Dips and Floors in Workplace Training: 
Gender Differences and Supervisors

Bernd Fitzenberger* and Grit Muehler†

September 2014

Abstract

This paper provides a detailed decomposition analysis of the gender differences in workplace training throughout the working life with a particular focus on parental leave and supervisors using personnel records from a large German firm. Females obtain less training during the early career, and more at higher age. The timing of the training gap seems to be driven by diverging career paths associated with employment interruptions. However, we find no evidence for catching-up effects after parental leave. Furthermore, including supervisor-fixed effects cannot explain the gender differences in training. The training of both male and female employees is positively associated with the training of the supervisor.

Keywords: training, gender gap, company data
JEL-classification: M53, M12, J14

Acknowledgements: We are grateful to the anonymous referee for many useful comments and suggestions. We would like to thank the company that kindly provided the data. We also benefitted from helpful discussions with Michael Beckmann, Günther Schulze, and Thomas Zwick. Part of this work was financed by the German Research Foundation (DFG) through the research project “Personnel Management in Times of Technical and Organizational Change” within the DFG priority program “Flexibility in Heterogenous Labor Markets” (FSP 1169). All errors are ours.

*University of Freiburg, IFS, IZA, ROA, ZEW. Corresponding Author: Department of Economics, University of Freiburg, 79085 Freiburg, Germany, E-mail: bernd.fitzenberger@vwl.uni-freiburg.de.
†Centre for European Economic Research (ZEW). P.O. Box 103443, D-68034 Mannheim, E-mail: muehler@zew.de.
1 Introduction

Workplace training is important for the acquisition of human capital after the end of the primary education in school and university (OECD, 2003). By providing workplace training, firms aim at developing the skills and competencies of their employees during their professional careers in order to retain or increase their productivity. And by participating in workplace training, employees can expect positive effects not only on wages, but also on employment prospects or career development (see Groot, 1999; Asplund, 2005; Hansson, 2008, for comprehensive reviews). When starting a job after the initial secondary or tertiary education, workplace training provides the necessary job specific knowledge to build a career in the industry chosen. Workplace training may be necessary for career progression and for catch up after employment interruptions.

The gender training gap may evolve in different ways. On the one hand, gender differences in career paths may be associated with a gender training gap, which grows with age, potentially reinforcing gender differences in outcomes. Individual incentives to participate in training are the higher the longer the pay-back period (Becker, 1964). Hence, older employees and female employees are likely to participate less in training (Bishop, 1996; OECD, 2003). Anticipating longer employment interruptions or a higher incidence of part-time work, management might refrain from providing and financing workplace training for women (Barron et al., 1993). On the other hand, workplace training may also be a means for female or older employees to catch up. When returning full time to the labor market, female employees might find it necessary to engage more into workplace training than males of equal age due to greater training needs. Females may also use workplace training to signal their high labor force attachment. Bishop (1996) finds evidence for greater training needs the shorter the tenure within a job.

Lower workplace training by females could also be due to the fact that a supervisor may prefer higher training of employees of the same gender (Shakeshaft et al., 1991; Rothstein, 1997; Melero, 2004). The supervisor’s observed training behavior may reflect his/her views on the needs of training for employees. If positions of supervisors are predominantly filled
by males, this would not only result in initial differences but also contribute to a widening of the gender training gap as workers age in their job.

Using personnel records from a single large German firm, this paper takes a life cycle perspective to analyze the age profile of the gender training gap by applying an age-specific decomposition approach. The paper makes two substantive and two small methodological contributions. First, we use very reliable panel data on all employees in a large firm in the financial industry to analyze both the incidence and the duration of training. Our analysis is an example of insider econometrics holding firm characteristics and the management strategies constant (Bartel et al., 2004; Ichniowski and Shaw, 2003). Second, our study is the first to analyze empirically the impact of both the training behavior and the gender of the supervisor on the amount of training and the gender training gap among the subordinates. As a first methodological contribution, our paper is the first to decompose the continuous age profile of the gender differences in workplace training. We estimate the age specific characteristics effect, the age specific coefficient effect, and the age composition effect. As a second methodological contribution, we use a weighted block bootstrap approach (Barbe and Bertail, 1995) for the decomposition results based on probit regressions for the incidence of training. We adapt the approach to estimate clustered standard errors for the decomposition results.

Our empirical results show a gender training gap close to zero at the age of 25. The gap grows afterwards and peaks at age 35, when females have a 0.75 days shorter training duration per year. This divergence in training behavior can be mainly attributed to differences in characteristics (such as wage, working time, or position in the firm). There is no catching-up effect after child birth. Furthermore, including supervisor-fixed effects cannot explain the gender differences in training. However, the training of both male and female employees is positively associated with the training duration of the supervisor.

The remainder of this paper is organized as follows. The next section motivates our analysis by giving some background information and a brief literature review. Section 3 describes the data. The decomposition approach is outlined in Section 4. Section 5 presents results and discussion, and Section 6 concludes. The appendix involves addi-
tional material and results, and the additional online appendix provides complementary empirical results.

2 Literature and Hypotheses

Germany’s vocational education and training (VET) system allows for a smooth transition into the labor market for the vast majority of the non-college bound youth after graduation from secondary school. The skills acquired during vocational training in the VET system are in high demand in the German labor market (Quintini, 2011). The dual vocational training system in Germany involves training both in a company and in a vocational training school (see Wolter and Ryan (2011) and Franz and Soskice (1995)). The good outcomes for upper secondary VET provide a particularly strong basis for future upskilling and workplace training (Fazekas and Field, 2013; Pischke, 2001).

Possibly because of the VET system, the participation in training among employees, who are not in an apprenticeship, is somewhat lower in Germany than in the Nordic countries, the UK, or the U.S., but it is still slightly above the OECD average of 41%. In 2007, 45% of 25-64-year-olds and 53% within the group of employed adults took part in formal and/or non-formal training in Germany (OECD, 2012). Those working full-time and those with higher level qualifications are much more likely to participate.

Participation in training can be viewed both as an investment of the firm into the training of the employee and as an investment of the employee into his/her career. Training differences between men and women might exist from the beginning of the working life onwards. A large part of the gender wage gap is due to the segregation of men and women into different occupations which starts with the apprenticeship system (Kunze, 2005; Kunze and Fitzenberger, 2005). This might also hold for the gender training gap, as workplace training is an important means to provide the necessary on-the-job skills.
2.1 The Gap in Workplace Training

Although, age and gender effects in overall training are well documented\(^1\), little is known about the gender training gap at the workplace and its development as employees age. When restricting training to formal, employer-provided training courses which take place during paid working time, studies show that female employees are equally likely to receive training as males (or slightly less), but their training durations are unambiguously shorter (O’Halloran, 2008), in particular when they are young (Bassanini et al., 2007).\(^2\) And, the gender differences in training volumes are more pronounced, when taking a life cycle perspective (Bishop, 1996; OECD, 2003). To our knowledge, there exists no study which analyzes the age profile of the gender training gap and which performs a comprehensive decomposition analysis thereof.

Studies that analyze employer-provided training based on personnel records of single firms are scarce. Due to lack of employer-reported data, most evidence for Europe is based on training information, which is self-reported by employees (Bassanini et al., 2007; Grund and Martin, 2012). In-depth information on e.g. the content of the training, duration of training, or the background of the participants is difficult to obtain in household survey data (such as the European Community Household Panel, the National Longitudinal Survey of Youth, or the German Socioeconomic Panel) or in multi-firm data (such as the Continuing Vocational Training Survey) as there are often long periods of recall time and a limited amount of waves for which the information is observed.\(^3\) Also, because of recall bias, training information, which is self-reported by employees, is generally considered less reliable than employer-reported training data.\(^4\) This holds in particular when one

---

\(^1\)Gender differences in overall training are the subject of a large literature where females are found to participate less (Lynch, 1992; Royalty, 1996; Evertsson, 2004; Grund and Martin, 2012), to the same extent (Altonji and Spletzer, 1991; Barron et al., 1993; Veum, 1996; Arulampalam et al., 2004; Frazis and Loewenstein, 2006) or more often (Green and Zanchi, 1997; OECD, 1999, 2003; Simpson and Stroh, 2002; Bassanini et al., 2007; O’Halloran, 2008) in workplace training compared to males.

