

Gender, Race, and Meritocracy in Organizational Careers¹

Emilio J. Castilla
Massachusetts Institute of Technology

This study helps to fill a significant gap in the literature on organizations and inequality by investigating the central role of merit-based reward systems in shaping gender and racial disparities in wages and promotions. The author develops and tests a set of propositions isolating processes of *performance-reward bias*, whereby women and minorities receive less compensation than white men with equal scores on performance evaluations. Using personnel data from a large service organization, the author empirically establishes the existence of this bias and shows that gender, race, and nationality differences continue to affect salary growth after performance ratings are taken into account, *ceteris paribus*. This finding demonstrates a critical challenge faced by the many contemporary employers who adopt merit-based practices and policies. Although these policies are often adopted in the hope of motivating employees and ensuring meritocracy, policies with limited transparency and accountability can actually increase ascriptive bias and reduce equity in the workplace.

An extensive body of research on organizations and stratification has established that organizations play a key role in generating and perpetuating inequality in employment outcomes (Baron and Bielby 1980; Baron 1984; Bielby and Baron 1986; Reskin 1993; Phillips 2005). To date, the

¹ I am grateful for the financial support received from the Center for Human Resources and the Center for Leadership and Change Management at the Wharton School, University of Pennsylvania. Earlier versions of this article were presented at the 2005 American Sociological Association annual conference in Philadelphia, the 2005 Academy of Management annual meeting in Honolulu, the Eleventh Organizational Behavior conference at Wharton, and the Center for the Study of the Economy and Society at Cornell University. I thank Roberto Fernandez, Ezra Zuckerman, Jesper Sorensen, and Mauro F. Guillen for their help. I have also benefited from the extensive and detailed comments made by Lucio Baccaro, Lotte Bailyn, Diane Burton, Jesús M. De Miguel, Isabel Fernández-Mateo, Kate Kellogg, Anne Marie Knott, Thomas Kochan, Denise Loyd, Mark Mortensen, Paul Osterman, Michael Piore, Nancy Roth-

research has mainly focused on identifying and testing the particular mechanisms that account for gender and racial inequality in wages and career attainment within organizations. Recently, Petersen and Saporta (2004) provided a useful and compelling framework for organizing our thinking around three main processes that lead to employer discrimination. Many empirical studies have looked at the first process, which they call *allocative discrimination*, whereby women and minorities are sorted into different kinds of jobs with different pay, whether through hiring, promotion, or termination (e.g., Rosenfeld 1992; Marsden 1994a, 1994b; Baldi and McBrier 1997; Barnett, Baron, and Stuart 2000; Petersen, Saporta, and Seidel 2000; Elvira and Zatzick 2002; Petersen and Saporta 2004; Fernandez and Sosa 2005; Fernandez and Fernandez-Mateo 2006). Other studies have examined *within-job wage disparities*, whereby women and minorities receive lower salaries than their white male counterparts within a given occupation and establishment (e.g., Jacobs 1989, 1995; England 1992; Tomaskovic-Devey 1993; Petersen and Morgan 1995). Finally, some empirical work has explored the third process, identified as *valuative discrimination*, in which female- and minority-dominated occupations with equal skill requirements and other wage-relevant factors are paid lower salaries because they are valued less (e.g., Bridges and Nelson 1989; Baron and Newman 1990; for a review, see England [1992] and Nelson and Bridges [1999]).

Despite substantial progress on clarifying the mechanisms that shape discrimination, researchers have paid less attention to current employer practices that might counteract such discrimination and remediate workplace inequality (for a recent step in this direction, see Kalev, Dobbin, and Kelly [2006]). One approach that has gained considerable popularity is the use of merit-based practices in organizations. Under the old employment system, lifetime jobs with predictable career advancement and stable pay were virtually guaranteed. Pay raises were given on the basis of seniority or granted automatically to all employees at the same per-

bard, Sean Safford, Christopher Wheat, Steffanie Wilk, JoAnne Yates, Mark Zbaracki, and the *AJS* reviewers. I owe many thanks to Robin Kietlinski, Aneri H. Jambusaria, and Anjani Trivedi for their invaluable research assistance. I also want to thank my colleagues at the MIT Sloan School and at the Wharton School, especially Peter Cappelli, Sarah Kaplan, Gerald McDermott, Lori Rosenkopf, Daniel Levinthal, John Paul Macduffie, and Katherine Klein. Several members of the organization I study also provided great help during the data collection process. I want to thank the people who held the following positions at that time: the director and the vice president of human resources, the head of the research unit within the human resources department, and several midlevel human resource managers. The views expressed here are exclusively my own. Direct correspondence to Emilio J. Castilla, MIT, Sloan School of Management, 50 Memorial Drive, Room E52-568, Cambridge, Massachusetts 02142. E-mail: ecastilla@mit.edu

centage levels (Kochan, Katz, and McKersie 1986). However, this traditional model of employment has gradually been replaced by market-driven employment strategies, including merit-based reward systems and other performance management practices (Cappelli et al. 1997; Cappelli 1999). Perhaps organizations are increasingly adopting these merit-based practices and standards in the hope of ensuring that rewards are allocated meritocratically and eliminating inequity (Jackson 1998). Indeed, many workers find that these practices give them greater opportunities (Ospina 1996; Osterman 1999). However, there is a growing sense that individuals' career chances are becoming less equal (Frank and Cook 1995), and several scholars who study the transformation of the employment relationship have already raised equity and fairness concerns about the use of such practices (e.g., Osterman et al. 2001). Some studies have even suggested that the formalization of such practices may mask inequality in the distribution of rewards and may generate discrimination at the workplace (see Reskin 2000; Elvira and Graham 2002).

One of the key aspects of this market-driven way of organizing work has been the adoption of pay-for-performance and performance-management systems to measure and reward employees' merit and contributions to the company. Organizations frequently implement formal and informal performance evaluations that, in the end, affect major employee career outcomes such as task assignments, training opportunities, salary increases, and promotions (Cleveland, Murphy, and Williams 1989). Even though the study of pay-for-performance programs promises to contribute to our understanding of whether contemporary organizations that adopt merit-based practices remedy gender and racial inequality, we know little to date about how these policies influence the distribution of salaries and other rewards among employees. A few recent studies have looked at employee wages and careers within organizations, but in doing so they have ignored the role of merit and performance evaluations (for a review, see Petersen and Saporta [2004]). The same omission occurs in the line of research on organizations and inequality in employee attainment (for recent reviews, see Phillips [2005] and Roth [2006]). In addition, this body of research is incomplete because it has not examined in depth how these organizational merit-based practices can create the "opportunity structure" for gender and race discrimination (Petersen and Saporta 2004).

In order to make progress in our understanding of organizations and social stratification, I investigate in this article the role formal merit-based reward systems play in shaping gender and racial disparities in the distribution of wages in one work organization. Specifically, I examine the relationships between performance evaluations and two key outcomes—wage growth and promotions—using personnel data from a large service organization in the United States that started a performance evaluation

program as the basis for employee salary increase decisions. As widely advocated by employers and human resource specialists (see, e.g., Campbell, Campbell, and Chia 1998; Gerhart and Rynes 2003; Mathis and Jackson 2003), the organization I study decided to separate performance appraisals from pay decisions for two main reasons: (1) to facilitate the provision of feedback to employees for their future development and (2) to make pay decisions by strengthening the connection between employee performance evaluations and the size of employee pay increases at the end of the year.²

I argue, however, that by decoupling the performance evaluation and wage-setting processes, organizations may introduce the structural conditions for bias to occur at two distinct stages, as summarized in figure 1. The first stage is the performance evaluation stage, where *performance evaluation bias* can occur; in such cases, the performance rating process itself, because of its subjectivity, is affected by some gender, race, or nationality bias (arrow 1 of fig. 1). Significant progress has been made in understanding this type of bias, with important lab- and field-based studies comprehensively documenting the existence (and persistence) of performance evaluation bias (for a review of work in this tradition, see Bartol [1999], Elvira and Town [2001], Roth, Huffcutt, and Bobko [2003], and McKay and McDaniel [2006]). But even assuming that there is no bias in this first stage, or that it can be remedied, there is a second way in which performance evaluations may fail to ensure equal returns to employees: bias can affect the direct link between performance evaluations and employee career outcomes such as salary increases, promotions, or terminations (stage 2 of fig. 1). So, with the addition of this second stage in the performance-reward system, organizations might introduce discretion, which can result in the work of minority employees receiving less compensation over time even when they are evaluated as performing the same jobs at the same level as nonminority employees (arrows 2 and 3 of fig. 1).³

² In practice, the separation of performance appraisals from pay decisions can be carried out in three ways: (1) temporal separation only (the same individual makes the two decisions, but at different points in time); (2) interpersonal separation (two different people make the decisions, but at close points in time); and (3) temporal and interpersonal separation (the pay allocator makes the decision about compensation after the appraisal has been completed, but bases the decision on information provided in the appraisal). The process I study here is of the third type.

³ Arrows 2 and 3 of fig. 1 represent the two ways in which gender and race can affect stage 2. Arrow 2 illustrates the potential direct effect of ascriptive characteristics (gender, race, or country of origin) on employee career outcomes such as salary, salary increases, or promotions, net of employee performance evaluations. Arrow 3 illustrates the potential interaction effects between performance evaluations and ascriptive characteristics on employee career outcomes. I later argue that if merit-based practices

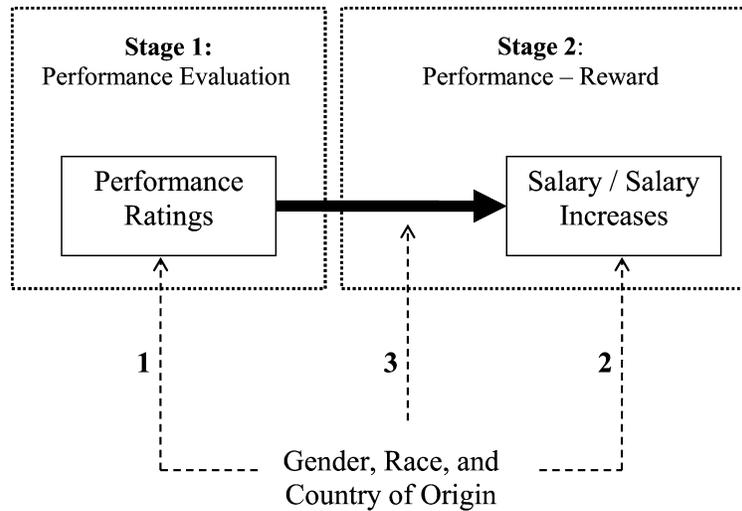


FIG. 1.—The theoretical performance-reward model under study

For the first time in the literature on organizations and gender/racial inequality, I develop and test two theoretical propositions that isolate processes of what I call *performance-reward bias*, whereby, even after merit is constructed in the performance evaluation stage, employers consciously or unconsciously discount the performance ratings of employees because of their gender, race, or nationality, *ceteris paribus*. This is a new form of valuative discrimination, which is independent of other processes generating ascriptive inequality, such as allocative discrimination or within-job wage discrimination (as described in Petersen and Saporta [2004]).⁴ Because equal merit results in equal rewards in any truly meritocratic system, a key challenge of these systems is how to measure merit.

linking performance evaluations to employee rewards work the way advocates of meritocracy claim, then the inclusion of stage 2 (that is, the performance-reward process) should also explain away both the direct effect of ascriptive characteristics (arrow 2) as well as the interaction effects between ratings and ascriptive characteristics (arrow 3) on employee outcomes over time.

⁴ Research on valuative discrimination has most frequently been done at the macro-level, establishing that female- and minority-dominated occupations with equal skills and wage-relevant characteristics are valued less than white male-dominated ones. Consistent with the valuative discrimination literature, the performance-reward bias is a more precise mechanism under which, once the merit or performance score has been constructed for each employee in the evaluation process, some workers still obtain different rewards for the same score as others. This performance-reward bias mechanism is independent of whether the occupation itself is valued less because it is dominated by women or minorities.

Once merit is measured for each employee, though, any differential reward for the same merit score is evidence of this performance-reward bias.⁵ This study uses a unique organizational setting to test this performance-reward bias in depth, making it possible to explore whether including employee performance ratings in any of the empirical models explains away the effect of gender or race on employee wage growth after controlling for job, work unit, supervisor, and human capital characteristics. This organization keeps rich personnel data; supervisors evaluate employees using dyadic performance evaluations, which constitute the primary basis for employee salary increases each year, and supervisors' evaluations and salary increase decisions are typically two distinct organizational processes.

Ultimately, this article helps to fill a significant gap in the research into the role of organizations in shaping inequality, by demonstrating that the formalization of performance management systems can introduce organizational processes and routines that make it possible for bias and discriminatory judgments to occur at several stages. My study focuses on one of these stages, namely, the link between performance evaluations and wage determination. I show that bias is likely to occur when the structural conditions are such that there is more discretion, less accountability, and less transparency. Ironically, while the practice of linking salary increases to performance ratings can create the appearance of meritocracy, it also creates the second (as well as the first) stage of performance management and thereby introduces the possibility of bias and discrimination. In my analyses, I find that women and minorities do not receive lower starting salaries or performance ratings than white men once I control for job and work-unit fixed effects. However, in the long run, my longitudinal analyses provide evidence of performance-reward bias and

⁵ The main empirical question of the article is whether similar measures of "merit" lead to similar levels of reward. I treat meritocracy as a process in which merit is somehow measured and then compensated. Meritocracy is thus one possible way of assigning rewards (nepotism and seniority, e.g., are other ways). This is a definition of meritocracy as a process, not as a value. Seen through this lens, the question at issue in the article is whether the process is consistent and, therefore, whether employees get the same reward for the same level of merit regardless of their gender, race, or nationality. If rewards are not consistent, they are either arbitrary (no telling who gets what rewards), discriminatory (some groups systematically get more or less rewards than others for equal levels of merit), or both at the same time. Meritocracy is also, however, an ideology that justifies the distribution of rewards. Sometimes I may seem to equate meritocracy with fairness, because these two concepts are popularly equated, but what I study is not fairness by some substantive standard, or in the perception of the individuals being judged, but the consistency of the formal process of assigning rewards that we call merit based. Unpacking what is actually happening inside a performance evaluation system described as meritocratic is the point of the study. I thank an anonymous reviewer for suggesting this clarification of terms.

show that different salary increases are granted for observationally equivalent employees (i.e., those in the same job and work unit, with the same supervisor and same human capital) who receive the same performance evaluation scores. This finding of performance-reward bias is robust after controlling for a number of complicating factors, including employee turnover.

