DISCUSSION PAPER SERIES

IZA

IZA DP No. 3533

Analysing the Gender Wage Gap Using Personnel Records of a Large German Company

Christian Pfeifer
Tatjana Sohr

June 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Analysing the Gender Wage Gap Using Personnel Records of a Large German Company

Christian Pfeifer
Leibniz University Hannover and IZA

Tatjana Sohr
Leibniz University Hannover

Discussion Paper No. 3533
June 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3533
June 2008

ABSTRACT

Analysing the Gender Wage Gap Using Personnel Records of a Large German Company*

We use monthly personnel records of a large German company to analyse the gender wage gap (GWG). Main findings are: (1) the unconditional GWG is 15 percent for blue-collar and 26 percent for white-collar workers; (2) conditional on tenure, entry age, schooling, and working hours, the GWG is 13 percent for blue-collar as well as for white-collar workers; (3) after additionally controlling for hierarchical levels, the GWG is less than 4 percent for blue-collar and 8 percent for white-collar workers; (4) Oaxaca decompositions reveal that the unexplained part of the GWG is 87 percent for blue-collar workers and 46 percent for white-collar workers; (5) males have larger absolute wage growths than females; (6) the relative GWG gets larger with tenure for blue-collar but smaller for white-collar workers; (7) individual absenteeism has no significant impact on the GWG; (8) the gender gap in absenteeism is between 26 and 46 percent. Overall, the results are consistent with statistical discrimination explanations of the gender wage gap, though we cannot rule out other forms of discrimination. A simple model within the context of absenteeism and statistical discrimination is offered.

JEL Classification: J16, J3, J71, M5

Keywords: absenteeism, gender, personnel data, statistical discrimination, wage differentials

Corresponding author:

Christian Pfeifer
Institute of Labour Economics
Leibniz University Hannover
Königsworther Platz 1
30167 Hannover
Germany
E-mail: pfeifer@aoek.uni-hannover.de

* The authors would like to thank Knut Gerlach and participants at the Colloquium in Personnel Economics 2008 in Bonn and at research seminars at Aarhus School of Business and Free University Berlin for helpful comments. Any remaining shortcomings are, of course, our own. Financial support from the Fritz Thyssen Foundation and the Federal State of Lower Saxony are gratefully acknowledged.

1. Introduction

Research on wages of men and women report consistently a gender wage gap (GWG in the following) unconditional and conditional on important control variables as well as for different countries, industries, and decades (e.g., Cain, 1986; Gunderson, 1989; Altonji and Blank, 1999; Blau and Kahn, 2000; Blau and Kahn, 2003). Neumark (1999, p. 414) states: “There are two dominant explanations of these wage differentials that are the basis for most empirical research on this topic. The first is that employers discriminate against women or minorities in some fashion that results in lower wages for them even when they are equally productive. The second is that women or minorities come to the market with productivity shortfalls.” Discrimination can take several forms (Cain, 1986). For example, a firm might pay females lower wages than males because of employer, employee, or customer taste discrimination (Becker, 1971). Another form of discrimination is the use of group statistics (e.g., lower average productivity of women) as a proxy for unobserved productivity. Such statistical discrimination, however, is hard to measure and “has received relatively little empirical attention with regard to wage differentials by race and sex” (Neumark, 1999, p. 416).

Several studies note that the GWG is smaller under piece-rate than under time-rate based remuneration schemes (Jirjahn and Stephan, 2004; Petersen, Snartland, and Meyersson Milgrom, 2005). As wages largely reflect productivity under piece-rates, the GWG rather reflects differences in productivity than discrimination. Under time-rates all sources of discrimination might add to the initial GWG due to productivity differences. However, this kind of studies cannot distinguish between different types of taste discrimination and statistical discrimination. Other studies looked at entry wages and the evolution of the wage gap (Neumark, 1999; Altonji and Pierret, 2001). Statistical discrimination predicts that the impact of the group variable (gender or race) gets smaller with tenure because the employer learns about individual productivity. The results of Altonji and Pierret (2001), however, are not very supportive of this prediction because the wage gap between blacks and whites is virtually zero when entering the labour market and rises with experience.

This paper adds a new perspective to the explanation of the persistence of the GWG by looking at monthly personnel records of a large German company from January 1999 to

December 2005.[1] As a special feature, these personnel records include information on contractual as well as on absent (sickness) working hours. The latter might be used as a proxy for productivity. However, since the firm has to pay the contractual working hours regardless of absenteeism, it seems reasonable to interpret absenteeism as costs.

We find that the GWG is remarkably reduced but still significant after controlling for differences in worker characteristics (tenure, entry age, schooling, and monthly working hours). Moreover, Oaxaca decompositions reveal that a substantial part of the GWG remains unexplained. Analysing wage careers and the evolution of the GWG, we find that absolute wage growth is larger for males than females. The relative GWG, however, gets smaller with tenure for white-collar workers but larger for blue-collar workers. This difference might be explained by the firm's training decisions ('self-fulfilling prophecy') and the type of skills acquired during training, which differ between production (firm specific skills) and administration work (general skills). Including individual absenteeism in the estimates does not reduce the GWG significantly. The sizable GWG might be explained by statistical discrimination against women due to their higher average absenteeism rates, which have also a higher variance than among men, – a point that was also raised in previous studies (Corcoran and Duncan, 1979; Osterman, 1979; Kahn, 1981; Osterman, 1981). We explain this finding partly in a simple microeconomic model, which combines the cost aspect of absenteeism with gender based statistical discrimination by a profit-maximizing firm. Policy implications are drawn from the empirical and theoretical findings.

The subsequent paper is structured as follows: Section 2 gives a literature review. In Section 3, the personnel data set and the firm are briefly described, which is accompanied by descriptive statistics about employment, wages, and absenteeism. The regression analyses follow in Section 4, which contains estimates of the GWG, an

[1] Earlier studies examining personnel records of a single firm are Malkiel and Malkiel (1979), who analysed small cross section samples of high qualified professionals in the late 1960s, and Osterman (1979), who analysed sex discrimination in a cross section of professional employees from 1976. Ichino and Moretti (2006) recently analysed absenteeism and wages with personnel data of an Italian bank. Their research focus, however, was quite special (menstrual cycle) and different from ours. Dohmen, Lehmann, and Zaiceva (2008) use personnel records of a Russian firm to analyse the effect of transition on the GWG. Absenteeism is not incorporated in their study.

Oaxaca decomposition of the GWG, gender differences in wage careers, evolution of the GWG over tenure, wage effect of absenteeism, and gender gap in absenteeism. Section 5 presents a simple economic model, which combines the wage cost aspect of absenteeism with statistical discrimination, and gives some policy implications.

2. Literature Review

Women in Germany – and not just here – still earn significantly less than men, even though the difference diminished clearly as studies with the German socio-economic panel (GSOEP) (Prey, 1999) and the German IAB employment sample (Fitzenberger and Wunderlich, 2002; Gartner and Rässler, 2005) show. Apart from the influence of individual characteristics, research has put especially the effects from employer discrimination on the GWG to the fore. On the one hand, discrimination of women can be the consequence of preferences and incomplete competition. If there is a 'taste for discrimination' the management or the employees of the firm prefer to cooperate with men (Becker, 1971; Blau, Ferber, and Winkler, 2006). On the other hand, there can be statistical discrimination. In our context, firms use the gender of an employee to predict his or her productivity in case of incomplete information about productivity (Aigner and Cain, 1977). For example, unscheduled absenteeism could have a negative effect on productivity.

Absenteeism is costly for firms due to various reasons (Allen, 1981a). First, the absent worker has to be replaced by someone who is probably less productive. Second, in case of unexpected absenteeism managerial resources have to be shifted from more productive uses to level out the failure. Against this background and proceeding from the assumption that women on average miss their jobs more often and for a longer time a profit-maximizing employer could decide to hire less women or to pay them lower wages.

Most existing studies about absenteeism at the workplace conclude that women are absent more often than men. Mastekaasa and Modesta Olsen (1998) show that there are incentives for statistical discrimination of women due to their higher absence times with data from civil servants in Norway. They find that the stronger responsibilities of women for child care have no influence on differences in absence times between women and men, in contrast to Vistnes (1997) who works with data from the US National Medical Expenditure Study. Corresponding with studies from Leigh (1983) and Paringer (1983), Vistnes (1997) shows that the health status is a more important predictor of absenteeism than economic factors (e.g., granting of sick leave, existence of health insurance). Allen (1981b) analyses the probability of absenteeism in dependence from various economic and socio-demographic characteristics. He concludes that the observed higher absence times of women can be explained almost completely by differences in marginal incomes and working hour flexibility.

One of a few studies on absenteeism in context of the GWG has been carried out by Corcoran and Duncan (1979). They use the ninth wave of the US Panel Study of Income Dynamics (PSID), a household survey which includes detailed information on work history, on-the job-training, and absenteeism. By estimating earnings functions and using the decomposition method developed by Oaxaca (1973), they find that the higher absent working hours of women account for a quite low share of the wage differential between the sexes.

While most studies on absenteeism use survey data, only a few analyse personnel data. Malkiel and Malkiel (1973) estimate earnings functions and use Oaxaca decomposition on a sample of professional employees from a single corporation. They find a negative influence of the absence rate on the wage level, which is significant only for female employees. Osterman (1979) analyses sex discrimination in a large publishing firm in New York City by estimating wage differentials within similar jobs. In his study, the negative influence of absenteeism on wages also turned out as significant only for women. Ichino and Moretti (2006) use personnel data of a large Italian bank to analyse if differences between women and men with regard to absenteeism are due to biological differences (menstrual cycle), which would have consequences for wages and salaries of women. They find that the probability of missing work increases significantly for women relative to men 28 days after the preceding absenteeism spell. The effect disappears for employees older than 45 years. In their model, the authors assume that the employer cannot observe individual productivity directly and instead uses observable characteristics (e.g., absenteeism) to predict productivity. The results confirm the existence of statistical discrimination due to the higher absenteeism rate of women: The GWG which is 13.5 percent in the sample would be reduced by 1.6 percentage points – or 11.8 percent – if the average woman would not suffer from menstrual symptoms, at least in this organization.

For Germany, we find only one study analysing causes of absenteeism on basis of personnel data: Stephan (1991) uses data from more than 10,000 employees of a large German company. She investigates the hypothesis that absenteeism which is not due to illness is used to realize the optimal labour supply. It turns out that individual characteristics – among them gender – have a significant influence on absence times. However, there are hints that especially the higher absenteeism of women could be due to the endowment of their workplaces.

Few international comparative studies analyse the influence from the institutional framework in which absenteeism occurs. Osterkamp and Röhn (2007) construct a panel of 20 OECD countries for the years 1996 to 2002 from the OECD Health Data-base and the WHO Health For All Database. Frick and Malo (2005) use data from the 2000 European Survey on Working Conditions on 12 countries belonging to the EU. Both studies develop an index measuring the 'generosity' of sickness benefits systems. They find this generosity to be an important predictor for the number of days absent.

The German sickness benefits system belongs apart from the Swedish and the Norwegian one to the most 'generous' among OECD and EU countries respectively. Sick pay in Germany is regulated in the Act on Continued Payment of Remuneration ('Lohnfortzahlungsgesetz'). An employee who is sick for more than three days has to present a medical certificate of sickness from his physician. From October 1996 to 1998 the law provided an option to limit the continued pay specified by the law, but in many collective wage agreements the level was fixed to 100 percent. Therefore, this option was cancelled after a change of government in 1998 (Thalmaier, 2002). Since 1999 sick employees have a legitimate claim to 100 percent wage replacement from their employer, for a period up to six weeks. In case of longer sickness absence due to the same disease, there is a claim of 70 percent of the previous regular remuneration for 78 weeks within three years, which is paid by the health insurance. Although the German sickness benefits system is quite generous, we find Germany in the centre of the distribution of absence days (with an average of 6.89 days absent) for the year 2000 (Frick and Malo, 2005). In the 2000 wave of the GSOEP, 50 percent of respondents stated to have been absent due to sickness for at least one day (Statistisches Bundesamt, 2006).

3. Data and Descriptive Statistics

3.1 The Firm

The data set was extracted from computerised personnel records of a large German limited company which produces innovative products for the world market. The company has a works council and is subject to an industry wide collective contract. The personnel records contain information for all employees in the company's headquarter on a monthly basis from January 1999 to December 2005. The used sample includes blue-collar and white-collar workers without apprentices, workers in early retirement schemes, and long-term absent workers (at least one month absent). Moreover, observations with missing values in any of the used variables are dropped. In sum, the total number of observations (individuals) during the 84 month observation period is 50,722 (786) for blue-collar workers and 73,174 (1,250) for white-collar workers in an unbalanced panel design.

3.2 Employment

Figure 1 gives an impression of the firm's overall employment development from January 1999 (month 1) to December 2005 (month 84). Total employment was about 1,300 in early 1999 and increased to 1,600 in 2003. Since then, employment decreased and was about 1,400 in the end of 2005. The firm employs more white-collar than blue-collar workers. Both types of workers follow the same overall employment trend.

- ***Insert Figure 1 about here***

Figure 2 depicts the share of female employment. The share of females in total employment increased from 23 percent in 1999 to 25 percent in 2005. This trend can be observed for blue-collar as well as for white-collar workers. However, the share of females is higher among white-collar workers than among blue-collar workers.

- ***Insert Figure 2 about here***

The firm has a hierarchy according to the wage groups of the collective contract (see Table 1). Table 2 and 3 indicate that females work on average at lower levels. Especially at the highest levels few or no females are employed.

- ***Insert Table 1 about here***
- ***Insert Table 2 about here***
- ***Insert Table 3 about here***

Further information on the characteristics of male and female blue-collar and white-collar workers are summarised in Table 4. Men have on average longer tenure than women. While male blue-collar workers are on average younger than female blue-collar workers when they enter the firm, the opposite is found for white-collar workers. Men have on average more schooling than women. For example, 40 percent of male but only 23 percent of female white-collar workers have a university degree. The significant differences between average male and female worker characteristics make it necessary to control for these variables in regression analyses. A complete list of the variables used in the regression analyses and detailed descriptive statistics can be found in Table A.1 and A.2 in the Appendix.

- ***Insert Table 4 about here***

3.3 Wages

Table 5 informs about average nominal gross hourly wages for male and female workers and the differences between them. Average wages are 14.94 Euros for blue-collar workers and 23.41 Euros for white-collar workers. Male blue-collar (white-collar) workers earn on average 2.16 (5.31) Euros more than female blue-collar (white-collar) workers. The gender differences are highly significant in a two-sample t-test. A further description of the wage distributions of male and female workers can be found in Figures 3 and 4, in which kernel density estimates are presented.

- ***Insert Table 5 about here***
- ***Insert Figure 3 about here***
- ***Insert Figure 4 about here***

The development of mean nominal wages in Euros by months for male and female blue-collar and white-collar workers is given in Figure 5. There is a positive trend with some spikes which indicate the agreed wage increases in collective contracts. Male workers earn on average more than female workers and white-collar workers earn on average more than blue-collar workers. Note that male blue-collar workers also earn less than female white-collar workers.

- ***Insert Figure 5 about here***

The female-male wage ratios are depicted in Figure 6. Female blue-collar workers earn on average between 14 and 15 percent less than male blue-collar workers. While the GWG is quite stable over time for blue-collar workers, the female-male wage ratio is increasing for white-collar workers from 77 percent in 1999 to 80 percent in 2005.

- ***Insert Figure 6 about here***

3.4 Absenteeism

Figure 7 gives an impression of the amount of lost working hours due to absenteeism and the associated wage costs. An advantage of using absent working hours instead of absent working days is that females work on average fewer working hours. Therefore, the associated wage costs for women would be overestimated when using absent working days. The sum of total absent working hours ranges from 4,000 hours to more than 12,000 hours per month. Most of the spikes appear during the winter months (e.g., cold, flu). As a lower boundary for the costs of absenteeism, the wage costs are calculated by multiplying the hourly wage with the number of absent working hours for every worker in every month.[2] The sums of total monthly wage costs due to absenteeism range from 50,000 Euros to 250,000 Euros. To get an impression of the relative magnitude of theses costs, we divided the monthly wage costs due to absenteeism by total monthly wage costs. Figure 8 shows that absenteeism is responsible for 2 to 5 percent of the wage bill. Thus, absenteeism is quite costly for the firm.

