Fall 1-8-2010

The Impact of Performance Ratings on Federal Personnel Decisions

Seong Soo Oh

THE IMPACT OF PERFORMANCE RATINGS ON FEDERAL PERSONNEL DECISIONS

A Dissertation
Presented to
The Academic Faculty

by

Seong Soo Oh

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Public Policy

Georgia State University and Georgia Institute of Technology
December 2009

THE IMPACT OF PERFORMANCE RATINGS ON FEDERAL PERSONNEL DECISIONS

Approved by:

Dr. Gregory B. Lewis, Advisor
Andrew Young School of Policy Studies
Georgia State University

Dr. Gordon Kingsley
School of Public Policy
Georgia Institute of Technology

Dr. William L. Waugh Jr.
Andrew Young School of Policy Studies
Georgia State University

Dr. Julia E. Melkers
School of Public Policy
Georgia Institute of Technology

Dr. Katherine G. Willoughby
Andrew Young School of Policy Studies
Georgia State University

Date Approved: August 31, 2009

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

CPDF	Central Personnel Data File
CSRA	Civil Service Reform Act
EPA	Environmental Protection Agency
GAO	Government Accountability Office
GS	General Schedule
NASA	National Aeronautics and Space Administration
OLS	Ordinary Least Square
OPM	Office of Personnel Management
PATCO	Professional, Administrative, Technical, Clerical, and Other
PMRS	Performance Management and Recognition System

SUMMARY

Can pay-for-performance increase the motivation of public employees? By providing a basis for personnel decisions, particularly linking rewards to performance, performance appraisals aim to increase employees' work motivation and ultimately to improve their work performance and organizational productivity. With the emphasis on results-oriented management, performance appraisals have become a key managerial tool in the public sector. Critics charge, however, that pay-for-performance is ineffective in the public sector, largely because the link between performance and rewards is weak. However, no one has empirically measured the strength of the linkage.

If performance ratings do have an impact on career success in the federal service, they might contribute to race and gender inequality. Although many studies have examined factors affecting gender and racial differences in career success, studies that try to connect gender and racial inequalities to managerial tools are scarce.

Using a one percent sample of federal personnel records, the first essay examines the impact of performance ratings on salary increases and promotion probabilities, and the second essay explores whether women and minorities receive lower ratings than comparable white males, and women and minorities receive lower returns on the same level of performance ratings than comparable white males. The first essay finds that performance ratings have only limited impact on salary increases, but that they significantly affect promotion probability. Thus, the argument that performance-rewards link is weak could be partially correct, if it considers only pay-performance relationships. The second essay finds that women receive equal or higher performance ratings than

comparable white men, but some minority male groups, particularly black men, tend to receive lower ratings than comparable white men. On the other hand, the returns on outstanding ratings do not differ between women and minority male groups and white men, though women groups seem to have disadvantages in promotion with the same higher ratings as comparable men in highly male-dominant occupations.

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Many observers believe that the link between performance and rewards is weak in the federal service. Paul Light criticizes the lack of compensation for outstanding performers in the federal service, arguing that "the very definition of outstanding demands comparison. If the term is to have merit, it cannot apply to more than a handful" (Light, 1999). Rynes et al. (2005) also argue that "Given the typically *weak* link between performance and pay as well as problems with getting supervisors to provide credible measures of performance for administrative purposes, perhaps any positive results at all should be regarded as impressive" (p. 588, my italics). That weak link between performance and rewards may be a primary reason why critics argue performance-based compensation is not effective and has failed to achieve desired outcomes (Ingraham, 1993; Kellough & Nigro, 2002; Pearce & Perry, 1983).

Critics, however, do not provide empirical evidence on how weak the link between performance and rewards is. Performance appraisal systems stand on the belief that by providing a basis for personnel decisions, particularly linking pay to performance, performance appraisals motivate employees to increase their work efforts, and hence improve individual performance and organizational productivity (Daley, 1992; Milkovich & Wigdor, 1991). Therefore, for performance-based rewards to be effective, they must link rewards to employee performance, but many studies have found that the link

between performance and rewards is weak (Pearce & Perry, 1983; Rynes et al., 2005; Smith & Rupp, 2003).

However, the arguments that performance-rewards link is weak are mostly based on surveys or less systematic perceptions. Ingraham (1993), for example, cited two surveys conducted by the U.S. Merit Systems Protection Board and argued, "[A] majority of those responding do *not* perceive a link between performance and reward" (p. 352, italics in original). In response to the 1989 Merit Principles Survey question "If you perform better in your present job, how likely is it that you will receive more pay?", for instance, only 37.1 percent said it was "likely" or "very likely." This survey gives us employee perceptions on the practice of pay-for-performance in the federal civil service, but it still does not provide quantifiable evidence on the strength of the link between performance and pay. If there is a meaningful difference in performance pay between top and average performers but only a few employees are evaluated as top performers, for instance, then most average performers cannot improve their performance enough to receive that performance pay and their answers would primary be "unlikely." On the contrary, if most employees receive performance pay, though the amount is small, the employees would choose "very likely." Does this mean that performance-pay relationship is very strong, so that performance-based pay is very effective to motivate employess? Other problems with this kind of survey are that employees may not recognize the size of difference in performance rewards between high and average performers or, because most federal employees dislike performance appraisal and pay-for-performance (Ingraham, 1993), they may recognize pay differences between high and average performers but not want to admit it. Despite the significance of the topic, little is

known about the strength of the linkage between performance and rewards. The dissertation provides empirical evidence by examining the impact of performance ratings on salary growth and promotion probabilities.

If agencies use performance appraisals to make pay and promotion decisions, performance ratings should be associated with an individual's career success. According to Sherk (2007), pay inequality among individuals rose in recent years due to increased use of performance pay. Also, performance appraisals could be associated with gender and racial disparities in career success. Castilla (2008) presents two possible explanations of how performance ratings can contribute to the disadvantages of women and minorities in career success. First, women and minorities can receive lower performance ratings than comparable men and whites with the same performance level ("performance-evaluation bias"). Second, women and minorities can receive lower returns on the same performance ratings ("performance-reward bias"). Although, many studies have examined factors affecting gender and racial differences in career success, few try to connect gender and racial inequalities to managerial tools. The dissertation explores how performance appraisals could be associated with gender and racial disparities in career success.

The dissertation is composed of two essays. Using a one percent sample of the Central Personnel Data File, which is maintained by the U.S. Office of Personnel Management, essay 1 examines how strongly performance ratings affect salary growth and promotion in the federal civil service, and essay 2 explores how performance ratings are associated with gender and racial inequalities in federal career success.

1.2 Performance Appraisals and Performance-Based Rewards

Although many researchers have studied performance appraisal, the practice of performance evaluation varies among organizations and the term "performance appraisal" is used a bit differently from researcher to researcher. Some researchers define performance appraisal broadly and use performance appraisal and performance management interchangeably. But in general, performance appraisal refers to "the process in which, for a specified period of time, employees' work performance, behaviors, or traits are rated, judged, or described by a person other than the rated employee and the results are kept by the organization" (Coens & Jenkins, 2000, p. 14).

Typically, employees' performance is evaluated by their immediate supervisors (Daley, 1992). Immediate supervisors know their subordinates best and can use performance appraisal as a managerial tool. However, many kinds of rater errors can occur, and that reduces the credibility of the evaluation results. Examples of these types of errors include the halo effect (inappropriate generalization of one aspect of an employee' performance to all other aspects of the job), first impression error (influence of initial positive or negative impression on an employee to subsequent evaluation on the employee), the similar-to-me effect (giving higher ratings to those who resembled raters), the comparison or contrast effect (evaluating based on the relative performance of each other not on the actual performance), and central tendency error (evaluating all employees in the middle of the scale) (Kellough, 2002; Murphy & Cleveland, 1995). To reduce rater error, many studies have concentrated on how to improve performance appraisal or have suggested alternatives (Bretz et al., 1992). Frequently suggested prescriptions include 360 degree evaluations, which combine information from

supervisors, peers, subordinates, and self-evaluation to evaluate employee performance (Rynes et al., 2005), and rater training to improve accuracy (Daley, 1992).

Ideally, raters should evaluate job outcomes, but because performance measurement is so difficult, particularly in the public sector (Montoya & Graham, 2007), most organizations evaluate employees based on their traits, such as dependability and cooperativeness, and on some behaviors (Kellough, 2002).

The results of performance appraisal are used for numerous purposes. The common purposes of performance appraisals can be grouped into two broad categories: developmental and judgmental (Boswell & Boudreau, 2000; Daley, 1992, 2005; Rynes et al., 2005). Developmental purposes focus on distinguishing *within* individuals whereas judgmental purposes are related to distinguishing *among* individuals (Cleveland et al., 1989). According to Boswell & Boudreau (2002), development is "any effort concerned with enriching attitudes, experiences, and skills which improve the effectiveness of employees" (p. 392). Examples of developmental uses are identifying strengths and weaknesses and identifying training needs. Thus, developmental purposes focus more on individuals' potential than on current levels of skills and abilities.

Judgmental purposes are well known for merit pay or pay-for-performance. Performance appraisals compare individuals' performance to other employees or to set standards in the organization, and therefore help to identify good and bad performers. Based on the performance appraisal results, organizations make administrative decisions. According to Daley (1992), "promotion and merit pay are the two most widely known and used of the judgmental appraisal purposes" (p. 17). Other judgmental purposes include decisions on retention, demotion, and reassignment.

In theory, performance appraisals are supposed to be used for both developmental and judgmental purposes. In practice, judgmental appraisals that affect pay and promotion decisions are common than developmental appraisals (Daley, 1992; Murphy & Cleveland, 1995).

1.3 History of Federal Performance Appraisals

The history of performance appraisal in the federal government begins in the early 1950s with the passage of the Performance Rating Act of 1950, which aimed to identify the best and worst employees (U.S. Office of Personnel Management, 2008b, 2008c). However, both supervisors and subordinates were dissatisfied with performance appraisal and it was not emphasized in the federal government throughout the 1960s and most of the 1970s (Riccucci & Naff, 2008). The Civil Service Reform Act (CSRA) of 1978 changed the performance appraisal system dramatically. Agencies were required to develop appraisal systems for all federal employees and were told that "results of the appraisal must be used as a basis for training, rewarding, reassigning, promoting, reducing in grade, retaining, and removing employees" (U.S. Office of Personnel Management, 2008c). The CSRA also created the Merit Pay System which covers supervisors and managers in General Schedule (GS) 13-15. Under the Merit Pay System, which became mandatory on October 1, 1981, agencies pooled half of the general comparability adjustment and within-grade and quality-step increase monies and distributed the fund to the employees based on their performance ratings (Milkovich & Wigdor, 1991; Nigro et al., 2007). However, the Merit Pay System did not work well; funding for merit pay was cut and employees complained about "meaningless merit

increases" (Nigro et al., 2007). Pearce and Perry (1983) found that the Merit Pay System was not better than the previous seniority-based compensation in motivating employees.

The federal government introduced the Performance Management and Recognition System (PMRS) in 1985 and replaced the Merit Pay System with a new pay-for-performance system for GS 13-15 employees. Under this system, middle managers were assigned one of five summary ratings and received performance awards based on those ratings. PMRS covered about 8 percent of federal white-collar employees in 1993 (Eisenberg & Ingraham, 1993).

In 1995, the federal government substantially decentralized performance appraisal and pay-for-performance. Agencies can now select eight permissible summary rating patterns, allowing from two (pass or fail) to five levels of summary ratings (U.S. Office of Personnel Management, 2008c). The two largest federal departments (Homeland Security and Defense) were allowed to develop their own compensation systems, using pay banding to widen the gap between the pay of high and low performers (Hyde, 2005; Naff & Newman, 2004).

Federal experience with performance-based pay is not positive, however. Studies found that federal supervisors did not differentiate enough between high and poor performers in performance ratings, which made it hard for high performers to receive enough rewards; they also found that the systems failed to motivate employees (Gaertner & Gaertner, 1985; Nigro et al., 2007; Pearce & Perry, 1983). Nevertheless, performance-based pay is always a popular reform. The Obama Administration is going to introduce a governmentwide pay-for-performance system that "takes employees' performance into account when compensating them" (Losey, 2009). The Office of Personnel Management

(OPM) director announced the reform plan for making effective pay-for-performance, which allows large pay increases for a small number of truly outstanding performers. On the other hand, Kunreuther (2009) argues that the precondition for effective pay-for-performance is an unbiased and realistic evaluation system but that decades of federal experience show the precondition is hard to achieve. He also argues that rewarding only a few stellar performers with substantial pay-out may demotivate a large number of mediocre employees, so that the system cannot achieve improved organizational productivity. However, little is known about whether the federal government has ever rewarded outstanding performance with "substantial pay-outs," whether the amount of merit pay was enough to motivate outstanding performers, and whether rewarding outstanding performers with "substantial pay out" demotivates average performers. Indeed, there exists a widespread perception that the link between pay and performance is weak, and therefore reform efforts have focused on strengthening it. However, in order to achieve the goal of improving organizational productivity, a reform initiative needs not just perception but empirical data on the relationship between pay and performance.

CHAPTER 2

THE IMPACT OF RATINGS ON CAREER ADVANCEMENT

2.1 Introduction

With the emphasis on results-oriented public management, performance appraisal systems have become a key tool for managing public agencies' performance. By providing a basis for personnel decisions, particularly by linking rewards to performance, performance appraisals seek to motivate poor performers to increase their productivity, and strong performers to maintain or enhance theirs. Because of the theoretical promise of improving individual performance and ultimately organizational productivity, organizations have made much effort to improve performance-based compensation as well as to develop performance appraisal systems (Bretz et al., 1992).

One of the major issues of the federal reform effort was the extent to which performance appraisal results would be used for personnel decisions. The Civil Service Reform Act of 1978 states that federal agencies should develop performance appraisals for all employees and use the results "as a basis for training, rewarding, reassigning, promoting, reducing in grade, retaining, and removing employees" (U.S. Office of Personnel Management, 2008c). However, critics argue that performance appraisals in the public sector are broken, particularly because the linkage between performance and rewards is so weak that performance appraisals fail to serve their purpose of motivating employees and improving performance (Light, 1999).

Because of the central role of performance-based rewards in federal human resource management, the Obama Administration is pursuing reform by establishing a

governmentwide pay-for-performance system that "should include large raises for a small group of truly outstanding performers" (Rosenberg, 2009). How large do performance pay differences between the best and worst performers need to be to motivate them? Although reform efforts to make effective performance appraisal systems should start from a knowledge of the relationship between performance appraisal results and rewards, little is known about the strength of the linkage. In addition, results of performance appraisals are supposed to be used for numerous purposes including promotion decisions, but most studies have focused primarily on merit pay or pay-for-performance.

This study examines the impact of performance ratings on salary increases and promotions in the federal civil service. For performance-based compensation to be effective, according to equity theory, there should be meaningful reward differences among employees depending on their performance. This study does not judge whether reward differences are "meaningful" but measures the differences in rewards by performance rating levels and provides a basis for judging whether the differences are large enough to motivate employees. Expectancy theory argues that employees are motivated when they perceive that their work effort leads to higher performance (expectancy), that higher performance leads to rewards (instrumentality), and that the rewards are valuable to them (valence). This study does not cover the expectancy and valence parts of expectancy theory. In other words, the present study does not examine whether performance ratings reflect employees' real performance or whether governments' rewards are valuable to public employees. This study focuses only on one part of instrumentality: whether higher performance ratings lead to higher rewards. In particular, because pay and promotion are primary rewards for government employees,

this study examines how strongly performance ratings are tied to pay increases and promotions.

This paper is organized as follows. Section 2.2 reviews theories that provide a basis for performance-based rewards and previous studies that are related to performance-rewards relationship. Section 2.3 reviews factors that affect career advancement. Section 2.4 develops hypotheses on pay increase and promotion probabilities. Section 2.5 describes the data, variables, and research design. Section 2.6 presents the findings and discusses what they tell us about the hypotheses. Section 2.7 concludes the paper by summarizing the study and discussing its implications.

2.2 Literature Review

Theoretical Bases of Performance-Based Rewards

According to equity theory, employees' work motivation depends on a balance between their perceived inputs to organizations and outcomes from the organizations to them (Adams, 1965). Employees consider both internal and external equity. Internal equity means that employees feel inequity if their input is greater or smaller than the rewards. External equity means that employees compare their outcome/input ratio with those of their coworkers, and they feel inequity if their outcome/input ratio is lower or higher than their coworkers'. Equity theory argues that employees try to reach a balance either by reducing their input (work effort) if they are under-compensated or by raising their input if they are over-compensated, either in terms of internal or external equity. Many empirical studies support the theory, finding that undercompensated workers lower their input to achieve equity (Ambrose & Kulik, 1999; Griffeth et al., 1989; Heneman &

Werner, 2005). Although critics argue that employees' responses on overcompensation differ from those on undercompensation (Mowday, 1991) and that perceptions of fairness are influenced not only by "the ends achieved" (distributive justice) but by "the means used to achieve those ends" (procedural justice) (Greenberg, 1990), equity theory provides a useful framework for understanding worker behavior in organizations (Bolino & Turnley, 2008).

In particular, equity theory gives us some implications for performance-based rewards (Heneman & Werner, 2005). For performance-based rewards to be effective, employees must perceive the amount of rewards to be equitable both compared to their input and compared to other employees (Gabris & Ihrke, 2000). Therefore, in order to retain high performers and keep their performance level high, there should be meaningful reward differences between high and average performers. Because employees' perceived input/outcome ratio is as important as the actual input/outcome ratio but equity sensitivity varies among employees (Rainey, 2003), managers need to better communicate with employees on performance rewards.

Another theory that provides a rationale for performance-contingent rewards is expectancy theory. Vroom (1964) argues that an employee's work effort is a function of the multiplication of expectancy, instrumentality, and valence. The theory indicates that in order for rewards to motivate employees, individuals should perceive that more effort leads to higher performance (expectancy), that improved performance leads to higher rewards (instrumentality), and that the rewards are valuable to them (valence). Many researchers propose that expectancy theory is a useful framework to understand individual and organizational motivations and apply the theory to a variety of settings,

including faculty motivation to conduct research (Chen et al., 2006; Tien, 2000), student motivation to learn (Hancock, 1995), and organizations' strategic decisions (Chen & Miller, 1994).

Expectancy theory has some implications for performance-based rewards (Heneman & Werner, 2005). First, there should be an opportunity for employees to improve their performance. If not, expectancy cannot be present. Second, the amount of reward should be contingent upon the level of performance, so that employees can understand that performance is instrumental in achieving performance-based rewards. Third, rewards must be a valued outcome to the employees.

Several issues occur when the theory is applied to performance-based rewards. In many cases, employees work together with their coworkers so that individual effort may not result in increased performance. Because of resource limitations, organizations often can not provide high enough rewards for high performance. One of the controversial issues is how effective the extrinsic rewards are. Organizations largely rely on monetary rewards to motivate employees. It is clear that monetary rewards are strong motivators (Bucklin & Dickinson, 2001; Lawler, 1971; Locke et al., 1980; Rynes et al., 2004; Stajkovic & Luthans, 2003). For example, Locke et al. (1980) argue that "money is the crucial incentive" and that "no other incentive or motivational technique comes even close to money with respect to its instrumental value" (p. 379). However, public employees tend to place great value on the intrinsic rewards (Houston, 2000) and on helping others and serving society rather than monetary rewards (Perry, 1996). Oh and Lewis (2009) found that performance appraisal systems, which rely on extrinsic rewards, are less effective with intrinsically motivated public employees. In addition, in some

circumstances, extrinsic rewards can drive out intrinsic motivation, particularly when a majority of employees are intrinsically motivated (Canton, 2005; Frey, 1997; James, 2005; Ryan & Deci, 2000).

Linkage between Performance Appraisal Results and Rewards

Agencies introduce performance appraisal as a tool to motivate employees. Some studies show that merit pay results in positive outcomes, particularly for individual performance (Heneman & Werner, 2005; Milkovich & Wigdor, 1991). In theory, performance-based rewards are promising and some positive experience in the private sector provides good justification of performance-based rewards in the public sector. In practice, however, many empirical studies show that performance appraisal is ineffective and that the impact of performance ratings on salary decision is minimal in the public sector (Pearce & Perry, 1983; Smith & Rupp, 2003).

The ineffectiveness of performance appraisal and performance-based pay comes from a variety of sources. In order for performance appraisal results to be used for pay decisions, performance ratings should reflect employee performance accurately so that everyone can trust the results. However, rater error is unavoidable, particularly in the public sector where performance is not easily measured. Not only because of measurement itself, but because of supervisors' rational behavior, performance ratings do not reflect real employee performance (Klingner & Nalbandian, 1998; Murphy, 2008; Nigro et al., 2007; Rynes et al., 2005). For example, Murphy (Murphy, 2008) argues that "performance appraisal is a complex event that occurs in environments that often push raters to distort their ratings to accomplish valued goals or to avoid the negative

repercussions of giving ratings their subordinates or superiors will find objectionable" (p. 198). Nigro et al. (2007) also argue that accurate ratings are not rewarded by the organization, and "negative ratings often yield nothing more than stressful interpersonal conflict and time-consuming appeals by resentful workers to suspicious civil service boards" (p. 170). Under these circumstances, it is hard for organizations to rely much on the performance ratings when they make pay decisions.

Another reason for the weak relationship between performance ratings and rewards is a shortage of resources. Pay and promotion are limited. In particular, it is hard for government agencies to have large enough budgets to give meaningful pay differences between good and poor performers. Rating errors and limited resources result in inflated performance ratings (Riccucci & Naff, 2008). According to the U.S. Office of Personnel Management, over two thirds of all federal employees received performance ratings of 4 or 5, the two top ratings, during the fiscal years of 1991-1996 (U.S. Office of Personnel Management, 1999).

Whether performance ratings reflect employee performance accurately or not, performance ratings are used as proxy of employee productivity in many studies (Alonso & Lewis, 2000; Medoff & Abraham, 1981). Whether performance ratings are inflated or not, individual performance ratings are supposed to be linked to organizational rewards under pay-for-performance systems (Heneman & Werner, 2005). This study aims to determine whether and how strongly performance ratings are related to rewards.

2.3 Other Factors Affecting Career Advancement

Studies have shown that individual characteristics and job-related factors are associated with career advancement. To correctly measure the impact of performance

ratings on salary increases and promotion probabilities, this study controls for gender, race, education, federal experience, age, supervisory status, grade, occupation type, and agencies. This section briefly reviews how these factors are associated with pay and promotion.

Gender and Racial Minority Status

Gender and race are important factors in predicting career success. Many studies show that women's and minorities' average salaries are lower than men's and whites', though the gender pay gap has been decreasing. Glass ceiling theory argues women and minorities have fewer opportunities for advancement to upper-level positions than do men and whites, but studies on gender/racial differences in promotion probabilities provide mixed results. While some studies have found that promotion probabilities are lower for women and minorities than for white men (Blau & DeVaro, 2006; Cobb-Clark, 2001; Paulin & Mellor, 1996), others find that women's promotion probabilities are as high or higher than men's (Hersch & Viscusi, 1996; Krull, 2006; Lewis, 1986, 1992). Hersch and Viscusi (1996) explain that because women tend to be hired at lower level women have higher promotion rates than men. Gerhart and Milkovich (1989) also found women received more promotions and larger salary increases than men, and explained that one reason is that women work at lower positions which have large promotion opportunities.

Education

Education is a key element of human capital. According to human capital theory, education increases individual productivity, so more educated employees earn more than

less educated employees (Becker 1993; Mincer 1993). Screening theory, on the other hand, argues that education does not enhance productivity but that higher earnings reflect innate ability, which leads individuals to obtain more education, and employers use educational attainment as a proxy for likely productivity (Belman & Heywood, 1991; Hungerford & Solon, 1987; Spence, 1974). Empirical studies show that education has no impact on salary increase (Lewis, 1997a) and has only minimal effect on promotion probabilities in the federal government (Lewis, 1986).

Federal Experience

Work experience is a human capital factor that increases salary level. However, studies show that work experience is curvilinearly associated with salary increases (e.g., Lewis & Oh, 2009), probably because workers invest in skills more early in their careers than late in their careers (Becker, 1973). Empirical studies have found that salary increases and promotion probabilities drop with federal experiences (Lewis, 1986, 1997a).

Age

Most labor economics literature that has examined the determinants of salary show that age has positive impact on salary. However, a positive coefficient on age does not necessarily (?) mean that older employees are more productive than younger employees (Hellerstein et al., 1999). Studies have found that the relationship between age and productivity is negative (Rhodes, 1983; Cox & Nkomo, 1992) or that age is not related to productivity (McEvoy & Cascio, 1989). On the other hand, Sturman (2003) showed that age-productivity relationship is not linear but inverted U-shape, indicating

age is positively associated with productivity up to a certain age, and then negatively associated with productivity. Consistent with Sturman (2003), Lewis and Oh (2008) showed that the relationship between age and salary level is curvilinear, which means salary rises at a decreasing rate and then decreases at a increasing rate with age. Lewis (1997a) also found that salary increase and promotion chances drop with age in the federal agencies.

Supervisory Status

Federal supervisors earn more than nonsupervisors. Taylor (1979) analyzed 15,912 federal employees drawn from the U.S. Civil Service Commission's data file and showed supervisors' salaries were 11.5 to 17.2 percent higher than nonsupervisors in 1977, after controlling for age, years of service, veterans status, years of education, and PATCO category. Using data from the Central Personnel Data File, Johnson and Libecap (1989) found supervisors earned 14.2 percent more than nonsupervisors in 1985, holding constant level of education, occupational group, and PATCO category. Asch (2001) analyzed data from the Department of Defense civil service personnel and found supervisors earned 6.0 to 11.4 percentage points more than nonsupervisors, controlling for a variety of variables including performance rating.

Although supervisors have a salary advantage, they do not differ in promotion probabilities from comparable federal employees who are not supervisors. Asch (2001) showed that supervisors' promotion speed is not faster than nonsupervisors in the Department of Defense and the three military departments when she controlled for individual and job-related factors.

Grade Level

In the federal civil service, salary level is largely a function of grade level. Job evaluation determines the value of the job and assigns the grade for the job (U.S. Office of Personnel Management, 2002b). Advancement to a higher grade leads to about a 10 percent increase in salary (Spyropoulos, 2005). Promotion chances drops with grade. In his study on middle-level federal managers, Lewis (1992) showed men's average promotion rates were 7.6 percent for GS 13, 6.1 percent for GS 14, and 2.4 percent for GS 15 in 1979-1989. Naff (1994) showed the average promotion rate was about 10 percent or below for GS 13-GS 15 whereas the rate was 33 to 44 percent for GS 9 in 1991-1992. Lewis (1997a) compared promotion probability among grade level groups in the federal agencies and found promotion probabilities were highest at GS 1-GS 3 and decreased with the grade, holding gender, age, federal experience, and education constant.

