UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

FACULTY OF BUSINESS ADMINISTRATION AND ECONOMICS

Working Paper Series

Subjective Appraisals of Employee Potential: Do Gender and Managerial Level Matter?

Anica Rose

January 2017

WiWi

Subjective Appraisals of Career Potential: Do Gender and Managerial Level Matter?

Anica Rose

Paderborn University – Warburger Str. 100 – 33098 Paderborn

anica.rose@upb.de – 00495251602087 – 00495251603242

Abstract

While a growing number of empirical studies have analyzed gender differences at various career stages, there is a dearth of studies about formal appraisals of men's and women's career potential, i.e., their promotability. In this paper, I will empirically analyze whether female employees' promotability assessments are systematically inferior to their equally qualified male colleagues. In doing so, I use detailed personnel data of a large global German company that has a formal promotability evaluation process in place. I consider a wide range of contextual variables that have been neglected in the past, such as information on employees' demographic (i.e., gender, age, tenure) and job-related characteristics (i.e., pay grade, working hours, performance assessments), additional information on the employees' direct supervisors, and the composition of the department. I find women's likelihood of receiving an evaluation that qualifies them as promotable to be around 5 percentage points lower than for their male counterparts – the probability of receiving an outstanding assessment being only 20 percent per se. The gap is even more pronounced at around the age of 30, i.e., the average childbearing age in Germany. Furthermore, gender gaps persist at managerial levels, which points to the existence of systematic gender differences in formal promotability evaluation processes.

1 Introduction

Women remain underrepresented in the higher echelons of companies worldwide (Noland et al. 2016). Even though 45 percent of the Standard & Poor's 500 labor force are female, there are only 25 percent of women in executive positions and only 4.4 percent hold CEO positions (Catalyst 2016). In particularly male industries and occupations, there are far fewer women employed (Catalyst 2015). A closer consideration of one of Germany's key industrial sectors – the metal and electrical industry – reveals that female employees account for only 20 percent in this traditionally male-dominated field. While the female ratio is higher in the industry's clerical or administrative occupations, it lies below 20 percent in technical or engineering jobs. Since the sector's management staff is typically recruited from the latter group of employees, the share of women in managerial positions amounts to merely 10 percent (Gesamtmetall 2011). The low representation of women in higher-paying jobs extends through all domains and has, thus, led to various affirmative action policies, such as quotas on boards of directors – not only in Germany, but also in other countries around the globe (Terjesen et al. 2015).[1]

The considerable gender gap has elicited an intense discussion among researchers and politicians about its driving forces. Most often, the persistent disadvantage of women in terms of career success is explained by the "glass ceiling" phenomenon: Women cannot move beyond a certain hierarchical level due to vertical segregation by gender (Boll et al. 2016). Past research demonstrates that several barriers may inhibit women's progression at various stages of their post-hiring career. For instance, empirical research has focused on gender differences in performance evaluations (Bartol 1999; Bretz et al. 1992; Levy and Williams 2004), promotions (Blau and DeVaro 2007; Cobb-Clark 2001; Ransom and Oaxaca 2005), and wages (Barnett et al. 2000; Booth et al. 2003; Gerhart and Milkovich 1989; Grund 2015; Hill et al. 2015). It is found that gender differences are much more pronounced in rewards (including promotion rates and wages) than in performance (Joshi et al. 2015; Roth et al. 2012). With regard to promotions, empirical findings suggest sometimes non-significant (Acosta 2010; Booth et al. 2003; Lewis 1986; Petersen and Saporta 2004; Spilerman and Petersen, 1999) yet sometimes detrimental effects for female employees of being promoted (Blau and DeVaro 2007; Cannings and Montmarquette 1991; Cobb-Clark 2001; Ransom and Oaxaca 2005; Roth et al. 2012). Even though

---

[1] The non-profit organization Catalyst offers a list of countries' most current legislative, regulatory or voluntary approaches to increase gender diversity on boards (Catalyst 2014).

promotability appraisals (i.e., the assessment of the suitability for being promoted) as a crucial feature of career progression seem to be lower for women than for men (Roth et al. 2012), the evidence is clearly limited – undoubtedly because of a lack of data.

From a theoretical perspective, the reasons for women's lower promotion opportunities are twofold: On the one hand, they might reflect intentional discriminatory practices of supervisors (Becker 1971), while, on the other hand, the gender gap might be rationally explained by the (perceived) differences in men's and women's labor-market preferences (Lazear and Rosen 1990). In other words, we do not know, for example, whether supervisors' evaluations of women's career potential are systematically more unfavorable than those of their male colleagues or whether the bias reflects the female employees' labor-market preferences. Promotability appraisals represent the stage between performance evaluations and actual promotions. This phase is crucial because, first, formal systems of subjective appraisals of employees' potential are a widespread and growing HR practice in Western European organizations[2] and, second, they serve as a basis for pivotal personnel decisions: Mentoring, training, and – most importantly – promotions. Receiving favorable evaluations of future potential is, therefore, a necessary (though not sufficient) condition of subsequent career success. Examining this practice as a possible mechanism leading to, or at least reflecting, subjective biases through discrimination helps to explain and alleviate vertical segregation by gender.

In this paper, I contribute to the debate of gender discrimination in promotion opportunities using unique and extensive personnel data from a large German company that has a formal evaluation process of employees' promotability in place. The company is representative of other businesses operating in the German metal and electrical industry. Gender differences in promotion opportunities – if existing – are more likely to become visible in this highly competitive environment than in non-private sector organizations: Due to the fundamental focus on egalitarianism and procedural fairness as well as the rationalized and rigid promotion processes, gender differences were typically found to be insignificant in federal agencies or public services (Barnett et al. 2000; Lewis 1986; Powell and Butterfield 1994). The dataset's extensive individual-level information enables controlling for a wide range of further demographic and job-specific differences between

---

[2] Whilst being a standard HR application in U.S. companies already (Cappelli and Conyon 2016; Joshi et al. 2015; Prendergast 1999), employee assessments have particularly gained importance in the German business context during the past decade (Kampkötter 2016).

men and women that may influence promotability evaluations, such as age, wage level, functional units (i.e., specific work settings) and a formal evaluation of the most recent performance. Moreover, the separation between employees at non-managerial versus managerial levels of the organizational hierarchy makes it possible to study gender differences in two pools of employees with presumably distinct human capital endowments.

The present data allow inferring with a high degree of confidence that discrimination by gender in formal promotability appraisals can indeed contribute to the reported disadvantage in career success. In non-managerial levels, women's career potential assessments are found to be less favorable than those of their male counterparts. The disadvantageous treatment is particularly striking for women at around the age of 30. Moreover, the gender gap seems to be driven by male superiors and persists in the full-time sample – even though I control for individual performance using a non-gender biased indicator. Apart from that, I am not able to rule out gender gaps in managerial levels, in which men and women are assumed to have similar preferences and abilities. These results, which support the existence of taste-based discrimination in formal promotability evaluation processes, can have far-reaching implications for other organizations with similar subjective appraisal systems.

The remainder of this chapter is structured as follows: In the subsequent section 2, I will derive a series of hypotheses based on the relevant theoretical and empirical research on gender differences in promotion opportunities. Section 3 outlines the data, first descriptive findings as well as the methodological approach, whereas the empirical results are provided in section 4. I conclude with a discussion of the results and the implications in section 5.

2 Gender Differences in Promotion Opportunities: Related Literature and Theoretical Background

Previous research suggests that women are less likely than men to make it to the upper organizational echelons (Haslam and Ryan 2008; Ryan and Haslam 2007). The low representation of women in top positions has been attributed to discriminatory practices along the career progression line. So far, however, there is a paucity of studies assessing the impact of the adverse treatment from evaluators concerning employees' promotability assessments. Even though there are a few studies, in which supervisors are asked to assess

their subordinates' potential to move up within the organization[3], the results differ depending on the organizational setting (Roth et al. 2012): While women are found to be disadvantaged in a large transportation firm (Hoobler et al. 2009), there seems to be a positive discrimination in favor of women in non-profit (Thacker and Wayne 1995) and governmental organizations (Harris et al. 2006). These studies' main limitations are that the superiors' assessments do not have any consequences for the subordinates' actual promotion prospects. Despite the low presence of research on employees' career potential, past researchers found evidence of women's career disadvantages concerning their promotion rates (for example Cobb-Clark 2001; Joshi et al. 2015) – the subsequent stage of promotability assessments. These results are statistically significant even when controlling for job-specific performance ratings (Blau and DeVaro 2007; Cannings and Montmarquette 1991), firm characteristics (for example non-profit versus for-profit organizations as in Blau and DeVaro 2007), or the fact that a company has already experienced a class-action lawsuit against gender discrimination (Ransom and Oaxaca 2005).

The economic theories of taste-based (Becker 1971) and statistical discrimination (Aigner and Cain 1977; Arrow 1973; Phelps 1972) provide well-established explanations for the gender differences found in labor-market outcomes. Therefore, the concepts can also be used in the context of promotability evaluations to explain the observable relationships between employees and their supervisors. Even though both theories predict negative outcomes for minority groups, such as women, their underlying explanatory approaches differ. On the one hand, taste-based gender discrimination refers to an employer's negative preferences towards female employees. The systematic taste for discrimination is typically fueled by prejudices towards women's personal attributes. Nevertheless, the discriminatory behavior does not necessarily have to originate from the supervisor's personal distaste, but might also be based on the anticipated disutility of co-workers or customers who would have to cooperate with the minority group. In the long run, discrimination will induce costs to the employer, as the demand for women will decrease, despite the fact that they are as productive as men (Becker 1971). The concept of statistical discrimination, on the other hand, explains the disadvantageous treatment of women on the basis of employers' incomplete information about employees' preferences,

---

[3] In all studies, supervisors have to respond to three items using a 7-point Likert scale: (i) "I believe that this employee will have a successful career", (ii) "If I had to select a successor for my position, it would be this subordinate", and (iii) "I believe that this employee has high potential" (Thacker and Wayne 1995).

effort levels, and future productivity. In case of asymmetric information on employees' individual attitudes and behavior, decisions will be made upon past statistical experience with the minority group or upon prevailing group characteristics associated with certain kinds of role models (Aigner and Cain 1977; Arrow 1973; Phelps 1972). Typical "female attributes" are benevolence, sensitiveness, or cooperativeness – traits that have evolved over centuries and that qualify women to be the "housewives" rather than the "moneymakers". In contrast, men are stereotyped as being more aggressive, dominant and self-confident – attributes that are perceived to correspond better with managerial roles (Azmat and Petrongolo 2014; Niederle 2014). According to the role congruity theory, inferior promotability ratings for female leaders are evoked by traits typically associated with women that are perceived to be incompatible with leadership positions (Eagly and Karau 2002). In line with this, women are often assumed to put less effort into labor-market activities than men or to prefer female-dominated occupations that often allow to work part-time and to be more flexible (Becker 1985). Consequently, biases in the information sets of labor-market preferences or productivities can lead to unintentional discriminatory treatment of women who are in fact as capable and willing as men to strive for success and promotions into top positions. Based on the notion of statistical discrimination, Lazear and Rosen (1990) establish a "jobs theory of discrimination" in order to rationalize the low representation of women in higher-paying jobs. The authors assume men and women to be equally qualified and to receive identical wages in the same occupation. From this equity perspective, men and women with identical qualifications and performance levels should not be evaluated differently with respect to their promotability. However, as soon as employers have to decide upon costly job assignments and the investment in firm-specific human capital, they do not only have to consider employees' abilities, but also have to weigh the costs for promotions against the expected returns (Lazear and Rosen 1990). As an example, men were found to have better social networks, which can even substitute a worse performance (Cannings and Montmarquette 1991). Apart from that, employers have to assess an employee's likelihood of job exit. In order to maximize the organization's predicted returns on investment, it is rational for the employer to prefer an employee who is more likely to remain in the company for as long as possible, even though men and women might not differ in their performance or skill sets. Employers expect women's risk to leave the company to be larger than men's, since female employees are assumed to have higher non-market (i.e., household sector) opportunities than their male counterparts (Lazear and Rosen 1990). This assumption is

generally well accepted: While men have a mostly uninterrupted employment history and consistently work full-time, women's career paths are often discontinuous and they more often work part-time (Destatis 2008). The previous empirical findings as well as the theoretical predictions of both taste-based as well as statistical discrimination lead to the following baseline hypothesis:

Hypothesis 1: Female employees will be evaluated less favorably with respect to their promotability than male employees.

In line with the productivity-/preference-based arguments, gender differences in promotability assessments should be greatest, when the likelihood of women leaving the company is particularly high (Lazear and Rosen 1990). Women's probability to resign from their jobs is assumed to be greatest in times of expected maternity. In Germany, the average age of mothers at the time of birth is 31.7 years[4] (Destatis 2015a). Career interruptions were found to drive the "family gap in career progression", as they decrease mothers' work experience and tenure but increase their likelihood to be employed under a part-time contract (Kunze 2015). This tendency is also reflected in the gender pay gap that holds even among highly educated employees with children (Grund 2015). Since in case of incomplete contracts, employers make decisions based on group averages, I argue that the gender gap in promotability appraisals will be most pronounced in younger age cohorts, and that this gap closes afterwards. From this, the following hypothesis is derived:

Hypothesis 2: Female employees will be evaluated less favorably with respect to their promotability than male employees, in particular at around the age 30.