- ***Insert Figure 7 about here***
- ***Insert Figure 8 about here***

Table 6 shows that the average numbers of absent working hours per month are 7.14 for blue-collar workers and 3.42 for white-collar workers. Female blue-collar (white-collar) workers are on average 2.20 (0.72) hours more absent than male blue-collar (white-collar) workers. The differences are highly significant in a two-sample t-test. Moreover, it appears that standard deviations are larger for females. Figures 9 and 10 illustrate graphically that means of absenteeism are larger for females than for males. This finding gives rise to the issue of statistical discrimination, which is discussed later.

- ***Insert Table 6 about here***
- ***Insert Figure 9 about here***
- ***Insert Figure 10 about here***

4. Regression Analyses

4.1 Gender Wage Gap

To exploit the panel character of our data set, we estimate log-linear earnings functions with random effects GLS (general least squares) models and heteroskedasticity robust

[2] Additional costs of absenteeism might occur, for example, due to replacement costs (e.g., temporary help service) and quasi-fixed employment cost (e.g., investments in firm-specific human capital).

standard errors.[3] The dependent variable is the logarithm of hourly gross wages in Euros in a given month. Female and month dummies are the only explanatory variables in our first specification. The results in the second columns of Table 7 and 8 indicate a highly significant GWG of approximately 15 percent for blue-collar workers and 26 percent for white-collar workers. Including additional variables to control for tenure, entry age, schooling, and working hours reduces the GWG to approximately 13 percent for blue-collar as well as for white-collar workers (see Pfeifer 2008 for a discussion on the impact of the different control variables). After additionally controlling for hierarchical levels, the GWG is still significant but reduced to less than 4 percent for blue-collar workers and 8 percent for white-collar workers.

- ***Insert Table 7 about here***
- ***Insert Table 8 about here***

While the inclusion of tenure, entry age, schooling, and working hours reflects the importance to control for differences in worker characteristics between men and women, the results for hierarchical levels need a short discussion. After levels are included in the estimates, female wages are still lower than male wages but the GWG is remarkably reduced. This finding indicates that females self-select into jobs at lower levels and the firm selects them less often to higher levels (Winter-Ebmer and Zweimüller, 1997; Pfeifer, 2007). It is consistent with results from previous studies, which report that segregation of men and women in different occupations account for a large part of the GWG (e.g., Meyersson Milgrom, Petersen, and Snartland, 2001; Datta Gupta and Rothstein, 2005; Dohmen, Lehmann, and Zaiceva, 2008). Since different jobs of females and males might already reflect discrimination, the total GWG is underestimated if the estimates control for jobs (e.g., Gunderson, 1989; Kidd and Shannon, 1996; Barnet-Verzat and Wolff, 2008). Thus, we use specifications without levels in most of the subsequent estimates.

[3] The Breusch and Pagan Lagrange multiplier test for random effects shows that the random effects model is more appropriate than OLS because the variance of the individual error term is significantly different from zero.

4.2 Decomposition of the Gender Wage Gap

In this section, we use the decomposition method introduced by Oaxaca (1973), which decomposes the wage differential in two components. The first component comprises differences in the explanatory variables (explained effect), and the second one measures the effect of a different treatment of men and women with otherwise equal characteristics (unexplained effect). As reference group for weighting differences in individual characteristics we choose men. This implies the assumption that the male wage structure would arise if there were no discrimination. The decomposition equation then results as follows in (1):

$$(1) \qquad \overline{\ln(w_M)} - \overline{\ln(w_F)} = \underbrace{(\overline{X'_M} - \overline{X'_F})\hat{\beta}_M}_{\textit{explained differential}} + \underbrace{(\hat{\beta}_m - \hat{\beta}_F)\overline{X'_F}}_{\textit{unexplained differential}}$$

$\overline{\ln(w_M)}$ is the average log male wage and $\overline{\ln(w_M)}$ is the average log female wage. $\overline{X'_M}$ and $\overline{X'_F}$ are vectors which contain the means of individual characteristics for each gender. $\hat{\beta}_M$ and $\hat{\beta}_F$ are the estimated coefficients. Oaxaca (1973) refers to the unexplained part of the differential as discrimination. But when interpreting the results one must have in mind that parts of the wage differential, which are not explained by individual characteristics, do not necessarily come from discrimination. They could also be due to differences in unobserved characteristics between men and women (Achatz, Gartner, and Glück, 2005). Table 9 and 10 show the estimates for blue-collar and white-collar workers respectively. The first column in each table contains the estimates for men and the second column the estimates for women. Columns three to six present the results of the decomposition.

- ***Insert Table 9 about here***
- ***Insert Table 10 about here***

The rates of return to tenure are in favour of men in both groups, while for entry age this holds only for white-collar workers. One reason might be that firm-specific human capital is more important for blue-collar workers, while administrative work typically requires rather general qualifications. Therefore, higher entry age as a proxy for general skills is advantageous for white-collar but not for blue-collar workers. Apprenticeship training increases wages of male blue-collar workers by 7.8 percent, while we find no significant effect for females. Conversely, the rates of return to high school and university degrees among white-collar workers are larger for females. Moreover, we find no hint for a wage penalty for working fewer hours for blue-collar women, which might reflect a selection of women into jobs suitable for part-time work. In case of white-collar workers, the number of working hours has no effect on women's wages, whereas we find a positive effect for men.

The results of the decomposition support these findings. The number of monthly working hours has almost no influence on the size of the GWG for blue-collar workers. For white collar-workers the differences in the coefficients contribute with about 22 percent to the GWG. We find no significant effect of entry age on the GWG for blue-collar workers. In case of white-collar workers both effects, explained and unexplained, are highly significant, but the explained effect is in contrast to the unexplained effect rather inferior. This also holds after controlling for hierarchical levels, even if the effects decrease as the results in Table 11 show.[4] A reason might be that men and women have different employment histories when entering the firm, with more gaps in female employment histories due to family responsibilities. Corcoran and Duncan (1979) find that a different work history contributes with 28 percent to the GWG in the PSID data.

- ***Insert Table 11 about here***

The coefficients for tenure, entry age, and university degree decrease considerably after controlling for hierarchical levels, which shows the high correlation of these variables. At the same time, the coefficient for monthly working hours increases for both, men and women, and the explained differential becomes significant in the decomposition equation. Differences in mean characteristics have a highly significant influence on the GWG at all levels. We find a slight advantage for women at level 2, while from level 3 upwards there is an advantage for men, which increases with level. Differences in level

[4] As already mentioned in Section 3.2, there are no men working at level 1 and no women working at level 6 and 7 among blue-collar workers. Therefore, we do not carry out estimations with hierarchical levels for blue-collar workers.

coefficients do not contribute to the size of the GWG. This finding supports the inferences from Section 4.1 that females self-select into lower levels and are less often selected into higher levels by the firm.

Table 12 summarises the decomposition results. The explained differential contributes with only 1.9 percent to the GWG for blue-collar workers. Consequently, 87 percent of the total GWG are not explained by differences in individual characteristics. For white-collar workers about 13.8 and 17.5 percent of the GWG, respectively, can be attributed to differences in mean characteristics. As expected from the previous results the part of the GWG due to discrimination decreases substantially when controlling for hierarchical levels.

- ***Insert Table 12 about here***

Because the unexplained part of the GWG remains substantially large after controlling for several individual characteristics (87 percent for blue-collar workers, about 46 and 28 percent for white-collar workers, respectively), it is important to analyse further possible explanations. One explanation might be statistical discrimination due to on average higher absenteeism rates of women. Before focusing on this issue, we take a closer look at different wage careers of men and women and at the evolution of the GWG.

4.3 Wage Careers and Evolution of the Gender Wage Gap

The coefficients of tenure, squared tenure, and cubed tenure[5] from the previous estimates in Section 4.2 are used to predict the wage-tenure profiles in Figures 11 and 12, in which entry wages are predicted at tenure of zero and means of all other covariates. Male blue-collar workers start on average with a predicted wage of 13.5 Euros and female blue-collar workers with wages below 13 Euros. Male white-collar workers start on average also with higher wages (18 Euros) than females (15 Euros). The lower gender gap in entry wages among blue-collar workers might be explained by

[5] Higher order terms of tenure were not significant and did not change the predicted wage profiles.

differences in entry age (see Table 4 in Section 3.2), which is a proxy for acquired general human capital and positively correlated with wages – at least for white-collar workers (see previous estimates). Female blue-collar workers are on average 1.5 years older than males when entering the firm, whereas female white-collar workers are on average 2.5 years younger than males. Except for female blue-collar workers, the wage-tenure profiles are quite steep until 20 years of tenure. Overall, the wage-tenure profiles are steeper for white-collar than for blue-collar workers, which might be explained by deferred compensation schemes that are less important in production jobs (Lazear, 1979). The absolute wage differentials between males and females increase with tenure because males have larger absolute wage growths.

- ***Insert Figure 11 about here***
- ***Insert Figure 12 about here***

To analyse the evolution of the relative GWG over tenure, earnings functions are estimated with interaction terms of the female dummy with tenure, squared tenure, and cubed tenure (see Table 13). For an easier interpretation of the results, Figure 13 plots the predicted GWG in dependence of tenure. The wage gap of female blue-collar workers is about minus 7 percent in their early career and decreasing with tenure. After 10 years the GWG is about minus 12 percent and the GWG reaches its minimum with nearly minus 18 percent after 26 years of tenure. Among white-collar workers the GWG in the early career is about minus 15 percent and, thus, twice as large as for blue-collar workers. The evolution of the GWG is also quite different, because it continuously increases to approximately minus 9 percent after 30 years of tenure. We estimated additional specifications which control for hierarchical levels (see Table 13). Although the GWGs are smaller (see Section 4.1), the evolution of the GWGs is quite similar (see Figure 14).[6]

- ***Insert Table 13 about here***

[6] Since the general GWG has been reduced over the last decades and women with more tenure have entered the labour market earlier, cohort effects might play an important role. However, our results proved to be robust after controlling for the year of birth or age instead of entry age. Additionally, we estimated fixed effects models (interaction female and tenure but no female dummy) that produced the same results as the interaction terms in the random effects models.

- ***Insert Figure 13 about here***
- ***Insert Figure 14 about here***

The reduction of the GWG among white-collar workers is in line with predictions from statistical discrimination and employer learning (Altonji and Pierret, 2001). Because the firm has more uncertainty about females' than males' individual productivity, the firm pays lower starting wages to females. If the firm learns that women are not less productive than males, it will adjust wages and, hence, the GWG will be reduced.

The story of statistical discrimination does not seem to make sense for blue-collar workers, because the GWG gets larger with tenure. An explanation for this finding in a world of statistical discrimination could be the 'self-fulfilling prophecy' concept (Farmer and Terrell, 1996). If a firm expects women to be less productive (or more absent) than men, it will spent less resources to train female workers and, consequently, women will be less productive than men in the long-run.[7] The early career disadvantage should be persistent and lead to an increasing GWG, because the career opportunities for women are worse than for their male colleagues. This might also explain why female blue-collar workers are stuck to lower hierarchical levels as shown in the descriptive analysis in Section 3.2 (for a discussion of lower promotion probabilities of females see Pfeifer, 2007).

The different developments of the GWGs among blue-collar and white-collar workers fit both into theories of statistical discrimination. As the explanations are somewhat different – employer learning vs. investments in human capital – and we analyse the workforce of a single company, it needs to be discussed why the firm treats blue-collar and white-collar workers differently. Considering employer learning, the firm might even find it easier to observe the productivity in production work (blue-collar) than in administration work (white-collar). Thus, we think it is more likely that investments in human capital in the 'self-fulfilling prophecy' concept can explain the differences. Administration work typically requires more general than firm-specific qualifications,

[7] Consistent with this, Fahr and Sunde (2008) report evidence for Germany that women are less likely than men to participate in formal training after finishing apprenticeship.

whereas it is the opposite way around in production work.[8] The type of skills has implications for cost coverage of training (Becker, 1975) and, consequently, on training decisions. The firm is more likely to invest in specific skills of workers, who have a higher expected productivity, who are likely to provide more working hours, and who are likely to have higher employment stability. Therefore, the firm will rather invest in male than female blue-collar workers. Female white-collar workers, however, can largely decide on their own if they invest in human capital because skills are mostly general and, consequently, the firm will not cover the training costs anyway.

Another explanation might be related to differences in absenteeism, because the gender gap in absenteeism follows the same trend as the evolution of the GWG. As we will show in Section 4.5, the evolution of the gender gap in absent working hours differs between blue-collar and white-collar workers. Whereas the gender gap in absenteeism increases with age for blue-collar workers, the gender gap decreases with age for white-collar workers. A reason might be that physical fitness is more important in production than in administration work and female physical fitness is more sensitive to ageing than male physical fitness.

4.4 Wage Effect of Individual Absenteeism

Individual absenteeism should affect negatively individual wages for several reasons. Not only quasi-fixed labour costs like adjustment costs and investments in human capital are explanations for a wage penalty for workers who are more absent, but also wage costs. If firms have to cover the costs of wage replacement for absent (sick) workers as in Germany, firms will pay workers, who are more likely to be absent, lower average wages because they supply less effective working hours. In addition to a negative wage effect of absenteeism, the inclusion of control variables for individual absenteeism might reduce the GWG because absenteeism might have been an omitted variable in the previous estimates.

[8] Note that the analyzed firm is very specialized and has also a unique production technology.

Several specifications of the earnings functions are estimated for blue-collar workers (see Table 14) and white-collar workers (see Table 15).[9] In the upper part (see panel A) of Tables 14 and 15, the estimates for the total sample are repeated and an additional explanatory variable measuring the individual number of absent working hours in a given month is included. Contradicting our hypothesis of a negative impact, absenteeism is positively correlated with the hourly wage and has a stronger impact for white-collar than for blue-collar workers. However, the coefficients are small and not significant for blue-collar workers. Overall, the inclusion of absenteeism does not change the GWG. To check this result, several robustness checks are performed.

- ***Insert Table 14 about here***
- ***Insert Table 15 about here***

An endogeneity problem might exist, because workers might not earn less or more because they are more absent, but workers react with more absenteeism to lower or higher wages (adjustment to the equilibrium in the leisure-consumption framework of labour supply, income and substitution effects). Therefore, our first robustness check is an IV (instrumental variable) estimator (see last estimate in panel A). For this purpose, an instrument was constructed which is highly correlated with the individual number of absent working hours in a given month but not with the individual wage. The mean number of absent working hours in a given month fulfils this condition, since it measures an exogenous shock to absenteeism (e.g., cold, flu) but not to wages. The wage effect of absenteeism in the IV estimates is positive and weakly significant for blue-collar workers, whereas it is negative and highly significant for white-collar workers.

The next robustness check is concerned with the measurement of absenteeism. As it is plausible that wages are not adjusted every month for actual absenteeism, the firm might use longer measurement (learning) periods. Thus, we use a sample of workers with at least 24 months observed employment history and different measurement spells (sum in one month, one year, and two years) for absenteeism as explanatory variables

[9] Tables 14 and 15 report only coefficients for female dummies and absenteeism measures. Complete estimation results can be requested from the corresponding author.

(see panel B). None of these absenteeism variables is significant and no impact on the GWG can be detected. In the next step, we use only observations from December 2005 (see panel C) and as an additional measurement of absenteeism the mean monthly number of absent working hours in the total observed employment history. This long term measure, which uses all available information on absenteeism, finally shows the expected impact. Workers with on average more absent working hours earn significantly lower wages and the GWG is reduced by 0.33 percentage points for blue-collar workers and 0.67 percentage points for white-collar workers.