PATCO Category

All federal white-collar occupations are classified into five categories: Professional, Administrative, Technical, Clerical, and Other (PATCO), based on "the subject matter of work, the level of difficulty or responsibility involved, and the educational requirements" (U.S. Office of Personnel Management, 2002a, p. 242). Because the PATCO system requires a higher level of education (a bachelor's degree) for positions in Professional and Administrative than for positions in the three other categories (U.S. Office of Personnel Management, 2002a), employees in Professional and Administrative categories have a career advantage over employees in the other three categories. In general, Professional and Administrative occupations are "two-grade interval" series, because promotions are granted two grades at a time from General

Schedule 5 through 11, while Technical and Clerical occupations are "one-grade interval" series, which means employees are promoted one grade at a time (U.S. Office of Personnel Management, 2002b). Because promotion to a higher grade results in higher salary increases than normal annual salary increases, employees in Professional and Administrative also have a salary advantage. Studies show that women's disadvantages in pay in federal agencies are associated with their relatively high concentration on Technical or Clerical occupations (e.g., DiPrete & Soule, 1986; Lewis, 1996).

Agency

Career patterns of employees vary with the agencies in which they work. Borjas (1980) found that employee salary rate was negatively related to the size of the agency in the federal government and that in 1977 salaries for comparable employees were highest in Community Service Administration and lowest in Agriculture. Lewis (1997a) showed that Defense employees' salary increases and promotion chances were lower than comparable employees in domestic agencies.

As discussed above, federal career advancement is not simply a function of some individual characteristics, but numerous factors come into play to in determining pay increases and promotion probabilities. In addition, those factors could be positively or negatively associated with performance ratings that employees receive. For example, Lewis (1997b) shows that older employees are more likely than younger ones to receive outstanding ratings in federal agencies. Therefore, it is necessary to separate these factors for measuring the relationship between performance ratings and career advancement.

2.4 Hypotheses

Pay Increase

Performance appraisal systems are designed to evaluate employee performance and agencies use the results for pay decisions. Based on equity and expectancy theories, higher performers should receive higher merit increases and poor performers should receive low or no merit increases. For performance-based pay to be effective, some portion of the pay increase should be based on employee performance and there should be meaningful differences in performance pay among employees. Mitra et al. (1995) argue that if the merit portion of the annual salary increase is too small, merit pay could decrease employee motivation and morale; therefore, they argue, merit pay should be at least 6 to 7 percent of base pay if it is to be effective. Fuller and Tinkham (2002) suggest that, in a firm with an effective reward differentiation program, outstanding performers (defined as the top 5 to 15 percent) receive performance pay twice that of average performers and poor performers receive 40 percent less than average performers. Medoff and Abraham (1981) measured the impact of performance ratings on the log of annual salary in two major U.S. corporations and found that differences in salary between those who received the highest and lowest performance rating were 7.8 and 6.2 percent, after controlling for level of education, years of work experience, region, and grade level.

Under the Performance Management and Recognition System introduced in 1985, federal employees in GS 13-15 who received level 5 (outstanding) were to receive performance awards of 2 to 10 percent of their pay in addition to the general pay increase (a percentage increase to base pay and a locality pay adjustment which is recommended by Congress and approved by the President), whereas those who received level 2 and 1

were to be penalized: level 2 employees were to receive only 50 percent of the general pay increase and level 1 employees were to receive no pay increase (Riccucci & Naff, 2008).

Hypothesis 1.1: Salary increases will be higher for highly rated employees than for poorly rated employees.

In general, the amounts of performance pay do not seem to vary much among employees either in the private or public sector, however (Heneman & Werner, 2005; Murphy, 1992; Zenger, 1992). According to Teel (1986), "most of the organizations surveyed awarded 60 percent or more of their employees raises that varied no more than 2 percent from the average" (p. 90). Employees in the Federal Senior Executive Service do not think pay and bonuses are meaningfully different among executives. Only 26 percent of respondents agreed or strongly agreed that "Pay distinctions are meaningfully different among executives" (U.S. Office of Personnel Management, 2008e, p. 2). OPM (2008b) explains that the reason such a low percentage of federal executives believe pay differs meaningfully may be because the rating pattern used in the agencies does have enough rating levels to distinguish employee performance or agencies do not have enough funding to make difference between good and poor performers (U.S. Office of Personnel Management, 2008b).

Hypothesis 1.2: Differences in salary increases between highly rated and poorly rated employees will be small.

Many scholars and practitioners point out that one of the serious problems with performance appraisals is rating inflation (Eisenberg & Ingraham, 1993; Light, 1999;

Murphy & Cleveland, 1995): "Everybody gets a satisfactory rating or above" (Taylor et al., 1995, p. 506). Ratings inflation jeopardizes the performance appraisal systems (Bretz et al., 1992; Hyde, 2005), they argue, because inflated ratings make it hard to measure employees' work performance accurately, and thus it is difficult for organizations to use performance ratings for either administrative or developmental purposes (Roch, 2005).

Many studies have tried to uncover the source of rating inflation. Murphy and Cleveland (1995) argue that, inflated ratings are mostly a result of raters' rational behavior, not an inability to evaluate performance accurately. Raters give inflated ratings because high evaluations may indicate that the raters are effective in employee development (Greenberg, 1991) and because "they want to be personally popular" (Gerhart & Rynes, 2003, p. 168). Murphy and Cleveland (1995), focusing on negative consequences of accurate ratings, add more reasons why raters inflate their ratings. Because a low rating can result in an unpleasant situation, raters inflate to avoid that situation. In the worst case, "accuracy in evaluation could make it difficult for supervisors and subordinates to work together" (p. 266), and therefore accurate rating could lower organizational productivity in the long run. They argue that if accurate ratings are an important goal, organizations must reward accurate ratings and punish inaccuracy; otherwise rational supervisors are likely to give inflated ratings to their subordinates.

However, little is known about how rating inflation affects rating effectiveness. There are two opposing arguments. Some researchers argue that ratings used for personnel decisions are significantly higher than ratings used for feedback or research (Jawahar & Williams, 1997; Milkovich & Wigdor, 1991; Murphy & Cleveland, 1995).

On the contrary, Montoya and Graham (2007) argue that it is easy for supervisors to give higher ratings when the highest rating does not necessarily lead to additional pay.

Although both explanations are possible depending on the organizational context (Daley, 1992), performance ratings should be less inflated if organizations use performance ratings to make personnel decisions, because they have limited resources to reward high performers. The experience of the Government Accountability Office (GAO) was that performance ratings are not inflated but differentiate among employees when personnel decisions rely heavily on performance evaluations. In January 2002, GAO implemented a new competency-based performance management system, which emphasizes links between performance and merit pay increases. As a result, "for fiscal year 2002, the GAO-wide average performance appraisal rating was 2.19 (out of 5) compared with 4.26 (out of 5) for fiscal year 2001. Similarly, under the new system, no employees received a score of 4.7 or higher, while 19 percent of employees received a score of 4.7 or higher for fiscal year 2001" (U.S. General Accounting Office, 2003, pp. 9-10). If performance ratings are not inflated, it may imply that performance ratings have a substantial impact on personnel decisions.

Examples of private companies that employ forced ranking systems, in which raters must rate employees according to the assigned number of employees for each grade, also support that performance rating-personnel decision relationship is stronger when performance rating is less inflated. In the private sector, the use of forced ranking system has become popular in recent years. For example, the General Electric system limits the highest performers to 20 percent and classifies 10 percent of employees as the poorest performers. Studies show that companies employing the forced ranking system aim not

only to avoid rating inflation but to strongly tie personnel decisions to performance (Kräkel, 2008; Scullen et al., 2005).

Average performance ratings vary widely among federal agencies. Employees in the Army were six times as likely as employees in Agriculture to receive outstanding ratings in 1995 (67 vs. 11 percent), and employees in Defense Department received high ratings whereas employees in Agriculture, Labor, Treasury, Interior, Transportation, and Veterans Affairs received low ratings (Lewis, 1997b). Also, the average performance rating rose between 1991 and 1996 and the rise over time is pretty steady (U.S. Office of Personnel Management, 1999). The average rating for General Schedule 13-15 supervisors and management officials, increased from 4.16 (in 1991) to 4.45 (in 1996), and the percentage receiving outstanding ratings increased from 34.7 to 56.0 percent. In General Schedule 1-12, the average rating rose from 3.95 to 4.16, and the percentage receiving outstanding ratings was 27.2 percent in 1991 but 39.2 percent in 1996 (U.S. Office of Personnel Management, 1999).

Hypothesis 1.3: As performance ratings have risen over time, the impact of performance ratings on salary increase has decreased.

Hypothesis 1.4: The impact of performance ratings on salary increase was smaller in agencies where mean performance ratings were higher.

Promotion Probabilities

Informing promotion decisions is another important purpose of performance appraisals (Milkovich & Newman, 1996; Murphy & Cleveland, 1995). Buckley et al. (1995) write that "Performance appraisals are tied to many individual and organizational

outcomes, that is, pay and promotion" (p. 105). Some federal agencies also clearly state that performance appraisals serve as a basis for promotion decisions (National Institute of Standards and Technology, 2009; U.S. Office of Personnel Management, 2008d). For example, the National Institute of Standards and Technology (2009) states that the objectives of its performance management are "To use performance appraisal results as a basis for paying, rewarding, [and] promoting" (National Institute of Standards and Technology, 2009).

Compared to pay decisions, promotion decisions have several different characteristics, however. First, while performance-based pay is contingent on past performance, promotion decisions not only serve as a reward for past performance but consider who are the best fit for the advanced job (Banks & Murphy, 1985). Second, organizations make performance-based pay decisions annually, but make promotion decisions based on the past several years' work (London & Stumpf, 1983). Third, unlike pay decisions, employees in higher grade levels have fewer chances to be promoted, because position openings are fewer in high grades (Lewis, 1986, 1992).

Some studies found weaker linkages between performance ratings and promotions than between performance ratings and salary increases (Castilla, 2008). In his analysis of a large private firm, Mobley (1982) found that the firm's promotion decisions relied 30 percent on an employee's performance rating. The GAO (2003) also reported that performance ratings have only a "modest" impact on promotion chances of federal workers. Part of the reason is that, as Asch (2005) argues, the promotion system requires the supervisor to determine who are the best performers but performance ratings do not specifically provide performance rankings. Individual employees' performance can be

rank-ordered, although they are rated as same level of performance, and therefore it is possible that some outstanding performers are promoted whereas other outstanding performers are not, even if all other factors are equal.

Although very few studies have empirically tested the relationship between performance ratings and promotions, it is clear that performance ratings are positively related to promotion probabilities. Using logit analysis, Powell and Butterfield (1994) examined whether gender and other job-related factors including performance ratings affected promotion decisions in a cabinet level department where 39 SES positions were filled in the periods of 1987-1994, and found that performance ratings had a significant effect on promotion decisions. Lyness and Heilman (2006) analyzed 448 upper-level managers in a private company and found that as performance rating[1] increased by one level, the odds of promotion rose 53 percent for female managers and 13 percent for male managers, holding constant age, organizational tenure, education, and organizational level.

Hypothesis 1.5: Highly rated employees are more likely than poorly rated employees to be promoted.

Hypothesis 1.6: Difference in promotion probabilities between highly rated and poorly rated employees will be small.

Rating inflation may affect not only salary increase but promotion probabilities.

---

[1] The variable performance rating has three values from 1 (low) to 3 (high).

Hypothesis 1.7: As mean performance ratings have risen over time, the impact of performance ratings on promotion probabilities has decreased.

Hypothesis 1.8: The impact of performance ratings on promotion probabilities is smaller in agencies where mean performance ratings are higher.

2.5 Data and Methods

Data

This study relies on a one percent random sample of the Central Personnel Data File (CPDF). The U.S. Office of Personnel Management (OPM) collects personnel data from the federal agencies, except special agencies including U.S. Postal Service, the Central Intelligence Agency, the Federal Bureau of Investigation, the Defense Imagery and Mapping Agency, and the White House Office. The CPDF contains demographic information, job information, the agency of employment and location of employees (U.S. Office of Personnel Management, 2008a).

Each year OPM draws a one percent random sample from the CPDF for study purposes based on their social security numbers. Because employees in the sample one year remain in the sample for all years in which they work for the government in April, it allows us to track employees over time and to identify promotions. Because of the privacy concerns, OPM declined to provide updated information on employee performance ratings, which is the key variable for the study. So the latest available data on individual performance ratings is 2003 and the earliest is 1988. Thus, this study analyzes full-time, white-collar employees from 1988 through 2003. Because a substantial number of federal employees have gone to a pass/fail rating system since

1996 but the pass/fail system is too crude to provide a basis for pay and promotion decisions (Liff, 2007; Montoya & Graham, 2007), this study restricted the sample to employees under five-level rating systems.[2]

Dependent Variables

To measure the impact of performance ratings on personnel decisions, this study uses the *percentage change in salary* and *promotion* as dependent variables. Because salary level depends largely on employees' grade, the dollar amount of salary growth is higher for higher-grade employees than for lower-grade employees. That is, even a small percentage salary increase for higher-grade employees increases salary by more dollars than a high percentage increase for lower-grade employees. Thus, instead of using actual dollar amount for salary growth, this study uses percentage change in salary. Salary growth is measured by the difference in salary between the next year and this year, divided by this year's salary and multiplied by 100. The variable *promotion* is coded 1 if his or her next year's grade is higher than this year's grade and coded 0 for everyone else. As shown in Table A.1, the mean percent change in salary is 6.72, and 14.6 percent were promoted employees over the 1988-2003 period.

---

[2] Although Liff (2007) writes that the goal of pass/fail system is to "weed out" poor performers, analysis on a one percent sample of CPDF shows that among those who were under pass/fail system only 0.12 percent of federal employees received "fail" ratings and that there is no significant difference in exit rate between those who received "pass" and those who received "fail." Only a few federal employees are under three-level rating system and among those, only a few or no (in some years) employees received the lowest rating. In addition, there is no evidence that the highest rating lead to higher salary increase or higher promotion probabilities than mid-level rating, meaning three-level rating system is not much different from pass/fail system in terms of the use of the results.

Independent Variables

Critics argue that performance ratings may not accurately reflect employee performance. However, by definition, pay-for-performance is an incentive plan that "reward[s] employees based on subjectively rated measures of performance" (Heneman & Werner, 2005, p. 6). The key independent variable is a set of dummy variables for performance ratings. The U.S. Office of Personnel Management categorizes summary ratings into five levels: "1 Unacceptable, 2 Level between Fully Successful and Unacceptable, 3 Fully Successful or equivalent, 4 Level between Outstanding and Fully Successful, and 5 Outstanding or equivalent." Because few employees receive the lowest two ratings, level 1 or 2, I combined the two into a single category and labeled it *less than fully successful*. The variables *outstanding*, *exceeds fully successful*, and *fully successful* represent level 5, 4, and 3, respectively. The reference group in the regression and logit models is those who received *fully successful* ratings. Thus, coefficients on performance rating dummy variables show whether the named groups are different from those who received *fully successful* in salary growth and promotion probabilities.

This study controls for human capital factors. Education, federal experience, and age are measured in years. To capture the curvilinear impact of experience on salary growth and promotion probability (e.g., salary growth will be faster and promotion probability will drop faster early in their careers than later in their careers), I add squared term for federal service experience and age.

Other individual characteristics also have an impact on career success. Gender and race are common factors that explain pay and promotion inequalities. The CPDF contains five racial categories including American Indians. Because the sample size for

American Indians is small, I dropped the group and created 7 dummy variables for seven race and gender group: *black male, Hispanic male, Asian male, white female, black female, Hispanic female, and Asian female*, and use *white male* as the reference group.

Career patterns vary with job-related factors and the agencies in which employees work. Salary rises with grade but promotion probabilities drop. I create a set of dummy variables for each grade level and for the Senior Executive Service, which is labeled as grade 16 for convenience. But because only a few employees are in grades 1 and 2 in the sample, I combine grade 1, 2, and 3, and label them as grade 3. The reference group is *grade 8*. Following the OPM classification, I create dummy variables for *Professional, Administrative, Clerical*, and *Technical* occupations with the *Other* category as the reference group. A dummy variable *supervisory status* is coded 1 for those whose position is a supervisor or manager, and 0 for everyone else. Organizational factors are some of the most important factors affecting performance ratings and personnel decisions (Condrey & Brudney, 1992). Performance ratings vary dramatically by agency (Lewis, 1997b), so I create agency dummy variables for all cabinet-level departments and the four largest independent agencies (the Environmental Protection Agency, the General Services Administration, the National Aeronautics and Space Administration, and the Social Security Administration). All other agencies are combined and are used as the reference group, *other agency*.

Methods

To test the hypotheses, this study first analyzes the full panel data set. To understand the overall picture of performance ratings, salary increases, and promotions, I

present the descriptive statistics. As reviewed above, many factors influence career advancement in the federal service, and those factors could be associated with performance ratings. Understanding basic patterns of differences of career advancement between performance groups helps to understand how individual and job-related factors affect pay increases and promotion probabilities. Therefore, before separating those factors from the pure rating effect, I run t-tests to test for differences in career advancement between those receiving outstanding and fully successful ratings and between those receiving exceeds fully successful and fully successful ratings. Using panel data analysis, I measure the impact of performance ratings on salary increase. Since the dependent variable *promotion* is dichotomous, this study employs panel data logit analysis.

Two common methods for panel data analysis are fixed effect and random effect models. The random effect model assumes that the unobserved individual effects are not correlated with the any of the other regressors, whereas the fixed effect model allows the unobserved effect to be correlated with the other explanatory variables. Because a Hausman test for this study rejected the null hypothesis that there is no correlation between unobserved individual effects and the other regressors, this study uses the fixed effect model.

The panel data analysis for this study has weaknesses, however. The panel data set used for this study is unbalanced, because different employees are in the panel for different lengths of time, depending on when they entered federal employment and how many years they stayed. In general, unbalanced panel data works with panel data analysis as long as the reasons for missing values are random (Wooldridge, 2003), but in

this sample, the reason that some observations are dropped is not random but related to the purpose of performance appraisals; that is, analysis on the CPDF shows that linkages between performance ratings and salary growth are stronger for dropped observations, which moved from five-level rating system to pass/fail or three-levels, than for existing observations. The other weakness is that the model includes time invariant variables such as race and gender, but the fixed effect model drops variables that do not vary within individuals (e.g., gender and race) so that we cannot estimate parameters (Greene, 2003). Also, outstanding performers who consistently receive outstanding ratingsare excluded from the estimation when we use fixed effect model, which may leads to a biased results. Therefore, the primary analysis will be cross-sectional ordinary least square (OLS) and logistic regressions.

To understand how performance ratings, salary increases, and promotions have changed over time, I present the descriptive statistics by year from 1988 to 2003. To examine whether differences in salary increases and promotion probabilities between outstanding and fully successful performers are constant over time or rising, falling, fluctuating, I run a set of t-tests by year comparing those who received outstanding and fully successful ratings. And then, I conduct ordinary least square regressions year-by-year during 1988-2002 to examine whether the patterns of salary increases in t-tests hold up to multivariate analyses, which control for individual and job-related factors.

To estimate the impact of performance ratings on promotion probabilities by each year, I use logistic regression analysis. Because logit analysis assumes that the log-odds of promotion are a linear function of the independent variables, the percentage change implied by a logit coefficient depends on the values of all the independent variables. One

of this study's purposes is to measure the size of the impact of performance ratings on promotion probabilities, so I use the *prchange* command in Stata program (Long & Freese, 2006), which allows us to calculate the impact of the variable into a percentage change, holding all the other variables at their means.

Analysis of the full panel data and the OLS and logit analyses allow us to examine the impact of ratings on salary increases and promotion probabilities in the federal civil service, after controlling for race and gender, education level, federal experience, age, supervisory status, grade, PATCO category, and agency. However, because the distribution of performance ratings and the practices regarding salary increase and promotion decisions vary with individual and job-related factors, we need further analysis to examine how the impact of ratings differs within each subcategory of factors. Among the most important factors are occupation type (PATCO), grade level, and agency. As in the full panel analysis and the cross-sectional analysis by year, I present first the descriptive statistics, and then run a set of t-tests, and finally conduct the panel data analysis for each category of occupation type (PATCO), grade level, and agency.

2.6 Findings

Analysis of the Full Data Set

Table 2.1 presents the descriptive statistics of performance ratings and career advancement in the period of 1988-2003. Federal employees' performance rating is pretty high. The average performance rating was 4.06, and 73 percent of employees received either *outstanding* (34.6 percent) or *exceeds fully successful* (38.4 percent)

ratings. The average salary increase was 6.72 percent, and on average, 14.6 percent of employees were promoted annually.

Table 2.1: Descriptive Statistics of Performance Ratings and Career Advancement

	Mean	Standard Deviation	Minimum	Maximum
Performance Rating	4.06	.792	1	5
% Outstanding	34.6	47.6	0	100
% Exceeds Fully Successful	38.4	48.6	0	100
% Change in Salary	6.72	4.74	0	49.7
% Promoted	14.6	35.3	0	100

Table 2.2: Career Advancement Comparison: T-test

Mean % Salary Change		% Promoted	
Fully Successful	Outstanding	Fully successful	Outstanding
7.04**	6.59	15.4**	14.6
Fully Successful	Exceeds Fully Successful	Fully Successful	Exceeds Fully Successful
7.04**	6.62	15.4**	14.1

Table 2.2 shows the results of t-tests to compare the differences in career advancement between outstanding and fully successful performers and between exceeds fully successful and fully successful performers. Surprisingly, outstanding performers' mean salary increase rate and the percentage of them promoted were lower than for fully successful performers. Mean salary increases were 0.45 percentage points lower and the percentage promoted was 0.8 percentage point lower for outstanding performers than for fully successful performers. Fully successful ratings also have advantages over exceeds

fully successful ratings in salary increase rate (7.04 vs. 6.62) and percentage promoted (15.4 vs. 14.1).

However, the t-tests results could be misleading both in terms of the size and sign of the difference, because many individual and job-related factors are associated with salary increases and promotions, and those factors are also related to the performance ratings, but the t-tests do not control for these factors. In fact, when I ran a regression with percent change in salary as the dependent variable and with performance ratings as the only independent variable, higher ratings were negatively associated with salary increase; but when I added federal experience to the model, the coefficient on performance ratings became positive, which indicates that if employees have the same years of federal experience, higher performers' salary increases are higher than lower performers'.[3]

Therefore, the panel data analysis tests whether the t-test results hold true even after controlling for individual and job-related factors. As shown in Table 2.3, outstanding ratings have a positive impact on salary increases and promotions, when the model controls for gender and race, years of education, federal experience, age, supervisory status, PATCO category, and agency. Holding the other variables constant, outstanding performers' salary increases are .06 percentage points higher than fully successful performers, but the result is not statistically significant. The advantage of outstanding ratings is clear in promotion. Outstanding performers' log-odds of

---

[3] Additional analysis (not shown) indicates that years of federal service is negatively associated with salary increase, but more experienced employees were more likely to receive higher performance ratings than less experienced employees. Therefore, the bivariate (or trivariate in this paper) analysis could give us a misleading picture of the impact of ratings on salary increases and promotion probabilities.

promotion are .268 higher than fully successful performers' log-odds of promotion, holding the other variable constant. Those who received exceeds fully successful ratings also appear to be more likely than those who received fully successful ratings to be promoted, though this coefficient is only significant at .05 level in a one-tailed test. Coefficients on *less than fully successful* are negative, as expected, but the results are not statistically significant. The multivariate analysis results contradict the t-test results, implying that we need to consider individual and job-related factors to measure the impact of ratings on salary increases and promotion probabilities correctly.

Table 2.3: Impact of Performance Ratings on Career Advancement: Panel Model

	Salary Increase	**Promotion**
Outstanding	0.059	0.268**
	(1.56)	(8.15)
Exceeds fully	-0.040	0.047
	(1.24)	(1.69)
Less than fully	-0.132	-0.152
	(0.69)	(0.80)

Absolute value of t (for salary increase) and z (for promotion) statistics in parentheses
Model includes individual and organizational characteristics
* significant at 5%; ** significant at 1%

Analysis by Year

Average performance ratings increased steadily until 1999 and the pattern is similar to percent of outstanding rating (Table A.2). However, salary increase and promotion do not change consistently with the year, which rejects the Hypothesis 1.3 and 1.7 that as performance ratings have risen over time, the impact of performance ratings on salary increases and promotion probabilities has decreased. Both average salary increase and the percentage promoted were highest in 1988-1990, but average

performance rating was lowest in those three years. Performance ratings were highest in 1999 with a mean rating of 4.3 and 49.6 percent receiving outstanding rating. They have decreased since then, probably because a substantial number of federal employees moved from the five-level rating system to pass/fail or to three-level rating systems.

Table A.3 presents the t-test results comparing outstanding and fully successful ratings by year. Differences in salary increases and promotions between outstanding and fully successful are statistically significant only in four years. In addition, outstanding performers' salary increase and promotion rate are higher than fully successful performers only in one year and two years, respectively. To some extent, t-test results may be useful to understand basic patterns in the differences of salary increase and promotion rate between outstanding and fully successful ratings. However, as shown in above, a variety of individual and job-related factors are associated with salary increases, promotions, and performance ratings. Therefore, the analytical models need to consider those factors for more accurate measurement.

Table A.4 shows the multiple regression results for percentage change in salary. The results partially support Hypothesis 1.1, that salary increase would be higher for highly rated employees than for poorly rated employees. Outstanding ratings do seem to have a positive impact on salary increases, but the coefficients on other two rating variables are mostly statistically insignificant. The reference group for performance ratings is *fully successful*, so coefficients on each performance rating show whether the named group had higher rates of salary increase than those who received fully successful ratings. Holding the other variables constant, salary increases were 0.2 to 0.6 percentage points higher for those who received *outstanding* than for those who received *fully*

successful in the 1988-2002 period. The coefficient on *outstanding* is largest in 1991, when the average salary increase was 7.6 percent; thus, a 0.6 percentage point difference between those who received outstanding ratings and those who received fully successful ratings is quite small. The results support Hypothesis 1.2, that differences in salary increases between highly rated and poorly rated employees would be small.

Although the average performance rating increased until 1999, the size of coefficients on *outstanding* does not vary much by year. The *outstanding* coefficients were 0.2 to 0.3 in 1988-1990, and increased to 0.6 in 1991, and then were 0.2 to 0.5 in 1992-1999. The three most recent years' coefficients were not statistically significant. Because a substantial number of federal employees went to pass/fail or three rating systems, that change could affect the relationship between outstanding ratings and salary increase. While rating was the least inflated in earlier years and was the most inflated in late 1990s, the impact of outstanding ratings on salary increase does not differ between two periods. This finding rejects Hypothesis 1.3, which hypothesized a negative relationship between rating inflation and the impact of ratings on salary increase.