In their meta-analysis on gender differences in performance and rewards, Joshi et al. (2015) emphasize the moderating impact of occupation-, industry- and job-level characteristics of an employee's workplace. As an example, an employee's job level moderates promotion probabilities in the sense that women's promotion opportunities are worse in medium pay grades but better than men's in senior ranks, i.e., above clerical ceilings (Spilerman and Petersen 1999). Similarly, when focusing on middle managers rather than lower-level occupations, women are promoted faster than their male counterparts (Tsui and Gutek 1984). These findings suggest that the net disadvantages found for female employees in lower organizational levels are less prevalent or even non-

---

[4] In more detail, the average age of German mothers is 30.5 years for the first child, 32.3 years for the second child, and 33.5 years for the third child (Destatis 2015a).

existent in higher hierarchical levels. In other words, there seem to be fewer biases from role stereotyping as soon as women have already advanced to a higher managerial level (Petersen and Saporta 2004). Extending the theory of statistical discrimination and the assumed differences between men's and women's preferences helps to explain why the employer's assessment behavior differs with regard to an employee's current hierarchical job level. Information asymmetries between female employees and their superiors are reduced, once women have advanced to a managerial position. Employers do no longer have to base their decisions on average group values that typically hold for female employees. More specifically, women's occupation of (or self-selection into) managerial roles already signals their labor-market preferences. Moreover, these preferences as well as the typically higher age are indicators for a lower risk of leaving the company in the subsequent years. In a similar vein, Spilerman and Petersen (1999) note that women in higher corporate ranks are in a better position than female employees in lower salary grades to afford options that reduce their household responsibilities (for example the employment of housekeepers or babysitters). Apart from that, a prior advancement signals employees' effort levels and achievements. Hence, women who have been promoted before and managed to rise to the upper levels despite the assumed biases in the non-managerial levels, tend to be the better-performing employees. Thus, men and women in higher organizational levels can no longer be expected to differ significantly in their educational attainment or leadership skills (Lazear and Rosen 1990; Tsui and Gutek 1984). In their meta-analysis, Koch et al. (2015) find that the gender-role congruity bias will decrease, if additional information clearly indicates the women's high competence. Consequently, in higher managerial levels there should be no room left for statistical discrimination. Apart from that, the resource-based view contends that women in managerial positions are a particularly "valuable" firm resource, since female leaders are considered to be rare and can, thus, lead to a sustained competitive advantage (Barney 1991). The inclusion of this resource may not only bring in different perspectives or decision-making approaches (Beekun et al. 2010), but also satisfies the external pressure from various stakeholder groups, such as politicians or investors (Lee and James 2007). In the European Union, for example, there is an ongoing debate on introducing a female quota on corporate boards of listed companies, which is, however, not yet legally binding (Catalyst 2014). Taken together, it is reasonable to hypothesize that women in upper hierarchical levels are no longer disadvantaged, but may even be favored with respect to their promotability evaluations:

Hypothesis 3: In managerial levels of the job hierarchy, female employees will be evaluated more favorably with respect to their promotability than male employees.

Summarizing the theoretical and empirical literature of gender differences in career developments and promotion rates in particular, indicates a lack of appropriate studies on employees' promotability appraisals. I address this paucity using unique personnel data drawn from a German company's formal evaluation process. Hereby, the present study tackles the limitations of previous research. Rather than using survey data (Harris et al. 2006; Hoobler et al. 2009; Thacker and Wayne 1995) or subjective expectations as outcome variables (Blau and DeVaro 2007), the dataset includes binding promotability evaluations that have real consequences for the employees' future careers. As an example, Blau and DeVaro (2007) use survey data from the mid-1990s, in which U.S. managers were asked if the most recently hired employee has been promoted or is expected to be promoted within the next five years. Even though they find that female workers have lower chances of being promoted than male employees, this effect might even be underestimated due to the self-reported character of the outcome variable. Moreover, my data are not restricted to a specific age cohort (Blau and DeVaro 2007; Cobb-Clark 2001) or hierarchical level (Tsui and Gutek 1984), but rather encompass the employee population as a whole over a time period of three consecutive years. This study further enriches the previous literature on actual promotion rates, as I can control for a performance-related measure (as used in Cannings and Montmarquette 1991) as well as further demographic and job-specific characteristics of employees and their direct supervisors.

3   Dataset and Descriptive Statistics

In order to examine gender gaps in employees' promotability assessments, I draw on field data collected in a large global German company operating in the metal and electrical industry. The company's formal evaluation process requires supervisors to evaluate both their subordinates' *past* performance (skills) and their *future* (i.e., medium-term) career potential (promotability) in a face-to-face interview. The family company pursues a "promoting from within" policy[5] and ascribes a binding character to the promotability assessments: Evaluation scores are electronically recorded and serve to build a talent pool out of the best-rated employees that is tapped, as soon as vacancies need to be filled.

---

[5] According to company insiders, 70 percent of managerial positions are, on average, filled internally.

Employees are aware of the fact that above-average evaluations actually imply promotions within the next two to three years.

The introduction of the company's formal rating process in 2012 allows drawing on cross-sectional data pooled across three consecutive years (2012, 2013, and 2014). In order to account for unobservable person- and job-related differences between employees with and without leadership responsibilities (such as the educational attainment or the organization of careers along the hierarchy), I separate between 4,297 employee-year observations at the *non-managerial* and 880 employee-year observations at the *managerial* hierarchical level.[6] Representative of the German metal and electrical industry, the dataset is male-dominated: On average, 25.11 percent of the employees without any leadership responsibilities are women, while only 8.64 percent of the employees with leader responsibility are female.

3.1 Promotability Evaluations

Employees' promotability is evaluated on a 5-point Likert scale. An evaluation of "1" means that the employee is not able to meet the requirements of the current position, while category "2" adds that requirements could be met with sufficient training. The intermediate rank "3" reflects a condition in which the employee is described as a high-performer who will also be able to meet the requirements in the following year. Still, the employee will only be classified as promotable, if he or she is evaluated with either "4" ("Employee has the potential for taking more responsibility in the medium-term future") or "5" ("Employee has a great potential for medium-term career moves"). While "4" indicates the prospect of getting more responsibilities (for example in a project) and, thus, a higher remuneration, "5" denotes the prospect for career moves to the next higher hierarchy. Consequently, I use two alternative variables to define employees' future career success. The first dependent variable is the employee's promotability score (*SCORES*), which ranges between 1 and 5. Since employees are only likely to receive a medium-term promotion when being evaluated with either "4" or "5", the likelihood of being evaluated as *PROMOTABLE* (1=Getting an above-average evaluation of 4 or 5, 0=Otherwise) is used as an alternative dependent variable.

---

[6] In more detail, over the observation period of three years there are 1,805 (401) distinct employees working at the non-managerial (managerial) level. Nevertheless, I refrain from treating the observations as panel data due to several limitations: First, no detailed information on career interruptions (for example, due to maternity, parental or educational leaves, sabbaticals or actual dismissals) is available. Second, even if I balanced the sample and analyzed only those employees who "survive" in the dataset over the entire three years, I might unintentionally conduct a positive selection of the better-performing employees.
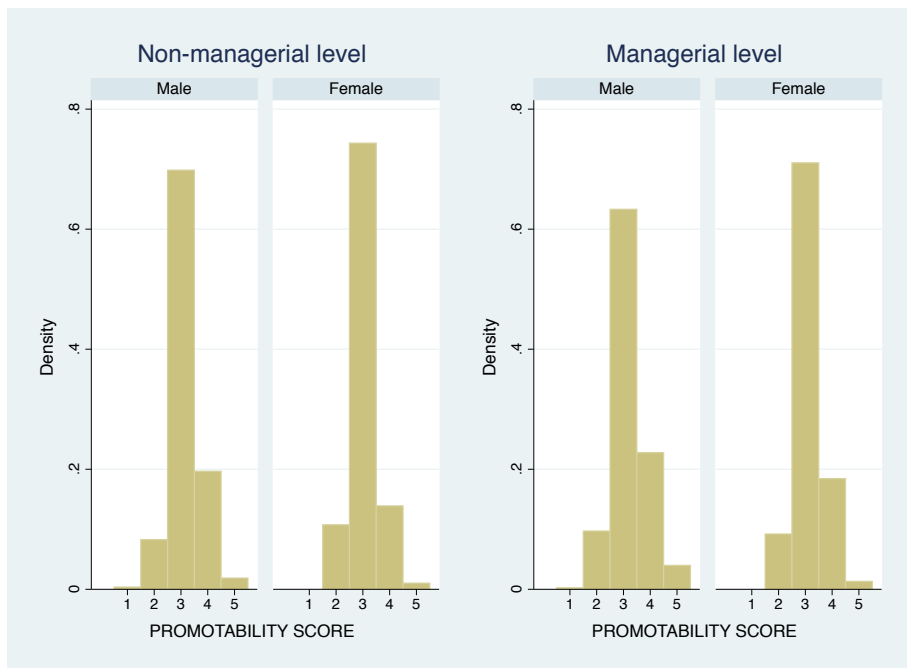
Notes. Promotability scores range between 1="Not able to meet the requirements of current position" and 5="Great potential for medium-term career moves". Distributions refer to 4,297 (880) employee-year observations at the non-managerial (managerial) level.

Figure 1 illustrates the frequency distributions of employees' promotability assessment scores at the non-managerial (left panel) and managerial (right panel) organizational levels by gender. At both hierarchical levels the distributions clearly indicate a tendency towards evaluating employees according to the intermediate category "3". The lowest category "1" is the least represented evaluation score (12 observations at the non-managerial and 2 observations at the managerial level), followed by category "5" (71 observations at the non-managerial and 33 observations at the managerial level). At the non-managerial level, a descriptive Wilcoxon-Mann-Whitney test reveals that there is a statistically significant difference between the underlying distributions of men's and women's promotability evaluation scores: On average, the male group achieves higher scores than the female group (z=4.765, p<0.001). Only in 19.87 percent of the observations are employees without leadership responsibilities evaluated as being promotable. While 21.54 percent of the male employees receive an outstanding promotability evaluation ("4" or "5"), significantly fewer women (14.92 percent) are qualified as promotable ($\chi^2(1)$=22.197, p<0.001). At the managerial level, the difference in men's and women's distributions of their promotability scores is strong, too, but not statistically significant (z=1.046, p=0.2957). On average, 26.14 percent of all managerial employees receive an above-average evaluation. Even though 26.74 percent of the male and only 19.74 percent of the

female employees are classified as promotable, Chi^2-Tests reveal no statistically significant difference ($\chi^2(1)$=1.765, p=0.184).

In order to draw sound conclusions concerning the preliminary tendencies towards gender differences in promotability evaluations, further person- and job-related variables will be introduced in the following. Descriptive statistics of these covariates for both hierarchical levels are illustrated in Table 2.[7]

3.2 Employee Characteristics

Apart from gender, laboratory as well as field studies point to further potential sources of biased evaluations: Researchers studied the impact of ratees' demographic characteristics that are supposed to be job-irrelevant, such as race[8] (Elvira and Town 2001; Giuliano et al. 2011; Greenhaus et al. 1990; Kraiger and Ford 1985), age (Cox and Nkomo 1992; Ferris et al. 1985), and appearance (Bento et al. 2012; Czajka and DeNisi 1988; Galin and Benolie 1990; Heilman and Stopeck 1985), but also job-relevant criteria, such as employees' tenure, work experience, or past performance (De Pater et al. 2009; Jawahar and Ferris 2011).

In non-managerial levels, employees are on average in their early forties (min=20, max=65). Prior research suggests age discrimination in promotability among very old cohorts (Ferris et al., 1985), but also among very young employees (Cox and Nkomo 1992). As a consequence, I assume an inverse u-shaped rather than a linear relationship between age and promotability scores. Moreover, with regard to the theoretical assumptions, gender discrimination is likely to vary over the lifecycle: Put differently, women's promotability scores should be particularly unfavorable during expected times of maternity leaves, i.e., around the age of 30. Non-managerial employees have worked within the company for about 14 years and are paid according to pay grade 10. In the sample, pay grades range between 3 and 15, while 15 represents a non-tariff salary. Pay grades are representative of other German companies that operate in the metal and electrical industry and, thus, comply with a collective bargaining agreement (ERA-TV).[9]

---

[7] See Table 6 and Table 7 in the Appendix for the correlation matrix of the variables included in the final models for the managerial and non-managerial levels, respectively.

[8] In my data, only 2 percent of the observed employees and less than 1 percent of the raters have a non-German background. Since there is almost no variability in the ethnic background of employees and raters, "origin" is not included as a demographic variable in the analyses.

[9] In Germany, the ERA-TV is guiding (though not necessarily binding) for more than 6,900 companies and 2.2 million employees working in the metal and electrical industry (Gesamtmetall 2015).

Mean comparison tests show that female employees are on average older and have also worked significantly longer for the company than their male counterparts. There is a lower fluctuation among female employees, which might indicate women's tendency to prefer a secure occupation within one company. On the one hand, a longer tenure could be associated with a broader professional network, an increase in company-specific know-how and might, thus, induce better evaluations. On the other hand, however, high-performers will be discovered and promoted even without a long tenure. On average, men are represented in higher pay grades than women. As an illustration, even at non-managerial hierarchical levels, more than 25 percent of all male employees are in pay grade 13 or higher, while only around 11 percent of female employees are paid based on the wage grades 13 onwards. In the dataset, an increase in pay grade (for example from 12 to 13) presents a promotion. Rises in sub-grades (for example from pay grade 12.1 to 12.2), however, are typically due to seniority. In the analyses, I will account for rises in "integer" pay grades only. Even though I lack information on employees' exact wages, I can approximate the hourly wages in each pay grade in order to demonstrate the monetary incentive of promotions (Metall NRW 2015). As illustrated in Table 1, wage increases from promotions are rather small in lower pay grades, whereas it makes a significant difference to be paid according to pay grade 13 or 14 instead of pay grade 12 or lower.

