

Appendix: Methodology (Measurement Detail)

A1.1 Overall Grading Methodology

Industry final grades are calculated from the average of the grades for overall representation, managerial representation, segregation, and wage gaps. They are calculated in a grade point average format. We use the conventional U.S. education A-F scale, plus a “super” failing grade of Z for industries that are so far below passing on some dimension that we felt that they must be singled out. The letter grades for the four measures of disparity were assigned the following numeric values: 4 for A’s, 3 for B’s, 2 for C’s, 1 for D’s, 0 for F’s, and -1 for Z’s.

The mean score of these gives the grade point average. The grade point average (GPA) is then converted into letter grades based on the following scale:

- A GPA greater than 3.67
- A- GPA greater than 3.33 and less than or equal to 3.67
- B+ GPA greater than 3.00 and less than or equal to 3.33
- B GPA greater than 2.67 and less than or equal to 3.00
- B- GPA greater than 2.33 and less than or equal to 2.67
- C+ GPA greater than 2.00 and less than or equal to 2.33
- C GPA greater than 1.67 and less than or equal to 2.00
- C- GPA greater than 1.33 and less than or equal to 1.67
- D+ GPA greater than 1.00 and less than or equal to 1.33
- D GPA greater than 0.67 and less than or equal to 1.00
- F GPA greater than 0.00 and less than or equal to 0.67
- Z GPA less than or equal to 0.00

For display in tables, industries are sorted by the mean of the industries’ rankings from the four primary measures of disparity, within their overall grade.

A1.2 Overall Representation

Overall representation is the proportion of all employees in a workplace that are of a particular demographic group. Focusing on establishments with at least 50 employees, we measured disparities in each group's overall representation by comparing representation in 2012 EEO-1 reporting establishments to representation in labor markets defined using the 2007-2011 ACS PUMS 5-year estimates.

To create the baselines, we limited the ACS sample to civilians who were engaged in paid work at the time of the survey. We used two complementary labor market baselines: the state labor market and the state-industry labor market. State labor markets include all workers in each state. State-industry labor markets include workers within a particular industry within each state. We used NAICS 2007 two-digit codes to determine baseline industry. For the purposes of calculating overall representation disparities, these labor markets were also divided into ten EEOC occupational categories. For each labor market, the baseline was calculated by taking the proportion of workers in that labor market in each occupation who were from the target demographic group.

We calculated workplace overall representational disparities separately using the state labor market baselines and the state-industry labor market baselines. To calculate within workplace overall representation disparities we first calculated the proportion of a particular demographic group of workers within each EEOC occupation in that workplace. Second, the occupation's state or state-industry labor market baseline was subtracted from the proportion; this gave the raw occupational representation disparity. Third, this raw occupational representation disparity was weighted by the number of employees in that occupation in the workplace. Finally, the weighted occupational representation disparities for all occupations were summed and divided by the total number of employees in the workplace.

Overall representational disparity can be represented algebraically as

$$(\bar{X} - \mu_0) = \left[\sum_{o=1}^{10} n_o (\pi_o - \pi_{ACS}) \right] / n_{EEO}$$

where $(\bar{X} - \mu_0)$ is the representational disparity, n_o is the weight for number of workers in one of ten occupations within the establishment, π_o is the proportion of each demographic group workers in an occupation within the establishment, π_{ACS} is the proportion of workers in that demographic group within the ACS baseline labor market, and n_{EEO} is the total number of workers in the establishment.

After state level and state-industry level representational disparities were calculated for each workplace, these two measures were averaged to create the overall representation scale. Positive values indicate greater representation of the demographic group in the workplace.

Description of Tables: Table A.2.1, A.3.1, A.4.1, A.5.1, A.6.1 display the mean values of these scales for each of 92 industries, along with their rankings and measures of statistical significance for female, African American, Hispanic, Asian, and Native American respectively. The industries that under-represent each demographic group are listed first. Below are descriptions of the measures included in the tables.

- At the top of each table, we include the number of establishments analyzed, and the

mean and standard deviation for the industries' mean values on the overall representation scale.