\(^2\)Unfortunately, little is known about informal training. However, there seems to be a strong positive correlation between formal and informal training at the firm level (OECD, 1999; Bishop, 1996).

\(^3\)The GSOEP contains information on further training for the years 1989, 1993, 2000, 2004, and 2008. However, respondents need to recall all their trainings in the last three years with durations from a few hours up to several months and further information such as course content is only available for two out of the five years (1989, 1993).

\(^4\)A great deal of the literature, especially for the US, is concerned with the reliability of the training information used (Bartel, 1995; Bishop, 1996; OECD, 1999).
wants to distinguish between the incidence and the duration (intensity) of training by employees.

We are aware of only three studies focusing on personnel records for estimating the effects of employer-provided training on wages and other outcomes. Krueger and Rouse (1998) assess the impact of a specific employer-based education program, Bartel (1995) focuses on the nature of training, i.e. whether it is remedial or used to stimulate career advancement, and Xiangmin and Batt (2007) examine the productivity effects of informal training. All of the studies find positive impacts of training on either productivity or wages. However, none of them focuses on training participation over age or on gender differences in the participation and duration of training.

Barron et al. (1993) provide a theoretical model to explain the gender training gap. The model assumes that female employees are less attached to the firm and therefore the firm invests less into workplace training of female employees. Our hypothesis is that females participate less in workplace training than males because firms expect that employment interruptions or quits are more likely for female employees compared to male employees. Furthermore, we expect that the gender training gap to increase during the early career because it may become more likely over time that a female will have a child. Accordingly, firms will invest less into the careers of their female employees, and female employees may also be less interested in career development, if they expect to have a child soon. In contrast, we expect a convergence in training after having returned to their previous jobs (catching-up of mothers) and/or after the cessation of fertility. This convergence may relate to firms or/and females becoming more interested in training after a certain age. An alternative hypothesis relates to the fact that male employees may be more mobile across firms. Thus, firms may be less inclined to invest into the training of unstable employees.

Furthermore, observed gender differences in training may be affected by selection effects caused by the sorting of male and female employees into different professions, industries,
or firms. Because female supervisors and female employees may be strongly selected into specific professions, industries, or firms, it is important to have access to firm level data, which allows to control pertinent selection effects (Rothstein, 1997). By using personnel records for all employees of a single large firm, our analysis allows to hold firm characteristics and the management strategies constant. Note that the financial industry implements a larger amount of formal training in comparison to other industries (OECD, 1999).

2.2 Gender and Training Policy of the Supervisor

The gender and the training of supervisors is often neglected in the training literature. A first hypothesis to be tested relates to the interaction with the same gender (Rothstein, 1997). Supervisors may have a preference to allow for more training (and thus to promote the career) among subordinates of the same gender because their cooperation involves less frictions. For instance, Cardoso and Winter-Ebmer (2010) find for Portugal that female employee benefit from working in female-led firms compared to male-led firms. Since a majority of supervisors is male, there may on average exist a preference by supervisors for training of male employees. However, the effect may also be reversed if groups of the lower status individuals, e.g. females, tend to identify with members of the higher status group, e.g. males (Cardoso and Winter-Ebmer, 2010, p.145).

A second hypothesis focuses on gender differences in supervisor training policy. Following Shakeshaft et al. (1991), female supervisors may prefer more training than male supervisors (irrespective of the gender of the subordinates) because female supervisors value competence more strongly and male supervisors value trust more strongly. Melero (2004) emphasizes the importance of gender stereotypes in the management style of supervisors. According to this view, female supervisors favor a more interpersonal and interactive relationship at the workplace, and they engage more in employee-mentoring. In contrast,

---

6The study finds that female employees show higher wages in female-led firms. Cardoso and Winter-Ebmer (2010, p.144) suspect that preferences may result in mentoring one group of employees by help in career progression and access to on-the-job training and networks. For this effect, Cardoso and Winter-Ebmer (2010, p. 145) refer to the similarity-attraction paradigm in social psychology (Schneider, 1987).
male supervisors are more task-oriented and follow an authoritarian management style. This line of argument suggests that female supervisors prefer more formal training than male supervisors, irrespective of the gender of the subordinates. Borghans et al. (2006) argue that female employees have better interactive skills (higher social capital). Therefore, female supervisors may be better able to assess training needs, and they have a better supervision style than male supervisors.

A third hypothesis says that the amount of training a supervisor receives may itself show a positive association with the training of the subordinates, possibly reflecting the needs at the workplace. Bandura’s (1977) social learning theory suggests that employees tend to emulate the behavior of their supervisor to ensure that their behavior is consistent with accepted work norms and expectations. Schneider (1987) suggests a attraction-selection-attrition perspective implying that supervisors who value the benefits of training are more likely to attract and select subordinates of the same orientation.

Our analysis includes supervisor fixed effects to analyze if supervisors are treating male and female employees differently, and we analyze the gender match between supervisor and employee.

3 Data

We use data from a high-paying firm with a high-skilled workforce in the financial industry in Germany. From the human resources department we received administrative personnel records on all of the firm’s employees based in Germany for the years 2004 to 2007. In addition, we obtained for each of these years the training records on those employees who participated in one of the training courses provided by the firm. The continuous training programs offered by the employer involve around 3,000 courses. Participation is organized using special software. The software was introduced in 2003 and it is used to carry out all the steps necessary for participation in the firm’s training programs: information research, signing up, and approval of the supervisor. The firm leaves training

---

7The reporting date is 31st of December in each year.
decisions to the responsibility of its employees and thereby aims at encouraging an active interest in one’s own lifelong learning. At the same time, we think that employees typically choose training in agreement with their supervisors. As we do not have any information on informal training such as learning-by-doing, observing co-workers or by simply asking colleagues (which are in addition very difficult to measure), we restrict the analysis to formal training courses, provided and paid by the employer. In total, we have 131,130 person-year observations (69,907 [53.3%] among females and 61,223 [46.7%] among males).

We make some adjustments regarding courses and individuals included. As the firm operates in the financial sector, it is obliged by law to guarantee that the employees can always comply with legal requirements. We drop participation in the corresponding mandatory courses, which represent 7.6% of the total number of courses. On average, a mandatory course lasts slightly longer than half a day (0.6 days), and participation in all mandatory courses adds up on average to 0.78 days per year and employee. Furthermore, we include only completed courses.

Regarding the individuals included, we drop all individuals who are younger than 25 years or older than 50 years to avoid unusual training behavior due to the start or end of the professional career. For the same reason, we also drop trainees, interns, and expatriates. Our final sample consists of 101,889 observations, 54,793 (53.8%) females and 47,096 (46.2%) males. The adjustment with respect to the age range considered does not change the gender ratio compared to the original data set. The share of females among employees falls from 60% at age 25 to 48% at age 50 (detailed results are available upon request).

Note that due to the short panel dimension these may reflect cohort differences.

Training characteristics covered in the data are training incidence, training intensity (number, length of course) and training category such as business, information technology, professional qualifications. Our main outcome variables of interest are training participation and number of training days per year. Training participation is a dummy variable which takes the value one when the employee participates in at least one training course per year.

---

8Most job training is informal, however (Bishop, 1996). But as Veum (1995) points out, informal training such as learning-by-doing, observing co-workers or by simply asking colleagues is very difficult to measure.

8
and zero otherwise. Training duration by calendar year is calculated based on the days and hours spent in formal employer provided training. If the training period is longer than five consecutive days, the calculation is adjusted to weekdays.

[Table 1 about here.]

The average training gap between males and females (gender training gap) is 6 percentage points for training incidence and 0.38 training days per year for training duration (see Table 1). A closer look shows that the gap in training participation and duration varies considerably by age. Fig. 1 shows the average training duration by age and gender (‘Duration females’ and ‘Duration males’). Starting from nearly equal participation rates and durations for employees in their mid twenties, the gender specific profiles diverge during the first 10 to 15 years of working life. The gap in training duration between males and females is largest at age 35 when females have 0.75 days per year less training than males. Above age 35, the gap starts to close again, and at age 50, it amounts to 0.2 days per year. Thus, on average, females receive less training than males, with the highest gap around age 35.