Salary increases do not differ significantly between *exceeds fully successful* and *fully successful*. Coefficients on *exceeds fully successful* are statistically significant only in 2 of the 15 years. Less than fully successful performers tend to receive lower salary increase than fully successful performers, but the results fall short of statistical significance.

The remaining coefficients show the impact of individual and job-related characteristics on salary increases when they received same performance ratings. Most gender and race coefficients are not statistically significant except for *white female*.

Holding the other variables constant, white women tend to receive 0.2 to 0.5 percentage points higher salary increase than comparable white men.

Education does not seem to have an impact on salary increase. All the coefficients on education (except in 1992) are not statistically significant, which may indicate more educated employees are not more productive than less educated employees in the same grade, supervisory status, and PATCO category. In 1992, one additional year of education raises the expected salary increase by 0.08 percentage points, holding the other independent variables constant.

Federal work experience is consistently negatively associated with salary increases. All the coefficients on years of federal service are negative, and they are statistically significant in all eight years, ranging from -0.29 to -0.43 percentage points. Salary increase rate also drops with age.

Supervisors seem to have a salary increase advantage over nonsupervisors. The rate of salary increase is 0.3 to 0.5 percentage points higher for supervisors than for nonsupervisors (coefficients are significant in 5 years). This percentage difference may not be comparable with the findings of Johnson and Libecap (1989), which measured difference in annual salary between supervisors and nonsupervisors, but the size of supervisors' advantage in salary increase seems to be small compared to 14.2 percentage points advantage in annual salary in Johnson and Libecap's study (1989).

Consistent with the previous studies (Lewis, 1997a), the salary increase rate is higher in lower grades than in higher grades. The reference group for grade variables is Grade 8, so coefficients on each grade level show the differences in salary increase

between the named groups and Grade 8. Salary increase rate for GS 7 or below is higher than Grade 8, but the rate for GS 9 or above is lower than Grade 8.

Employees in Professional and Administrative occupations have an advantage in salary increases. Those who are in Professional category are 0.7 to 4.0 percentage points higher than those who are in the Other category in their salary increase during 1988-2002 periods, holding the other variables constant. Salary increase rates are 1.2 to 3.8 percentage points higher for employees in the Administrative category than for employees in the Other category. The differences in salary increase rate are highest in 1990 and lowest in 2002; coefficients on *Professional* and *Administrative* are 4.0 and 3.8 in 1990, and 0.7 and 1.2 in 2002, respectively. The Technical and Clerical categories do not differ significantly from the Other category in salary increase in most years. Coefficients on Technical are statistically significant only in two years (0.9 in 1990 and 0.9 in 2000). Coefficients on *clerical* are in general negative, but statically significant only in two years (-1.4 in 1988 and -0.9 in 1992). Professional and Administrative occupations in the federal civil service have advantages not only in the level of salary but salary increase rate.

The rate of salary increase does not vary much across federal agencies; on average, the annual salary increase rate is 6.7 percent, with the rate lowest (5.5 percent) in the Social Security Administration and highest (7.8 percent) in the Justice Department. However, salary increase rate is lower in some agencies, particularly Army, Air Force, Navy, and other Defense departments, than in other agencies. Consistent with Lewis (1997a), employees in the Department of Defense and the Army, Air Force, and Navy tend to have lower salary increases. The reference group for agency dummy variables is

other agencies, which includes all the federal agencies that are not included in the model. Air Force employees' salary increase rates are 0.5 to 1.2 percentage points lower than the reference group, holding individual and other position-related characteristics constant. Compared to the employees in *other agencies*, Army employees have disadvantage in salary increase, rating from 0.5 in 1998 to 1.5 percentage points in 1988. Coefficients on *Navy* are also negative and are statistically significant in four years, indicating Navy employees' salary increase is lower than employees in *other agencies*.

Table A.5 shows the results of logit analysis for promotion probability by year. Supporting Hypothesis 1.5, higher ratings are expected to result in higher promotion probability. Outstanding performers are 2.6 to 5.4 percentage points more likely than fully successful performers to be promoted, when all the other variables have mean characteristics. Since, on average, only 14.8 percent of employees are promoted each year during 1988 to 2003, the differences in promotion probability between outstanding performers and fully successful performers are substantial. Therefore, the results do not clearly support Hypothesis 1.6, which hypothesized that difference in promotion probabilities between highly rated and poorly rated employees would be small. Although the results are not consistent across years - coefficients are significant in 9 years but the sign is negative in 2002 - those who received exceeds fully successful ratings also have more chances to be promoted. Coefficients on *exceeds fully successful* of .179 in 1992 and .280 in 1990 (the smallest and greatest among significant coefficients) can be translated to expectation that those who received exceeds fully successful are 1.6 to 3.4 percentage points more likely to be promoted than comparable employees those who received fully successful ratings, among employees with mean characteristics. The

coefficient on less than fully successful is consistently negative, although the coefficients fall short of statistical significance. Castilla (2008) argues that performance ratings have a weaker impact on promotions than on salary increases, but the results of this study show that while only outstanding ratings lead to salary increases higher than those for with fully successful ratings, both exceeds fully successful and outstanding ratings raise promotion probabilities relative to fully successful ratings.

Findings on other control variables are as follows. Race and gender do not seem to have an impact on promotion chances when performance ratings and other individual and job-related factors are controlled. But, unexpectedly, years of education are positively associated with promotion probabilities. Coefficients on education are positive and are statistically significant in 8 years. One additional year of education around its mean raises promotion probability by 0.4 (in 1990) to 0.8 percent points, when the other variables are set their means.

Consistent with Lewis (1997a), federal experience and age have negative impacts on promotion probabilities. The coefficients on federal experience are negative and statistically significant in all 15 years. Age coefficients are negative and significant in four years. Consistent with Asch (2001), supervisors' promotion probabilities do not differ from those of nonsupervisors; the coefficients are consistently positive but significant only in three years.

Promotion probability drops with grade level. The patterns are similar to the relationship between grade level and salary increases. The coefficients are positive until grade 7, but are negative from grade 11 to 15. The reference group for grade dummy variables is Grade 8, so the results can be interpreted that those who are in GS 3-7 are

more likely but those who are in GS 11-15 are less likely than those who are in GS 8 to be promoted.

The percentage of employees who were promoted was lowest in Professional and Administrative occupations. However, when the model controlled for performance ratings and individual and job-related factors, promotion probabilities were the highest among Professional and Administrative occupations. Employees in Professional occupations are 3.9 to 20.6 percentage points more likely to be promoted than employees in the Other category, holding the other variables at their means. Promotion probabilities are 8.3 to 18.6 percentage points higher for those who are in the Administrative category than for those who are in the Other category, when they have mean characteristics. As in the case of salary increase, there is no significant difference in promotion probability among comparable employees in the Technical, Clerical, and Other categories.

Promotion probabilities vary across federal agencies. In particular, consistent with Lewis (1997a), promotion chances are lower in the Department of Defense and the three military departments. Promotion chances are lowest in the Veterans Affairs and Agriculture but are highest in Transportation.

Analysis by Subcategory

Analysis by PATCO Category

The average salary increase is higher for Administrative (7.1 percent) and Professional (6.7 percent) than Technical (6.3 percent) and Clerical (6.5 percent) employees (Table A.6), probably because Professional and Administrative occupations

are two-grade interval series. Promotion probabilities, on the other hand, are higher for Technical and Clerical than for Professional and Administrative occupations.

The t-tests results are shown in Table A.7. Also oddly, the mean salary increase is higher for those who received fully successful than for those who received outstanding in each of five PATCO categories, although the differences are small (from 0.3 to 0.6 percentage points). Promotion probabilities are higher for fully successful performers than for outstanding performers, but the difference is statistically significant only in Professional and Administrative occupations. However, simple comparison of mean percent salary change and percentage promoted between two groups might be misleading, so multivariate analysis is used to measure the impact of ratings in each of five PATCO categories.

Table A.8 shows the results of panel data analysis for percent change in salary increase by PATCO category. Surprisingly, outstanding ratings have significant impact on salary increase only for Technical employees. Holding the other variables constant, outstanding performers' salary increase rate is 0.2 percent points higher than fully successful performers. Coefficients on outstanding in the other four groups are statistically insignificant, implying that performance ratings are not determinants in salary increase within same occupational category, when employees are the same race and gender and supervisory status, have the same levels of education, federal experience, age, and grade, and work in the same agencies.

Table A.9 shows the results of rating impact on promotion probability by PATCO category. Unlike salary increases, for which outstanding ratings had a significant impact only for Technical employees , promotion chances are higher for outstanding performers

than for fully successful employees in all five PATCO categories. The log-odds of promotion for outstanding is .20 and .45 higher than the log-odds of promotion for fully successful in Professional and Other occupations, respectively, holding the other variables constant. Even exceeds fully successful ratings have a significant impact on promotion probabilities in Clerical occupations. The findings indicate that performance ratings are predictors of promotion probability within each occupation group.

Analysis by Grade Level

Salary increase rate and promotion probabilities drop with grade level, except for some middle grades (Table A.10). GS 7 and GS 9 employees' salary increase and promotion probabilities are higher than employees who are one-grade lower than them. It may be because college graduates and master's degree holders who are in Professional and Administrative occupations typically start their careers at GS 7 and GS 9, respectively, and are promoted two levels at a time until GS 11. On the other hand, performance ratings are higher at higher grades.

Table A.11 shows t-test results by grade level. Outstanding performers' salary increase rate is lower than fully successful performers until grade 8, and is consistently higher from grade 12 to 15. The pattern is similar to promotion probabilities. Outstanding performers do not receive consistently higher promotion probabilities than fully successful performers until grade 10, but outstanding ratings have a promotion advantage over fully successful ratings from grade 11 to grade 15.

Table A.12 shows the results of panel data analysis for percent change in salary increase by grade. The coefficients on outstanding are all positive, but are statistically

significant in only five of 14 grades. Receiving an outstanding rating has a significant impact on salary increases for the low grades (GS 1-3 and 5) and middle grades (GS 9, 11 and 12). The impact of outstanding ratings is greatest for employees who are GS 1-3. Outstanding performers' salary increase rate is 1.4 percentage points higher than fully successful performers, holding the other variables constant. This advantage is substantial, considering the mean percent change in salary for GS 1-3 is 7.8. Other than that group, the size of the outstanding rating impact is small or is not statistically significant. In particular, salary increase does not seem to be a function of performance ratings for the top 4 grades. Exceeds fully successful ratings have advantages in salary increase in three grades (GS 5, 9, and 11). For those groups, salary increase rate is 0.2 to 0.4 percentage points higher for exceeds fully successful performers than for fully successful performers. The results indicate that in most grades, outstanding ratings have a small impact on salary increases within same grade, when employees' individual characteristics and job-related factors are controlled for.

The results of rating impact on promotion probability by agency are shown in Table A.13. The patterns are a bit different from those for salary increase. Coefficients on outstanding are positive (except for GS 6), and are statistically significant in five years, indicating outstanding ratings have advantages over fully successful ratings in promotions. Similar to the salary increase, outstanding performers' promotion probability is greater than fully successful performers, when they are low grades (GS 4 and 5) and middle-levels (GS 9, 11, and 10). Most coefficients on *exceeds fully successful* and *less than fully successful* are not statistically significant, implying that

only outstanding ratings have an impact on promotion probability among employees who are in the same grades.

Analysis by Agency

Salary increase and promotion probabilities vary across federal agencies (Table A.14). The Department of Justice's mean percent change in salary is 7.8 percent, whereas that of the Social Security Administration is 5.6 percent. On average, 22.3 percent of Commerce employees were promoted, but only 11.9 percent of Veterans Affairs employees were promoted in 1988-2003 periods. Performance ratings vary dramatically across agencies. While the mean performance rating is 4.5 and 60.6 percent of Army employees received outstanding ratings, mean performance rating is 3.4 and only 11.2 percent of Social Security Administration employees are identified as outstanding performers.

Table A.15 shows the t-test results by agency. Outstanding performers' salary increase and promotion rate are not consistently higher or lower than fully successful performers in federal agencies. For example, the mean percentage change in salary is 1.2 percentage points lower for outstanding performers than for fully successful performers in the Army, but the salary increase rate is 0.6 percentage points higher for outstanding performers than for fully successful performers in the Department of Energy. Because individual characteristics as well as organizational factors are associated with salary increase, promotion, and performance ratings, these t-test results could be misleading.

Table A.16 shows the results of panel data analysis for percent change in salary increase by agency. Oddly, many coefficients on *outstanding* are negative, though the

results are not statistically significant. The coefficients on *outstanding* are significant only in Commerce Department and Other Agencies. Outstanding performers' salary increase rate is 0.7 and 0.6 percent points higher than fully successful performers in Commerce Department and Other Agencies, respectively, holding the other variables constant.

Table A.17 presents the rating impact on promotion probability by agency. Again, patterns for promotion probability are quite a bit different from the case of salary increase. Most coefficients are positive, and outstanding ratings have a significant impact on promotion probability in seven federal agencies. In particular, although the results are statistically insignificant, outstanding ratings were negatively associated with salary increase in three federal agencies (Air Force, Agriculture, and Health and Human Services), but outstanding ratings are significant factors for promotion in those three agencies. On the other hand, differences in salary increases between outstanding and fully successful performers were greatest in the Commerce Department with 0.7 percent point difference, but differences in promotion probabilities are not significant for the department .

In sum, Hypothesis 1.1 indicated that higher ratings would lead to higher salary increases than lower ratings, but the results of panel data analysis only partially support the hypothesis. Supporting Hypothesis 1.2, differences in salary increases between highly rated and poorly rated employees were small, even when the results were statistically significant. In Hypothesis 1.3, I hypothesized a negative relationship between rating inflation and the impact of ratings on salary increase, but the results reject the hypothesis. The results reject Hypothesis 1.4, that the impact of performance ratings

on salary increase would be smaller in agencies where performance ratings were more inflated. Mean ratings were average in Commerce and Other Agencies but their impact on salary increases were greatest in the two agencies, whereas Social Security Administration, Veterans Affairs, and Transportation were the least inflated but outstanding ratings were not associated with salary increase in these agencies.

Overall, the results of panel analysis support Hypothesis 1.5, that highly rated employees were more likely than poorly rated employees to be promoted. The outstanding ratings have substantial advantage over fully successful ratings in promotion probabilities, so the results do not support Hypothesis 1.6, Hypothesis 1.6, which hypothesized that difference in promotion probabilities between highly rated and poorly rated employees would be small. The results reject 1.7 that as performance ratings have risen over time, the impact of performance ratings on salary increases and promotion probabilities has decreased. Hypothesis 1.8 suggested that the impact of ratings on promotion probabilities would be small in inflated agencies, but the results do not support the hypothesis. There is no consistent finding for the relationships between rating inflation of agencies and the rating-promotion decision linkage.

2.7 Conclusions

Linking rewards to performance is a key element for successful performance-based rewards system. However, many observers argue that performance-pay relationship is weak in the federal service, often without providing empirical data. This study provides empirical evidence on the impact of performance ratings on pay increases and promotion probabilities in the federal service. In general, performance ratings do not seem to have much impact on salary increases. Outstanding ratings only have a

significant impact on salary increases in some low and middle grades, in Technical occupations, and in the Commerce department. Although, the results of cross-sectional analysis are statistically significant in most years, the size of the difference in salary increases between outstanding and fully successful performers is small. However, because so many factors come into play in determining salary increases, and because the performance rating-salary increase relationship is not consistent across grade levels, occupation types, and agencies, we must not simply generalize that performance ratings have weak impact on salary increase. For employees in Grade 5, for instance, the t-tests showed that salary increase was significantly higher for those who received fully successful than for those who received outstanding ratings, but when the panel model controlled for individual and job-related factors, outstanding ratings have significant advantages over fully successful ratings in salary increase. Also, while in most occupation types and federal agencies, outstanding ratings did not significantly impact salary increases, the benefit of outstanding ratings in salary increase was significant in Technical occupation and in Commerce Department.

The patterns for promotion probabilities differed substantially from those for salary increases. Outstanding ratings were valuable for promotion in many grades, in all occupation categories, and in some federal agencies. In addition, those receiving outstanding and exceeds fully successful ratings had significant advantages in promotion in most years. These results imply that federal agencies use performance ratings more seriously for promotion decisions than for salary increases, probably because promotion decisions have a greater impact on individual career advancement and have an impact on organizational performance for longer periods of time than annual salary increase

decisions. Therefore, the argument that performance-rewards link is weak could be partially correct, if it considers only pay-performance relationships.

The findings suggest that future studies need to analyze a variety of aspects of the use of performance ratings. According to the CSRA of 1978, federal agencies should use performance appraisal results not only for pay and promotion decisions but for other purposes, including retention and termination. A new OPM director, John Berry, appointed by the President Obama, said that one goal of pay-for-performance reform plans is to retaining good performers by providing them with large enough rewards compared to the private sector companies. However, little is known about whether performance-based reward is effective in retaining good performers.

Although this study provides starting points for understanding the relationship between performance ratings and rewards in a quantifiable manner, the most recent data used for this study is 2003. The federal government has experienced dramatic changes in performance-based reward systems since 2003. Therefore, future study should use more recent data which reflect most important change including pay-for-performance systems in Homeland Security and Defense Departments.

CHAPTER 3

PERFORMANCE RATINGS AND GENDER AND RACIAL INEQUALITIES IN CAREER SUCCESS

3.1 Introduction

Despite governments' efforts to achieve equal employment and affirmative action, gender and racial disparities in career success persist in the U.S. workforce. Shrek (2007) argues that increased use of performance pay is increasing income inequality. Inequality rises, he says, because good performers are rewarded on their high productivity. If performance appraisal provides an unbiased and objective measurement of employee performance, the disadvantages of women and minorities in career success could be explained by their lower performance ratings, which arise from their lower productivity. However, studies found that linkages between performance ratings and rewards are weak (Pearce & Perry, 1983; Rynes et al., 2005; Smith & Rupp, 2003), suggesting the possibility that women and minorities may not be recognized from the organizations in their high performance so that women and minorities' high performance ratings are not connected to the rewards.

Castilla (2008) presents two possible ways in which performance ratings could contribute to the disadvantages women and minorities face in career success. First, discrimination could occur in the performance appraisal process: women and minorities could receive lower performance ratings than comparable white males with the same performance level, which he calls "performance-evaluation bias." Second, women and minorities could receive lower returns on the same level of performance ratings than white males ("performance-reward bias").

Although many studies have examined factors affecting gender and racial differences in career success, studies that try to connect gender and racial inequalities to managerial tools are scarce. This study explores whether performance ratings help explain women's and minorities' disadvantages in career success. Specifically, this study addresses two research questions: 1) Do women and minorities receive lower performance ratings than comparable white males in the federal service? and 2) Do high performance ratings increase salaries and promotion probabilities less for women than men and for minorities than whites? By showing whether discrimination occurs at the performance evaluation stage or at the rewards stage or both, this study adds new perspective on gender and racial inequality in career success.

This paper is composed of seven sections. Section 3.2 presents theories that explain why women and minorities do not achieve the same career success as men and whites, and reviews literature on whether performance ratings are objective and accurate. Section 3.3 reviews how individual and job-related characteristics are related to performance ratings, salary increases, and promotions. Section 3.4 develops hypotheses on the performance-evaluation bias and performance-reward bias. Section 3.5 explains data, variables, and research method. In Section 3.6, I present the findings and discuss hypotheses. Section 3.7 summarizes the findings and discusses the implication of the study.

3.2 Literature Review

Determinants of Career Success

According to human capital theory, investment in education and training raises the productivity of workers, and hence increases their earnings by providing skills and

knowledge that employers reward them for (Becker, 1993; Mincer, 1993). A number of empirical studies support this theory (e.g., Mincer, 1993; Murphy & Welch, 1992). For example, Murphy and Welch (1992) found that college-educated employees earned 40 percent and 58 percent more than high school graduates in 1963 and in 1989, respectively. Barkley et al. (1999) showed that master's degree holders earned 20.3 percent more than bachelor's degree holders in starting salary, after controlling for college major, career experience, family and demographic variables. In fact, empirical studies explained the reduced gender and race-based pay gap due to the increased years of schooling (O'Neill & Polachek, 1993) and to the improved contents of schooling (e.g., high-paying college majors such as engineering and business) for women and minorities (Loury, 1997). However, many studies show the existence of gender and racial disparities in earnings even after controlling for human capital (e.g., Alkadry & Tower, 2006; Ortiz & Roscigno, 2009; Padavic & Reskin, 2002).

Labor market discrimination theory attributes pay disparities and occupational segregation to employers' prejudice against the disadvantaged groups (Becker, 1971; Blau et al., 2006). Employers are reluctant to hire some group of people who are subject to the taste for discrimination from employers themselves, employees, and customers. Under imperfect information on the productivity of job applicants, employers, particularly risk-averse employers, make hiring decision based on past statistics to avoid costly mistakes, so that employers are unwilling to hire some group of people who were on average less productive than others (Phelps, 1972). Employees may not want to work with a certain group of people, and customers also may prefer to be served by some group of people than others. From employers' point of view, hiring people who are

discriminated against by their coworkers and customers is not rational because cost for hiring the disadvantaged group is their wage plus "discrimination coefficient," which measures the strength of discrimination in terms of money whereas cost for hiring the privileged group is just their wage (Blau et al., 2006).

Glass ceiling theory, on the other hand, focuses on the limited advancement of women and minorities in hierarchical organizations. According to Cotter et al. (2001), an "unseen, yet unbreachable barrier … keeps minorities and women from rising to the upper rungs of the corporate ladder, regardless of their qualifications and achievements" (p. 656). Many empirical studies support the glass ceiling theory. Kelly et al. (1991) compared career advancement between men and women in six state governments and found that mobility into the upper-level positions occurs at a higher rate for men than for women. Miller et al. (1999) analyzed city government data, provided by the U.S. Equal Employment Opportunity Commission, and found that women are highly concentrated on redistributive agencies where average pay level is lower than distributive and regulatory agencies. Using three sources of information: the CPDF, focus group interviews, and a survey conducted by the U.S. Merit Systems Protection Board, Naff (1994) analyzed federal agencies and argued that experience and education only partly explained gender differences in advancement. However, in general, the glass ceiling is less prevalent in the public sector than in the private sector. Compared to the private sector, a higher percentage of senior managers in the federal government are women (Stivers, 2002) and a higher percentage of professionals in the public sector are women (Peterson & Lewis, 1999).

Each of these theories explains part of the reason why there is difference in career success between men and women and between whites and minorities, but studies based on these theories do not focus on the human resource management tools by which pay and promotion decisions are made.

Accuracy of Performance Ratings

Performance appraisals stand on the assumption that supervisors can measure subordinates' performance accurately and objectively. According to Milkovich and Wigdor (1991), "supervisors can give reliable ratings of employee performance under controlled conditions and with a carefully developed rating scales. In addition, there is indirect evidence that supervisors can make moderately accurate performance ratings" (p. 66). Attitudes of federal employees on the accuracy of performance ratings are overall positive but not very high. According to the 2005 Merit Principles Survey, 54 percent of federal employees agreed or strongly agreed that "In my work unit, performance ratings accurately reflect job performance," and 67 percent of respondents positively answered that "My supervisor rates my performance fairly and accurately" (U.S. Merit Systems Protection Board, 2005).

However, many kinds of rater errors are hard to avoid. The performance appraisal process involves raters applying performance criteria and standards to individual employees, and different raters may interpret the criteria and standards differently. Not only because of measurement itself, but because of rational behavior of supervisors, performance ratings may not accurately reflect employee performance (Klingner & Nalbandian, 1998; Murphy, 2008; Nigro et al., 2007; Rynes et al., 2005). For example, Murphy (2008) argues that "performance appraisal is a complex event that

occurs in environments that often push raters to distort their ratings to accomplish valued goals or to avoid the negative repercussions of giving ratings their subordinates or superiors will find objectionable" (p. 198). Nigro et al. (2007) also argue that accurate ratings are not rewarded by the organization, and "negative ratings often yield nothing more than stressful interpersonal conflict and time-consuming appeals by resentful workers to suspicious civil service boards" (p. 170).

Some types of rater errors do not particularly distinguish a certain group of employees from others: the halo effect is an inappropriate generalization of one aspect of an employee's performance to other areas, and central tendency error is an inclination to rate employees in the middle of the scale even when their performance is higher or lower than the average. However, other types of errors are related to race, gender, or other nontask factor-based errors. For example, the similar-to-me effect, which refers to the tendency of raters to give higher rating to those who resemble themselves, indicates that nontask factors such as gender, race, and age can affect performance ratings (Boyd, 2004; Kellough, 2002). Some empirical studies have found that performance ratings are lower for women (Griffeth & Bedeian, 1989; Williams & Walker, 1985), racial minorities (Elvira & Town, 2001; McKinney & Collins, 1991), and older employees (Cleveland & Landy, 1981), though others do not.

3.3 Hypotheses

Performance-Evaluation Bias

Many studies have found that the gender and race of the ratees have an impact on employee performance ratings, but have suggested mixed results on performance-

evaluation bias, both on gender and race. While some studies showed that women receive lower performance ratings than men, other studies found no difference in performance ratings between men and women, or found women received higher ratings than men. Miller and Judith (2002) concluded that women consistently received lower performance ratings than men, particularly in male-dominant organizations . Similarly, several studies found that racial minorities receive lower performance ratings than whites because raters give higher ratings to ratees of the same race (Stauffer & Buckley, 2005). Lewis (1997b) found that minorities are less likely than whites to receive outstanding ratings, the highest ratings offered by the federal government.

McKinney and Collins (1991), on the other hand, showed that women received higher performance ratings than men. They tested the possibility of performance appraisal bias in major metropolitan park and recreation agencies by analyzing the combination of raters' and ratees' gender, race, and age, and found that both male and female raters gave a higher ratings to female ratees. In his analysis of federal agencies, Lewis (1997b) found that women received higher performance ratings than white men at the same grades in the same agencies. Drazin and Auster (1987) analyzed a single large financial services organization and found no significant difference in performance ratings between men and women at the same level.

However, DiTomaso and Smith (1996) argued that "there is not clear evidence that white women and minorities are rated lower on performance evaluations" (p. 97). Waldman & Avolio (1991) also argued that there is little or no racial bias in performance ratings and the lower ratings of blacks simply reflect their lower job-related ability and experiences. Based on the review, I test the following hypotheses:

Hypothesis 2.1: Women and minorities received the same performance ratings as comparable white males.

