3. Methodology

Ranking industries in terms of employment opportunities requires us to define what equal opportunity might look like. Since opportunity is not observable, we follow legal and social science precedence and focus on disparity in outcomes, after statistically adjusting for opportunity to employ. We focus on four dimensions of employment opportunity – overall employment, employment in managerial jobs, segregation across jobs, and wage disparity.

For each dimension, we construct a series of baseline comparisons to gauge the degree of disparity. Thus, disparity is measured as deviation from equality, given a reasonable employee availability baseline. In legal proceedings around employment discrimination, the most common approaches for evaluating disparities in employment are comparisons to employment representation in the local labor marker and establishments in similar industries. These baselines conceptualize opportunity to hire as constrained by local labor market and industry representation of the target group. For example, if a workplace has significantly fewer women in management than other establishments in the same geographic area or the same industry within the same geographic area, then we can say that the disparity is indicative of failing to recruit or hire women that are available in the local labor market.

We describe these baselines in detail below. In brief, the baselines for our two indicators of employment representation – overall and managerial – are the occupation specific supply of the target group in the same state and the same state-industry. We use industry baselines at the two-digit NAICS level, but calculate disparity at the three-digit level. Thus, the industry baseline is to workplaces with similar technology and labor requirements. For segregation, the baseline is the internal distribution of groups across occupations within the workplace and the workplace specific deviation from segregation in its two-digit industry group. For earnings, the baseline is the target group relative to reference group wage, after statistical adjustment for industry specific group differences in education, labor market experiences, and local labor market wage levels.

As in any statistical exercise, we measure disparity with error. There are three basic sources of error in these estimates. The first is measurement error associated with the baselines. Following classical measurement theory, we use multiple baselines for each dimension to minimize the impact of measurement error. Appendices report scale reliabilities, all of which are quite high at the indicator level, building confidence in the scales we have computed. Sampling error is the second source of uncertainty in our estimates. For each measure, we report levels of statistical significance. Tests of statistical significance adjust for sample size and so become more conservative when samples are small.

The final source of error is unmeasured, but consequential, variables. The use of local labor market labor supply baselines does not directly observe job applications and so can mask group level self-selection away from particular employees. The EEO-1 data underestimate true segregation because only aggregate occupational distinctions are observed. Our estimates of wage gaps are based on Census, not workplace, data and incomplete adjustment measures of prior individual labor market experience. Thus, the wage gaps may be produced by disparity within establishments or by sorting between establishments. Both may be produced by establishment level bias in hiring, pay or promotion, self-selection across establishments, and unmeasured differences in employment or training.

As a result of all three sources of error, we pay less attention to the estimated absolute disparity levels, and more to the relative ranking of industries in assigning grades. We describe how grades were assigned in Appendix A1. We use the conventional educational grading system, with one exception: we differentiate failing grades into two categories, F and Z. Z grades go to industries whose disparity levels are extremely high.