

**Studying Race/Ethnic and Sex Segregation at the Establishment Level:
Methodological Issues and Substantive Opportunities Using EEO-1 Reports***

Corre L. Robinson

Tiffany Taylor

Donald Tomaskovic-Devey

Department of Sociology

North Carolina State University

Catherine Zimmer

The Odum Institute and Department of Sociology

University of North Carolina at Chapel Hill

Matthew W. Irvin, Jr.

Department of Sociology

North Carolina State University

* The research for this paper was supported by a grant from the National Science Foundation (SES-0216424). We would like to thank Kevin Stainback, Tricia McTague, Bliss Cartwright, Ron Edwards, Alexandra Kalev, Julie Kmec, and Barbara Reskin for comments on the paper and general discussion of the EEO-1 reports. An earlier version of this paper was presented at the 2003 Annual Meeting of the American Sociological Association. Please send all correspondence to Corre Robinson, Department of Sociology, North Carolina State University, Raleigh, NC 27529-8107 or clrobins@server.sasw.ncsu.edu.

**Studying Race/Ethnic and Sex Segregation at the Establishment Level:
Methodological Issues and Substantive Opportunities Using EEO-1 Reports**

Abstract

Scholars of employment segregation now recognize that gender, race, and class processes are mutually constitutive. Coupled with new data collection strategies, understanding of the organization of work and distribution of inequality will improve. We explore the strengths and weaknesses of longitudinal establishment data collected by the EEOC, comparing these to other data used to study workplace status processes. Our findings both confirm and dispute well-known occupation-based analyses of workplace segregation and lead to similar substantive conclusions. EEOC data are useful for trends in segregation, locating segregation in spatial, temporal and industrial contexts, and combining with organizational data to uncover mechanisms.

Keywords

Sex, gender, race, workplace segregation, inequality, stratification, status segregation

The sociology of work has traditionally described and understood work activity, labor processes, and inequality through a class-based lens. Sociologists of work have typically collected data, often qualitative or historical, on workplace activities, conflicts and change. The sociology of occupations has traditionally focused on inequalities in distributional outcomes of rewards such as prestige, income, or the quality of employment. In contrast to the sociology of work, researchers in this arena often utilize quantitative survey data on individuals containing indicators of individual resources such as family background and education to explain variation in access to occupation-based rewards. While both literatures arose in the context of various market-oriented class theories (e.g., Marxian, status attainment, neoclassical economic), both have increasingly focused on how status processes, particularly gender, are constituent parts of the labor process and the production of inequality. This focus on gender, race and other status processes has encouraged workplace scholars to broaden their observational lens to recognize that status and class processes are often mutually constitutive. It has encouraged occupational researchers to observe the importance of status segregation and integration for both the structure of inequality and its distribution.

Theoretical advances require that we revisit our observational tools. This paper has three goals, all linked to this methodological lesson. The first goal reflects the fundamental shift in the sociology of work and occupations from class-based conceptualizations of the labor process and status attainment to conceptualizations that broaden the actors and interests to include status distinctions such as gender or race. Hence, our first goal is to provide a comparative analysis of the strengths and weaknesses of quantitative data currently in use to reveal these status dynamics in workplaces. Our second and central goal is to describe and evaluate a major data resource for workplace analyses of status processes that has recently become available in the U.S. These data

have been collected by the U.S. Equal Employment Opportunity Commission [EEOC] since 1966 and describe, at the workplace-level, the occupational and race/gender composition of employment for private sector employees. Our third goal is to evaluate the quality and potential of the EEO data relative to theoretical goals and alternative data sources. We hope to provide the research community with a set of evaluative tools for choosing and developing data resources for the study of workplace status processes and introduce them to a major new data resource.

Studying Gender and Racial Inequality in Workplaces

Over twenty years ago Baron and Bielby (1980) urged researchers to “bring the firm back in” to analyses of stratification. At that time, most stratification research ignored the organizational context in which inequality took place. Baron and Bielby’s argument, while compelling, did little to change how data were collected. Twenty plus years after their publication, organizational context, a central feature of employment inequality, remains relatively underdeveloped in quantitative data collection strategies (Reskin, McBrier, & Kmec, 1999).

The rise of gendered and race/ethnic analyses of both workplaces and inequality have encouraged researchers to pay more attention to organizations as units where gender and ethnicity are enacted. Gender and ethnic employment segregation have been tied to the basic organization of production (Acker 1990), labor relations with management (Cockburn, 1991; McIlwee & Robinson, 1992) and other workers (Padavic, 1991; Reskin & Padavic, 1988), inequalities in the distribution of rewards (Kmec 2003a; Tomaskovic-Devey, 1993a and b), and the political struggles over status distinctions in state employment policy (Bridges & Nelson, 1989), private employers’ personnel policy (Collins, 1997), and organizational practices and politics (Nelson & Bridges 1999; Tomaskovic-Devey, Kalleberg, & Marsden, 1996).

This scholarship has encouraged us to theorize about a multitude of linked phenomena, most conspicuously the level and consequences of gender and ethnic segregation in employment (Reskin, McBrier, & Kmec 1999; Tomaskovic-Devey, 1993a). Many questions about status segregation remain unanswered. For instance, is there any change as a result of the U.S. civil rights movement and associated legal, regulatory, and normative pressures? Which types of organizations are most likely to desegregate or exhibit high or low gender or ethnic segregation? What organizational personnel practices encourage segregation or integration? What are the consequences of status segregation and integration for organizational conflict, labor process organization, and individual employment rewards? While these questions have been raised we are just beginning to scratch the surface for answers. In order to advance scholarship in both the sociology of work and in stratification, we need to develop better comparative organizational data that allow us to view class-gender-ethnic dynamics and outcomes across and within workplaces. In the next section of this paper, we list several data sources researchers have used to study status processes in workplaces and compare the strengths and weakness of these methodological strategies. We follow with a comparison of the EEO-1 surveys to the three most prominent methodological approaches – surveys of individuals, surveys of organizations, and organizational personnel records.

Comparing Data Collection Strategies

Table 1 compares the EEO-1 reports we focus on in the rest of this paper with other prominent data collection strategies for studying status segregation and inequality. The columns illustrate various data available for studying segregation, such as occupational data collected from individuals, firm level data (such as data used by Baron & Bielby, 1980, 1984, 1986 or

Peterson & Morgan, 1995), other surveys of organizations, and finally EEO-1 data collected by the EEOC. The rows represent basic criteria researchers might apply to these data.

(Insert Table 1 about here.)

The most common data collection strategy has been to use large-scale surveys of individuals such as national censuses to estimate the race or sex composition of occupations. Occupation-based analyses of workplace segregation are common in the exploration of post-civil rights trends in U.S. sex and racial segregation (Baunach, 2002; Beller, 1985; Blau, 1988; Carlson, 1992; Cotter, DeFiore, Hermesen, Kowalewski, & Vanneman, 1995; Jacobs, 1989; Jacobsen, 1994; King, 1992). They indicate post-1960 declines in segregation between men and women and blacks and whites (other race and ethnic comparisons are rare). Other studies documenting the increased representation of women and minorities in managerial roles are also based on occupational analyses as well (see the review in Smith, 2002). Because these studies typically lack data on the actual physical location of work, it is possible that much of what we think we know is incorrect or exaggerates the extent of integration. Clearly there has been an increased representation of women and minorities in managerial occupations. If these female or minority managers are typically found in one type of firm (e.g., retail trade, personal service sector) and white men remain managers in other firms (e.g., durable manufacturing, business services sectors) the integration of managerial occupations might mask a lived reality closer to industrial apartheid.

Some of the most prominent work in this area focuses on the sex composition of occupations as an explanation for individual stratification outcomes (England, Farkas, Kilbourne, & Dou, 1988; Groshen, 1991; Tam, 1998). This research builds directly on the human capital and status attainment inspired data collection strategy of surveys of individuals. Studies in this

tradition typically lack information on the jobs people actually occupy but have appended information such as the national average sex composition and skill requirements of occupations. While theoretically far removed from the advice to *bring the firm back in*, this research tradition has been influential for two reasons. The first is the widespread availability of surveys of individuals with information on individual human capital and a central job reward – employment earnings. In addition to hundreds of specialized surveys of individuals there are many major national surveys inspired by some combination of the status attainment tradition in sociology and/or the human capital tradition in economics. The U.S. Census, as well as monthly Current Population Surveys (CPS), have very large samples with measures of education, sex, race, occupation, industry, and earnings. Major longitudinal surveys such as the National Longitudinal Survey of Youth (NLSY) and the Panel Study of Income Dynamics (PSID) follow the same people in the labor force for up to thirty years with the same measures as the Census and CPS, as well as, detailed information on labor force experience and family dynamics. Cohen and Huffman (2003) may have given new life to census-based occupational studies by examining spatial variation in the degree of occupational segregation and its connection to the gender wage gap.

While surveys of individuals have been overwhelmingly dominated by an occupational conceptualization of stratification, some scholars have adapted the method to collect job and organizational data from employees. Villimez and Bridges (1988), Tomaskovic-Devey (1993a and b), and Kmec (2003b) have all shown that researchers can learn about status segregation at the job level simply by asking employees to report on the sex or race composition of their job or workplace in surveys of individuals. Surveys of individuals are limited to a single person within a workplace and so can describe between workplace segregation and inequality, but not

processes occurring within workplaces that perpetuate workplace inequality. While this approach does not collect true organizational data, information on jobs and organizations are clearly preferable to workplace scholars to national occupational analyses (Tomaskovic-Devey & Skaggs 2002).

Baron and Bielby's (1980, 1984, 1986) early research was unique in that they had access to detailed data on the population of jobs within a sample of 290 workplaces. Their findings showed that job-level sex segregation within establishments is significantly higher than estimates based on the distribution of men and women across Census occupational categories. They also showed that the labor process and various work rewards were tied to the sex composition of jobs in these organizations. Since then we have seen scholars begin to develop quantitative data that include information on organizations and/or the sex composition of employment. The limited availability of workplace data continues to hamper efforts to study employment segregation and the consequences of gender and other status distinctions for the social organization of work.