- *Mean Value* is the mean score on the overall representation scale among establishments in the industry (potential range -1.0 to 1.0). The Cronbach's alpha measure of scale reliability is 0.87 for female, 0.97 for African American, 0.97 for Hispanic, 0.98 for Asian, and 0.98 for Native American.
- *S.D.* is the standard deviation of the Mean Value.
- *% Establishments Underrepresenting* is the percent of establishments in the industry with an overall representation scale score below zero.
- *Sign Test Probability of Negative Comparison* is a one-sided binomial significance test. The null hypothesis is that the number of negative cases could have happened by chance in a normal distribution centered at zero.¹
- *% Establishments Significantly Lower than Baseline* reports results from a one-tailed z-tests for difference of proportions between each establishment's representation of each demographic group and the labor markets' representation of the group. We use 0.05 as the alpha level.²

Grades and Rankings: Grades for overall representational disparity are based on the number of standard deviations (SD) away from zero (equal representation) overall representation scale score. The follow standards were applied:

- A is awarded if the industry's score is more than 1.14 SD above zero.³
- B is awarded if the industry's score is more than .39 SD above zero, but no more than 1.14 SD above zero.⁴
- C is awarded if the industry's score is no more than .39 SD above zero and no more than .39 below zero.⁵
- D is awarded if the industry's score is more than .39 SD below zero, but no more than 1.14 SD below zero.⁶
- F is awarded if the industry's score is more than 1.14 SD below zero, but no more than 1.96 SD below zero.⁷
- Z is awarded if the industry's score is more than 1.96 below zero.⁸

¹ This measure tests whether the distribution of establishments over- and under-representing in the industry follows the binomial distribution, $X \sim B(n, p)$. Here X is the distribution of positively coded, over-representing establishments, and negatively coded, under-representing establishments, n is the number of establishments in the industry, and p is 0.5. When the value for the Sign Test Probability is at or below 0.05, we can say that we have statistically significant under-representation in the distribution.

² The formula for the z-test is $Z = (\bar{X} - \mu_0) / s$, where $(\bar{X} - \mu_0)$ is the representational disparity (defined in text), and

$$s = \sqrt{\frac{\pi_{ACS} * (1 - \pi_{ACS})}{n_{ACS}} + \frac{\pi_{EEO} * (1 - \pi_{EEO})}{n_{EEO}}}$$

where π_{ACS} is the proportion of workers of the demographic group within the ACS baseline labor market, n_{ACS} is the total number of workers in the ACS baseline, π_{EEO} is the proportion of workers of the demographic group within the establishment, and n_{EEO} is the total number of workers in the establishment.

³ In a normal distribution, a score of 1.14 would realize 12.7% of the sample in the A range.

⁴ In a normal distribution, a score between .39 and 1.14 would realize 22.3% of the sample in the B range.

⁵ In a normal distribution, a score between -.39 and .39 would realize 30% of the sample in the C range.

⁶ In a normal distribution, a score between -.39 and -1.14 would realize 22.3% of the sample in the D range.

⁷ In a normal distribution, a score between -1.14 and -1.96 would realize 7.7% of the sample in the F range.

⁸ In a normal distribution, a score of -1.96 would realize 5% of the sample in the Z range.

Rankings for overall representational disparity are based on scores from the overall representation scale. A ranking of one indicates relatively high representation of each demographic group. The rankings align with the grades. Industries with higher ranks get better grades.

A1.3 Managerial Representation

Managerial representation is the proportion of managers in a workplace who are of a particular demographic group. We measured disparities in managerial representation through comparing the managerial representation of women in 2012 EEO-1 reporting establishments to their representation in labor markets defined using the 2007-2011 ACS PUMS 5-year estimates. We created a managerial representation scale for this report. The scale was constructed from three component variables:

- 1) Workplace managerial representation versus a state labor market baseline
- 2) Workplace managerial representation versus a state-manager labor market baseline
- 3) Workplace managerial representation versus a state-industry-manager labor market baseline

The labor market baselines from the ACS PUMS data are the following:

- 1) State labor market baselines are the proportion of the state's labor market in each demographic group
- 2) State-manager labor market baselines are the proportion of managers within the state in each demographic group
- 3) State-industry-management labor market baselines are the proportion of managers within the state and in the same NAICS 2-digit industry in each demographic group

Each workplace's managerial representational disparity was calculated by first subtracting the labor market baseline from the proportion of managers in each demographic group. This was done for the three labor market baselines, and then, the average of those three values became the managerial representation scale. Positive values indicate greater representation in management.