Finally, because the results of this study imply that merit-based policies with high transparency and accountability may reduce bias and increase equity, this is an important contribution to our thinking about how employer practices can counteract discrimination and remediate bias. Drawing on my research, I suggest that the lack of both *accountability* and *transparency* behind the implementation of the second stage explains why, in an organization such as this, neither employees nor administrators seem to be aware of performance-reward bias. According to experimental research on accountability, when decision makers know they will be held accountable for their decisions, bias is less likely to occur (Tetlock 1983, 1985; Tetlock and Kim 1987). In this setting, accountability is less salient at the second stage, so performance-reward bias can be expected. The relative lack of formalization and transparency at the second stage also results in (or at least does not eliminate) performance-reward bias. In the discussion section of this article, I provide evidence that these theoretical mechanisms account for the existence of performance-reward bias in this particular organization. I also propose a few future research strategies for the continued investigation of the role of organizational practices (and the structural conditions they create) in the generation and reproduction of gender, race, or other non-performance related gaps in wages and careers.

WHY PERFORMANCE EVALUATIONS?

It has become standard practice for large organizations to establish a performance appraisal/reward system that attracts, retains, and motivates employees. Performance appraisal is the process of evaluating how well employees do their jobs in comparison to a set of standards and then communicating that evaluation to the employees (Mathis and Jackson 2003). These evaluations—also called employee ratings, performance reviews, or results appraisals—are widely used in contemporary organizations. According to the U.S. Census Bureau's 1994 National Employer Survey, one of the most comprehensive surveys of employers in the United States, supervisors conducted posttraining performance appraisals in 66% of 4,000 private sector establishments. Recently, Compensation Resources, Inc. (2004), surveyed 571 companies and found that almost 80% of U.S.

respondents conducted performance evaluations at least once a year (with 16% of the companies conducting them twice a year).

Organizations generally use performance evaluations for two primary, often conflicting purposes. The first is administrative. Organizations measure performance for the purpose of making administrative decisions about employees (e.g., pay, promotions, terminations, layoffs, and transfer assignments). In a recent compensation study by the Society for Human Resource Management, 69% of human resource (HR) professionals indicated that their organizations offer incentive compensation or variable bonuses based on performance (Burke 2005). Similarly, according to data from Hewitt Associates, as increases to base pay remained stable in 2007, more companies have been reported to rely on performance-related rewards (that must be earned anew each year) to motivate employees (Miller 2006).⁶

The second use of performance evaluations is developmental. Supervisors provide key information and feedback to their employees for future development. In such cases, supervisors act more as coaches than as judges, since they can inculcate in workers the desire to improve their job performance. Practically speaking, the administrative and developmental uses are often intertwined and difficult to distinguish when performance evaluations are implemented in real organizations. Historically, supervisors and managers have evaluated the performance of individual employees and have also made compensation recommendations for the same employees. However, many practitioners have advocated for the separation of performance appraisals and salary discussions, for several reasons. One reason is that decoupling these two processes and strengthening the tie between the performance evaluations of employees and their career outcomes encourages employees' perception of merit, increases job satisfaction, and motivates them to work hard (Martocchio 2004; Milkovich and Newman 2004). Second, employees often focus more on the monetary amount received than on the feedback. A third reason is that managers can manipulate performance ratings to justify the salary increases they wish to give specific individuals (Mathis and Jackson 2003).

⁶ Various studies, along with reports from particular companies, show a significant relationship between incentive plans and improved organizational performance (for a review, see Bohlander and Snell [2007] and Mathis and Jackson [2003]). Lawler's (2003) research, e.g., shows that a performance system is more effective when there is a clear connection between the performance management system and the reward system of the organization. Based on questionnaire data on performance management practices at 55 Fortune 500 companies, one of the main findings is that tying appraisal results to salary increases and bonuses is "a positive with respect to the effectiveness of the appraisal system" (pp. 398–99). In addition, reports by consulting firms indicate that higher-performing companies give out far more merit pay to their top performers than do lower-performing companies (IOMA 2000).

And last, supervisors do not like to complete performance evaluations (Pfeffer 1994), and they are reluctant to differentiate among employees, sometimes giving inflated ratings because they want to be popular (Gerhart and Rynes 2003).

To address some of these issues, many organizations have chosen to separate performance appraisals from salary discussions while strengthening the connection between employee performance and the size of employee pay increases. Some organizations have managers first conduct performance appraisals and discuss the results with employees at a later time. Others have introduced separate organizational processes, with different organizational actors in charge of the appraisal and compensation stages. Organizations typically decouple the performance evaluation process from the wage-setting process when they want to use the performance evaluation for both employee development and administrative purposes. Consequently, if any part of the performance appraisal system fails, better-performing employees may not receive larger pay increases, and the result is unfairness in the distribution of rewards. Despite wide interest in the issues of equity and fairness in the use of merit-based practices and their importance in helping us understand wage inequality in organizations, little research has explored how performance programs directly impact employees' wages and careers. In the remainder of this section, I discuss the body of literature on gender and racial inequality in organizations and present the main theoretical propositions of this article.

Gender and Racial Inequality and Bias in Organizations

Much sociological research has examined the link between ascriptive or personal characteristics and career outcomes (e.g., England 1992; Petersen and Morgan 1995; Nelson and Bridges 1999; and Petersen and Saporta 2004; this link represents a reduced form of arrow 2 of fig. 1, since these studies do not control for performance ratings). Discrimination seems to be pervasive in organizations, and many studies have documented different patterns and trends in discrimination across jobs and over time (for a review, see Petersen and Saporta [2004]). As mentioned above, Petersen and Saporta (2004) have proposed that gender disparities in wages and attainment caused by employer discrimination can come about through three different processes: allocative discrimination, within-job wage discrimination, and valuative discrimination. Empirical studies of these three processes have been undertaken in the gender discrimination and segregation literature. For example, the Petersen and Morgan (1995) study claims that within-job wage discrimination is currently unimportant. England (1992) and Nelson and Bridges (1999) show that valuative discrimination probably accounts for a substantial part of the gender wage

gap. Petersen and Saporta (2004) find that the largest gender differential exists in conditions at hire, with a 15% wage difference between men and women. These initial differences in job levels and salaries decrease to the point of disappearing over time, as employees remain in the organization and attain seniority.

There is also a large lab- and field-based literature on employer evaluation bias (e.g., Mobley 1982; Tsui and Gutek 1984; Pulakos et al. 1989) and even biased self-assessments (Ridgeway 1997; Correll 2001).⁷ Individual-level accounts, common in early psychological and organizational behavior studies, argue that, for various reasons, the demographic characteristics of the individuals doing the ratings (raters) and the individuals being rated (ratees) matter (e.g., Hamner et al. 1974; Hall and Hall 1976; Lee and Alvares 1977; Schmitt and Lippin 1980; London and Stumpf 1983). Although relatively few field studies have assessed such effects, the question of how the rater's and/or ratee's demographics impact performance evaluations remains inconclusive (see, e.g., Tsui and Gutek 1984; Thompson and Thompson 1985; Yammarino and Dubinsky 1988; Griffeth and Bedeian 1989; Tsui and O'Reilly 1989). Beyond the effects of simple individual-level demographics, several studies of gender and racial bias in performance evaluations focus on the employee-supervisor dyad as well as examine the group, team, and even organization levels (e.g., Wagner, Pfeffer, and O'Reilly 1984; Williams and O'Reilly 1998; Pelled, Eisenhardt, and Xin 1999; Elvira and Town 2001).⁸ More recently, there has been a discussion in the literature about discretion in weighing evaluative criteria (e.g., Hodson, Dovidio, and Gaertner 2002; Norton, Vandello, and Darley 2004; Uhlmann and Cohen 2005). In sociology, theoretical approaches to gender and racial inequality have emphasized cognitive psychological processes such as stereotypes operating in the workplace (e.g., Reskin 1998, 2000; Valian 1998; Gorman 2005).

Although past studies have addressed different parts of this puzzle, none of this prior research has examined whether there is any relationship between performance appraisal, wages, and wage growth. This potential link is especially important to study in organizations that adopt merit-

⁷ As noted above, recent reviews of this work can be found in Bartol (1999), Elvira and Town (2001), Roth et al. (2003), and McKay and McDaniel (2006).

⁸ For example, many studies have highlighted the importance of ratee-rater similarity in predicting employee performance ratings—what in sociology has been referred to as “homophily,” the tendency for employees to associate with and to “like” people similar to themselves (see McPherson, Smith-Lovin, and Cook 2001; see also, e.g., Tsui and O'Reilly 1989, Liden, Wayne, and Stilwell 1993). Some theories have also stressed that both actual and perceived similarity between rater and ratee perpetuate bias in evaluations (e.g., Turban and Jones 1988).

based reward systems.⁹ Because of this gap in the literature, it is less understood at this point how gender, race, and performance—specifically, subjective performance evaluations aimed at measuring employee merit and contribution—influence employee career outcomes within organizations. In addition, little attention has been given to how performance programs may create the structural conditions for bias and discrimination to appear. Studying the processes underpinning the link between the evaluation of merit and reward allocation is therefore vital if we are to understand inequality in today's organizations.

The Performance-Reward Bias Process

The most important challenge of meritocracy is measuring merit so that equal merit results in equal rewards. But once merit is measured for each employee, it is also crucial that there not be any differential rewards for the same merit scores. With this article, rather than looking into the specific motivations or determinants of discriminatory behaviors or whether the performance evaluations themselves are biased, I seek to explore how employee performance evaluations (used as a way of measuring employee merit and contribution) are associated with two important career-related outcomes, namely, salary increases and promotions, in one large service organization. More specifically, I examine the ways in which gender and race affect the performance-reward process and seek to empirically establish the existence of performance-reward bias. In empirical terms, evidence of either or both of the following scenarios supports the existence of performance-reward bias in organizations: (1) disparity in salary increases by race and gender net of performance ratings (i.e., the direct effects of gender and race on salary increases, controlling for performance, as illustrated by arrow 2 of fig. 1); and (2) disparity in the effect of performance ratings on salary increases by gender and race (i.e., the interaction effects between ratings and gender or race, or arrow 3 of fig. 1).¹⁰

⁹ On a related note, labor economists long accepted the assumption that observed higher relative earnings reflect higher relative productivity. Medoff and Abraham's (1981) study was the first one to start providing evidence of whether experience-earnings differentials can be explained by experience-productivity differentials—in other words, whether those paid more are more productive. Interestingly enough, the study uses computerized personnel data *only* on white male managers and professionals at a major U.S. manufacturing corporation (so-called Company C). Also, this study assumes that the performance ratings that supervisors give to their white male managerial and professional subordinates adequately reflect the subordinates' relative productivity in the year of assessment.

¹⁰ This is consistent with the proposed analyses for testing direct and indirect effects in Baron and Kenny (1986) and Judd and Kenny (1981).

In order to study this performance-reward process in depth, I structure my analyses as follows. I start by testing whether the gender, race, or national origin of employees have any significant effects on salary growth and promotions over time, controlling for the performance ratings given to employees by supervisors. Formal merit-based practices linking performance evaluations to employee compensation and promotions are meritocratic when the following theoretical proposition is supported:

PROPOSITION 1.—*After controlling for key human capital and job characteristics, equally performing employees are equally likely to obtain a performance-based reward, earn similar amounts in salary increases, and be promoted regardless of their non-performance-related demographic characteristics such as gender, race, or country of origin.*

By “equally performing employees,” I refer to employees who get the same performance ratings as the result of a performance evaluation process. If proposition 1 is true, in organizations where performance evaluations are used as the primary basis for employee rewards the inclusion of performance ratings in empirical models should explain away the effect of ascriptive characteristics on wage growth or promotion over time (i.e., there should be no such effect when employee performance is the only factor used to make compensation and promotion decisions; or, arrow 2 of fig. 1 should disappear).¹¹

Second, in addition to any difference in the payoff to performance ratings by gender and race, there can be some significant interaction effects between ratings themselves and race or gender. In this article, I also test these interaction effects, with the purpose of investigating whether performance evaluations are less effective at generating rewards for women and minorities. If formal merit-based practices linking performance evaluations to employee compensation and promotions are meritocratic, then the following proposition should be supported:

PROPOSITION 2.—*The effects of performance ratings on the likelihood of obtaining a performance-based reward, earning similar amounts in salary increases, and being promoted are the same for all employees regardless of their non-performance-related demographic characteristics such as gender, race, or country of origin.*

If proposition 2 is true, the interaction effects between performance ratings and race or gender in empirical models should not be significant

¹¹ I emphasize here that this is the prediction when the design and implementation of this performance-based reward is meritocratic—i.e., when it ensures that performance (or the subjective evaluation of it) is the main predictor in the distribution of rewards. This can also be expected because of the benefits associated with salary increases and promotions in general; from an economics standpoint, such promotions and salary increases help to motivate and retain high-quality employees (for a review of some of these economic theories, see Lazear [1998]).