One important finding demonstrated in occupation-based studies and Bielby and Baron's (1986) establishment-level research is that as the level of job or occupational detail increases, so does measured segregation. Hence, more detail in describing the employment context is preferable to less. The best data in this regard are firm personnel records or specialized studies that capture information on all employees and all jobs in specific workplaces. These data are rare and typically for only one point in time, but produce the most accurate estimates of actual status segregation and its association with labor process organization. Bielby and Baron's 1986 paper remains the exemplar of this type of data, since it contains many workplaces and a good deal of information on the labor process. Petersen and Morgan (1995) and Petersen, Saporta, and Seidel (2000) are the only other prominent examples employing personnel records, but they are much

more limited in the information available. Later in this paper we compare EEO-1 estimates of sex segregation to Bielby and Baron (1986), Petersen and Morgan (1995) and some additional proprietary establishment personnel records. In general, firm personnel records are great data sources for studying workplace status processes and should be developed aggressively in the future. They have great potential for revealing spatial and temporal processes. Although past research has not had such data, they clearly exist in the records of many corporations.

The third data collection strategy that researchers have developed is specialized surveys of workplaces (Goux & Maurin 1997; Kalleberg & Lincoln, 1988; Kalleberg, Knoke, Marsden, & Spaeth, 1996; Kmec, 2003b; Le Grand, Hedstrom, Szulkin, & Tahlin, 1990; Mastekaasa, 1992; Villemez & Bridges, 1988). These surveys tend to be quite detailed in terms of organizational characteristics of interest to their designers but they are typically limited to descriptive information on a few jobs within the organization. The U.S. National Organizations Survey [NOS] for example, contains sex composition information on managerial jobs, the core production job in the establishments, and a third job tied to the original sampling frame from a survey of individuals (Kalleberg, Knoke, Marsden, & Spaeth 1996). Specialized organizational surveys are typically limited by small samples of establishments, geographically-specific findings (i.e., cities or countries), time-specific findings (most are cross-sectional), and a description of only a few jobs within the organization. While most specialized organizational surveys have been cross-sectional, panel studies, (like Norway's NSOE), are one strategy for exploring the dynamics of status segregation processes over time. If researchers follow the sex or ethnic composition of a stable set of jobs within a sample of organizations over time they could examine what aspects of organizational policy and practice affect status segregation.

Researchers are unlikely to agree on a single “gold standard” for data, but these comparisons reveal some data characteristics, which can be maximized. For example, if we want to study status and class dynamics in workplaces, data on jobs is preferable to data on occupations. Jobs embedded within organizations are preferable to jobs attached only to individuals. Data that allow us to track status and other changes in organizations over time are preferable to cross-sectional estimates. Data on multiple organizations that allow comparisons over time, across space, and with other organizational characteristics would be particularly useful for examining questions of both causality and social change. No one data collection strategy will contain all of these characteristics, if for no other reason than prohibitively high costs (Mastekaasa, 1992; Kmec, 2003b). However, these comparisons of basic data characteristics can help researchers determine which data may be best suited or developed for their research questions. We now turn to the EEO-1 survey data and evaluate them relative to the standards outlined in Table 1.

EEO-1 Survey Data

The main purpose of this paper is to evaluate the quality and utility of EEO-1 data, the private sector establishment surveys collected by the U.S. Equal Employment Opportunity Commission [EEOC] since 1966. As we will see, the EEO-1 surveys combine most of the desirable traits for studying workplace status processes outlined in Table 1. Compared to other organizational surveys, EEO-1 surveys have a very large number of observations that can be linked over time. Like all surveys some concepts are well measured (workplace, firm, industry, geographic location, and federal contractor status), for some we could imagine better measures (jobs rather than occupations), and other concepts are completely absent. The major limitation of

EEO-1 surveys relative to actual firm personnel records is the absence of job level detail and data on individuals. We now describe the EEO-1 survey reports in more detail.

Title VII of the US Civil Rights Act of 1964 mandates that private employers submit yearly reports on the sex and race/ethnic makeup of their employees to the EEOC. Reports include Employer Information Reports (EEO-1), Apprenticeship Information Reports (EEO-2 & 2E), Local Union Reports (EEO-3), State and Local Government Information Reports (EEO-4), Elementary-Secondary Staff Information Reports (EEO-5), and Higher Education Staff Information Reports (EEO-6) (Equal Employment Opportunity Commission, 1981). In the past, these data have rarely been available to the scientific community. For this project, access to confidential EEO-1 reports was gained through the use of an Intergovernmental Personnel Act agreement where the senior project member became an unpaid employee of the EEOC for the purposes of conducting this research.¹

EEO-1 reports contain establishment employment counts of sex by five race/ethnic groups (White, Black, Hispanic, Asian/ Pacific Islander, American Indian/ Alaskan Native) distributed across nine occupational categories (officials and managers, professionals, technicians, sales workers, office and clerical workers, craft workers, operatives, laborers, and service workers).² These reports also include information on the establishment's parent company, industry, and geographic location. Each record states whether or not the parent company is a federal contractor. Federal contractors must report affirmative action plans and progress to the Office of Federal Contract Compliance Programs (OFCCP). Lastly, the data identification numbers which make it possible to link workplaces to themselves and to their parent firm over time.

Coverage is currently limited to all private firms with 50 or more employees if federal contractors and 100 or more employers if non-contractors. Prior to 1983, separate reports were required for contractor firms with 25 or more employees and non-contractor firms with 50 or more employees.³ Firms are instructed that employees do not include temporary or casual employees, but do include leased employees as well as both part-time and full-time employees.

EEO-1 reports have the advantage of providing organizational detail to the level of specific workplaces. EEO-1 reports are, however, limited in coverage because they use standardized occupational groups rather than the job titles in use in specific workplaces to describe establishment employment distributions. Thus it is important to evaluate the gains and losses in using EEO-1 reports to study workplace segregation. To accomplish this, we compare estimates of segregation based on EEO-1 reports to previous research that use job titles within establishments as well as U.S. Census and Current Population Survey occupation-based analyses. We also compare EEO-1 data to other organizational datasets with regards to their respective strengths and weaknesses in measuring employment segregation.

The EEO-1 survey shares the advantages of other organizational surveys – true workplace data and consistent indicators across workplaces. EEO-1 surveys also provide the advantage of being able to track changes for multiple status groups and embed them in multiple contexts. Compared to more specialized surveys, such as the NOS, they have relatively little descriptive information beyond industry and location. Compared to personnel records they lack job level detail. EEO-1 data do not include any attributes of employees or specific organizational practices like personnel policies. As Mastekaasa (1992) notes, cost constraints dictate the extent to which collectors of organizational data are able to maximize the number of establishments

surveyed relative to the level of detailed information about the employees in those establishments.

An Example Using the EEO-1 Surveys of Establishments to Describe Employment

Segregation Trends

Previous research concludes that segregation in U.S. firms and occupations has decreased since the passage of the Civil Rights Act (e.g., Jacobs, 1989; Tomaskovic-Devey, 1993a and b). Remarkably, there are no workplace-based estimates of these trends. Using all available EEO-1 reports since 1966, Figure 1 reports the first such estimates, and confirms a decrease in employment segregation within firms.⁴ Figure 1 suggests that the processes of segregation and integration are different between sex and race groups and across time. Sex segregation declines continuously since 1966, while Black-White and Hispanic-White segregation trends are essentially flat from about 1980 to the early 1990s, and have barely declined since then. Black-Hispanic segregation shows a third pattern of essentially no change in segregation until after 1992.⁵ We hope that Figure 1 makes clear the potential of these data for opening many new avenues for investigation.

(Insert Figure 1 about here.)

The next section of the paper reviews EEO-1 data origins and history. We describe changes in data coverage, collection, and quality over time. This discussion is informed by correspondence with those responsible for collecting and processing the data at the U.S. Equal Employment Opportunity Commission as well as our own analyses of the data files. We then review the few previous studies that have used EEO-1 data and compare EEO-1-based estimates of employment segregation with previous research using establishment and occupation-level data. These comparisons show that EEO-1 reports are likely to produce results comparable to

more detailed job-level data and will lead to new insights beyond Census-based trend analyses.

We conclude with a discussion of the possibilities, problems, and limitations of EEO-1 data relative to other data, and offer suggestions for the research community.

EEO-1 Survey Protocols and Quality

In accordance with guidelines established in Title VII of the Civil Rights Act of 1964, employers use various forms of EEOC reporting. Reporting is contingent on the layout and structure of the firm. A single establishment firm with 100 or more employees (or 50 or more employees and a federal contract) files a single establishment report. In the year 2000 there were 14,065 single establishment reports. A multi-establishment firm with 100 or more employees (50 or more employees and a federal contract) is required to file a company-wide consolidated report, a headquarters unit report, and individual establishment reports for each establishment with 50 or more employees. There were 25,410 such firm and headquarter reports in the year 2000, with individual establishment reports for 158,250 unique workplaces. In total, there were 224,471 establishment observations for the year 2000. Between 1966 and 2000 there are over four million establishment observations, which means that by the year 2000 there are over four million workplace observations that can be linked into a panel, and embedded in their firm, industrial, and geographic context. Thus these data are excellent for tracking change in status segregation in its corporate, industrial, geographic, and government regulatory contexts.

Consolidated reports list total firm employment by race, sex and occupation and include employment in establishments too small to meet individual filing criteria. Consolidated reports do not specify industry or geographic locations of particular establishments, since these are treated as attributes of workplaces, not firms. Employment distributions for establishments with less than 50 employees are typically only counted in consolidated reports.⁶ Headquarter and

individual establishment reports contain geographic, industry, and employment information for each unit in the firm. There is also a category referred to as special reports in the EEO-1 files. These special reports include the employment counts for multiple small establishments (fewer than fifty employees) that are part of a larger firm. There were only 1059 special reports filed in the year 2000.

The EEOC estimated in 1966 that 75 percent of all employers with 100 or more employees who were required to file, actually filed EEO-1 reports (Equal Employment Opportunity Commission, 1967). We have no idea how that response rate was calculated. Later EEOC reports do not include response rate estimates but do contain coverage estimates. Figure 2 compares EEO-1 reported employment to annual average national employment estimates from the U.S. Bureau of Labor Statistics (BLS). Across all sectors, coverage drops from over 50 percent of all private sector employment in 1966 to just over 40 percent in 2000. All of the drop in coverage takes place between 1982 and 1985 and is probably a function of changes in EEO firm size reporting requirements in 1983. Taking into account the dip in 1983, EEO coverage has remained relatively stable over time. While a smaller proportion of all private sector establishments are captured post-1983, coverage rates gradually increase for most industries after 1983. This suggests that, with the exception of mining, where almost all firms are very large, there has been a slight increase in EEO-1 response rates and/or a secular increase in firm size overtime.