Overall, we find no conclusive evidence that individual absenteeism has a strong negative wage effect and that its inclusion in the earnings functions explains a larger part of the GWG. Corcoran and Duncan (1979) and Duncan and Corcoran (1984) also find that individual absenteeism explains almost none of the GWG although differences in mean absent working hours between men and women are large. Osterman (1979) reports evidence that average absenteeism in previous years has a negative effect on wages of females but that the impact on the GWG is limited. Malkiel and Malkiel (1979) also find a negative correlation between average absenteeism and annual income, which is only significant for females.

4.5 Gender Differences in Absenteeism

In this section, we use random effects GLS estimates for the number of absent working hours per month to determine the gender gap in absenteeism. Table 16 informs about the results. The first estimates include only the female dummy and month dummies. Female blue-collar (white-collar) workers are on average 2.1 (0.8) hours more absent than male blue-collar (white-collar) workers, which concludes in a gender gap in absenteeism of 31 (26) percent. After including control variables for tenure, entry age, schooling, and working hours, the gender gap in absenteeism increases to 2.2 hours (32 percent) for blue-collar workers and 1.5 hours (46 percent) for white-collar workers.[10] The estimated gender gap in absenteeism in our personnel data falls into the range of

[10] Ichino and Moretti (2006) also find for four different data sets that the gender gap in absenteeism increases after control variables are included in the estimates.

estimates in previous studies (e.g., Ichino and Moretti, 2006). As Ichino and Moretti (2006) point out, females are exposed to an additional health shock in absenteeism, namely the menstrual cycle. Therefore, independent of differences in worker characteristics, a sizable gender gap in absenteeism remains which is partly determined by biological gender differences. Other sources of higher female absenteeism might stem from family responsibilities (e.g., care for sick children).

- ***Insert Table 16 about here***

Ichino and Moretti (2006) point out that the menstrual cycle effect should disappear after the age of 45 years and report convincing evidence in support of this hypothesis. Thus, we also look at the evolution of the gender gap in absent working hours. As we are now interested in the age effect, we replace the tenure and entry age variables with age variables in the following estimates. Moreover, interaction terms between the female dummy and age, squared age, and cubed age are incorporated. Table 17 contains the results. To ease interpretation, the evolution of the gender gap in absent working hours for blue-collar and white-collar workers is depicted in Figure 15. Female blue-collar (white-collar) workers are about 1.4 (1.7) days more absent than male blue-collar (white-collar) workers at the age of 25. The gender gap increases with age for blue-collar workers to more than three absent working hours after the age of 50. For white-collar workers the gender gap is relatively constant until the age of 45, but decreasing afterwards. The latter finding is line with Ichino and Moretti (2006) who analysed white-collar workers and the menstrual cycle effect. An explanation for the steadily increasing gender gap for blue-collar workers might be that physical fitness is more important in production than in administration work and female physical fitness is more sensitive to ageing than male physical fitness. The differences in the gender gap in absenteeism between blue-collar and white-collar workers might also partly explain the different evolutions of the GWGs (see Section 4.3).

- ***Insert Table 17 about here***
- ***Insert Figure 15 about here***

As shown in the previous section, individual absenteeism has only a small impact on the GWG. Therefore, it seems not to be straightforward to interpret the costs associated with individual absenteeism as the major determinant of the observed GWG. However, combining the cost approach with statistical discrimination might explain the persistence of the GWG. Recalling the findings, the gender gap in absenteeism is between 26 and 46 percent and larger than the GWGs of about 13 percent estimated in the previous sections. These findings are the basis for the subsequent theoretical model and discussion.

5. A Simple Model and Discussion

5.1 The Basic Model with Absenteeism and Homogenous Workers

The subsequent model combines wage costs due to absenteeism and statistical discrimination when determining female and male wages to explain the GWG.[11] We use three scenarios in which a firm maximizes profits under perfect competition. The first scenario (A) comprises no wage replacement for absent workers, i.e., the firm pays the market wage only for effective working hours (contractual working hours minus absent working hours). In the second scenario (B), absent workers receive 100 percent wage replacement and the firm cannot adjust the market wage, i.e., the firm pays the market wage for contractual working hours regardless of absent working hours. The third scenario (C) also comprises 100 percent wage replacement. In this scenario, however, the firm can adjust the market wage according to absent working hours, i.e., the firm pays an adjusted wage for contractual working hours.

At first, we analyse a simple setting of homogeneous workers. Scenario A (no wage replacement) can be described as follows. The competitive firm maximizes its profits ($\prod$) in (2) if the difference between the value of total output (pQ) and costs is maximised. We assume market output prices (p), a fixed production technology (Q), constant capital (K), market capital prices (r), and market wages (w). Moreover, total labour consists of the number of workers (N) times the effective working hours, which

[11] The model developed in Ichino and Moretti (2006) is quite different from ours, because their "key insight is that absenteeism is a noisier signal of shirking attitudes for females than males" (Ichino and Moretti, 2006, p. 17).

are contractual working hours (H_C) minus absent working hours ($H_A \leq H_C$). As all workers are homogenous, all workers provide the same number of working hours. Furthermore, absent working hours are exogenously chosen by workers. Thus, the firm's only choice variable is the number of hired workers, which is similar to the standard textbook model of the profit-maximizing firm in the short run. The first order condition (FOC) in (3) yields the standard result that a firm hires workers up to the point in which wage costs per worker equals the value of marginal product. If workers are more absent, the firm simply hires more workers.

$$(2) \qquad \max_{N} \; \Pi = pQ\left[\left(H_C - H_A\right)N, K\right] - w\left(H_C - H_A\right)N - rK$$

$$(3) \qquad w\left(H_C - H_A\right) = p\frac{\partial Q}{\partial N}$$

In scenario B (100 percent wage replacement without wage adjustment), the profit function and FOC slightly change. As the firm has to pay the market wage regardless of absent working hours, absent working hours are dropped from labour costs in the profit function in (4) and are not part of the LHS of the FOC in (5). Compared with scenario A, wage replacement in scenario B implies less employment because the firm has not to pay for effective but for contractual working hours.

$$(4) \qquad \max_{N} \; \Pi = pQ\left[\left(H_C - H_A\right)N, K\right] - wH_C N - rK$$

$$(5) \qquad wH_C = p\frac{\partial Q}{\partial N}$$

Scenario C comprises 100 percent wage replacement and wage adjustment. The firm adjusts market wages according to absenteeism so that total labour costs equal total labour costs as if there is no wage replacement (scenario A). The effective wage for contractual working hours is denoted with w_E and calculated as follows in (6) and (7).

$$(6) \qquad w_E H_C = w\left(H_C - H_A\right)$$

$$(7) \qquad w_E = \frac{\left(H_C - H_A\right)}{H_C}w = \left(1 - \frac{H_A}{H_C}\right)w$$

The new profit function and FOC in (8) and (9) show that the firm hires the same number of workers in scenario C as in scenario A, if wages can be adjusted according to absenteeism – even in case of a 100 percent wage replacement.

(8) $$\max_{N} \ \Pi = pQ\left[\left(H_C - H_A\right)N, K\right] - w_E H_C N - rK$$

(9) $$w_E H_C = p\frac{\partial Q}{\partial N}$$

5.2 Introducing Heterogeneous Groups, Statistical Discrimination, and Employer Learning

The model with homogenous workers is very basic and intuitive. Let us now assume that workers are heterogeneous with respect to absenteeism, but have the same number of contractual working hours and are equally productive. Moreover, we assume that the firm has no information about future individual absenteeism. However, the firm has information on average absent working hours among two groups – females (F) and males (M). These group averages are used to form expectations about individual absenteeism of workers (statistical discrimination). Expected absenteeism of a newly hired female is $E\left[H_{AF}\right]$ and expected absenteeism of a newly hired male is $E\left[H_{AM}\right]$. From the previous empirical analysis, we know (like the firm) that females are on average more hours absent than males ($E\left[H_{AF}\right] > E\left[H_{AM}\right]$).

In scenario A (no wage replacement), such statistical discrimination has no impact on wages at first. As the firm has only to pay for effective working hours, the firm is indifferent between hiring females and males. However, the firm would have to hire more females. In case of quasi-fixed employment costs, this leads to the conclusion that the firm prefers to hire males. Consequently, females might reduce their reservation wages in the long run and a GWG might arise. Scenario B (100 percent wage replacement without wage adjustment) implies also no impact on wages at first, because the firm prefers to hire males with lower expected absenteeism and associated wage costs than females.[12] Like in scenario A, females might reduce their reservation wages and a GWG might arise.

The most interesting scenario is scenario C with 100 percent wage replacement and the possibility to adjust wages according to expected absenteeism. In such a setting, the firm is also willing to hire females. However, the firm only hires females with higher expected absenteeism than males if female wages are adjusted downwards. Consequently, a GWG arises. Recall the definition of the effective wage after wage adjustment in (10). We can also define the effective wage for females and males, from which the female-male wage ratio can be calculated in (11). To simplify notification and to allow for differences in working hours, we introduce the expected absenteeism rate which is $E\left[A\right]=\frac{E\left[H_A\right]}{H_C}$.

$$(10) \qquad w_E=\left(1-\frac{E\left[H_A\right]}{H_C}\right)w=\left(1-E\left[A\right]\right)w$$

$$(11) \qquad \frac{w_{EF}}{w_{EM}}=\frac{\left(1-\frac{E\left[H_{AF}\right]}{H_C}\right)w}{\left(1-\frac{E\left[H_{AM}\right]}{H_C}\right)w}=\frac{\left(1-E\left[A_F\right]\right)}{\left(1-E\left[A_M\right]\right)}$$

The first derivate of the female-male wage ratio with respect to the expected female absenteeism rate in (12) is negative, i.e., if, ceteris paribus, female absenteeism increases, female wages and, consequently, the female-male wage ratio decreases. It can be seen that the female-male wage ratio reacts more strongly to a change in female absenteeism if male absenteeism is high. The rationale is that the wage cost argument gains in importance for higher absenteeism rates, for which expected male absenteeism is the reference point.

[12] Renes and Ridder (1995) and Sattinger (1998) show analogously that profit-maximizing firms set higher hiring standards and interview fewer workers from a group which has a higher average quit rate, i.e., firms are more likely to hire men.

(12) $$\frac{\partial\left(\frac{w_{EF}}{w_{EM}}\right)}{\partial E\left[A_F\right]} = -\frac{1}{\left(1-E\left[A_M\right]\right)} < 0$$

Figure 16 illustrates this finding graphically for two cases. In the first case, men are not absent at all ($E\left[A_M\right]=0$) and, therefore, the slope is minus one. The second case depicts the general case in which expected male absenteeism rates are larger zero ($E\left[A_M\right]>0$). The axis intercept is larger one and the slope is steeper than in the first case. If $E\left[A_F\right]>E\left[A_M\right]$, female wages are lower than male wages.

- ***Insert Figure 16 about here***

We use a simple example from our descriptive statistics (see Section 3) to demonstrate the quantitative dimension of absenteeism on the female-male wage ratio in our analysed company. Table 18 displays the means of absenteeism and contractual working hours for females and males, from which the female-male wage ratio is calculated according to (11). In a world without any other forms of discrimination and under the assumption of equal productivity of the genders, the female-male wage ratios are 98.3 percent for blue-collar workers and 99.17 percent for white-collar workers. Thus, the GWG should be relatively small between women and men based on differences in absenteeism rates.[13] This is because the overall absenteeism rates are small, though differences between men and women are large. Nevertheless, statistical discrimination seems to have a not negligible impact on the GWG because we only account for statistical discrimination on one factor, namely wage costs due to absenteeism. There are, however, other factors (e.g., quasi-fixed employment costs, quit rates) which are correlated with employment costs and from which statistical discrimination is likely to occur (Renes and Ridder, 1995; Sattinger, 1998). All these factors might add up to a sizeable GWG explained by statistical discrimination.

- ***Insert Table 18 about here***

[13] Allen (1983, pp. 391-392) also reports empirical evidence that the differences in average absence rates between men and women should account only for a small fraction of the GWG.

If the firm learns about individual absenteeism, statistical discrimination based on average absenteeism in the gender groups might get less important when setting wages. In scenario A, however, employer learning has no impact on the GWG because absent working hours are not paid. In scenario B (100 percent wage replacement without wage adjustment), employer learning has also no impact on wages because the firm hires no females. Even if females reduce their reservation wages to get hired, learning has no impact on the GWG because the GWG is induced by differences in reservation wages.

In scenario C (100 percent wage replacement with wage adjustment), the firm statistical discriminates females when they enter the firm so that they receive lower wages than males, whose absenteeism rates are on average lower. After learning about individual absenteeism, the firm could adjust wages according to expected future individual absenteeism. However, females with more absenteeism than males still receive lower wages. Since wage setting is a complex and sensitive process, the firm might prefer to stay with the initial wage differential so that the impact of individual absenteeism on wages is limited.

5.3 Implications

The results of the theoretical model, which is based on our empirical findings, have implications for the wage structure. If the GWG is partly determined by the cost argument of higher absenteeism rates of females, this part of the wage differential between men and women will be persistent over time. Moreover, statistical discrimination by the firm leads to redistribution of wages. Women with lower than average female absenteeism are underpaid relatively to their individual effective working hours, whereas women with higher than average female absenteeism are overpaid. Thus, females with low absenteeism subsidise females with high absenteeism. If the gender gap in wages does not completely account for the gender gap in absenteeism, a second redistribution from males to females exists. Winner of these redistribution processes are especially women with high absenteeism. Statistical discrimination, however, makes also men with higher than average male absenteeism to winners because they are paid higher wages than in a world in which every worker is paid according to his effective number of working hours.

In the setting of our model, policy interventions to reduce the GWG are market interventions. Note that the wage differentials between men and women are no market failure and that statistical discrimination might be efficient (Schwab, 1986; Norman, 2003). Nevertheless, policy interventions can be justified on equity instead of efficiency grounds (Cain, 1985). First, policy can implement equal pay laws, which in the framework of our model are not the same as anti discrimination laws. Such an equal pay intervention would advise the firm to pay male and female workers the same wage for a given number of contractual working hours regardless of absenteeism, i.e., the firm is not allowed to statistical discriminate females in making wage adjustments. Because this results into higher wage costs for effective working hours of women, the firm will hire no or fewer women and the hired female workers might be a positive selection (e.g., women need a higher productivity than men to equalize the higher absent working hours). Second, policy can subsidize female wages to reduce the GWG in paid hourly wages (Cain, 1985). This need not to be done explicitly but can also be done implicitly by different taxation of men and women. A third policy could directly aim at absenteeism of females. For example, special health care policy for women might reduce the gender gap in absenteeism. Fourth, nursery schools in Germany often suffer from inflexible opening hours and long closing periods during school holidays. More financial and personnel support for childcare facilities could improve the situation for mothers. Fifth, working time flexibility as well as possibilities to work from home (e.g., telework) could help to combine market work with family responsibilities.

Even though we discussed our results in the light of statistical discrimination and find them largely consistent with it, we cannot rule out other explanations of the GWG like taste or societal discrimination (e.g., education, family responsibilities). Therefore, the findings should not be interpreted as a ‘magic bullet’ which explains – or even excuses – all disadvantages of women in the labour market. Furthermore, it should be kept in mind that our results are not representative because we analysed only data of one single company.

References

Achatz, J., Gartner, H., Glück, T. (2005). Bonus oder Bias? Mechanismen geschlechtsspezifischer Entlohnung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 57, 466-493.

Aigner, G.J., Cain, G.G. (1977). Statistical theories of discrimination in labor markets. Industrial and Labor Relations Review 30, 175-187.

Allen, S.G. (1981a). Compensation, safety, and absenteeism: evidence from the paper industry. Industrial and Labor Relations Review 34, 207-218.

Allen, S.G. (1981b). An empirical model of work attendance. Review of Economics and Statistics 63, 77-87.