(i.e., arrow 3 of fig. 1 should disappear). In the rest of this article, I test these two theoretical propositions, which isolate processes of performance-reward bias. The central finding of this study is that gender, racial, and nationality differences in salary growth persist even after controlling for performance evaluations (i.e., proposition 1 is rejected).¹² This study also supports the finding that performance ratings have a significantly lower effect on annual salary increases for African-American employees, *ceteris paribus* (i.e., proposition 2 is also rejected).

RESEARCH SETTING

The organization I study (henceforth referred to as ServiCo) is a large private employer with a total workforce of over 20,000 employees. ServiCo is primarily a service-sector organization, with several offices located in a competitive urban labor market in North America. This organization is particularly proud to offer a diverse work community. It is at the cutting edge in research and information technology. The organization offers health care and education benefits for employees and their families; it also offers professional development opportunities and flexible work options. Indeed, according to a new hire survey conducted by the company's HR division in early 2000, 40% of all respondents chose to work for ServiCo for reasons related to "professional development." About 11% of the respondents had "heard that ServiCo is a good place to work." In an exit survey in 2002, 74% of all respondents reported that they "still recommend this organization as a good place to work." ServiCo has employees in a variety of full-time, part-time, and temporary positions.

This organization offers a number of practical advantages for the current analysis. The HR department keeps detailed databases on the education and demographics of their employees (and supervisors), including gender, race, and nationality (U.S.-born vs. not U.S.-born). In addition to these computer databases, the HR department keeps electronic and paper files on each employee's performance evaluations and career outcomes such as salary increases, promotions, and terminations since 1996, including a standardized performance evaluation form on which the su-

¹² The main goal of this study is *not* to test whether supervisors tend to give women and/or minorities lower ratings than white men (fig. 1, arrow 1). This is because this path in particular has already been well studied in the performance bias literature (as reviewed above). Instead, my purpose is to examine whether there is bias affecting the link between performance evaluations and salary increases over time, after controlling for supervisor and work unit differences in salary increases. In a preliminary analysis of the performance data, however, I found that in this organization the distribution of performance evaluations looks the same regardless of gender, race, or nationality.

supervisor's and employee's names are clearly written. Employees at this company are evaluated at least once a year, which is typical in large organizations. From evaluation forms and other archival sources, I was able to construct a longitudinal database containing the job history and performance evaluations for all employees during the 1996–2003 period, a total of 8,898 employees. This database also includes the demographics, education, work history, and other human capital characteristics of those employees doing the evaluating, even when they were not included in the professional groups under analysis. Finally, wherever possible, I incorporate evidence gained through observations at a few work units in the organization and through reports, briefs, and other documents provided by the organization, as well as interviews with several individuals involved in different aspects of the performance-evaluation appraisal and salary decision making processes.

The Employees

The sample under study includes all of ServiCo's support staff, a total of 8,898 exempt and nonexempt nonexecutive and nonmanagement employees, two groups for which there has been great concern about inequity (Valian 1998; Petersen and Saporta 2004). For full-time, permanent jobs, there are five broad occupational groups: service (12% of positions), secretarial and clerical (23%), professional (52%), technical and semiprofessional (10%), and skilled crafts (3%). The organization did not allow me to access performance compensation data for top and middle managers, executives, or top professionals/consultants (this group constituted 36% of the total number of employees in the organization) nor data on unionized staff covered by the terms and conditions of a collective bargaining agreement (9% of all employees in the organization). Table 1 reports the descriptive statistics for the main variables analyzed in this study. To simplify the table, I omit descriptive statistics for the different job titles and work units or centers (there are 312 different job titles and 272 different work units during the period of study). As the table shows, this company is diverse; this is not surprising, given that the organization is "committed to recruiting and retaining a diverse workforce," as stated by one of the hiring managers. Among its employees, 67% are women, almost 19% are African-American, 9.5% are Asian American, 2.3% are Hispanic, and 5.5% are not U.S.-born. In 2003, the average annual salary was approximately \$41,000 (SD = \$28,000).

Gender, Race, and Meritocracy

TABLE 1
BASIC DESCRIPTIVE STATISTICS FOR VARIABLES OF INTEREST FOR ALL
EMPLOYEES, 2003

Variable	Mean (SD)	Percentage
Main demographics:		
Age (in years)	35.34 (10.21)	
Female		65.71
Male		34.29
African-American		18.78
Asian		9.54
Caucasian		69.10
Hispanic		2.30
Other race/ethnicity28
Single		50.84
Married		45.80
Divorced		2.94
Widowed42
Not U.S.-born		5.51
U.S.-born		94.49
Highest level of education achieved:		
Doctorate		9.51
Master's degree		12.34
Bachelor's degree		28.64
Some college		8.26
Associate's degree		4.60
Trade certificate		1.75
High school diploma		8.40
No education credentials		26.51
Salary and tenure:		
Tenure (in years)	2.65 (2.03)	
Salary (in dollars)	41,388.41 (28,243.10)	
Major occupational groups:		
Professional		51.90
Secretarial and clerical		23.10
Service		11.80
Skilled crafts		3.04
Technical and paraprofessional		10.16

NOTE.—N = 5,998 in the year 2003.

The Importance of Performance Evaluations

At ServiCo, the importance of performance evaluations is expressed through the company's Web site and other communication tools such as e-mail, memos, and pamphlets. According to the HR policy manual, performance is the "primary basis for all employee salary increases." Consequently, a performance appraisal must be completed for any employee obtaining a merit increase in order to validate the award. The overall performance appraisal return rate is on the order of 92% and continues

to improve over time.¹³ ServiCo started its “performance program” in the early 1990s in response to a report from a prominent consulting firm, which indicated that “most employees were reporting negative experiences, citing ‘poor management by supervisor’ and ‘lack of feedback from supervisor.’” Also, a small-scale survey of employees in 1993 suggested that 51% of respondents “did not feel that they were adequately recognized for their contributions.” One employee wrote in her survey response, “Although the quality of my work was consistently outstanding, my supervisor declared it to be of no value.” Another employee stated that, at that time (before the new performance program was implemented), “it does not matter how well I do; I feel we all get the same salary increase every year!” According to the director of HR, such responses highlighted the “need for ongoing supervisory training and the implementation of a new performance evaluation system.”

The new appraisal process was introduced in 1994. Its main purpose was (and remains) clear: to “facilitate constructive dialogue between employees and supervisors, to encourage the employee’s professional development, to clarify job duties and performance objectives, . . . and to make compensation decisions.” In pursuit of these goals, ServiCo has made efforts to separate the process of performance evaluation (the developmental use of appraisals) from the process of reward allocations (the administrative use). “The idea was to correct some of the problems the old performance evaluation system had, such as lack of feedback from supervisors,” according to the vice president of HR.

Figure 2 illustrates the performance appraisal process at ServiCo. All performance evaluations of staff members are dyadic: a supervisor (or a manager one level higher) evaluates a set of employees individually. The performance evaluation process is set up so that the supervisor meets with each employee annually to discuss and help the employee develop and improve his or her performance (fig. 2, step 1). On the basis of this evaluation, the employee might be recommended for a salary increase or bonus; typically, this recommendation comes from someone superior to the rater (step 2). According to the director of HR, the main purpose of implementing performance appraisals in this way is “to separate the pay decisions from the developmental use of the performance evaluation.” In the majority of cases, the head of supervisors (in large units) or the head of the unit (in smaller ones) recommends, on the basis of these performance evaluations, who will get a salary increase as well as the form and amount

¹³ The performance appraisal return rate is the percentage of employees whose performance evaluations are submitted to HR in a given year.

STEPS	TIME	LOCATION	ACTORS	TYPE OF DATA	PERFORMANCE-REWARD PROCESS	PERFORMANCE EVALUATION USE	STEP CHARACTERISTICS
1	Once a year	Unit in a given division	Supervisor	Electronic format	Evaluates employees once a year, using a short performance evaluation form (PE)	Developmental use (feedback)	Formalized with accountability and transparency
			Employee	Electronic format	Receives feedback and PE from main supervisor once a year		
2	End of calendar year	Unit in a given division	Head of the unit or head of supervisor(s) (depending on unit size)	Paper form	Based on the PE forms, recommends certain employees for an annual salary increase, as either (1) percentage or lump sum (more typical) or (2) amount (based on the unit budget each year; up to \$500)	Administrative use (pay decisions)	Formalized without much transparency or accountability
3	End of calendar year	HR division (separate building)	HR manager	Paper form	Signs off the recommendation to increase an employee's salary to ensure (1) only high performers are rewarded and (2) salary increases are justified appropriately		Formalized with transparency but without accountability

FIG. 2.—The performance management system at work at the organization under study

of each increase.¹⁴ There are two main ways of incrementing a worker's salary over time in this setting: (1) salary increases or adjustments as a percentage of the base salary or (2) bonuses as lump sums (these are small amounts of money assigned per unit or center—up to \$500, depending on the unit and year under study).¹⁵ All salary increases ultimately have to be approved by a member of the HR division (fig. 2, step 3). As one of the HR managers put it, "This way we can guarantee that supervisors work to enhance their employees' performance and development."

Compensation and Rewards

According to the HR policy manual, available online to all employees in the organization, the purpose of ServiCo's compensation system is to achieve the following three major goals: (1) to evaluate jobs consistently and fairly, (2) to pay competitive salaries, and (3) to regularly adjust the pay structures after considering the external market value for comparable jobs. As one of the HR staff members put it, "We want to support a system which provides flexibility in pay administration and career development." The salary range is posted online so that it is possible for employees to see where in the range their salaries fall. The position of the company is firm with regard to setting salaries: "The base salary for a new hire is set using factors that relate both to the individual's qualifications (education, experience and overall competence) as well as to ServiCo's current organizational and job needs" (quoted from the HR policy manual).

In addition to the base salary, ServiCo has extra compensation resources, including bonuses and salary increases. As the policy manual explicitly indicates, "Each work unit may choose among the many extra compensation programs available at ServiCo, [in order] to reward employees fairly and equitably and to recognize productive behaviors and attitudes." Certain extra compensation arrangements, as defined in the HR policy manual, require consultation with the HR division. Throughout the manual and other company documents, it is clearly stated that any

¹⁴ The allocation of the salary increase budget across work units is decided each year by several budget subcommittees, and the heads of the units and of supervisors are typically the ones who recommend who receives salary increases as well as the amount of those increases.

¹⁵ I am not capable of empirically distinguishing between these two different methods of incrementing a worker's salary over time. Unfortunately, I did not have additional information about the superior making each salary increase recommendation, nor about the person in HR making the final decision on whether to grant the salary increase or bonus. Such information is not kept in electronic format; the files containing this information are kept in file cabinets and are only accessed (if at all) at the time of the salary decision, as explained to me by one HR staff member.

extra compensation payment for an employee requires signed approval by a member of HR.

Employees are recommended for a salary increase or bonus by someone superior to their evaluating supervisors. The instructions in the performance evaluation form state that “performance is the primary basis for all salary increases.” Other company documents contain similar statements; this organization is clearly concerned with ensuring a meritocratic link between good performance and rewards: “[Especially in years with budget constraints,] increases must be reserved for the most productive employees.” There are also explicit statements about when not to award salary increases: “No salary increase will be awarded to employees exhibiting unacceptable levels of performance,” according to the HR policy manual.

Measuring Employee Performance

During the period of analysis, from 1996 to 2003, 38,832 performance evaluations of 8,818 different employees were submitted to the HR division. The scores were recorded in electronic format and the evaluation forms were stored in secured file cabinets. Generally, supervisors prefer to fill out a one-page evaluation form (the “short form”) as opposed to a longer version in which more detail may be provided about the employee’s performance and developmental needs. The second section of the short form asks the supervisor to summarize the employee’s performance by selecting one of five scoring categories. The supervisor is instructed to choose the category that “best describes the employee’s overall performance.” Performance ratings range from 1 to 5, with the following qualitative statements assigned to each score: (1) “performance is unacceptable for the job and important improvement is required”; (2) “performance does not consistently meet all established expectations for the job and requires improvement”; (3) “performance consistently meets established expectations for the job”; (4) “performance is reliable and consistently meets and at times exceeds all established expectations for the job”; and (5) “performance is clearly and consistently outstanding in most aspects of current job responsibilities.”