(Insert Figure 2 about here.)

EEO-1 coverage is higher in industrial sectors with larger firms. Manufacturing has the highest coverage rates (65-75 percent over time) and construction the lowest (10-20 percent over time). EEO coverage in services is roughly one-third of that reported by the BLS. Since

industries characterized by small firms and small numbers of employees (i.e., construction, trade services, and agriculture) are not well represented, researchers using these data must be aware of sample selection issues. It is also worth reiterating that the EEOC's definition of "employee" excludes temporary and casual workers "hired for a specified period of time or for the duration of a specified job" which may further explain coverage issues for certain industries (Equal Employment Opportunity Commission, 1981, p. vii).

EEO-1 data for years 1966 through 2000 contain few missing data or out of range values across the 4 million plus records examined. Original reliability of firm reports is unknown. As in any organizational survey, there is likely to be measurement error associated with the method and quality of reporting used by firms. The EEO-1 survey instrument collects information on methods used by firms to assess employment composition, (i.e., visual inspection vs. payroll records). However, this variable is not in the current data files.⁷ The EEOC does not perform external validity checks on reporting. Smith and Welch (1984) provide the only suggestion we can find in the literature of systematic misreporting. They indirectly infer from aggregate comparisons with Current Population Survey data that, in the early years of EEO-1 reports, some firms may have reclassified some professionals to managers in order to appear more sex and ethnically integrated in their managerial ranks. This analysis was not conclusive, focusing on aggregate CPS and EEO-1 data comparison and ignoring alternative explanations such as possible changes in occupational composition resulting from secular shifts in industrial structure or firm composition over time.

We found a few problems in the EEO-1 data worth noting for researchers. In a very few cases (less than 0.1 percent), establishments lacked an assigned Standard Industrial Classification [SIC] code. By comparison, 3.6 percent of cases in the U.S. Current Population Survey have

insufficient industry and occupation data (U.S. Department of Labor, 2002). Although firms are responsible for specifically describing their “major activity,” the EEOC assigns SIC codes in the data. This suggests that, in almost all cases, firms provide enough information for the EEOC to assign industry codes.

A few establishments reported suspiciously high levels of employment although these cases were so rare as to be inconsequential in aggregate analyses. We investigated some of these cases by comparing the name and location of firms to corporate website information on employment. We concluded that, occasionally, consolidated firm reports are miscoded as headquarter or individual establishment reports. This type of miscoding, while rare, might be consequential for more focused studies of specific industries or communities.

Prior to 1990, EEO-1 survey forms did not contain separate counts for white employment. For these years, data analysts must estimate the white counts by totaling the number of non-whites in each occupation (i.e., Asian-American/ Pacific Islanders, African-Americans, American Indians/Alaskan Natives, and Hispanics) and then subtracting this number from grand totals in each occupational category to estimate the number of whites in each occupational category. When we did this we found that in data years prior to 1990, a very small proportion (less than 0.1 percent) of individual establishments contained erroneous grand totals. This could be a result of reporting errors or errors in the data entry process.⁸

The categories used by the EEOC for EEO-1 reports make no definitional distinction between race and ethnicity. The data collection instrument instructs respondents to report all Hispanics, regardless of race, as Hispanic. Thus, it is not possible with these data to treat race and ethnicity as overlapping or multiple identity or heritage categories. This reporting convention may be a source of measurement error, at least under some conceptualizations of race

and ethnicity. On the other hand, reported ethnic distinctions are from the point of view of employers and probably match the socially constructed conceptions of race/ethnicity in these workplaces. In any case, researchers must proceed under the assumption that these race/ethnic categories tend to be mutually exclusive.

There has been some speculation that budget and policy adjustments associated with Reagan-era political shifts may have reduced the quality of data collection and enforcement mandates (Smith & Welch 1984). What we find is that overall, EEOC data are easy to use and have improved over time. There is no obvious decline in either response rates or data coverage during the Reagan-Bush presidency.

The list of contextual variables available in the EEO-1 files grows over time to include more contextual identifiers, such as consistent 2-, 3-, and 5-digit FIPS geographic codes and four digit SIC as well as North American Industrial Classification [NAICS] codes in the 1990s.⁹ Data years 1966 through 2000 contain comparable 2-digit SIC codes.¹⁰ Year 1966 contains the 1957 version of 3-digit SIC coding. Years 1971 through 1973 contain the 1967 version, while years 1973 through 1991 contain the 1972 version of SIC codes. Years 1987 through 1989 and years 1992 through 2000 contain the 1987 version of codes. With the exception of 1997, years 1990 through 1998 contain a “pretest” version of 4-digit SIC codes.¹¹ Year 1998 also contains a “pretest” for NAICS codes. Post 1999, the data are standardized with 2-, 3-, and 4-digit 1987 SIC coding along with 2-, 3-, 4-, 5-, and 6-digit NAICS codes.

Overall, the number of cases and level of detail in EEO-1 data allow for unique analyses of establishment, firm, community, and industry effects on the race and sex composition of employment. Despite changes in technology and political context, EEOC data spanning thirty years contain almost no missing data, and contain few cases with extreme or unusual values. The

data quality seems to be as high as or higher than academic and other national surveys of individuals or firms. Response rates are probably comparable or superior to surveys of individuals or organizations, although this conclusion is based on the 1966 reported response rate (75 percent) and the relative stability of employment coverage over time.

Past Research Using EEO-1 Data

There has been very limited use of EEOC establishment-level data in the past. Those few studies of which we are aware have primarily used these data to measure trends in race and sex composition of occupations or industries (Adams, 1972; Becker, 1980; Cartwright & Edwards, 2002; Leonard, 1984, 1989; Skaggs, 2001; Smith & Welch, 1984). Two researchers have also evaluated the effectiveness of federal equal opportunity monitoring through comparisons of federal contractors, who face more monitoring pressures from the OFCCP, and non-federal contract holders whose only federal oversight is their yearly EEO-1 report to the EEOC (Leonard, 1984, 1989; Smith & Welch, 1984). Smith and Welch (1984) conclude that while minority employment in firms subject to both EEOC reporting and OFCCP monitoring increased, these gains occurred before federal enforcement efforts were in place. Leonard (1984) was more convinced of the effectiveness of OFCCP reporting on black advances in white collar and skilled occupations, particularly after 1978 when OFCCP enforcement increased. However, Leonard concedes the effects of greater enforcement were likely coupled with educational advances among African-Americans and increases in the average establishment size among firms hiring the most minorities. Leonard (1989) argues that contractor compliance programs initially emphasized race over sex in their enforcement regulations. After controlling for establishment size, industry, region, rate of growth, and occupational structure, OFCCP compliance seems to have played a minor role at best in advancing the status of white females (Leonard, 1989). What

is striking about these two studies is not their contradictory conclusions, as much as that in almost forty years of EEOC data collection these are the only published studies of federal enforcement influence.

EEO-1 data have also been useful for examining the institutional context in which women and minorities gain access to managerial jobs (Adams, 1972; Cartwright & Edwards, 2002; Skaggs, 2001). Adams' (1972) early study of 25 southern metropolitan areas made use of 1966 EEO-1 reports to examine the effects of institutional and labor market forces beyond employer's control on occupational distribution. Supplementing EEO-1 data with aggregate Census data, Adams models the effects of education, age, economic growth, demand for skilled labor, industry composition, and market size on occupational distributions for black men and women in 1966. Adams found that an institutional environment shaped by racial discrimination also shaped the racial distribution of occupations and this distribution differed by sex.

EEOC analyses have also provided evidence of industry-specific patterns in race and sex inequality (Cartwright & Edwards, 2002; Skaggs, 2001). Cartwright and Edwards (2002) examine the glass ceiling effect on women and minority managerial occupational attainment across detailed industries. Industry variables add contextual information to the nine broad occupation categories by placing the occupations within establishments in specific industries. They find dramatic sex differences in barriers to management across 164 industries.

Skaggs (2001) uses 1983-1999 EEO-1 data to examine the political, economic, legal, and social contexts affecting minority and female integration into management roles in the retail supermarket industry. Skaggs supplements EEO-1 data with firm-level information on whether or not the firm has had a class action employment discrimination lawsuit and community contextual data on local equal opportunity norms, federal district courts, and competition

between race and sex groups in local labor markets. She found that litigation increased female and minority access to management, most notably in liberal federal courts before and after the Reagan-Bush administrations. Integration progressed more rapidly in regions of the country with stronger pro-equal opportunity attitudes. Integration also progressed rapidly in labor markets with less direct competition for employment between white males and other groups. These analyses demonstrate the flexibility and potential of EEO-1 data for examining organizational change and inequality.

Our conclusion is that while very few researchers have used these data they have great potential. They can be used to investigate workplace trends in segregation as originally designed (Becker, 1980; Cartwright & Edwards, 2002; Leonard, 1984, 1989; Smith & Welch, 1984). They can also be used to investigate the spatial, industrial, and organizational context of segregation (Adams, 1972; Cartwright & Edwards, 2002; Skaggs, 2001). Finally, Adams (1972) and Skaggs (2001) demonstrate the potential for researchers to supplement EEO-1 data with survey and/or contextual data from other sources in order to answer sophisticated research questions.

EEO-1 Segregation Estimates Compared to Occupation-Based Estimates

Since most research on trends in segregation is based on Census and CPS derived occupational estimates (Baunach, 2002; Beller, 1985; Blau, 1988; Carlson, 1992; Jacobs, 1989; Jacobsen, 1994; King, 1992), a comparison of EEO-1 estimates with occupational estimates provides information about the degree to which a lack of information about establishments has distorted prior findings. Figure 3 examines trends in sex segregation comparing EEO-1 estimates of sex segregation with three well-known occupation-based estimates. While the trends are similar in all four sets of estimates, the EEO-1 estimates actually report a more dramatic decline in segregation. It is interesting to note that the levels of estimated segregation are roughly

comparable across the period. This suggests that EEO-1 information about precise establishment locations tends to offset the loss of occupational detail.