Table 1 Average Hourly Wages in ERA-Pay Grades

Pay grade	Average hourly wages in EUR (referring to monthly base salary)	Wage increase in EUR (referring to preceding pay grade)
3	15.00	0.17
4	15.24	0.24
5	15.56	0.32
6	15.96	0.40
7	16.46	0.50
8	17.32	0.86
9	18.72	1.40
10	20.57	1.85
11	23.06	2.49
12	25.09	2.03
13	28.64	3.55
14	33.48	4.84

Note. Average hourly wages refer to the remuneration tables published by the employers association of the metal and electrical industry in North Rhine-Westphalia, Germany (Metall NRW 2015).

Apart from that, there are significantly more female than male employees who work part-time. With regard to the employment status, the sample is, thus, representative of the German working population (Destatis 2015b). Although there is no information on the number of children, working part-time is a reliable indicator for employees' household responsibilities (Kunze 2015). Employees who self-select into part-time jobs may differ in their labor-market preferences, i.e., they might not even want to be promoted and, thus, prefer a job without leader responsibilities.

In comparison to non-managerial ranks, employees with (either dotted or solid line[10]) leadership responsibilities are on average four years older, show a longer mean tenure of 16 years and are on average in pay grade 13. In those higher organizational levels, female and male employees are more homogeneous, as they do not differ significantly in age or tenure. Even though men's and women's entry level is identical (pay grade 7), women are on average more represented in lower pay grades than their male colleagues: 75.25 percent of all men and only 65.79 percent of all women belong to pay scales 13, 14 or even non-tariff remuneration grades. These differences are statistically highly significant. Moreover, there are significantly more men working full-time than women – even though this gap is less pronounced than at the non-managerial level. In the subsequent multivariate analyses I will further account for the fact that an employee does not only receive an appraisal, but is also in the position to rate other employees. While this variable is neglectable in non-managerial hierarchies, in managerial tiers almost 60 percent of the employees are in such a dual role due to their leadership responsibilities – men significantly more often than women.

One of the most important job-related characteristics is an employee's productivity. Since differences in men's and women's promotability appraisals could simply be driven by gender differences in employees' skills, effort levels or achievements, it is crucial to control for performance. Due to the fact that it is usually not possible for a company to quantify the employees' individual contributions to the organizational success (for example with regard to clerical workers), previous studies have often failed to include a measure of an employee's true performance (Joshi et al. 2015). Alternatively, an employee's performance (rather than promotability) appraisal can be included as an

---

[10] The dataset comprises 591 employee-year observations with disciplinary leader responsibilities (i.e., solid-line reporting) and 289 employee-year observations with professional leader responsibilities (i.e., dotted-line reporting). Nevertheless, I will not further differentiate between both leadership styles in order to avoid losing too many observations.

appropriate performance indicator (Cannings and Montmarquette 1991; Prendergast 1999). In my data, the rater assesses employees' performance of the most recent business year on the basis of five different areas of competencies: Professional, methodological, personal, social, and leadership skills – will the latter be only measured for employees with leader responsibilities. Each competence area is measured on a 5-point Likert scale (from 1="Does not always come up to the requirements" to 5="Considerably exceeds the expectations"). Although the dataset provides the mean scores for each of the five skill areas, these are highly correlated. Consequently, the analyses will include the aggregated mean value of the five skill areas as an overall performance measure. The company uses this aggregated measure as an indicator to initiate feedback meetings and adequate training programs. Mean comparison tests (see Table 2) do not reveal any significant difference in the performance scores between men and women – independent of the hierarchical level.[11] Hence, I am confident to include a non-gender biased performance indicator. Preliminary analyses indicate a correlation between the performance score and the promotability score. Moreover, the difference between the promotability and the performance scores are statistically significantly greater for men (mean=0.11) than for women (mean=0.01). In other words, even when taking formal performance scores into account, men's promotability evaluations seem to be more positively biased than those of female employees – at least at the non-managerial level.

---

[11] See Figure 5 in the Appendix for the kernel density plots of the mean performance score by gender at the non-managerial and managerial level, respectively. Additional Kolmogorov-Smirnov tests for both hierarchical levels reveal that the mean performance scores across gender have the same distribution.

Table 2 Descriptive Statistics of Employees' Characteristics

Variable & definition	Total mean (sd)	Male employees mean (sd)	min	max	Female employees mean (sd)	min	max	Mean comparison t	df
AGE									
Age in years									
Non-managerial	41.68 (10.80)	41.05 (10.91)	20	65	43.57 (10.24)	21	65	-6.66***	4,295
Managerial	45.52 (8.15)	45.55 (8.15)	25	63	45.20 (8.27)	26	59	0.36	878
TENURE									
Tenure in years									
Non-managerial	13.66 (10.11)	12.71 (10.11)	0	44	16.48 (9.57)	0	35	-10.74***	4,295
Managerial	16.19 (10.14)	16.21 (10.11)	1	46	16.05 (10.56)	1	38	0.13	878
PAYGRADE									
Pay grade (15=Non-tariff)									
Non-managerial	10.16 (2.80)	10.52 (2.64)	3	15	9.10 (2.98)	3	15	14.72***	4,295
Managerial	13.43 (2.16)	13.49 (2.13)	7	15	12.78 (2.42)	7	15	2.69***	878
PARTTIME									
1=Part-time, 0=Otherwise									
Non-managerial	0.01	0.05	0	1	0.24	0	1	-18.64***	4,295
Managerial	0.06	0.06	0	1	0.12	0	1	-1.99**	878
RATER									
1=Employee also rater, 0=Otherwise									
Non-managerial	0.002	0.003	0	1	0.002	0	1	0.68	4,295
Managerial	0.59	0.60	0	1	0.46	0	1	2.40**	878
PERF									
Mean performance score									
Non-managerial	3.03 (0.24)	3.03 (0.24)	1.4	4.3	3.04 (0.25)	2.1	4.2	-0.78	4,295
Managerial	3.14 (0.25)	3.14 (0.26)	2.2	4.3	3.14 (0.24)	2.5	4	0.07	878

Note. Descriptive statistics refer to 3,218 (804) male and 1,079 (76) female employee-year observations at the non-managerial (managerial) level.

3.3 Rater Characteristics

Apart from employee-specific characteristics, information on the immediate supervisor will be included in the subsequent analyses. With regard to the moderating impact of the relative (dis)similarity of the gender-specific ratee-rater dyad, it has been shown that women anticipate more positive subjective evaluations as the probability of being

evaluated by a female instead of male superior increases (Maas and Torres-González 2011). Qualitative evidence from a U.S. survey of more than 1,000 employees supports the view that – when confronted with male supervisors – women prefer to self-select into occupations with objective (rather than subjective) reward systems, which are characterized by supposedly fewer discriminatory practices. Men's preferences, however, do not change with their superiors' gender. At the beginning of their careers, women are as likely as their male counterparts to strive for top management positions. Their motivation and ambitious behavior, however, disappears after five years of work experience, while men remain as goal-oriented as before. The main reason for women's significant drop in career ambition of more than 60 percent is the perceived lack of support and appreciation of their male superiors (Coffman and Neuenfeldt 2014). Given these outcomes, the influence of the rater's gender as well as the gender relation between ratee and rater on appraisal outcomes should not be neglected. Of those raters who evaluate non-managerial employees, only 5 percent are female (N=215). At the managerial level, 3.98 percent of the raters are women (N=35). This gap might already point to lower promotability opportunities for female employees. At the non-managerial level, female employees who are evaluated by female raters account for only 1.79 percent of the observations, while this percentage is even lower at the managerial level (0.45 percent). Moreover, raters who evaluate employees without leadership responsibilities are on average 47 years old, have worked for the company for 18 years and are remunerated according to pay grade 14. In higher hierarchical levels, raters' mean age is 49 years, mean tenure is 15 years, and average pay grade is 14. At both levels, female raters are significantly younger than male raters and, on average, in a lower pay grade. Finally, I control for raters who evaluate employees, but who are simultaneously assessed by their next higher supervisor. For those raters, promotability scores can be included to control for possible effects of envy.[12]

---

[12] See Table 8 in the Appendix for an overview of raters' descriptive statistics.

3.4 Methodological Approach

In order to identify the determinants of an employee's future career success, promotability evaluations will be empirically linked to the above-described set of employee and rater characteristics. The different scalings of the two alternative outcome variables require specific estimation techniques: While the ordinal scale of *SCORES* suggests the use of ordered probit regression analyses, probit regressions will be applied for the binary variable *PROMOTABLE*. In order to account for possible influences that might arise from being evaluated by either the same or a different rater as in the previous year, standard errors are clustered at dyad-level. The following equation is estimated separately for employees at non-managerial and managerial hierarchical levels:

$$Y_i = \beta_0 + \beta_1 * \text{EMPLOYEE}_i + \beta_2 * \text{RATER}_i + \beta_3 * \text{DEP}_i + \beta_4 * \text{UNIT}_i + \beta_5 * \text{YEAR}_i + \varepsilon_i .$$

EMPLOYEE and RATER represent vectors, which include the above-described demographic as well as job-specific variables for the employee and the immediate supervisor, respectively. Apart from that, further covariates, which control for the composition of the department (DEP) are included: Variables, such as the number of colleagues, teams' mean performance ratings, the share of female workers within the teams as well as teams' mean age and mean tenure, do not only account for the diversity among an employee's colleagues, but also for the diversity among the rater's subordinates. Furthermore, since the dataset comprises personnel data from a large company, it can be assumed that the seven functional units (finance, operations, HR, manufacturing, SCM, IT, and sales) vary in both their requirements, such as task complexity, and in unit-specific leadership styles. Hence, UNIT represents indicator variables for each section. The large majority of the observations stem from the manufacturing area, followed by IT and SCM. The majority of all female employees work in the highly male-dominated manufacturing and IT sections. The HR and sales departments both have a rather balanced share of male and female employees. These shares of male and female employees in the respective departments are similar to the occupational segregation by gender on the German labor market (Hausmann and Kleinert 2014).[13] Finally, the year of the rating (YEAR) is included.

---

[13] Separate regressions show that the adverse treatment of women is particularly pronounced in the manufacturing and IT sections, and less prevalent in gender-balanced units (HR and sales department). Nevertheless, due to the small number of observations in some of the functional units (see Table 9 in the Appendix), I refrain from drawing conclusions from separate analyses as suggested by Koch et al. (2015).

4 Econometric Analyses

Due to the assumed differences in the human capital endowments of employees working at the non-managerial versus managerial levels, the empirical results for both echelons will be presented separately.

4.1 Gender Differences at the Non-Managerial Level

Table 3 depicts the regression results of the gradually developed baseline model. In this baseline model, the five promotability assessment *SCORES* of employees at the non-managerial level serve as the dependent variable. Different model specifications subsequently include employees' demographic characteristics (1), employees' job-relevant information (2), information on an employee's corresponding rater (3), department-level controls (4), functional unit dummies (5), and, finally, year controls (6).[14]

According to the regression estimates, being female instead of male reduces the probability of receiving a higher promotability assessment score. This effect is statistically highly significant throughout all model specifications and even though I control for a wide range of covariates – the employee's past performance (*PERF*) in particular.

[14] See Table 10 in the Appendix for additional OLS regression results.