Managerial representational disparity can be represented algebraically as

$$(\bar{X} - \mu_0) = \pi_{EEO} - \pi_{ACS}$$

where π_{ACS} is the proportion of each demographic group within the ACS baseline labor market, and π_{EEO} is the proportion of managers in each demographic group within the establishment.

Description of Tables: Table A.2.2, A.3.2, A.4.2, A.5.2, A.6.2 display the mean values of these scales for each of 92 industries, along with their rankings and measures of statistical significance for female, African American, Hispanic, Asian, and Native American respectively. The industries that under-represent each demographic group the least are listed first. Below are descriptions of the measures included in the tables.

- At the top of each table, we include the number of establishments analyzed, and the mean and standard deviation for the industries' mean values on the managerial

representation scale.

- *Mean Value* is the mean score on the managerial representation scale among establishments in the industry (Potential range -1.0 to 1.0). The Cronbach's alpha measure of scale reliability was 0.96 for female, 0.99 for African American, 0.98 for Hispanic, 0.99 for Asian, and 0.99 for Native American.
- *S.D.* is the standard deviation of the Mean Value.
- *% Establishments Underrepresenting* is the percent of establishments in the industry with a managerial representation scale score below zero.
- *Sign Test Probability of Negative Comparison* is a one-sided binomial significance test. The null hypothesis is that the number of negative cases could have happened by chance in a normal distribution centered at zero.⁹
- *% Establishments Significantly Lower than Baseline* reports results for all establishments in the industry of a one-tailed z-test for difference of proportions between each establishment's representation of the demographic groups in management and the labor markets' representation of the demographic groups. We use 0.05 as the alpha level.¹⁰

Grades and Rankings: Grades for overall representational disparity are based on the number of standard deviations (SD) away from zero (equal representation to baseline) overall representation scale score. The follow standards were applied:

- A is awarded if the industry's score is more than 1.14 SD above zero.
- B is awarded if the industry's score is more than .39 SD above zero, but no more than 1.14 SD above zero.
- C is awarded if the industry's score is no more than .39 SD above zero and no more than .39 below zero.
- D is awarded if the industry's score is more than .39 SD below zero, but no more than 1.14 SD below zero.
- F is awarded if the industry's score is more than 1.14 SD below zero, but no more than 1.96 SD below zero.

Rankings for managerial representational disparity are based on scores from the managerial representation scale. A rank of one indicates relatively high in management. The rankings align with the grades. Industries with higher ranks get better grades.

A1.4 Workplace Segregation

We measure workplace segregation by examining how people from two demographic groups are distributed across occupations within establishments. Here we examine male-female, African American-White, Hispanic-White, Asian-White, and Native American-White segregation:

To measure segregation, we created a segregation scale composed of the mean of three measures: unadjusted segregation using the Duncan Index, adjusted segregation, and

⁹ See footnote 9.

¹⁰ See footnote 10 for formula. We conducted the same z-tests on only establishments that had at least 10 managers. Aggregated to the industry level, the correlation between the test used and the alternative test was always at least 0.95.

segregation by industry. Before compiling these measures to create the segregation scale, each was mean-centered.

Unadjusted Measure: The unadjusted measure is the Duncan index of dissimilarity, which measures the degree to which two mutually exclusive groups are distributed across units making up a whole. In this case, the “whole” is the workplace, and the units are the following EEO-1 occupational categories:

- 1) Top Management
- 2) Middle Management
- 3) Professionals
- 4) Technicians
- 5) Sales workers
- 6) Administrative support workers
- 7) Craft workers
- 8) Operatives
- 9) Laborers and helpers
- 10) Service workers

The index of dissimilarity formula is:

$$D = \frac{1}{2} \sum_{oe=1}^{10} \left| \frac{x_{oe}}{X} - \frac{y_{oe}}{Y} \right|$$

Where x_{oe} is the number of people from demographic group X (females; African Americans; Hispanics; Asians; Native Americans) in an occupation within an establishment, X is the number of people from demographic group X in an establishment, y_{oe} is the number of people from demographic group Y (males; whites) in an occupation within an establishment, and Y is the number of people from demographic group Y in an establishment.