(Insert Figure 3 about here.)

Figure 4 compares EEO-1 and Census and CPS occupational estimates (King, 1992) of race segregation trends within sex. King's occupation-level estimates suggest that race segregation declined substantially among women and moderately among men between 1970 and 1980, and both were virtually flat between 1980 and 1988. While female racial occupational segregation was substantially higher in 1970, by 1988, the occupation-level estimates of race segregation within sexes converged to essentially equivalent levels. EEO-1-based estimates indicate a different set of patterns in actual workplaces. While racial segregation among males also declined from 1970 to 1980 and was nearly flat between 1980 and 1988, the 1970 to 1980 decline is less pronounced than in the occupational estimates. The EEO-1 estimates indicate that racial segregation among women actually *increased* slightly between 1970 and 1988. EEO-1 estimates also suggest that racial segregation among men was substantially higher than among women across the same period, with substantial convergence by 1988. Because the central difference between the two data sources is the use of establishment-level information, this comparison indicates that, while black and white women are in increasingly similar occupations over the period 1970-1988, they are employed in increasingly dissimilar firms. The occupation-level picture exaggerates male progress between racial groups in achieving similar male occupational distributions.

(Insert Figure 4 about here.)

Figure 5 compares EEO-1 estimates with King's (1992) occupational estimates of sex segregation trends within race. The contrast in these two sets of estimates is also worth noting.

The occupational estimates suggest, for both blacks and whites, a pronounced decline in sex segregation between 1970 and 1980, followed by an upturn between 1980 and 1988. The EEO-1 estimates suggest a more linear decline in sex segregation across the whole period, although the decline is more dramatic among whites than blacks. The substantially higher estimates of sex segregation among whites compared to blacks suggest that white women and white men are less likely than black women and black men to work in the same workplace. Since the previous figure showed much higher male black/white segregation than female black/white segregation, the two figures strongly suggest that white men tend to work in different establishments than white women, black men, and black women, even when they have similar occupational titles. To our knowledge this is the first time this pattern of workplace segregation has been documented (see Reskin, McBrier, & Kmec, 1999 for speculation that this is the case).

(Insert Figure 5 about here.)

Comparing EEO-1 Trends to Job-Establishment Studies

Some scholars have criticized the use of occupational measures, like those found in EEO-1 data, for segregation analyses (Bielby & Baron, 1986; Tomaskovic-Devey, 1993a and b). They advocate using job-level measures embedded in specific workplaces. Because EEO-1 data allow for analyses of occupations (rather than jobs) within specific workplaces, research using these data will typically underestimate the degree of job segregation. In this section we directly compare EEO-1 estimates to three studies with estimates of sex segregation computed for specific establishments with job-level data.

Bielby and Baron (1986) produced the most specific analysis of the extent to which data with aggregated occupations understate workplace segregation. Using data collected between 1964 and 1979 for 290 establishments in California, Bielby and Baron use the Dictionary of

Occupational Titles (D.O.T.) to reconstruct the EEOC's classification scheme. First, they condense the nine EEOC categories into seven major groups (professional and technical workers, managers and officials, skilled production workers, unskilled and semiskilled production workers, clerical workers, sales workers, and service workers) to make them compatible with the U.S. Employment Service data they used. They then compared segregation distributions using these seven occupational categories with measures based on 10,525 unique job titles distributed across 290 workplaces. Their comparison showed that data using occupational categories within firms underestimate true employment segregation.

We replicate Bielby and Baron's (1986) analysis using EEO-1 reports. Table 2 shows a comparison of sex segregation indices of dissimilarity reported by Bielby and Baron (1986) and 1971 EEO-1 California establishment reports.¹² We computed the segregation statistic (Index of Dissimilarity or D) across the EEOC's nine occupational categories within establishments as well as across the more aggregated seven occupational groups used by Bielby and Baron. The median D 's across the seven occupational categories are comparable between EEO-1 ($D=64.3$) and Bielby and Baron's sample ($D=60.6$). This suggests their small sample is similar in levels of segregation to that of the EEOC's much larger sample. The D statistic increases slightly when computed across the nine EEOC occupation categories ($D=66.7$). Bielby and Baron report a significantly higher D statistic using detailed job titles given by the actual workplaces ($D=96.3$).

(Insert Table 2 about here.)

Values for the distribution of firms across different levels of D indicate that around 1970, 66 percent of California establishments had nearly complete job title segregation ($D \geq 95$). In the EEO-1 reports, approximately 66 percent of observations have values of $D \geq 60$. This suggests that an EEO-1 based estimate of $D \geq 60$ approximates nearly perfect job level segregation. If we

refer back to Figure 1, we see that median gender D 's ≥ 60 characterized the entire U.S. prior to 1974. Similarly, looking at the next 20 percent of cases in the two frequency distributions, D 's ≈ 40 -60 using EEO-1 reports may be comparable to segregation levels of $D \approx 60$ -95 at the job level. Figure 1 indicates that current sex and race/ethnic segregation D values are in the mid 30s range, suggesting real world segregation around $D \approx 50$. This suggests that after thirty years of equal opportunity law and personnel policy, at least half of workers would still have to change jobs to achieve an equal sex distribution of workers across employment contexts. On the other hand, it also suggests that sex segregation in 1964 was nearly complete.

Our second comparison is to the work of Petersen and Morgan (1995). They calculate sex segregation indices of dissimilarity with job-level data for approximately 5,000 establishments in 14 industries. In Table 3, we compare EEO-1 and Petersen and Morgan industry samples in terms of the percentage of females' employed and average firm size within industries, as well as the average D statistic within each industry. The average percentage of female employment across industries in the Petersen and Morgan sample is 44.2 vs. 42.02 percent for the EEO-1 data. The correlation between these two estimates across the 14 industries is .88. Petersen and Morgan's estimate of average establishment size across these industries is 130 employees, while the EEOC estimate is 304, reflecting the higher thresholds for EEO-1 reporting. However, the correlation between size estimates is also quite high ($r = .88$). Hence, the EEO-1 data samples from larger establishments with similar sex compositions and size distributions to those of Petersen and Morgan.

(Insert Table 3 about here.)

The average index of dissimilarity in the Petersen and Morgan set of industries is $D = 81$. The comparable median for the EEO-1 reports is $D = 57$. Their correlation is only .49 across the

fourteen industries. This suggests that EEO-1 occupation-based estimates may not be linear transformations of a job title-based true score. Our investigation of non-linear patterns yield only a slightly better fit for a cubic line ($r^2=.28$) versus a linear estimate ($r^2=.24$). Visually the linear and cubic lines appear identical. Thus the relationship appears linear across the 14 industries, with some fairly large variation within particular industries.

We suspect that industry variation in aggregation bias produced by having occupational groups, rather than job titles, is not random but a reflection of real world divisions of labor. In some industries, job distinctions cover more than one EEO-1 occupational group. Therefore, aggregation bias in estimates of segregation will be low. Other industries have more job distinctions within EEO-1 occupational groups and proportionally less between them. In these industries, EEO-1 estimates of segregation will tend to underestimate actual job level segregation dramatically. This implies that comparisons of segregation within industries are likely to be more accurate than comparisons between industries.

We also look at correlations between dissimilarity indices of sex segregation, average establishment size, and percent female employment within the two samples to see if differences in measures are likely to distort statistical relationships in explanatory analyses. The correlation between the segregation statistic and percent female employment in the industry is $-.45$ in the Petersen and Morgan job-level data and $-.49$ in the EEO-1 data. The correlation of establishment size with sex segregation is $-.35$ for Petersen and Morgan and $-.26$ for the EEO-1 data. These comparisons suggest that both types of data would lead to comparable substantive estimates, although the relationship with firm size might be underestimated in EEO-1 data, a possible artifact of the truncation of firm size in EEO-1 data collection protocols.

We also have access to unpublished estimates of sex segregation for a sample of six warehouse facilities in a single firm over six years during the 1990s. These estimates were based on the firm's employment records for each individual in the firm, which were aggregated to the job title-level within each establishment. One of the establishments was present in the data for only four years, so we can compare thirty-four establishment-year observations for their level of sex segregation measured at the job title level with similar measures using the same establishment's EEO-1 reports. The mean sex segregation computed across job titles is $D = 92.5$, suggesting very high levels of sex segregation. Using the EEO-1 reports the mean is 71.3. Thus although they are lower, the EEO-1 estimates are still very high, consistent with the true score computed across job titles. The correlation between D computed at the job title and D computed from the EEO-1 reports for the same establishments is .58. The employment size correlation is more nearly perfect at .94. Size is strongly negatively related to D for estimates based on employment records and job titles (-.72) and using the EEO-1 reports (-.66).

We also regressed D on employment size and year for both the EEO-1 and employment records estimates. In both cases, the substantive findings were identical and explained variances were nearly identical (R^2 for the job model =.58 and for the EEO-1 model =.56). For these establishments, sex segregation declined over time and was lower in larger establishments. It was the case, however, that coefficients were substantively larger in the EEO-1 data, reflecting the higher variance in D estimates based on EEO-1 reports (Std Dev=9.7) than on actual employment records with job title information (Std Dev =3.8). Again, these comparisons suggest that EEO-1 data will underestimate true segregation, but that substantive conclusions will tend to be accurate relative to what is actually going on in workplaces.

We make two conclusions based on these three comparisons of EEO-1 data estimates to estimates from studies based on job titles. First, EEO-1 reports underestimate true segregation because of occupational aggregation in the survey form, but this underestimate is not systematically tied to other measured (size, percent female, and time) characteristics associated with segregation. Second, we propose that an EEO-1 estimate of segregation, where $D \geq 60$, probably approximates near total job level segregation. Finally, we suspect that industry-specific divisions of labor may mute or exaggerate the degree to which segregation is underestimated using the EEO-1 reports. Understanding when this is happening probably requires substantive knowledge of typical divisions of labor in specific industries. It also implies that for demographic analyses, comparisons of segregation within industries may be more meaningful than comparisons between industries.