Hypothesis 2.2: Women and minorities received lower performance ratings than comparable men and whites when the positions were upper levels.

Hypothesis 2.3: Women received lower performance ratings than comparable men in male-dominant occupations.

Performance-Rewards Bias

Many studies support performance-reward bias. Gerhart (1990) showed that each additional average performance rating point resulted in a 16 percent increase in current salary for men but only a 10 percent increase for women. Lyness and Judiesch (1999) found that nonpromoted women received higher performance ratings than nonpromoted men and that after controlling for age, education, tenure, and level, promoted women also received higher ratings than promoted men. Lyness and Heilman (2006) suggested that gender explains 6 percent of the variance in performance ratings and that among promoted employees, women's performance ratings are higher than men's ratings. Using data from a large private company, Castilla (2008)[4] found that salary increase is lower for blacks and Hispanics than for comparable whites and that women's salary increase is lower than men's, after controlling for performance ratings. Drazin and Auster (1987) found that at lower levels each increase in performance rating point leads to almost the same salary increases for both men and women, but that at managerial levels each

[4] Castilla did not check performance-evaluation bias, but in his preliminary analysis he found that performance evaluation distributions do not differ by gender and race.

additional performance rating point led to higher salary increase for men than for women. On the contrary, Elvira and Town (2001) found that salary differences between blacks and whites disappeared when performance ratings were controlled.

Hypothesis 2.4: Higher performance ratings gave the same salary increases to men and women and to whites and minorities.

Hypothesis 2.5: Higher performance ratings increased salaries less for women and minorities than for men and whites in upper-level positions.

Hypothesis 2.6: Higher performance ratings increased salaries less for women than for men in male-dominant occupations.

With regard to gender differences in promotion probability, the results of studies are mixed. One group of studies found that promotion probabilities are lower for women and minorities than for white men (Blau & DeVaro, 2006; Cobb-Clark, 2001; Paulin & Mellor, 1996). For example, Paulin & Mellor (1996) found that white men are 17 percentage points more likely than minority men to be promoted, after controlling for performance ratings, and they argued that the system of rules governing men's promotion is different from that of women and minorities. In other words, higher performance ratings lead to higher promotion probabilities for women and minorities, but there is no evidence that performance ratings are related to promotion probabilities for men. Other studies show that women's promotion rate is as high or higher than men (Hersch & Viscusi, 1996; Krull, 2006; Lewis, 1986, 1992). According to Hersch and Viscusi (1996), a higher percentage of women than men receive promotions because women tend to be employed at lower levels of the organization. Using the Department of Defense

Civilian Personnel Data Files, Spyropoulos (2005) showed that women's promotion rate is 23 percent higher than men's, holding demographic and human capital factors constant. Lewis (1986) found that white men and women have similar promotion probabilities in the federal government after controlling for individual characteristics.

Hypothesis 2.7: Higher performance ratings gave the same promotion probabilities between men and women and between whites and minorities.

Hypothesis 2.8: Higher performance ratings had lower promotion probabilities for women and minorities than for men and whites when the positions are upper levels.

Hypothesis 2.9: Higher performance ratings had lower promotion probabilities for women than for men in male-dominant occupations.

3.4 Other Factors Affecting Performance Ratings, Salary Increases, and Promotions

Do women and minorities receive different rewards purely because of their gender and racial status? Performance ratings may vary with individual and position characteristics. Studies also show that individual and job-related factors affect salary (increases) and promotions. Therefore, we need to check whether performance-evaluation bias and performance-reward bias exist, when women and men and minorities and whites have similar individual characteristics and hold comparable positions. This section reviews how individual and job-related characteristics could be related to performance ratings and organizational rewards.

Education

Better-educated workers earn more than less-educated ones because education improves skills and productivity (Becker, 1993; Mincer, 1993) or because education signals a likelihood of productivity (Belman & Heywood, 1991; Hungerford & Solon, 1987). If education raises a worker's productivity, more educated workers' performance ratings would be higher than for those less educated. Lewis (1997b), however, found that probability of receiving outstanding ratings decreased with education in federal agencies. Nevertheless, years of schooling is positively associated with earning levels, though Lewis (1986; Lewis, 1997a) finds education is negatively associated with salary increase and has no or only trivial impact on promotion probability in the federal government.

Federal Experience

Work experience increases job knowledge, which in turn increases workers' performance (Schmidt et al., 1986). Therefore, more experienced workers should be more likely to receive higher performance ratings than less experienced workers. Lewis (1997b) found that the probability of receiving outstanding ratings increased up to 12 to 18 years of federal service, then declined, probably because federal employees' work motivation is higher in their early career than in their late careers. On the other hand, salary increases and promotion probabilities drop with federal experience (Lewis, 1986, 1997a)

Age

According to Vroom (1964), work performance is a function of ability and motivation, which may change over time. But, the results of the empirical studies on the

relationship between age and productivity are mixed. While cognitive ability declines with age, verbal ability is stable throughout working life (Skirbekk, 2003). Avolio et al. (1990) introduce conflict possibility that work motivation is low for older employees but that loyalty to the organization and turnover rate are low for older employees. Therefore, empirical studies have found that age is either negatively related (Cox & Nkomo, 1992; Rhodes, 1983) or not related (McEvoy & Cascio, 1989) or curvilinearly related (Sturman, 2003) with the productivity. Consistent with Struman (2003), Lewis (1997b) showed that likelihood of receiving outstanding ratings increase up to a certain age, then decrease in the federal civil service. Lewis (1997a) also found that salary increases and promotion chances drop with age in the federal agencies.

Sex-type of Occupation

Some federal occupations are female-dominant (Lewis, 1996) and the proportion of women in occupations are associated with performance rating distribution, salary increases, and promotion probability. Federal employees in female-dominant occupations are more likely to receive outstanding ratings than comparable employees in male-dominant occupations, although the relationship is not linear (Lewis, 1997b). Salary level is lower for female-dominant occupations (Catanzarite, 2003; Cohen & Huffman, 2003), but women's promotion probability is high in female-dominant occupations (Reskin et al., 1999).

Grade Level

Performance ratings rise with grade level in the federal civil service (Lewis, 1997b), but salary increase rate and promotion probabilities drop with grade level. Naff

(1994) showed that average promotion rate is about 10 percent or below for GS 13-GS 15, whereas the rate is 33 to 44 percents for GS 9. Lewis (1997a) found that promotion probabilities are highest for GS1-GS3, and decrease with grade level increases. Because advancement to a higher grade leads to about 10 percent increase in salary in the federal civil service (Spyropoulos, 2005), but promotion chances are lower for higher grade employees, salary increases rate would be lower for higher grade employees.

Supervisory Status

Federal supervisors receive higher performance ratings than nonsupervisors (Lewis, 1997b), probably because federal supervisory selection emphasizes technical skills (U.S. Office of Personnel Management, 2001) so that good performers are selected as supervisors, and because the second-line supervisors are reluctant to give lower ratings to the first-line supervisors who are in the leadership positions than nonsupervisors Although the federal government pays supervisors more than nonsupervisors (Asch, 2001; Johnson & Libecap, 1989), salary increase rate and promotion probability may not differ between supervisors and nonsupervisors. For example, Asch (2001) analyzed Department of Defense civil service personnel file from 1982 through 1996 and found that supervisors' promotion speed is not faster than nonsupervisors.

PATCO Category

Federal white-collar positions are classified into five broad occupational categories: Professional, Administrative, Technical, Clerical, and Other (PATCO), based on "the subject matter of work, the level of difficulty or responsibility involved, and the

educational requirements" (U.S. Office of Personnel Management, 2002b, p. 242). Because educational and experience requirements for entry are highest for Professional and Administrative occupations, employees in Professional and Administrative categories have advantages over employees in other three categories in career advancement. In general, promotions are granted two grades at a time from GS 5 through GS 11 for Professional and Administrative occupations, whereas Technical and Clerical employees are promoted one grade at a time (U.S. Office of Personnel Management, 2002a). Because promotion to a higher grade results in higher salary increase than normal annual salary increase, Professional and Administrative employees have a salary advantage. Performance ratings, on the other hand, are highest for Technical and Clerical employees. According to Lewis (1997b), employees in Technical and Clerical occupations are most likely to receive outstanding ratings.

Agency

Performance ratings vary greatly across federal agencies. Using a one percent sample of CPDF from 1990 through 1995, Lewis (1997b) found that Defense employees receive high ratings whereas employees in the departments of Agriculture, Labor, Treasury, Interior, Transportation, and Veterans Affair receive low ratings. Alonso and Lewis (2001) used 1996 Merit Principles Survey and 1991 Survey of Federal Employees and showed that Agriculture, Transportation, Treasury, OPM, and Labor employees are least likely to receive outstanding ratings and that State and Defense employees are most likely to receive outstanding ratings.

Career advancement practices also differ among federal agencies. Borjas (1980; Borjas, 1982) argued that federal salary rate is negatively associated with the size of the agency, and using a CPDF he showed that employee salary rate is highest in Labor Department and is lowest in Defense and Agriculture. Salary increases and promotion probabilities are lower in Defense agencies than in domestic agencies (Lewis, 1997a).

3.5 Data and Methods

Data

As a model employer, the federal government has been trying to rectify gender and racial inequalities in career success. So the federal government is a good site to analyze how performance ratings are associated with gender and racial inequalities in career success. This study analyzes a one percent sample of the Central Personnel Data File (CPDF). The U.S. Office of Personnel Management (OPM) systemically collects a variety of personnel data on federal employees and annually selects a one percent random sample based on the final three digits of the employee's Social Security Number for study purpose. Because an employee remains in the sample as long as she works for the federal government, the samples are random but not independent.

OPM collects employee performance rating information, but due to privacy concerns, it declined to provide updated information on performance rating. The most recent CPDF this study uses is 2003 and the oldest data that contains performance rating information is 1988. Thus, this study analyzes full-time federal employees from 1988 through 2003.

Federal performance appraisals have been decentralized since 1995 and a substantial number of federal employees have moved from five-level rating systems to pass/fail or other rating systems. In this sample, employees who are under pass/fail system increased from 5.5 percent in 1998 to 36.6 percent in 2003, whereas those who are under five-level ratings decreased to from 92.1 percent in 1998 to 53.6 percent in 2003. However, the results of pass/fail (or three level rating systems[5]) do not provide meaningful information with which to make personnel decisions (Liff, 2007; Montoya & Graham, 2007), this study restricted the sample to employees under the five-level rating system.

Dependent Variables

The dependent variable for testing performance-evaluation bias is *performance rating*. Federal employees receive one of five rating levels, ranging from level 1 (unacceptable) to level 5 (outstanding). However, because only a few employees receive levels 1 and 2, I combined the three lowest rating categories to enhance the appropriateness of the analysis.[6] Therefore, *performance rating* is an ordinal level variable with three categories.

To check performance-rewards bias, I use *percent change in salary* and *promotion* as dependent variables, because salary and promotion are the best-known

---

[5]Analysis on this data shows that most employees who are under three-level ratings received mid level or higher. For example, only one out of 762 employees under three-levels received the lowest rating in 2003. In addition, there was no difference in salary growth and promotion probabilities between the highest and the mid-level employees.

[6] In addition, no one in some race and gender groups received either level 1 or 2. This caused the Brant test of the proportionality of odds assumption to fail.

rewards for outstanding employees. One of the greatest salary determinants in the federal government is an employee's grade. On average, women's grade is lower than men, and therefore there is possibility that even with women's higher salary growth rate, salary growth in dollar amount can be lower for women than for men. So percentage salary change is the better measurement than dollar amount. The variable *percent change in salary* is created by the formula: ((next year's salary – this year's salary) / this year's salary) × 100. The variable *promotion* is coded 1 if his or her next year grade is higher than this year grade and coded 0 for everyone else.

Independent Variables

Model for Performance-Evaluation Bias

The key independent variables are a set of race and gender groups. OPM classifies employees into five racial groups: white, black, Hispanic, Asian, and Native American. However, because the sample included so few Native Americans, I dropped them and created seven race and gender dummy variables: *Black male, Hispanic male, Asian male, White female, Black female, Hispanic female, and Asian female*. The reference group is white male.

OPM categorizes occupational group and occupational series, which provide full structure of federal white collar occupations. Occupational group is a major category, which embraces "a group of associated or related occupations," and occupational series are "a subdivision of an occupational group or job family consisting of positions similar as to specialized line of work and qualification requirements" (U.S. Office of Personnel Management, 2009). Some occupational series are male-dominant and the gender composition of an occupation may affect performance rating distribution. Based on the

OPM classification on occupational group and series, I create 9 dummy variables for the sex-type of the occupational series, from *less than 10 percent male* to *80 to 89 percent male* in occupation. The reference group is occupational series where *90 percent or higher are male*.

Performance ratings vary greatly by grade and agency, implying that gender and racial differences in performance ratings can arise from the differences in their grade and the agency in which they work. To control for grade, I create 14 dummy variables; one variable for each grade from grade 4 to 16[7] and grades 1 through 3 are combined due to small sample size. I also create 21 agency dummy variables including all cabinet-level departments and four largest independent agencies in terms of the number of employees, which are Environmental Protection Agency, General Services Administration, National Aeronautics and Space Administration, and Social Security Administration. All the other agencies are combined as *other agencies* and used as the reference group.

Federal white-collar occupations are classified into five categories: Professional, Administrative, Technical, Clerical, and Other (PATCO). Because the practice of personnel management varies with PATCO categories, I include four PATCO dummy variables, with *Other* as the reference group.

This study controls for individual characteristics. The CPDF does not have information on evaluators' characteristics such as race, gender, and age, which may affect evaluatees' levels of performance ratings, so I include only evaluatee's characteristics. The variable *supervisory status* is a dummy variable coded 1 for supervisor or manager

[7] The General Schedule has 15 grades, but for convenience, employees in the Senior Executive Service are coded as grade 16.

and 0 for everyone else. *Education* is an interval level variable, measured in years. Federal experience and age are measured in years, and squared terms for both variables are included in the model to capture any curvilinear effects.

Model for Performance-Reward Bias

Basically the models for testing performance-reward bias are same as the model for performance-evaluation bias. To test performance-reward bias, which means women and minorities would receive lower returns on the same performance ratings, seven race and gender dummy variables are used: *black male, Hispanic male, Asian male, white female, black female, Hispanic female, and Asian female*. I compare each of these race and gender groups with *white male*, the reference group, in salary change and promotion probability.

Some studies argue that there is occupational sex segregation and that "women's work" pays worse (Catanzarite, 2003; Cohen & Huffman, 2003). To examine whether occupational sex-type affects racial and gender differences in rewards, I include nine dummy variables from *less than 10 percent* male to *80-89 percent male in occupation*. PATCO categories also are associated with male-female composition and pay level. Less women than men work in Professional and Administrative occupations, although the gap has been decreasing in recent years, and Professional and Administrative occupations pay better than three other categories. So, four dummy variables for PATCO categories are included in the model and *Other* category is used as the reference group.

An employee's salary level is largely determined by his grade, but percent change of salary may differ among grade levels. Promotion probability also drops with grade

levels. I create a dummy variable grade 3 for those who are in grade 1, 2, or 3, and a set of dummy variables for each of every grade level from grade 4 to 16. Each coefficient will be compares with grade 8, the reference group.

As Condrey and Brudney (1992) showed, the practice of performance-related pay varies with agencies. I create dummy variables for all cabinet level departments and four independent agencies: Environmental Protection Agency, General Services Administration, National Aeronautics and Space Administration, and Social Security Administration, agencies that employ the largest number of workers.

Individual characteristics are important factors in determining pay level and promotion probabilities. Supervisory status is a dummy variable coded 1 for supervisors or managers and 0 for everyone else. Education, federal experience, and age are all interval level variables and measured in years. To capture curvilinear effect, I include squared terms for federal experience and age.

Methods

Models for Performance-Evaluation Bias

Since the dependent variable *performance rating* is an ordinal level variable, ordered logit is the appropriate model to test performance-evaluation bias. Although *performance rating* has three values (1 for fully successful or lower, 2 for level between fully successful or lower and outstanding, and 3 for outstanding), the difference between 1 and 2 is not same as the difference between 2 and 3. Therefore, ordinary least square regression, which assumes a linear relationship between the dependent variable and the independent variables, is not appropriate for this study. The multinomial logit analysis

can be used for discrete outcomes but has a weakness in that it cannot account for the ordinal nature of dependent variable.

The ordered logit model assumes the proportionality of odds, that is, the odds ratios are the same at all cut points. For example, the first column (1988) in Table 22 implies that the odds that a white woman will receive an *outstanding* rather than a lower rating is 1.5 times as high as the odds for a comparable white man and that the odds that she will receive an *outstanding* or *exceeds fully successful* rather than a lower rating is 1.5 times as high as his odds. Brant test results show that the ordered logit model for this study violates the proportionality of odds assumption. However, because the sample size is so large that the Brant test can detect even minor violations of the assumption and most key independent variables are not statistically significant in most years, I use the ordered logit analysis.

To test the hypothesis that women and minorities received the same performance ratings as comparable men and whites, I run a set of cross-sectional ordered logit models by year from 1988 to 2003. Because position and organizational characteristics may affect performance ratings so that gender and racial differences in performance ratings vary with those factors, I conduct additional analyses, which compare performance ratings among race and gender groups, by PATCO category and agency. In order to conduct the additional analysis, it is necessary to have large enough sample sizes. Therefore, I combine 15 years of cross-sectional data and use pooled data analysis. To test the hypothesis that women and minorities received lower performance ratings than comparable men and whites when the positions were upper levels, I run ordered logit models by grade level. To test whether women receive lower performance ratings than

men in male-dominant occupations, I run another series of ordered logit analyses by sex-type of occupation.

Models for Performance-Rewards Bias

To test the possibility of performance-rewards bias, this study analyzes the impact of performance ratings on salary increases and on promotions separately. First, I compare salary increases among each race and gender group in each of three performance ratings (fully successful and lower, exceeds fully successful, and outstanding), after controlling for individual and job-related characteristics. This analysis, which is pooled OLS, will tell us whether the impact of higher ratings on salary increases is greater or smaller for women and minorities than for comparable white men. In addition, analysis with interaction terms allows us whether the differences in the impact of the higher ratings on salary increases among race and gender groups are statistically significant.

This study test the hypothesis that the return on an *outstanding* rating is lower for women and minorities in upper-level positions and in male-dominant occupations . Using cross-sectional OLS by year, I examine whether women receive lower salary increases than men in male-dominant occupations.

Following the same process as in salary increases, this study first examines the impact of performance ratings on promotion probabilities by race and gender groups, and confirms whether the difference of the impact among groups are statistically significant using additional analysis with interaction terms. Then, I check whether women and minorities have lower promotion probabilities than comparable men and whites when the positions are upper levels. Finally, to test the hypotheses that women had lower

promotion probability than comparable men in male-dominant occupations, I restrict the sample to those who are under male-dominant occupations where over 70 percents are male employees.

3.6 Findings

Basic Patterns

Table 3.1 shows the differences in mean characteristics between white males and each of seven race and gender groups in the 1988-2003 period. Numbers in the *White males* row represent the mean values for white males. Thus, 30.9 percent of white men received outstanding ratings over this period, 12.8 percent received promotions, and their mean salary increase was 6.4 percent. The differences of the values from white males are obtained using regression analysis and are mostly statistically significant. Thus, black males were 4.5 percentage points less likely to receive outstanding ratings but 2.2 percentage points more likely to be promoted, and their mean salary increase (6.7 percent) was 0.3 percentage points larger than white males. Interestingly, percentage of receiving outstanding ratings is lower for other male groups than for white men, but the percentage is higher for women (except for black women) than for white men. Blacks and Hispanic men were about 4.5 percentage points less likely to receive outstanding ratings than white men, whereas white and Asian women were about 7.7 percentage points more likely to do so than white men.

Table 3.1: Differences in Mean Characteristics by Race and gender

	% Out-standing	% Salary Change	% Promo-tion	Grade	Education	Federal experience	Age	Supervisor	Professional Administrative
White males[a)]	30.9	6.4	12.8	10.8	15.2	14.7	45.3	23.1	71.4
Differs from white males for:									
Black males	-4.5**	0.3**	2.2**	-1.7**	-1.0**	-0.8**	-0.6**	-7.2**	-19.2**
Hispanic males	-4.6**	0.6**	4.1**	-1.1**	-0.7**	-2.3**	-2.1**	-6.9**	-15.5**
Asian males	-0.9	0.1	1.6**	-0.1	0.4**	-2.5**	0.8**	-8.3**	5.4**
White females	7.8**	0.7**	5.0**	-2.4**	-1.3**	-1.2**	-1.2**	-11.8**	-25.3**
Black females	-0.9**	0.6**	5.8**	-3.2**	-1.8**	-0.6**	-3.8**	-14.7**	-37.7**
Hispanic females	1.2	1.0**	9.0**	-3.1**	-1.7**	-2.4**	-4.1**	-15.1**	-34.1**
Asian females	7.7**	0.7**	5.6**	-2.6**	-0.5**	-3.4**	-2.0**	-15.0**	-21.3**
Observations	228,409	143,015	214,487	228,409	228,394	228,409	228,407	228,307	228,409

a) Numbers in White males row represent mean values for white male in each column, and coefficients on other race and gender groups show the differences of the values from white male

* significant at 5%; ** significant at 1%

Average salary increases and promotion rates were higher for other race and gender groups than for white men. Except for Asian men, coefficients on two male race groups and all four female race groups are positive (ranging from 0.3 to 1.0 percentage points) and statistically significant. Hispanic women's salary increase rate was the highest, which was 1 percentage point higher than white men's. Likewise, all seven race and gender groups have higher average promotion rates than white men; the differences range from 1.6 (Asian men) to 9.0 (Hispanic women) percentage points.

The mean grade is higher for white men than for other groups, however. White men's mean grade is 12.8, and six race and gender groups' mean grades are 1.1 (Hispanic men) to 3.2 grades (black women) lower than white men's. In general, women's mean grades are lower than men's.

The seven race and gender groups' higher mean values on the three dependent variables but lower mean grades than white men might be related to their lower human capital levels. They are 0.5 to 1.8 years less educated, 0.6 to 3.4 years less experienced, and 0.6 to 4.1 years younger than white men, except for Asian men, who are a bit more educated and older than white men. In general, mean years of education and mean age are lower for women than for the three male minority groups.

White men are more likely than other race and gender groups to be supervisors and to work in Professional and Administrative occupations. The percentage of white men who are supervisors is 23 percent and other race and gender groups are 7 to 15 percentage points less likely than white men to be supervisors. Altogether, 75.5 percent of white men work in either Professional or Administrative categories, but black and Hispanic men are 19 and 15 percentage points less likely than white men to work in those

occupational categories. Women are much less likely than white men and three other male minority groups to work in Professional and Administrative categories, with coefficients ranging from 21 to 38.

In sum, white men have lower average ratings, salary increases, and promotion rates than the other seven race and gender groups but higher mean grades, percentages who are supervisors, and percentages who work in Professional and Administrative. The following analysis examine whether the patterns exist even after controlling for individual and job-related characteristics.

Performance-Evaluation Bias

Table 3.2 reports the ordered logit results of the race and gender differences in performance ratings by year from 1988 to 2003. The model includes individual and organizational factors (but these are not shown). The reference group for the set of race and gender dummy variables is white men. So, positive logit coefficients indicate that a group is more likely to receive higher ratings than white men, the reference group.

Asian and white women are the most likely and black men are the least likely to receive higher ratings. Holding the other variables constant, white women's log-odds of receiving outstanding ratings are 0.3 to 0.4 higher than the log-odds of white men over the period of 16 years. The logit coefficients can be translated to expectations that white women were 6.8 percentage points more likely than white men to receive outstanding ratings and 8.6 percentage points less likely than white men to receive less than fully successful or lower ratings in 1988, holding the other variables at their means. The differences between white women and men in receiving outstanding ratings do not

change much over time. White women were 6.8 percentage points more likely to get outstanding ratings in 1988 and 6.6 percentage points more likely to do so in 2002. Asian women were 6.9 and 14.9 percentage points more likely to receive outstanding ratings than white men in 1988 and in 2002, respectively, when the other variables have mean characteristics. On the other hand, black men's log-odds of receiving outstanding ratings are 0.19 to 0.29 lower than white men's log-odds. Other groups are not different from white men in receiving higher performance ratings. Therefore, the results do not support Hypothesis 2.1, which hypothesized that women and minorities received the same performance ratings as comparable white men. Although average performance ratings have increased over time, difference in receiving outstanding ratings among race and gender groups do not vary much with time.

Table 3.2: Race and gender Differences in Performance Appraisal Ratings by Year: Ordered Logit Model

	1988	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**
White females	0.405**	0.358**	0.389**	0.319**	0.419**	0.380**	0.367**	0.356**
	(7.25)	(6.63)	(7.40)	(6.16)	(8.21)	(7.51)	(7.15)	(6.84)
Black females	0.026	-0.089	-0.080	-0.042	-0.054	0.056	-0.017	0.019
	(0.35)	(1.21)	(1.12)	(0.60)	(0.78)	(0.81)	(0.25)	(0.27)
Hispanic females	-0.309*	-0.071	0.208	0.094	0.238	0.045	0.110	0.027
	(1.98)	(0.48)	(1.49)	(0.67)	(1.77)	(0.34)	(0.83)	(0.20)
Asian females	0.384*	0.561**	0.138	-0.028	0.263	0.177	0.110	0.203
	(2.13)	(3.25)	(0.84)	(0.18)	(1.67)	(1.14)	(0.71)	(1.34)
Black males	-0.185*	-0.221*	-0.256**	-0.289**	-0.206*	-0.244**	-0.224**	-0.230**
	(2.03)	(2.51)	(2.98)	(3.43)	(2.52)	(2.96)	(2.72)	(2.73)
Hispanic males	-0.214	-0.096	-0.055	-0.290*	-0.235	-0.095	-0.181	-0.194
	(1.53)	(0.74)	(0.43)	(2.28)	(1.92)	(0.79)	(1.52)	(1.66)
Asian males	0.218	0.097	0.032	0.099	0.139	0.093	0.121	0.131
	(1.33)	(0.62)	(0.22)	(0.68)	(1.00)	(0.68)	(0.89)	(0.95)
Observations	10197	10719	11038	11194	11519	11559	11321	11071

Table 3.2 Continued

	1996	1997	1998	1999	2000	2001	2002	2003
White females	0.382**	0.364**	0.331**	0.368**	0.391**	0.421**	0.281**	0.396**
	(7.04)	(6.62)	(5.57)	(5.16)	(5.19)	(5.22)	(3.14)	(4.48)
Black females	-0.017	0.050	0.053	0.029	0.109	0.027	-0.149	-0.034
	(0.23)	(0.70)	(0.68)	(0.31)	(1.13)	(0.27)	(1.34)	(0.31)
Hispanic females	0.140	0.323*	0.267	0.303	0.522**	0.430*	-0.187	-0.042
	(1.08)	(2.41)	(1.88)	(1.81)	(3.00)	(2.32)	(0.95)	(0.23)
Asian females	0.539**	0.519**	0.482**	0.314	0.438*	0.713**	0.639*	0.608*
	(3.41)	(3.25)	(2.85)	(1.61)	(2.09)	(3.02)	(2.44)	(2.40)
Black males	-0.260**	-0.168	-0.192*	-0.256*	-0.295*	-0.219	-0.229	-0.293*
	(2.96)	(1.92)	(2.07)	(2.33)	(2.51)	(1.71)	(1.62)	(2.16)
Hispanic males	-0.160	-0.032	0.125	-0.015	0.060	0.019	0.001	-0.030
	(1.29)	(0.26)	(0.91)	(0.10)	(0.37)	(0.11)	(0.01)	(0.17)
Asian males	0.077	0.040	-0.008	-0.044	-0.000	0.041	-0.029	0.034
	(0.52)	(0.26)	(0.05)	(0.24)	(0.00)	(0.19)	(0.12)	(0.14)
Observations	10282	9958	8906	6563	5959	5591	4732	5099

Absolute value of z statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1

Table A.18 shows the ordered logit results for performance evaluation bias by PATCO category. The patterns of gender and racial differences in performance ratings vary a bit among PATCO categories. In Professional occupations, black men and women are expected to receive lower ratings than comparable white men, whereas white women tend to receive higher ratings than comparable white men. Three race and gender groups (white women, Asian women, and Asian men) are more likely than comparable white men to receive higher ratings in Administrative occupations. Interestingly, all four female groups are expected to receive higher performance ratings than white men, whereas two male groups, black and Hispanic, are expected to receive lower ratings than white men in Technical occupations. In the Other category, performance ratings do not differ among gender and racial groups, except that Hispanic men are less likely than white men to receive higher ratings.