As most of the divergence in training occurs between age 25 and 40, childbirth followed by a period of parental leave could rationalize the divergence. However, training behavior could as well just reflect divergence in other career related factors such as wage, working time or hierarchical status. Table 1 shows the differences between males and females in these characteristics. The average earnings are 58,000 Euro for males and 39,000 Euro for females. 95% of the male and 55% of the female employees work full-time. The corporate title is an important indicator for the hierarchical level of the employees’ position. 73% of the male employees and 41% of the female employees hold a corporate title. Possessing

---

9Training duration is set to zero for individuals who do not participate in firm-provided training in a specific year. The summary statistics and the regressions for training duration include these individuals. As a robustness check regarding the sensitivity to outliers, estimation of training duration was also implemented using a tobit regression, where long training durations are artificially censored at 20 days per year because of observations with very long training durations. As the results of the linear regression and the tobit regression are similar, we report only the linear regressions for the estimation of training duration.

10The variable describes the basis wage, which does not include bonuses or other extra-payments. In the regressions in the following section, wage is expressed in dimension 1,000 Euro.
a title is not necessarily associated with supervisory responsibility. 20% of the males but only 6% of the females in our sample are supervisors.

Further evidence shows that the gender differences in wages, working time, and hierarchical status are widening during the first ten years of the professional career (detailed graphical evidence is available in the additional online appendix). The working time decreases for females until age 35. At this age, the average woman works 62% of a full-time contract, whereas the average male of the same age works 98%. The development of this and the other individual and job characteristics are likely to be associated with differences in training participation and duration. We decompose the outcome variables by age to analyse in more detail to which extent the gap is driven by differences in characteristics and by differences in coefficients associated with these characteristics.

[Figure 1 about here.]

4 Decomposition Approach

Investigating the gender gap in training behavior requires a decomposition approach which is able to capture the dynamics along the age profile, rather than decomposing mean differences. Our outcome variables of interest are the probability to participate in firm-provided formal training and the length of training per year. As we are going to draw age profiles of the gender specific training behavior, we need to extend the standard Blinder-Oaxaca approach and decompose our two outcomes – training probability and training duration – into three terms: an age-specific coefficients effect, an age-specific characteristics effect, and an age composition effect. This last effect captures the changing gender composition over the age groups.

We consider individuals from age 25 to age 50 grouped in 26 age cells. This decomposition approach extends the standard Blinder (1973) and Oaxaca (1973) decomposition by decomposing the characteristics and the coefficients effects along the age dimension and by accounting for a separate age “composition effect”. Our two outcomes – training
probability and training duration – are decomposed into the age profiles of three terms: the age-specific coefficients effect, the age-specific characteristics effect, and the age composition effect. This last effect captures the changing gender composition over the age groups. In the following, we discuss our decomposition approach. The details of the derivation can be found in appendix A.1.

4.1 Decomposition of Training Duration

The decomposition of training duration is based on separate linear regressions for males and females. The decomposition for the number of training days per year can be written in a formal way as follows. The two groups males ($m$) and females ($f$) and the difference in the mean outcomes of females ($\bar{Y}_f$) and males ($\bar{Y}_m$) are of interest to us. The difference in mean outcomes is given by

$$\bar{Y}_f - \bar{Y}_m = \bar{Z}_f \hat{\gamma}_f - \bar{Z}_m \hat{\gamma}_m.$$  \hspace{1cm} (1)

The set of regressors $Z^g$ ($g \in \{f,m\}$) includes a full set of age dummy variables and further regressors $x^g_i$. $\bar{Z}^g$ comprises the sample means of the regressors by gender. $\gamma^g$ are the regression coefficients. Accounting for each year of age between 25 and 50 separately, our age profile is based on 26 different age cells $a = 25, \ldots, 50$.

The difference in outcomes can be written as

$$\bar{Y}_f - \bar{Y}_m = \sum_{a=25}^{50} h^f_a \bar{Y}_a^f - \sum_{a=25}^{50} h^m_a \bar{Y}_a^m = \sum_{a=25}^{50} h^f_a \bar{Z}_a^f \hat{\gamma}_f - \sum_{a=25}^{50} h^m_a \bar{Z}_a^m \hat{\gamma}_m,$$  \hspace{1cm} (2)

where $Y^g_a$ and $Z^g_a$ are gender specific means by age and $h^g_a$ denotes the gender specific age distribution.
As an extension of a standard Blinder-Oaxaca decomposition, which we derive in the appendix, one can decompose the total difference in the mean training duration into the age profiles of the age specific decomposition and an age composition effect as

\[
\sum_{a=25}^{50} h_a^f \bar{Y}_a^f - \sum_{a=25}^{50} h_a^m \bar{Y}_a^m = \sum_{a=25}^{50} h_a^m \left[ (\hat{\alpha}_a^f - \hat{\alpha}_a^m) + \bar{X}_a^m (\hat{\beta}_f - \hat{\beta}_m) + (\bar{X}_a^f - \bar{X}_a^m) \hat{\beta}_f + \left( \frac{h_a^f - h_a^m}{h_a^m} (\hat{\alpha}_a^f + \bar{X}_a^f \hat{\beta}_f) \right) \right].
\]

Here, \( \hat{\alpha}_a^g \) denotes the coefficient of the age dummy variable and \( \bar{X}_a^g \) are the age specific means of the remaining regressors.

Equation (3) shows that the average gender gap in training days per year involves the age profiles of three terms, namely of the age specific coefficients effect (coef), the age specific characteristics effect (char), and the age composition effect (comp). It is well known that decompositions are not unique in the sense that they can be based on different counterfactuals (Fortin et al., 2011). The overall contributions of the three age profiles to the total gender gap in training duration are calculated based on \( h_a^m \), i.e. holding the age composition of males constant. The first two terms the use the counterfactual based on male characteristics and female coefficients, thus, evaluating differences in coefficients at male characteristics and differences in characteristics at female coefficients. In addition, the age composition effect accounts for the differences in the age composition at average female training variables. Our calculation of age-specific effects relies on two counterfactual outcomes instead of one as the standard Blinder-Oaxaca decomposition. We denote the counterfactual outcome that is needed to calculate the coefficients and characteristics effect by age as \( \text{CharM-CoefF-AgeM} \). This is the outcome that would prevail if males kept their characteristics but participated in training in the same way as females – or, to put it differently, if females were assigned male characteristics but assignment conditional on characteristics was still according to female coefficients. The coefficients effect explains the differences in training participation due to gender specific differences in coefficients.
The age specific coefficients effect aggregates to the coefficients effect known from the standard Blinder-Oaxaca decomposition of the total gender gap (see appendix). The age specific characteristics effect accounts for differences in the characteristics within one age group. As shown in Section 3, females and males differ, for instance, in terms of wage, working time, supervisory status, and the age composition.

Three outcomes – the actual male and female outcome and the counterfactual $CharM-CoefF-AgeM$ – are calculated within an age cell. To fully assess the age-specific decomposition, these age cells have to be weighted to reflect the importance of an age group with regard to the total number individuals for each gender. The fourth outcome, i.e. the counterfactual $CharF-CoefF-AgeM$ accounts for the changing gender composition across age groups (the share of female employees falls with age). This “counterfactual female outcome” is the male-weighted outcome for females in each age cell.\textsuperscript{11} The composition effect is calculated as the difference of the counterfactual female outcome and the actual outcome observed for females. Thus, the composition effect picks up the change in characteristics, which occurs due to the changing gender composition along the age profile.