Table 3 Regression Estimates of Gender on Promotability Scores at the Non-Managerial Level

SCORES	(1)	(2)	(3)	(4)	(5)	(6)
EMPLOYEE						
FEMALE	**-0.189***	**-0.222***	**-0.206***	**-0.201***	**-0.193***	**-0.192***
	(0.051)	**(0.059)**	**(0.061)**	**(0.074)**	**(0.074)**	**(0.074)**
AGE	1.067***	0.518***	0.566***	0.438**	0.435**	0.431**
	(0.168)	(0.174)	(0.174)	(0.177)	(0.178)	(0.178)
AGE^2	-0.015***	-0.008***	-0.009***	-0.007***	-0.007***	-0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
TENURE	0.068**	0.048*	0.036	0.007	0.008	0.008
	(0.027)	(0.029)	(0.030)	(0.032)	(0.033)	(0.033)
PAYGRADE		0.009	0.013	0.013	0.020	0.020*
		(0.009)	(0.011)	(0.012)	(0.012)	(0.012)
PARTTIME		-0.122*	-0.140**	-0.117*	-0.113	-0.110
		(0.070)	(0.071)	(0.070)	(0.070)	(0.072)
PERF		3.094***	3.112***	3.696***	3.694***	3.694***
		(0.118)	(0.120)	(0.147)	(0.147)	(0.147)
RATER						
FEMALE			-0.036	-0.088	-0.096	-0.090
			(0.117)	(0.119)	(0.121)	(0.121)
AGE			-0.673*	-0.575	-0.557	-0.569
			(0.358)	(0.352)	(0.352)	(0.353)
AGE^2			0.006	0.005	0.005	0.005
			(0.004)	(0.004)	(0.004)	(0.004)
TENURE			0.081***	0.069**	0.067**	0.066**
			(0.028)	(0.031)	(0.031)	(0.031)
PAYGRADE			0.017	0.033*	0.047**	0.047**
			(0.020)	(0.020)	(0.021)	(0.021)
PARTTIME			-0.347	-0.269	-0.258	-0.251
			(0.222)	(0.221)	(0.223)	(0.224)
EMPLOYEE			0.062	0.037	0.053	0.072
			(0.080)	(0.079)	(0.080)	(0.083)
DEP CONTROLS				Yes	Yes	Yes
UNIT CONTROLS					Yes	Yes
YEAR CONTROLS						Yes
cut1	-1.080***	6.348***	5.173***	2.287**	3.042***	3.065***
Constant	(0.332)	(0.457)	(0.957)	(1.041)	(1.082)	(1.084)
cut2	0.371	8.455***	7.295***	4.480***	5.239***	5.264***
Constant	(0.337)	(0.476)	(0.967)	(1.045)	(1.089)	(1.091)
cut3	2.602***	11.210***	10.062***	7.276***	8.043***	8.068***
Constant	(0.338)	(0.492)	(0.972)	(1.049)	(1.094)	(1.096)
cut4	3.921***	12.827***	11.687***	8.923***	9.694***	9.719***
Constant	(0.346)	(0.511)	(0.979)	(1.055)	(1.101)	(1.103)
Observations	4,297	4,297	4,297	4,297	4,297	4,297
Number of dyads	2,228	2,228	2,228	2,228	2,228	2,228
Pseudo R^2	0.2019	0.2052	0.2113	0.2159	0.2179	0.2181

Notes. Table reports coefficients after ordered probit regressions with SCORES as the dependent variable. Age, age^2, and tenure multiplied by ten for ease of interpretation. Robust standard errors clustered at dyad-level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Average marginal effects (AMEs) after ordered probit regressions provide more detailed information on how gender influences the probability of receiving each of the five assessment categories. Figure 2 illustrates the differences between otherwise equal male and female employees in their predicted margins of receiving each of the five promotability scores. While there are no economically relevant gender differences in receiving the worst rating ("1"), female employees are, on average, more likely than their male counterparts to receive a low ("2") or mediocre ("3") evaluation. This tendency switches as soon as above-average scores are considered: Women are less likely than men to get an outstanding evaluation ("4"). Gender differences in receiving the best promotability evaluation ("5") are small, yet still statistically significantly disadvantageous for women.

Fig. 2 Average Marginal Effects of gender on promotability Scores
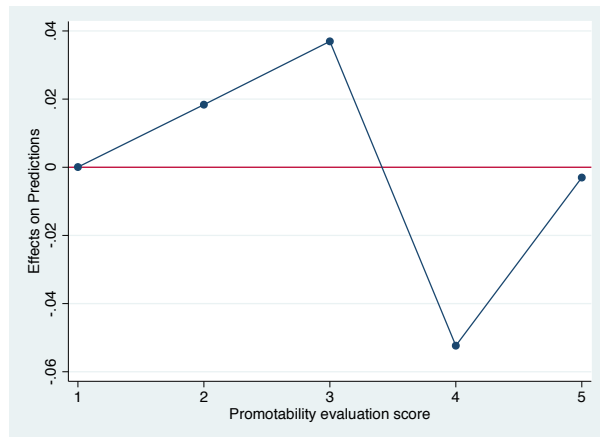
Note. Blue dots represent the average marginal effects of being female instead of male on the distinct promotability score. Promotability scores range between 1="Not able to meet the requirements of current position" and 5="Great potential for medium-term career moves".

In order to come to more detailed conclusions concerning gender gaps in actual promotability classifications, Table 4 uses the probability of being classified as *PROMOTABLE* (score "4" or "5" versus lower scores) as the dependent variable. While column 1 provides the coefficients of the probit regression, AMEs are reported in column 2. The average probability for a female employee to receive an outstanding evaluation and, thus, to be qualified as promotable is 5.1 percentage points lower than for otherwise comparable men – although the overall probability of being classified as promotable is only 19.84 percent. Summarizing these baseline findings so far, one cannot reject that female employees are disadvantaged in that they receive less favorable promotability evaluations than equally qualified men. Hypothesis 1, thus, cannot be rejected.

**Table 4 Regression Estimates of Gender on Probability of Promotable Evaluation at the Non-Managerial Level**

PROMOTABLE	COEFFICIENT	AME
EMPLOYEE		
FEMALE	**-0.245****	**-0.051****
	(0.100)	**(0.021)**
AGE	0.672**	-0.005
	(0.264)	(0.001)
AGE^2	-0.116***	./.
	(0.033)	
TENURE	-0.147***	-0.030***
	(0.048)	(0.010)
PAYGRADE	0.069***	0.014***
	(0.017)	(0.003)
PARTTIME	-0.299***	-0.062***
	(0.115)	(0.024)
PERF	3.633***	0.751***
	(0.209)	(0.038)
RATER		
FEMALE	-0.030	-0.006
	(0.150)	(0.031)
AGE	-1.419***	-0.001
	(0.435)	(0.001)
AGE^2	0.147***	./.
	(0.048)	
TENURE	-0.007	-0.002
	(0.039)	(0.008)
PAYGRADE	0.096***	0.020***
	(0.031)	(0.006)
PARTTIME	0.030	0.006
	(0.296)	(0.061)
EMPLOYEE	0.014	0.003
	(0.105)	(0.022)
CONTROLS	INCL.	
Constant	-6.570***	./.
	(1.367)	
Observations	4,297	4,297
Number of dyads	2,228	2,228
R^2	0.2587	./.

Notes. Table reports coefficients (column 1) and average marginal effects (column 2) after probit regressions with PROMOTABLE as the dependent variable. Age, age^2, and tenure multiplied by ten for ease of interpretation. No separate marginal effect for age^2 can be estimated. Robust standard errors clustered at dyad-level in parentheses. Estimations include department, unit, and year controls. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Because of the model's non-linearity, the detrimental female effect is likely to vary depending on the values of the remaining predictors. A further look at the effects of the covariates in Table 4 indicates that employees receive more favorable promotability assessments with increasing age. This effect, however, is non-linear and diminishes after an optimal age of around 33.5 years. Due to the fact that age^2 cannot change regardless of the values of its linear term, one cannot estimate a separate marginal effect for the

interaction. Hence, more detailed analyses are necessary. With regard to the theoretically derived differences in men's and women's risks of leaving the company during certain periods of life, the age effect is likely to show a distinctive trend for female and for male employees. The left panel of Figure 3 illustrates the predicted margins of being classified as *PROMOTABLE* for female and male employees taking different age levels into account. The graph indicates that even though the inverse u-shaped trend is more pronounced for men than for women, the predicted probabilities of being classified as promotable are largest during younger ages and decrease afterwards for both genders. Hence, there exists a relatively short period in which promotion opportunities are greatest. The right panel of Figure 3 illustrates that the gender gap in the probability of being *PROMOTABLE* is statistically significant across all age levels, but largest for men and women aged 20 to 35. At the age of 30, female employees have a 6.2 percentage points lower promotion probability than their male colleagues. This gender gap declines after the age of 30 but remains statistically significant. This leads me to conclude that hypothesis 2 cannot be rejected.

Fig. 3 Gender Differences in Promotable Evaluation at the Non-Managerial Level by Age
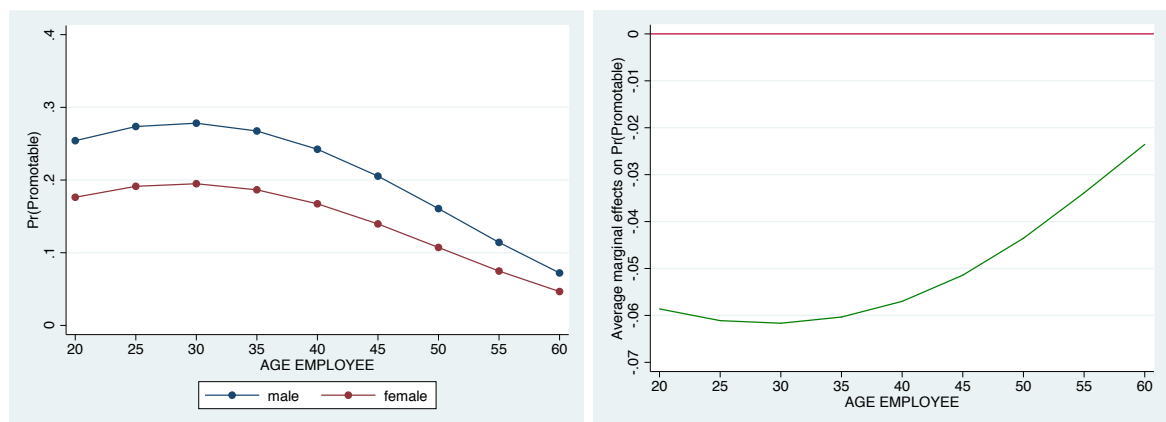
Notes. Predicted probabilities of men and women (left-hand side) and average marginal effects of gender (right-hand side) on being classified as PROMOTABLE on y-axis, respectively. Age levels on x-axes.

Apart from these moderating age effects, Table 4 depicts that the probability of receiving an outstanding promotability evaluation is influenced by employees' working hours: Being employed part-time instead of full-time decreases the likelihood of being evaluated as *PROMOTABLE* by 6.2 percentage points. Women are significantly more likely to work under a part-time contract: Between the ages of 20 to 29, almost all female employees work full-time. This ratio drops to around 70 percent for female employees between 30 and 49 years. Above the age of 50 years, the percentage of female full-time workers rises again. This distribution clearly indicates that women self-select into part-time jobs after

returning to the company due to maternity or parental leaves. Furthermore, it might be an indicator for employees' preferences, i.e., people working part-time may not even aim to be promoted. Predicted margins in Figure 4 show that the promotability probability decreases for both genders as soon as part-time contracts are observed. Furthermore, men's predicted probabilities are higher than women's likelihood to receive a *PROMOTABLE* assessment – even slightly (but not significantly) when comparing part-time working men with full-time working women. While part-time working men have a 5.84 percentage points lower probability of receiving a *PROMOTABLE* evaluation than full-time working men, women working part-time have a 4.97 percentage points lower promotability probability than women working full-time. However, the differences between part-time male and female employees are not statistically significant, while female employees working full-time have a 4.92 percentage points lower probability than their full-time male counterparts. Restricting the sample to full-time employees only does not change the baseline results – even when taking different age levels into account.[15] Hence, even among employees working full-time, who signal their labor-market preferences, women remain disadvantaged, i.e., they receive less favorable promotability evaluations.

Fig. 4 Gender Differences in Promotable Evaluation at the Non-Managerial Level by Working Hours
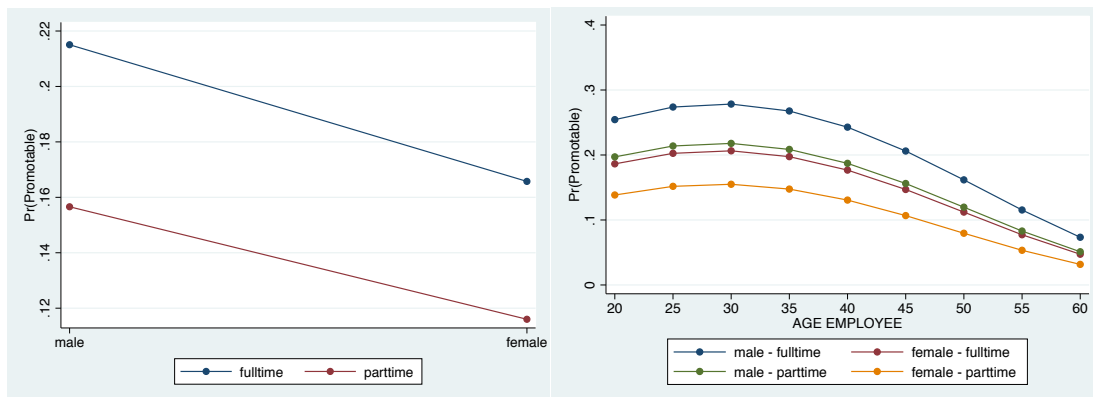
Notes. Predicted probabilities of full-time and part-time men and women, respectively, on being classified as PROMOTABLE on left-hand side. Different age levels added at the right-hand side.

Moreover, Table 4 reveals that tenure has a small yet negative effect on employees' promotability. On average, an increase in tenure by one year leads to a 0.3 percentage points lower probability of being evaluated as *PROMOTABLE* (3 percentage points for ten years). This finding supports the argument that high-performing talents will stand out

---

[15] See Figure 6 in the Appendix for an illustration of the predicted probabilities and AMEs for the sample restricted to full-time employees.

immediately. The gender gap is most pronounced shortly after joining the company (5.7 percentage points) and decreases with longer tenure. Nevertheless, the gender gap is statistically significant at all levels of an employee's tenure (see Figure 7 in the Appendix). Furthermore, pay grades enter positively into the equation indicating enhanced promotion opportunities in higher pay grades. An increase in pay grade by one unit leads to a 1.4 percentage points higher probability of being classified as *PROMOTABLE*. The effect, however, is relatively small and might be ascribed to the fact that a rise in pay grade should usually be the consequence (rather than the antecedent) of a promotion. Nevertheless, the upward trend applies for both male and female employees, although the men's lies above women's wage level (see Figure 8 in the Appendix). These gender differences are statistically significant across all pay grades. Nevertheless, they are more pronounced in higher than in lower pay grades: While women have a 3.1 percentage points lower probability of being classified as *PROMOTABLE* in pay grade 3, this disadvantage doubles to a 6.1 lower probability in non-tariff contracts (=pay grade 15). The final and expectedly most decisive influencing variable is the employee's mean performance score. A 0.1-unit increase leads to a 7.51 percentage points higher probability of receiving an outstanding promotability score. The gender gap is largest at a mean performance score of 3.3, whereas there are no significant differences between men's and women's promotability evaluations when the performance score lies below 2.4 or above 4. This, however, applies to only 1 percent of the observations (see Figure 1).