Allen, S.G. (1983). How much does absenteeism cost? Journal of Human Resources 18, 379-393.

Altonji, J.G., Blank, R.M. (1999). Race and gender in the labor market. Handbook of Labor Economics 3, 3143-3259.

Altonji, J.G., Pierret, C.R. (2001). Employer learning and statistical discrimination. Quarterly Journal of Economics 116, 313-350.

Barnet-Verzat, C., Wolff, F.-C. (2008). Gender wage gap and the glass ceiling effect: a firm-level investigation. International Journal of Manpower (forthcoming).

Becker, G.S. (1971). The economics of discrimination. 2nd Edition, Chicago: University of Chicago Press.

Becker, G.S. (1975). Human capital: a theoretical and empirical analysis, with special reference to education. 2nd edition, New York et al.: Columbia University Press.

Blau, F.D., Ferber, M.A., Winkler, A.E. (2006). The economics of women, men, and work. 5th edition, Upper Saddle River, NJ: Prentice Hall.

Blau, F.D., Kahn, L.M. (2000). Gender differences in pay. Journal of Economic Perspectives 14, 75-99.

Blau, F.D., Kahn, L.M. (2003). Understanding international differences in the gender pay gap. Journal of Labor Economics 21, 106-144.

Cain, G.G. (1985). Welfare economics of policies toward women. Journal of Labor Economics 3: S375-S396.

Cain, G.G. (1986). The economic analysis of labor market discrimination: a survey. Handbook of Labor Economics 1, 693-785.

Corcoran, M., Duncan, G.J. (1979). Work history, labor force attachment, and earnings differences between races and sexes. Journal of Human Resources 14, 3-20.

Datta Gupta, N., Rothstein, D.S. (2005). The impact of worker and establishment characteristics on male-female wage differentials: evidence from Danish matched employee-employer data. Labour 19, 1-34.

Dohmen, T., Lehmann, H., Zaiceva, A. (2008). The gender wage gap inside a Russian firm: first evidence from personnel data – 1997 to 2002. IZA Discussion Paper No. 3428.

Duncan, G.J., Corcoran, M. (1984). Do women 'deserve' to earn less than men? In: Duncan, G.J. (ed.): Year of poverty, years of plenty. Ann Arbor, Mich.: The University of Michigan.

Fahr, R., Sunde, U. (2008). Gender differentials in skill use and skill formation in the aftermath of vocational training. Applied Economics Letters (forthcoming).

Farmer, A., Terrell, D. (1996). Discrimination, Bayesian updating of employer beliefs, and human capital accumulation. Economic Inquiry 34, 204-219.

Fitzenberger, B., Wunderlich, G. (2002). Gender wage differences in West Germany: a cohort analysis. German Economic Review 3, 379-414.

Frick, B., Malo, M.Á. (2005). Labour market institutions and individual absenteeism in the European Union: the relative importance of sickness benefit systems and employment protection legislation. Faculty of Management and Economics, Witten/Herdecke University, Mimeo.

Gartner, H., Rässler, S. (2005). Analyzing the changing gender wage gap based on multiply imputed right censored wages. IAB Discussion Paper No. 5/2005.

Gunderson, M. (1989). Male-female wage differentials and policy responses. Journal of Economic Literature 27, 46-72.

Ichino, A., Moretti, E. (2006). Biological gender differences, absenteeism and the earning gap. IZA Discussion Paper No. 2207.

Jirjahn, U., Stephan, G. (2004). Gender, piece rates and wages: evidence from matched employer-employee data. Cambridge Journal of Economics 28, 683-704.

Kahn, L.M. (1981). Sex discrimination in professional employment: a case study – comment. Industrial and Labor Relations Review 34, 273-275.

Kidd, M.P., Shannon, M. (1996). Does the level of occupational aggregation affect estimates of the gender wage gap? Industrial and Labor Relations Review 49, 317-329.

Lazear, E.P. (1979). Why is there mandatory retirement? Journal of Political Economy 87, 1261-1284.

Leigh, J.P. (1983). Sex differences in absenteeism. Industrial Relations 22, 349-361.

Malkiel, B.G., Malkiel, J.A. (1973). Male-female pay differentials in professional employment. American Economic Review 63, 693-705.

Mastekaasa, A., Modesta Olsen, K. (1998). Gender, absenteeism, and job characteristics – a fixed effects approach. Work and Occupations 25, 195-228.

Meyersson Milgrom, E.M., Petersen, T., Snartland, V. (2001). Equal pay for equal work? Evidence from Sweden and a comparison with Norway and the U.S. Scandinavian Journal of Economics 103, 559-583.

Neumark, D. (1999). Wage differentials by race and sex: the roles of taste discrimination and labor market information. Industrial Relations 38, 414-445.

Norman, P. (2003). Statistical discrimination and efficiency. Review of Economic Studies 70, 615-627.

Oaxaca, R.M, Ransom, M.R. (1999). Identification in detailed wage decompositions. Review of Economics and Statistics 81, 154-157.

Oaxaca, R.M. (1973). Male-female wage differentials in urban labor markets. International Economic Review 9, 693-709.

Osterkamp, R., Röhn, O. (2007). Being on sick leave: possible explanations for differences of sick-leave days across countries. CESifo Economic Studies 53, 97-114.

Osterman, P. (1979). Sex discrimination in professional employment: a case study. Industrial and Labor Relations Review 32, 451-464.

Osterman, P. (1981). Sex discrimination in professional employment: a case study – reply. Industrial and Labor Relations Review 34, 275-276.

Paringer, L. (1983). Women and absenteeism: health or economics? American Economic Review (Papers and Proceedings) 73, 123-127.

Petersen, T., Snartland, V., Meyersson Milgrom, E.M. (2005). Are female workers less productive than male workers? Productivity and the gender wage gap. Mimeo.

Pfeifer, C. (2007). Determinants of promotions in an internal labour market: testing implications from tournament theory and efficient allocation. Paper presented at the 22nd Congress of the European Economic Association.

Pfeifer, C. (2008). An empirical note on wages in an internal labour market. Economics Letters 99, 570-573..

Prey, H. (1999). Die Entwicklung der geschlechtsspezifischen Lohndifferenz in Deutschland 1984 –1996. Diskussionspapiere des Forschungsinstituts für Arbeit und Arbeitsrecht an der Universität St. Gallen Nr. 57.

Renes, G., Ridder, G. (1995). Are women overqualified? Labour Economics 2, 3-18.

Sattinger, M. (1998). Statistical discrimination with employment criteria. International Economic Review 39, 205-237.

Schwab, S. (1986). Is statistical discrimination efficient? American Economic Review 76, 228-234.

Statistisches Bundesamt (2006). Datenreport 2006. Bundeszentrale für politische Bildung.

Stephan, G. (1991). Fehlzeiten: Eine theoretische und empirische Untersuchung mit Individualdaten. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 24, 583-594.

Thalmaier, A. (2002). Eine ökonomische Analyse von Fehlzeiten. Frankfurt et al.: Lang.

Vistnes, J.P. (1997). Gender differences in days lost from work due to illness. Industrial and Labor Relations Review 50, 304-323.

Winter-Ebmer, R., Zweimüller, J. (1997). Unequal assignment and unequal promotion in job ladders. Journal of Labour Economics 15, 43-71.

Appendix

Table A.1: Descriptive statistics blue-collar workers

	obs	mean	std. dev.	min	max
hourly gross wage (Euros)	50722	14.9392	1.9362	8.8765	33.5824
log wage	50722	2.6957	0.1286	2.1834	3.5140
absent working hours	50722	7.1422	19.2652	0.0000	152.0000
female (dummy)	50722	0.1898	0.3922	0.0000	1.0000
tenure (years)	50722	15.0591	9.1817	0.0658	49.1096
tenure squared / 100	50722	3.1108	3.2535	0.0000	24.1175
tenure cubed / 1000	50722	7.5076	11.5804	0.0000	118.4401
entry age (years)	50722	27.1737	7.4866	13.7041	55.5315
entry age squared / 100	50722	7.9446	4.6050	1.8780	30.8375
age (years)	50722	42.2328	8.6482	19.1973	65.0767
age squared / 100	50722	18.5840	7.3360	3.6853	42.3498
age cubed / 1000	50722	84.7773	49.0392	7.0749	275.5985
apprenticeship degree (dummy)	50722	0.7231	0.4475	0.0000	1.0000
monthly working hours	50722	150.7820	6.7302	43.5000	174.0000
monthly working hours squared / 100	50722	227.8050	17.3817	18.9225	302.7600
level 1 (dummy)	50722	0.0874	0.2824	0.0000	1.0000
level 2 (dummy)	50722	0.0638	0.2445	0.0000	1.0000
level 3 (dummy)	50722	0.0997	0.2996	0.0000	1.0000
level 4 (dummy)	50722	0.4055	0.4910	0.0000	1.0000
level 5 (dummy)	50722	0.1869	0.3899	0.0000	1.0000
level 6 (dummy)	50722	0.1114	0.3147	0.0000	1.0000
level 7 (dummy)	50722	0.0452	0.2078	0.0000	1.0000

Table A.2: Descriptive statistics white-collar workers

	obs	mean	std. dev.	min	max
hourly gross wage (Euros)	73174	23.4121	6.8717	8.7356	83.3333
log wage	73174	3.1153	0.2710	2.1674	4.4228
absent working hours	73174	3.4159	12.9307	0.0000	168.0000
female (dummy)	73174	0.2751	0.4466	0.0000	1.0000
tenure (years)	73174	13.8197	9.5887	0.0055	48.2000
tenure squared / 100	73174	2.8293	3.2501	0.0000	23.2324
tenure cubed / 1000	73174	6.8592	11.2046	0.0000	111.9802
entry age (years)	73174	28.3615	7.3982	13.7041	62.2575
entry age squared / 100	73174	8.5911	4.6121	1.8780	38.7600
age (years)	73174	42.1812	9.4360	18.6603	65.9753
age squared / 100	73174	18.6829	8.0275	3.4821	43.5275
age cubed / 1000	73174	86.3187	53.9495	6.4976	287.1739
low school degree (dummy)	73174	0.4847	0.4998	0.0000	1.0000
high school degree (dummy)	73174	0.1632	0.3696	0.0000	1.0000
university degree (dummy)	73174	0.3521	0.4776	0.0000	1.0000
monthly working hours	73174	151.6139	19.5006	17.3200	174.3000
monthly working hours squared / 100	73174	233.6703	50.5451	2.9998	303.8049
level 1 (dummy)	73174	0.1029	0.3038	0.0000	1.0000
level 2 (dummy)	73174	0.2351	0.4241	0.0000	1.0000
level 3 (dummy)	73174	0.2543	0.4355	0.0000	1.0000
level 4 (dummy)	73174	0.1480	0.3551	0.0000	1.0000
level 5 (dummy)	73174	0.1122	0.3156	0.0000	1.0000
level 6 (dummy)	73174	0.1476	0.3547	0.0000	1.0000

Figures and Tables included in Text

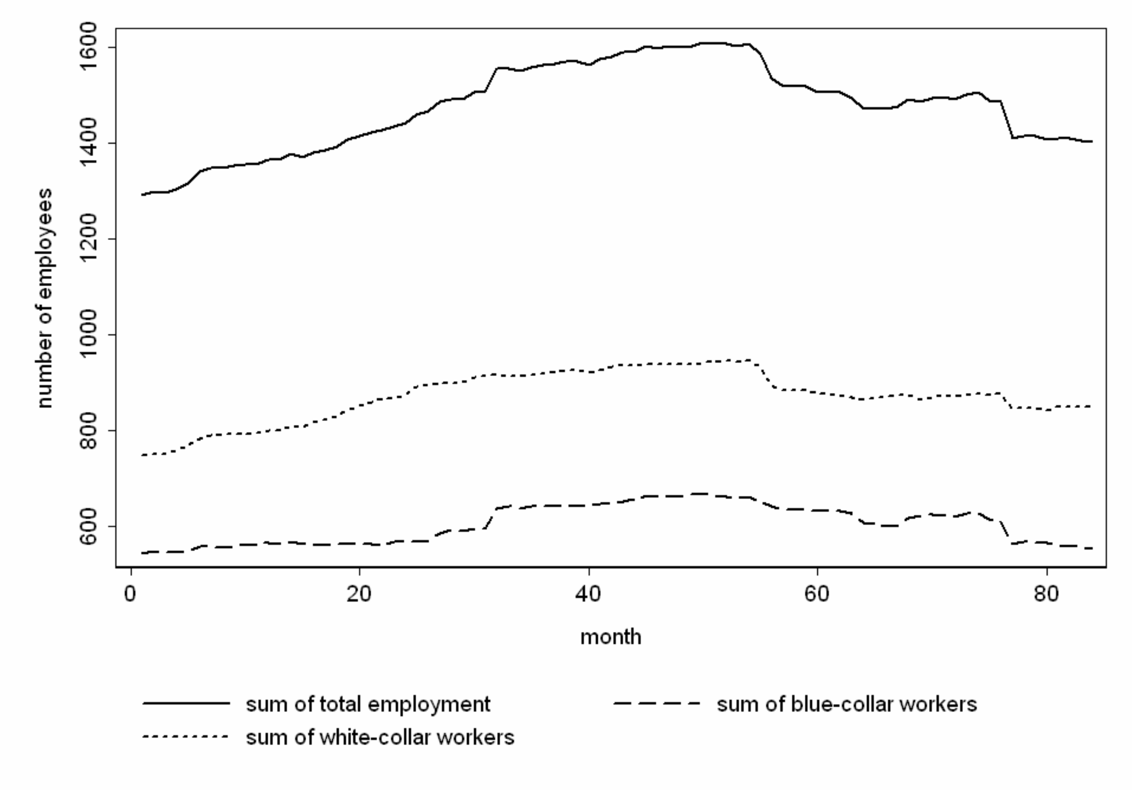

Figure 1: Trends in employment

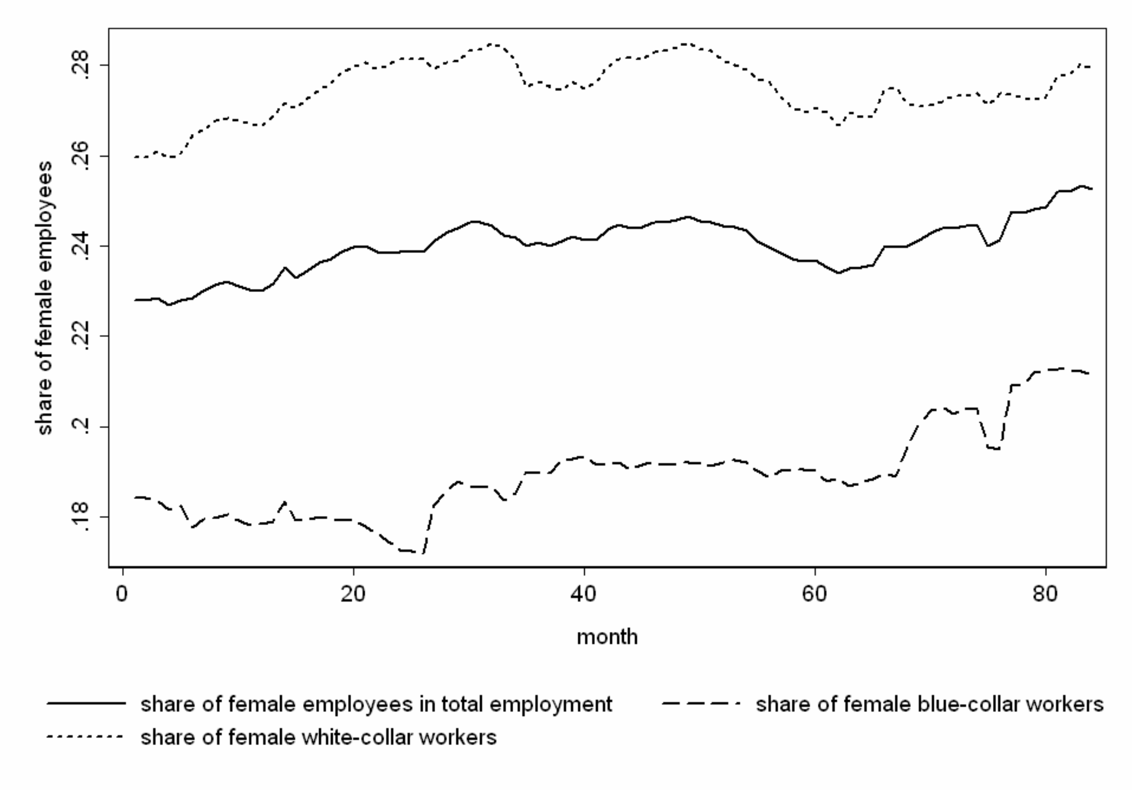

Figure 2: Trends in share of female employment

Table 1: Level descriptions from collective contract

level	blue-collar workers	white-collar workers
1	unskilled work (instruction)	simple tasks (instruction or basic training)
2	semi-skilled work (basic training)	somewhat difficult tasks (three-year apprenticeship)
3	semi-skilled work (two-year apprenticeship)	moderately difficult tasks (university of applied science degree)
4	somewhat difficult skilled work (three-year apprenticeship)	difficult tasks, making decisions of limited scope (university degree)
5	moderately difficult skilled work (three-year apprenticeship)	very difficult tasks, making decisions of broader scope (university degree)
6	difficult skilled work (three-year apprenticeship)	upper management tasks, non-pay-scale (not subject to collective contract)
7	very difficult skilled work (three-year apprenticeship)	