Specifically, the decomposition in Equation (3) rests on two counterfactuals. First,

$$
\sum_{a=25}^{50} h_m^a (\hat{\alpha}_a^f + \bar{X}_a^m \hat{\beta}^f),
$$

the counterfactual based on male characteristics (by age), female coefficients, and the male age composition ($CharM-CoefF-AgeM$). And second,

$$
\sum_{a=25}^{50} h_m^a (\hat{\alpha}_a^f + \bar{X}_a^f \hat{\beta}^f),
$$

the counterfactual based on female characteristics (by age), female coefficients, and the male age composition ($CharF-CoefF-AgeM$). Adding and substracting these counterfactuals

\textsuperscript{11}The outcomes are weighted with the number of males respective age group in relation to the total number of males.
tuals yields the decomposition (see appendix A.1.1). As part of the empirical results, we will also discuss the age profiles of counterfactuals.

Standard errors and confidence intervals for the age-specific effects are obtained by bootstrapping individuals with 1,000 replications to obtain clustered standard errors at the individual level, which account for the correlation of the observations of the same individual over time (cluster bootstrap, see Cameron et al., 2008).

4.2 Decomposition of Training Probability

The decomposition for training participation applies the Fairlie (2005) decomposition for a probit model of the training probability in the two groups. The decomposition parallels the decomposition in the linear case with the slight difference that probabilities instead of linear predictions are used for calculating the counterfactual outcomes.

Analogous to the previous subsection, we account for the age-specific characteristics effect, the age-specific coefficients effect and the age composition effect. The probit model is estimated with age specific dummy coefficients $\hat{\alpha}_g^a$ ($g$ gender). Based on the derivation in appendix A.1.2 we obtain the following decomposition:

$$Y_f - Y_m \approx \Phi_f - \Phi_m = \sum_{a=25}^{50} h_m^a \left[ \Phi \left( \hat{\alpha}_f^a + X_f^a \hat{\beta}_f \right) - \Phi \left( \hat{\alpha}_m^a + X_m^a \hat{\beta}_m \right) \right]$$

(4)

Age Specific Coefficients Effect

$$+ \left[ \Phi \left( \hat{\alpha}_f^a + X_f^a \hat{\beta}_f \right) - \Phi \left( \hat{\alpha}_f^a + X_m^a \hat{\beta}_f \right) \right] + \left[ \frac{h_f^a - h_m^a}{h_m^a} \Phi \left( \hat{\alpha}_f^a + X_f^a \hat{\beta}_f \right) \right]$$

Age Specific Characteristics Effect

Age Composition Effect
Here, $\Phi_g$ denotes the sample averages of the fitted probabilities by gender and $\Phi(\hat{\alpha}_a + X_a^g \hat{\beta})$ is the sample mean for age $a$ based on coefficient vector $(\hat{\alpha}_a, \hat{\beta})$. The decomposition in equation (4) is analogous to the decomposition for the training duration in equation (3).

Because of the underlying probit model, we suggest to estimate standard errors using a weighted bootstrap approach instead of a standard bootstrap (for a formal description of the basic weighted bootstrap, see Barbe and Bertail, 1995). The standard errors and confidence intervals are calculated based on 1,000 bootstrap replications. The weighted bootstrap avoids numerical problems as the procedure prevents resamples with perfect collinearity of regressors or perfect predictions caused by the use of dummy variables. Both problems would occur for a standard pairwise bootstrap when accidentally only observations are drawn into the resample for which the dependent variable is zero (or one) for all observations in certain cells (such that the probit would try to predict an exact zero [or an exact one] for such observations) or one of the explanatory dummy variables are all zero (or one) for all observations.

Technically, our weighted bootstrap procedure assigns a randomly drawn weight to each individual to include all (weighted) observations into the estimation instead of drawing only a subset of observations, which are drawn at least once, as down in the standard pairwise bootstrap. Specifically, for the weighted bootstrap, we draw the weights from a uniform distribution on the interval $[0, 2]$. Therefore, the weights have a mean of one and a variance of $1/3$. As drawing from this interval underestimates the variance by a factor of three, the obtained bootstrap variance-covariance matrix has to be multiplied by three.

We suggest to assign the same weight for one individual over time, which implies that standard errors are clustered at the individual level.

---

12 Note that the sample average of the fitted probabilities based on the sample coefficients ($\Phi_g$) only approximates the sample training rate ($\overline{Y}^a$).
5 Results

The results of the decomposition are presented in three steps. First, we show the age-specific decomposition outcomes for training duration and training incidence. Second, we exclude individuals on parental leave to assess how the training gap and the age profile of the decomposed effect changes once child birth and leave periods are considered. Third, we repeat the first two decompositions including supervisor fixed effects in training of their subordinates.

5.1 Estimates of the Gender Training Gap

The decomposition starts by estimating the underlying regressions for training participation and training duration separately for males and females. We include yearly age dummies, controls for individual and job characteristics (as described in Table 1), as well as dummies for years and missing values. In addition, we include interactions between the covariates and age to obtain a sufficiently flexible fit to the data.\(^{13}\) The idea behind this is, that age is a proxy for the career development, and different employees (with regard to e.g. education, wage, tenure) will take different career and training decisions. For example, employees with a university degree will probably invest more into training in their early years with the company compared to their colleagues who already spend the practical part of their dual vocational training already in the company (see Section 2). The regression results are provided as Tables AOA.1 and AOA.2 in the online appendix.

The results show that education and age, besides gender, are the most important factors for training participation and duration, besides gender. Both the probability to participate in workplace training and the duration of training decrease severely with age. Regarding economic significance, the effects increase the older the employees are. Compared to the under 25-year-olds, male employees aged 50 participate three days less in training per year, female employees 2.7 days less. Furthermore, our results suggest that effects

\(^{13}\)The interaction terms include linear and quadratic interactions between age and the other characteristics (education, wage, tenure, full-time, job title, supervisor), age and the missing dummies as well as between age and the year dummies.
are not the same for males and females. As this cannot easily be seen when comparing separate regressions, we perform a joint regression using the same specification and add a gender dummy and gender-age dummies in five year intervals. Results are displayed in Table A.3.1. The interaction terms on gender and age reveal age specific differences between males and females. Females who are older than 41 are more likely to participate in firm-provided training than males in their age group. For training duration, this effect is present at all ages. However, the gender dummy is larger and negative, resulting in the training gap as shown in Table 1. We conclude that females and males exhibit different age profiles in their training participation. A simple comparison of mean outcomes as shown in Table 1 reveals the gender differences, but it cannot show the development by age. Henceforth, our decomposition takes that into account by estimating age specific coefficients and characteristics effects.

We apply the decomposition in Equation (3) for training duration and in Equation (4) for training incidence. The actual and counterfactual outcomes used to calculate the three effects are displayed in Fig. 1.\textsuperscript{14} The triple-dashed line at the top is the counterfactual male outcome, i.e. the duration if males had female coefficients, and the dotted line at the bottom is the counterfactual female outcome, i.e. the female duration reweighted by the relative age-composition difference between males and females.\textsuperscript{15} The figure shows that males would do more days of training, especially from age 37 onwards if assigned female coefficients (triple-dashed line). This indicates a coefficient effect which rises with age. Similarly, the counterfactual reweighted female duration (CharF-CoeffF-AgeM, dotted line) lies above the actual female outcome for females younger than age 37 and below for females from age 37 onwards. The rise in the share of male workers by age results in a positive (negative) effect below (above) age 37.

\textsuperscript{14}The underlying calculations for this figure (and all following figures) do not weight the outcome with $h_m^{\alpha}$. This allows to directly read the number of training days (the training probability) from the vertical axis. Only the counterfactual outcome for CharF-CoeffF-AgeM is reweighted to accommodate the fact that Equation (3) weights by the age composition of males $h_m^{\alpha}$, where the age composition of females $h_f^{\alpha}$ is needed to calculate the outcome.

\textsuperscript{15} The counterfactual outcome for CharF-CoeffF-AgeM is displayed in a reweighted way to accommodate the fact that Equation (3) weights by the age composition of males $h_m^{\alpha}$. The profile for CharF-CoeffF-AgeM in Fig. 1 is calculated as $\frac{h_f^{\alpha}}{h_m^{\alpha}}X_f^{\beta f}$ (see also Footnote 14).
As displayed in Equation (3), the three effects of interest can be obtained by drawing the differences between the hypothetical and actual outcomes. Fig. 2 displays the results of plotting the values of these three effects and their 95% confidence intervals over age (see Table AOA.3 in the online appendix for further details).