In all PATCO categories except Other, white women receive higher ratings than comparable white men but black men receive lower ratings than comparable white men. Asian women also are more likely than white men to receive higher ratings in Administrative, Technical, and Clerical occupations. The results again reject the Hypothesis 2.1 that women and minorities received the same performance ratings as comparable white men.

As shown in Table A.19, differences in performance ratings among gender and racial groups vary across federal agencies. For example, while all four female groups were expected to receive higher ratings than comparable white men in the Army, no female group received higher ratings than white men in the Social Security Administration and NASA. In general, female groups tend to receive higher ratings but

minority male groups tend to receive lower ratings than comparable white men. White women receive higher ratings than white men in 17 of 24 agencies. Holding the other variables at their means, white women are 16.2 and 5.8 percentage points more likely than white men to receive outstanding ratings in the Labor and Health and Human Services, respectively. Asian women and Hispanic women are expected to receive higher ratings than comparable white men in 9 and 7 of 24 agencies, respectively. However, black women receive higher ratings only in three agencies but receive lower ratings in four agencies.

On the contrary, minority male groups, particularly black men, consistently receive lower performance ratings. While black men do not receive higher ratings than comparable white men in any federal agency, they receive significantly lower ratings than white men in half of the federal agencies. In particular, black men are 44.2 and 17.0 percentage points more likely to receive fully successful or lower ratings than white men in the EPA and NASA, respectively, holding the other variables at their means.

Table A.20 shows the ordered logit results that test Hypothesis 2.2, that women and minorities received lower performance ratings than comparable men and whites in upper-level positions. There is no clear tendency that as grade level increases women and minorities' probability of receiving higher ratings would decrease. However, all four female groups receive higher ratings than comparable white men in grade 5 through 7, whereas only white women receive higher ratings in grades 4, 14, and 16. No minority male group receives lower ratings than white men until grade 6, but above grade 6, black and Hispanic males receive lower ratings than white males in 6 and 4 grade levels, respectively.

As shown in Table A.21, there is also no clear evidence that female groups receive lower ratings than comparable white men in male-dominant occupations. White and Asian women are 9.6 and 3.7 percentage points more likely to receive outstanding ratings than white men when 71 to 80 percent of workers in an occupation are mean, holding the other variables at their means. White women are expected to receive higher ratings than white men even when the occupations are 91 percent or above men. On the other hand, female-dominant occupations do not guarantee that women could receive the same performance ratings as comparable men. In female-dominant occupations, where over 90 percent are women, black women are 7.0 percentage points less likely to receive outstanding ratings than white men and other three female groups do not differ in receiving outstanding ratings from white men, when the other variables have mean characteristics. Therefore, the results do not support the Hypothesis 2.3 that women received lower performance ratings than comparable men in male-dominant occupations. Regardless of male-female composition in the occupations, black men and Hispanic men receive lower ratings than comparable white men.

Performance-Rewards Bias

Table 3.3 shows pooled OLS results of the impact of performance ratings on percent change in salary increases for eight race and gender groups. The reference group is those who received fully successful or lower ratings, so the coefficients on *outstanding* and *exceeds fully successful* show the percentage difference in salary increases from those who received the three lowest ratings. Among white males, outstanding performers tend to receive 0.3 percentage points bigger pay increases than fully successful or lower

performers, holding the other variables constant. Unexpectedly, the impact of outstanding ratings on salary increase is greater for female groups than for white men. Compared to fully successful or lower performers, outstanding performers' salary increases are 1.3, 0.9, and 0.5 percentage points larger for Hispanic, Asian, and black females, respectively. White women who received outstanding ratings tend to receive 0.4 percentage points more salary increases than comparable white women who received fully successful or lower ratings. While salary increases are 0.5 percentage points higher for outstanding performers than for fully successful or lower performers for black males, there is no significant difference in salary increase rate between those who received outstanding and those who received the three lowest ratings among Hispanic men and among Asian men.

Are the differences in the impact of outstanding ratings on salary increases between white men and other groups significant? The results of the OLS model with interaction terms show that outstanding ratings give an additional 1.0 and 0.6 percentage points salary increase for Hispanic women and Asian women than for white men. The impact of outstanding ratings is 0.2 percentage points larger for black women than for white men. The coefficients for other groups are not statistically significant, indicating the advantages of outstanding ratings over fully successful or lower ratings in salary increases do not differ significantly between white men and other four groups. Thus, the results reject the Hypothesis 2.4, that higher performance ratings increase salaries by similar amounts for men and women and for whites and minorities.

Table 3.3: Difference in the Impact of Performance Ratings on Percent Change in Salary Growth: Pooled OLS Model

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.283**	0.415**	0.480** [B]	1.289** [A]	0.884** [B]	0.471**	0.347	0.111
	(6.15)	(7.16)	(5.36)	(5.66)	(3.13)	(3.50)	(1.54)	(0.42)
Exceeds fully successful	-0.012	0.064	0.206* [B]	0.590** [A]	0.215	0.013	0.087	-0.125
	(0.29)	(1.13)	(2.53)	(2.76)	(0.78)	(0.11)	(0.45)	(0.54)

Absolute value of t statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%
A: Differences in coefficients between white male and other groups are statistically significant at .01 level, B: significant at .05 level

The coefficients on *exceeds fully successful* are positive and significant for black and Hispanic women. Holding the other variables constant, those who receive exceeds fully successful ratings tend to receive 0.6 and 0.2 percentage points more salary increase than those who received fully successful or lower ratings among Hispanic and black women, respectively.

As shown in Table A.22, the patterns are a bit different in the upper grades (grade 14 or above). Outstanding ratings have significant impact on salary increases only for white men and black men. Holding the other variables constant, white men who received outstanding ratings are 0.4 percentage points bigger salary increases than white men who received fully successful or lower ratings. Among black men, the salary increase is 0.4 percentage points bigger for outstanding performers than for fully successful or lower performers. The other six groups, on the other hand, do not have advantages in salary increases when they received outstanding ratings rather than fully successful or lower ratings. Exceeds fully successful ratings do not have impact on salary increases in any of the eight groups. Therefore, the results partially support the Hypothesis 2.5, that higher performance ratings had lower salary increases for women and minorities than for men and whites when the positions are upper levels.

In the lower level grades (grade 5 or below), receiving outstanding ratings rather than fully successful or lower ratings have advantages in salary increases for three female groups and white men. Outstanding performers receive 1.2 percentage points, 0.5 percentage points, and 0.3 percentage points more salary increases than fully successful or lower performers, among Hispanic women, black women, and white women, respectively. While outstanding ratings do not have an impact on salary increases for the

other three male groups, outstanding performers receive 0.6 percentage points higher salary increases than fully successful performers among white men. In other words, white men have advantages in salary increases when they receive outstanding ratings either in the upper level positions or lower grade levels. For Hispanic women and black women, even exceeds fully successful performers receive higher salary increase than fully successful or lower performers when they in the lower level grades.

Table A.23 presents the OLS results that test the Hypothesis 2.6, that higher performance ratings increased salaries less for women than men in male-dominant occupations. The results do not support the hypothesis; however, women may have had disadvantages in male-dominant occupations but advantages in female-dominant occupations. Among female groups, outstanding ratings had a significant impact on salary increases only for white women in male-dominant occupations, but for Hispanics, blacks, and whites in female-dominant occupations. On the other hand, outstanding performers receive higher salary increases than fully successful performers among white men and among black men in male-dominant occupations, but only white men receive higher salary increases when they receive outstanding ratings rather than fully successful ratings in female-dominant occupations. However, differences in the impact of outstanding ratings on salary increases among race and gender groups are not statistically significant.

Table 3.4 shows the results of pooled logit analysis on the impact of performance ratings on promotion. Outstanding ratings have a substantial impact on promotions in most groups (except for Asian males). Coefficients on *outstanding* are largest for Hispanic women and smallest for white men. Outstanding performers are 12.9 and 4.6

Table 3.4: Difference in the Impact of Performance Ratings on Promotion: Pooled Logit Model

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.449**	0.372**	0.424**	0.937** [A]	0.596**	0.482**	0.584**	0.051
	(11.45)	(10.02)	(7.29)	(6.60)	(3.27)	(5.00)	(3.75)	(0.24)
Exceeds fully Successful	0.131**	0.122**	0.238**	0.456** [B]	0.102	0.169*	0.292*	0.047
	(3.73)	(3.41)	(4.53)	(3.46)	(0.57)	(2.00)	(2.20)	(0.26)
Observations	59,913	44,364	17,465	3,062	2,171	7,948	3,349	2,323

Absolute value of z statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%
A: Differences in coefficients between white male and other groups are statistically significant at .01 level, B: significant at .05 level

percentage points more likely to be promoted than fully successful performers among Hispanic women and white women, respectively, holding the other variables at their means. Exceeds fully successful performers also have higher promotion chances than those who received the three lowest ratings in six groups. The *outstanding* coefficient is larger for Hispanic men and women, Asian women, and black men than for white men, but the difference is statistically significant only for Hispanic women. Although coefficients on *exceeds fully successful* vary among race and gender groups, the difference is significant only between white men and Hispanic women. Therefore, the impact of higher performance ratings on promotions does not differ between white men and other six race and gender groups, which partially supports the Hypothesis 2.7, that higher performance ratings increased promotion probabilities by the same amount for men and women and for whites and minorities.

As shown in Table A.24, outstanding ratings have significant impact on promotion only for black men and white men in the upper levels. Holding the other variables at their mean, those who received outstanding ratings are 6.5 and 3.6 percentage points more likely to receive promotions than those who received the three lowest ratings among black men and white men, respectively. In the upper levels, even exceeds fully successful ratings have advantage over fully successful or lower ratings in promotion among white men. Therefore, the results partially support the Hypothesis 2.8, which hypothesized that higher performance ratings increased promotion probabilities less for women and minorities than for men and whites in upper-level positions.

The patterns are quite different in the lower grades. Outstanding ratings give a higher chance of promotion than fully successful or lower ratings for three female groups

(Hispanic, black, and white) as well as for white men and Hispanic men. Although coefficients on outstanding are greater for Hispanic women and men than for white men, the differences are not statistically significant, indicating the impact of outstanding ratings does not differ among race and gender groups.

Table A.25 shows the results of logit analysis, which tested Hypothesis 2.9, that higher performance ratings increased promotion probabilities less for women than men in male-dominant occupations. Both outstanding ratings and exceeds fully successful ratings have significant impact on promotion probabilities for three male groups but for only one female group, which is white women, in male-dominant occupations. On the contrary, outstanding performers have higher promotion probabilities than fully successful or lower performers for three female groups (Hispanic, black, and white) but for only one male group, white men, in female-dominant occupations. In addition, compared to the lowest three ratings, exceeds fully successful ratings have advantages in promotion among Hispanic women and among black women in female-dominant occupations. Although differences in the impact of higher ratings on promotion among race and gender groups are not statistically significant, which do not support the Hypothesis 2.9, most women groups seem to have disadvantages in promotion with the same higher ratings as comparable men in highly male-dominant occupations.

3.7 Conclusions

This study examined whether performance appraisals contribute to race and gender inequalities in career success. To understand whether discrimination occurs in the evaluation process or reward process, this study tested both possibilities separately. First,

women receive equal or higher performance ratings than comparable white men. Women, particularly white and Asian women, are most likely to receive higher performance ratings, after controlling for individual and job-related characteristics. On the other hand, some minority male groups, particularly black men, tend to receive lower ratings than comparable white men. However, the patterns vary with PATCO category, agency, and grade level. For example, black women are expected to receive lower ratings than comparable white men in Professional occupations, but they tend to receive higher ratings than white men in Technical occupations. Black and Hispanic men are equally likely to receive higher ratings as comparable white men until grade 6, but they are expected to receive lower ratings in many grade categories than white men in grade 7-16. Second, the returns on outstanding ratings do not differ between women and minority male groups and white men. Rather, outstanding ratings give additional salary increases for Hispanic, Asian and black women than for comparable white men. The impact of outstanding ratings on promotion probabilities is greater for Hispanic women than for white men. But women groups seem to have disadvantages in salary increases and promotion with the same higher ratings as comparable men in upper-level positions and highly male-dominant occupations.

This study suggests that performance appraisals do not contribute to the disadvantages of women and minorities in career success. In fact, performance ratings may have weak impact on salary increases and promotion probabilities. When I didn't control for performance ratings and compared salary increases and promotion probabilities among race and gender groups, the results did not much differ from the models which controlled for performance ratings. The managerial tools, performance

appraisals, do not seem to be related to gender and racial inequality in career success in the federal civil service.

CHAPTER 4

CONCLUSION

For performance-based rewards to be successful, there should be a clear link between performance ratings and rewards, and a meaningful difference in rewards between good and poor performers. Although many observers attribute the ineffectiveness of performance-based rewards in the public sector to the weak link between performance and rewards, little is known about the strength of the linkage. This dissertation attempts to provide empirical evidence on the impact of performance ratings on personnel decisions in two essays.

The first essay finds that performance ratings have only a limited impact on salary increases. Those who received outstanding ratings are expected to receive higher salary increases than those who received fully successful ratings only in Technical occupations, in five of 14 grades, and in the Commerce Department. In addition, the size of the impact on salary increases is very small. This study measured the impact of performance ratings on salary increases separately by year, PATCO category, grade level, and agency, but even when coefficients were significant, salary increases differed by only 0.2 to 1.4 percentage points between those who received outstanding ratings and comparable employees those who received fully successful ratings. The impact of *exceeds fully successful*, the second highest rating in the federal civil service, on salary increases is also minimal. Coefficients on *fully successful* are not significant any of PATCO category, and are significant only 2 of 15 years and in 3 of 14 grade groups.

The impact of outstanding ratings on promotion probabilities is larger. Promotion probabilities are significantly higher for those who received outstanding ratings than for comparable employees who received fully successful ratings in 11 of 15 years and in all PATCO categories. The results imply that federal agencies use performance appraisal results more seriously for promotion decisions than for salary increase decisions. However, the strength of the linkage between performance ratings and promotion varies among grade levels and agencies. While outstanding ratings have significant impact on promotion probability in some low (GS 4 and GS 5) and middle grades (GS 9, 11, 12), promotion probabilities do not differ between outstanding performers and fully successful performers in other grade level groups. Also, outstanding ratings are valuable for promotion in some agencies such as Army, Air Force, and Treasury.

To examine whether performance appraisals contribute to gender and racial inequalities in career success, the second essay tested performance evaluation bias and performance reward bias separately. The study finds that women receive equal or higher performance ratings than comparable white men. In particular, white women receive consistently higher ratings than white men, after controlling for individual and job-related characteristics. However, some minority male groups, particularly black men, tend to receive lower ratings than comparable white men. On the other hand, the returns on outstanding ratings do not differ between minority male groups and white men. Outstanding ratings increase salaries and chances of promotion more for some female groups.

This study provides some important policy implications. First, the federal government needs to strengthen the linkage between performance ratings and rewards, if

performance-based rewards to be effective. Although many observers pointed out that rewards for top performers are not much greater than average performers, they didn't provide evidence in a quantifiable manner. This study demonstrates that even the highest performance ratings, outstanding ratings, do not have a significant advantage in career advancement over middle level ratings, fully successful ratings, in the federal civil service. Therefore, if agencies want to have effective performance-based reward systems, they should provide top performers with "meaningfully" different rewards, so that employees perceive a stronger linkage between performance ratings and rewards.

Second, although the overall impact of performance appraisal ratings on career advancement seems to be weak, outstanding ratings clearly give higher chances of promotion to the employees in the federal civil service. In other words, despite rating inflation in the federal government, performance appraisals still serve for personnel decisions. Thus, the argument that federal performance appraisal is useless (Light, 1999) is not true. Half of federal agencies moved from traditional five level ratings to pass or fail systems (Henry, 2009), but the pass or fail system is not an effective tool to serve personnel decisions (Liff, 2007; Montoya & Graham, 2007). Instead of giving up an important purpose of performance appraisals, judgmental purposes, federal agencies need to improve their performance appraisal systems to become a helpful tool to make personnel decisions.

Lastly, this study provides useful insights for reducing gender inequalities in career success. Disadvantages of women in career success are not associated with performance appraisals, which is a key managerial tool for human resource management. Rather, performance ratings for women are as high as or higher than those for white men.

In addition, once women receive the same outstanding ratings as comparable white men, they are equally likely to receive salary increases and promotion. Therefore, if agencies strongly use performance appraisal results for their personnel decisions, differences in career success between men and women would be reduced.

By providing empirical evidence on the impact of performance ratings on personnel decisions, this study expands our understanding of performance appraisals and performance-based rewards in the federal civil service. However, this study does not reflect a dramatic change of federal performance appraisals and performance-based rewards system in recent years, because the most recent data used for this study is 2003. Therefore, future study should use more recent data which includes new pay-for-performance systems in Departments of Defense and Homeland Security.

Performance-based rewards are not unique in the federal government but are used in many state and local governments. However, the findings of this study may not be generalizable to the other levels of governments. In order to understand the relationship between performance ratings and rewards in the public sector in general, future research needs to analyze data for state and local governments.

APPENDIX

TABLES

Table A.1: Performance Appraisal Rating Distribution, 1988-2003 (In Percentage)

Rating	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
5	21.9	22.9	24.6	27.5	29.5	31.0	33.7	38.1	41.3	41.9	41.8	49.6	45.4	43.8	42.6	44.0
4	38.0	38.6	39.9	42.6	42.8	43.0	42.7	40.7	38.0	36.7	31.8	32.8	29.2	27.1	25.4	25.8
3	39.3	37.9	35.0	29.4	27.4	25.7	23.2	20.8	20.2	21.0	26.1	17.3	25.2	28.7	31.6	29.8
2	0.7	0.4	0.4	0.5	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.1	0.2	0.3	0.3
1	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.0	0.1	0.2	0.1	0.1
Mean	3.81	3.84	3.89	3.97	4.01	4.05	4.10	4.17	4.20	4.20	4.15	4.31	4.20	4.14	4.10	4.13
N	11,964	12,284	12,593	12,527	12,954	13,110	12,982	12,501	12,151	11,210	10,105	7,486	6,711	6,330	5,669	5,146

Due to the fact that a substantial number of federal employees have gone to rating systems other than five rating levels but this study restricted sample to those under five-level ratings, the sample size is decreased since 1996.

Table A.2: Career Advancement by Year

	Mean Performance Rating	% Outstanding	Mean % Change in Salary	% Promoted	Sample Size
1988	3.8	21.4	8.4	21.4	11,194
1989	3.8	23.0	7.5	19.2	11,537
1990	3.9	24.6	8.7	17.9	11,877
1991	4.0	27.5	7.8	17.1	12,024
1992	4.0	29.5	6.8	14.0	12,472
1993	4.0	30.9	6.1	13.5	12,496
1994	4.1	33.9	5.6	12.6	12,247
1995	4.2	38.3	5.2	12.7	11,902
1996	4.2	41.3	5.7	12.1	11,277
1997	4.2	41.8	5.4	12.2	10,707
1998	4.2	41.9	6.2	12.9	9,596
1999	4.3	49.6	7.4	13.1	7,093
2000	4.2	45.7	6.7	13.8	6,396
2001	4.1	44.0	7.5	14.1	6,004
2002	4.1	43.0	6.9	14.6	5,238
2003	4.1	44.0	N/A	N/A	5,146

Table A.3: T-test by Year: Career Advancement Comparison between Fully Successful and Outstanding

	Mean % salary change		% promoted	
	Fully successful	Outstanding	Fully successful	Outstanding
1988	8.5**	7.8	21.7**	18.5
1989	7.9**	7.5	19.4	20.2
1990	8.9*	8.6	17.1	18.4
1991	7.8	7.8	16.4	17.7
1992	7.0	6.8	13.9	14.0
1993	6.2	6.0	15.9**	12.3
1994	5.7	5.6	12.2	13.8
1995	5.1	5.3	11.5	14.1**
1996	5.6	5.8	11.5	13.3*
1997	5.5	5.4	12.3	12.4
1998	6.2	6.2	12.0	13.3
1999	7.5	7.4	13.4	13.2
2000	6.6	6.7	13.9	13.7
2001	7.3	7.6*	13.5	14.2
2002	6.9	6.9	15.0	14.6

* significant at .05; ** significant at .01

Table A.4: Impact of Performance Ratings on Percent Change in Salary Growth by Year : OLS Model

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Outstanding	0.227	0.396**	0.245	0.609**	0.402**	0.185	0.467**	0.578**	0.538**	0.254*	0.406**	0.445**	0.219	0.170	-0.107
	(1.69)	(3.07)	(1.77)	(5.22)	(3.80)	(1.67)	(4.41)	(5.50)	(5.00)	(2.33)	(3.58)	(2.94)	(1.25)	(0.94)	(0.54)
Exceeds fully	-0.154	0.003	0.086	0.254*	0.071	0.076	0.145	0.297**	0.105	0.027	0.031	0.075	0.111	-0.181	-0.338
	(1.37)	(0.03)	(0.73)	(2.50)	(0.76)	(0.77)	(1.49)	(2.99)	(1.03)	(0.25)	(0.27)	(0.50)	(0.63)	(0.98)	(1.67)
Less than fully	-1.168	-1.092	-0.212	-0.221	-0.861	-1.140	-0.410	-0.661	-0.853	-0.434	-0.004	0.704	0.948	-0.055	1.929
	(1.73)	(1.69)	(0.25)	(0.39)	(1.27)	(1.73)	(0.66)	(1.09)	(1.52)	(0.73)	(0.01)	(0.71)	(0.81)	(0.06)	(1.93)
Black males	-0.205	-0.008	0.538*	-0.235	0.131	-0.004	0.243	0.180	0.145	0.286	0.423*	0.292	-0.486	0.322	-0.165
	(0.92)	(0.04)	(2.38)	(1.24)	(0.78)	(0.02)	(1.49)	(1.12)	(0.90)	(1.75)	(2.48)	(1.37)	(1.93)	(1.31)	(0.65)
Hispanic males	0.258	0.709*	0.600	0.479	0.328	-0.001	0.298	-0.175	-0.492*	0.010	-0.147	0.008	0.375	0.300	0.522
	(0.75)	(2.19)	(1.71)	(1.66)	(1.30)	(0.00)	(1.26)	(0.76)	(2.10)	(0.04)	(0.58)	(0.03)	(1.08)	(0.87)	(1.52)
Asian males	-0.524	-0.809*	0.445	0.567	-0.437	-1.430**	-0.191	0.187	-0.142	-0.639*	0.292	-0.083	0.187	0.138	-1.223**
	(1.24)	(2.04)	(1.09)	(1.69)	(1.46)	(4.75)	(0.68)	(0.68)	(0.50)	(2.16)	(0.93)	(0.23)	(0.46)	(0.36)	(3.07)
White females	0.476**	0.569**	0.432**	0.358**	0.327**	0.587**	0.404**	0.321**	0.210*	0.249**	0.359**	0.319**	0.288	0.590**	0.528**
	(3.64)	(4.59)	(3.25)	(3.25)	(3.33)	(5.81)	(4.24)	(3.45)	(2.25)	(2.60)	(3.57)	(2.58)	(1.95)	(4.12)	(3.54)
Black females	-0.011	0.344*	0.518**	0.175	0.063	-0.020	-0.199	-0.053	0.117	0.209	0.504**	0.412*	0.375	0.646**	0.418*
	(0.06)	(2.00)	(2.80)	(1.14)	(0.46)	(0.14)	(1.52)	(0.42)	(0.91)	(1.60)	(3.68)	(2.44)	(1.89)	(3.36)	(2.12)
Hispanic females	0.039	0.478	0.801*	0.540	0.221	0.747**	0.158	0.335	0.415	0.173	0.849**	0.075	0.121	0.925*	1.267**
	(0.11)	(1.35)	(2.15)	(1.73)	(0.80)	(2.70)	(0.61)	(1.31)	(1.75)	(0.71)	(3.32)	(0.24)	(0.33)	(2.54)	(3.49)
Asian females	0.318	0.022	1.195**	0.325	0.015	-1.234**	-0.240	-0.038	-0.326	0.269	0.500	0.012	0.720	0.419	0.460
	(0.70)	(0.05)	(2.71)	(0.87)	(0.05)	(3.74)	(0.79)	(0.13)	(1.16)	(0.94)	(1.71)	(0.03)	(1.71)	(1.04)	(1.15)
Education (year)	0.016	0.008	0.028	0.001	0.068**	0.034	-0.001	-0.001	-0.017	0.006	0.012	0.011	-0.007	-0.042	-0.003
	(0.52)	(0.28)	(0.88)	(0.05)	(2.91)	(1.40)	(0.04)	(0.04)	(0.76)	(0.27)	(0.49)	(0.36)	(0.19)	(1.23)	(0.07)
Years of service	-0.438**	-0.463**	-0.427**	-0.386**	-0.335**	-0.272**	-0.302**	-0.274**	-0.285**	-0.313**	-0.287**	-0.370**	-0.316**	-0.327**	-0.424**
	(19.67)	(21.61)	(18.50)	(19.87)	(19.29)	(14.64)	(16.50)	(14.70)	(14.75)	(15.66)	(13.59)	(14.59)	(10.74)	(11.81)	(15.26)
Years of service squared	0.009**	0.009**	0.009**	0.008**	0.007**	0.005**	0.006**	0.005**	0.006**	0.006**	0.005**	0.007**	0.006**	0.006**	0.008**
	(14.72)	(15.98)	(13.69)	(14.58)	(14.14)	(10.51)	(12.21)	(10.55)	(11.14)	(11.91)	(9.80)	(11.11)	(7.82)	(8.32)	(11.76)
Age (year)	-0.246**	-0.205**	-0.338**	-0.262**	-0.238**	-0.144**	-0.284**	-0.201**	-0.264**	-0.264**	-0.198**	-0.312**	-0.216**	-0.235**	-0.362**
	(6.32)	(5.41)	(8.26)	(7.66)	(7.74)	(4.46)	(9.01)	(6.38)	(7.91)	(7.51)	(5.22)	(6.68)	(3.92)	(4.32)	(6.48)
Age squared	0.002**	0.001**	0.003**	0.002**	0.002**	0.001*	0.002**	0.002**	0.002**	0.002**	0.001**	0.003**	0.001*	0.002**	0.003**
	(4.27)	(3.41)	(6.71)	(5.57)	(5.63)	(2.48)	(6.87)	(4.47)	(5.96)	(5.54)	(3.56)	(5.35)	(2.54)	(3.13)	(5.16)
Supervisor	0.448**	0.515**	0.357*	0.353**	0.456**	0.331**	-0.019	0.572**	0.281*	0.277*	0.270*	0.448**	-0.001	0.437**	0.123
	(3.13)	(3.72)	(2.37)	(2.81)	(4.02)	(2.81)	(0.17)	(5.16)	(2.49)	(2.37)	(2.16)	(3.05)	(0.01)	(2.61)	(0.70)

