more generally, the data should be carefully reviewed and cleaned. Filtering on employee counts and on hours worked would be beneficial, but some issues would be best addressed by modifying the basic data collection methodology. [Recommendation 5-1]

If EEOC continues to collect pay data in bands, narrower pay bands should be adopted, and the number of bands should be expanded for top earners to better capture variation in pay. [Recommendation 3-4]

EEOC should develop, test, and (if found acceptable) implement modifications to the Component 2 instrument to collect individual-level employee pay data, which reflects employers' current reporting practice to state and federal agencies. EEOC's transition to individual-level pay data should be informed by the Occupational Employment and Wage Survey instrument and protocol. Field testing should estimate respondent burden relative to alternative methods and assess confidentiality protections to be applied. [Recommendation 3-7]

EEOC should improve the coverage of its master list, perhaps using an interagency agreement between EEOC and Bureau of Labor Statistics to appropriately maintain business registers. [Recommendation 4-1]

EEOC should distinguish between Fair Labor Standards Act-exempt and non-exempt workers, and between part-time and full-time workers. A measure for the number of weeks worked should be included to account for part-year employment. EEOC should only collect hours worked for non-exempt employees. [Recommendation 3-9]

EEOC should explore the measurement of pay gaps for additional groups protected under its authority or policy equities, including persons age 40 and older, persons with disabilities, and veterans. To do this robustly while minimizing respondent burden, other federal data collections measuring pay of these groups, such as the American Community Survey, may be instructive. [Recommendation 3-10]

EEOC should update instructions to filers to conform to the federal standard on measuring race/ethnicity. This standard offers solutions for reporting race/ethnicity data in a combined format. [Recommendation 3-5]

EEOC should work with other federal agencies to develop and test ways to measure employees' sex, gender identity, and sexual orientation in a manner appropriate for EEOC data collections. [Recommendation 3-6]

EEOC should work with employer groups and federal data collection agencies to explore ways to collect individual-level data, such as education, job experience, and tenure, which will support detailed pay-disparity analyses and employer self-assessments. [Recommendation 3-11]

PANEL ON THE EVALUATE THE QUALITY OF COMPENSATION DATA COLLECTED FROM U.S. EMPLOYERS BY THE U.S. EQUAL EMPLOYMENT OPPORTUNITY COMMISSION THROUGH THE EEO-1 FORM **William M. Rodgers III** (*Chair*), Federal Reserve Bank of St. Louis; **Paul P. Biemer**, RTI International; **Lisa Catanzarite**, UNITE-LA and Associate of Chamber of Commerce Executives; **Siwei Cheng**, New York University; **Rebecca Dixon**, National Employment Law Project; **Lisette Garcia**, The Pennsylvania State University; **Claudia Goldin**, Harvard University; **Judith K. Hellerstein**, University of Maryland; **Elizabeth Hirsh**, University of British Columbia; **M.V. Lee Badgett**, University of Massachusetts; **Kristen M. Olson**, University of Nebraska-Lincoln; **Donald Tomaskovic-Devey**, University of Massachusetts; **Valerie Rawlston Wilson**, Economic Policy Institute

STUDY STAFF **Jennifer Park**, Study Director; **Bradford Chaney**, Senior Program Officer; **Eric Grimes**, Senior Program Assistant

FOR MORE INFORMATION

The Issue Brief was prepared based on the Consensus Study Report *Evaluation of Compensation Data Collected Through the EEO-1 Form* (2022). Copies of the Consensus Study Report are available from the National Academies Press at www.nap.edu. Recordings of the workshops organized by the committee are available on the National Academies website at <https://www.nationalacademies.org/our-work/panel-to-evaluate-the-quality-of-compensation-data-collected-from-us-employers-by-the-equal-employment-opportunity-commission-through-the-eo-1-form>.

The study was sponsored by the U.S. Equal Employment Opportunity Commission. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

Division of Behavioral and Social Sciences and Education

**NATIONAL
ACADEMIES** *Sciences
Engineering
Medicine*

Copyright 2022 by the National Academy of Sciences. All rights reserved.