Conclusion

EEO-1 data are a powerful source of information for understanding organizational trends and processes that affect workplace sex and race segregation. As our comparisons suggest, these data are the only nationally representative, time series organizational data that exist for the United States. Because these data were originally designed to monitor compliance with Equal Employment Opportunity laws, they will be most powerful for studying workplace segregation trends and processes. They may assist us in answering important but virtually unresearched questions. For example, what is the degree of sex and race segregation between firms? Has it declined over time? If some employers simply refuse to hire women or minorities, how much do they contribute to the level of segregation in the U.S. labor force? Has this changed over time? Has the documented decline in occupational segregation been mirrored by similar declines at the

firm level? Are these patterns similar across industries and local labor markets? If not, what is it about industrial or local labor market context, which determines the trajectory of employment segregation over time? To what extent do declines in segregation reflect declines in segregation within establishments or the replacement of more segregated establishments with less segregated ones over time?

These data might also be used to establish reasonable baseline estimates of the available population of minority or female workers for general occupations (e.g., managers) in specific industries or communities. Such information would be very useful in equal employment opportunity lawsuits, where one of the first questions that must be answered when a firm is charged with discrimination is whether or not the available minority or female labor force is underrepresented in the accused firms' workforce. Similarly, estimates of industry or community segregation can be compared to an accused firm's segregation levels to see if the firm seems unusually high in its level of race or sex segregation.

Because EEO-1 data contain information on Native Americans, Hispanics and Asian-Americans researchers have the potential to move race/ethnic research beyond its traditional emphasis on black-white segregation. The separate trend line for black-Hispanic segregation in Figure 1 shows that EEO-1 reports allow researchers to confront the issue of competition among minority groups for employment roles.

Conversely, these data are not without limitations. They have some of the normal problems associated with any survey-based approach to data collection. First, we do not know precisely how accurately they reflect the actual level of segregation in the reporting firms. We suspect that firms are fairly accurate in their reporting since misreporting to the federal government might be presumed to have serious consequences. Certainly, response quality should

be higher than in surveys of individuals or academic surveys of organizations. On the other hand, there is some speculation that in the late 1960s and early 1970s some companies may have reclassified professional jobs to managerial ones in order to appear more integrated in their managerial ranks (Smith & Welch, 1984). From a sociological point of view, this may not be error so much as an interesting organizational phenomenon to be explored (see Baron, Davis-Blake, & Bielby, 1986; Strang & Baron, 1990). Since it is possible to link workplaces over time, EEO-1 data are suitable for studying changes in organizational divisions of labor at the level of occupations.

Another obvious limitation of these data is the lack of information on very small firms. Where other national data on organizations capture the smallest of establishments (i.e., the SLLS and NSOE), these units are an anomaly in EEO-1 data. Small establishments that are part of larger firms are sometimes included in special and consolidated reports, but most are not counted. This means that EEO-1 based estimates miss segregation trends in smaller independent firms and small establishments in larger firms if they are different from those in larger establishments and larger firms.

The most obvious limitation of these data is that sex/race distributions are aggregated into occupational groups, not the actual jobs used to organize work and its rewards. There is a large substantive difference between a plant manager and an acting foreman. Our comparisons to Bielby and Baron (1986), Petersen and Morgan (1995), and a small sample of establishments, show that when using EEO-1 estimates we would expect to underestimate segregation relative to its true level. These underestimates seem to be at about the same level as underestimates based on the most detailed occupational data available from Census and CPS sources. Our comparisons suggest that these underestimates are not systematically related to the true score, nor do they

distort associations with important covariates like firm size or status composition. On the other hand, we see large residual variation across industry in estimates of D based on job-title and occupational aggregations within establishments. We suspect this reflects variations in real industry-based divisions of labor. This suggests that these data are more powerful for segregation trend comparisons across industries and for segregation level comparisons within industries. Use of these data will clearly be enhanced by institutional knowledge of industry-specific divisions of labor and employment practices (see especially Skaggs, 2001).

It is important that EEO-1 data lead to conclusions about segregation trends that differ from those reached by previous occupation-level research. While EEO-1 based estimates agree with the general conclusion that black-white and male-female segregation declined since 1966, they disagree strongly as to when and for whom. Our comparison with King (1992) is especially instructive here. In comparing estimates of race segregation within sexes, the occupation-level data show a dramatic decline in segregation between black and white women and a gradual decline between black and white men. EEO-1 data show a quite different pattern, a slight increase in segregation among women and a more gradual decline in race segregation among men. This suggests that black and white women may work in the same occupations, but they increasingly work in different establishments. This comparison also suggests that the estimated decline in race segregation among men may be exaggerated in occupational data. In comparing sex segregation within races, the occupational data show a dramatic decline from 1960-1980 and a substantial increase from 1980-1988 while EEO-1 data show a more gradual decrease over the entire period. This finding, when considered with the comparisons in race segregation within sexes, makes it clear that white males often work in different establishments than white females, black females, or black males. This is certainly a promising focus for future research.

There is great potential in the use of EEO-1 data on its own, but we further our understanding of social mechanisms if we use such data in combination with other data that provide more information on social context (Kmec, 2003b). Skaggs (2001) uses the EEO-1 reports to investigate the influence of lawsuits, federal circuit courts, community norms, and status competition processes upon female, black, and Hispanic access to managerial jobs. These are the kind of mechanisms that Reskin (2003) encourages inequality researchers to address. The survey work of Kalev, Dobbin, and Kelly (2003) represents another opportunity for expanding the intellectual reach of EEO-1 reports. They use EEO-1 reports to generate a sampling frame of organizations, survey those organizations, and combine the survey responses with EEO-1 trend analyses of managerial integration. Similarly, one could imagine a retrospective matching of EEO-1 reports to existing organizational surveys such as the North Carolina Business Needs Survey (Tomaskovic-Devey & Skaggs, 1999a and b), the NOS (Kalleberg, Knoke, Marsden, & Spaeth, 1996), or the MCSUI (Holzer, 1996; Kmec 2003b). We recently gained access to North Carolina Employment Security establishment records of employment and payroll levels. Matching these data with EEO-1 reports could be used to study the influence of segregation and composition change on wage rates. One could also imagine matching EEO-1 reports to individual earnings and job records in cases where workplace identifiers were present such as the NOS-General Social Survey file (Kalleberg et al., 1996) or Lincoln and Kalleberg's (1985) sample of 8000 workers in 100 US and Japanese manufacturing plants. Recent work matching census and firm data could also be supplemented with EEO-1 Reports (Bayard, Hellerstein, Neumark, & Troske, 1999; Carrington & Troske, 1998).

Combining EEO-1 reports with other establishment and individual-level data could advance our knowledge of the mechanisms of inequality and be quite useful to the EEOC's

efforts to enforce anti-discrimination laws. No one method of analysis or source of data can provide the total picture. We believe that EEO-1 records are a valuable addition to the social science data arsenal for research both on equal opportunity issues and more general organizational dynamics.

Footnotes

¹ As a special federal employee, the senior project member becomes liable for any unauthorized release of the data. EEOC's Office of Program, Research and Information has entered into a number of such agreements to foster the development of research in this policy area. Questions regarding this effort can be directed to Ron Edwards (ronald.edwards@eoc.gov).

² Firms determine the methods used to tally and classify employees within occupations. They have the option of performing a visual inspection, referring to payroll records, or using some other procedure to generate the employment distributions reported to the EEOC.

³ Federal contractors are defined as firms that (1) have a federal contract or first-tier subcontract worth \$50,000 or more; or (2) act as depositories of federal funds in any amount, or (3) act as issuing and paying agents for U.S. Savings Bonds and Notes (from “Data Notes,” EEOC, 1999).

⁴ We use the index of dissimilarity (D) in this paper because it allows us to compare EEO-1 estimates to previous work. In addition, it is easy to compute and has a simple interpretation as the proportion of workers who would have to switch jobs to create full integration. The index of dissimilarity is computed across occupations within establishments as follows:

$$D = \frac{1}{2} \sum_{oe=1}^{N_{oe}} | P_{oe1x} - P_{oe1y} |$$

where P_{oe1x} and P_{oe1y} are the proportions of group x and y, respectively, within an occupation in an establishment. EEO-1 data can also be used to compute any of the two and multiple group segregation measures available in the literature (see, for example, James & Tauber, 1985; Reardon & Firebaugh, 2002) and can be used in log linear modeling approaches to describe

segregation (Charles & Grusky, 1995). They can also be used to compute simple percent distributions of groups in occupations, as well as odds ratios of group representation between any two occupations or groups of occupations.

These represent all available reports from all years since 1966, which are currently available in machine-readable form from the EEOC. Data were not available for years 1967-1970, 1974, and 1976. Because of confidentiality requirements, no public use data on establishments, firms, or even industries within local geographic areas can be released.

⁵ Reporting requirements excluded some smaller firms from reporting in 1983. Hispanics tend to work in smaller workplaces than African Americans. The shift in reporting requirements probably produced the dip in Black-Hispanic segregation in 1983.

⁶ It is the case, however, that many firms file establishment reports for workplaces with less than fifty employees and these observations appear among establishment reports.

⁷ Based on our discussions with representatives at the EEOC, no one recalls ever using this item; therefore analysis files were standardized such that this item was deleted from all years of data. We have begun discussions to see if this item might be added back into analysis files, since it is the only indicator we have of potential data quality as well as the organizational capacity to monitor equal opportunity goals.

⁸ We also computed totals for race and sex within and across occupational categories and compared these to the totals given in the datasets to see if these matched. As expected, there

were some discrepancies in early years of data prior to 1990. After 1993, the data are virtually free of this problem. This suggests that EEOC data cleaning procedures improved in the early 1990s.

⁹ The 1966 EEO-1 data file contained FIPS codes unique to the EEOC. We prepared SAS code to standardize these codes over time. This file is available at:

http://sasw.chass.ncsu.edu/s&a/faculty/tomaskovic-devey_d/don_tomaskovic_documenting_desegregation.html.

¹⁰ 1997 is an exception. Data in this year do not contain a two-digit SIC variable. Using the SAS substring command, it can be created from the 3-digit SIC code (For example, `SIC2=substr(sic3,1,2,))`. SIC codes changed in years 1959, 1972, 1977, 1987. The North American Industrial Classification System [NAICS] has currently replaced the SIC system.

¹¹ We have examined these pretest codes and have found that in early years they tend to be unreliable, slowly converging to reliable codes. We recommend that researchers use pretest codes with extreme caution.

¹² Bielby and Baron's (1996) data combine years 1964 through 1979. Unfortunately, we were unable to ascertain how these years of data were combined.