In 2001, the average employee performance rating was approximately 4 (SD = 1.15). In 2003, the average rating was slightly lower (3.96) with a higher standard deviation (1.37). Figure 3 shows the frequency of all possible evaluation outcomes for the 5,904 employees evaluated in the sample under study in 2003: about 72% of the evaluations fall into the two top performance categories. The distribution does not change shape

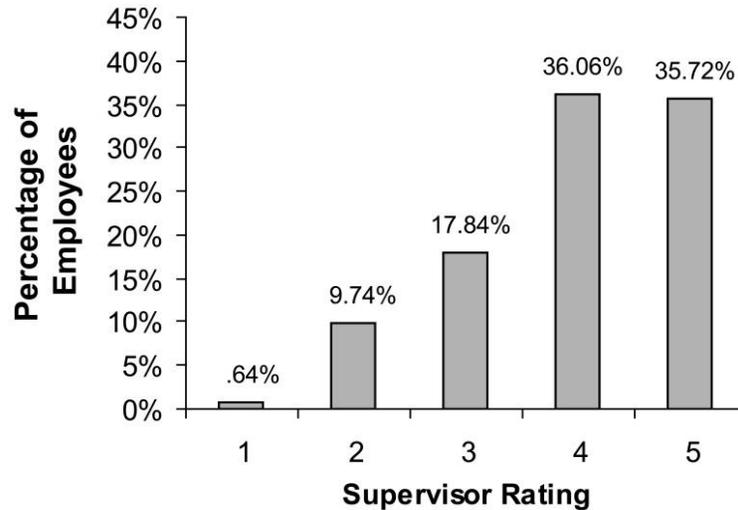


FIG. 3.—Performance histogram for all employees in 2003 ($N=5,904$)

when it is examined by employee gender, race, or nationality (these analyses are available upon request).¹⁶

Below the performance summary, the supervisor can also write additional comments and recommend follow-up actions. At the very end, the form has space for the signatures of the staff member being evaluated, the supervisor, and the administrative representative in the work unit overseeing the performance evaluation process. The document is then forwarded to the HR division to become part of the staff member's official personnel file. Of all the employees who received performance evaluations, about 9,191 (23.7%) were recommended for and subsequently granted a salary increase by someone superior to the supervisor evaluating the employee. "Outstanding" evaluations represented 4.5% of the total. All requests for a salary review must be submitted to the HR compensation office with a letter of request from the "appropriate administrative unit head." One of the HR managers in the compensation office noted that "this is typically done by the head of the supervisors, in the case of large centers, or by the head of the unit, in the case of smaller ones." As a final check, all salary raises or bonuses require a final sign-off by a member

¹⁶ Again, I emphasize that the main goal of this article is not to test whether supervisors tend to give women and/or minorities lower ratings in comparison to white men in the performance evaluation stage. Instead, my intent is to focus on whether there is bias affecting the link between performance evaluations and salary increases over time—even when there may be no bias in the first stage.

of the HR division, which is independent of any other divisions: “We are basically avoiding any allocation of these bonuses to be unjustified,” according to the manager of one of the units at ServiCo. The increases are then effective in the first pay period of the new fiscal year, regardless of when the performance evaluation is submitted (which is almost always at the end of the fiscal year).

RESEARCH METHODOLOGY

Using personnel data on performance and compensation from this one organizational setting, I start by testing whether ascriptive characteristics such as gender, race, or national origin influence salary growth and promotions over the tenure of an employee after I control for the level of employee performance (proposition 1). All of these models are estimated with additional controls for tenure in the job, part-time status, and level of education as well as job title, unit/center, and supervisor fixed effects. These control variables allow me to test whether observationally equivalent employees with different demographic characteristics get different salary increases over time, even after they receive the same performance evaluation scores. I also test interaction effects between performance evaluations and employees’ ascriptive characteristics (proposition 2). In the next two subsections, I explain the regression equations estimated in this study.

Salary Growth

In order to test whether ascriptive characteristics such as gender and race have an effect on salary growth, I specify and estimate the regression equation displayed below. The performance–salary growth data structure is a pooled cross-sectional time series (yearly). The data are unbalanced, and consequently, the number of observations varies among employees because some individuals leave the organization earlier than others (while many workers stay). Research studies typically model such data with fixed-effects estimators, which analyze only the within-individual, over-time variation. This choice is unappealing in this context, because the majority of the independent variables (i.e., ascriptive characteristics) do not vary over time. I estimate various cross-sectional time-series linear models using the method of generalized estimating equations (GEE).¹⁷ I

¹⁷ The method of GEE deals with the correlation of the errors directly by specifying and estimating the variance-covariance error matrix. The results do not change when estimating models imposing less structure on the variance-covariance error matrix (e.g., the traditional OLS model, or other regression models clustering around employees).

report the robust estimators that analyze both between-individual and within-individual variation. Specifically, I use the method of GEE developed by Liang and Zeger (1986). This method also requires specifying and estimating a correlation structure when estimating these models:

$$\ln(w_{i,t}) = \alpha + \beta_0 \ln(w_{i,t-1}) + \beta_1' P_{i,t-1} + \beta_2' X_{i,t-1} + \beta_3' D_{i,t-1} + \varepsilon_{i,t}, \quad (1)$$

where the dependent variable is the natural logarithm of annual salary at time t (and salary at time $t - 1$ is one of the main independent variables in the model).¹⁸ I include three different vectors of independent variables. The first one ($P_{i,t}$) includes a set of dummy variables for four of the five possible employee performance ratings in a given year—the omitted category is 3, “performance consistently meets established expectations for the job.”¹⁹ The second vector of demographic variables ($X_{i,t}$) includes dummy variables for female, African-American, Asian American, Hispanic, and non-U.S.-born employees (the omitted category is U.S.-born white male) and dummy variables for marital status (married, divorced, and widowed; single is the omitted category).²⁰ The vector also includes a set of dummy variables controlling for the highest level of education achieved (where the omitted category is college degree), age, and part-time status. The third vector ($D_{i,t}$) includes dummy variables for job title, unit, and supervisor.²¹ Adding this vector of variables to the equation allows me to examine the impact of performance evaluations on salary

Under mild regularity conditions, GEE estimators are consistent and asymptotically normal, and they are therefore more appropriate for cross-sectional time-series data structures.

¹⁸ In preliminary analyses, I estimated a set of growth models where the dependent variable was the natural logarithm of annual salary *growth*; i.e., $\ln(w_{i,t} - w_{i,t-1})$. I also estimated several autoregressive models to account for the potential heteroscedastic autocorrelated behavior of the error terms (using the *xtgls* command in Stata, as I have previously described in detail [Castilla 2007]). I always found results very consistent with the ones I report in this article. These models, not shown here, are available upon request.

¹⁹ Results do not change substantially if performance is included as a continuous variable ranging from 1 to 5. Given that performance evaluations are not normally distributed (as shown above in fig. 3), including a set of dummy variables in the model is the appropriate choice in this case.

²⁰ In earlier regression analyses, I also included two dummy variables to account for Native American and Pacific Islander employees, who represent less than 0.28% of all staff employees in any given year during the period under study. None of the substantive results changed when I included those two variables in all the models estimated in this article. The estimated coefficients for these two dummy variables were always close to zero and insignificant.

²¹ In the case of small work units (those in which there is only one supervisor), the work unit and supervisor fixed effects are obviously redundant. In such cases only one control was introduced.

increases, controlling for important work-level variables.²² The model is estimated for all workers in the population under study during the 1996–2003 period.

As a robustness check, because gender and racial differences in salary growth may also reflect gender and racial differences in turnover rates, the above estimated salary growth models are also estimated correcting for employee turnover. Since turnover may change the gender and racial composition of the workplace, the observed disparities in salary growth when measured across a cohort of workers (not for any particular individual) could be entirely due to population heterogeneity (Tuma 1976; Price 1977; Jovanovic 1979). Any attempt to assess the relationship between demographic characteristics and salary increases requires separating these two processes and taking into account the risk of employee turnover. Following Lee (1979, 1983), Lee, Maddala, and Trost (1980), and Lee and Maddala (1985), I control for the retention of employees over time by including the previously estimated turnover hazard when I estimate such longitudinal models. This results in a two-stage estimation procedure as follows:

$$\begin{aligned} \ln(w_{i,t}) = & \alpha + \beta_0 \ln(w_{i,t-1}) + \beta_1 P_{i,t-1} + \beta_2 X_{i,t-1} + \beta_3 D_{i,t-1} \\ & + \delta \bar{\pi}(t-1, Z_{i,t-1}) + \varepsilon_{i,t}, \end{aligned} \quad (2)$$

where the dependent variable is still the natural logarithm of annual salary at time t (as in model [1]); however, now the model includes $\bar{\pi}$, which is the estimated turnover hazard rate (using event history analysis):

$$\pi(t, Z_{i,t}) = \exp[\Gamma'Z_{i,t}] q(t), \quad (3)$$

where π is the instantaneous turnover rate. This rate is commonly specified as an exponential function of covariates multiplied by some function of time, $q(t)$. Z is a vector of covariates that affect the hazard rate of turnover for any given hire. Z is indexed by i to indicate heterogeneity by individual case and by t to make clear that the values of explanatory variables may change over time. I estimate the effect of the variables in model (3) using the Cox model, which does not require any particular assumption about the functional form of $q(t)$ (Cox 1972, 1975). Because of potential issues

²² Because this organization has a detailed job classification system, I am able to control for the native job title, a six-digit code indicating job classification. Examples of job titles include specific job positions, ranging from non-IT jobs such as regular clerks, sales clerks, administrative coordinators, and assistants, to IT jobs such as computer technicians and support specialists (there are 312 different job titles in the employee population under study). This is an improvement over past studies, in which the analysis is at the job-grade level. With this methodology, I am therefore examining both job titles and the level at which the employee is performing his or her job duties in a given job, in a given work unit, under the supervision of a given superior.

relating to the inclusion of the predicted value from a nonlinear model into any model (see Hausman 2001), I also tested for the effect of turnover across other different specifications, functional forms, and measures of turnover; I always found similar results.

Salary Increase Decisions and Promotions

In addition to the models estimating salary increases over time, I estimate various event history models that explore the impact of performance evaluations on both salary increases and promotion decisions. Thus, I estimate two sets of Cox regression models as specified in the following equation:

$$\pi(t) = \alpha + \beta_0 \ln(w_{i,t-1}) + \beta_1 P_{i,t-1} + \beta_2 X_{i,t-1} + \beta_3 D_{i,t-1} + \varepsilon_{i,t}. \quad (4)$$

In the first set of models, the dependent variable is the instantaneous hazard rate of a salary increase decision. This variable takes the value of 1 if the employee is awarded a bonus or a salary increase and is 0 otherwise. In the second set, the dependent variable measures whether the employee is promoted (the value is 1 if the employee is promoted, 0 otherwise).²³ Given that an employee can be promoted and/or his or her salary can be increased several times during his or her tenure in the organization, these two processes are modeled as a series of repeatable events. The main purpose of these analyses is to test whether bias occurs in the link between performance evaluations and more visible career outcomes such as salary increases (regardless of quantity) and promotions. I argue that both salary increase decisions and promotions are visible organizational processes at work; employees may notice who gets promoted and who gets salary increases over time. This is in contrast to the *amount* by which an employee's salary might be increased every year (modeled in the previous subsection), which is typically unobservable or unknown information to the rest of employees in the organization, so that any concrete salary comparisons among employees are eliminated.²⁴

²³ These models were also estimated controlling for turnover, as explained in detail above. Similar results were found.

²⁴ Closely related to this, some empirical and theoretical work in organizational behavior has explored the role of information processing in perceptions of discrimination (Crosby 1982, 1984; for a review, see Major et al. [2002]). I come back to this work in the discussion section below.

RESULTS

Starting Salary and Salary Growth

I begin by testing whether ascriptive characteristics influence the process of allocating starting salaries to new employees. In order to avoid any salary decisions affecting exclusively internal employees, I only analyze those individuals hired from outside the company from 1996 to 2003—8,298 employee hires. The first column of table 2 presents the coefficients of the starting salary model controlling for year of hire, job title, and hiring unit fixed effects (these coefficients are omitted to facilitate the reading of the table).²⁵ The dependent variable is the natural logarithm of annual salary in the year of hire. I find no evidence in this organization of significant differences in appointment salary by gender or race after controlling for year of hire, job title, and work unit, and other important human capital characteristics such as the highest education level attained. Asian Americans, however, seem to earn a starting annual salary 1.5% lower than whites (the coefficient is barely significant at the .10 level). Non-U.S.-born employees, however, make 4.5% less money than U.S.-born employees ($P < .001$).²⁶

I next test whether ascriptive characteristics influence salary growth over the tenure of an employee after I control for the level of employee performance evaluation (proposition 1). The rest of the columns in table 2 report the results of several salary growth models; all models control for job title, unit, and supervisor fixed effects.²⁷ By comparing the coefficients of models 1 and 2, I find that the effects of demographic characteristics such as gender and race do not change much (either in magnitude or significance) when the four performance rating dummy variables are introduced in the analyses. Similar results are found in models 3 and

²⁵ There are 312 different job titles within the study sample. ServiCo has 12 divisions, each of which comprises different units. Each work unit/center typically has a head of unit, a few staff supervisors, and sometimes some supervisor assistants. Several staff members per supervisor (or supervisor assistants) support a group of top professionals, consultants, and researchers. There are 272 different units during the period of study. The HR division is independent from the rest of the divisions, and has the typical offices or units—namely, compensation and benefits, training and development, information systems, and payroll. To simplify the table, I omit the different fixed-effects coefficients included in the estimation of the models.

²⁶ The main demographic coefficients did not change substantively when I introduced interaction terms among gender, race, nationality, and marital status. In these interaction effect models, one important finding is worth describing: non-U.S.-born males earn 11% less than U.S.-born males; the difference in starting earnings between U.S.-born and non-U.S.-born females is 2.7%.

²⁷ The coefficient estimates reported in table 2 do not change much whether the models include or exclude the job title, unit or center, and supervisor fixed effects in the models. I chose not to report the models without fixed effects for simplification purposes.