Table 2: Division in levels for blue-collar workers

level	statistics	total	male	female
1	number of obs.	4431	0	4431
	share in row	100	0	100
	share in column	8.74	0	46.02
2	number of obs.	3238	635	2603
	share in row	100	19.61	80.39
	share in column	6.38	1.55	27.04
3	number of obs.	5056	3450	1606
	share in row	100	68.24	31.76
	share in column	9.97	8.4	16.68
4	number of obs.	20569	19749	820
	share in row	100	96.01	3.99
	share in column	40.55	48.06	8.52
5	number of obs.	9482	9314	168
	share in row	100	98.23	1.77
	share in column	18.69	22.67	1.74
6	number of obs.	5652	5652	0
	share in row	100	100	0
	share in column	11.14	13.75	0
7	number of obs.	2294	2294	0
	share in row	100	100	0
	share in column	4.52	5.58	0
total	number of obs.	50722	41094	9628
	share in row	100	81.02	18.98
	share in column	100	100	100

Table 3: Division in levels for white-collar workers

level	statistics	total	male	female
1	number of obs.	7527	2619	4908
	share in row	100	34.79	65.21
	share in column	10.29	4.94	24.38
2	number of obs.	17205	10242	6963
	share in row	100	59.53	40.47
	share in column	23.51	19.31	34.59
3	number of obs.	18606	13641	4965
	share in row	100	73.32	26.68
	share in column	25.43	25.72	24.67
4	number of obs.	10828	9093	1735
	share in row	100	83.98	16.02
	share in column	14.8	17.14	8.62
5	number of obs.	8211	7465	746
	share in row	100	90.91	9.09
	share in column	11.22	14.07	3.71
6	number of obs.	10797	9985	812
	share in row	100	92.48	7.52
	share in column	14.76	18.82	4.03
total	number of obs.	73174	53045	20129
	share in row	100	72.49	27.51
	share in column	100	100	100

Table 4: Composition of the workforce and gender differences

(A) blue-collar workers	total	male	female	M-F
tenure in years	15.0591	15.2288	14.3348	0.8940
entry age in years	27.1737	26.8965	28.3568	-1.4603
apprenticeship dummy	0.7231	0.7777	0.4901	0.2876
contractual working hours	150.7820	151.5719	147.4104	4.1615
number of observations	50722	41094	9628	

(B) white-collar workers	total	male	female	M-F
tenure in years	13.8197	14.3583	12.4003	1.9580
entry age in years	28.3615	29.0581	26.5259	2.5322
low school degree dummy	0.4847	0.4909	0.4682	0.0227
high school degree dummy	0.1632	0.1090	0.3062	-0.1972
university degree dummy	0.3521	0.4001	0.2255	0.1745
contractual working hours	151.6139	156.7344	138.1199	18.6145
number of observations	73174	53045	20129	

Note: All gender differences are significant at the 1%-level in a two sample t-test.

Table 5: Mean wages and gender differences

(A) blue-collar workers

	obs	mean	std. err.	std. dev.
male	41094	15.3500	0.0090	1.8284
female	9628	13.1860	0.0132	1.2993
combined	50722	14.9392	0.0086	1.9362
difference		2.1640	0.0197	

(B) white-collar workers

	obs	mean	std. err.	std. dev.
male	53045	24.8737	0.0305	7.0297
female	20129	19.5605	0.0323	4.5790
combined	73174	23.4121	0.0254	6.8717
difference		5.3133	0.0534	

Note: Mean gross hourly wages in Euros. Gender differences are significant at the 1%-level in a two sample t-test.