References

- Acker, J. (1990). Hierarchies, Jobs, Bodies: A Theory of Gendered Organizations. *Gender & Society* 4,139-158.
- Adams, A. (1972). Black-White Occupational Differentials in Southern Metropolitan Employment. *Journal of Human Resources*, 7, 500-517.
- Baron, J. N. & Bielby, W.T. (1980). Bringing the Firm Back In: Stratification, Segmentation, and the Organization of Work. *American Sociological Review*, 45, 737-65.
- Baron, J. N. & Bielby, W.T. (1984). The Organization of Work in a Segmented Economy. *American Sociological Review*, 49, 454-473.
- Baron, J. N. & Bielby, W.T. (1986). The Proliferation of Job Titles in Organizations. *Administrative Science Quarterly*, 31, 561-586.
- Baron, J.N., Davis-Blake, A. & Bielby, W.T. (1986). The Structure of Opportunity: How Promotion Ladders Vary Within and Among Organizations. *Administrative Science Quarterly*, 31, 248-273.
- Baunach, D.M. (2002). Trends in Occupational Sex Segregation and Inequality, 1950 to 1990. *Social Science Research*, 31,77-98.
- Bayard, K., Hellerstein, J., Neumark, D. & Troske, K.R. (1999). New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data. National Bureau of Economic Research. Forthcoming.
- Becker, H.J. (1980). Racial Segregation among Places of Employment. *Social Forces*, 58, 761-776.
- Beller, A.H. (1985). Changes in the Sex Composition of U.S. Occupations. *Journal of Human Resources*, 20, 235-250.

- Bielby, W.T. & Baron, J.N. (1986). Men and Women at Work: Sex Segregation and Statistical Discrimination. *American Journal of Sociology*, 91, 759-799.
- Blau, F.D. (1988) *Occupational Segregation by Gender: A Look at the 1980s*. Paper presented at the annual meeting of the American Economic Association. New York.
- Bridges, W.P. & Nelson, R.L. (1989). Markets in Hierarchies: Organizational and Market Influences on Gender Inequality in a State Pay System. *American Journal of Sociology*, 95, 616-59.
- Carlson, S.M. (1992). Trends in Race/Sex Occupational Inequality: Conceptual and Measurement Issues. *Social Problems* 19, 268-285.
- Carrington, W.J. & Troske, K.R. (1998). Interfirm Segregation and the Black/White Wage Gap. *Journal of Labor Economics*, 16, 231-260.
- Cartwright, B. & Edwards, P.R. (2002). Using EEO-1 Data to Examine Gender-Based Occupation Segregation by Industry.
<https://www.eeoc.gov/stats/reports/glassceiling/index.html>
- Charles, M. & Grusky, D.B. (1995). Models for Describing the Underlying Structure of Sex Segregation. *American Journal of Sociology*, 100, 931-71.
- Cockburn, C. (1991). *In the Way of Women: Men's Resistance to Sex Equality in Organizations*. Ithaca, NY: ILR Press.
- Cohen, P.N. & Huffman, M.L. (2003). Occupational Segregation and the Devaluation of Women's Work Across U.S. Labor Markets. *Social Forces*, 81, 881-890.
- Collins, P.H. (1997). African-American Women and Economic Justice: A Preliminary Analysis of Wealth, Family and African-American Social Class. *University of Cincinnati Law Review*, 65(3), 825-852.

- Cotter, D.A., DeFiore, J.M., Hermesen, J.M., Kowalewski, B.M., Vanneman, R. (1995). Occupational Sex Desegregation in the 1980s. *Work and Occupations*, 22, 3-21.
- England, P. Farkas, G., Kilbourne, B.S. & Dou, T. (1988). Explaining Occupational Sex Segregation and Wages: Findings From a Model with Fixed Effects. *American Sociological Review*, 53, 544-558.
- Equal Employment Opportunity Commission. (1967). *Job Patterns for Minorities and Women in Private Industry, 1967*. U.S.GPO:DC.
- Equal Employment Opportunity Commission. (1981). *Job Patterns for Minorities and Women in Private Industry, 1980*. U.S.GPO:DC.
- Equal Employment Opportunity Commission. (1999). *Job Patterns for Minorities and Women in Private Industry, 1999*. U.S.GPO:DC.
- Goux, D. & Maurin, E. (1997). Meritocracy and Social Heredity in France: Some Aspects and Trends. *European Sociological Review*, 13, 159-177.
- Groshen, E.L. 1991. The Structure of Female/Male Wage Differential: Is it Who You Are, What You Do, or Where You Work? *Journal of Human Resources*, 26, 457-72.
- Holzer, H.J. (1996). *What Employers Want: Job Prospects For Less-Educated Workers*. New York: Russell Sage Foundation.
- Hultin, M. & Szulkin, R. (1999). Wages and Unequal Access to Organizational Power: An Empirical Test of Gender Discrimination. *Administrative Science Quarterly* 44 (3), 453-472.
- Jacobs, J.A. (1989). Long-Term Trends in Occupational Segregation by Sex. *American Journal of Sociology*, 95, 160-73.
- Jacobsen, J. P. (1994). Sex Segregation at Work: Trends and Predictions. *The Social Science*

- Journal*, 31, 153-169.
- James, D.R. & Tauber, K.E. (1985). Measures of Segregation. *Sociological Methodology*, 14, 1-32.
- Kalleberg, A.L., Knoke, D., Marsden, P. & Spaeth, J. (1996). *Organizations in America*. Thousand Oaks, CA: Sage.
- Kalleberg, A.L. & Lincoln, J.R. (1988). The Structure of Earnings Inequality in the United States and Japan. *American Journal of Sociology*, 94, 121-153.
- Kalleberg, A.L. & Mastekaasa, A. (1998). Organizational Size, Layoffs, and Quits in Norway. *Social Forces*, 76(4), 1243-1273.
- Kalev, A., Dobbin, F. & Kelly, E. (2003). Is Affirmative Action Obsolete?: Anti-Discrimination Measures and the Entrance of Women and African-Americans into Management. Manuscript, Department of Sociology, Harvard University.
- King, M.C. (1992). Occupational Segregation by Race and Sex, 1940-88. *Monthly Labor Review*, 115, 30-36.
- Kmec, J.A. (2003a). Minority Job Concentration and Wages. *Social Problems* 50(1), 38-59.
- Kmec, J.A. (2003b). Collecting and Using Employer-Worker Matched Data. *Sociological Focus*, 36(1), 81-96.
- Le Grand, C., Hedstrom, P., Szulkin, R. & Tahlin, M. (1990). Samspelet mellan individ och struktur pa den svenska arbetsmarknaden. Paper presented at the NAUT conference. Oslo. March 1990.
- Leonard, J.S. (1984). Employment and Occupational Advance Under Affirmative Action. *The Review of Economics and Statistics*, 66, 377-85.

- Leonard, J.S. (1989). Women and Affirmative Action. *The Journal of Economic Perspectives*, 3(1), 61-75.
- Lincoln, J. R. & Kalleberg, A.L. (1985). Work Organization and Workforce Commitment: A Study of Plants and Employees in the U.S. and Japan. *American Sociological Review*, 50, 738-760.
- Mastekaasa, A. (1992). Organizational Contexts and Individual Behavior: Potential and Limitations of the Norwegian Study of Organizations. *Acta Sociologica*, 35, 141-150.
- McIlwee, J. & Robinson, J.G. (1992). *Women in Engineering: Gender, Power and the Workplace*, Albany, NY: State University of New York Press.
- Nelson, R.L. & Bridges, W.P. (1999). *Legalizing Gender Inequality: Courts, Markets and Unequal Pay for Women in America*, Cambridge University Press.
- Padavic I. (1991). Attraction of Male Blue Collar Jobs for Black and White Women: Economic Need, Exposure and Attitudes. *Social Science Quarterly*, 72, 33-49.
- Petersen, T. & Morgan, L.A. (1995). Separate and Unequal: Occupation-Establishment Sex Segregation and the Gender Wage Gap. *American Journal of Sociology*, 101, 329-365.
- Petersen, T., Saporta, I. & Marc Seidel, D.L. (2000). Offering a Job: Meritocracy and Social Networks. *American Journal of Sociology*, 106, 763-816.
- Reardon, S. F. & Firebaugh, G. (2002). Measures of Multigroup Segregation. *Sociological Methodology*, 32, 33-101.
- Reskin, B. F. (2003). Including Mechanisms in Our Models of Ascriptive Inequality. *American Sociological Review*, 68, 1-21.
- Reskin, B.F., McBrier, D., & Kmec, J.A. (1999). The Determinants and Consequences of Workplace Sex and Race Composition. *Annual Review of Sociology*, 25, 335-61.

- Reskin, B.F. & Padavic, I. (1988). Supervisors as Gatekeepers: Male Supervisors' Response to Women's Integration in Plant Jobs. *Social Problems*, 35, 536-550.
- Skaggs, S. (2001). *Discriminating Litigation: Implications for Women and Minorities IN Retail Supermarket Management*. Doctoral dissertation, North Carolina State University.
- Smith, R.A. (2002). Race, Gender, and Authority in the Workplace: Theory and Research. *Annual Review of Sociology*, 28, 509-542.
- Smith, J.P. & Welch, F. (1984). Affirmative Action and Labor Markets. *Journal of Labor Economics*, 2, 269-301.
- Strang, D. & Baron, J.N. (1990). Categorical Imperatives: The Structure of Job Titles in California State Agencies. *American Sociological Review*, 55, 479-495.
- Tam, T. (1998). Sex Segregation and Occupational Gender Inequality in the United States: Devaluation or Specialized Training? *American Journal of Sociology*, 102, 1652-92.
- Tomaskovic-Devey, D. (1993a). *Gender & Racial Inequality at Work: The Sources and Consequences of Job Segregation*. Ithaca, NY: ILR Press.
- Tomaskovic-Devey, D. (1993b). The Gender and Race Composition of Jobs and the Male/Female, White/Black Pay Gaps. *Social Forces*, 72, 45-76.
- Tomaskovic-Devey, D., Kalleberg, A.L. & Marsden, P.V. (1996). Organizational Patterns of Gender Segregation. In A.L. Kalleberg, D. Knoke, P.V. Marsden, and J.L. Spaeth (Eds.), *Organizations in America: Analyzing Their Structures and Human Resource Practices*. (Pp. 276-301). Thousand Oaks, Ca: Sage Publications.
- Tomaskovic-Devey, D. & Skaggs, S. (1999a). An Establishment-Level Test of the Statistical Discrimination Hypothesis. *Work and Occupations* 26, 422-445.