[Figure 2 about here.]

As Fig. 2 shows, the three effects are not constant over age. The coefficients effect (coeff, solid line) reveals that if males were assigned female coefficients, their training duration would rise, driven by an increase in the female coefficients effect starting from age 37 onwards. However, this effect is by far outweighed by the characteristics effect (char, dashed line). When males are assigned female characteristics, training duration sharply declines until age 35 and then stabilizes until age 50. The characteristics of males and females are diverging during the first ten years of their working life (cumulating in a maximal characteristics effect of -0.85 training days at age 35) but apparently not beyond that period. Nevertheless, this difference strongly affects the duration of training. The shape of the coefficients effect suggests that either the females themselves or the firm try to compensate the adverse characteristics effect. The dotted line (comp) denotes the age composition effect showing a positive effect at young age (due to the higher share of females) and a negative effect at older age (due to the lower share of females). The overall gap is widened between ages 25 and 35 due to the sharp increase in the characteristics effect and narrowed from age 37 on due to the increase in the coefficients effect. This suggests that females respond differently in their training investments later in their careers compared to males.

The overall average training gap of -0.38 days (female minus male) can be decomposed by using the weighted sums of the three terms as shows in Equation (3). The characteristics effect largely outweighs the two other terms and accounts for -0.64 days. The average coefficients effect is 0.23 days, and the average composition effect accounts for 0.03 days. As Fig. 2 shows, the age-specific effects are not constant over time. Thus, a standard Blinder-Oaxaca decomposition would miss the age dynamics.
The results for the participation equation are similar although largely insignificant. Fig. 3 shows the decomposed gap in the probability to participate in firm-provided training (see Table AOA.4 in the online appendix for further details). The overall difference of 6 percentage points is on average nearly fully explained by the characteristics effect. The average coefficients effect is close to zero as well as the average composition effect. Again, the age-specific effects are not constant. In particular, the coefficients effect rises substantially and significantly from age 39 onwards. This again suggests that particularly training investments of females change late in the career.

Given the importance of the age-specific characteristics effect, it would be interesting to know which characteristics are the most important ones. This is difficult to judge from the detailed regression results because the specifications involve interaction effects of all covariates with age. The age profiles of covariates (see Fig. AOA.1-AOA.3 in the online appendix) reveal that the working time indicator (part-time employment among females increases strongly between age 25 and 35 and declines strongly afterwards) closely follows the age profile of the age-specific characteristics effect. Furthermore, the gender wage gap and the gender gap in supervisory positions grows up to age 40, which also contributes to the increase of the gender training gap during the early working life but these covariates can not explain the decline in the gender training gap after age 35 (detailed results are available upon request). Note that the positive coefficients effect after age 40 indicates that the impact of the observed age-specific gender gaps in covariates, which are associated with lower training of females below age 40, is mitigated considerably after age 40.

Summing up the results from the first part of the analysis, female training investments are lower, especially in the thirties and forties, because of an evolving gender gap in characteristics between age 25 and 35 and because, relative to males, training of females is postponed to some extent to a later age. The major part of the gender training gap is explained by the age profile of part-time employment among females and by the combina-
tion of a growing gender career gap up to age 40 together with a mitigation of the impact of differences in career related variables after age 40.

5.2 Catching-up after Parental Leave?

One reason why females invest later in training compared to males could be a catching-up process after parental leave. Although both, men and women, probably start with the same investments in training, family formation may cause female career paths and training participation to fall behind those of males. When returning full time to the labor market after an employment interruption or a period of part-time work due to child birth, female employees might find it necessary to engage more into workplace training than males of equal age due to greater training needs.

Assuming that fixing a job and fixing tenure in the job holds effective career development constant, we stipulate that training needs are higher after an employment interruption by female employees compared to male employees of the same age but with higher tenure Bishop (1996). This reflects either a catching-up of workplace relevant knowledge and competencies or a (re-)start of the career as the investment horizon is now expected to be uninterrupted (Bassanini et al., 2007). Hence, gender differences in career paths could result in the higher participation and longer duration in firm–provided training during that stage of professional life as reflected by the positive coefficients effect in Fig. 2 (see also the descriptive evidence in Fig. 1).

As the gender characteristics gap develops between age 25 and 35, it is plausible that diverging female career paths lead to the substantial training investments from the mid-thirties onwards. To analyse whether such a “catching-up” process plays a role in explaining the training behavior over the career, we exclude individuals on parental leave. Fig. 4 shows the results for the decomposed duration of training when individual-year-observations on parental leave are excluded.

---

16Our data records individuals on parental leave whose job in the firm is protected for a maximum of 3 years after childbirth. Individuals on parental leave where included before, as it is possible to work part-time during parental leave by German legislation. About 30% of the women on parental leave within the company work part-time.
During the years 2004 to 2007, we observe 5,637 individuals (5.5%) in parental leave. Nearly all of them are females (98.6%). Only 98 individuals (1.8% of individuals in parental leave) have observations on working time that are non-zero. 277 (5%) take part in training activities during that time. As Fig. 4 shows, excluding these observations flattens the characteristics effect in the first part of the age profile. Thus, the impact of characteristics is lower (the average is -0.51 days per year) compared to the characteristics effect in the non-restricted sample (where the average was -0.64 days per year). As we are mainly excluding observations with zero working time, the impact of the characteristics effect is lower. This occurs because working time is one of the characteristics diverging the most between males and females during the early years of a career (see Fig. AOA.2 in the online appendix). The coefficients (0.25 days on average) and composition effect (0.03 days on average) are slightly more pronounced during that first part of the age profile. However, the development of the three effects is nearly unaffected over the second part of the profiles. In particular from age 41 onwards, the characteristics effect lies at about the same level as before, and the coefficients effect rises as before.

We estimate linear regressions as above to compare the duration of training of the individuals before and after parental leave to females employees not in parental leave (detailed results are available upon request).\footnote{Unfortunately, we have only four years of data, so that we cannot follow females for a longer period after they return. Further, a large part of females is in parental leave during the whole period we observe so that we only have very few observations (147) which return at all during the four years.} The regressions include dummy variables for periods shortly before and shortly after parental leave. The regression results show that females who are going to leave in the next one, two, or three years receive about the same amount of training (1.85 days per year) as other female employees. Looking at observations in the first or second period after return from parental leave reveals that their training is substantially shorter (0.4 days per year). Including dummy variables measuring the years before and after parental leave in the linear regression on training duration for females shows no significant rise in training before the leave but a significant drop after returning.
Thus, females do not seem to accumulate human capital “in advance” anticipating a drop in human capital investment during parental leave but they are not investing strongly directly after returning from parental leave either. We conclude that although observations on parental leave seems to be responsible for diverging characteristics between males and females during the late twenties and early thirties, there is no evidence for catching-up effects that would explain the substantial investment in training of females during the second half of the working life.

5.3 Supervisor Fixed Effects in Employees’ Training

One might suspect that the lower training investment of females is associated with a gender bias of supervisors (Rothstein, 1997; Shakeshaft et al., 1991; Melero, 2004). Since careers of females are more likely to involve parental leave and (periods of) part-time work, supervisors may be biased towards higher workplace training for their male subordinates. Supervisors may prefer employees of the same gender and the own training of the supervisor may reflect his/her training policy. In the following, we explore the effect of the supervisor and his/her characteristics (including gender) on the duration of training received by males and females. To do this, we first estimate supervisor fixed effects by gender of employee, i.e. in the regression explaining an employee’s training we add a dummy variable for each supervisor. In a second step, we analyse how the size of the fixed effects depends upon supervisor characteristics and the gender match between supervisor and employee. As the last step, we again decompose the gender training gap including supervisor fixed effects.

To estimate supervisor fixed effects for training of the subordinates, we estimate the OLS training regressions by gender on the sample of employee observations for which we have information on the supervisor of the employee (see Table AOA.2 in online appendix). We interpret the supervisor fixed effect as the “baseline propensity” of training, which differs by supervisor and by gender of the employee, irrespective of other characteristics.