With regard to the raters' characteristics, Table 4 provides no reason to assume that male and female raters differ in their evaluation behavior. Nevertheless, there are only 77 female and 138 male employees who are evaluated by women. Hence, the small number of observations leads to large confidence intervals and, thus, reduces the explanatory power of differentiating between the evaluation behavior of male and female raters. Still, the restriction of the dataset to those employees who are evaluated by a male rater can at least not rule out the observed gender differences in employees' promotability evaluations of the baseline regressions.[16] In more detail, female employees are almost 6 percentage points less likely than their male colleagues to receive a *PROMOTABLE* evaluation when being evaluated by a male supervisor (as compared to 5.1 percentage points in the full sample). This finding is in line with Koch et al. (2015) who indicate that male raters

---

[16] See Table 11 in the Appendix for the results of the ordered probit and probit regressions applied to the restricted sample.

exhibit a greater tendency towards gender-role stereotyping – in particular in male-dominated occupations. Table 4 further indicates that the impact of raters' age follows a u-shaped trend, i.e., raters seem to evaluate more leniently when being relatively young, yet become stricter with increasing age. With regard to the moderating role of raters' age, Figure 9 in the Appendix illustrates that female employees are around 5 percentage points more likely to be disadvantaged by raters between 40 and 60. Nevertheless, only a small number of very young supervisors is observed and the majority of the raters are between 40 and 60 years old. Moreover, raters in higher pay grades evaluate employees on average more favorably. An increase in a rater's pay grade by one unit leads to a 2 percentage points higher probability of receiving the evaluation *PROMOTABLE*. Around 60 percent of the observed employees are evaluated by raters remunerated according to pay grade 13 or higher. Gender differences are largest in those higher pay grades (see Figure 10 in the Appendix): While female employees who are evaluated by raters in pay grade 12 have a 4.4 percentage points lower probability of receiving a *PROMOTABLE* evaluation than their male colleagues, the difference when looking at raters who receive a non-tariff salary is even more pronounced (5.4 percentage points lower probability). Finally, there is no significant impact of the fact that the raters themselves are evaluated by their next higher supervisors. Even if raters' own promotability scores are included, which restricts the sample to 3,311 employee-year observations, no statistically significant effect on employees' probability to receive an above-average evaluation can be identified.

4.2   Gender Differences at the Managerial Level

The results presented so far refer to employees who have not yet taken on managerial responsibilities. In order to disentangle the effects of statistical discrimination from taste-based discrimination, however, regressions need to be repeated for the 880 employee-year observations at the managerial level. Table 5 depicts the regression estimations using the promotability *SCORES* (column 1) as well as the *PROMOTABLE* dummy (column 2) as two alternative outcome variables. At first sight, the estimation results reveal no statistically significant gender differences. Nevertheless, even though the female coefficient is just insignificant, the effect remains negative and the size is comparable to the ones estimated for the non-managerial level (see Table 3 and Table 4). Only the standard errors are slightly higher, which is, in all likelihood, due to the low number of observations: At the managerial level, there are 76 women of whom only 15 received a *PROMOTABLE* evaluation. As the analyses do not merely rely on a sample selection, but

rather include the company's full employee population, I argue that the gender gap still persists at higher levels of the organizational hierarchy and should not be neglected. AMEs after probit regressions show that the average woman has a 6.9 percentage points lower probability than her male colleague to receive a *PROMOTABLE* evaluation (z=-1.30). More formally, hypothesis 3, which proposes a positive discrimination of female employees with leadership responsibilities, cannot be supported.

Table 5 Regression Estimates of Gender on Promotability Evaluation at the Managerial Level

	SCORES	PROMOTABLE
EMPLOYEE		
FEMALE	**-0.144**	**-0.315**
	(0.168)	**(0.243)**
AGE	0.123	1.649*
	(0.650)	(0.920)
AGE^2	-0.053	-0.255**
	(0.070)	(0.106)
TENURE	-0.018	-0.148
	(0.069)	(0.106)
PAYGRADE	0.054	0.012
	(0.047)	(0.053)
PARTTIME	0.023	0.056
	(0.166)	(0.238)
RATER	0.034	-0.020
	(0.132)	(0.174)
PERF	3.333***	3.339***
	(0.281)	(0.356)
CONTROLS	INCL.	
Observations	880	880
Number of dyads	474	474
Pseudo R^2	0.2320	0.3129

Notes. Table reports coefficients after ordered probit (column 1) and probit regressions (column 2) with SCORES or PROMOTABLE as the dependent variables, respectively. Cut points for ordered probit regression (column 1) not shown. Age, age^2, and tenure multiplied by ten for ease of interpretation. All estimations include rater, department, unit, and year controls. Robust standard errors clustered at dyad-level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

5 Discussion and Conclusions

Much has been said about the "invisible ceiling" that prevents women from reaching the organizations' upper echelons, although they progress at an equal pace as men at the beginnings of their careers. Prior research focuses on women's career disadvantages in personnel decisions, such as performance appraisals, promotions, and wages (Joshi et al., 2015). So far, however, the decisive role of formal promotability evaluation processes, which have a crucial impact on employees' advancement opportunities, has not been considered.

In this paper, I analyze whether female employees are systematically disadvantaged in the formal promotability assessments they receive from their (mainly male) supervisors. With regard to the non-managerial levels, findings confirm evidence of the existence of gender discrimination in promotability evaluations, even though a rich set of employee-, rater-, and work-specific covariates is included. Compared to men, women who do not hold a managing position have, on average, a 5.1 percentage points lower probability of receiving an evaluation that qualifies them as being promoted within the next two to three years. This effect is statistically significant even though I control for individual performance using a non-gender biased indicator.

As theoretically established in section 2, different concepts can explain the observed disadvantages. On the one hand, gender differences might be unwittingly caused by information asymmetries between the supervisor and the employee, which induces employers to refer to gender statistics in their evaluation processes (statistical discrimination). The fact that men and women differ in their attitudes towards competition or risk preferences is already well documented (Azmat and Petrongolo 2014; Niederle 2014) and might, thus, be reflected in both genders' preferences of being promoted. As an example, there might be incomplete contracts with regard to an employee's probability of leaving the labor market. This risk is perceived to be greater for women than for men – and in particular at around the age of 30 due to childbearing and -rearing. Although not all female employees will indeed leave the company, it is rational for the employer to refer to these group characteristics when assessing the subordinates' career potential. In other words, the supervisor will prefer to invest in male rather than female employees by giving men better evaluations irrespective of their performance or other job-related variables. Indeed, the observed gender gap is found to be largest around the age of 30 (6.2 percentage points) and to decrease gradually as employees age. I proxied employees' preferences by the inclusion of their working hours. Women are likely to self-select into part-time jobs when having greater responsibilities in the household sector. Furthermore, career interruptions or part-time contracts for mothers lead to lower levels of work experience and, thus, significantly decrease promotion probabilities (Kunze 2015). According to the results, women in part-time jobs receive lower promotability evaluations than women in full-time jobs. In all likelihood, these differences may indeed indicate different preferences of women working under part-time versus full-time contracts. However, when restricting the sample to full-time employees only, the significant gender

differences in employees' promotability assessments persist. Hence, employers' intentional taste against women (i.e., taste-based discrimination) cannot be excluded as a possible explanation for the observed gender gaps. Preference-driven biases in the promotability assessments that female employees receive from their male supervisors could, for instance, be evoked by misogynist stereotypes or prejudices that hold in traditional male domains, such as the metal and electrical industry. As an example, employers often do not see a match between women's and leaders' attitudes, as women are characterized as less competent in terms of being promoted. Even though an analysis of the gender-specific superior-subordinate dyads would have helped to further differentiate between taste-based and statistical discrimination, the rating behavior of male and female superiors could not be compared due to the low number of female raters. Although the mere fact that there are mainly male raters does at least not rule out the existence of taste-based discrimination, the analysis of managerial employees further helps to disentangle the effects of taste-based versus statistical discrimination. It is argued that superiors might deem the probability that women will leave the company in exchange for the household sector as a low-risk threat, once they have already climbed the corporate ladder. Prior promotions, hence, reduce information asymmetries, as they serve as a signal for women's labor-market preferences and future productivity (Petersen and Saporta 2004; Spilerman and Petersen 1999). The results, however, reveal that the disadvantageous treatment of female employees persists in the upper echelons of the company. As suggested by Hoobler et al.'s (2009) findings, managers may still perceive women's family-work conflict to be larger than it actually is. This bias is found to hold for male and female supervisors.

Even though the results support a systematic "mistreatment" of women in formalized evaluation processes, it remains challenging to differentiate between supply- and demand-side effects, as both explanations are not mutually exclusive and may simultaneously contribute to women's unfavorable evaluations. Although I proxy employees' preferences by differentiating between full-time and part-time as well non-managerial and managerial employees, women could still be less motivated about being promoted. A follow-up study with individual interviews of employees at different hierarchical levels would help to shed more light on supervisors' expectations as well as men's and women's actual promotion preferences.

A possible drawback of the study is that the observed promotability assessments cannot be accurately linked to actual promotions or wage gains attached to career advancements due

to the short observation period. The company introduced employee appraisals in 2012; Promotability assessments, however, refer to employees' medium-term future and the availability of vacant positions in the next two to three years. Due to the fact that the data is restricted to three subsequent years (2012-2014), medium-term promotion effects cannot be demonstrated empirically. Although I do find a statistically significant positive effect of above-average promotability assessments on employees' pay grades, this effect might not be causal.

As the findings refer to observations from a single company, the study benefits from a high level of internal validity: In contrast to prior studies on gender differences in promotion rates, which often base their analyses on survey data, I observe employees' demographic characteristics, job-related criteria, such as their performance as well as supervisors' actual behavior. Hence, I am confident that the findings on gender discrimination offer interesting insights and decision-making bases for companies relying on similar formal promotability evaluation processes. Both organizations and policy makers can learn from the findings of this study that affirmative action policies, such as the implementation of a female quota in managerial levels or even top management positions, will only be efficient, if women indeed wish to be promoted. Furthermore, in order to reduce unintentional discriminatory practices against female employees, supervisors should aim to decrease information asymmetries about women's promotion preferences. Since formal employee evaluations are used as a pivotal managerial decision-making tool in order to establish training programs and determine performance-related pay as well as promotions (see for example Cappelli and Conyon 2016; Cleveland et al. 1989; Landy and Farr 1983; Murphy and Cleveland 1995), it is crucial for employers to base their decisions on unbiased evaluations. Despite possible negative consequences for the employer, biases might seriously influence employees' organizational commitment, motivation, job satisfaction and, thus, performance – in particular when linked to one's future career success (Kuvaas 2006). That is why raters should be held accountable for their evaluation behavior and should be motivated to focus their appraisals on job-related rather than demographic employee characteristics in order to prevent gender discrimination in the appraisal process to become a "business as usual" practice (Powell and Graves 2003).

6 Appendix

Fig. 5 Kernel Density Estimate of Performance Evaluation by Gender and Managerial Level
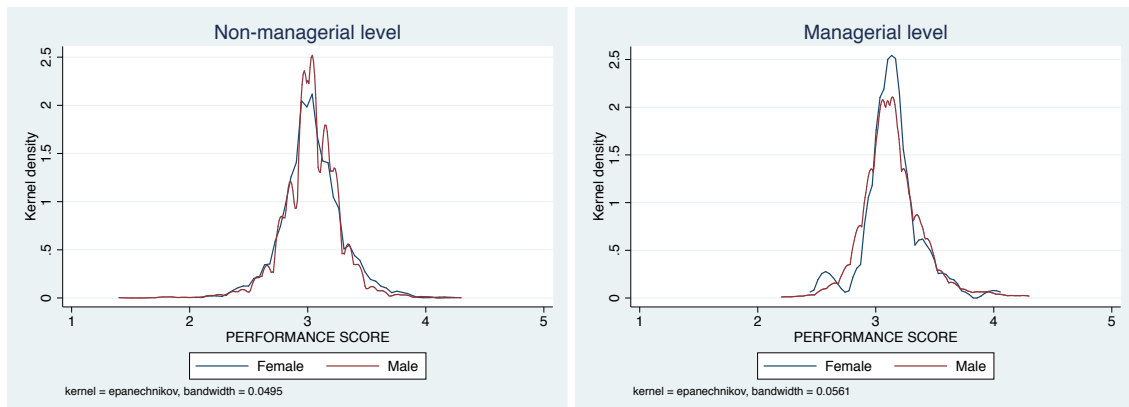
Note. Subsequent two-sample Kolmogorov-Smirnov tests for equality of men's and women's distribution functions do not reveal any statistically significant differences at both hierarchical levels, respectively.