Tomaskovic-Devey, D. & Skaggs, S. (1999b). Degendered Jobs?: Organizational Processes and Gender Segregated Employment. *Research in Social Stratification and Mobility*, 17, 139-172.

Tomaskovic-Devey, D. & Skaggs, S. (2002). Sex Segregation, Labor Process Organization and Gender Earnings Inequality. *American Journal of Sociology*, 108(1), 102-128.

Tomaskovic-Devey, D., Zimmer, C. & Harding, S. (2003). What Do Markets Really Look Like?: Evaluating Neoclassical, Embeddedness and Transaction Costs Images. Paper presented at the 2003 American Sociological Association meetings.

U.S. Department of Labor, Bureau of Labor Statistics and U.S. Census Bureau. (2002). Current Population Survey: Design and Methodology, Technical Paper 63RV.
www.census.gov/prod/2002pubs/tp63RV.pdf

Villemez, W.J. & Bridges, W.P. (1988). When Bigger is Better: Differences in the Individual Level Effect of Firm and Establishment Size. *American Sociological Review*, 53, 237-255.

Table 1. Summary Advantages and Disadvantages of Available Data Collection Strategies for Studying Workplace Status Inequality and Segregation

Data Characteristics	Occupational Surveys of Individuals^a	Firm Personnel Records^b	Other Surveys of Organizations^c	EEO-1 Establishment Surveys
National?	Yes	No	Generally	Yes
Spatial Analysis?	Limited	No	Rarely	Yes
Longitudinal-Trend or Panel?	Trend for Survey Years	Rarely	Rarely	Panel & Trend 1966-Present
Employee Characteristics?	Yes	Limited	No	No
Organizational Characteristics?	No	Some	Some	Some
Job Characteristics?	No	Yes	Some	Some
Measure of Within Establishment Segregation?	No	Yes	Yes	Yes
Measure of Between Establishment Segregation?	No	Yes	Limited	Yes
Who Reports?	Employee	Employer	Employer	Employer

^aOccupational surveys of individuals include the U.S. Census data & Current Population Surveys [CPS] (Baunach, 2002; Beller, 1985; Blau, 1988; Carlson, 1992; Jacobs, 1989; Jacobsen, 1994; King, 1992), the French Education and Qualifications Survey [FQP] (Goux & Maurin, 1997), Sweden's 1991 Level of Living Survey [SLLS] (Hultin & Szulkin, 1999), the North Carolina Employment Survey [NCES] (Tomaskovic-Devey, 1993a and b), and the Multi-City Study of Urban Inequality [MCSUI] (Kmec, 2003b).

^bFirm-level/ personnel record data include Baron and Bielby (1980, 1984, 1986), and Peterson and Morgan (1995).

^cOther surveys of organizations include the National Organization Survey [NOS] (Kalleberg et al., 1996, the Australian National Organization Survey [AUSNOS] (Tomaskovic-Devey, Zimmer, & Harding, 2003), Sweden's 1991 Establishment Survey [SES] (Hultin & Szulkin, 1999), Norway's 1989 and 1993 Survey of Organizations and Employees [NSOE] (Kalleberg & Mastekaasa; 1998; Mastekaasa, 1992), and the Multi-City Telephone Employer Survey [MCETS] (Kmec, 2003b).

Table 2. Estimated Sex Segregation in Employment using Bielby and Baron (1986) and 1971 EEO-1 Reports^a

Dissimilarity Index Category	% of Establishments			
	EEO-1 9 Occupations	EEO-1 7 Occupations ^b	B & B 7 Occupations	B & B Across 10,525 Job Titles
<i>D</i> lt 20	8.4	9.32	8	0
20 to 39	15.29	17.71	20	2
40 to 59	19.88	19.24	21	2
60 to 79	19.85	18.5	13	8
80 to 89	12.64	12.02	10	8
90 to 94	7.54	7.26	7	6
95 to 99	9.53	9.01	5	16
<i>D</i> = 100	7.04	7.11	9	50
Median <i>D</i>	66.7	64.3	60.6	96.3
Total Establishments	12,329	12,329	290	290
Interquartile Range	47.9	51.2	57.6	5.9

Sources: Bielby and Baron (1986), Table 4 and our computations from 1971 EEO-1 data.

^aBielby and Baron combined observations for years 1964 to 1979. We use 1971 for comparison because it is the midpoint of 1964-1979.

^bOccupational categories in 1971 EEO-1 data were condensed from 9 to 7 groups to be consistent with Bielby and Baron (1986) data. Operatives and laborers were combined to form one category. Professionals and technicians were also combined.

**Table 3. Comparisons of Establishment-Level Descriptive Statistics and Dissimilarity Indices [D]
for Peterson and Morgan (1995) Industry Wage Survey [IWS] Data and EEO-1 Data**

P & M IWS Data (1)							EEO-1				
SIC (2)	Industry (3)	Yr (4)	% Female	Total Empl.	Total Establ.	D _{oe} (5)	Yr (6)	% Female	Total Empl.	Total Establ.	D _{oe} (7)
232	Men's and boys' shirts	1978	92.5	40,068	220	.949	1978	81.7	314,639	1,243	.567
806	Hospitals	1981	84.8	151,187	250	.640	1981	80	2,441,697	3,466	.379
614	Banking	1980	82.8	74,481	579	.686	1980	54.2	31,860	227	.623
631	Life insurance	1980	75.9	30,976	221	.792	1980	55.5	398,767	3,845	.667
235	Wool textiles	1975	55.4	5,087	57	.874	1975	77.9	5,240	26	.485
226	Cotton & synthetic fiber textiles	1975	54.6	86,327	340	.824	1975	28	44,206	164	.51
701	Hotels and motels	1978	51.5	80,204	629	.820	1978	50.9	348,803	1,834	.257
737	Computer & data processing services	1982	44.5	14,520	355	.626	1982	35.1	318,825	1,250	.493
251	Wood household furniture	1974	35.9	36,140	332	.658	1975	31.2	187,821	828	.368
356	Machinery	1983	15.5	54,873	796	.859	1983	23.5	189,496	671	.645
336	Nonferrous foundries	1975	9.4	18,218	364	.922	1975	20.2	53,238	261	.585
285	Paints and varnishes	1976	5.6	10,939	292	.895	1975	22.1	36,982	264	.778
281	Industrial chemicals	1976	2.5	71,400	266	.866	1975	14.6	321,950	888	.849
344	Fabricated structural steel	1974	.6	23,231	332	.970	1975	13.4	255,906	1,293	.84
	Unweighted average across industries		44.23	768,006	5,909	.814		42.02	4,949,430	16,260	0.57

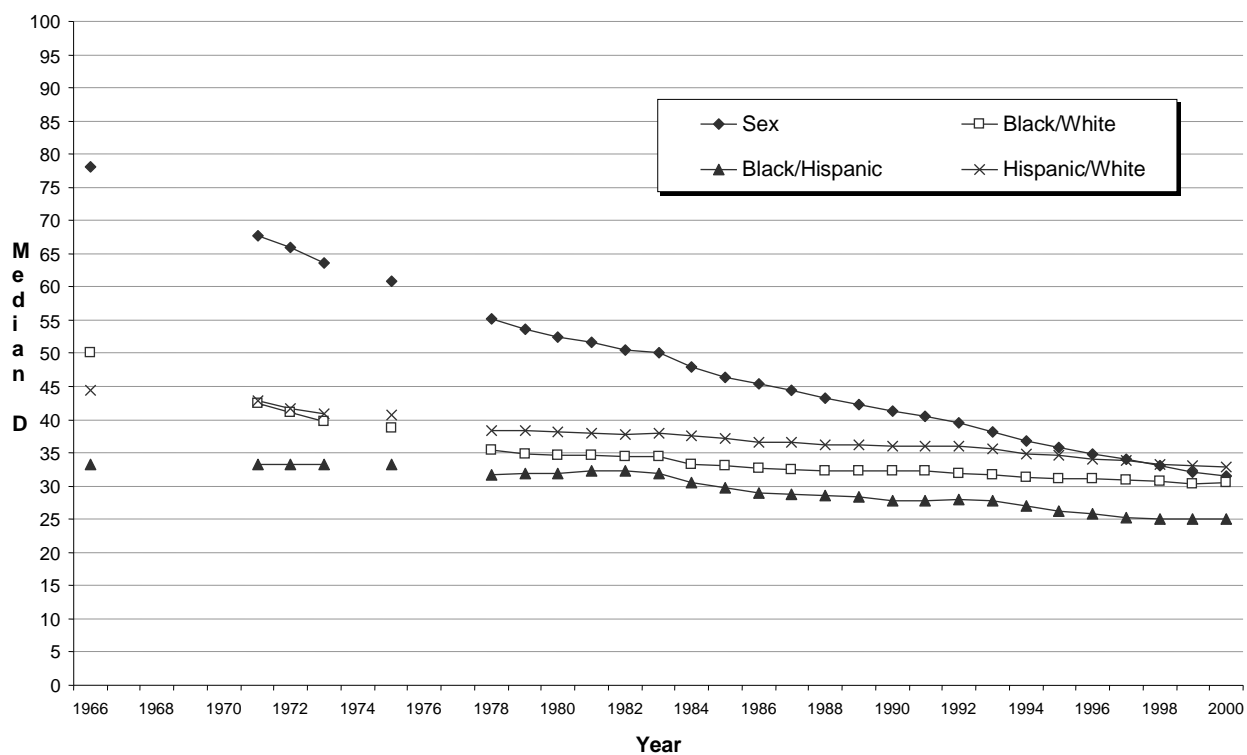
(1) *Sources:* Peterson and Morgan (1995), Table 1 & Table 4, Pp.333 & 350 and our computations from various years of EEO-1 data files.