Note that this is not a panel fixed effects regression for the supervisors’ training investment.
of the employees. We center the explanatory variables around their respective means in the overall sample to make the fixed effects comparable between males and females. The 5,372 supervisors we observe in the sample can have solely male employees, solely female employees, or employees of both genders. Thus, we end up with a (stacked) sample of 22,925 supervisor-year observations, for which we have fixed effects and additional information on the supervisor. We truncate these observations from below and from above (at the 2nd and 98th percentile) to assure that our results are robust with regard to outliers.\textsuperscript{19} The subsequent analysis uses a sample with 22,015 observations.

We are primarily interested in the gender congruence between supervisor and employee, i.e. we investigate if the behavior of male and female supervisors is different towards female and male employees. Table 2 involves a contingency table of average fixed effects by gender of the employee and the supervisor. The average fixed effect for female supervisors is uniformly higher than for male supervisors. The average fixed effect for female employees is higher than for male employees, i.e. \textit{ceteris paribus} supervisors allow a larger number of training days for female employees than for male employees.

Turning to the multivariate analysis, Table 3 reports the results of a linear regression of the fixed effects on the type of fixed effect (referring to male or female employees) and on the characteristics of the supervisor based on supervisor–employee matches.\textsuperscript{20} As we are primarily interested in the relationship between the gender of the supervisor and the gender of the employee, we include a variable for the gender of the supervisor, a dummy for a fixed effect for female employees, $f_{\text{fem}}$ (the reference group involves the fixed effects for males), and an interaction term, which is one if both the supervisor and the employee are female, $f_{\text{fem}} \times b_{\text{fem}}$.\textsuperscript{19}\textsuperscript{20}

\textsuperscript{19}This eliminates 452 female and 458 male observations from the sample.

\textsuperscript{20}Note that the regression is estimated for our sample of supervisors, which explains why the sample size is about one fifth of the entire sample, see Table 1.
The regression results show that all three gender effects are insignificant (the effect for female supervisors is only significant at the 10% level). Thus, supervisor fixed effects do neither depend upon the gender of the employee nor upon the gender congruence between supervisor and employee. A closer look at the other characteristics reveals that especially the training duration of the supervisor influences the size of the fixed effects and, thus, the duration of training of the employees. Employees receive more training when the supervisor’s own training duration is longer. Other supervisor characteristics such as wage, tenure, full-time, and corporate title prove significant as well.

For our decomposition analysis, we attribute the supervisor fixed effect to the age-specific characteristics effect assuming that the supervisor is part of the “job environment” of a worker. The resulting age profiles are displayed in Fig. 5. The decomposition of the training duration is based on the sample of observations for which we have information on the supervisor.21 The profiles of the decomposed effects reveals that the characteristics and the coefficients effect are more pronounced compared to the estimation without fixed effects. These changes are mainly due to a change in the estimates for males in the underlying linear regression leading to different predictions in the counterfactual $CharM-CoefF-AgeM$. However, the shape of the age profile remains unaffected.

To sum up, the results for our particular firm show that supervisors do not treat male and female employees differently. The training of both male and female employees is positively associated with the training duration of the supervisor.

6 Conclusions

Based on personnel records of a large German firm, this paper finds substantial gender and age differences in workplace training, both regarding participation and duration.

---

21 As a robustness check, we estimated the fixed effects regression on the full sample, including individual fixed effects for those observations with missing information on the supervisor. The decomposition yields similar results.
We provide a detailed decomposition analysis of the gender training gap by age with a particular focus on parental leave and supervisors.

Our results show that the gender training gap differs by age. The average gender gap in training duration is 0.38 days per year. Being close to zero at age 25, the gap grows during the first 10 years and peaks at age 35 with females of this age having a 0.75 days shorter training duration per year. After age 35, the gap closes to 0.2 days per year at age 50. Female training investments are lower and seem to take place at a higher age.

These age-related gender differences in training can mainly be attributed to differences in characteristics. A major part of the gender training gap is explained by the age profile of part-time employment among females and by the combination of a growing gender career gap up to age 40 together with a mitigation of the impact of differences in career related variables after age 40. We cannot find evidence for a difference in training right before birth and for a catching-up effect after child birth. In fact, female employees participate less in training shortly after child birth. Thus, we conclude that female employees in general are affected by the age pattern in the gender training gap described above, irrespective of whether they are going to have a child soon.

Furthermore, we investigate the link between the supervisor and the gender training gap. The analysis shows that the gender congruence of employee and supervisor, i.e. that the supervisor is of the same gender as the employee, does not explain gender differences in training. The training of both male and female employees is positively associated with the training duration of the supervisor. Supervisor training could be a proxy for the unobserved training needs in a job environment.

There are important limitations to our findings. First and foremost, the external validity of our results is limited because we use single-firm data. Also, other potentially important variables such as employment and career experiences before entering the firm, the personal circumstances and future career aspirations of the employees, and (expected) productivity are unobserved. However, the administrative character of the data, the possibility to match employees to supervisors as well as comprehensive training data recorded by the company provide valuable insights on formal training at the workplace which cannot be
obtained using household survey data. Furthermore, our analysis is limited by its descriptive nature and might be affected selection and endogeneity effects. However, the detailed results show how the gender training gap is associated with individual characteristics as well as with the gender and training behavior of the supervisor, which is of interest by itself. Finally, with four years of data, we can not exclude the possibility that some of the results regarding the age profiles in training are driven by cohort effects. While we suspect that the age profile of training for males has not changed strongly across cohorts, it is not inconceivable that the age profile may have changed for females. In general, the labor force attachment of females has increased for younger cohorts, thus one would expect that females of younger cohorts invest more into training. However, the strong extensions of parental leave periods in Germany from the 1970s to the 1990s may actually reduce the incentive for training investment into female employees who are highly attached to the labor market but who plan to have a child (see Puhani and Sonderhof (2011) for evidence supporting this line of argument). Thus, a priori it is not clear in what direction cohort effects could bias our decomposition results.

Overall, our results show that gender differences exist in workplace training. Our key insights are that the gender training gap depends on age, and the training participation is influenced by the training behavior of the supervisor. To actively support the career development of women, especially when they work part-time or when their career has not developed up to their potential, could be a useful means to narrow the gender training gap.
References


Quintini, Glenda (2011), ‘Right for the job: Over-qualified or under-skilled?’, OECD Social, Employment and Migration Working Papers No. 120.


Figure 1: Age profile of actual and hypothetical training durations per year

CharM-CoefF-AgeM is the counterfactual duration, if males had female coefficients. CharF-CoefF-AgeM is the counterfactual duration of females reweighted to the share of males in the respective age group \(a\), see Footnotes 14 and 15 for a detailed explanation.
Figure 2: Age profile of gender decomposition in training days per year

The coefficients effect (coeff) is indicated by the solid line, the characteristics effect (char) by the dashed line, and the composition effect (comp) by the dotted line. The confidence intervals are depicted by the grey shades, with the respective upper (95up) and lower (95low) confidence bounds in the same shape as the graphs for the effects.
Figure 3: Age profile of gender decomposition in training incidence

The coefficients effect (coeff) is indicated by the solid line, the characteristics effect (char) by the dashed line and the composition effect (comp) by the dotted line. The confidence intervals are depicted by the grey shades, with the respective upper (95up) and lower (95low) confidence bounds in the same shape as the graphs for the effects.
Figure 4: Age profile of gender decomposition in training duration (excluding individuals on parental leave)

The coefficients effect for the sample excluding individuals on parental leave (PL) is indicated by the solid line (coeff_noPL), the characteristics effect (char_noPL) by the dashed line and the composition effect (comp_noPL) by the dotted line.
Figure 5: Age profile of gender decomposition in training duration (including supervisor fixed effects)

The coefficients effect for the estimation with fixed effects (FE) is indicated by the solid line (coeff_FE), the characteristics effect (char_FE) by the dashed line and the composition effect (comp_FE) by the dotted line.
Table 1: Means of variables by gender