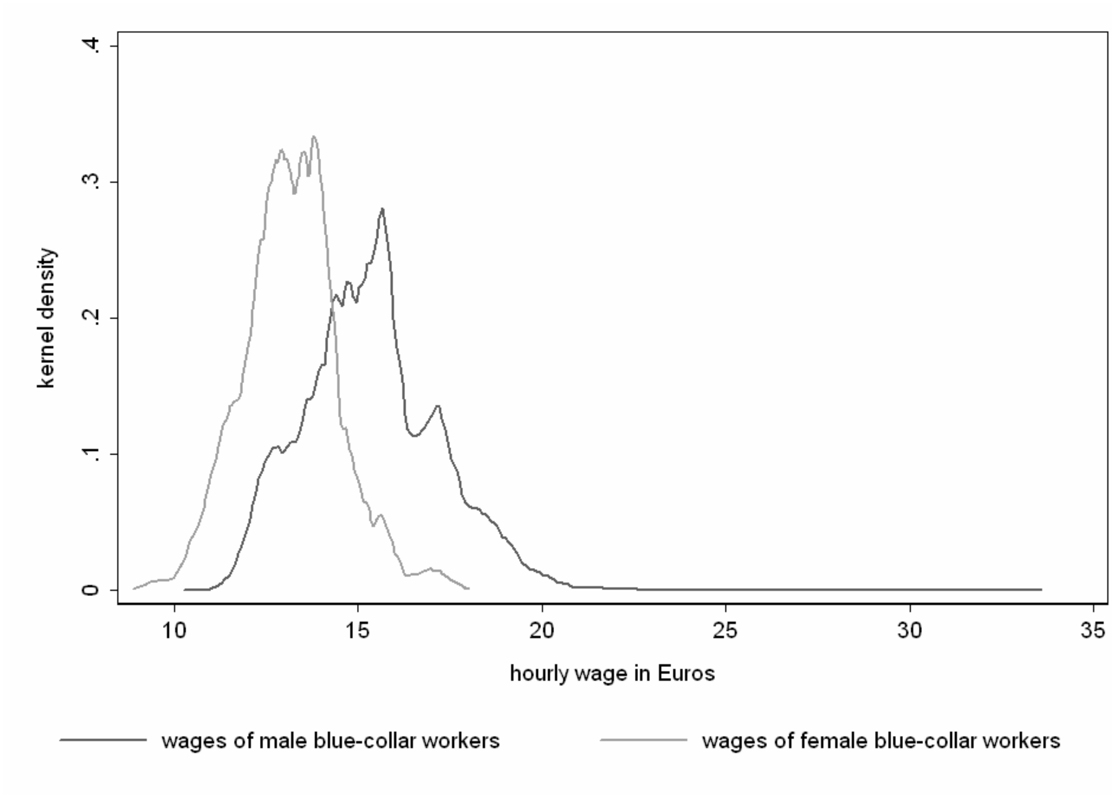

Figure 3: Female and male wage distribution of blue-collar workers

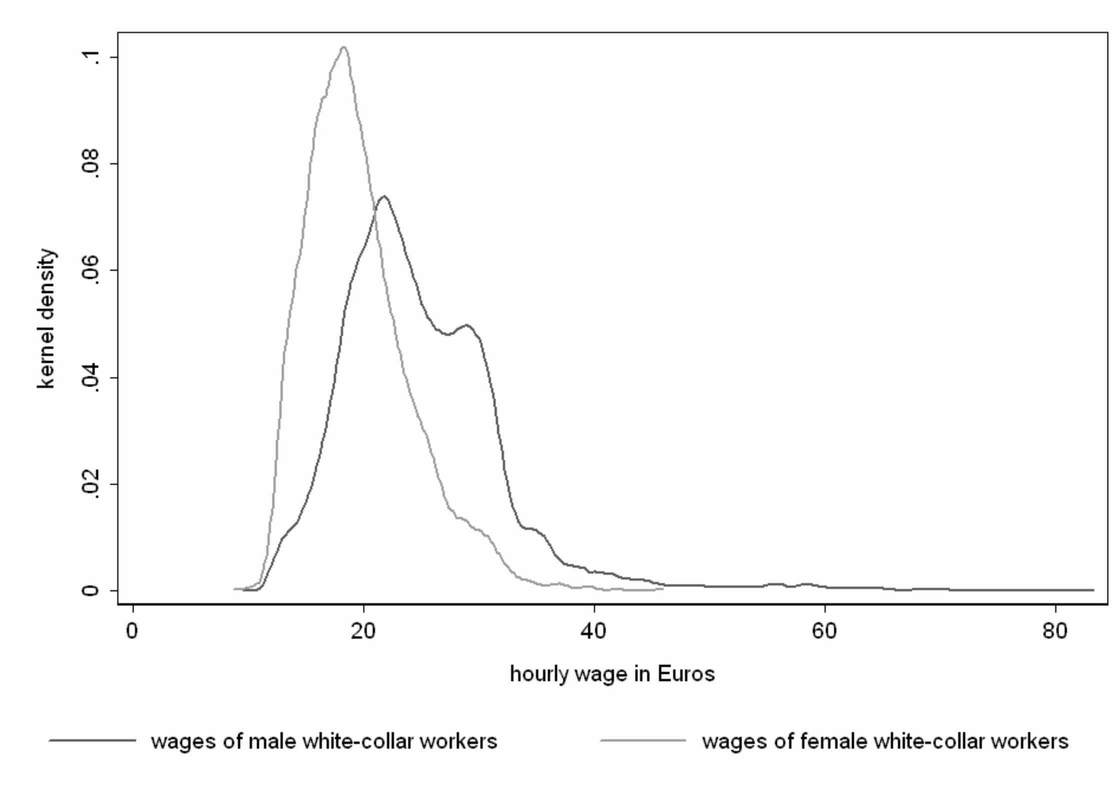

Figure 4: Female and male wage distribution of white-collar workers

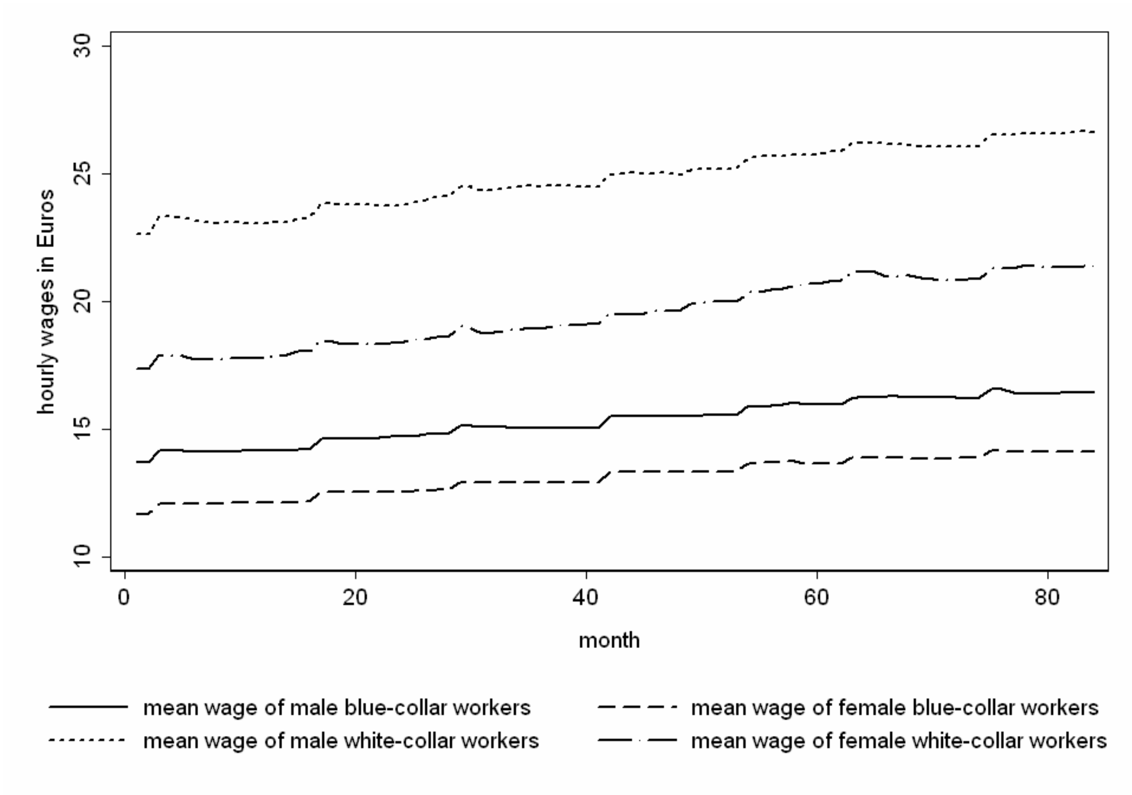

Figure 5: Trends in wages

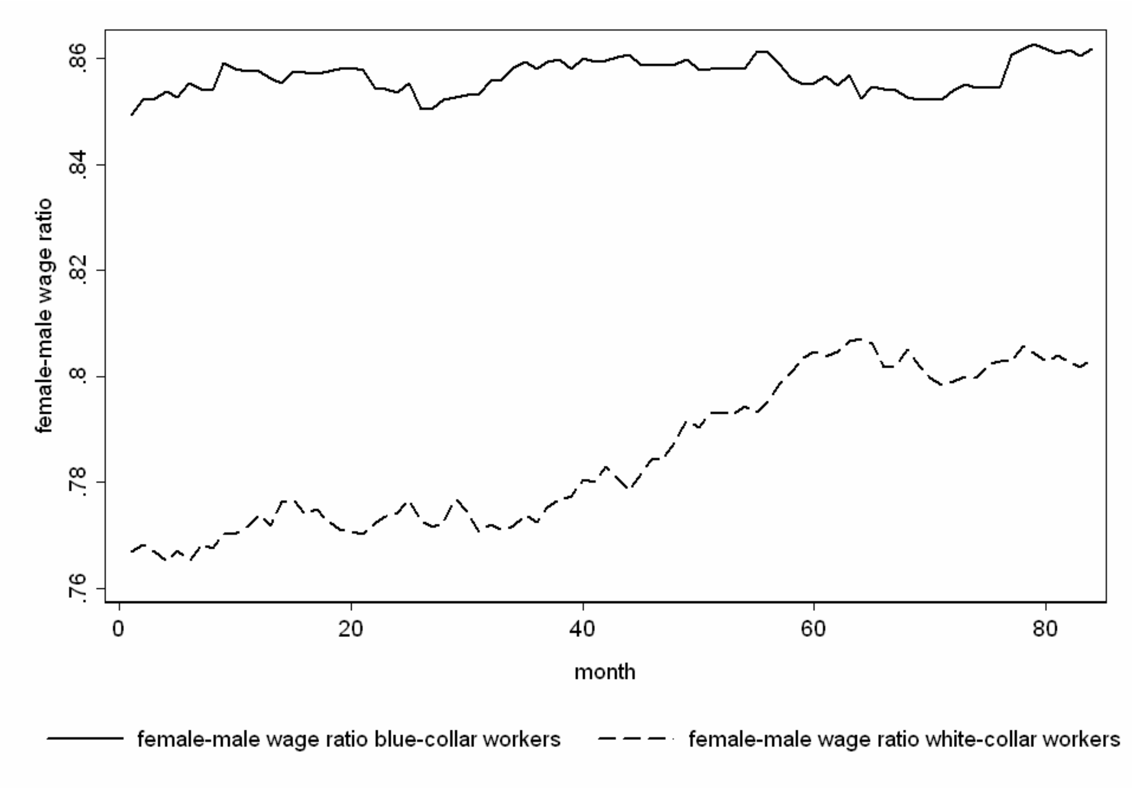

Figure 6: Trends in female-male wage ratio

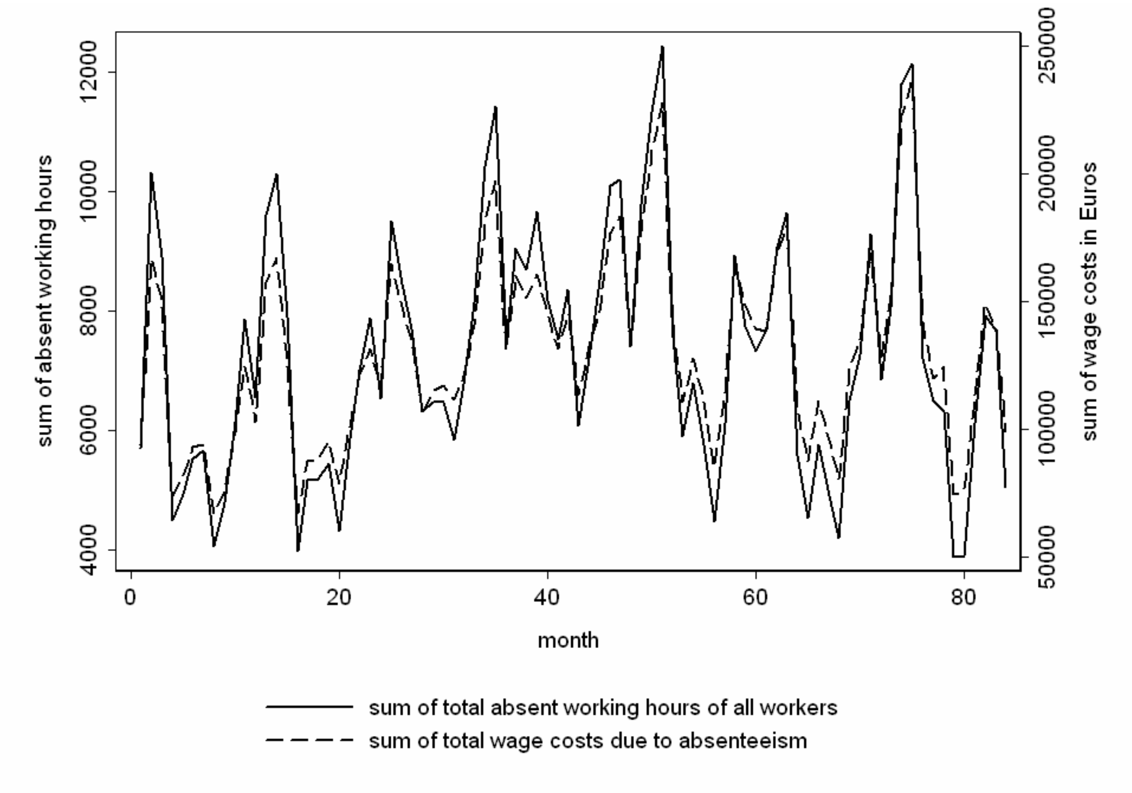

Figure 7: Absent working hours and associated wage costs

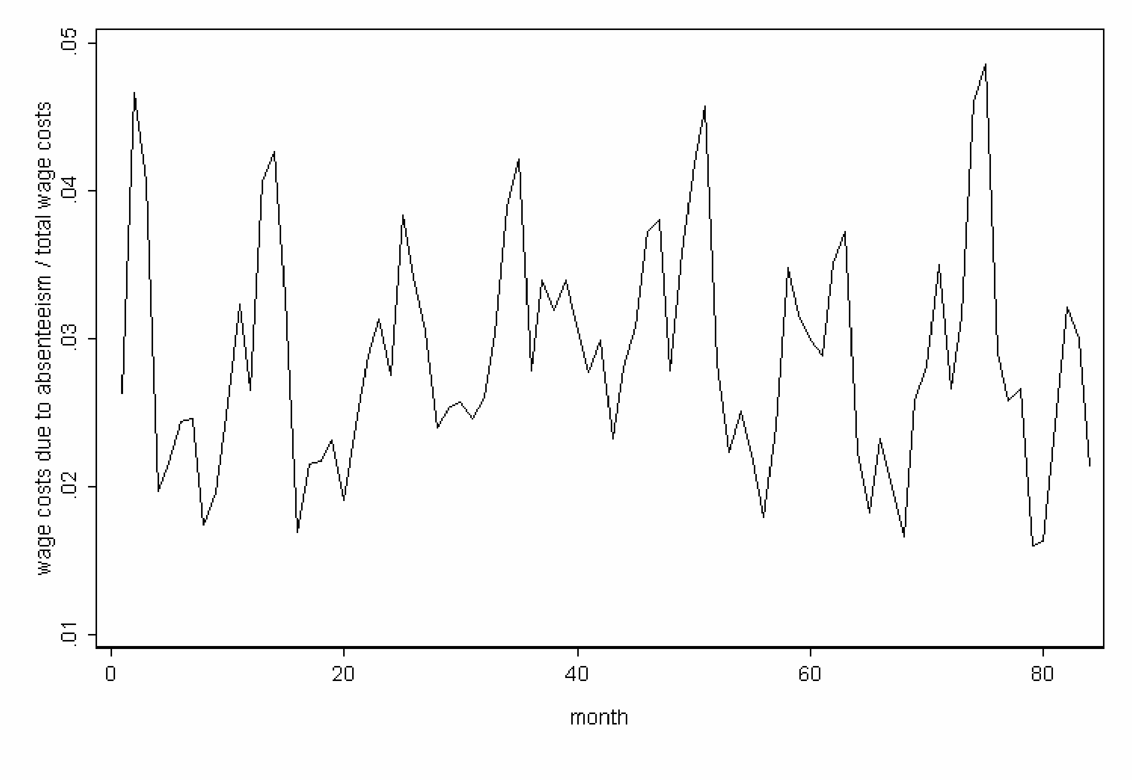

Figure 8: Relative magnitude of wage costs due to absenteeism

Table 6: Mean absenteeism and gender differences

(A) blue-collar workers

	obs	mean	std. err.	std. dev.
male	41094	6.7253	0.0930	18.8515
female	9628	8.9216	0.2125	20.8466
combined	50722	7.1422	0.0855	19.2652
difference		-2.1963	0.2179	

(B) white-collar workers

	obs	mean	std. err.	std. dev.
male	53045	3.2186	0.0554	12.7700
female	20129	3.9358	0.0940	13.3312
combined	73174	3.4159	0.0478	12.9307
difference		-0.7172	0.1070	

Note: Mean absent working hours per month. Gender differences are significant at the 1%-level in a two sample t-test.

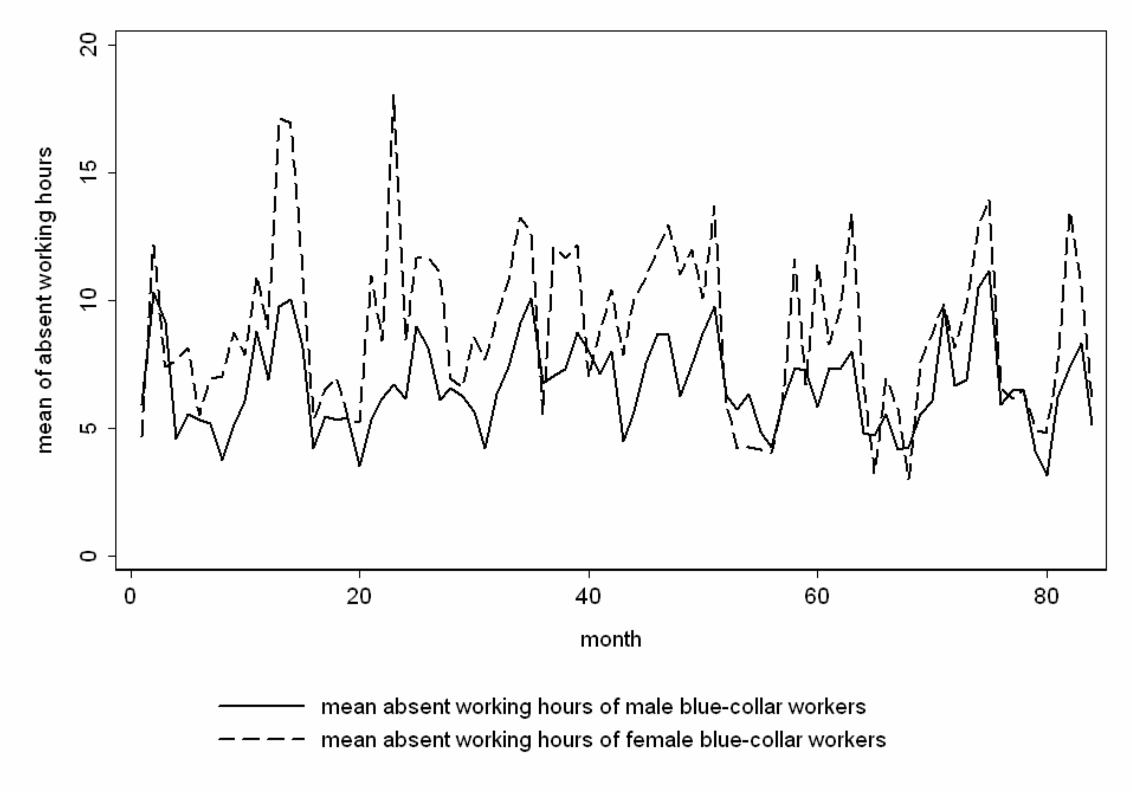

Figure 9: Mean absent working hours of male and female blue-collar workers

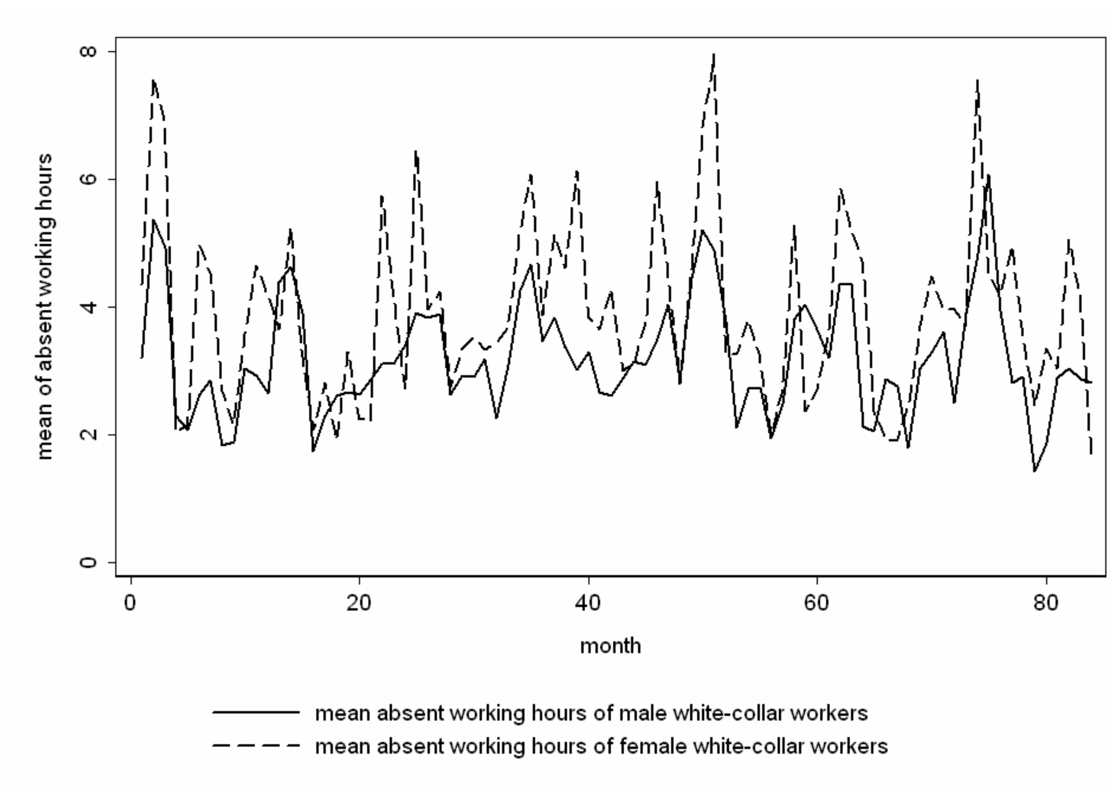

Figure 10: Mean absent working hours of male and female white-collar workers

Table 7: Estimates of the gender wag gap for blue-collar workers

	blue-collar workers		
female (dummy)	-0.1471***	-0.1288***	-0.0353***
	[0.0073]	[0.0073]	[0.0058]
tenure (years)		0.0155***	0.0101***
		[0.0004]	[0.0003]
tenure squared / 100		-0.0528***	-0.0452***
		[0.0014]	[0.0012]
tenure cubed / 1000		0.0058***	0.0058***
		[0.0002]	[0.0002]
entry age (years)		-0.0073***	-0.0037**
		[0.0020]	[0.0015]
entry age squared / 100		0.0114***	0.0060***
		[0.0031]	[0.0023]
apprenticeship degree (dummy)		0.0602***	0.0147***
		[0.0068]	[0.0049]
monthly working hours		0.0013***	0.0013***
		[0.0001]	[0.0001]
monthly working hours squared / 100		-0.0006***	-0.0007***
		[0.0001]	[0.0001]
level 2 (dummy)			0.0206***
			[0.0013]
level 3 (dummy)			0.0359***
			[0.0024]
level 4 (dummy)			0.0815***
			[0.0027]
level 5 (dummy)			0.1849***
			[0.0032]
level 6 (dummy)			0.2394***
			[0.0037]
level 7 (dummy)			0.3043***
			[0.0043]
month	yes	yes	yes
constant	2.5882***	2.5169***	2.4182***
	[0.0042]	[0.0354]	[0.0265]
observations	50722	50722	50722
individuals	786	786	786
R-squared (overall)	0.3810	0.4907	0.7458

Note: Random effects GLS estimates for log hourly wage. Heteroskedasticity robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Estimates of the gender wag gap for white-collar workers

	white-collar workers		
female (dummy)	-0.2592***	-0.1306***	-0.0819***
	[0.0142]	[0.0102]	[0.0064]
tenure (years)		0.0396***	0.0290***
		[0.0005]	[0.0004]
tenure squared / 100		-0.1269***	-0.0944***
		[0.0020]	[0.0017]
tenure cubed / 1000		0.0128***	0.0096***
		[0.0003]	[0.0003]
entry age (years)		0.0378***	0.0120***
		[0.0029]	[0.0013]
entry age squared / 100		-0.0278***	-0.0006
		[0.0041]	[0.0019]
high school degree (dummy)		0.0283**	0.0126
		[0.0141]	[0.0089]
university degree (dummy)		0.1971***	0.1141***
		[0.0108]	[0.0068]
monthly working hours		0.0002	0.0010***
		[0.0001]	[0.0001]
monthly working hours squared / 100		0.0000	-0.0004***
		[0.0000]	[0.0000]
level 2 (dummy)			0.0977***
			[0.0017]
level 3 (dummy)			0.1769***
			[0.0023]
level 4 (dummy)			0.2521***
			[0.0028]
level 5 (dummy)			0.3183***
			[0.0032]
level 6 (dummy)			0.3800***
			[0.0038]
month	yes	yes	yes
constant	3.0066***	1.8118***	2.2082***
	[0.0084]	[0.0486]	[0.0238]
observations	73174	73174	73174
individuals	1250	1250	1250
R-squared (overall)	0.1769	0.4351	0.7451

Note: See Table 7.