The index provides the proportion of a group that would need to change occupations in order for there to be no segregation in that workplace. The index ranges from 0 to 1, with 1 indicating total segregation. Establishments do not receive an index of dissimilarity score if they have no members of one or both groups being compared.

Adjusted Measure: Because the EEOC occupation categories ignore within-occupation job-title segregation, the index of dissimilarity underestimates actual workplace segregation. To partially correct for this underestimation, the segregation value for each workplace is adjusted upward to take into account the degree of occupational heterogeneity within the workplace.

To do this, for each type of segregation we first calculated the Gibbs-Martin index of occupational heterogeneity. The formula for the Gibbs-Martin index of heterogeneity is:

$$H = 1 - \sum_{o=1}^{10} p_o^2$$

where p_o is the proportion of employees in an establishment within an occupation. Next, we regressed unadjusted segregation on the index of heterogeneity for all EEO-1 establishments.

We multiplied the estimated regression coefficient by the difference between .99 (the maximum value of the Gibbs-Martin index) and the observed workplace index of heterogeneity. This value was added to the unadjusted measure of segregation and values above 1 were recoded to 1.

Segregation by Industry: For segregation by industry, the mean adjusted segregation value for the industry is subtracted from the workplace's segregation value. This measure asks, how segregated is this workplace relative to other workplaces in the same two-digit industry?

Description of Table: Table A3 displays the mean values of this scale for each of 92 industries, along with their rankings and measures of significance. Higher levels of segregation are listed first.

- At the top of each table, we include the number of establishments analyzed, and the mean and standard deviation for the industries' mean values on the segregation scale.
- *Mean Value* is the mean score on the segregation scale among establishments in the industry (Potential range 0 to 1.0). The Cronbach's alpha measure of scale reliability is 0.90 for female, 0.93 for African American, 0.93 for Hispanic, 0.94 for Asian, and 0.88 for Native American.
- *S.D.* is the standard deviation of the Mean Value.
- *% of Establishments Significantly Higher than National Average* measures come from one-tailed t-tests for difference between the each establishment's segregation level and the national mean level of segregation. This is done separately for Observed Segregation (unadjusted measure) and Adjusted Segregation. We use 0.05 as the alpha level.¹¹

Grades and Rankings: For segregation, grades are based on the percent of establishments in the industry that have adjusted segregation scores (among the three segregation measures: unadjusted segregation using the Duncan Index, adjusted segregation, and segregation by industry) significantly higher than the national average (see above for description of measure). The grade scale is constructed as follows:

- A is awarded if the percent of establishments with adjusted segregation scores significantly higher than the national average is 15% or less.
- B is awarded if the percent of establishments with adjusted segregation scores significantly higher than the national average is more than 15%, but no more than 25%.
- C is awarded if the percent of establishments with adjusted segregation scores significantly higher than the national average is more than 25%, but no more than 35%.
- D is awarded if the percent of establishments with adjusted segregation scores significantly higher than the national average is more than 35%, but no more than 45%.
- F is awarded if the percent of establishments with adjusted segregation scores significantly higher than the national average is more than 45%, but no more than 55%.

¹¹ The formula for the t-test is

$$t = \frac{X - \bar{X}}{s/\sqrt{n}}$$

Where X is the segregation for the establishment, \bar{X} is the national mean level of segregation, s is the standard deviation of the segregation across all establishments, and n is the total number of establishments.

- Z is awarded if the percent of establishments with adjusted segregation scores significantly higher than the national average is more than 55%.

The rankings for segregation are based on the grades above and segregation scale scores. This allows more than one measure to influence the rankings. Industries are ranked within grades based on their mean rank score. Lower segregation results in higher grades and better rankings.