<table>
<thead>
<tr>
<th>variables</th>
<th>males</th>
<th>females</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>outcome variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>participation rate</td>
<td>0.58</td>
<td>0.52</td>
</tr>
<tr>
<td>training days per year</td>
<td>2.19</td>
<td>1.81</td>
</tr>
<tr>
<td><strong>explanatory variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>38.06</td>
<td>37.45</td>
</tr>
<tr>
<td>school degree (in %)</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>vocational degree (in %)</td>
<td>0.34</td>
<td>0.51</td>
</tr>
<tr>
<td>university degree (in %)</td>
<td>0.51</td>
<td>0.31</td>
</tr>
<tr>
<td>wage (in 1,000 Euro)</td>
<td>58.37</td>
<td>38.76</td>
</tr>
<tr>
<td>tenure (in years)</td>
<td>14.86</td>
<td>15.44</td>
</tr>
<tr>
<td>working part-time (in %)</td>
<td>0.05</td>
<td>0.45</td>
</tr>
<tr>
<td>corporate title (in %)</td>
<td>0.73</td>
<td>0.41</td>
</tr>
<tr>
<td>supervisor (in %)</td>
<td>0.20</td>
<td>0.06</td>
</tr>
</tbody>
</table>

| N                              | 47,096 | 54,793  |
Table 2: Contingency table of duration fixed effects

<table>
<thead>
<tr>
<th>fixed effects</th>
<th>employees</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>supervisor</td>
<td>males</td>
<td>females</td>
</tr>
<tr>
<td>male</td>
<td>2.13</td>
<td>2.17</td>
</tr>
<tr>
<td>female</td>
<td>2.17</td>
<td>2.26</td>
</tr>
</tbody>
</table>
Table 3: Regression of fixed effects on gender of employee and supervisor characteristics

<table>
<thead>
<tr>
<th>variables</th>
<th>coefficient</th>
<th>(std. error)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>gender employee</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fem</td>
<td>0.020</td>
<td>(0.029)</td>
</tr>
<tr>
<td>fem × b fem</td>
<td>0.077</td>
<td>(0.059)</td>
</tr>
<tr>
<td><strong>supervisor characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>-0.104</td>
<td>(0.060)</td>
</tr>
<tr>
<td>dur. training</td>
<td>0.112</td>
<td>(0.005)</td>
</tr>
<tr>
<td>voc. degree</td>
<td>0.066</td>
<td>(0.121)</td>
</tr>
<tr>
<td>univ. degree</td>
<td>0.082</td>
<td>(0.118)</td>
</tr>
<tr>
<td>wage (in 1,000 Euro)</td>
<td>-0.010</td>
<td>(0.001)</td>
</tr>
<tr>
<td>tenure</td>
<td>0.026</td>
<td>(0.003)</td>
</tr>
<tr>
<td>full-time</td>
<td>-0.465</td>
<td>(0.096)</td>
</tr>
<tr>
<td>corporate title</td>
<td>0.658</td>
<td>(0.066)</td>
</tr>
<tr>
<td>missing dummies</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>1.581</td>
<td>(0.137)</td>
</tr>
</tbody>
</table>

N: 22,015
A Appendix

A.1 Age Specific Decomposition

The aim of our econometric analysis is to investigate the gender gap in training probability (incidence) and in training duration by decomposing the total gap into the age-specific differences by gender and the differences in the age decomposition.

A.1.1 OLS: Training Duration

The decomposition of training duration is based on separate linear regressions for males and females. We consider individuals from age 25 to age 50 grouped in 26 age cells. These regressions are fully saturated by dummy variables for each year of age, i.e. the regression include a dummy variable \(D_{a,i}, a = 25, \ldots, 50\) for each of the 26 age cells. The two groups males \((m)\) and females \((f)\) and the difference in the mean outcomes of females \((\bar{Y}^f)\) and males \((\bar{Y}^m)\) are of interest to us.

The regression model for gender \(g (g \in \{m, f\})\) and worker \(i\) is

\[
y_i^g = \sum_{a=25}^{50} D_{a,i}^g \alpha_a^g + x_i^g \beta^g + u_i = z_i^g \gamma^g + u_i, \tag{5}
\]

where \(D_{a,i}\) denote the age dummy variables and \(x_i^g\) are the remaining regressors, which could also depend upon age \(a\) through interaction effects. The full regressor vector is given by \(z_i^g = (D_{25,i}^g, \ldots, D_{50,i}^g, x_i^g)\), and \(\gamma^g = (\alpha_{25}^g, \ldots, \alpha_{50}^g, \beta^g)\) is the full coefficient vector. We estimate the model without intercept.

Let \(N^g\) denote the number of observations for gender \(g\). Then, the gender specific averages of the regressors and the dependent variable are given by \(h_a^g = \sum_{i \in g} D_{a,i}^g / N^g\) and \(\bar{W}^g = \sum_{i \in g} w_i^g / N^g\) (with \(W \in \{X, Z, Y\}\)), respectively. \(h_a^g\) represents the age distribution by gender.

**Standard Blinder-Oaxaca-Decomposition:** The difference in mean outcomes in the entire sample can be decomposed as follows

\[
\bar{Y}^f - \bar{Y}^m = \bar{Z}^f \hat{\gamma}^f - \bar{Z}^m \hat{\gamma}^m = \bar{Z}^m (\hat{\gamma}^f - \hat{\gamma}^m) + (\bar{Z}^f - \bar{Z}^m) \hat{\gamma}_f, \tag{6}
\]

where the first term denotes the coefficients effect and the second term denotes the characteristics effect. As it is well known, there are different ways to implement the decomposition depending upon which counterfactual is used. Here, and in the following, we use the counterfactual based on male characteristics and female coefficients.
Accounting for age, the decomposition in equation (6) can be rewritten as

\[ \bar{Y}_f - \bar{Y}_m = \sum_{a=25}^{50} h_a^m \left[ \hat{\alpha}_a^f - \hat{\alpha}_a^m + \bar{X}_a^m (\hat{\beta}_f - \hat{\beta}_m) \right] \] (7)

\[ + (\sum_{a=25}^{50} h_a^f \bar{X}_a^f - \sum_{a=25}^{50} h_a^m \bar{X}_a^m) \hat{\beta}_f + \sum_{a=25}^{50} (h_a^f - h_a^m) \hat{\alpha}_a^f \]

where the right-hand-side in the first line corresponds to \( \bar{Z}^m (\hat{\gamma}_f - \hat{\gamma}_m) \) (coefficients effect) and the expression in the second line corresponds to \( (\bar{Z}_f - \bar{Z}_m) \hat{\beta}_f \) (characteristics effect) in equation (6). Here, \( \bar{X}_a^g \) denote the age-specific mean of regressor \( x_i^g \) by gender \( g \).

**Standard Blinder-Oaxaca- Decomposition by age**: The first order conditions of OLS estimation imply that the regression model (5) fits exactly the age specific mean \( \bar{Y}_a^g \), i.e.

\[ \bar{Y}_a^g = \hat{\alpha}_a^g + \bar{X}_a^g \hat{\beta}_g, \]

where \( \bar{Y}_a^g \) denote the age-specific mean of regressor \( y_i^g \) by gender \( g \). This is because the model includes age-specific dummy variables.

Using this result, the age-specific mean gender gap can be decomposed by a standard Blinder-Oaxaca decomposition as follows

\[ \bar{Y}_a^f - \bar{Y}_a^m = \hat{\alpha}_a^f + \bar{X}_a^f \hat{\beta}_f - \hat{\alpha}_a^m - \bar{X}_a^m \hat{\beta}_m \]

\[ = \left[ \hat{\alpha}_a^f - \hat{\alpha}_a^m + \bar{X}_a^m (\hat{\beta}_f - \hat{\beta}_m) \right] + \left[ (\bar{X}_a^f - \bar{X}_a^m) \hat{\beta}_f \right], \]

where, in the last line, the first term represents the age specific coefficients effect including the coefficient of the age dummy and the second term afterwards represents the age specific characteristics effect.