Table A.4 Continued

GS 1-3	1.993**	1.325**	1.296*	0.379	0.966*	1.385**	0.089	0.354	0.666	0.064	0.406	-0.053	-1.071	-0.028	-1.707
	(4.06)	(2.72)	(2.31)	(0.77)	(2.19)	(2.91)	(0.20)	(0.78)	(1.45)	(0.13)	(0.74)	(0.07)	(1.17)	(0.03)	(1.63)
GS 4	1.195**	0.348	0.172	0.211	0.082	0.495	-0.027	0.429	0.369	0.476	-0.558	-0.062	-0.818	0.519	-0.017
	(3.12)	(0.97)	(0.43)	(0.64)	(0.28)	(1.57)	(0.09)	(1.53)	(1.32)	(1.63)	(1.83)	(0.16)	(1.74)	(1.12)	(0.04)
GS 5	1.415**	0.641	0.440	0.240	0.321	0.189	0.091	0.229	0.056	0.121	0.150	0.150	-0.396	0.448	0.154
	(3.97)	(1.95)	(1.20)	(0.80)	(1.20)	(0.66)	(0.36)	(0.95)	(0.24)	(0.50)	(0.60)	(0.48)	(1.04)	(1.21)	(0.42)
GS 6	1.011**	0.458	0.551	0.005	0.224	-0.076	0.043	0.478*	0.558*	0.398	0.162	0.742*	-0.302	1.075**	0.532
	(2.75)	(1.36)	(1.48)	(0.02)	(0.83)	(0.27)	(0.17)	(1.98)	(2.36)	(1.64)	(0.66)	(2.46)	(0.81)	(2.97)	(1.48)
GS 7	2.149**	1.553**	1.248**	0.770**	0.606*	0.919**	0.721**	0.576*	0.736**	0.649**	0.711**	0.042	0.164	0.189	0.588
	(6.13)	(4.84)	(3.48)	(2.64)	(2.33)	(3.32)	(2.99)	(2.49)	(3.25)	(2.84)	(3.05)	(0.15)	(0.49)	(0.58)	(1.84)
GS 9	0.827*	0.400	0.025	0.129	-0.244	0.510	0.235	0.065	0.043	0.584*	0.261	0.292	0.447	0.899**	1.091**
	(2.32)	(1.22)	(0.07)	(0.43)	(0.91)	(1.78)	(0.92)	(0.27)	(0.18)	(2.41)	(1.06)	(0.99)	(1.25)	(2.60)	(3.19)
GS 10	-1.560**	-1.608**	-1.274**	-1.802**	-2.110**	0.246	-0.442	0.124	-0.258	-0.401	-0.350	-0.302	-0.743	-0.061	0.544
	(3.36)	(3.62)	(2.67)	(4.49)	(6.00)	(0.67)	(1.16)	(0.33)	(0.68)	(1.02)	(0.89)	(0.58)	(1.25)	(0.10)	(0.87)
GS 11	-1.185**	-1.237**	-1.405**	-1.108**	-1.141**	-0.769**	-0.890**	-0.673**	-0.563*	-0.632*	-0.304	0.002	-0.336	0.471	0.121
	(3.26)	(3.72)	(3.79)	(3.66)	(4.21)	(2.62)	(3.44)	(2.69)	(2.28)	(2.52)	(1.19)	(0.01)	(0.91)	(1.32)	(0.34)
GS 12	-1.902**	-2.225**	-2.048**	-2.213**	-2.051**	-2.013**	-1.635**	-1.455**	-1.423**	-1.269**	-0.890**	-1.130**	-1.367**	-0.058	-0.757*
	(5.18)	(6.58)	(5.45)	(7.21)	(7.47)	(6.79)	(6.21)	(5.70)	(5.66)	(4.97)	(3.39)	(3.64)	(3.65)	(0.16)	(2.09)
GS 13	-1.992**	-2.512**	-2.517**	-2.231**	-2.230**	-1.466**	-2.109**	-1.805**	-1.553**	-1.713**	-1.143**	-1.393**	-1.782**	-0.773*	-0.957*
	(5.10)	(6.98)	(6.34)	(6.86)	(7.68)	(4.69)	(7.54)	(6.67)	(5.82)	(6.33)	(4.11)	(4.26)	(4.55)	(2.03)	(2.53)
GS 14	-2.441**	-2.865**	-2.899**	-2.453**	-2.457**	-1.437**	-2.218**	-1.978**	-1.516**	-1.602**	-1.140**	-1.570**	-1.445**	-0.784	-0.585
	(5.75)	(7.30)	(6.73)	(6.96)	(7.80)	(4.27)	(7.27)	(6.69)	(5.08)	(5.32)	(3.66)	(4.26)	(3.27)	(1.80)	(1.34)
GS 15	-2.599**	-3.237**	-2.986**	-2.955**	-2.631**	-1.567**	-2.537**	-2.607**	-1.846**	-2.055**	-1.266**	-1.609**	-1.439**	-1.332**	-0.801
	(5.21)	(6.91)	(5.83)	(7.08)	(7.06)	(4.01)	(6.98)	(7.47)	(5.23)	(5.84)	(3.43)	(3.69)	(2.80)	(2.65)	(1.54)
GS 16	-3.579**	-4.397**	16.368**	-4.020**	-4.483**	-1.624*	-4.038**	-3.729**	-2.874**	-2.365**	-2.817**	-2.876**	-3.029**	-2.383**	-2.738**
	(3.40)	(4.96)	(16.62)	(5.32)	(6.67)	(2.35)	(5.99)	(5.84)	(4.14)	(3.27)	(3.88)	(3.53)	(3.27)	(2.83)	(2.81)
Professional	1.766**	2.668**	3.404**	0.690*	1.736**	1.596**	1.800**	1.653**	1.728**	1.849**	0.617*	1.646**	2.097**	2.005**	0.886*
	(4.47)	(7.10)	(8.27)	(2.03)	(5.86)	(5.24)	(6.31)	(5.79)	(6.13)	(6.52)	(2.14)	(5.14)	(5.67)	(5.57)	(2.47)
Administrative	2.339**	2.912**	3.359**	1.187**	2.361**	2.862**	2.192**	1.769**	1.982**	2.413**	1.517**	2.148**	3.218**	1.969**	1.371**
	(6.28)	(8.22)	(8.52)	(3.66)	(8.36)	(9.83)	(8.03)	(6.46)	(7.36)	(8.95)	(5.57)	(7.18)	(9.40)	(5.89)	(4.19)
Technical	-0.421	0.132	0.697	-1.289**	0.071	0.367	0.263	-0.218	0.178	0.474	-0.022	0.284	0.829**	0.502	-0.487
	(1.21)	(0.40)	(1.90)	(4.28)	(0.27)	(1.38)	(1.06)	(0.88)	(0.73)	(1.94)	(0.09)	(1.06)	(2.67)	(1.66)	(1.63)
Clerical	-1.345**	-1.046**	-0.678	-2.427**	-0.824**	-0.390	-0.397	-0.495	-0.329	0.215	-0.451	-0.013	0.329	0.036	-0.637
	(3.78)	(3.09)	(1.80)	(7.83)	(3.01)	(1.37)	(1.47)	(1.85)	(1.23)	(0.80)	(1.66)	(0.04)	(0.90)	(0.10)	(1.73)

Table A.4 Continued

Air Force	-1.039**	-0.974**	-1.129**	-1.148**	-1.189**	-0.670**	-0.596*	-0.806**	-0.906**	-0.319	-0.636**	-0.783**	-0.729*	-0.832**	-0.261
	(3.39)	(3.16)	(3.34)	(4.05)	(4.77)	(2.61)	(2.51)	(3.40)	(5.09)	(1.74)	(3.24)	(2.66)	(2.21)	(2.62)	(0.72)
Agriculture	-1.152**	-0.717*	-0.853*	-0.523	-0.325	-0.179	-0.362	-0.180	-0.144	0.430*	0.136	-0.273	-1.072**	-0.777*	-0.954**
	(3.51)	(2.19)	(2.37)	(1.76)	(1.26)	(0.68)	(1.48)	(0.74)	(0.78)	(2.24)	(0.67)	(0.89)	(2.97)	(2.17)	(2.59)
Army	-1.536**	-0.673*	0.122	-0.910**	-1.180**	-0.724**	-0.747**	-0.740**	-0.707**	-0.469**	-0.368*	-0.704*	-0.700*	-0.765*	-0.310
	(5.23)	(2.26)	(0.37)	(3.35)	(4.99)	(2.99)	(3.35)	(3.30)	(4.37)	(2.77)	(1.99)	(2.50)	(2.22)	(2.53)	(0.98)
Commerce	0.054	-0.512	-0.496	0.245	0.451	1.274**	1.020**	1.078**	0.715*	0.657*	0.901*	0.701	0.230	4.429**	-1.444**
	(0.12)	(1.17)	(1.02)	(0.62)	(1.29)	(3.58)	(3.11)	(3.36)	(2.57)	(2.33)	(2.51)	(1.66)	(0.44)	(8.76)	(2.73)
Defense	-0.618	-1.037**	0.055	-0.242	-0.552*	0.092	-0.144	-0.290	-0.338	-0.182	-0.207	-0.445	-0.922*	-0.909*	-0.447
	(1.69)	(2.98)	(0.14)	(0.80)	(2.10)	(0.35)	(0.59)	(1.20)	(1.82)	(0.95)	(1.03)	(1.34)	(2.45)	(2.50)	(1.18)
Justice	0.217	0.683	0.876*	2.385**	0.745**	-0.419	0.146	-0.170	-0.236	0.131	-0.007	0.313	-0.885**	-0.826*	0.032
	(0.53)	(1.73)	(2.08)	(7.06)	(2.58)	(1.45)	(0.56)	(0.65)	(1.14)	(0.63)	(0.03)	(1.03)	(2.62)	(2.42)	(0.09)
Labor	-0.367	-0.309	0.412	-0.190	-0.596	-0.198	-0.164	0.532	0.392	0.438	0.040	0.841	0.131	0.041	-0.000
	(0.73)	(0.62)	(0.77)	(0.42)	(1.45)	(0.47)	(0.42)	(1.36)	(1.16)	(1.17)	(0.10)	(1.50)	(0.17)	(0.06)	(0.00)
Energy	-0.355	-0.712	-0.134	-0.383	0.254	-0.196	0.772*	0.889*	0.037	0.405	0.687	1.282*	-2.296**	1.775*	-0.514
	(0.70)	(1.43)	(0.24)	(0.85)	(0.65)	(0.51)	(2.17)	(2.48)	(0.11)	(1.23)	(1.93)	(2.24)	(3.48)	(2.30)	(0.40)
Education	-0.933	0.888	-0.734	-0.266	0.534	0.517	0.418	0.879	2.115**	1.158*	0.124	N/A	-1.803	-0.034	N/A
	(1.20)	(1.16)	(0.89)	(0.40)	(0.94)	(0.89)	(0.73)	(1.60)	(4.00)	(2.20)	(0.11)		(0.83)	(0.01)	
EPA	0.239	0.270	1.011	1.036	0.712	0.389	-0.027	-1.254**	0.935**	1.671**	0.742*	0.678	-1.839	0.978	-0.695
	(0.38)	(0.46)	(1.56)	(1.86)	(1.46)	(0.78)	(0.06)	(3.30)	(2.75)	(4.92)	(2.18)	(1.52)	(1.42)	(0.83)	(0.29)
GSA	-0.760	-0.223	0.884	-0.359	0.117	-0.018	0.438	0.745	0.939**	1.130**	0.279	-0.450	N/A	4.313	0.458
	(1.42)	(0.42)	(1.53)	(0.75)	(0.29)	(0.04)	(1.08)	(1.89)	(2.60)	(2.99)	(0.57)	(0.56)		(1.05)	(0.36)
HHS	-1.384**	-1.429**	-0.246	-0.461	-0.383	1.647**	-0.195	-0.018	0.016	-0.010	0.548	-0.332	-0.732	-0.425	12.080**
	(4.26)	(4.37)	(0.69)	(1.56)	(1.50)	(6.34)	(0.83)	(0.08)	(0.06)	(0.04)	(1.95)	(0.88)	(1.78)	(1.11)	(3.08)
HUD	-0.652	0.320	0.375	0.019	-0.465	0.386	0.287	1.361**	0.579	0.072	0.759*	0.407	1.015	-0.831	-0.115
	(0.77)	(0.65)	(0.69)	(0.04)	(1.15)	(0.93)	(0.72)	(3.45)	(1.64)	(0.19)	(1.97)	(0.88)	(1.92)	(1.62)	(0.21)
Interior	-0.977**	-1.273**	-1.383**	-0.339	-0.408	0.164	0.199	0.039	0.114	0.293	-0.017	0.352	0.666	-0.842	-0.403
	(2.59)	(3.36)	(3.37)	(0.99)	(1.36)	(0.54)	(0.69)	(0.14)	(0.49)	(1.20)	(0.06)	(0.77)	(0.90)	(1.00)	(0.48)
NASA	0.253	-0.618	-0.859	-0.647	-0.764*	0.008	0.139	-0.150	0.105	-0.031	0.469	-0.957	-0.184	-1.493	-2.037
	(0.54)	(1.34)	(1.76)	(1.58)	(2.09)	(0.02)	(0.39)	(0.42)	(0.33)	(0.09)	(0.77)	(0.97)	(0.16)	(0.88)	(0.72)
Navy	-0.656*	-0.668*	0.303	-0.799**	-0.822**	-0.980**	-0.881**	-1.004**	-0.769**	-0.370*	-0.376*	-0.615*	-0.490	-0.839*	-0.604
	(2.12)	(2.17)	(0.91)	(2.92)	(3.44)	(4.01)	(3.91)	(4.44)	(4.66)	(2.15)	(2.00)	(2.15)	(1.48)	(2.52)	(1.73)
State	0.593	0.855	-0.702	0.645	-0.789	0.939	-0.131	-0.189	-0.352	0.372	0.699	1.009	N/A	-1.240	-1.297
	(0.72)	(1.08)	(0.82)	(0.92)	(1.31)	(1.48)	(0.22)	(0.32)	(0.63)	(0.64)	(1.21)	(1.19)		(0.52)	(0.32)
Transportation	-0.122	0.854*	0.992**	1.016**	0.499	0.341	0.562*	0.667**	-0.111	0.632	0.179	0.232	-1.121	-1.462*	-0.514
	(0.35)	(2.43)	(2.59)	(3.24)	(1.83)	(1.23)	(2.18)	(2.61)	(0.30)	(1.71)	(0.46)	(0.38)	(1.55)	(2.16)	(0.83)
Treasury	0.258	-0.383	0.755*	0.001	-0.647**	-0.819**	-0.283	-0.151	-0.170	-0.153	0.751**	-0.064	-0.162	0.760*	0.281
	(0.82)	(1.22)	(2.20)	(0.00)	(2.65)	(3.32)	(1.25)	(0.66)	(1.03)	(0.90)	(4.03)	(0.22)	(0.51)	(2.47)	(0.88)

Table A.4 Continued

Veterans Affairs	-1.659**	-1.510**	-0.902**	-1.480**	-1.213**	-1.342**	-1.038**	-1.084**	-1.115**	-0.570**	-0.453*	-0.733*	-0.879*	-1.493**	-0.982*
	(5.24)	(4.72)	(2.58)	(5.12)	(4.83)	(5.28)	(4.43)	(4.63)	(6.44)	(3.19)	(2.44)	(2.07)	(1.97)	(3.14)	(1.97)
Constant	18.578**	16.978**	19.137**	19.381**	15.528**	11.796**	15.593**	13.150**	15.348**	14.702**	14.316**	18.707**	16.099**	17.333**	21.336**
	(17.84)	(16.71)	(17.33)	(21.19)	(19.20)	(13.88)	(19.35)	(16.42)	(18.77)	(17.14)	(15.49)	(16.30)	(11.82)	(12.86)	(15.39)
Observations	10,197	10,719	11,038	11,194	11,519	11,559	11,321	11,071	10,282	9,958	8,906	6,563	5,959	5,591	4,732
R-squared	0.24	0.24	0.21	0.23	0.20	0.18	0.19	0.16	0.16	0.17	0.15	0.18	0.15	0.18	0.23

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

Table A.5: Impact of Performance Ratings on Promotion Probability by Year: Logit Model

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Outstanding	0.203**	0.468**	0.516**	0.493**	0.441**	0.076	0.566**	0.638**	0.627**	0.474**
	(2.63)	(6.20)	(6.79)	(6.38)	(5.27)	(0.88)	(6.36)	(7.06)	(6.47)	(4.93)
Exceeds fully	0.052	0.238**	0.280**	0.237**	0.179*	0.006	0.184*	0.273**	0.209*	0.220*
	(0.80)	(3.66)	(4.28)	(3.48)	(2.42)	(0.07)	(2.22)	(3.14)	(2.21)	(2.36)
Less than fully	-0.622	-1.172	-0.062	-0.546	-0.420	-1.591*	-0.943	-0.911	-0.173	-0.833
	(1.50)	(1.91)	(0.13)	(1.35)	(0.73)	(2.00)	(1.20)	(1.21)	(0.30)	(1.12)
Observations	10261	10783	11085	11260	11608	11632	11413	11175	10352	10024

	1998	1999	2000	2001	2002
Outstanding	0.401**	0.455**	0.172	0.182	0.008
	(4.14)	(3.61)	(1.42)	(1.38)	(0.06)
Exceeds fully	0.138	0.048	-0.050	-0.073	-0.435**
	(1.43)	(0.38)	(0.41)	(0.55)	(2.85)
Less than fully	-0.107	0.276	0.400	0.036	0.237
	(0.19)	(0.40)	(0.58)	(0.05)	(0.36)
Observations	8977	6612	5994	5619	4627

Absolute value of z statistics in parentheses
Model includes individual and organizational characteristics
* significant at 5%; ** significant at 1%

Table A.6: Career Advancement by PATCO Category

	Mean Performance Rating	% Outstanding	Mean % Change in Salary	% Promoted	Sample Size
Professional	4.0	31.7	6.7	12.7	33,609
Administrative	4.1	36.4	7.1	14.3	49,175
Technical	3.9	29.8	6.3	14.9	34,342
Clerical	4.0	34.8	6.5	17.8	21,595
Other	3.9	24.3	6.9	16.8	4,127

Table A.7: T-test by PACTO Category: Career Advancement Comparison between Fully Successful and Outstanding

	Mean % salary change		% promoted	
	Fully successful	Outstanding	Fully successful	Outstanding
Professional	7.1**	6.5	15.3**	13.7
Administrative	7.6**	7.0	17.4**	15.2
Technical	6.5**	6.2	15.4	15.7
Clerical	6.8**	6.2	18.3	17.8
Other	7.2**	6.6	18.6	17.1

* significant at .05; ** significant at .01

Table A.8: Impact of Performance Ratings on Percent Change in Salary Growth by PATCO Category: Panel Model

	Prof.	Admin.	Tech.	Cler.	Other
Outstanding	0.048	0.135	0.211**	0.143	-0.081
	(0.64)	(1.93)	(2.89)	(1.53)	(0.38)
Exceeds fully	-0.051	-0.019	0.069	0.046	0.035
	(0.80)	(0.749)	(1.16)	(0.58)	(0.20)
Less than fully	0.540	-0.484	0.016	-0.178	-2.307
	(1.36)	(1.24)	(0.04)	(0.54)	(1.36)
Observations	33295	48631	33463	21141	4071
R-squared	0.31	0.27	0.16	0.15	0.22

Absolute value of t statistics in parentheses
Model includes individual and organizational characteristics
* significant at 5%; ** significant at 1%

Table A.9: Impact of Performance Ratings on Promotion Probability by PATCO Category: Panel Model

	Prof.	Admin.	Tech.	Cler.	Other
Outstanding	0.200*	0.375**	0.316**	0.396**	0.450*
	(2.08)	(5.96)	(4.27)	(4.63)	(2.25)
Exceeds fully	0.009	0.073	0.062	0.144*	0.076
	(0.52)	(1.34)	(1.02)	(1.98)	(0.48)
Less than fully	0.231	-0.102	-1.096	0.032	-6.319
	(1.16)	(0.26)	(1.95)	(0.09)	(0.32)
Observations	20048	31893	20333	13795	3100

Absolute value of z statistics in parentheses
Model includes individual and organizational characteristics
* significant at 5%; ** significant at 1%

Table A.10: Career Advancement by Grade Level

	Mean Performance Rating	% Outstanding	Mean % Change in Salary	% Promoted	Sample Size
Grade 3[a)]	3.6	17.9	7.8	30.7	2,644
Grade 4	3.8	26.0	6.8	20.6	13,842
Grade 5	3.9	31.1	6.6	18.0	25,598
Grade 6	4.0	34.9	6.5	19.0	17,189
Grade 7	4.0	32.5	7.7	22.0	22,360
Grade 8	4.1	34.3	6.1	14.5	7,070
Grade 9	4.0	29.5	8.1	22.4	24,024
Grade 10	3.9	25.8	6.3	16.6	3,414
Grade 11	4.0	30.1	7.2	15.3	31,510
Grade 12	4.1	33.9	6.3	9.0	36,802
Grade 13	4.1	36.9	6.0	6.8	25,150
Grade 14	4.2	41.7	5.8	5.7	12,398
Grade 15	4.3	52.3	5.4	1.4	5,415
Grade 16	4.4	55.0	5.5	N/A	740

a) Grade 1, 2 and 3 are combined

Table A.11: T-test by Grade Level: Career Advancement Comparison between Fully Successful and Outstanding

	Mean % salary change		% promoted	
	Fully successful	Outstanding	Fully successful	Outstanding
Grade 3[a)]	8.3*	7.3	31.9	33.3
Grade 4	6.9**	6.5	19.7	21.8*
Grade 5	6.8**	6.4	17.1	19.0**
Grade 6	6.7**	6.4	19.9**	17.9
Grade 7	8.7**	7.0	27.0**	20.4
Grade 8	6.2	6.1	13.6	15.5
Grade 9	8.6**	7.6	25.0**	22.3
Grade 10	6.5	6.2	17.4	15.7
Grade 11	7.2	7.2	15.3	17.8**
Grade 12	6.2	6.5**	8.8	12.3**
Grade 13	5.9	6.2**	6.1	8.9**
Grade 14	5.6	6.1**	3.7	8.2**
Grade 15	5.2	5.5*	0.1	2.3**
Grade 16	5.0	5.6	N/A	N/A

a) Grade 1, 2 and 3 are combined
* significant at .05; ** significant at .01

Table A.12: Impact of Performance Ratings on Percent Change in Salary Growth by Grade: Panel Model

	GS 1-3	GS 4	GS 5	GS 6	GS 7	GS 8	GS 9	GS 10	GS 11	GS 12
Outstanding	1.367**	0.252	0.335**	0.153	0.171	0.199	0.311*	0.036	0.269**	0.186*
	(2.71)	(1.50)	(2.99)	(1.10)	(1.24)	(0.84)	(2.29)	(0.11)	(2.59)	(2.31)
Exceeds fully	0.292	0.215	0.379**	0.188	0.159	-0.001	0.260*	-0.191	0.179*	0.073
	(0.84)	(1.56)	(4.08)	(1.61)	(1.40)	(0.01)	(2.34)	(0.81)	(2.20)	(1.09)
Less than fully	0.239	-0.633	-0.325	0.345	-0.089	3.065	0.516	-1.926	0.061	-0.409
	(0.27)	(1.14)	(0.66)	(0.58)	(0.15)	(1.62)	(0.81)	(1.10)	(0.13)	(0.99)
Observations	1383	7575	14248	10014	13470	4315	14565	2330	19825	23997
R-squared	0.09	0.06	0.08	0.08	0.08	0.07	0.05	0.10	0.06	0.10

	GS 13	GS 14	GS 15	GS 16
Outstanding	.020	0.127	0.251	0.037
	(0.23)	(1.10)	(1.28)	(0.07)
Exceeds fully	-0.013	0.152	0.222	-0.170
	(0.19)	(1.47)	(1.20)	(0.35)
Less than fully	-0.133	0.367	3.098	N/A
	(0.23)	(0.57)	(1.74)	
Observations	16410	8327	3645	505
R-squared	0.11	0.12	0.14	0.92

Absolute value of t statistics in parentheses
Model includes individual and organizational characteristics
* significant at 5%; ** significant at 1%

Table A.13: Impact of Performance Ratings on Promotion Probability by Grade: Panel Model