Table 6 Correlation Matrix – Non-Managerial Level

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1 PROMOTABILITY SCORE	1																												
2 PROMOTABLE	0.83	1																											
3 FEMALE_EMPLOYEE	-0.07	-0.07	1																										
4 AGE_EMPLOYEE	-0.12	-0.19	0.1	1																									
5 TENURE_EMPLOYEE	-0.05	-0.14	0.16	0.62	1																								
6 PAYGRADE_EMPLOYEE	0.14	0.18	-0.22	0.01	-0.04	1																							
7 PARTTIME_EMPLOYEE	-0.03	-0.05	0.27	0.09	0.12	0	1																						
8 PERF_EMPLOYEE	0.53	0.37	0.01	-0.04	0	0.21	0.03	1																					
9 FEMALE_RATER	-0.03	0	0.06	-0.05	-0.04	0.04	0.02	-0.05	1																				
10 AGE_RATER	0.02	-0.02	-0.04	-0.02	0.01	-0.06	-0.02	0.08	-0.18	1																			
11 TENURE_RATER	0.02	-0.05	-0.1	0.01	0.13	-0.29	0	0	-0.15	0.48	1																		
12 PAYGRADE_RATER	0.07	0.11	0.07	-0.03	-0.08	0.48	0.02	0.12	-0.07	0.12	-0.27	1																	
13 PARTTIME_RATER	-0.02	0	-0.04	-0.05	-0.02	-0.01	-0.01	0	0.24	-0.12	-0.04	-0.19	1																
14 RATEREMPLOYEE	0	-0.01	-0.04	-0.02	-0.02	-0.04	0.02	-0.02	-0.07	0	0.09	-0.07	0.03	1															
15 EMPLOYEES	-0.04	-0.11	-0.07	-0.03	0.01	-0.49	-0.07	-0.07	-0.12	0.18	0.39	-0.28	-0.01	-0.04	1														
16 S_FEMALES	-0.03	-0.01	0.58	0.09	0.11	-0.12	0.21	0.03	0.1	-0.1	-0.14	0.07	-0.05	-0.02	-0.1	1													
17 AGE_MEAN	-0.06	-0.07	0.13	0.43	0.31	-0.12	0.08	-0.01	-0.12	-0.01	-0.01	-0.04	-0.1	-0.08	-0.07	0.19	1												
18 TENURE_MEAN	-0.01	-0.04	0.1	0.27	0.5	-0.19	0.07	-0.01	0.07	0.29	-0.16	-0.04	-0.04	0.06	0.18	0.61	1												
19 PERF_MEAN	0.23	0.16	0.05	-0.01	0	0.2	0.06	0.63	-0.05	0.13	-0.01	0.2	0.01	-0.03	-0.11	0.06	-0.01	0	1										
20 FINANCES	0	-0.02	0.04	0.03	0.02	0.13	0.11	0.07	0.02	-0.01	-0.05	0.12	-0.02	0.05	-0.12	0.06	0.06	0.02	0.11	1									
21 OPERATIONS	0	0	0	0.05	0.03	-0.02	0.08	0	0.01	-0.05	-0.08	-0.12	-0.02	0.02	0	0.01	0.09	0.03	-0.02	-0.04	1								
22 HR	0.02	0.02	0.15	0	-0.01	0.09	0.03	0.09	0.11	0.07	-0.08	0.1	0	-0.04	-0.13	0.23	0	-0.03	0.15	-0.03	-0.04	1							
23 MANUFACTURING	-0.05	-0.06	-0.08	0	0.05	-0.41	-0.1	-0.14	-0.08	0.1	0.36	-0.26	0.06	0.07	0.46	-0.12	0	0.15	-0.2	-0.19	-0.27	-0.17	1						
24 SCM	0.01	0	0.06	0.04	0.04	0.02	0.01	-0.01	0.07	-0.1	-0.1	-0.09	-0.04	-0.11	-0.14	0.1	0.07	0.05	-0.05	-0.06	-0.09	-0.06	-0.41	1					
25 IT	0.03	0.06	-0.06	-0.05	-0.1	0.38	-0.01	0.11	-0.07	-0.04	-0.21	0.35	-0.05	0.02	-0.3	-0.12	-0.09	-0.2	0.16	-0.08	-0.12	-0.08	-0.55	-0.18	1				
26 SALES	0.02	0.03	0.12	-0.06	-0.06	0.12	0.06	0.06	0.15	0	-0.15	0.08	0.06	-0.09	-0.13	0.2	-0.11	-0.12	0.09	-0.03	-0.04	-0.02	-0.17	-0.06	-0.08	1			
27 2012	-0.01	0	0	-0.03	-0.01	-0.04	-0.09	-0.1	-0.03	-0.06	0	0	-0.05	-0.24	0	0.01	-0.06	-0.02	-0.16	-0.12	-0.17	0	0.06	0.11	-0.01	-0.01	1		
28 2013	-0.01	-0.01	0.01	0	0	0.02	-0.09	-0.01	0.03	0.04	0.01	0.07	-0.07	0.09	0	0.01	0.01	0.01	-0.01	-0.12	-0.17	0.01	-0.02	0.18	0.02	0.01	-0.47	1	
29 2014	0.02	0.01	-0.01	0.02	0.01	0.02	0.17	0.1	0	0.02	-0.01	-0.06	0.12	0.15	0	-0.03	0.05	0.01	0.16	0.23	0.32	-0.01	-0.03	-0.28	-0.01	0	-0.52	-0.51	1

Note. Correlations refer to 4,297 employee-year observations at the non-managerial level.

Table 7 Correlation Matrix – Managerial Level

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1 PROMOTABILITY SCORE	1																													
2 PROMOTABLE	0.52	1																												
3 FEMALE_EMPLOYEE	-0.04	-0.02	1																											
4 AGE_EMPLOYEE	-0.19	-0.24	-0.01	1																										
5 TENURE_EMPLOYEE	-0.19	-0.19	0	0.52	1																									
6 PAYGRADE_EMPLOYEE	0.13	0.07	-0.09	0.19	-0.16	1																								
7 PARTTIME_EMPLOYEE	0	-0.05	0.07	0.03	-0.01	0.07	1																							
8 RATER	0.1	0.05	-0.08	0.19	-0.07	0.56	0.13	1																						
9 PERF_EMPLOYEE	0.48	0.35	0	0.05	-0.03	0.24	0.04	0.11	1																					
10 FEMALE_RATER	-0.08	-0.05	0.02	0	0.02	0.01	-0.03	-0.09	-0.02	1																				
11 AGE_RATER	0.05	-0.02	0.01	0.16	0.04	0.15	0.17	0.07	0.16	-0.08	1																			
12 TENURE_RATER	-0.1	-0.1	-0.02	-0.08	0.24	-0.37	-0.1	-0.36	-0.15	0.11	0.25	1																		
13 PAYGRADE_RATER	0.03	0.04	-0.02	0.06	-0.07	0.67	0.04	0.41	0.11	-0.15	0.22	-0.21	1																	
14 PARTTIME_RATER	-0.02	0	0.03	-0.04	0	-0.04	-0.02	-0.09	0.03	0.29	-0.14	-0.01	-0.22	1																
15 RATEREMPLOYEE	-0.05	0.01	0	-0.11	0.12	-0.19	0.06	-0.14	-0.19	0.05	-0.18	0.25	-0.13	0.03	1															
16 EMPLOYEES	-0.14	-0.03	0	-0.07	0.15	-0.56	0.12	-0.42	-0.03	-0.02	0.06	0.32	-0.38	0.03	0.14	1														
17 S_FEMALES	0.03	-0.02	0.45	0.04	-0.04	0.11	0.01	0	0.08	0.06	0.05	-0.05	0.07	-0.04	-0.06	0.02	1													
18 AGE_MEAN	-0.07	-0.12	0.07	0.42	0.19	0.22	0.1	0.29	-0.07	-0.07	0.27	-0.08	0.13	-0.14	-0.2	-0.18	0.1	1												
19 TENURE_MEAN	-0.16	-0.11	0.04	0.2	0.53	-0.08	0.03	-0.01	-0.08	-0.02	0.06	0.3	0.02	-0.06	0.13	-0.01	0.04	0.52	1											
20 PERF_MEAN	0.25	0.22	0.02	0.07	-0.15	0.35	0.08	0.27	0.6	-0.14	0.27	-0.26	0.23	-0.04	-0.34	-0.18	0.04	0.16	-0.15	1										
21 FINANCES	0.04	-0.05	0.03	-0.01	-0.09	0.16	-0.06	0.05	-0.04	0.15	-0.05	0	0.09	-0.02	0	-0.11	0.09	0.07	-0.06	-0.05	1									
22 OPERATIONS	0.04	-0.02	0.01	-0.03	-0.07	-0.18	0.01	-0.01	-0.07	0.01	-0.03	0.03	-0.19	-0.02	-0.05	0.04	-0.01	0	-0.08	-0.07	-0.05	1								
23 HR	-0.06	-0.11	0.21	0.04	0.06	0.04	0	-0.07	0.01	0.04	0.15	0.05	0.09	0.04	-0.11	-0.13	0.31	0.1	0.19	0.02	-0.06	-0.06	1							
24 MANUFACTURING	-0.1	-0.06	-0.04	0.03	-0.1	-0.29	0.2	-0.1	-0.04	-0.01	0.01	0.26	-0.09	0.06	0.23	0.42	-0.12	0.01	0.28	-0.18	-0.19	-0.18	-0.23	1						
25 SCM	0.04	0.03	-0.03	-0.02	-0.1	-0.1	-0.1	0.05	-0.18	0.05	-0.09	-0.05	-0.22	-0.03	-0.04	-0.07	-0.01	-0.01	-0.08	-0.16	-0.08	-0.08	-0.1	-0.32	1					
26 IT	0.06	0.14	-0.12	-0.01	-0.16	0.34	-0.14	0.07	0.18	-0.11	-0.05	-0.21	0.23	-0.04	-0.06	-0.26	-0.13	-0.05	-0.21	0.27	-0.13	-0.13	-0.15	-0.5	-0.22	1				
27 SALES	0.04	0	0.11	-0.01	-0.13	0.12	0.01	0.04	0.12	-0.04	0.13	-0.18	0.08	-0.02	-0.2	-0.12	0.14	-0.07	-0.19	0.2	-0.05	-0.05	-0.06	-0.18	-0.08	-0.12	1			
28 2012	-0.05	0.04	0.02	-0.02	0.01	-0.04	-0.02	-0.01	-0.07	-0.05	-0.07	-0.04	-0.04	-0.05	-0.18	-0.02	0.02	-0.1	-0.03	-0.11	-0.14	-0.14	0.01	0	0.13	0.03	-0.01	1		
29 2013	-0.04	0	-0.03	-0.01	0	0.02	-0.02	0	0	0.01	0.05	0.01	0.06	-0.05	0.16	0.03	-0.02	0.02	0.01	-0.03	-0.15	-0.15	-0.01	0.03	0.19	-0.02	0	-0.45	1	
30 2014	0.08	-0.05	0.02	0.03	-0.01	0.02	0.04	0.01	0.07	0.04	0.02	0.03	-0.02	0.1	0.01	-0.01	0	0.08	0.02	0.14	0.28	0.27	0	-0.03	-0.3	-0.01	0.02	-0.51	-0.5	1

Note. Correlations refer to 880 employee-year observations at the managerial level.

Table 8 Descriptive Statistics of Raters' Characteristics

Variable & definition	Total mean (sd)	Male raters mean (sd)	min	max	Female raters mean (sd)	min	max	Mean comparison t	df
AGE									
Age in years									
Non-managerial	47.21 (7.18)	47.51 (6.97)	25	63	41.67 (8.60)	27	56	11.81***	4,295
Managerial	49.46 (6.55)	49.56 (6.43)	30	63	47 (8.66)	29	56	2.27**	878
TENURE									
Tenure in years									
Non-managerial	18.37 (10.27)	18.74 (10.18)	0	44	11.50 (9.49)	1	38	10.20***	4,295
Managerial	15.11 (10.17)	14.88 (9.91)	0	39	20.83 (14.31)	2	38	-3.41***	878
PAYGRADE									
Pay grade (15=Non-tariff)									
Non-managerial	14.01 (1.38)	14.03 (1.35)	9	15	13.60 (1.84)	10	15	4.43***	4,295
Managerial	14.53 (1.20)	14.57 (1.15)	9	15	13.63 (1.88)	10	15	4.60***	878
PARTTIME									
1=Part-time, 0=Otherwise									
Non-managerial	0.01	0.01	0	1	0.12	0	1	-16.33***	4,295
Managerial	0.01	0.01	0	1	0.11	0	1	-9.12***	878
EMPLOYEE									
1=Rater also employee, 0=Otherwise									
Non-managerial	0.92	0.93	0	1	0.84	0	1	4.79***	4,295
Managerial	0.83	0.82	0	1	0.91	0	1	-1.39	878
SCORE									
Promotability score									
Non-managerial	3.15 (0.61)	3.15 (0.62)	2	5	3.17 (0.40)	2	4	-0.31	3,309
Managerial	3.41 (0.68)	3.43 (0.26)	2	5	2.95 (0.24)	2	3	3.34***	618

Notes. Descriptive statistics refer to observations of 4,082 (845) male and 215 (35) female rater-year observations at the non-managerial (managerial) level. Since not all raters are evaluated by their next higher supervisor, the number of observations for which raters' promotability scores are available reduces the number of rater-year observations to 3,200 (596) for male and 111 (24) for female raters at the non-managerial (managerial) level.