Table 9: Estimates by gender for blue-collar workers

	men	women	decomposition explained	as %	unexplained	as %
tenure (years)	0.0165***	0.0112***	0.0100***	1.5	0.0805***	7.6
	[0.0005]	[0.0010]	[0.0014]		[0.01607]	
tenure squared / 100	-0.0524***	-0.0673***	-0.0338***	-2.6	0.0476***	4.0
	[0.0015]	[0.0041]	[0.0029]		[0.0142]	
tenure cubed / 1000	0.0054***	0.0113***	0.0252***	1.2	-0.0466***	-3.4
	[0.0002]	[0.0008]	[0.0021]		[0.0069]	
entry age (years)	-0.0076***	-0.0031	0.0045	1.1	-0.1225	-12.9
	[0.0024]	[0.0033]	[0.0048]		[0.1097]	
entry age squared / 100	0.0113***	0.0042	-0.0034	-0.9	0.0554	6.1
	[0.0038]	[0.0050]	[0.0041]		[0.0490]	
apprenticeship (dummy)	0.0728***	0.011	0.0032	2.1	0.0481***	3.0
	[0.0079]	[0.0118]	[0.0034]		[0.0111]	
monthly working hours	0.0010***	0.0015***	0.0061***	0.4	-0.0709	-6.9
	[0.0002]	[0.0002]	[0.0007]		[0.0446]	
monthly working hours squared / 100	-0.0005***	-0.0008***	-0.0086***	-0.5	0.0671**	6.4
	[0.0001]	[0.0001]	[0.0009]		[0.0286]	
month	yes	yes				
constant	2.5090***	2.4060***				
	[0.0426]	[0.0543]				
observations	41094	9628				
individuals	632	154				
R-squared (overall)	0.3853	0.3102				

Note: See Table 7.

Table 10: Estimates by gender for white-collar workers

			decomposition			
	men	women	explained	as %	unexplained	as %
tenure (years)	0.0411***	0.0356***	0.0696***	8.0	0.0792***	6.8
	[0.0006]	[0.0009]	[0.0033]		[0.0164]	
tenure squared / 100	-0.1281***	-0.1252***	-0.0752***	-7.7	-0.0088	-0.7
	[0.0022]	[0.0046]	[0.0041]		[0.0152]	
tenure cubed / 1000	0.0129***	0.0132***	0.0254***	2.5	-0.0025	-0.2
	[0.0004]	[0.0008]	[0.0019]		[0.0068]	
entry age (years)	0.0450***	0.0201***	0.0510***	11.4	0.7226***	66.0
	[0.0022]	[0.0047]	[0.0120]		[0.1504]	
entry age squared / 100	-0.0345***	-0.0126*	-0.0180*	-4.9	-0.1972***	-16.6
	[0.0029]	[0.0069]	[0.0099]		[0.0669]	
high school (dummy)	0.0113	0.0342*	-0.0067*	-0.2	-0.0025	-0.7
	[0.0191]	[0.0186]	[0.0037]		[0.0029]	
university (dummy)	0.1869***	0.2280***	0.0398***	3.3	-0.0164*	-0.9
	[0.0117]	[0.0207]	[0.0037]		[0.0095]	
monthly working hours	0.0030***	0.0000	0.0007	5.6	0.4624***	40.8
	[0.0006]	[0.0001]	[0.0024]		[0.0887]	
monthly working hours squared / 100	-0.0009***	0.0001	0.0042	-4.1	-0.2314***	-18.6
	[0.0002]	[0.0001]	[0.0028]		[0.0480]	
month	yes	yes				
constant	1.4383***	2.0606***				
	[0.0579]	[0.0735]				
observations	53045	20129				
individuals	887	363				
R-squared (overall)	0.373	0.3222				

Note: See Table 7.

Table 11: Estimates by gender for white-collar workers with hierarchical levels

	men	women	decomposition			
			explained	as %	unexplained	as %
tenure (years)	0.0299***	0.0270***	0.0528***	5.9	0.0430***	3.7
	[0.0005]	[0.0006]	[0.0024]		[0.0115]	
tenure squared / 100	-0.0957***	-0.0946***	-0.0568***	-5.7	-0.0032	-0.3
	[0.0019]	[0.0037]	[0.0032]		[0.0125]	
tenure cubed / 1000	0.0096***	0.0105***	0.0202***	1.8	-0.0066	-0.5
	[0.0003]	[0.0007]	[0.0016]		[0.0056]	
entry age (years)	0.0161***	0.0041**	0.0104**	4.1	0.3484***	31.8
	[0.0017]	[0.0018]	[0.0046]		[0.0726]	
entry age squared / 100	-0.0037	0.0032	0.0046	-0.5	-0.0627**	-5.3
	[0.0023]	[0.0027]	[0.0038]		[0.0318]	
high school (dummy)	0.0059	0.0081	-0.0016	-0.1	-0.0002	-0.1
	[0.0137]	[0.0081]	[0.0016]		[0.0017]	
university (dummy)	0.1114***	0.1170***	0.0204***	1.9	-0.0022	-0.1
	[0.0084]	[0.0098]	[0.0018]		[0.0051]	
monthly working hours	0.0043***	0.0003***	0.0063***	8.0	0.6204***	54.7
	[0.0004]	[0.0001]	[0.0021]		[0.0665]	
monthly working hours squared / 100	-0.0016***	-0.0001***	-0.0071***	-7.5	-0.3469***	-27.9
	[0.0001]	[0.0001]	[0.0025]		[0.0369]	
level 2 (dummy)	0.1041***	0.0898***	-0.0137***	-1.6	0.0028***	0.5
	[0.0023]	[0.0023]	[0.0005]		[0.0006]	
level 3 (dummy)	0.1790***	0.1777***	0.0019***	0.2	0.0003	0.0
	[0.0030]	[0.0036]	[0.0006]		[0.0012]	
level 4 (dummy)	0.2506***	0.2710***	0.0231***	2.1	-0.0035***	-0.2
	[0.0033]	[0.0055]	[0.0008]		[0.0011]	
level 5 (dummy)	0.3115***	0.3682***	0.0382***	3.2	-0.0080***	-0.2
	[0.0037]	[0.0066]	[0.0010]		[0.0011]	
level 6 (dummy)	0.3855***	0.3816***	0.0564***	5.7	0.0007	0.0
	[0.0043]	[0.0081]	[0.0015]		[0.0017]	
month	yes	yes				
constant	1.8717***	2.3580***				
	[0.0431]	[0.0288]				
observations	53045	20129				
individuals	887	363				
R-squared (overall)	0.6878	0.7807				

Note: See Table 7.

Table 12: Summary of decomposition results (as percent)

	blue-collar workers	white-collar workers	white-collar workers with hierarchical levels
raw differential (R)	15.0	25.3	24.4
explained differential (E)	1.9	13.8	17.5
unexplained differential (U)	13.1	11.6	6.9
explained as % total (E/R):	12.9	54.3	71.7
unexplained as % total (U/R):	87.1	45.7	28.3

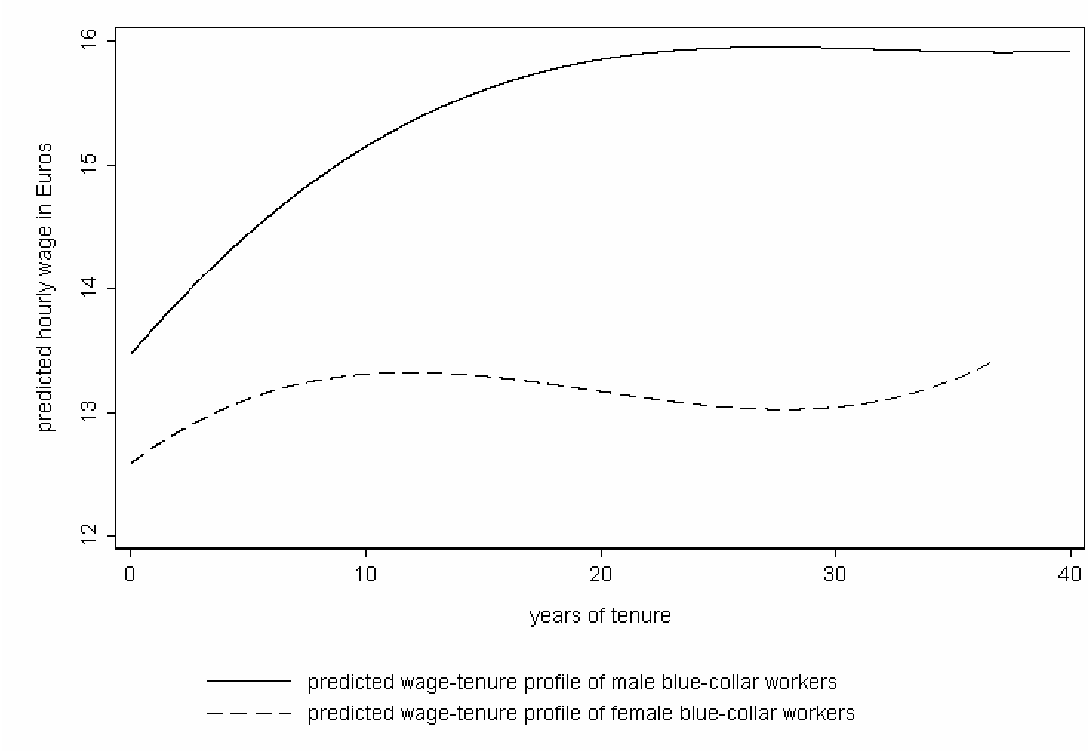

Figure 11: Predicted wage-tenure profile for male and female blue-collar workers

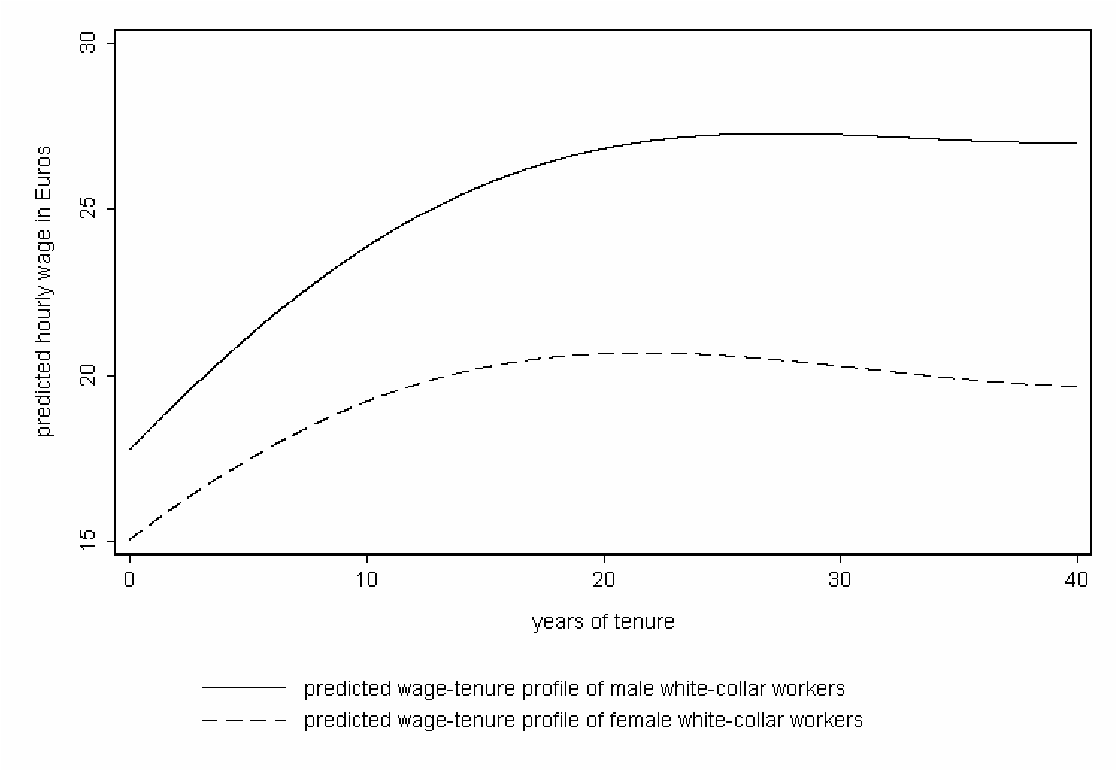

Figure 12: Predicted wage-tenure profile for male and female white collar workers

Table 13: Gender wage gap and tenure

	without control for levels		with control for levels	
	blue-collar	white-collar	blue-collar	white-collar
female (dummy)	-0.0714***	-0.1492***	0.0117	-0.0942***
	[0.0089]	[0.0111]	[0.0081]	[0.0071]
interaction female * tenure	-0.0041***	0.0013	-0.0033***	0.0009
	[0.0007]	[0.0008]	[0.0006]	[0.0007]
interaction female * squared tenure	-0.0144***	0.0005	-0.0076**	-0.0014
	[0.0042]	[0.0051]	[0.0038]	[0.0042]
interaction female * cubed tenure	0.0057***	0.0007	0.0033***	0.0012
	[0.0008]	[0.0009]	[0.0008]	[0.0008]
tenure (years)	0.0164***	0.0393***	0.0107***	0.0289***
	[0.0004]	[0.0006]	[0.0003]	[0.0004]
tenure squared / 100	-0.0523***	-0.1276***	-0.0446***	-0.0950***
	[0.0015]	[0.0022]	[0.0013]	[0.0019]
tenure cubed / 1000	0.0054***	0.0128***	0.0054***	0.0096***
	[0.0002]	[0.0004]	[0.0002]	[0.0003]
controlled for levels	No	No	Yes	Yes
observations	50722	73174	50722	73174
individuals	786	1250	786	1250
R-squared (overall)	0.4978	0.4331	0.7453	0.7442

Note: Random effects GLS estimates for log hourly wage. Heteroskedasticity robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Earnings functions control for entry age, squared entry age, schooling, working hours, squared working hours, and month.