TABLE 2
REGRESSION MODELS PREDICTING ANNUAL STARTING SALARY AND SALARY GROWTH

Independent Variable	Starting Salary (ln) ^a	Salary at Time <i>t</i> (ln) ^b				Turnover Rate Model (Cox Regression)
		Model 1	Model 2	Model 3	Model 4	
Constant	11.006*** (.2095)	2.357*** (.3746)	2.284*** (.3641)	.255*** (.0609)	.254*** (.0572)	
ln annual salary (<i>t</i> -1)		.785*** (.0337)	.791*** (.0327)	.980*** (.0059)	.979*** (.0056)	-.185*** (.0135)
Tenure		.002***	.002***	.002***	.002***	-9.482*** (.2938)
Age	.004*** (.0003)	.001***	.001***	.000	.000	-.031 (.0103)
Part-time	-.011* (.0055)	-.005**	-.003**	.001	.003	5.526 (10.2546)
Performance rating:						
Unacceptable (1)			-.027 (.0547)		-.036 [†] (.0204)	2.613*** (.3907)
Requires improvement (2)			-.017*		-.013*	1.176***
Good and reliable (4)			.014*** (.0022)		.010*** (.0029)	-.134 (.1555)
Outstanding (5)			.024*** (.0022)		.018*** (.0027)	-.045 (.1589)
Demographics:						
Female	.006 (.0052)	-.004*** (.0013)	-.004** (.0014)	-.004** (.0012)	-.004** (.0012)	.072 (.1217)

African-American	-.007 (.0067)	-.006*** (.0012)	-.004*** (.0013)	-.005*** (.0013)	-.005*** (.0015)	.083 (.1516)
Asian American	-.015 [†] (.0080)	.001 (.0025)	.001 (.0026)	-.001 (.0036)	-.001 (.0037)	.525* (.1759)
Hispanic	-.015 (.0141)	-.006*** (.0015)	-.006*** (.0016)	-.005*** (.0015)	-.005*** (.0016)	.664 [†] (.2820)
Not U.S.-born	-.045*** (.0099)	-.005*** (.0013)	-.004*** (.0013)	-.006*** (.0015)	-.006*** (.0015)	.211 (.2664)
Marital status:						
Married022*** (.0052)	-.001 (.0018)	.000 (.0018)	-.001 (.0018)	-.002 (.0018)	-.783*** (.1330)
Divorced010 (.0151)	-.010* (.0041)	-.004 (.0034)	.002 (.0037)	.003 (.0035)	-.241 (.3852)
Widowed	-.028 (.0537)	-.019 (.0125)	-.016 (.0112)	-.011* (.0053)	-.009 [†] (.0054)	-5.462 (22.2098)
Turnover hazard rate ($t-1$) ^c						
χ^2 statistic		666,281***	798,202***	140,695***	167,422***	2,104***

NOTES.—Numbers in parentheses are SDs. All models include dummy variables for highest education level achieved. The omitted category for highest education level achieved is “college”; for performance rating, “(3) performance consistently meets established expectations for the job”; for demographics, “U.S.-born white male”; and for marital status, “single.” All salary growth models control for job title, unit/center, and supervisor fixed effects.

^a N of external hires = 8,298. Model controls for year of hire, job class, and unit/center (but does not control for supervisor fixed effects, because initial salaries are determined prehire). F -test = 57.57 ($P < .001$); $R^2 = 0.82$.

^b N of employees = 5,104.

^c The turnover hazard rate is estimated from the Cox regression model reported in the last column of the table.

[†] $P < .10$ (all two-sided t -tests).

* $P < .05$.

** $P < .01$.

*** $P < .001$.

4 when the salary growth models correct for employee turnover—specifically, the fact that females and minority employees might have different propensities to turn over.²⁸ Because gender and racial differences in salary growth can also reflect (among other things) gender and racial differences in promotion rates, I also estimated the salary growth models correcting for employee promotion and found similar results.²⁹

Tenure has the expected positive sign ($P < .001$).³⁰ Age and part-time status are only significant variables when not controlling for turnover: older employees obtain higher salary increases ($P < .001$) and part-time employees get lower increases than full-time employees (the coefficient is significant at the .01 level in models 1 and 2).³¹ Most importantly, the performance rating dummy variables in models 2 and 4 also have the expected signs when organizational practices are in place to reward performance. Thus, according to model 2, those employees whose performance was judged as requiring improvement received a salary increase 1.7% lower than employees whose performance was average ($P < .05$). Employees whose performance was good and reliable received an increase 1.4% higher than employees with average performance ($P < .001$); finally, employees whose performance was outstanding received an increase 2.4% higher than those with average performance ($P < .001$).

The key finding in this study is reported in models 2 and 4 of table 2, where I present the effects of demographic characteristics on salary growth after the employee level of performance is introduced in the regression equations. These models do not support proposition 1: I find significant effects for certain individual characteristics on salary growth after controlling for employee performance levels. More specifically, from model 4 I find that, *ceteris paribus*, the salary growth is 0.4% lower for women than for men, even after performance evaluations are introduced into the model ($P < .01$). African-American employees get a salary increase 0.5%

²⁸ I tested for the effect of turnover across different specifications, functional forms, and measures of turnover and always found comparable results. Similarly, results do not change across a variety of parametric transition rate models (i.e., the Cox model presented in the table, the proportional exponential model, or the proportional Weibull model).

²⁹ Results were similar to those reported in models 3 and 4. The promotion dummy variable included in the model (with a value of 1 if the employee was promoted, 0 otherwise) was positive but insignificant in all models.

³⁰ Work in the human capital tradition argues that work experience declines over time. In preliminary analyses, I captured this effect by entering a squared term for tenure. I found no evidence of diminishing returns to tenure (the *t*-value of the squared term for tenure was always less than 1). I therefore present the model with the simple linear effect of tenure because it is the best-fitting specification.

³¹ As with the tenure effect, I experimented with various nonlinear specifications and found that the simple linear effect of age has the best fit.

lower than equally performing white employees. In addition, Hispanic Americans get a salary increase 0.5% lower than whites ($P < .001$ for both coefficients). A significant salary increase discrepancy is also found for non-U.S.-born employees, who get a salary increase 0.6% lower than native employees, other things being equal ($P < .001$). I thus demonstrate that observationally equivalent employees with different demographic characteristics get different salary increases even after they receive the same performance evaluation scores.³²

Model 4 presents the results of the salary growth regression model, correcting for the turnover rate.³³ Looking at the coefficient for the estimated hazard rate in the salary growth model, I find that the likelihood of turnover is associated with lower salary growth over time. In other words, the more likely an employee is to leave the organization, the lower his or her salary increase is (the coefficient is negative and barely significant at the .10 level). In looking at the results of the turnover hazard rate analysis in models 3 and 4 (reported in the last column of table 2), I do not find statistically reliable evidence that high performers are more likely to leave this organization. Instead, I find that “unacceptably performing” workers and employees whose “performance requires improvement” are, respectively, 13 and 3 times more likely to leave than employees whose performance “meets established expectations for the position” ($P < .001$ for both coefficients).³⁴ As in the salary growth models, tenure and annual salary in the previous year have significant effects on turnover ($P < .001$ for both coefficients). The longer the tenure of the employee, the less likely he or she is to leave the organization. Higher-paid employees are also less likely to turn over. The model reports that Asian American and Hispanic employees are more likely to turn over in the organization ($P < .05$ and $P < .10$, respectively).³⁵

³² These results do not change much after controlling for employees’ earlier promotions, bonuses, and even past performance evaluations. In some additional models, I also controlled for past salary increases and found that salary increase decisions in the current year are unrelated to the previous year’s salary increase recommendations. In other words, the coefficient for last year’s decision to increase salary, as well as the amount of the raise, is unrelated to the current year’s decisions. This empirical finding is consistent with the company’s HR policy manual and culture, which encourage using current performance ratings as the primary determinants of bonus and raise decisions each year. To avoid too much detail about different salary growth model specifications, these additional regression analyses are not shown here.

³³ Similar results were found when controlling for employee promotions.

³⁴ $13 = \exp(2.6)$; $3.24 = \exp(1.17)$.

³⁵ I also estimated a set of models including different interaction terms, not shown here. In general, including two-way and three-way interactions among demographic variables did not improve the fit of the model (all probabilities of the incremental χ^2 statistics are insignificant at the .01 level).

In order to test proposition 2, in table 3 I reestimate model 4 of table 2 including different sets of interaction terms. For presentation purposes, table 3 only reports the main interactions between demographic variables and performance ratings. The purpose of these interaction effects models is to test whether performance evaluations themselves are less effective at generating rewards for women and minorities (proposition 2). The different models in this table show that most of these interactions are insignificant and that, overall, the fit of the model does not seem to improve much when these interaction terms are included.³⁶ Gender does not seem to significantly change the impact of performance ratings on salary growth. However, when it comes to race, the interactions for African-American employees are found to be significant: the positive effect of ratings on annual salary increases is lower for African-Americans (coefficients are significant at least at the .05 level). I therefore reject proposition 2; this organization rewards the same performance score differently for certain demographic groups (in this case, less generously for African-American employees). The main effects do not change much when these demographic-performance interactions are added to the model.³⁷

Finally, in order to evaluate the magnitude of these salary increase differences, I calculated the lifetime cumulative effects of this performance-reward bias over a 10-year period for several employees in this organization, using the coefficients of model 4 in table 2. For example, if equally performing white men get a 10% salary increase each year, white women are predicted to get a 9.96% increase—less, but not substantially so. Thus, if men and women both start at \$10 per hour (or \$20,000 a year, assuming that full-time employees work 50 weeks a year, 40 hours a week), men get a 10% increase per year (resulting in an annual salary of \$22,000 after one year), and women get a 9.96% increase (to \$21,992 after one year), then after 10 years men will make \$25.94 per hour (almost \$51,900 a year), while women will make \$25.84 per hour (approximately \$51,700 a year).³⁸ So from an initial parity in wages—a wage gap of 1 (10/10)—

³⁶ These results are not surprising, given my analyses of employee performance evaluations (discussed above): I found that the distribution of ratings does not change shape when examined by gender, race, or nationality.

³⁷ To simplify the presentation of these models, I do not report constant terms and the main effects of variables included in model 4, table 2. These omitted coefficients do not differ appreciably from the values reported for that model.

³⁸ To compute these numbers, I used the following compound interest (future value) formula: $W_T = W_0(1 + r)^T$, where W_T is the hourly wage at time t , W_0 is the starting hourly wage, T is the unit of time, and r is the rate of wage increase.

TABLE 3
REGRESSION MODELS PREDICTING ANNUAL SALARY GROWTH: PERFORMANCE AND DEMOGRAPHIC INTERACTION EFFECTS

Main Interaction Models	Performance Rating				Model χ^2 Test (df)	Incremental χ^2 Test (df)	Prob > χ^2
	Unacceptable (1)	Improvement (2)	Good (4)	Outstanding (5)			
Gender × performance:							
Female	.038 (.0486)	.013 (.0095)	-.006 (.0043)	-.007 (.0047)	8,413,169*** (616)	6.82 (4)	.146
Race × performance:							
African-American	-.017 (.0321) dropped	-.031** (.0111) .003 (.0136)	-.013* (.0057) -.011 (.0007)	-.020** (.0073) -.011 (.0079)			
Asian American							
Hispanic	dropped	-.083 (.0607)	.005 (.0137)	-.001 (.0115)	26,400,000*** (622)	16.83 (10)	.078
Nationality × performance:							
Not U.S.-born	dropped	-.015 (.0210)	-.003 (.0099)	-.006 (.0075)	8,802,348*** (615)	.90 (3)	.826
Gender × race × performance ...					5,194,613*** (626)	25.60* (14)	.029
Gender × race × nationality × performance					1,287,463*** (629)	26.60* (17)	.064

NOTES.—N = 5,104. Only interaction terms are reported in this table. All models include all main effects and independent variables as in model 4 in table 2. Numbers in parentheses in the performance rating models are SDs. The omitted category for performance rating is “(3) performance consistently meets established expectations for the job.”

* $P < .05$ (all two-sided t -tests).

** $P < .01$.

*** $P < .001$.

there will be a wage gap of 0.996 a decade later.³⁹ After 10 years in the company, the largest wage gaps are found for African-American and Hispanic women (0.992) and for non-U.S.-born women (0.991).

Salary Increase Decisions and Promotions

In the previous section, I examined the amount of salary increases for employees by gender, race, and nationality. In this section, I present the models testing whether demographic characteristics influence decisions to increase salary regardless of the magnitude (examining a total of 9,191 salary decisions) and to award promotions (examining a total of 262 promotion decisions) over the tenure of an employee, after controlling for the level of employee performance. Table 4 reports the results of several Cox regression models. Regarding annual salary increase or bonus decisions (1 if the employee gets a bonus or salary increase, 0 otherwise), I find that performance evaluation ratings are the most significant predictors at ServiCo—coefficients are reported in the first two columns of table 4. Employees whose performance is “unacceptable” or “requires improvement” are, respectively, approximately 71% and 12% less likely to get a salary increase than employees whose performance “consistently meets established expectations for the position” ($P < .05$).⁴⁰ Employees with “good and reliable” and “outstanding performance” evaluations are more likely to get a salary increase over time—5% more likely if performance is good and reliable ($P < .05$) and 23% more likely if it is outstanding ($P < .001$). Clearly, when it comes to the decision to increase salary or to award a bonus (regardless of the amount), ascriptive characteristics do not seem to matter much—with the exception of non-U.S.-born employees, who are 14.5% less likely to get a salary increase ($P < .01$).

The last two columns of table 4 present the coefficients for employee promotion (1 if the employee is promoted, 0 otherwise). In this case, performance is not as significant in predicting promotions as it was in predicting salary increases. This is consistent with what I learned from my interviews with a few supervisors, who indicated that these annual

³⁹ I performed several supplementary analyses to ensure that the results presented in this article are robust. I estimated the models presented in tables 2 and 3 separately by divisions (12 in total), by the five broad occupational groups as well as by the exempt and nonexempt worker categories, and found substantially similar results. Additionally, I analyzed the data separately by gender, race, and nationality. The results of all these additional analyses still demonstrated that after controlling for key job and human capital characteristics, both performance evaluations and non-performance-related ascriptive characteristics do explain variation in annual salary increases in this research site. Results are available upon request.

⁴⁰ $-70.5\% = 100\% \times [\exp(-1.22) - 1]$; $-11.7\% = 100\% \times [\exp(-.124) - 1]$.