Table 9 Regression Estimates of Gender on Probability of Promotable Evaluation at the Non-Managerial Level by Functional Area

| PROMOTABLE	Finances	Operations	HR	Manufacturing	SCM	IT	Sales	Male-dominated	Gender-balanced
EMPLOYEE									
FEMALE	0.734	-0.001	0.254	**-0.418****	0.329	**-0.54****	0.062	**-0.261****	-0.197
	(0.486)	(0.408)	(0.648)	**(0.164)**	(0.226)	**(0.248)**	(0.544)	**(0.104)**	(0.395)
AGE	-2.820*	4.801***	-2.704	0.628*	1.247	0.678	-0.817	0.743***	-1.230
	(1.607)	(1.333)	(2.269)	(0.343)	(0.879)	(0.698)	(2.139)	(0.273)	(1.349)
AGE^2	0.029	-0.057***	0.032	-0.011***	-0.021*	-0.014*	0.013	-0.013***	0.017
	(0.019)	(0.016)	(0.027)	(0.004)	(0.011)	(0.008)	(0.029)	(0.003)	(0.017)
TENURE	-0.039	-0.576***	-0.419	-0.133**	-0.240*	0.085	-0.71**	-0.121**	-0.356
	(0.195)	(0.198)	(0.381)	(0.065)	(0.128)	(0.107)	(0.305)	(0.049)	(0.261)
PAYGRADE	0.094	-0.060	0.234*	0.081***	0.069	0.089*	0.373**	0.051***	0.071
	(0.089)	(0.096)	(0.137)	(0.024)	(0.045)	(0.054)	(0.163)	(0.016)	(0.091)
PARTTIME	-0.400	-0.288	0.203	-0.238	-0.78**	-0.241	-0.321	-0.300***	-0.395
	(0.377)	(0.313)	(0.704)	(0.168)	(0.382)	(0.245)	(0.485)	(0.116)	(0.406)
PERF	2.143	5.066***	4.57***	3.603***	3.54***	4.49***	7.99***	3.621***	4.629***
	(1.309)	(0.893)	(1.302)	(0.294)	(0.553)	(0.474)	(2.201)	(0.215)	(0.941)
CONTROLS				INCL.					
Observations	124	241	100	2,396	510	820	96	4,095	202
Number of dyads	124	241	59	1,245	345	439	67	2,100	128
Pseudo R^2	0.2762	0.4957	0.4253	0.2809	0.3105	0.2841	0.5241	0.2542	0.3651

Notes. Table reports coefficients after probit regressions with PROMOTABLE as the dependent variable. Age, age^2, and tenure multiplied by ten for ease of interpretation. All estimations include rater, department, unit, and year controls. Robust standard errors clustered at dyad-level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 10 OLS Regression Estimates of Gender on Promotability Scores at the Non-Managerial Level

| SCORES	(1)	(2)	(3)	(4)	(5)	(6)
EMPLOYEE						
FEMALE	**-0.089*****	**-0.080*****	**-0.075*****	**-0.075*****	**-0.071****	**-0.071****
	(0.024)	**(0.023)**	**(0.024)**	**(0.028)**	**(0.028)**	**(0.028)**
AGE	0.485***	0.180***	0.197***	0.144**	0.144**	0.142**
	(0.081)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)
AGE^2	-0.007***	-0.003***	-0.003***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
TENURE	0.029**	0.014	0.010	-0.003	-0.003	-0.003
	(0.013)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)
PAYGRADE		0.003	0.005	0.005	0.007	0.007
		(0.003)	(0.004)	(0.005)	(0.005)	(0.005)
PARTTIME		-0.049*	-0.055**	-0.046*	-0.043	-0.041
		(0.027)	(0.027)	(0.026)	(0.026)	(0.027)
PERF		1.227***	1.231***	1.453***	1.449***	1.449***
		(0.045)	(0.045)	(0.055)	(0.055)	(0.055)
RATER						
FEMALE			-0.016	-0.033	-0.036	-0.033
			(0.044)	(0.044)	(0.045)	(0.046)
AGE			-0.260*	-0.208	-0.203	-0.209
			(0.141)	(0.136)	(0.136)	(0.136)
AGE^2			0.002	0.002	0.002	0.002
			(0.002)	(0.001)	(0.001)	(0.001)
TENURE			0.028***	0.024**	0.024**	0.023**
			(0.011)	(0.012)	(0.011)	(0.011)
PAYGRADE			0.005	0.011	0.016**	0.016**
			(0.007)	(0.008)	(0.008)	(0.008)
PARTTIME			-0.152*	-0.114	-0.110	-0.106
			(0.083)	(0.082)	(0.083)	(0.083)
EMPLOYEE			0.019	0.010	0.017	0.026
			(0.031)	(0.030)	(0.030)	(0.031)
CONTROLS						
DEP				Yes	Yes	Yes
UNIT					Yes	Yes
YEAR						Yes
Constant	2.346***	-0.900***	-0.553	0.608	0.325	0.319
	(0.163)	(0.184)	(0.391)	(0.413)	(0.428)	(0.429)
Observations	4,297	4,297	4,297	4,297	4,297	4,297
Number of dyads	2,228	2,228	2,228	2,228	2,228	2,228
R^2	0.0338	0.2976	0.2983	0.3123	0.3145	0.3148

Notes. Table reports OLS coefficients with SCORES as the dependent variable. Age, age^2, and tenure multiplied by ten for ease of interpretation. Robust standard errors clustered at dyad-level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Fig. 6 Gender Differences in Promotable Evaluation at the Non-Managerial Level by Age - Restricted to Full-Time Employees
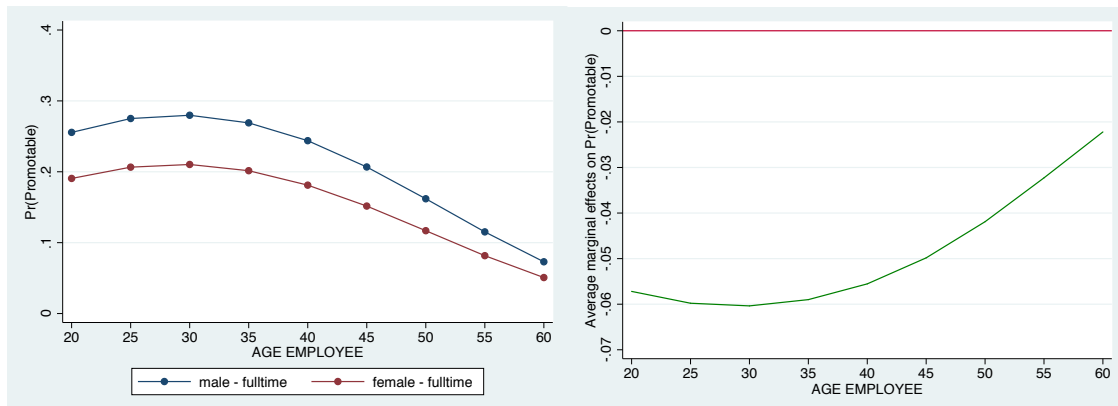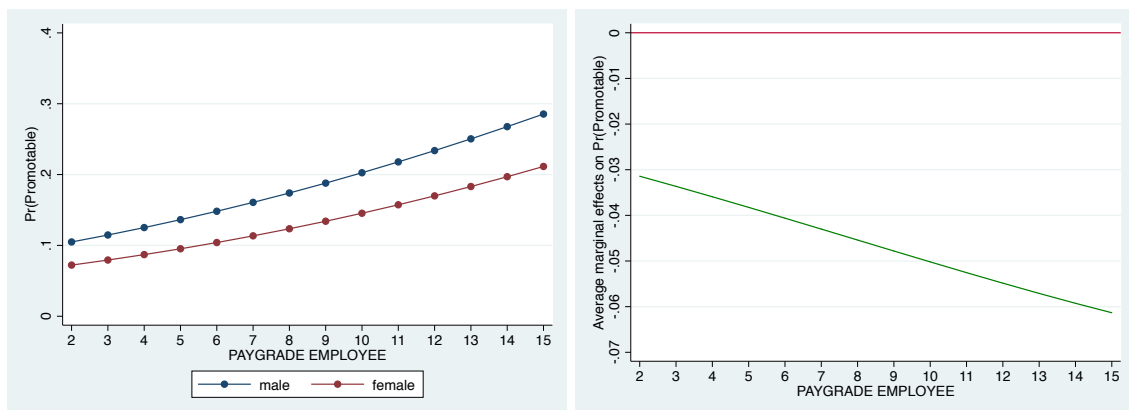

Notes. Predicted probabilities of men and women (left-hand side) and average marginal effects of gender (right-hand side) on being classified as PROMOTABLE on y-axis, respectively. Age levels on x-axes.

Fig. 7 Gender Differences in Promotable Evaluation at the Non-Managerial Level by Tenure
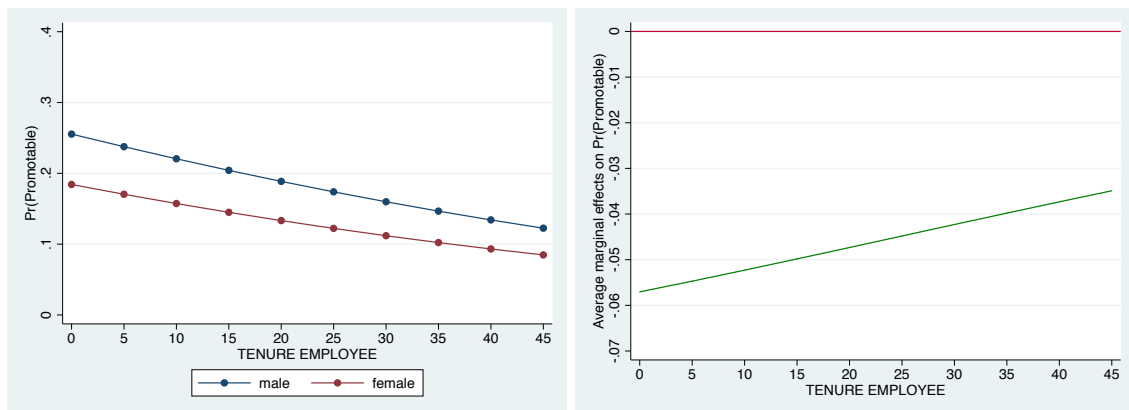
Notes. Predicted probabilities of men and women (left-hand side) and average marginal effects of gender (right-hand side) of being classified as PROMOTABLE on y-axis, respectively. Employees' tenure on x-axes.

Fig. 8 Gender Differences in Promotable Evaluation at the Non-Managerial Level by Pay Grade

Notes. Predicted probabilities of men and women (left-hand side) and average marginal effects of gender (right-hand side) of being classified as PROMOTABLE on y-axis, respectively. Employees' pay grades on x-axes.

Table 11 Regression Estimates of Gender on Probability of Promotable Evaluation at the Non-Managerial Level – Restricted to Male Raters

PROMOTABLE	COEFFICIENTS	AME
EMPLOYEE		
FEMALE	**-0.289*****	**-0.060*****
	(0.103)	**(0.021)**
AGE	0.741***	
	(0.269)	
AGE^2	-0.124***	
	(0.034)	
TENURE	-0.143***	-0.030***
	(0.049)	(0.010)
PAYGRADE	0.067***	0.014***
	(0.017)	(0.004)
PARTTIME	-0.254**	-0.053**
	(0.117)	(0.024)
PERF	3.612***	0.746***
	(0.213)	(0.038)
CONTROLS	INCL.	
Constant	-7.516***	
	(1.519)	
Observations	4,082	4,082
Number of dyads	2,106	
R^2	0.2578	

Notes. Table reports coefficients (column 1) and average marginal effects (column 2) after probit regressions with PROMOTABLE as the dependent variable. Age, age^2, and tenure multiplied by ten for ease of interpretation. No separate marginal effect for age^2 can be estimated. Robust standard errors clustered at dyad-level in parentheses. Estimations include rater, department, unit, and year controls. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Fig. 9 Gender Differences in Promotable Evaluation at the Non-Managerial Level by Raters' Age
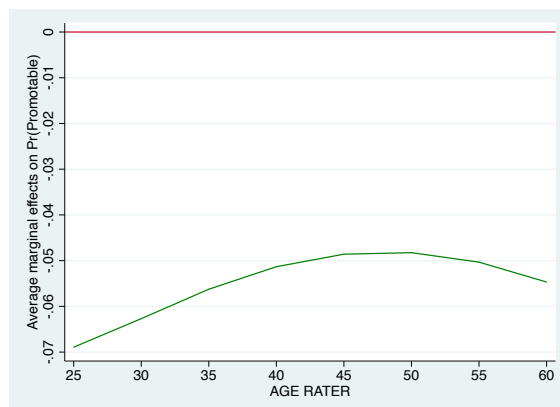
Note. Average marginal effects of gender on being classified as PROMOTABLE on y-axis. Raters' age levels on x-axis.

Fig. 10 Gender Differences in Promotable Evaluation at the Non-Managerial Level by Raters' Pay Grade
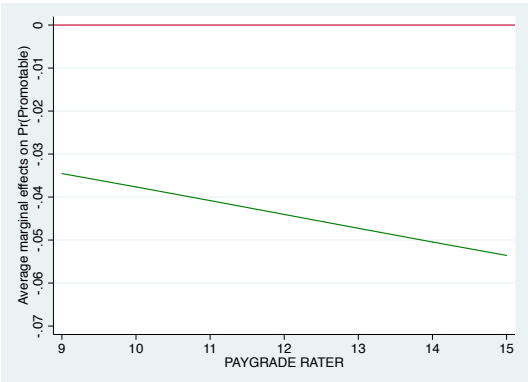
Notes. Average marginal effects of gender on being classified as PROMOTABLE on y-axis. Raters' pay grades on x-axis.