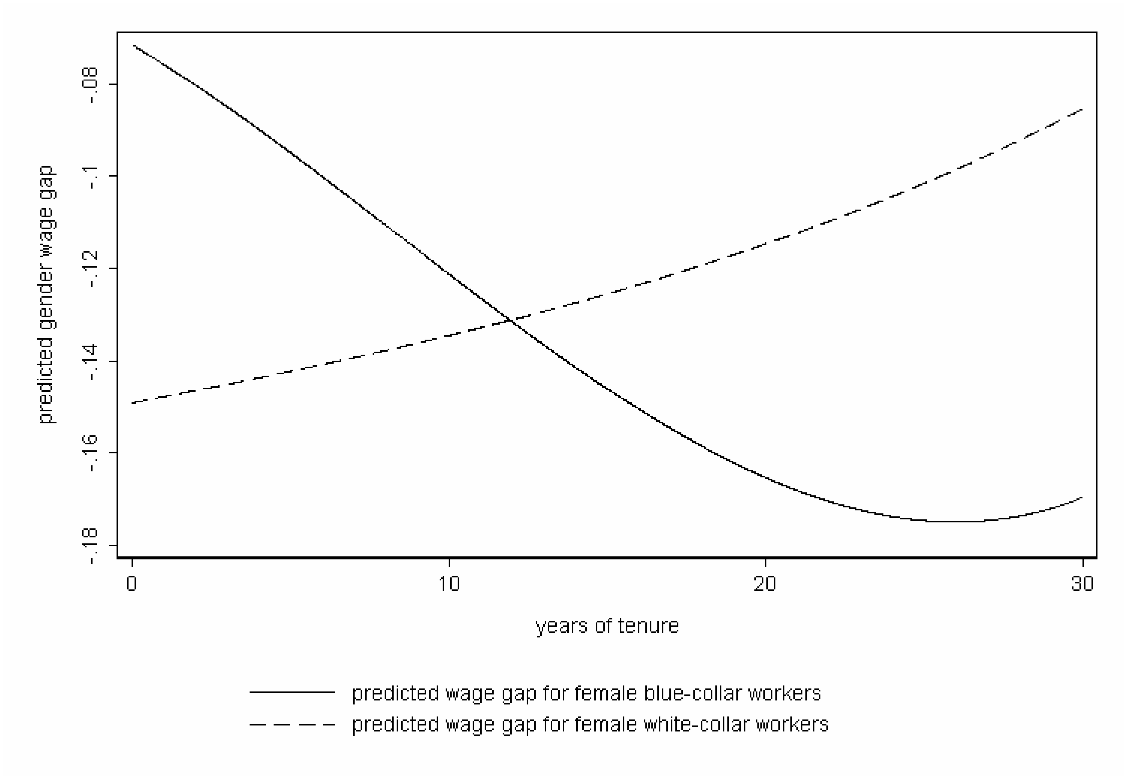

Figure 13: Predicted gender wage gaps

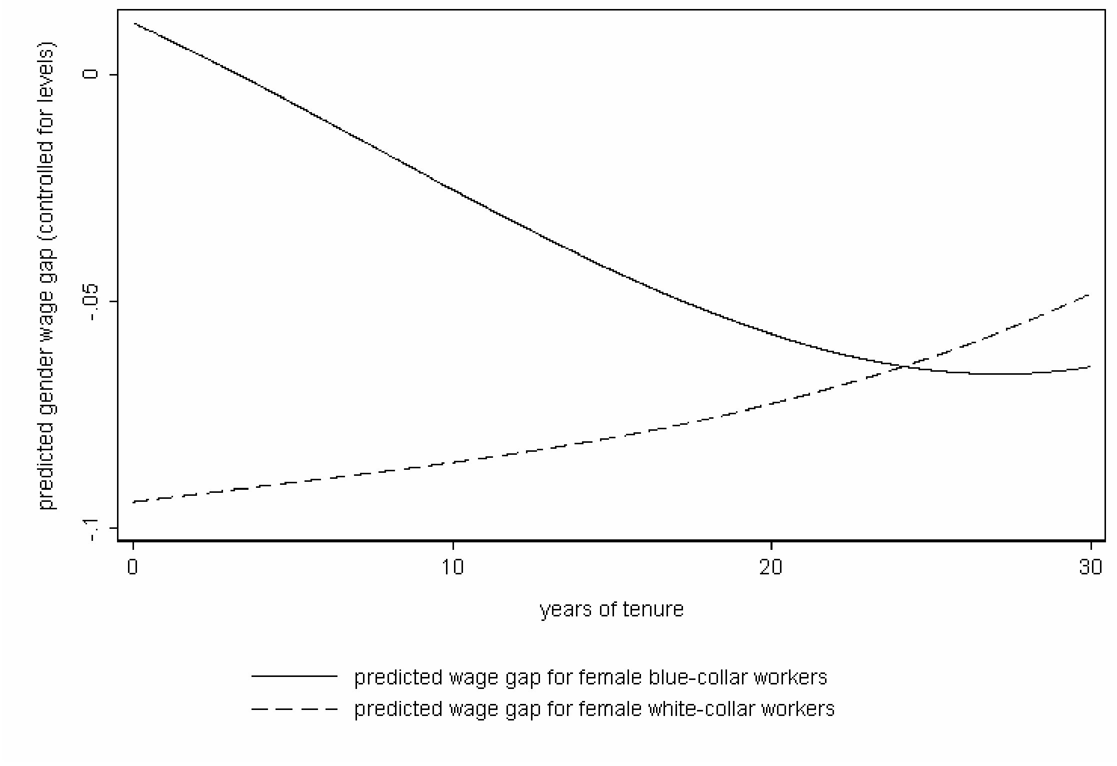

Figure 14: Predicted gender wage gaps with control for hierarchical levels

Table 14: Wage effect of absenteeism for blue-collar workers

(A) sample: all blue-collar workers (N=50722, n=786)

female (dummy)	-0.1288***	-0.1288***	-0.1296***	
	[0.0073]	[0.0072]	[0.0081]	
absent working hours in month / 100		0.0004		
		[0.0006]		
absent working hours in month (IV)			0.0003*	
			[0.0002]	
R-squared (overall)	0.4907	0.4907	0.4864	

(B) sample: all blue-collar workers with at least 24 month of observed employment history with the firm (N=31931, n=676)

female (dummy)	-0.1373***	-0.1373***	-0.1374***	-0.1372***
	[0.0076]	[0.0076]	[0.0076]	[0.0076]
absent working hours in month / 100		0.0008		
		[0.0006]		
sum of absent hours in last year / 100			0.0003	
			[0.0002]	
sum of absent hours in last two years / 100				-0.0002
				[0.0002]
R-squared (overall)	0.4403	0.4403	0.4401	0.4406

(C) sample: all blue-collar workers in December 2005 (N=553, n=553)

female (dummy)	-0.1287***	-0.1286***	-0.1254***	
	[0.0087]	[0.0087]	[0.0088]	
absent working hours in month		-0.0002		
		[0.0002]		
mean absent monthly hours (own employment history)			-0.0015**	
			[0.0006]	
R-squared	0.4539	0.4546	0.4592	

Note: Random effects GLS for sample A and B. OLS for sample C. Heteroskedasticity robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. All estimates control for tenure, squared tenure, cubed tenure, entry age, squared entry age, schooling, working hours, squared working hours, and month.

Table 15: Wage effect of absenteeism for white-collar workers

(A) sample: all white-collar workers (N=73174, n=1250)

female (dummy)	-0.1306***	-0.1306***	-0.1273***	
	[0.0102]	[0.0101]	[0.0191]	
absent working hours in month / 100		0.0021**		
		[0.0010]		
absent working hours in month (IV)			-0.0013***	
			[0.0005]	
R-squared (overall)	0.4351	0.4356	0.4042	

(B) sample: all white-collar workers with at least 24 month of observed employment history with the firm (N=46165, n=1021)

female (dummy)	-0.1373***	-0.1373***	-0.1373***	-0.1372***
	[0.0116]	[0.0115]	[0.0115]	[0.0114]
absent working hours in month / 100		-0.0001		
		[0.0011]		
sum of absent hours in last year / 1000			0.0001	
			[0.0033]	
sum of absent hours in last two years / 100				-0.0002
				[0.0002]
R-squared	0.4567	0.4569	0.4569	0.4571

(C) sample: all white-collar workers in December 2005 (N=849, n=849)

female (dummy)	-0.0779***	-0.0779***	-0.0712***	
	[0.0139]	[0.0140]	[0.0139]	
absent working hours in month		0.0001		
		[0.0006]		
mean absent monthly hours			-0.0075***	
(own employment history)			[0.0019]	
R-squared	0.6236	0.6236	0.6296	

Note: See Table 14.

Table 16: Gender gap in absent working hours

	blue-collar workers		white-collar workers	
female (dummy)	2.0739***	2.2047***	0.8361***	1.4528***
	[0.5640]	[0.5825]	[0.2789]	[0.3004]
tenure (years)		0.4475***		0.2575***
		[0.1105]		[0.0531]
tenure squared / 100		-2.4967***		-1.3872***
		[0.7093]		[0.3474]
tenure cubed / 1000		0.4251***		0.2171***
		[0.1290]		[0.0610]
entry age (years)		0.8447***		0.4208***
		[0.1677]		[0.1020]
entry age squared / 100		-1.2860***		-0.5454***
		[0.2605]		[0.1554]
apprenticeship degree (dummy)		-0.4796		
		[0.5141]		
high school degree (dummy)				-1.3246***
				[0.3355]
university degree (dummy)				-2.1722***
				[0.3089]
monthly working hours		-0.1249		0.0642***
		[0.1171]		[0.0165]
monthly working hours squared / 100		0.0742		-0.0172**
		[0.0482]		[0.0072]
month	yes	yes	yes	yes
constant	5.1335***	-7.3855	3.1192***	-10.0474***
	[0.7381]	[7.4094]	[0.5100]	[1.8421]
observations	50722	50722	73174	73174
individuals	786	786	1250	1250
R-squared (overall)	0.0120	0.0149	0.0060	0.0128
gender gap in absenteeism in percent	30.82%	32.33%	25.73%	45.90%

Note: Random effects GLS estimates for number of absent working hours. Heteroskedasticity robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Gender gap in percent is calculated by: 100*(predicted absent working hours females - predicted absent working hours males)/(predicted absent working hours males).

Table 17: Evolution of the gender gap in absent working hours

	blue-collar		white-collar	
female (dummy)	2.2246***	4.9161	1.3236***	7.8579
	[0.5837]	[34.3195]	[0.3000]	[10.6110]
interaction female * age		-0.2966		-0.5242
		[2.4925]		[0.8411]
interaction female * squared age		0.7415		1.4424
		[5.9314]		[2.1418]
interaction female * cubed age		-0.0464		-0.1331
		[0.4621]		[0.1760]
age (years)	0.8954	1.028	0.8811**	1.0023**
	[0.6496]	[0.6816]	[0.3902]	[0.4626]
age squared / 100	-2.4966	-2.8151	-2.1829**	-2.4704**
	[1.6306]	[1.7220]	[0.9767]	[1.1533]
age cubed / 1000	0.2270*	0.2498*	0.1801**	0.2036**
	[0.1322]	[0.1403]	[0.0789]	[0.0927]
observations	50722	50722	73174	73174
individuals	786	786	1250	1250
R-squared (overall)	0.0128	0.0132	0.0109	0.0110

Note: Random effects GLS estimates for absent working hours. Heteroskedasticity robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Estimates control for schooling, working hours, squared working hours, and month.

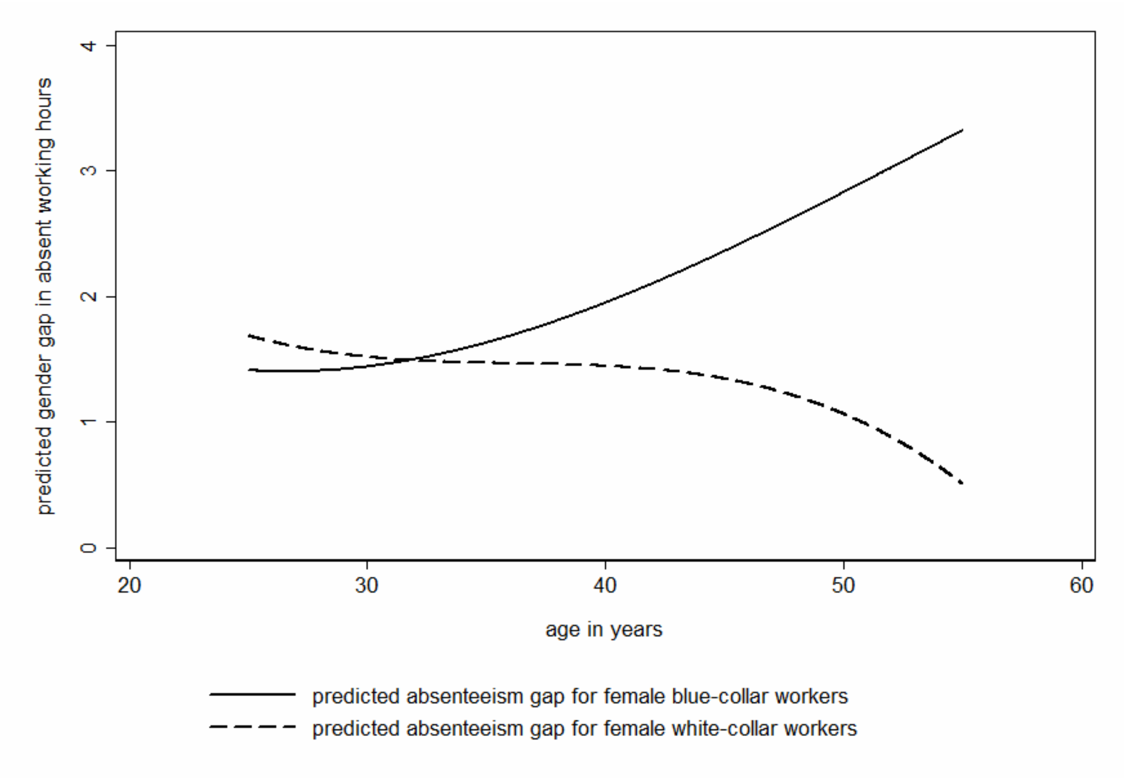

Figure 15: Predicted gender gaps in absent working hours

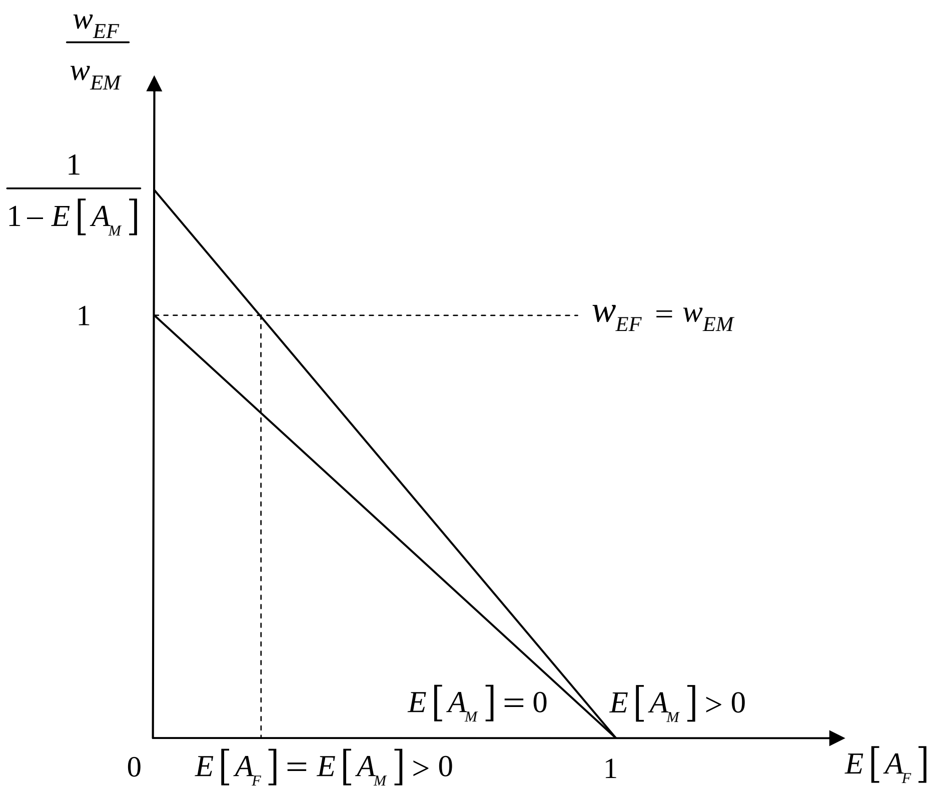

Figure 16: Absenteeism and the female-male wage ratio

Table 18: Numerical example for GWG effect of statistical discrimination

	blue-collar workers	white-collar workers
female mean absent working hours	8.92	3.94
female mean contractual working hours	147.41	138.12
female mean absenteeism rate	6.05%	2.85%
male mean absent working hours	6.73	3.22
male mean contractual working hours	151.57	156.73
male mean absenteeism rate	4.44%	2.05%
female-male wage ratio	98.32%	99.18%
gender wage gap	-1.7%	-0.8%