	GS 1-3	GS 4	GS 5	GS 6	GS 7	GS 8	GS 9	GS 10	GS 11	GS 12
Outstanding	0.637	0.429*	0.472**	-0.082	0.165	0.559	0.400*	0.243	0.398*	0.395*
	(1.11)	(2.13)	(2.82)	(0.34)	(0.83)	(1.37)	(2.22)	(0.27)	(2.34)	(2.43)
Exceeds fully	-0.263	0.193	0.377**	0.108	0.059	0.229	0.271	-0.786	0.189	0.190
	(0.64)	(1.19)	(2.73)	(0.54)	(0.37)	(0.70)	(1.79)	(1.24)	(1.33)	(1.28)
Less than fully	2.400	-0.327	0.511	0.155	-0.130	3.428	-0.211	-1.410	0.463	-0.625
	(1.62)	(0.33)	(0.52)	(0.15)	(0.11)	(0.83)	(0.24)	(0.73)	(0.50)	(0.53)
Observations	600	3583	5909	3823	5408	1551	5883	1252	7387	7340

	GS 13	GS 14
Outstanding	0.217	0.602
	(0.93)	(1.70)
Exceeds fully	0.086	0.514
	(0.40)	(1.47)
Less than fully	-0.078	-3.314
	(0.05)	(0.20)
Observations	3971	1955

Absolute value of z statistics in parentheses
Model includes individual and organizational characteristics
* significant at 5%; ** significant at 1%

Table A.14: Career Advancement by Agency

	Mean Performance Rating	% Outstanding	Mean % Change in Salary	% Promoted	Sample Size
Army	4.5	60.6	6.5	13.5	35,041
State	4.4	54.0	7.2	18.7	857
Justice	4.2	38.3	7.8	20.5	11,952
HUD	4.2	40.9	7.3	17.1	2,419
Air Force	4.1	41.3	6.5	13.8	21,197
Defense	4.1	34.6	6.7	14.4	13,050
EPA	4.1	33.8	7.5	19.9	2,233
GSA	4.1	31.7	7.5	20.7	1,951
NASA	4.1	30.7	6.6	17.3	2,912
Navy	4.1	35.3	6.4	14.3	28,157
Other Agency	4.1	34.7	7.0	17.1	7,554
Commerce	4.0	37.7	7.7	22.3	4,119
Energy	4.0	31.5	6.7	15.9	2,750
Education	4.0	37.8	7.0	18.8	891
HHS	3.9	26.7	6.8	18.7	12,646
Labor	3.8	18.1	7.1	16.6	2,500
Treasury	3.8	19.2	7.2	18.5	22,669
Agriculture	3.7	10.6	6.7	15.2	15,434
Interior	3.7	15.2	6.5	15.3	7,145
Transportation	3.7	18.1	7.5	19.8	7,422
Vet. Affairs	3.7	21.1	6.0	11.9	21,875
SSA	3.4	11.2	5.6	14.4	3,363

Table A.15: T-test by Agency: Career Advancement Comparison between Fully Successful and Outstanding

	Mean % salary change		% promoted	
	Fully successful	Outstanding	Fully successful	Outstanding
Air Force	7.0**	6.2	14.0	13.0
Agriculture	7.0	6.8	15.3	16.5
Army	7.5**	6.3	16.5**	12.1
Commerce	7.6	7.8	19.6	19.3
Defense	7.0	6.7	13.8	14.9
Justice	8.9**	7.5	26.6**	17.4
Labor	7.2	7.4	13.2	18.3
Energy	6.3	6.9*	14.0	17.2
Education	6.3	7.4	11.0	20.7*
EPA	8.2	7.4	24.3*	17.0
GSA	7.6	7.3	20.1	20.7
HHS	6.9	6.6	17.5	17.7
HUD	7.3	7.2	18.2	16.8
Interior	6.4	6.8*	12.9	17.2**
NASA	7.1	6.5	18.5	15.0
Navy	6.7**	6.3	13.3	13.4
State	7.6	7.0	19.0	15.0
SSA	5.8	5.5	10.5	17.1**
Transportation	8.3**	6.7	20.5**	16.1
Treasury	7.8**	7.1	19.0	19.7
Vet. Affairs	6.0	6.2	10.4	13.2**
Other Agency	7.1	7.2	15.8	17.0

* significant at .05; ** significant at .01

Table A.16: Impact of Performance Ratings on Percent Change in Salary Growth by Agency: Panel Model

	Commerce	Other	Labor	GSA	Navy	Energy	Vet. Aff.	Treasury
Outstanding	0.747*	0.571*	0.621	0.596	0.176	0.158	0.155	0.153
	(2.35)	(2.55)	(1.47)	(1.08)	(1.62)	(0.42)	(1.42)	(1.13)
Exceeds fully	0.635*	0.461*	0.047	0.726	-0.017	0.242	0.083	-0.004
	(2.33)	(2.39)	(0.17)	(1.52)	(0.18)	(0.79)	(0.91)	(0.04)
Less than fully	0.519	-0.289	0.701	1.283	-0.235	-4.068*	0.964	0.023
	(0.47)	(0.41)	(0.41)	(0.44)	(0.42)	(2.05)	(1.45)	(0.06)
Observations	2545	4527	1510	1239	18028	1697	12131	14618
R-squared	0.30	0.22	0.33	0.29	0.26	0.22	0.21	0.26

	Army	NASA	Agricult.	Interior	Transport.	HUD	HHS	Education
Outstanding	0.056	0.02	-0.002	-0.05	-0.058	-0.085	-0.119	-0.162
	(0.50)	(0.05)	(0.01)	(0.21)	(0.23)	(0.22)	(0.72)	(0.22)
Exceeds fully	-0.035	-0.146	-0.005	0.002	-0.354	0.023	0.169	-0.31
	(0.32)	(0.41)	(0.04)	(0.01)	(1.92)	(0.07)	(1.26)	(0.50)
Less than fully	0.15	0.037	-0.843	-0.631	-0.828	1.137	-0.719	-3.178
	(0.25)	(0.02)	(1.19)	(0.64)	(0.61)	(0.59)	(0.87)	(1.26)
Observations	22313	1782	9751	3896	4791	1603	7883	511
R-squared	0.19	0.26	0.23	0.20	0.31	0.32	0.22	0.28

Table A.16 Continued

	Air Force	Defense	Justice	EPA	SSA	State
Outstanding	-0.212	-0.247	-0.283	-0.464	-0.648	-0.954
	(1.74)	(1.49)	(1.20)	(0.94)	(1.52)	(1.38)
Exceeds fully	-0.329**	-0.272*	-0.262	-0.100	-0.751*	-0.901
	(3.18)	(1.96)	(1.27)	(0.24)	(2.12)	(1.50)
Less than fully	0.649	0.406	-0.404	3.189	8.966*	-1.966
	(0.54)	(0.60)	(0.32)	(0.79)	(2.54)	(0.77)
Observations	13385	7982	7059	1238	1621	499
R-squared	0.21	0.26	0.28	0.37	0.22	0.43

Absolute value of t statistics in parentheses
Model includes individual and organizational characteristics
* significant at 5%; ** significant at 1%

Table A.17: Impact of Performance Ratings on Promotion Probability by Agency: Panel Model

	Other	Treasury	Air Force	Vet. Aff.	Agriculture	Army	HHS	Labor
Outstanding	0.417*	0.381**	0.380**	0.364**	0.293*	0.277**	0.275*	0.673
	(2.08)	(3.71)	(3.13)	(3.33)	(2.02)	(2.82)	(2.04)	(1.59)
Exceeds fully	0.400*	0.079	0.132	0.214*	0.070	-0.017	0.163	0.609*
	(2.29)	(1.04)	(1.24)	(2.25)	(0.73)	(0.18)	(1.51)	(2.02)
Less than fully	0.188	0.023	-0.374	0.284	-0.730	-1.473	-0.893	N/A
	(0.30)	(0.06)	(0.26)	(0.40)	(0.69)	(1.58)	(0.80)	
Observations	2948	11295	8578	7058	6933	14894	5689	1014

	Commerce	GSA	Transport.	Energy	Justice	NASA	Defense	SSA
Outstanding	0.359	0.354	0.247	0.218	0.165	0.154	0.152	0.150
	(1.40)	(0.92)	(1.39)	(0.69)	(1.04)	(0.43)	(1.03)	(0.21)
Exceeds fully	0.348	0.058	0.083	-0.137	-0.098	0.049	-0.122	-0.499
	(1.55)	(0.17)	(0.64)	(0.50)	(0.73)	(0.16)	(0.94)	(0.94)
Less than fully	1.808*	N/A	-4.131	-2.755	-0.627	N/A	0.420	N/A
	(2.54)		(0.70)	(0.20)	(0.56)		(0.70)	
Observations	1705	938	3131	1131	5515	1169	4666	397

Table A.17 Continued

	Navy	HUD	Interior	Education	State	EPA
Outstanding	0.144	0.142	0.065	0.040	-0.448	-0.477
	(1.27)	(0.46)	(0.29)	(0.06)	(0.60)	(1.14)
Exceeds fully	-0.180	-0.147	0.092	-0.303	-0.058	-0.289
	(1.74)	(0.54)	(0.55)	(0.47)	(0.09)	(0.83)
Less than fully	-1.038	-0.017	-0.643	-4.403	-4.216	N/A
	(0.87)	(0.01)	(0.53)	(0.42)	(0.32)	
Observations	11406	1192	2472	363	374	870

Absolute value of z statistics in parentheses
Model includes individual and organizational characteristics
* significant at 5%; ** significant at 1%

Table A.18: Race and gender Differences in Performance Appraisal Ratings by PATCO Category: Ordered Logit Model

	Professional	Administrative	Technical	Clerical	Other
White females	0.363**	0.339**	0.312**	0.366**	-0.051
	(11.82)	(14.72)	(10.10)	(6.92)	(0.25)
Black females	-0.229**	0.008	0.106**	-0.022	-0.381
	(4.12)	(0.23)	(2.86)	(0.38)	(1.46)
Hispanic females	0.093	-0.003	0.357**	0.137	-0.388
	(0.80)	(0.05)	(5.50)	(1.62)	(1.19)
Asian females	-0.065	0.391**	0.470**	0.345**	1.634
	(0.75)	(4.62)	(4.78)	(3.80)	(1.41)
Black males	-0.434**	-0.278**	-0.122*	-0.177*	-0.114
	(7.81)	(7.47)	(2.51)	(2.22)	(1.23)
Hispanic males	-0.004	-0.024	-0.268**	-0.132	-0.323**
	(0.05)	(0.43)	(3.99)	(0.85)	(2.86)
Asian males	-0.022	0.312**	0.054	-0.080	0.274
	(0.39)	(3.88)	(0.54)	(0.37)	(1.14)
Observations	33,295	48,631	33,463	21,141	4,071

Absolute value of z statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%

Table A.19: Race and gender Differences in Performance Appraisal Ratings by Agency : Ordered Logit Model

	Air Force	Agriculture	Army	Commerce	Defense	Justice	SSA	Labor
White females	0.494**	0.481**	0.673**	0.351**	0.460**	0.624**	0.086	1.088**
	(10.98)	(9.12)	(17.54)	(3.17)	(8.10)	(8.86)	(0.53)	(7.70)
Black females	0.151	0.170	0.330**	1.069**	0.105	0.073	0.061	0.210
	(1.85)	(1.87)	(5.78)	(7.68)	(1.44)	(0.79)	(0.32)	(1.25)
Hispanic females	0.436**	0.493*	0.390**	-0.387	0.270	0.360**	0.060	1.282**
	(4.29)	(2.47)	(3.37)	(1.33)	(1.39)	(2.75)	(0.21)	(3.22)
Asian females	0.649**	0.460	0.664**	-0.204	0.537**	-0.033	-0.302	1.481*
	(4.92)	(1.85)	(4.99)	(0.52)	(2.82)	(0.16)	(0.49)	(2.47)
Black males	-0.323**	-0.404**	-0.119	0.111	-0.204*	-0.121	-0.777**	-0.340
	(3.60)	(3.65)	(1.92)	(0.65)	(2.24)	(1.36)	(2.59)	(1.37)
Hispanic males	0.217*	0.133	-0.227*	-1.137**	0.111	-0.304**	-0.942*	-1.623
	(2.31)	(1.04)	(2.22)	(2.72)	(0.57)	(3.30)	(2.18)	(1.88)
Asian males	0.144	0.266	-0.603**	-0.204	0.055	0.748**	-31.112	3.622**
	(0.98)	(1.33)	(5.21)	(0.94)	(0.31)	(3.27)	(0.00)	(4.25)
Observations	13,385	9,751	22,313	2,545	7,982	7,059	1,621	1,510

Table A.19 Continued

	Energy	Education	EPA	GSA	HHS	HUD	Interior	NASA
White females	0.305*	0.333	0.337*	0.660**	0.311**	0.164	0.866**	0.222
	(2.07)	(1.11)	(2.09)	(4.44)	(5.16)	(1.15)	(8.99)	(1.50)
Black females	0.226	0.073	-0.164	0.301	-0.134	-0.383*	-0.402*	-1.322**
	(1.17)	(0.24)	(0.84)	(1.40)	(1.87)	(2.48)	(2.27)	(4.09)
Hispanic females	1.010**	0.985	0.769	-0.246	-0.189	-1.057*	0.422	-0.753
	(2.96)	(0.61)	(1.87)	(0.43)	(1.50)	(2.14)	(1.76)	(1.87)
Asian females	0.241	N/A	0.720	0.549	-0.093	1.628**	-0.054	-0.369
	(0.32)		(1.80)	(1.31)	(0.48)	(3.83)	(0.16)	(0.67)
Black males	-0.285	-0.493	-2.358**	-0.307	-0.418**	-0.619**	0.143	-1.202**
	(1.18)	(1.36)	(5.61)	(1.23)	(4.28)	(3.03)	(0.52)	(4.53)
Hispanic males	0.130	-1.714	-2.126*	1.223*	-0.328	-0.776*	0.097	-0.543*
	(0.47)	(1.86)	(2.42)	(2.25)	(1.80)	(2.30)	(0.44)	(2.18)
Asian males	0.459	N/A	-0.722	0.684*	-0.550*	1.073*	-0.153	0.362
	(0.53)		(1.16)	(1.98)	(2.22)	(2.47)	(0.60)	(0.83)
Observations	1,697	511	1,238	1,239	7,883	1,603	3,896	1,782

Table A.19 Continued

	Navy	State	Transport.	Treasury	Vet. Aff.	Other Agency
White females	0.514**	0.175	0.438**	0.387**	0.324**	0.768**
	(12.00)	(0.59)	(5.23)	(8.96)	(6.94)	(9.59)
Black females	-0.012	0.579	-0.158	-0.115*	-0.043	0.515**
	(0.18)	(1.81)	(1.24)	(2.16)	(0.76)	(5.42)
Hispanic females	-0.138	N/A	0.217	0.115	0.577**	0.082
	(0.87)		(0.77)	(1.20)	(4.56)	(0.43)
Asian females	0.347**	-0.438	1.521**	0.450**	-0.256	0.781**
	(3.91)	(0.57)	(4.00)	(3.00)	(1.34)	(3.00)
Black males	0.022	-0.068	-0.782**	-0.322**	-0.088	-0.499**
	(0.31)	(0.12)	(5.19)	(3.97)	(1.31)	(4.01)
Hispanic males	-0.388**	N/A	-0.254	0.009	-0.051	0.566**
	(3.47)		(1.40)	(0.08)	(0.41)	(2.72)
Asian males	-0.092	0.242	0.350	0.192	0.109	1.409**
	(1.25)	(0.32)	(1.66)	(1.30)	(0.58)	(5.13)
Observations	18,028	499	4,791	14,618	12,131	4,527

Absolute value of z statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%

Table A.20: Race and gender Differences in Performance Appraisal Ratings by Grade Level: Ordered Logit Model

	GS 1-3	GS 4	GS 5	GS 6	GS 7	GS 8	GS 9	GS 10
White females	0.151	0.339**	0.496**	0.400**	0.630**	0.383**	0.424**	0.346**
	(0.93)	(4.78)	(9.34)	(6.12)	(13.46)	(4.61)	(10.76)	(3.03)
Black females	-0.052	0.069	0.173**	0.178*	0.233**	0.162	-0.006	0.333*
	(0.30)	(0.87)	(2.95)	(2.41)	(4.13)	(1.59)	(0.10)	(2.18)
Hispanic females	0.255	0.214	0.190*	0.425**	0.756**	0.026	0.210	-0.014
	(0.80)	(1.68)	(2.00)	(3.46)	(7.84)	(0.15)	(1.78)	(0.06)
Asian females	-0.148	0.173	0.470**	0.798**	0.885**	0.639	0.313*	0.162
	(0.49)	(1.21)	(4.27)	(5.22)	(6.62)	(1.74)	(2.44)	(0.40)
Black males	-0.092	-0.146	0.143	-0.180	-0.235**	-0.255	-0.259**	0.743**
	(0.38)	(1.45)	(1.85)	(1.96)	(2.78)	(1.92)	(3.80)	(2.76)
Hispanic males	0.224	-0.275	-0.076	-0.057	-0.464**	-0.818**	-0.253**	-0.161
	(0.68)	(1.27)	(0.56)	(0.38)	(3.77)	(3.69)	(3.02)	(0.67)
Asian males	N/A	-0.124	-0.238	0.458	0.227	-0.667	0.017	3.373**
		(0.48)	(1.35)	(1.89)	(1.34)	(1.95)	(0.13)	(2.79)
Observations	1,383	7,575	14,248	10,014	13,470	4,315	14,565	2,330

Table A.20 Continued

	GS 11	GS 12	GS 13	GS 14	GS 15	GS 16
White females	0.373**	0.433**	0.310**	0.578**	0.558**	1.137**
	(10.98)	(13.11)	(7.40)	(8.76)	(5.14)	(3.09)
Black females	0.219**	-0.031	-0.241**	-0.036	1.123**	-1.251
	(3.95)	(0.60)	(3.46)	(0.29)	(3.94)	(1.72)
Hispanic females	-0.165	0.354**	0.165	0.328	-1.941	N/A
	(1.65)	(2.92)	(0.93)	(1.02)	(1.89)	
Asian females	-0.097	0.311*	0.760**	-0.085	-0.662	N/A
	(0.85)	(2.50)	(4.30)	(0.38)	(1.13)	
Black males	-0.377**	-0.278**	-0.184**	-0.214*	-0.260	0.498
	(6.28)	(4.79)	(2.64)	(2.11)	(1.39)	(1.15)
Hispanic males	0.178*	-0.115	-0.203	0.226	-0.365	-5.778**
	(2.25)	(1.53)	(1.88)	(1.40)	(1.15)	(4.90)
Asian males	0.071	0.063	0.021	0.200	-0.714**	-1.933
	(0.80)	(0.81)	(0.19)	(1.42)	(2.96)	(1.81)
Observations	19,825	23,997	16,410	8,327	3,645	505

Absolute value of z statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%

Table A.21: Race and gender Differences in Performance Appraisal Ratings by Sex-type of Occupation: Ordered Logit Model

	% male: <11	11-20	21-30	31-40	41-50
White females	0.190	0.440**	0.524**	0.259**	0.364**
	(1.57)	(7.09)	(11.36)	(6.73)	(9.34)
Black females	-0.289*	0.138*	0.248**	0.073	-0.049
	(2.32)	(2.03)	(4.54)	(1.53)	(0.97)
Hispanic females	-0.001	0.192	0.461**	0.014	0.211
	(0.01)	(1.89)	(5.16)	(0.15)	(1.84)
Asian females	0.100	0.407**	0.694**	0.120	0.715**
	(0.66)	(3.13)	(6.24)	(0.99)	(4.57)
Black males	-0.203	0.103	-0.219**	-0.157**	-0.338**
	(1.00)	(0.84)	(2.73)	(2.63)	(5.20)
Hispanic males	-0.658	-0.068	-0.173	-0.095	-0.107
	(1.23)	(0.30)	(1.04)	(0.83)	(0.90)
Asian males	0.374	-0.708	0.068	-0.402**	0.069
	(0.57)	(1.89)	(0.36)	(2.60)	(0.40)
Observations	12,486	10,283	15,567	17,819	15,678

Table A.21 Continued

	51-60	61-70	71-80	81-90	91 or above
White females	0.341**	0.360**	0.420**	0.011	0.182*
	(6.56)	(9.10)	(8.64)	(0.21)	(2.37)
Black females	-0.149*	-0.218**	-0.161	-0.501**	-0.166
	(2.07)	(3.08)	(1.80)	(3.79)	(0.78)
Hispanic females	0.154	0.181	0.483*	-0.092	-0.716
	(0.95)	(1.49)	(2.50)	(0.51)	(1.64)
Asian females	0.005	0.486**	0.166	-0.209	-0.322
	(0.03)	(2.66)	(1.27)	(0.96)	(1.34)
Black males	-0.501**	-0.198**	-0.006	-0.359**	-0.133*
	(5.71)	(2.82)	(0.06)	(5.76)	(2.06)
Hispanic males	-0.003	-0.016	0.045	-0.177*	-0.213**
	(0.02)	(0.14)	(0.46)	(2.33)	(2.99)
Asian males	0.105	0.463**	0.075	0.286**	-0.001
	(0.71)	(3.65)	(0.64)	(2.83)	(0.02)
Observations	8,795	14,630	12,559	15,666	17,126

Absolute value of z statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%

Table A.22: Difference in the Impact of Performance Ratings on Percent Change in Salary Growth in the Upper and Lower Grade Levels

Upper Level (Grade 14 or Above)

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.401**	0.366	0.327	-5.413	0.287	0.400**	-0.324	-0.435
	(4.35)	(1.40)	(0.59)	(1.32)	(0.27)	(4.59)	(0.38)	(0.64)
Exceeds fully successful	0.162	0.069	-0.001	-1.753	0.463	0.162	0.029	0.221
	(1.81)	(0.25)	(0.00)	(0.48)	(0.42)	(1.91)	(0.04)	(0.38)

Lower Level (Grade 5 or Below)

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.611**	0.275**	0.551**	1.191**	0.484	0.054	0.564	-0.884 [B]
	(3.11)	(2.71)	(3.62)	(3.31)	(1.12)	(0.19)	(0.91)	(1.88)
Exceeds fully successful	-0.037	0.096	0.503** [A]	0.914** [A]	0.617	-0.002	0.540	-0.713
	(0.21)	(0.99)	(3.70)	(2.68)	(1.53)	(0.01)	(1.09)	(0.84)

Absolute value of t statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%
A: Differences in coefficients between white male and other groups are statistically significant at .01 level, B: significant at .05 level

Table A.23: Difference in the Impact of Performance Ratings on Percent Change in Salary Growth in the Male-Dominant and Female-Dominant Occupations

Male-Dominant (71% or Above Male)

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.191**	0.491*	-0.244	1.634	0.645	0.533*	0.407	0.229
	(3.02)	(2.15)	(0.49)	(1.39)	(0.88)	(2.29)	(1.29)	(0.68)
Exceeds fully successful	-0.056	-0.373	-0.060	1.021	1.054	0.426*	0.170 [B]	-0.194
	(1.02)	(1.85)	(0.14)	(1.00)	(1.58)	(2.11)	(0.64)	(0.68)

Female-Dominant (30% or Below Male)

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.446**	0.320**	0.495**	0.759**	0.732	0.241	0.241	1.248
	(2.68)	(4.33)	(4.35)	(2.72)	(1.90)	(0.74)	(0.22)	(1.16)
Exceeds fully successful	0.035	0.083	0.264*	0.346	0.337	-0.355	-0.330	1.076
	(0.23)	(1.14)	(2.51)	(1.27)	(0.91)	(1.25)	(0.43)	(1.07)

Absolute value of t statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%
A: Differences in coefficients between white male and other groups are statistically significant at .01 level, B: significant at .05 level

Table A.24: Difference in the Impact of Performance Ratings on Promotion in the Upper and Lower Grade Levels

Upper Level (Grade 14 or Above)

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	1.313**	0.661	1.508	N/A	N/A	1.824*	N/A	-6.628
	(5.90)	(1.68)	(1.19)			(2.52)		(0.15)
Exceeds fully Successful	0.788**	0.278	1.214	N/A	N/A	0.499	N/A	-3.495
	(3.50)	(0.67)	(0.92)			(0.69)		(0.03)

N/A: Coefficients are not available because either all the observations have the same value (i.e., nobody is promoted) or the sample size is zero

Lower Level (Grade 5 or Below)

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.526**	0.257**	0.510**	0.992**	0.411	0.094	1.209*	-0.999
	(4.05)	(3.74)	(5.04)	(3.60)	(1.11)	(0.45)	(2.31)	(0.87)
Exceeds fully Successful	0.150	0.060	0.354**	0.892** [B]	0.458	0.078	0.213	-0.327
	(1.31)	(0.92)	(4.00)	(3.51)	(1.37)	(0.44)	(0.50)	(0.35)

Absolute value of z statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%
A: Differences in coefficients between white male and other groups are statistically significant at .01 level, B: significant at .05 level

Table A.25: Difference in the Impact of Performance Ratings on Promotion in the Male-Dominant and Female-Dominant Occupations

Male-Dominant (71% or Above Male)

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.420**	0.592**	0.093	0.719	0.986	0.451*	0.646**	0.361
	(7.78)	(4.71)	(0.29)	(0.75)	(1.78)	(2.41)	(3.05)	(1.29)
Exceeds fully Successful	0.148**	0.214**	0.096	-0.070	0.462	0.367*	0.289*	0.047
	(3.16)	(1.94)	(0.37)	(0.09)	(0.89)	(2.35)	(1.67)	(0.29)

Female-Dominant (30% of Below Male)

	White males	**White females**	**Black females**	**Hispanic females**	**Asian females**	**Black males**	**Hispanic males**	**Asian males**
Outstanding	0.419**	0.279**	0.429**	0.749**	0.520	0.308	0.461	0.888
	(3.41)	(5.34)	(5.59)	(3.89)	(1.87)	(1.32)	(0.49)	(0.87)
Exceeds fully Successful	0.129	0.073	0.256**	0.384*	0.086	0.166	0.017	-0.873
	(1.13)	(1.41)	(3.65)	(2.12)	(0.33)	(0.84)	(0.03)	(0.95)

Absolute value of z statistics in parentheses
Model includes individual and job-related characteristics
* significant at 5%; ** significant at 1%
A: Differences in coefficients between white male and other groups are statistically significant at .01 level, B: significant at .05 level

REFERENCES

Adams, J. S. (1965). Inequity in Social Exchange. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2). New York: Academic Press.

Alkadry, M. G., & Tower, L. E. (2006). Unequal Pay: The Role of Gender. *Public Administration Review, 66*(6), 888-898.

Alonso, P., & Lewis, G. (2001). Public Service Motivation and Job Performance: Evidence from the Federal Sector. *The American Review of Public Administration, 31*(4), 363-380.