TABLE 4
COX REGRESSION MODELS PREDICTING SALARY INCREASE DECISIONS AND PROMOTIONS

Independent Variable	Salary Increase Decision		Promotion	
	Model 1	Model 2	Model 1	Model 2
ln annual salary ($t-1$)	-2.294*** (.0823)	-2.332*** (.0827)	-1.170*** (.2534)	-1.263*** (.2569)
Tenure	-.011 (.0078)	-.015* (.0078)	-.080 [†] (.0454)	-.096* (.0461)
Age	.009*** (.0016)	.010*** (.0016)	-.011 (.0089)	-.010 (.0089)
Part-time	.151*** (.0369)	.153*** (.0370)	.113 (.1848)	.092 (.1872)
Performance rating:				
Unacceptable (1)		-1.222* (.5977)		-5.756 (33.6376)
Requires improvement (2)		-.124* (.0604)		.497 [†] (.2788)
Good and reliable (4)		.051* (.0251)		.153 (.1881)
Outstanding (5)		.210*** (.0370)		.477* (.2022)
Demographics:				
Female	.037 (.0293)	.029 (.0294)	-.189 (.1480)	-.222 (.1486)
African-American	-.010 (.0364)	.014 (.0367)	-.059 (.1879)	-.034 (.1898)
Asian American	-.029 (.0488)	-.016 (.0490)	-.313 (.3083)	-.332 (.3106)
Hispanic	-.117 (.0806)	-.119 (.0808)	-.193 (.4660)	-.231 (.4657)
Not U.S.-born	-.150* (.0601)	-.157** (.0602)	-.708 (.4657)	-.734 (.4708)
Marital status:				
Married	.035 (.0279)	.029 (.0280)	.212 (.1533)	.221 (.1537)
Divorced	.170* (.0759)	.167* (.0762)	.405 (.3697)	.377 (.3713)
Widowed	-.143 (.2846)	-.106 (.2845)	-2.555 (10.1264)	-5.089 (23.0167)
χ^2 statistic	3,074***	3,120***	709***	715***
Log likelihood	-70,173	-70,038	-1,743	-1,738
Number of events	9,191	9,191	262	262

NOTES.— N of employees = 8,898. Numbers in parentheses are SDs. All models include dummy variables for highest education level achieved. The omitted category for highest education level achieved is “college”; for performance rating, “(3) Performance consistently meets established expectations for the job”; for demographics, “U.S.-born white male”; and for marital status, “single”. All models control for job title, unit/center, and supervisor fixed effects.

[†] $P < .10$ (all two-sided t -tests).

* $P < .05$.

** $P < .01$.

*** $P < .001$.

performance evaluations are not enough to justify promotion decisions at ServiCo. The “outstanding performance” rating is positive and significant ($P < .05$). In addition, demographic characteristics do not seem to be significant in explaining variation of promotion rates.⁴¹

DISCUSSION

This article empirically examines the relationship between performance evaluations and two key employee outcomes—salary increase decisions and promotions—in one large service organization in the United States, which introduced a formal performance evaluation program to try to encourage “constructive dialogue between employees and supervisors” and to “make compensation decisions.” Theoretically, I claim that the use of merit-based reward systems such as this one can result in organizations introducing bias at two different stages (as summarized in fig. 1). The first stage is the performance evaluation stage, where performance evaluation bias can occur; this implies that the performance rating process is affected by gender, race, or nationality bias.⁴² Even if one assumes that there is no bias in this first stage, bias can be introduced in the second stage, the link between performance evaluations and employee outcomes. This is what I have termed performance-reward bias. In this article, I have tested the two scenarios in which such bias can be detected. The first is when there is disparity in salary increases by race and gender net of performance ratings (proposition 1). The second is when there is disparity in the effect of ratings on salary increases by gender and race (proposition 2). In my analyses, I find empirical evidence that both of these scenarios exist at the organization under study, leading me to reject the meritocratic claims for this performance-reward system.

Figure 4 summarizes what transpires in this organization. Even in a work organization that institutionally values and supports the allocation of compensation on the basis of merit, I show bias in the translation of performance evaluation scores into amounts of salary increases over time: different salary increases are granted for observationally equivalent employees (i.e., those in the same job and work unit, with the same supervisor

⁴¹ With only 262 promotion events in the sample and the numerous fixed effects included in the equation, the promotion model lacks statistical power. So it can be argued that these nonsignificant race and gender effects are due to either lack of race and gender differences in promotions or lack of statistical power in the model.

⁴² As indicated above, there is ample evidence of the existence of performance evaluation bias (for a review, see Bartol [1999], Elvira and Town [2001], Roth et al. [2003], and McKay and McDaniel [2006]).

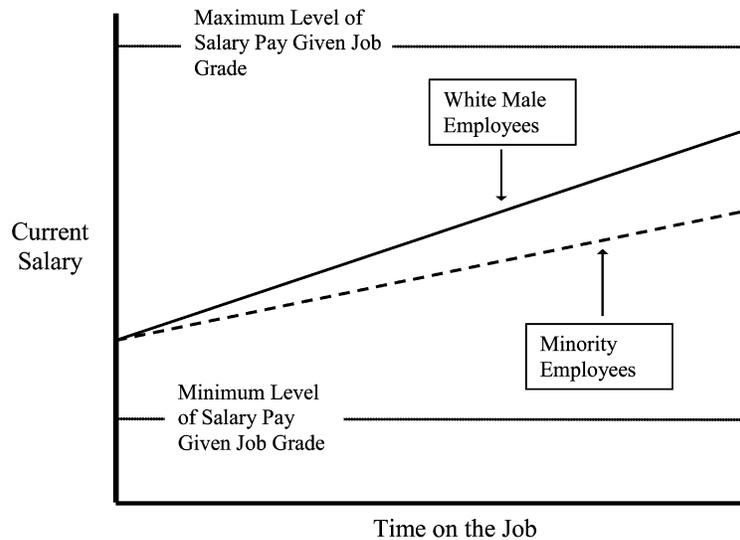


FIG. 4.—Summary graph of the findings, assuming the same starting annual salary for all employees. Note that this figure provides a stylized representation of my findings—it is not a plot of the model parameters.

and same human capital) who receive the same performance evaluation scores.

In order to understand how such performance-reward bias could occur in this organization, I interviewed some personnel at different levels in the organizational hierarchy. I found that this bias can be introduced at two points in the performance appraisal process. First, it can occur when the head of a unit (or head of supervisors) recommends to HR a particular annual salary increase amount for a given employee (fig. 2, step 2). The unit head may put forth a lower salary raise for an equally performing minority employee than for a nonminority employee, yielding a lower average for minorities than nonminorities in reward recommendations going to HR. Second, it can occur when an HR personnel member makes the decision to approve or reject a given salary increase recommendation generated by the head of the unit (fig. 2, step 3). HR may reject more minority rewards than nonminority rewards, even if unit heads are unbiased in their recommendations.

While investigating step 2 of the appraisal process was beyond the scope of this study, I conducted interviews with several HR personnel to learn how they make their decisions. These interviews suggested that the bias does not likely occur at step 3: HR members do not turn down any bonus recommendation decisions, nor do they adjust the magnitudes recom-

mended by unit heads. The HR director explained that “in a given year we might consider thousands of salary review proposals; we tend to tell our HR managers to look for high performers when it comes to approving such salary increases.”⁴³ All HR compensation office members I interviewed reiterated the explicit message in the HR policy manual, that all salary increases must be based on the individual’s job performance. When I asked whether they look at any information about the employee besides the short performance evaluation form filled out by the unit head, one HR staff member stated that “there is not much time to do so; we look at the form and ensure that the unit has submitted the required documentation for salary increase approval.” This required documentation consists of a letter included with the salary award recommendation. Altogether, this suggests that the performance-reward bias is likely introduced at step 2 of the appraisal process by the higher-level actors who recommend, based on the employee performance evaluations, whether a raise should be granted as well as the amount of the raise. This bias is not corrected by HR in step 3. I found step 2 in the performance appraisal process to be the least transparent; the heads of units and of supervisors are not accountable for their decisions regarding salary increase amounts either.

Why Is There Performance-Reward Bias?

Many social mechanisms can explain why, in an organization such as this, employees and administrators may be unaware of the existence of performance-reward bias. On the basis of my research, I believe there are two main theoretical mechanisms accounting for the performance-reward bias in this particular organization. The first main mechanism can be found in the social-psychological theory about the role of *accountability* in reducing (and perhaps even eliminating) bias (e.g., Tetlock 1983, 1985; Tetlock and Kim 1987; Lerner and Tetlock 1999). According to this mechanism, when decision makers know they will be held accountable for making fair decisions, bias is less likely to occur. This work on accountability points to two important conclusions. First, accountability motivates decision makers to process information and make decisions in more analytical and complex ways, which can help reduce judgmental biases. Second, the timing of accountability is crucial, because accountability appears to be much more effective in preventing rather than in reversing biases (Tetlock 1985, p. 233). In this large organizational setting, accountability is more salient at the first stage of the performance appraisal pro-

⁴³ An average of 1,149 proposals per year were reviewed and accepted, according to the data collected for this study.

cess, where there is no evidence of ascriptive bias. This is consistent with the notion that formalization reduces bias and increases equity—a central notion in the development of employers' compliance with Title VII and of internal labor markets and the HR profession since the late 1960s (see Anderson and Tomaskovic-Devey 1995; Huffman and Velasco 1997; Reskin and McBrier 2000; Stinchcombe 2001; Kalev et al. 2006; Dobbin and Kelly 2007).⁴⁴ I find that ascriptive bias exists only in the less formalized second stage of performance appraisals, where administrators are not accountable for their decisions regarding the amounts of salary increases.⁴⁵

A second, complementary theoretical mechanism is *transparency*. The formalization of practices that increase transparency can make disparities more noticeable and therefore more easily corrected. This is closely related to the theoretical claim about the likely existence of *information-processing bias* in large organizations. Previous experimental research has shown how individuals are less able to perceive gender or race discrimination on a personal level than on an organizational or societal level. Crosby et al. (1986) demonstrate that this phenomenon is feasible in part because of an information-processing bias—that is, the perception of discrimination processes is more difficult when one makes case-by-case comparisons than when one encounters information in the aggregate.⁴⁶ My findings that the most visible aspects of employee career outcomes—such as salary increases (regardless of quantity) and promotions—are not subject to the performance-reward bias process conforms to this information-

⁴⁴ As one of the anonymous reviewers of this article pointed out, there is also an implicit (neo-Weberian) institutional argument to make here about the effects of ongoing rationalization in modern work organizations: bureaucracies function properly and curtail discrimination as long as their doings are clearly visible and administrators are accountable for their decisions.

⁴⁵ Recent work on accountability (e.g., Lerner and Tetlock 1999) has moved away from a pure focus on improving decision making to a more balanced view of how accountability can actually *introduce* bias. Specifically, if decision makers know the conclusion their audience wants them to reach, they tend to use processes that yield that conclusion so they can give their audience what it wants. In this case, supervisors may be aware that their performance ratings can easily be evaluated for gender and race bias and may therefore give a “balanced” set of ratings across demographic groups. The result would be that supervisors end up giving fabricated ratings just to keep their employees and superiors happy. Although the ratings appear unbiased, the salary increases based on these ratings reflect either information shared through social interaction or “dis-counting” due to shared knowledge that a rating of 5 for a woman, e.g., is not the same as a 5 for a man, because everyone is forced to look fair in their performance ratings.

⁴⁶ Crosby et al. (1986) show evidence of the existence of cognitive bias in perceptions of discrimination. Their experiment demonstrates the importance of formatting for the perception of discrimination: subjects perceived less discrimination when they encountered relevant information in small chunks than when they saw the total picture all at once.

awareness argument. This is in contrast to the finding that females and minorities are then disadvantaged when it comes to decisions about the *amount* by which their compensation is increased every year, which is typically unobservable or unknown to the rest of employees. The invisibility of salary increase amounts eliminates concrete salary comparisons among employees and thus has the potential to mask unfairness in the performance-compensation link in organizations. These findings parallel those of empirical and theoretical work in organizational behavior (Crosby 1982, 1984; for a review, see Major et al. [2002]).⁴⁷

In this particular research setting, several features make these salary gaps less pronounced and possibly even invisible to employees and administrators. First, in this organization, performance ratings govern the decision-making process for salary increases, as the coefficients for performance ratings are the only significant predictors of decisions to increase salary. This is consistent with the policy manual available to employees, which states that salary increases should only be given to high-performing employees. Second, because most salary increases in a given year are quite low, salary disparities among employees are so small that they are not noticeable overall. In this organization, salary increases never exceeded 8% of the base salary, and most bonuses given as a lump sum were small—up to \$500, depending on the year and the unit under study. And last, employees stay in the organization for about 2.65 years on average; this relatively short tenure of employment may also minimize the long-term impact of the small differences in salary increases.

Finally, this study provides some field evidence of the “lower minimum standards and higher ability standards” argument proposed by Biernat and Kobrynowicz (1997): even when it may be easier for women and minorities to meet low standards, these employees are still subject to higher ability standards than white men, and consequently they must work harder to prove that their ability is similar or greater. In my research setting, because I find no significant disparities in recommendations to grant a salary increase by race or gender, I argue that women and minorities are equally as likely as white men to meet the “minimum standards” in order to be recommended for a salary increase. However, after controlling for ratings (as well as jobs, work units, and supervisors), I find that women’s and minorities’ performance appraisals are significantly discounted, meaning that they need to work harder and obtain higher

⁴⁷ This concept of visibility is also related to the concept of accountability. Thus, the formalization of practices can curtail discrimination when actors are accountable for their decisions at the different stages in the design and implementation of organizational practices and their doings are clearly visible. I am thankful to one of the anonymous reviewers of this article for pointing out the link between these concepts (as well as the concept of information-processing bias).

performance scores in order to receive similar salary increases to white men. This finding is also consistent with the work on “double standards” (Foschi 1992, 1996; Foschi, Lai, and Sigerson 1994), which finds, in effect, that women (and minorities) need to display a higher level of performance before decision makers conclude that they are equally as competent as men (and consequently award them the same salary increases).⁴⁸ Distinguishing among these different explanations (and other feasible theoretical mechanisms operating in other settings) remains a task for future research.