A1.5 Wage Gap

The wage gap is the industry average difference in wages between the focal group and a comparison group (e.g. women relative to men), after statistically controlling for differences in education, potential labor market experience, actual labor supply behavior, English language skills, and local labor market average wage.

To estimate wage gaps, we used 2007-2011 ACS PUMS 5-year estimates, limiting the dataset to civilians engaged in paid labor at the time of the survey. Wages were calculated by dividing annual income in 2011 dollars by the number of hours the respondent worked in the past 12 months. For all calculations, the natural logarithm of wages was used to control for skewness in the wage distribution.

To obtain industry wage gaps, we performed 92 regressions, one for each NAICS three-digit industry.¹² By estimating separate industry wage regressions for each industry, we allow each industry to have industry specific coefficients. We regressed logged wages on race and gender variables, human capital and labor supply variables, and fixed effects for 1238 Place of Work Public Use Microdata Areas. For each industry, we estimated the following regression:

$$\begin{aligned}
 Y_{Industry\ 1-92} = & \beta_0 + \beta_1 x_{Female} + \beta_2 x_{Black} + \beta_3 x_{Hispanic} + \beta_4 x_{Asian} + \beta_5 x_{AIAN} \\
 & + \Sigma(\beta_{6-11} x_{Educ. Attmt.}) + \beta_{12} x_{Yrs.Educ.} + \beta_{13} x_{Experience} \\
 & + \beta_{14} x_{Experience Squared} + \Sigma(\beta_{15-18} x_{Eng.Ability}) + \beta_{19} x_{Weekly Hours Worked} \\
 & + \beta_{20} x_{Weeks Worked} + \Sigma(\beta_{21-1258} x_{Community}) + \varepsilon
 \end{aligned}$$

where x_{Female} , x_{Black} , $x_{Hispanic}$, x_{Asian} , and x_{AIAN} , are dichotomous variables for each demographic group (AIAN being American Indians/Alaska Natives), $x_{Educ. Attmt.}$ is a set of dichotomous variables for highest level of education achieved, $x_{Yrs.Educ.}$ is a continuous measure of years of education, $x_{Experience}$ is a continuous variable for estimated years of work experience, $x_{Experience Squared}$ is the square of $x_{Experience}$, $x_{Eng.Ability}$ is a set of dichotomous variables constructed from an ordinal measure of English speaking ability, $x_{Weekly Hours Worked}$ is the number of hours worked in the typical week in the past 12 months, $x_{Weeks Worked}$ is the number of weeks worked in the past 12 months, and $x_{Community}$ is a set of dichotomous variables, one for each Place of Work Public Use Microdata Area (with one community omitted). Years of work experience was estimated by taking age, minus years of education, minus 6.

¹² Ninety-eight three-digit NAICS industries are available in the ACS. However, for consistency, we only include the 92 industries we used from the EEO-1. See footnote 8.

For example of male-female wage gap, for each industry, we stored the coefficient and the p-value of a two-tailed t-test on the coefficient for female. The coefficients were exponentiated and multiplied by 100 to give female's wages as a percentage of male's wages.

Grades and Rankings: Wage gap grades are based on the wage gap estimates from the equation described above. We use the following grade scale:

- A is awarded if female wages are 95% or more of male wages.
- B is awarded if female wages are at least 90% of male wages, but less than 95% of males' wages.
- C is awarded if female wages are at least 85% of male wages, but less than 90% of male wages.
- D is awarded if female wages are at least 80% of male wages, but less than 85% of male wages.
- F is awarded if female wages are at least 75% of male wages, but less than 80% of male wages.
- Z is awarded if female wages are less than 75% of male wages.

Wage gap rankings are based on the wage gap estimates as described above. Higher values result in better rankings.

A1.6 Industry Size, Pay, and Federal Contractor Context

Employment opportunity disparities are associated with industry level context in important ways. We focus on industry density of federal contractors, and so OFCCP oversight, mean wages, and overall employment size. We estimated correlation coefficients between three industry level contextual variables and overall GPA and its four main components: overall representation, management representation, wage gap, and segregation.