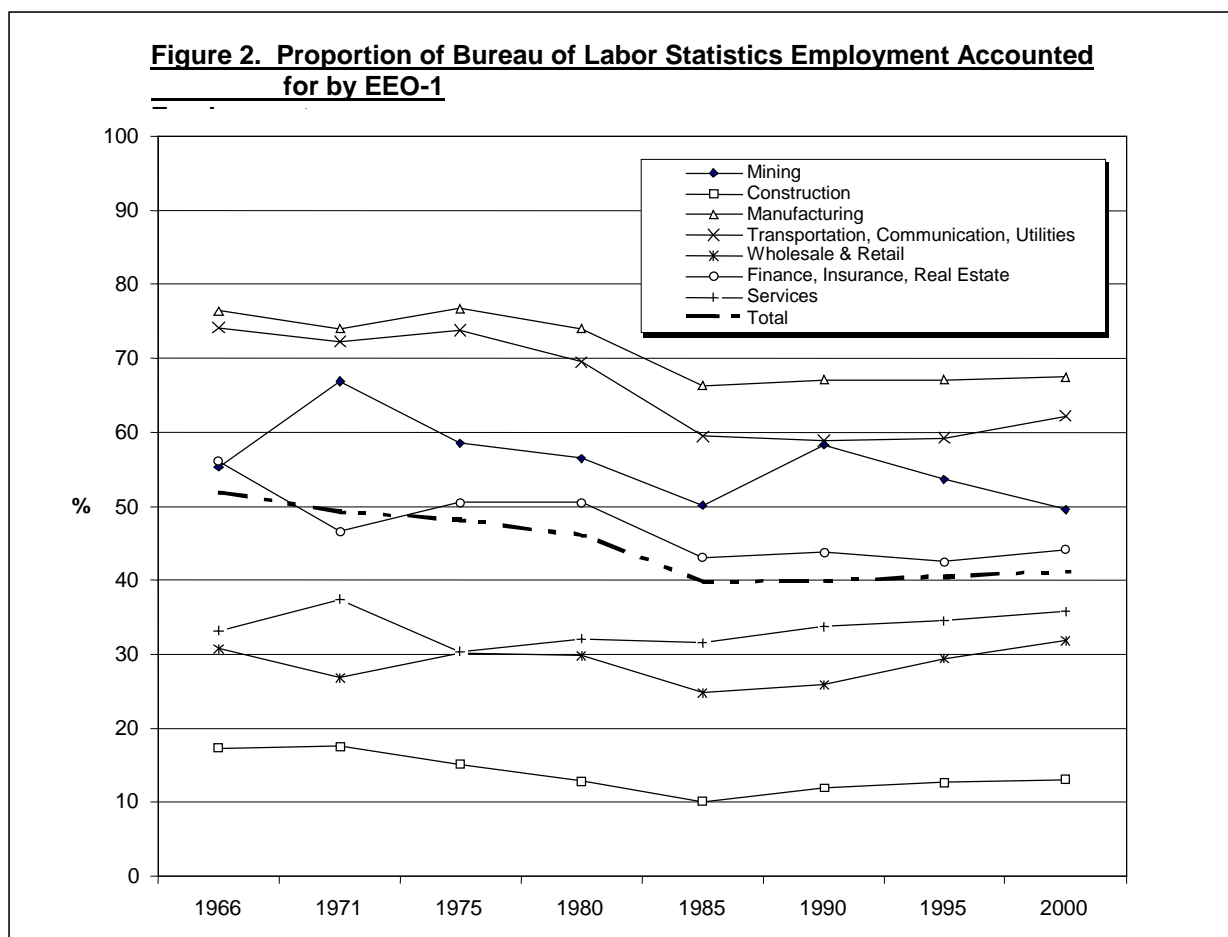
(2) Since Peterson and Morgan (1995) do not provide specific SIC codes in their analysis, we only show comparisons for those industries that we are sure match with their respective SIC codes.

(3) These industry labels are taken from Peterson and Morgan (1995). Actual SIC labels have been revised since their publication.

- (4) Years correspond to survey years for 16 industries chosen by the U.S. Bureau of Labor Statistics in the Department of Labor for the Industry Wage Survey [IWS] (Peterson and Morgan 1995).
- (5) Occupation-establishment segregation indices based on Duncan and Duncan formula.
- (6) EEOC data are unavailable for years 1974 and 1976. Data from 1975 were used for comparison.
- (7) Median D's are shown.

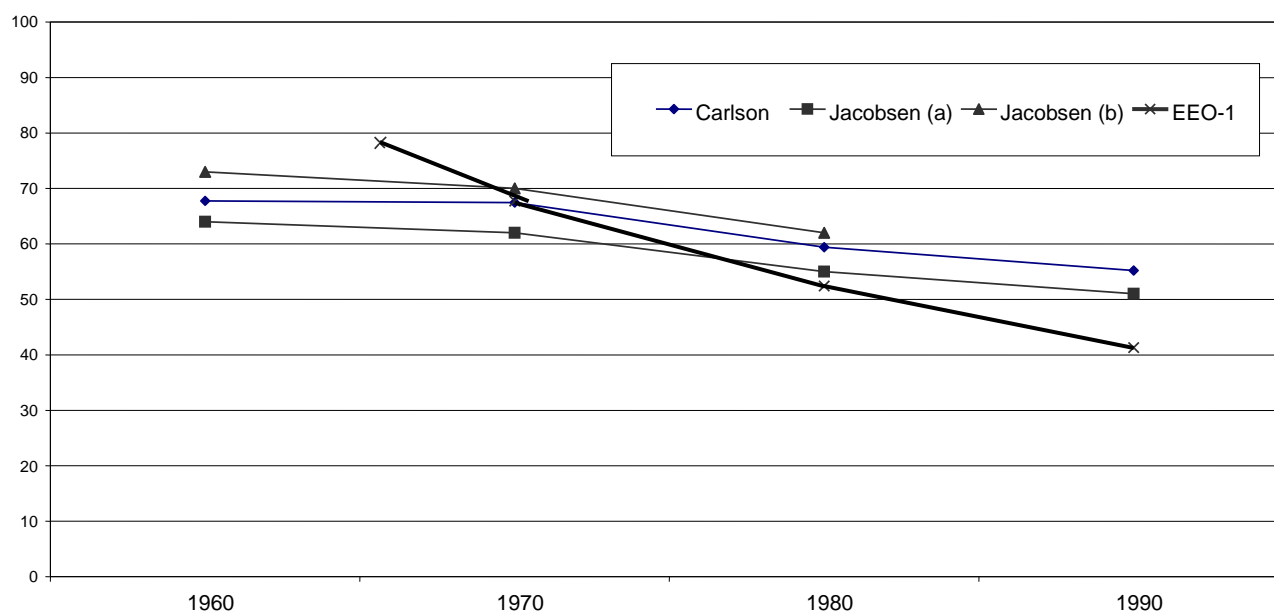
Figure 1. EEO-1 Establishment-level Dissimilarity Indices, 1966-2000





Sources: U.S. Equal Employment Opportunity Commission Annual Report (various years) and U.S. Department of Labor, Bureau of Labor Statistics, Employment and Earnings (various years).

**Figure 3. Aggregate Level Trends in Occupational Segregation by Sex:
Comparison of EEO-1, Carlson (1992), and Jacobsen (1994) Estimates^a**



^aCarlson (1992) uses 360 comparable occupation codes. Jacobsen (1994) (a) uses 238 occupation codes and (b) 12,850 occupations within industries. These studies are occupational-level whereas EEO-1 are establishment-level Ds based on 9 occupational categories within establishments.

Figure 4. Race Segregation Trends within Sex: Comparison of King (1992) CPS Occupational Data and EEO-1 Occupation-Establishment Data

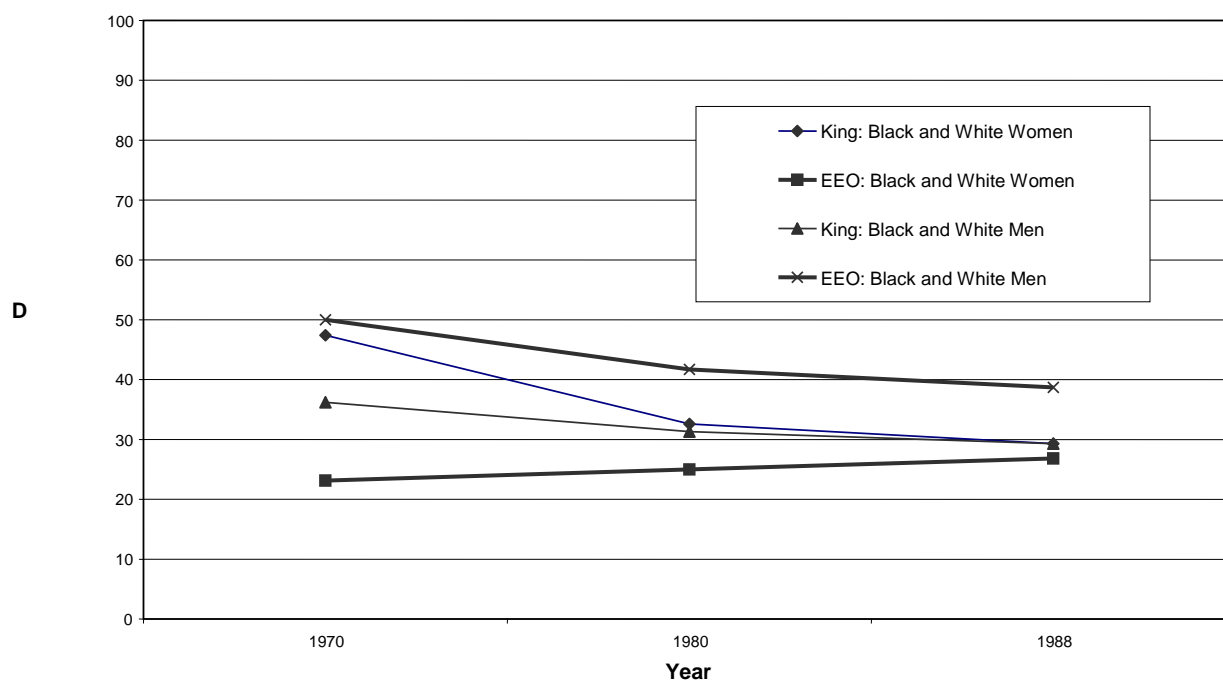


Figure 5. Sex Segregation Trends within Race: Comparison of King (1992) CPS Occupational Data and EEO-1 Occupation-Establishment Data