Limitations and Future Research

I believe that this research can be extended in several promising directions toward an understanding of race, gender, and meritocracy in organizational careers. The first and most obvious extension involves continued testing of the relationship between performance evaluations and compensation. While previous research has extensively addressed gender and racial biases through the performance-rating process, this article focuses on the performance-reward process. I demonstrate that this process, which has not been studied in depth before, may have become an important organizational process accounting for the persistence of gender, race, or other non-performance-related demographic differences in wages and career attainment within organizations. In my analyses, I do not find that initial salary allocations or the distribution of appraisals differ much by the race and gender of the employee. However, in other settings this might not be the case: there may also be bias in the process of evaluating employees, for example, exacerbating the bias in salary increases reported here. Future quantitative and qualitative studies should take a comprehensive approach and examine the multiple stages where bias can be introduced, contributing to the persistent growth in the cumulative disadvantage of women and minorities (Jacobs 1995).

The second extension of this research is closely related. In this article, I emphasize the effect of performance on salary increases and promotions only. Future research should examine other organizational processes at work in central aspects of the employment relationship, such as benefits, other types of promotions, and unit/job transfers. Emphasis should also

⁴⁸ Even though I am not able to measure ability directly in this study, I can measure performance in a given job and work unit. An alternative explanation for the finding that women and minorities receive lower salary increases net of performance evaluations could be that white men are truly more productive than women and minorities are and that the administrators recommending salary increases recognize and adjust for this fact. This would imply that white males are receiving lower performance ratings than their true performance merits—but we know from the experimental literature that this is not the case (see Biernat and Kobrynowicz 1997).

be given to other nonmonetary rewards in organizations, such as advantages in training opportunities, access to organizational resources, responsibility and authority, and, more generally, to the distribution of both advantages and disadvantages in organizations (e.g., DiTomaso et al. 2007).

Equally relevant to the literature on employee careers across organizations is the study of how previous employers' evaluations impact hiring decisions for employees across organizations. Extensive literature on hiring, including my own previous work, has examined the determinants of organizations' hiring practices, from human capital theories to social-network-focused perspectives (for a review, see Fernandez, Castilla, and Moore [2000] and Castilla [2005]). However, little is known about the ways in which employees' past performance experiences influence their future careers. Given that I do not have prehire data, I cannot evaluate in this article whether discrimination also occurs in the matching process at the point of hire. Without such data, many of the intervening mechanisms, as well as the long-term effects of past performance evaluations, are left open to future investigation on the connections among human resource practices, performance, and outcomes.

In this article, I was interested in testing whether demographic features are significant variables in predicting salary growth, after controlling for the level of performance rating in a given job in a given work unit. It is important to note that, under certain organizational arrangements, even when salary increases are not identical for minority employees, the unequal allocation of such increases might never result in large unequal outcomes. Thus, the use of performance-based bonuses may even appear to be quite meritocratic and unbiased. As I indicated above, several features of ServiCo may have contributed to employees and administrators' lack of awareness of this performance-reward bias. First, according to ServiCo's policy manual, performance ratings are the primary basis for any employee salary increase decisions. Second, since salary increases are low, any salary disparities among employees are unlikely to be noticeable. Finally, given the short average tenure of employees at ServiCo, few employees happen to be in the organization long enough for the wage differences to appear substantial. Future research should study whether the same evidence is found in organizations where salary increases are large or where seniority accounts for salary increases. Studies such as these could help to shed light on which types of performance management systems favor equality in the allocation of rewards today.

In this company setting, supervisors evaluate employees using dyadic performance evaluations. Future studies should also consider workplaces where evaluations are less dyadic. For example, researchers could look at organizations where a higher number of individuals participate in the

formal evaluation of employees. In fact, some of these less-dyadic options are already being implemented in companies. These performance appraisal processes are of particular interest in professional settings where the same employee is evaluated by co-workers, superiors, and subordinates—a 360-degree-style performance appraisal system. By looking at the mechanisms behind these evaluations as they impact key employee outcomes, future research can further our understanding of how organizational career processes can remedy biases in work organizations.

Finally, as with any study on this topic, there remains the question of the generalizability of these results. Although I study only a single organization, it is worth noting that this organization's human resource practices are not very different from those of current organizations that have chosen to adopt merit-based practices for distributing rewards among employees. According to Noe et al. (2006, p. 504), some type of merit-pay program "exists in almost all organizations (although evidence on merit pay effectiveness is surprisingly scarce)." Under the new system of market-driven employment practices (Cappelli 1999), organizations introduce performance-reward programs and other merit-based practices—*perhaps* in the hope of ensuring that rewards are allocated meritocratically and eliminating unfairness (Jackson 1998). However, this article focuses on the case of an organization that introduced a new performance appraisal process in order to encourage the development of its employees as well as to provide a basis for compensation decisions. And yet the formalization of this performance system created additional opportunities for discretion and biases to emerge, ultimately resulting in compensation differentials for women and minorities over time. Future research should take steps toward studying whether the patterns discovered in this organization can be generalized to other settings, by analyzing complete personnel data in other private organizations such as this one.

CONCLUSION

Organizations are increasingly using performance-pay programs that link the performance of employees to their compensation over time. Perhaps implicit in the creation and use of these programs is the presumption that today's organizational practices are based on merit and consequently that a significant positive relationship between performance, wages, and wage growth is institutionally valued and strongly supported. But since merit-based reward systems often introduce a sequence of organizational processes or routines, I argue that the nature and implementation of these programs may make it possible for bias and discriminatory judgments to occur at several stages. This article focuses on one of these stages, namely,

the link between performance evaluations and wage determination. I identify and provide evidence of what I call the performance-reward bias—the form of discrimination that happens when employers undervalue the work of certain minority employees.

Performance-reward bias is independent of other processes generating ascriptive inequality, as described in the Petersen and Saporta (2004) article on employer discrimination processes. Even assuming that (1) women and minorities are equally sorted into jobs, (2) women and minorities receive the same starting salary within a given occupation, within a given establishment, and (3) female- and minority-dominated occupations with equal skill requirements and other wage-relevant factors are valued the same as white-male-dominated occupations, I still find that the work of women and minorities can be discounted in organizations over time. This performance-reward bias is a new form of valuative discrimination, because once merit is measured in the appraisal process, women and minority employees still receive different rewards for the same merit scores as white men (after controlling for job, work unit, supervisor and other relevant human capital characteristics). This bias is also independent of the fact that the performance rating process itself might be affected by gender, race, or nationality bias.

This finding is of substantive significance because it demonstrates a critical challenge faced by employers who adopt merit-based practices to fairly reward and motivate their employees. Ironically, although these merit-reward policies create the appearance of meritocracy, this study shows that the less formalized, less transparent, and less accountable stages of the performance appraisal process can actually create a greater opportunity for subtle ascriptive biases to emerge, negatively affecting the fair distribution of rewards among employees in a way that is more or less invisible to everyone in the organizational setting.

Previous studies looking at wages and careers within organizations have not included performance or merit in their models, nor have they examined in depth the many organizational processes and stages at work behind these employee outcomes. The extensive research on the role of organizations in the distribution of salaries and rewards among employees commits the same omission (for a review, see Petersen and Saporta [2004], Phillips [2005], and Roth [2006]). This article is intended to be the first step toward correcting this imbalance in the literature on organizations and stratification and toward unpacking what is actually happening inside an organizational practice described as meritocratic. Future research should continue examining how the formalization and implementation of overall organizational merit-based practices may affect an individual's structures of opportunity and attainment in contemporary organizations.

REFERENCES

- Anderson, Cynthia D., and Donald Tomaskovic-Devey. 1995. "Patriarchal Pressures: An Exploration of Organizational Processes That Exacerbate and Erode Gender Earnings Inequality." *Work and Occupations* 22 (3): 328–57.
- Barnett, William P., James N. Baron, and Toby E. Stuart. 2000. "Avenues of Attainment: Occupational Demography and Organizational Careers in the California Civil Service." *American Journal of Sociology* 106 (1): 88–144.
- Baldi, Stéphane, and Debra B. McBrier. 1997. "Do the Determinants of Promotion Differ for Blacks and Whites?" *Work and Occupations* 24 (4): 478–97.
- Baron, James N. 1984. "Organizational Perspectives on Stratification." *Annual Review of Sociology* 10:37–69.
- Baron, James N., and William T. Bielby. 1980. "Bringing the Firms Back In: Stratification, Segmentation, and the Organization of Work." *American Sociological Review* 45 (5): 737–65.
- Baron, James N., and Andrew E. Newman. 1990. "For What It's Worth: Organizations, Occupations, and the Value of Work Done by Women and Non-whites." *American Sociological Review* 55 (2): 155–75.
- Baron, Reuben M., and David A. Kenny. 1986. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic and Statistical Considerations." *Journal of Personality and Social Psychology* 51:1173–82.
- Bartol, Kathryn M. 1999. "Gender Influences on Performance Evaluations." Pp. 165–78 in *Handbook of Gender and Work*, edited by Gary N. Powell. Thousand Oaks, Calif.: Sage.
- Bielby, William T., and James N. Baron. 1986. "Men and Women at Work: Sex Segregation and Statistical Discrimination." *American Journal of Sociology* 91 (4): 759–99.
- Biernat, Monica, and Diane Kobrynowicz. 1997. "Gender- and Race-Based Standards of Competence: Lower Minimum Standards but Higher Ability Standards for Devalued Groups." *Journal of Personality and Social Psychology* 72 (3): 544–57.
- Bohlander, George, and Scott Snell. 2007. *Managing Human Resources*, 14th ed. Mason, Ohio: Thomson South-Western.
- Bridges, William P., and Robert L. Nelson. 1989. "Markets in Hierarchies: Organizational and Market Influences on Gender Inequality in a State Pay System." *American Journal of Sociology* 95:616–58.
- Burke, Elizabeth. 2005. "2005 Reward Programs and Incentive Compensation: Survey Report." Society for Human Resource Management. <http://www.shrm.org/surveys>
- Campbell, Donald J., Kathleen M. Campbell, and Ho-Beng Chia. 1998. "Merit Pay, Performance Appraisal, and Individual Motivation: An Analysis and Alternative." *Human Resource Management* 37 (2): 131–46.
- Cappelli, Peter. 1999. *The New Deal at Work*. Boston: Harvard Business School Press.
- Cappelli, Peter, Laurie Bassi, Harry Katz, David Knoke, Paul Osterman, and Michael Useem. 1997. *Change at Work*. New York: Oxford University Press.
- Castilla, Emilio J. 2005. "Social Networks and Employee Performance in a Call Center." *American Journal of Sociology* 110 (5): 1243–83.
- . 2007. *Dynamic Analysis in the Social Sciences*. Oxford: Elsevier, Academic Press.
- Cleveland, Jeanette N., Kevin R. Murphy, and Richard E. Williams. 1989. "Multiple Uses of Performance Appraisal: Prevalence and Correlates." *Journal of Applied Psychology* 74 (1): 130–35.
- Compensation Resources, Inc. 2004. "How Often Does Your Organization Conduct Performance Evaluations?" Upper Saddle River, N.J.: Compensation Resources. <http://compensationresources.com/press-room/how-often-does-your-organization-conduct-performance-evaluations-.php>. Accessed April 1, 2006.

American Journal of Sociology

- Correll, Shelley J. 2001. "Gender and the Career Choice Process: The Role of Biased Self-Assessments." *American Journal of Sociology* 106:1691–1730.
- Cox, David R. 1972. "Regression Models and Life Tables (with Discussion)." *Journal of the Royal Statistical Society*, ser. B. 34 (2): 187–220.
- . 1975. "Partial Likelihood." *Biometrika* 62 (2): 269–76.
- Crosby, Faye. 1982. *Relative Deprivation and Working Women*. New York: Oxford University Press.
- . 1984. "Relative Deprivation in Organizational Settings." *Research in Organizational Behavior* 6:51–93.
- Crosby, Faye, Susan Clayton, Olaf Alksnis, and Kathryn Hemker. 1986. "Cognitive Biases in the Perception of Discrimination: The Importance of Format." *Sex Roles* 14 (11/12): 637–46.
- DiTomaso, Nancy, Corinne Post, D. Randall Smith, George F. Farris, and Rene Cordero. 2007. "Effects of Structural Position on Allocation and Evaluation Decisions for Scientists and Engineers in Industrial R&D." *Administrative Science Quarterly* 52 (2): 175–207.
- Dobbin, Frank, and Erin Kelly. 2007. "How to Stop Harassment: The Professional Construction of Legal Compliance in Organizations." *American Journal of Sociology* 112 (4): 1203–43.
- Elvira, Marta M., and Mary E. Graham. 2002. "Not Just a Formality: Pay System Formalization and Sex-Related Earnings Effects." *Organization Science* 13 (6): 601–17.
- Elvira, Marta M., and Robert Town. 2001. "The Effects of Race and Worker Productivity on Performance Evaluations." *Industrial Relations* 40 (4): 571–90.
- Elvira, Marta M., and Christopher D. Zatzick. 2002. "Who's Displaced First? The Role of Race in Layoff Decisions." *Industrial Relations* 41 (2): 329–61.
- England, Paula. 1992. *Comparable Worth: Theories and Evidence*. Hawthorne, N.Y.: Aldine de Gruyter.
- Fernandez, Roberto M., Emilio J. Castilla, and Paul Moore. 2000. "Social Capital at Work: Networks and Employment at a Phone Center." *American Journal of Sociology* 105 (5): 1288–1356.
- Fernandez, Roberto M., and Isabel Fernandez-Mateo. 2006. "Networks, Race and Hiring." *American Sociological Review* 71 (1): 42–71.
- Fernandez, Roberto M., and M. Lourdes Sosa. 2005. "Gendering the Job: Networks and Recruitment at a Call Center." *American Journal of Sociology* 111:859–904.
- Foschi, Martha. 1992. "Gender and Double Standards for Competence." Pp. 181–207 in *Gender, Interaction, and Inequality*, edited by Cecilia Ridgeway. New York: Springer-Verlag.
- . 1996. "Double Standards in the Evaluation of Men and Women." *Social Psychology Quarterly* 59:237–54.
- Foschi, Martha, Larissa Lai, and Kirsten Sigerson. 1994. "Gender and Double Standards in the Assessment of Job Applicants." *Social Psychology Quarterly* 57: 326–39.
- Frank, Robert H., and Phillip J. Cook. 1995. *The Winner-Take-All Society*. New York: Free Press.
- Gerhart, Barry A., and Sara L. Rynes. 2003. *Compensation: Theory, Evidence, and Strategic Implications*. Thousand Oaks, Calif.: Sage Publications.
- Gorman, Elizabeth. 2005. "Gender Stereotypes, Same-Gender Preferences, and Organizational Variation in the Hiring of Women: Evidence from Law Firms." *American Sociological Review* 70:702–28.
- Griffeth, Rodger W., and Arthur G. Bedeian. 1989. "Employee Performance Evaluations: Effects of Ratee Age, Rater Age, and Ratee Gender." *Journal of Organizational Behavior* 10 (1): 83–91.
- Hall, Francine S., and Douglas T. Hall. 1976. "Effects of Job Incumbents' Race and