Ambrose, M. L., & Kulik, C. T. (1999). Old Friends, New Faces: Motivation Research in the 1990s. *Journal of Management, 25*(3), 231–292.

Asch, B. (2005). The Economic Complexities of Incentive Reforms. In R. Klitgaard & P. C. Light (Eds.), *High Performance Government: Structure, Leadership, Incentives*. Santa Monica, CA: RAND Corporation.

Asch, B. J. (2001). *The Pay, Promotion, and Retention of Civil Service Workers in the Department of Defense*. Santa Monica, CA: RAND.

Avolio, B., Waldman, D., & McDaniel, M. (1990). Age and Work Performance in Nonmanagerial Jobs: The Effects of Experience and Occupational Type. *Academy of Management Journal*, 407-422.

Banks, C. G., & Murphy, K. R. (1985). Toward Narrowing the Research-Practice Gap in Performance Appraisal. *Personnel Psychology, 38*(2), 335-345.

Barkley, A. P., Stock, W. A., & Sylvius, C. K. (1999). Agricultural Graduate Earnings: The Impacts of College, Career, and Gender. *American Journal of Agricultural Economics, 81*(4), 785-800.

Becker, G. S. (1971). *The Economics of Discrimination*. Chicago, IL: University of Chicago Press.

Becker, G. S. (1993). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education* (Third ed. ed.). Chicago, IL: University of Chicago Press.

Belman, D., & Heywood, J. S. (1991). Sheepskin Effects in the Returns to Education: An Examination of Women and Minorities. *The Review of Economics and Statistics, 73*(4), 720-724.

Blau, F. D., & DeVaro, J. L. (2006). New Evidence on Gender Difference in Promotion Rates: An Empirical Analysis of a Sample of New Hires. *NBER Working Paper*.

Blau, F. D., Ferber, M. A., & Winkler, A. E. (2006). *The Economics of Women, Men, and Work*. Upper Saddle River: Pearson Prentice Hall.

Bolino, M. C., & Turnley, W. H. (2008). Old Faces, New Places: Equity Theory in Cross-Cultural Contexts. *Journal of Organizational Behavior, 29*(1), 29-50.

Borjas, G. J. (1980). Wage Determination in the Federal Government: The Role of Constituents and Bureaucrats. *The Journal of Political Economy, 88*(6), 1110-1147.

Borjas, G. J. (1982). Politics of Employment Discrimination in the Federal Bureaucracy. *Journal of Law & Economics, 25*(2), 271-299.

Boswell, W. R., & Boudreau, J. W. (2000). Employee Satisfaction with Performance Appraisals and Appraisers: The Role of Perceived Appraisal Use. *Human Resource Development Quarterly, 11*(3), 283-299.

Boswell, W. R., & Boudreau, J. W. (2002). Separating the Developmental and Evaluative Performance Appraisal Uses. *Journal of Business and Psychology, 16*(3), 391-412.

Boyd, N. M. (2004). Expanding the View of Performance Appraisal by Introducing Social Justice Concerns. *Administrative Theory & Praxis, 26*(3), 249-278.

Bretz, R. D., Milkovich, G. T., & Read, W. (1992). The Current State of Performance Appraisal Research and Practice: Concerns, Directions, and Implications. *Journal of Management, 18*(2), 321-352.

Buckley, M. R., Miceli, N. S., Purvis, E. S., & Mahoney, D. L. G. (1995). Perceptions of Inequity in Performance Appraisal Resulting from the Implementation of the Americans with Disabilities Act. In J. G. Veres & R. R. Sims (Eds.), *Human Resource Management and the Americans with Disabilities Act* (pp. 105). Westport, CT: Quorum Books.

Bucklin, B. R., & Dickinson, A. M. (2001). Individual Monetary Incentives: A Review of Different Types of Arrangements between Performance and Pay. *Journal of Organizational Behavior Management, 21*(3), 45-137.

Canton, E. (2005). Power of Incentives in Public Organizations When Employees Are Intrinsically Motivated. *Journal of Institutional and Theoretical Economics 161*(4), 664-680.

Castilla, E. J. (2008). Gender, Race, and Meritocracy in Organizational Careers. *American Journal of Sociology, 113*(6), 1479-1526.

Catanzarite, L. (2003). Race-Gender Composition and Occupational Pay Degradation. *Social Problems, 50*(1), 14-37.

Chen, M. J., & Miller, D. (1994). Competitive Attack, Retaliation and Performance: An Expectancy-Valence Framework. *Strategic Management Journal, 15*(2), 85-102.

Chen, Y., Gupta, A., & Hoshower, L. (2006). Factors That Motivate Business Faculty to Conduct Research: An Expectancy Theory Analysis. *The Journal of Education for Business, 81*(4), 179-189.

Cleveland, J. N., & Landy, F. J. (1981). The Influence of Rater and Ratee Age on Two Performance Judgments. *Personnel Psychology, 34*(1), 19-29.

Cleveland, J. N., Murphy, K. R., & Williams, R. E. (1989). Multiple Uses of Performance Appraisal: Prevalence and Correlates. *Journal of applied psychology, 74*(1), 130-135.

Cobb-Clark, D. (2001). Getting Ahead: The Determinants of and Payoffs to Internal Promotion for Young Men and Women. *Research in Labor Economics, 20*, 339-372.

Coens, T., & Jenkins, M. (2000). *Abolishing Performance Appraisals: Why They Backfire and What to Do Instead*. San Francisco: Berrett-Koehler Publishers.

Cohen, P. N., & Huffman, M. L. (2003). Occupational Segregation and the Devaluation of Women's Work across U.S. Labor Markets. *Social Forces, 81*(3), 881-908.

Condrey, S. E., & Brudney, J. L. (1992). Performance-Based Managerial Pay in the Federal Government: Does Agency Matter? *Journal of Public Administration Research and Theory, 2*(2), 157-174.

Cotter, D. A., Hermsen, J. M., Ovadia, S., & Vanneman, R. (2001). The Glass Ceiling Effect. *Social Forces, 80*(2), 655-681.

Cox, T., & Nkomo, S. M. (1992). Candidate Age as a Factor in Promotability Ratings. *Public Personnel Management, 21*(2), 197-210.

Daley, D. M. (1992). *Performance Appraisal in the Public Sector: Techniques and Applications*. Westport, CT: Quorum Books.

Daley, D. M. (2005). Designing Effective Performance Appraisal Systems. In S. E. Condrey (Ed.), *Handbook of Human Resource Management in Government*. San Francisco: Jossey-Bass.

DiPrete, T., & Soule, W. (1986). The Organization of Career Lines: Equal Employment Opportunity and Status Advancement in a Federal Bureaucracy. *American Sociological Review, 51*(3), 295-309.

DlTomaso, N., & Smith, S. A. (1996). Race and Ethnic Minorities and White Women in Management. In J. Tang & E. Smith (Eds.), *Women and Minorities in American Professions* (pp. 87-110). Albany, NY: State University of New York Press.

Drazin, R., & Auster, E. R. (1987). Wage Differences between Men and Women: Performance Appraisal Ratings Vs. Salary Allocation as the Locus of Bias. *Human Resource Management, 26*(2), 157-168.

Eisenberg, E. F., & Ingraham, P. W. (1993). Analyzing the Comparative Pay-for-Performance Experience: Are There Common Lessons? *Public Productivity & Management Review, 17*(2), 117-128.

Elvira, M., & Town, R. (2001). The Effects of Race and Worker Productivity on Performance Evaluations. *Journal of Economy and Society, 40*(4), 571-590.

Frey, B. S. (1997). On the Relationship between Intrinsic and Extrinsic Work Motivation. *International Journal of Industrial Organization, 15*(4), 427-439.

Fuller, J. J., & Tinkham, R. (2002). Making the Most of Scarce Reward Dollars: Why Differentiation Makes a Difference. *Employee Benefits Journal, 27*(3), 3-8.

Gabris, G. T., & Ihrke, D. M. (2000). Improving Employee Acceptance toward Performance Appraisal and Merit Pay System: The Role of Leadership Credibility. *Review of Public Personnel Administration, 20*(1), 41-53.

Gaertner, K. N., & Gaertner, G. H. (1985). Performance-Contingent Pay for Federal Aianagers. *Administration & Society, 17*(1), 7-20.

Gerhart, B. (1990). Gender Differences in Current and Starting Salaries: The Role of Performance, College Major, and Job Title. *Industrial and Labor Relations Review, 43*(4), 418-433.

Gerhart, B. A., & Rynes, S. (2003). *Compensation: Theory, Evidence, and Strategic Implications*. Thousand Oaks, CA: Sage Publications.

Greenberg, J. (1990). Organizational Justice: Yesterday, Today, and Tomorrow. *Journal of Management, 16*(2), 399-432.

Greenberg, J. (1991). Motivation to Inflate Performance Ratings: Perceptual Bias or Response Bias? *Motivation and Emotion, 15*(1), 81-97.

Greene, W. H. (2003). *Econometric Analysis* Upper Saddle River, NJ: Pearson Education.

Griffeth, R. W., & Bedeian, A. G. (1989). Employee Performance Evaluations: Effects of Ratee Age, Rater Age, and Ratee Gender. *Journal of Organizational Behavior, 10*, 83-90.

Griffeth, R. W., Vecchio, R. P., & Logan, J. W. (1989). Equity Theory and Interpersonal Attraction. *Journal of applied psychology, 74*(3), 394-401.

Hancock, D. R. (1995). What Teachers May Do to Influence Student Motivation: An Application of Expectancy Theory. *The Journal of General Education, 44*, 171-179.

Hellerstein, J. K., Neumark, D., & Troske, K. R. (1999). Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations. *Journal of Labor Economics, 17*(3), 409-446.

Heneman, R. L., & Werner, J. M. (2005). *Merit Pay: Linking Pay to Performance in a Changing World*. Greenwich, CT: Information Age Publishing.

Hersch, J., & Viscusi, W. (1996). Gender Differences in Promotions and Wages. *Industrial Relations, 35*(4), 461-472.

Houston, D. J. (2000). Public-Service Motivation: A Multivariate Test. *Journal of Public Administration Research and Theory, 10*(4), 713-727.

Hungerford, T., & Solon, G. (1987). Sheepskin Effects in the Returns to Education. *The Review of Economics and Statistics, 69*(1), 175-177.

Hyde, A. C. (2005). Pay for Performance. *Public Manager, 34*(1), 3-7.

Ingraham, P. W. (1993). Of Pigs in Pokes and Policy Diffusion: Another Look at Pay-for-Performance. *Public Administration Review, 53*(4), 348-356.

James, H. S. (2005). Why Did You Do That? An Economic Examination of the Effect of Extrinsic Compensation on Intrinsic Motivation and Performance. *Journal of Economic Psychology, 26*(4), 549-566.

Jawahar, I. M., & Williams, C. R. (1997). Where All the Children Are above Average: The Performance Appraisal Purpose Effect. *Personnel Psychology, 50*(4), 905-925.

Johnson, R. N., & Libecap, G. D. (1989). Bureaucratic Rules, Supervisor Behavior, and the Effect on Salaries in the Federal Government *Journal of Law, Economics, & Organization, 5*(1), 53-82.

Kellough, J. E. (2002). Employee Performance Appraisal and Pay for Performance in the Public Sector: A Critical Examination. In C. Ban & N. M. Riccucci (Eds.), *Public Personnel Management: Current Concerns, Future Challenges*. New York: Longman Publishers.

Kellough, J. E., & Nigro, L. G. (2002). Pay for Performance in Georgia State Government: Employee Perspectives on Georgiagain after 5 Years. *Review of Public Personnel Administration, 22*(2), 146-166.

Kelly, R. M., Guy, M. E., Bayes, J., Duerst-Lahti, G., Duke, L. L., Hale, M. M., et al. (1991). Public Managers in the States: A Comparison of Career Advancement by Sex. *Public Administration Review, 51*(5), 402-412.

Klingner, D. E., & Nalbandian, J. (1998). *Public Personnel Management: Contexts and Strategies*. Upper Saddle River: Prentice Hall.

Kräkel, M. (2008). Emotions in Tournaments. *Journal of Economic Behavior & Organization, 67*(1), 204-214.

Krull, H. (2006). *Earnings Disparities: The Role of Internal Promotions and Job Performance Evaluations*. Paper presented at the Center for Financial Research Seminar, Federal Deposit Insurance Corporation.

Lawler, E. E. (1971). *Pay and Organizational Effectiveness: A Psychological View*. New York: McGraw-Hill.

Lewis, G. B. (1986). Gender and Promotions: Promotion Chances of White Men and Women in Federal White-Collar Employment. *Journal of Human Resources, 21*(3), 406-419.

Lewis, G. B. (1992). Men and Women toward the Top: Backgrounds, Careers, and Potential of Federal Middle Managers. *Public Personnel Management, 21*(4), 473-491.

Lewis, G. B. (1996). Gender Integration of Occupations in the Federal Civil Service: Extent and Effects on Male-Female Earnings. *Industrial and Labor Relations Review, 49*(3), 472-483.

Lewis, G. B. (1997a). Guns, Butter, and Federal Careers: Growth, Decline, and Personnel in Defense and Domestic Agencies. *Journal of Public Administration Research and Theory, 7*(1), 59-84.

Lewis, G. B. (1997b). Race, Sex, and Performance Ratings in the Federal Service. *Public Administration Review 57*(6), 479-489.

Lewis, G. B., & Oh, S. S. (2008). Graduate Education and Federal Career Success: How Well Does the Mpa Stack Up? *Journal of Public Affairs Education, 14*(3), 463-474.

Liff, S. (2007). *Managing Government Employees: How to Motivate Your People, Deal with Difficult Issues, and Achieve Tangible Results*. New York: American Management Association.

Light, P. C. (1999). Behold Lake Wobegon East. *Government Executive*.

Locke, E., Feren, D., McCaleb, V., Shaw, K., & Denny, A. (1980). The Relative Effectiveness of Four Ways of Motivating Employee Performance. In K. Duncan, M. Gruenberg & D. Wallis (Eds.), *Changes in Working Life*. New York: Wiley.

London, M., & Stumpf, S. A. (1983). Effects of Candidate Characteristics on Management Promotion Decisions: An Experimental Study. *Personnel Psychology, 36*(2), 241-259.

Long, J. S., & Freese, J. (2006). *Regression Models for Categotical Dependent Variables Using Stata* (Second ed.). College Station: Stata Press.

Losey, S. (2009). Performance Pay for All Feds. *FederalTimes.Com*.

Loury, L. D. (1997). The Gender Earnings Gap among College-Educated Workers. *Industrial and Labor Relations Review, 50*(4), 580-593.

Lyness, K. S., & Heilman, M. E. (2006). When Fit Is Fundamental: Performance Evaluations and Promotions of Upper-Level Female and Male Managers. *Journal of Applied Psychology, 91*(4), 777-785.

Lyness, K. S., & Judiesch, M. K. (1999). Are Women More Likely to Be Hired or Promoted into Management Positions? *Journal of Vocational Behavior, 54*(1), 158-173.

McEvoy, G. M., & Cascio, W. F. (1989). Cumulative Evidence of Relationship between Employee Age and Job Performance. *Journal of applied psychology, 74*(1), 11-17.

McKinney, W. R., & Collins, J. R. (1991). The Influence of Race, Sex, and Age on Performance Appraisal Bias in Public Parks and Recreation. *Journal of Park and Recreation Administration, 9*(3).

Medoff, J. L., & Abraham, K. G. (1981). Experience, Performance, and Earnings. NBER Working Paper No. W0278.

Milkovich, G. T., & Newman, J. M. (1996). *Compensation*. Chicago, IL: Irwin.

Milkovich, G. T., & Wigdor, A. K. (Eds.). (1991). *Pay for Performance: Evaluating Performance Appraisal and Merit Pay*. Washington, D. C.: National Academy Press.

Miller, F. A., & Judith, H. K. (2002). *The Inclusion Breakthrough: Unleashing the Real Power of Diversity*. San Francisco, CA: Berrett-Koehler Publishers.

Miller, W., Kerr, B., & Reid, M. (1999). A National Study of Gender-Based Occupational Segregation in Municipal Bureaucracies: Persistence of Glass Walls? *Public Administration Review, 59*(3), 218-229.

Mincer, J. (1993). *Studies in Human Capital*. Brookfield, VT: Edward Elgar.

Mitra, A., Gupta, N., & Jenkins, G. D. (1995). The Case of the Invisible Merit Raise: How People See Their Pay Raises. *Compensation & Benefits Review, 27*(3), 71-76.

Mobley, W. H. (1982). Supervisor and Employee Race and Sex Effects on Performance Appraisals: A Field Study of Adverse Impact and Generalizability. *Academy of Management Journal, 25*(3), 598-606.

Montoya, S., & Graham, J. (2007). *Modernizing the Federal Government: Paying for Performance*. Santa Monica: RAND Corporation.

Mowday, R. T. (1991). Equity Theory Predictions of Behavior in Organizations. In R. M. Steers & L. W. Porter (Eds.), *Motivation and Work Behavior*. New York: McGraw-Hill.

Murphy, K. J. (1992). Performance Measurement and Appraisal: Motivating Managers to Identify and Reward Performance. In William J. Burns Jr. (Ed.), *Performance Measurement, Evaluation, and Incentives* (pp. 37-62). Boston: Harvard Business School Press.

Murphy, K. M., & Welch, F. (1992). The Structure of Wages. *The Quarterly Journal of Economics, 107*(1), 285-326.

Murphy, K. R. (2008). Perspectives on the Relationship between Job Performance and Ratings of Job Performance. *Industrial and Organizational Psychology, 1*, 197-205.

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*. Thousand Oaks: Sage.

Naff, K. C. (1994). Through the Glass Ceiling: Prospects for the Advancement of Women in the Federal Civil Service. *Public Administration Review, 54*(6), 507-514.

Naff, K. C., & Newman, M. A. (2004). Federal Civil Service Reform: Another Legacy of 9/11? *Review of Public Personnel Administration, 24*(3), 191-201

National Institute of Standards and Technology. (2009). Human Resource Systems: Performance Management. Retrieved February 15, 2009, from http://www.nist.gov/hrmd/projectops/hrsystems/performance.htm

Nigro, L. G., Nigro, F. A., & Kellough, J. E. (2007). *The New Public Personnel Administration*. Belmont, CA: Thomson-Wadsworth.

O'Neill, J., & Polachek, S. (1993). Why the Gender Gap in Wages Narrowed in the 1980s. *Journal of Labor Economics, 11*(1), 205-227.

Oh, S. S., & Lewis, G. B. (2009). Can Performance Appraisal Systems Inspire Intrinsically Motivated Employees? *Review of Public Personnel Administration, 29*(2), 158-167.

Ortiz, S. Y., & Roscigno, V. J. (2009). Discrimination, Women, and Work: Processes and Variations by Race and Class. *The Sociological Quarterly, 50*(2), 336-359.

Padavic, I., & Reskin, B. F. (2002). *Women and Men at Work*. Thousand Oaks, CA: Pine Forge Press.

Paulin, E. A., & Mellor, J. M. (1996). Gender, Race, and Promotions within a Private-Sector Firm. *Industrial Relations, 35*(2), 276-295.

Pearce, J. L., & Perry, J. L. (1983). Federal Merit Pay: A Longitudinal Analysis. *Public Administration Review, 43*(4), 315-325.

Perry, J. L. (1996). Measuring Public Service Motivation: An Assessment of Construct Reliability and Validity. *Journal of Public Administration Research and Theory, 6*(1), 5-22.

Peterson, J., & Lewis, M. (1999). *The Elgar Companion to Feminist Economics*. Northampton, MA: Edward Elgar.

Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *The American Economic Association, 4*, 659-661.

Powell, G. N., & Butterfield, D. A. (1994). Investigating The "Glass Ceiling" Phenomenon: An Empirical Study of Actual Promotions to Top Management. *Academy of Management Journal, 37*(1), 68-86.

Rainey, H. G. (2003). *Understanding and Managing Public Organizations* (Third ed.). San Francisco, CA: Jossey-Bass.

Reskin, B. F., McBrier, D. B., & Kmec, J. A. (1999). The Determinants and Consequences of Workplace Sex and Race Composition. *Annual Review of Sociology, 25*(1), 335-361.

Rhodes, S. R. (1983). Age-Related Differences in Work Attitudes and Behavior: A Review and Conceptual Analysis. *Psychological Bulletin, 93*(2), 328-367.

Riccucci, N. M., & Naff, K. C. (2008). *Personnel Management in Government Politics and Process*. Boca Raton, FL: CRC Press.

Roch, S. G. (2005). An Investigation of Motivational Factors Influencing Performance Ratings. *Journal of Managerial Psychology, 20*(8), 695-711.

Rosenberg, A. (2009). Administration Will Push for Governmentwide Pay for Performance. *Government Executive*.

Ryan, R. M., & Deci, E. L. (2000). When Rewards Compete with Nature: The Undermining of Intrinsic Motivation and Self-Regulation. In C. J. M. H. Sansone (Ed.), *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance*. San Diego: Academic Press.

Rynes, S. L., Gerhart, B., & Minette, K. A. (2004). The Importance of Pay in Employee Motivation: Discrepancies between What People Say and What They Do. *Human Resource Management, 43*(4), 381-394.

Rynes, S. L., Gerhart, B., & Parks, L. (2005). Personnel Psychology: Performance Evaluation and Pay for Performance. *Annual Review of Psychology, 56*, 571-600

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of Job Experience and Ability on Job Knowledge, Work Sample Performance, and Supervisory Ratings of Job Performance. *Journal of Applied Psychology, 71*(3), 432-439.

Scullen, S. E., Bergey, P. K., & Aiman-Smith, L. (2005). Forced Distribution Rating Systems and the Improvement of Workforce Potential: A Baseline Simulation. *Personnel Psychology, 58*(1), 1-32.

Sherk, J. (2007). Performance-Based Pay Driving Increase in Inequality, *WebMemo #1505*: The Heritage Foundation.

Skirbekk, V. (2003). Age and Individual Productivity: A Literature Survey. *Max Planck Institute for Demographic Research Working Paper 2003-028*.

Smith, A. D., & Rupp, W. T. (2003). Knowledge Workers: Exploring the Link among Performance Rating, Pay and Motivational Aspects. *Journal of Knowledge Management 7*(1), 107-124.

Spence, A. M. (1974). *Market Signalling: Information Transfer in Hiring and Related Screening Processes*. Cambridge, MA: Harvard University Press.

Spyropoulos, D. (2005). *Analysis of Career Progression and Job Performance in Internal Labor Markets: The Case of Federal Civil Service Employees.* Naval Postgraduate School, Monterey, CA.

Stajkovic, A. D., & Luthans, F. (2003). Behavioral Management and Task Performance in Organizations: Conceptual Background, Meta-Analysis, and Test of Alternative Models. *Personnel Psychology, 56*(1), 155-194.

Stauffer, J. M., & Buckley, M. R. (2005). The Existence and Nature of Racial Bias in Supervisory Ratings. *Journal of applied psychology, 90*(3), 586-591.

Stivers, C. (2002). *Gender Images in Public Administration: Legitimacy and the Administrative State*. Thousand Oaks: CA: Sage Publications.

Sturman, M. (2003). Searching for the Inverted U-Shaped Relationship between Time and Performance: Meta-Analyses of the Experience/Performance, Tenure/Performance, and Age/Performance Relationships. *Journal of Management, 29*(5), 609-640.

Taylor, M. S., Tracy, K. B., Renard, M. K., Harrison, J. K., & Carroll, S. J. (1995). Due Process in Performance Appraisal: A Quasi-Experiment in Procedural Justice. *Administrative Science Quarterly, 40*(3), 495-523.

Taylor, P. A. (1979). Income Inequality in the Federal Civilian Government. *American Sociological Review, 44*(3), 468-479.

Teel, K. S. (1986). Are Merit Raises Really Based on Merit? *Personnel Journal, 65*(3), 88-94.

Tien, F. F. (2000). To What Degree Does the Desire for Promotion Motivate Faculty to Perform Research? Testing the Expectancy Theory. *Research in Higher Education, 41*(6), 723-752.

U.S. General Accounting Office. (2003). *Results-Oriented Cultures: Modern Performance Management Systems Are Needed to Effectively Support Pay for Performance*. Washington, D.C.: U.S. General Accounting Office.

U.S. Merit Systems Protection Board. (2005). *The 2005 Merit Principles Survey*. Washington, D.C.: U.S. Merit Systems Protection Board.

U.S. Office of Personnel Management. (1999). *The Fact Book: Federal Civilian Workforce Statistics*. Washington, D.C.: U.S. Office of Personnel Management.

U.S. Office of Personnel Management. (2001). *Supervisors in the Federal Government: A Wake-up Call*. Washington, D.C.: U.S. Office of Personnel Management.

U.S. Office of Personnel Management. (2002a). *2002 Demographic Profile of the Federal Workforce*. Washington, D.C.: U.S. Office of Personnel Management.

U.S. Office of Personnel Management. (2002b). *A Fresh Start for Federal Pay: The Case for Modernization*. Washington, DC: U.S. Office of Personnel Management.

U.S. Office of Personnel Management. (2008a). About the Central Personnel Data File. Retrieved April 8, 2008, from http://www.opm.gov/feddata/html/acpdf.asp

U.S. Office of Personnel Management. (2008b). *Alternative Personnel Systems in the Federal Government*. Washington, D.C.: U.S. Office of Personnel Management.

U.S. Office of Personnel Management. (2008c). Chronology of Employee Performance Management in the Federal Government. Retrieved May 8, 2008, from http://www.opm.gov/perform/chron.asp

U.S. Office of Personnel Management. (2008d). Performance Management and Quality Step Increases Faqs. from http://www.opm.gov/perform/faqs/qualsi.asp

U.S. Office of Personnel Management. (2008e). *Senior Executive Service Survey Results*. Washington, D.C.: U.S. Office of Personnel Management.

U.S. Office of Personnel Management. (2009). *Handbook of Occupational Groups and Families*. Washington, D.C.: U.S. Office of Personnel Management.

Vroom, V. H. (1964). *Work and Motivation*. New York: Wiley and Sons.

Waldman, D. A., & Avolio, B. J. (1991). Race Effects in Performance Evaluations: Controlling for Ability, Education, and Experience. *Journal of applied psychology, 76*(6), 897–901.

Williams, R. S., & Walker, J. (1985). Sex Differences in Performance Rating: A Research Note. *Journal of Occupational Psychology, 58*(4), 331-337.

Wooldridge, J. M. (2003). *Introductory Econometrics*. Mason, OH: Thomson.

Zenger, T. R. (1992). Why Do Employers Only Reward Extreme Performance? Examining the Relationships among Performance, Pay, and Turnover. *Administrative Science Quarterly, 37*(2), 198-219.