**Total Blinder-Oaxaca- Decomposition by age**: Now, we show how the total Blinder-Oaxaca decomposition in (7) can be expressed as the sum of the age profiles of the age specific decomposition and an age-composition effect. By adding and subtracting \( \sum_{a=25}^{50} h_a^m \bar{X}_a^m \hat{\beta}_f \) in the second line of (7) and by expanding the numerator and denominator of \( (h_a^f - h_a^m) \) with \( h_a^m \), the difference in average outcomes can be written as

\[ \bar{Y}_f - \bar{Y}_m = \sum_{a=25}^{50} h_a^f \bar{Y}_a^f - \sum_{a=25}^{50} h_a^m \bar{Y}_a^m = \] (8)
This decomposition links the standard age specific Blinder Oaxaca decomposition and the total decomposition taking account of the gender differences in the age composition. Note that the first sum in the total decomposition (weighted average of age specific coefficients effects) corresponds to the total coefficients effect in the standard Blinder Oaxaca decomposition (6). The weighted average of the remaining two terms correspond to the total characteristics effect in (6).

### A.1.2 Probit: Training Probability

Analogous to the OLS case in the previous section, we now derive our decomposition of the gender in the training probability based Fairlie (2005) decomposition for a probit model.

The probit model for gender $g$ and worker $i$ (analogous to equation 5) is

$$
P(y_i^g = 1 | a, x_i^g) = \Phi \left( \sum_{a=25}^{50} D_{a,i}^g \alpha_a^g + x_i^g \beta^g \right) = \Phi \left( z_i^g \gamma^g \right),$$

where $\Phi(.)$ denotes the standard normal distribution function. Define $\Phi(Z^g \hat{\gamma})$ as the sample mean for gender $g$ of the fitted probabilities based on coefficient vector $\hat{\gamma}$ in the probit model, i.e.

$$\bar{\Phi}(Z^g \hat{\gamma}) = \frac{1}{N_g} \sum_{i \in g} \Phi \left( \sum_{a=25}^{50} D_{a,i}^g \hat{\alpha}_a + x_i^g \hat{\beta} \right) = \sum_{a=25}^{50} h_a^g \bar{\Phi}(\hat{\alpha}_a + X_a^g \hat{\beta}),$$

where $\bar{\Phi}(\hat{\alpha}_a + X_a^g \hat{\beta})$ is the sample mean for age $a$ based on coefficient vector $(\hat{\alpha}_a, \hat{\beta})$.

In contrast to OLS, the sample means of the fitted probabilities based on the gender specific probit coefficients correspond only approximately to the training rate in the samples $Y^g \approx \Phi(Z^g \hat{\gamma})$. This also holds for the age specific means, i.e. $Y_a^g \approx \Phi(\hat{\alpha}_a + X_a^g \hat{\beta})$.

The approximation in the entire sample is usually very good. The same holds here by age because specification (9) includes age specific dummy variables.

Analogous to the previous subsection, we account for the age-specific characteristics effect, the age-specific coefficients effect and the age composition effect. This leads to the following decomposition:
\[ Y^f - Y^m \approx \Phi_f - \Phi_m = \Phi(Z_f \gamma_f) - \Phi(Z_m \gamma_m) \] (10)

\[
= \sum_{a=25}^{50} h_a^m \left[ \Phi(\hat{\alpha}_a + X_a^m \beta_f) - \Phi(\hat{\alpha}_a + X_a^m \beta_m) \right] + \left[ \Phi(\hat{\alpha}_a + X_a^f \beta_f) - \Phi(\hat{\alpha}_a + X_a^m \beta_f) \right] \\
+ \left[ \frac{h_f^f \cdot h_m^m}{h_f^m} \cdot \Phi(\hat{\alpha}_a + X_a^f \beta_f) \right].
\]

The derivation follows analogous to the decomposition of the average fitted probabilities by gender analogous to the derivation for OLS in the previous subsection. This is based on the standard Fairlie (2005) decomposition.
Table A.3.1: Regression of training incidence and duration on age profile and gender

<table>
<thead>
<tr>
<th>variable</th>
<th>training participation (Probit)</th>
<th>training days (OLS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coefficient</td>
<td>std. err.</td>
</tr>
<tr>
<td>d_age26</td>
<td>0.229</td>
<td>0.145</td>
</tr>
<tr>
<td>d_age27</td>
<td>0.037</td>
<td>0.145</td>
</tr>
<tr>
<td>d_age28</td>
<td>-0.074</td>
<td>0.146</td>
</tr>
<tr>
<td>d_age29</td>
<td>-0.185</td>
<td>0.147</td>
</tr>
<tr>
<td>d_age30</td>
<td>-0.299</td>
<td>0.149</td>
</tr>
<tr>
<td>d_age31</td>
<td>0.089</td>
<td>0.150</td>
</tr>
<tr>
<td>d_age32</td>
<td>-0.001</td>
<td>0.151</td>
</tr>
<tr>
<td>d_age33</td>
<td>-0.050</td>
<td>0.152</td>
</tr>
<tr>
<td>d_age34</td>
<td>-0.141</td>
<td>0.153</td>
</tr>
<tr>
<td>d_age35</td>
<td>-0.144</td>
<td>0.155</td>
</tr>
<tr>
<td>d_age36</td>
<td>-0.012</td>
<td>0.155</td>
</tr>
<tr>
<td>d_age37</td>
<td>-0.076</td>
<td>0.155</td>
</tr>
<tr>
<td>d_age38</td>
<td>-0.100</td>
<td>0.156</td>
</tr>
<tr>
<td>d_age39</td>
<td>-0.151</td>
<td>0.157</td>
</tr>
<tr>
<td>d_age40</td>
<td>-0.153</td>
<td>0.159</td>
</tr>
<tr>
<td>d_age41</td>
<td>-0.329</td>
<td>0.161</td>
</tr>
<tr>
<td>d_age42</td>
<td>-0.380</td>
<td>0.162</td>
</tr>
<tr>
<td>d_age43</td>
<td>-0.386</td>
<td>0.163</td>
</tr>
<tr>
<td>d_age44</td>
<td>-0.391</td>
<td>0.165</td>
</tr>
<tr>
<td>d_age45</td>
<td>-0.385</td>
<td>0.168</td>
</tr>
<tr>
<td>d_age46</td>
<td>-0.672</td>
<td>0.172</td>
</tr>
<tr>
<td>d_age47</td>
<td>-0.698</td>
<td>0.173</td>
</tr>
<tr>
<td>d_age48</td>
<td>-0.681</td>
<td>0.175</td>
</tr>
<tr>
<td>d_age49</td>
<td>-0.666</td>
<td>0.178</td>
</tr>
<tr>
<td>d_age50</td>
<td>-0.690</td>
<td>0.183</td>
</tr>
<tr>
<td>female</td>
<td>-0.020</td>
<td>0.065</td>
</tr>
<tr>
<td>fage2630</td>
<td>-0.029</td>
<td>0.069</td>
</tr>
<tr>
<td>fage3135</td>
<td>-0.058</td>
<td>0.070</td>
</tr>
<tr>
<td>fage3640</td>
<td>-0.056</td>
<td>0.070</td>
</tr>
<tr>
<td>fage4145</td>
<td>0.069</td>
<td>0.071</td>
</tr>
<tr>
<td>fage4650</td>
<td>0.117</td>
<td>0.073</td>
</tr>
<tr>
<td>voc. degree</td>
<td>-0.780</td>
<td>0.136</td>
</tr>
<tr>
<td>tertiary degree</td>
<td>-0.837</td>
<td>0.368</td>
</tr>
<tr>
<td>wage (in 1,000 Euro)</td>
<td>-0.014</td>
<td>0.004</td>
</tr>
<tr>
<td>tenure</td>
<td>0.234</td>
<td>0.014</td>
</tr>
<tr>
<td>full-time</td>
<td>-2.567</td>
<td>0.475</td>
</tr>
<tr>
<td>corporate title</td>
<td>-0.751</td>
<td>0.070</td>
</tr>
<tr>
<td>supervisor</td>
<td>-0.324</td>
<td>0.717</td>
</tr>
</tbody>
</table>

interaction terms: yes
missing dummies: yes
year dummies: yes
interaction age year: yes
constant: -0.001

N: 101,889

Standard Errors are clustered at the individual level.
The reference individual is 25 years old, holds a vocational degree, works fulltime, and does neither have a corporate title nor a supervisory role.