Gender, Race, and Meritocracy

- Sex on Evaluations of Managerial Performance." *Academy of Management Journal* 19 (3): 476–81.
- Hamner, W. Clay, Jay S. Kim, Lloyd Baird, and William J. Bigoness. 1974. "Race and Sex as Determinants of Ratings by Potential Employers in a Simulated Work-Sampling Task." *Journal of Applied Psychology* 59 (6): 705–12.
- Hausman, Jerry. 2001. "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left." *Journal of Economic Perspectives* 15 (4): 57–67.
- Hodson G., J. F. Dovidio, and S. L. Gaertner. 2002. "Processes in Racial Discrimination: Differential Weighting of Conflicting Information." *Personality and Social Psychology Bulletin* 28:460–71.
- Huffman, Matt L., and Steven C. Velasco. 1997. "When More Is Less." *Work and Occupations* 24 (2): 214–45.
- IOMA (Institute of Management and Administration). 2000. "Towers Perrin Reveals How to Design the Most Effective Incentive Plans." *IOMA's Report on Salary Surveys*, March: p. 14.
- Jackson, Robert M. 1998. *Destined for Equality: The Inevitable Rise of Women's Status*. Cambridge, Mass.: Harvard University Press.
- Jacobs, Jerry A. 1989. *Revolving Doors: Sex Segregation and Women's Careers*. Stanford, Calif.: Stanford University Press.
- . 1995. *Gender Inequality at Work*. Beverly Hills, Calif.: Sage.
- Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87 (5): 972–90.
- Judd, Charles M., and David A. Kenny. 1981. "Process Analysis: Estimating Mediation in Treatment Evaluations." *Evaluation Review* 5:602–19.
- Kalev, Alexandra, Frank Dobbin, and Erin Kelly. 2006. "Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies." *American Sociological Review* 71:589–617.
- Kochan, Thomas A., Harry Katz, and Robert McKersie. 1986. *The Transformation of American Industrial Relations*. New York: Basic Books.
- Lawler, Edward. 2003. "Reward Practices and Performance Management System Effectiveness." *Organizational Dynamics* 32 (4): 396–404.
- Lazear, Edward. 1998. *Personnel Economics for Managers*. New York: John Wiley & Sons.
- Lee, Dennis M., and Kenneth M. Alvares. 1977. "Effects of Sex on Descriptions and Evaluations of Supervisory Behavior in a Simulated Industrial Setting." *Journal of Applied Psychology* 62 (4): 405–11.
- Lee, Lung-Fei. 1979. "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables." *Econometrica* 47:977–96.
- . 1983. "Notes and Comments: Generalized Econometric Models with Selectivity." *Econometrica* 51 (2): 507–13.
- Lee, Lung-Fei, and G. S. Maddala. 1985. "Sequential Selection Rules and Selectivity in Discrete Choice Econometric Models." *Econometric Methods and Applications* 2:311–29.
- Lee, Lung-Fei, G. S. Maddala, and R. P. Trost. 1980. "Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity." *Econometrica* 48:491–503.
- Lerner, Jennifer S., and Philip E. Tetlock. 1999. "Accounting for the Effects of Accountability." *Psychological Bulletin* 125 (2): 255–75.
- Liang, Kung-Yee, and Scott L. Zeger. 1986. "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika* 73:13–22.
- Liden, Robert C., Sandy J. Wayne, and Dean Stilwell. 1993. "A Longitudinal Study on the Early Development of Leader-Member Exchanges." *Journal of Applied Psychology* 78 (4): 62–74.

American Journal of Sociology

- London, Manuel, and Stephen A. Stumpf. 1983. "Effects of Candidate Characteristics on Management Promotion Decisions: An Experimental Study." *Personnel Psychology* 36:241–59.
- Major, Brenda, Richard H. Gramzow, Shannon K. McCoy, Shana Levin, Toni Schmader, and Jim Sidanius. 2002. "Perceiving Personal Discrimination: The Role of Group Status and Legitimizing Ideology." *Journal of Personality and Social Psychology* 82 (3): 269–82.
- Marsden, Peter V. 1994a. "The Hiring Process: Recruitment Methods." *American Behavioral Scientist* 37 (7): 979–91.
- . 1994b. "Selection Methods in U.S. Establishments." *Acta Sociologica* 37 (3): 287–301.
- Martocchio, Joseph J. 2004. *Strategic Compensation: A Human Resource Management Approach*, 4th ed. Upper Saddle River, N.J.: Prentice Hall.
- Mathis, Robert L., and John H. Jackson. 2003. *Human Resource Management*, 10th ed. Mason, Ohio: South-Western.
- McKay, Patrick F., and Michael A. McDaniel. 2006. "A Re-examination of Black-White Mean Differences in Work Performance." *Journal of Applied Psychology* 91 (3): 538–54.
- McPherson, Miller, Lynn Smith-Lovin, and James Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–44.
- Medoff, James L., and Katharine Abraham. 1981. "Are Those Paid More Really More Productive? The Case of Experience." *Journal of Human Resources* 16 (2): 186–216.
- Milkovich, George T., and Jerry M. Newman. 2004. *Compensation*, 8th ed. Boston: McGraw-Hill/Irwin.
- Miller, Stephen. 2006. "Employers Focus More on Performance-Based Rewards." Society for Human Resource Management. http://www.shrm.org/rewards/library_published/compensation/nonIC/CMS_018315.asp. Accessed April 1, 2006.
- Mobley, William H. 1982. "Supervisor and Employee Race Effects on Performance Appraisals: A Field Study of Adverse Impact and Generalizability." *Academy of Management Journal* 25 (3): 598–606.
- Nelson, Robert L., and William P. Bridges. 1999. *Legalizing Gender Inequality: Courts, Markets, and Unequal Pay for Women in America*. New York: Cambridge University Press.
- Noe, Raymond A., John R. Hollenbeck, Barry Gerhart, and Patrick M. Wright. 2006. *Human Resource Management: Gaining a Competitive Advantage*. New York: McGraw-Hill Irwin.
- Norton, Michael I., Joseph A. Vandello, and John M. Darley. 2004. "Casuistry and Social Category Bias." *Journal of Personality and Social Psychology* 87:817–31.
- Ospina, Sonia. 1996. *Illusions of Opportunity: Employee Expectations and Workplace Inequality*. Ithaca, N.Y.: ILR Press/Cornell University Press.
- Osterman, Paul. 1999. *Securing Prosperity: The American Labor Market; How It Has Changed and What to Do about It*. Princeton, N.J.: Princeton University Press.
- Osterman, Paul, Thomas Kochan, Richard Locke, and Michael Piore. 2001. *Working in America: A Blueprint for the New Labor Market*. Cambridge, Mass.: MIT Press.
- Pelled, Lisa Hope, Kathleen M. Eisenhardt, and Katherine R. Xin. 1999. "Exploring the Black Box: An Analysis of Work Group Diversity, Conflict, and Performance." *Administrative Science Quarterly* 44 (1): 1–29.
- Petersen, Trond, and Laurie Morgan. 1995. "Separate and Unequal: Occupation-Establishment Segregation and the Gender Wage Gap." *American Journal of Sociology* 101 (2): 329–65.
- Petersen, Trond, and Ishak Saporta. 2004. "The Opportunity Structure for Discrimination." *American Journal of Sociology* 109 (4): 852–902.
- Petersen, Trond, Ishak Saporta, and Mark-David Seidel. 2000. "Offering a Job: Meritocracy and Social Networks." *American Journal of Sociology* 106 (3): 763–817.

Gender, Race, and Meritocracy

- Pfeffer, Jeffrey. 1994. *Competitive Advantage through People: Unleashing the Power of the Work Force*. Boston: Harvard Business School Press.
- Phillips, Damon J. 2005. "Organization Genealogies and the Persistence of Gender Inequality: The Case of Silicon Valley Law Firms." *Administrative Science Quarterly* 50 (3): 440–72.
- Price, James L. 1977. *The Study of Turnover*. Ames: Iowa State University Press.
- Pulakos, Elaine D., Leonard A. White, Scott H. Oppler, and Walter C. Borman. 1989. "Examination of Race and Sex Effects on Performance Ratings." *Journal of Applied Psychology* 76 (5): 770–80.
- Reskin, Barbara F. 1993. "Sex Segregation in the Workplace." *Annual Review of Sociology* 19:241–70.
- . 1998. *The Realities of Affirmative Action in Employment*. Washington, D.C.: American Sociological Association.
- . 2000. "The Proximate Causes of Employment Discrimination." *Contemporary Sociology* 29 (2): 319–28.
- Reskin, Barbara F., and Debra B. McBrier. 2000. "Why Not Ascription? Organizations' Employment of Male and Female Managers." *American Sociological Review* 65 (2): 210–33.
- Ridgeway, Cecilia. 1997. "Interaction and the Conservation of Gender Inequality: Considering Employment." *American Sociological Review* 62 (2): 218–35.
- Rosenfeld, Rachel A. 1992. "Job Mobility and Career Processes." *Annual Review of Sociology* 18:39–61.
- Roth, Louise M. 2006. *Selling Women Short: Gender and Money on Wall Street*. Princeton, N.J.: Princeton University Press.
- Roth, Philip L., Allen I. Huffcutt, and Philip Bobko. 2003. "Ethnic Group Differences in Measures of Job Performance: A New Meta-analysis." *Journal of Applied Psychology* 88 (4): 694–706.
- Schmitt, Neal, and Martha Lappin. 1980. "Race and Sex as Determinants of the Mean and Variance of Performance Ratings." *Journal of Applied Psychology* 65 (4): 428–35.
- Stinchcombe, Arthur L. 2001. *When Formality Works: Authority and Abstraction in Law and Organizations*. Chicago: University of Chicago Press.
- Tetlock, Philip E. 1983. "Accountability and the Perseverance of First Impressions." *Social Psychology Quarterly* 46 (4): 285–92.
- . 1985. "Accountability: A Social Check on the Fundamental Attribution Error." *Social Psychology Quarterly* 48 (3): 227–36.
- Tetlock, Philip E., and Jae I. Kim. 1987. "Accountability and Judgment Processes in a Personality Prediction Task." *Journal of Personality and Social Psychology* 52 (4): 700–709.
- Thompson, Duane E., and Toni A. Thompson. 1985. "Task-Based Performance Appraisal for Blue-Collar Jobs: Evaluation of Race and Sex Effects." *Journal of Applied Psychology* 70 (4): 747–53.
- Tomaskovic-Devey, Donald. 1993. *Gender and Racial Inequality at Work: The Sources and Consequences of Job Segregation*. Ithaca, N.Y.: ILR Press.
- Tsui, Anne S., and Barbara A. Gutek. 1984. "A Role Set Analysis of Gender Differences in Performance, Affective Relationships, and Career Success of Industrial Middle Managers." *Academy of Management Journal* 27 (3): 619–35.
- Tsui, Anne S., and Charles A. O'Reilly III. 1989. "Beyond Simple Demographic Effects: The Importance of Relational Demography in Superior-Subordinate Analysis." *Academy of Management Journal* 32 (2): 402–23.
- Tuma, Nancy B. 1976. "Rewards, Resources, and the Rate of Mobility: A Nonstationary Multivariate Stochastic Model." *American Sociological Review* 41:330–38.
- Turban, Daniel B., and Allan P. Jones. 1988. "Supervisor-Subordinate Similarity: Types, Effects, and Mechanisms." *Journal of Applied Psychology* 73 (2): 228–34.

American Journal of Sociology

- Uhlmann, Eric L., and Geoffrey L. Cohen. 2005. "Constructed Criteria: Redefining Merit to Justify Discrimination." *Psychological Science* 16 (6): 474–80.
- Valian, Virginia. 1998. *Why So Slow? The Advancement of Women*. Cambridge, Mass.: MIT Press.
- Wagner, W. Gary, Jeffrey Pfeffer, and Charles A. O'Reilly III. 1984. "Organizational Demography and Turnover in Top Management Groups." *Administrative Science Quarterly* 29 (1): 74–92.
- Williams, Katherine Y., and Charles A. O'Reilly III. 1998. "Demography and Diversity in Organizations: A Review of 40 Years of Research." *Research in Organizational Behavior* 20:77–140.
- Yammarino, Francis J., and Alan J. Dubinsky. 1988. "Employee Responses: Gender- or Job-Related Differences?" *Journal of Vocational Behavior* 32:366–83.