7 References

Acosta, Pablo. 2010. Promotion Dynamics the Peter Principle: Incumbents vs. External Hires. *Labour Economics*, 17(6): 975-986.

Aigner, Dennis J., and Glen G. Cain 1977. Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review*, 30(2): 175-187.

Arrow, Kenneth J. 1973. The Theory of Discrimination. In *Discrimination in Labor Markets*, eds. Orley Ashenfelter and Albert Rees, 3-33. Princeton: University Press.

Azmat, Ghazala, and Barbara Petrongolo 2014. Gender and the Labor Market: What Have We Learned from Field and Lab Experiments? *Labour Economics*, 30(C): 32-40.

Barnett, William, James N. Baron, and Toby E. Stuart 2000. Avenues of Attainment: Occupational Demography and Organizational Careers in the California Civil Service. *American Journal of Sociology*, 106(1): 88-144.

Barney, Jay B. 1991. Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1): 99-120.

Bartol, Kathryn M. 1999. Gender Influences on Performance Evaluations. In *Handbook of Gender and Work*, ed. Gary N. Powell, 165-178. Thousand Oaks: Sage.

Becker, Gary S. 1971. *The Economics of Discrimination*. Chicago: University of Chicago Press.

Becker, Gary S. 1985. Human Capital, Effort, and Sexual Division of Labor. *Journal of Labor Economics*, 3(1): 33-58.

Beekun, Rafik I., Yvonne Stedham, James W. Westerman, and Jeanne H. Yamamura. 2010. Effects of Justice and Utilitarianism on Ethical Decision Making: A Cross-Cultural Examination of Gender Similarities and Differences. *Business Ethics: A European Review*, 19(4): 309-325.

Bento, Regina F., Lourdes F. White, and Susan R. Zacur. 2012. The Stigma of Obesity and Discrimination in Performance Appraisal: A Theoretical Model. *International Journal of Human Resource Management*, 23(5): 3196-3224.

Blau, Francine D., and Jed DeVaro. 2007. New Evidence on Gender Differences in Promotion Rates: An Empirical Analysis of a Sample of New Hires. *Industrial Relations*, 46(3): 511–550.

Blau, Peter M. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.

Booth, Alison L., Marco Francesconi and Jeff Frank. 2003. A Sticky Floors Model of Promotion, Pay, and Gender. *European Economic Review*, 47(2): 295-322.

Bretz, Robert D., George T. Milkovich, and Walter Read. 1992. The Current State of Performance Appraisals Research and Practice: Concerns, Directions, and Implications. *Journal of Management*, 18(2): 321-352.

Cannings, Kathleen, and Claude Montmarquette. 1991. Managerial Momentum: A Simultaneous Model of the Career Progress of Male and Female Managers. *Industrial and Labor Relations Review*, 44(2): 212-228.

Catalyst. 2014. *Increasing Gender Diversity on Boards: Current Index of Formal Approaches*. New York: Catalyst.

Catalyst. 2015. *Catalyst Quick Take: Women in Male-Dominated Industries and Occupations in U.S. and Canada*. New York: Catalyst.

Catalyst. 2016. *Pyramid: Women in S&P 500 Companies*. New York: Catalyst.

Cleveland, Jeanette N., Kevin R. Murphy, and Richard E. Williams. 1989. Multiple Uses of Performance Appraisal: Prevalence and Correlates. *Journal of Applied Psychology*, 74(1): 130-135.

Cobb-Clark, Deborah A. 2001. Getting Ahead: The Determinants of and Payoffs to Internal Promotion for Young U.S. Men and Women. In *Research in Labor Economics*, 339-372. Greenwich: JAI Press.

Coffman, Julie, and Bill Neuenfeldt. 2014. Everyday Moments of Truth: Frontline Managers Are Key to Women's Career Aspirations. http://www.bain.com/publications/articles/everyday-moments-of-truth.aspx. Accessed 20 October 2016.

Cox, Taylor, and Stella M. Nkomo. 1992. Candidate Age as a Factor in Promotability Ratings. *Public Personnel Management*, 21(2): 197-210.

Czajka, Joseph M., and Angelo S. DeNisi. 1988. Effects of Emotional Disability and Clear Performance Standards on Performance Ratings. *Academy of Management Journal*, 31(2): 394-404.

De Pater, Irene E., Annelies E. M. van Vianen, Myriam N. Bechtoldt, and Ute-Christine Klehe. 2009. Employees' Challenging Job Experiences and Supervisors' Evaluations of Promotability. *Personnel Psychology*, 62(2): 297-325.

Destatis. 2015a. Bevölkerung – Geburten. https://www.destatis.de/DE /ZahlenFakten/GesellschaftStaat/Bevoelkerung/Geburten/Tabellen/GeburtenMutter alter.html. Accessed 16 October 2016.

Destatis. 2015b. Parent Working Part-Time. https://www.destatis.de/ EN/FactsFigures/Indicators/QualityEmployment/Dimension3/3_7_ParentsWorking Parttime.html. Accessed 16 October 2016.

Eagly, Alice H., and Steve J. Karau. 2002. Role Congruity Theory of Prejudice Toward Female Leaders. *Psychological Review*, 109(3): 573-598.

Elvira, Marta, and Robert Town. 2001. The Effects of Race and Worker Productivity on Performance Evaluations. *Industrial Relations*, 40(4): 571-590.

Ferris, Gerald R., Valerie L. Yates, David C. Gilmore, and Kendrith M. Rowland. 1985. The Influence of Subordinate Age on Performance Ratings and Causal Attributions. *Personnel Psychology*, 38(3): 545–557.

Galin, Amira, and Barbara Benolie. 1990. Does the Way You Dress Affect Your Performance Rating? *Personnel*, 67(8): 49-53.

Gerhart, Barry A., and George T. Milkovich. 1989. Salaries, Salary Growth, and Promotions of Men and Women in a Large, Private Firm. In *Pay Equity: Empirical Inquiries*, eds. Robert T. Michael, Heidi I. Hartmann and Brigid O'Farell, 23-43. Washington: National Academy Press.

Gesamtmetall. 2011. Frauen in Führungspositionen in der M+E-Industrie. http://www.gesamtmetall.de/sites/default/files/downloads/frauen-in-fuehrungspositionen.pdf. Accessed 20 October 2016.

Giuliano, Laura, David I. Levine, and Jonathan Leonard. 2011. Racial Bias in the Manager-Employee Relationship: An Analysis of Quits, Dismissals, and Promotions at a Large Retail Firm. *Journal of Human Resources*, 46(1): 26-52.

Greenhouse, Jeffrey H., Saroj Parasuraman, and Wayne M. Wormley. 1990. Effects of Race on Organizational Experiences, Job Performance Evaluations, and Career Outcomes. *Academy of Management Journal*, 33(1): 64-86.

Grund, Christian. 2015. Gender Pay Gaps Among Highly Educated Professionals – Compensation Components Do Matter. *Labour Economics*, 34(C): 118-126.

Harris, Kenneth J., K. Michele Kacmar, and Dawn S. Carlson. 2006. An Examination of Temporal Variables and Relationship Quality on Promotability Ratings. *Group & Organization Management*, 31(6): 677-699.

Haslam, S. Alexander, and Michelle K. Ryan. 2008. The Road to the Glass Cliff: Differences in the Perceived Suitability of Men and Women for Leadership Positions in Succeeding and Failing Organizations. *Leadership Quarterly*, 19(5): 530-546.

Hausmann, Ann-Christin, and Corinna Kleinert. 2014. Berufliche Segregation auf dem Arbeitsmarkt – Männer- und Frauendomänen kaum verändert. *IAB Kurzbericht 09/2014*. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung.

Heilman, Madeline E., and Melanie H. Stopeck. 1985. Being Attractive, Advantage or Disadvantage? Performance-Based Evaluations and Recommended Personnel Actions as a Function of Appearance, Sex, and Job Type. *Organizational Behavior and Human Decision Processes*, 35(2): 202-215.

Hill, Aaron D., Arun D. Upadhyay, and Rafik I. Beekun. 2015. Do Female and Ethically Diverse Executives Endure Inequity in the CEO Position or Do They Benefit from Their Minority Status? An Empirical Examination. *Strategic Management Journal*, 36(8): 1115-1134.

Hoobler, Jenny M., Sandy J. Wayne, and Grace Lemmon. 2009. Bosses' Perceptions of Family-Work Conflict and Women's Promotability: Glass Ceiling Effects. *Academy of Management Journal*, 52(5): 939-957.

Jawahar, I. M., and Gerald R. Ferris. 2011. A Longitudinal Investigation of Task and Contextual Performance Influences on Promotability Judgments. *Human Performance*, 24(3): 251-269.

Joshi, Aparna, Jooyeon Son, and Hyuntak Roh. 2015. When Can Women Close the Gap? A Meta-Analytic Test of Sex Differences in Performance and Rewards. *Academy of Management Journal*, 58(5): 1516-1545.

Kampkötter, Patrick. 2016. Performance Appraisals and Job Satisfaction. *International Journal of Human Resource Management*. Published online, doi: 10.1080/095851 92.2015.1109538.

Koch, Amanda J., Susan D. D'Mello, and Paul R. Sackett. 2015. A Meta-Analysis of Gender Stereotypes and Bias in Experimental Simulations of Employment Decision Making. *Journal of Applied Psychology*, 100(1): 128-161.

Kraiger, Kurt, and J. Kevin Ford. 1985. A Meta-Analysis of Ratee Race Effects in Performance Ratings. *Journal of Applied Psychology*, 70(1): 56-65.

Kunze, Astrid. 2015. The Family Gap in Career Progression. In Gender Convergence in the Labor Market, *Research in Labor Economics*, 41, eds. Solomon W. Polacheck, Konstantinos Tatsiramos and Klaus F. Zimmermann, 115-142. Greenwich: JAI Press.

Kuvaas, Bard 2006. Performance Appraisal Satisfaction and Employee Outcomes: Mediating and Moderating Roles of Work Motivation. *International Journal of Human Resource Management*, 17(3): 504-522.

Landy, Frank J., and James L. Farr. 1983. *The Measurement of Work Performance: Methods, Theory, and Applications*. New York: Academic Press.

Lazear, Edward P., and Sherwin Rosen. 1990. Male-Female Wage Differentials in Job Ladders. *Journal of Labor Economics*, 8(1): 106-123.

Lee, Peggy M., and Erika H. James. 2007. She'-E-Os: Gender Effects and Investor Reactions to the Announcements of Top Executive Appointments. *Strategic Management Journal*, 28(3): 227-241.

Levy, Paul E., and Jane R. Williams. 2004. The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management*, 30(6): 881-905.

Lewis, Gregory 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.

Maas, Victor S., and Raquel Torres-González. 2011. Subjective Performance Evaluation and Gender Discrimination. *Journal of Business Ethics*, 101(4): 667-681.

Murphy, Kevin R., and Jeanette Cleveland. 1995. *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*. Thousand Oaks: Sage.

Niederle, Muriele 2014. Gender. *NBER Discussion Paper No. 20788*. Cambridge, National Bureau of Economic Research.

Petersen, Trond, and Ishak Saporta. 2004. The Opportunity Structure for Discrimination. *American Journal of Sociology*, 109(4): 852–902.

Phelps, Edmund S. 1972. The Statistical Theory of Racism and Sexism. *American Economic Review*, 62(4): 659-661.

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

Powell, Gary N., and Laura M. Graves. 2003. *Women and Men in Management*. Thousand Oaks: Sage.

Prendergast, Canice. 1999. The Provision of Incentives in Firms. *Journal of Economic Literature*, 37(1): 7-63.

Ransom, Michael, and Ronald L. Oaxaca. 2005. Intrafirm Mobility and Sex Differences in Wages. *Industrial and Labor Relations Review*, 58(2): 219-237.

Roth, Philip L., Kirsten L. Purvis, and Philip Bobko. 2012. A Meta-Analysis of Gender Group Differences for Measures of Job Performance in Field Studies. *Journal of Management*, 38(2): 719-739.

Ryan, Michelle K., and S. Alexander Haslam. 2007. The Glass Cliff: Exploring the Dynamics Surroundings the Appointment of Women to Precarious Leadership Positions. *Academy of Management Review*, 32(2): 549-572.

Spilerman, Seymour, and Trond Petersen. 1999. Organizational Structure, Determinants of Promotion, and Gender Differences in Attainment. *Social Science Research*, 28(2): 203-227.

Terjesen, Siri, Ruth V. Aguilera, and Ruth Lorenz. 2015. Legislating a Woman's Seat on the Board: Institutional Factors Driving Gender Quotas for Boards of Directors. *Journal of Business Ethics*, 128(2): 233-251.

Thacker, Rebecca A., and Sandy J. Wayne. 1995. An Examination of the Relationship Between Upward Influence Tactics and Assessments of Promotability. *Journal of Management*, 21(4): 739-756.

Tsui, Anne S., and Barbara A. Gutek. 1984. A Role Set Analysis of Gender Differences in Performance, Affective Relationships, and Career Success of Industrial Middle Managers. *Academy of Management Journal*, 27(3): 619-635.