

The gender wage gap - due to differences in efficiency wage effects or discrimination?

Citation for published version (APA):

Schwieren, C. A. A. (2003). *The gender wage gap - due to differences in efficiency wage effects or discrimination?* METEOR, Maastricht University School of Business and Economics. METEOR Research Memorandum No. 028 <https://doi.org/10.26481/umamet.2003028>

Document status and date:

Published: 01/01/2003

DOI:

[10.26481/umamet.2003028](https://doi.org/10.26481/umamet.2003028)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

The gender wage gap – due to differences in efficiency wage effects or discrimination?

Christiane Schwieren*

Abstract

Women often receive lower wages than men for comparable work. Many explanations are offered for this fact, ranging from women's lower negotiation skills to discrimination by employers. In this paper, an experiment, which was originally conceptualized to test efficiency-wage theory, has been applied to test whether women get paid less than men in an experimental market, and if this is the case, why. The experiment is a variant of Fehr & Falk's (1999) double auction with effort. Results are striking: Female workers receive significantly lower wages than male workers, no matter whether men or women are in the role of the firm. However, this does not pay for the firms, as women's reactions to low wages are equal to those of men: low effort. More specifically, a high discrepancy between the wage asked by a worker and the wage offered by the firm leads to low effort. Extrapolating from the experiment to the "real" labor market, the results are pointing towards a vicious cycle: Women are offered lower wages than they expect, and consequentially they exhibit low effort levels. Therefore, employers who do not realize that women – just like men – reciprocate might regard their productivity as lower.

Keywords: Discrimination, competition, efficiency wages, experiment

*Department of Economics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands, tel.: +31 43 3883652; fax: +31 43 3884878; e-mail: c.schwieren@algec.unimaas.nl

1. Introduction

It is still the case that women often earn less than men in the same jobs. Economics is providing a lot of empirical evidence that women earn lower wages than men, but the estimates of the wage gap vary widely with the variables included, the decomposition methods and the database used etc. (Stanley & Jarell, 1998). Most studies report a trend towards a narrowing of the wage gap during the last forty years (e.g., Blau & Kahn, 1994, 1997, 2000; Ashraf, 1996). Most of the research is done in the U.S., but studies dealing with European countries find similar patterns, with a somewhat better general situation in Scandinavian countries than in the U.S., but also more stagnation in the development of the wage gap in the 90s (e.g., Blau & Kahn, 1994; Datta Gupta et al., 1999), and a somewhat worse situation in Germany (e.g., Mavromaras & Rudolph, 1999). As explanatory variables for the change in the wage gap education and other human capital factors (e.g., Anderson & Shapiro, 1996), and more recently also institutional and situational variables (e.g., Fortin & Lemineux, 1998) are used. Still, in most estimates of the wage gap remains an unexplained part, which is usually interpreted as stemming from discrimination. Different economic and social psychological theories are used to explain discrimination (e.g., Becker 1957, Phelps 1972, Arrow 1973, Fiske 1998). One area of economic reasoning, however, has so far not been used to explain the wage gap, even if it seems quite obvious that it could contribute to an explanation: efficiency wage theory (see, however, Haagsma 1993, 1995).

Efficiency-wage theory basically states that employers don't pay the lowest possible wage, but rather pay a higher wage (the lowest possible wage in efficiency-units) to motivate workers to exhibit higher effort, to keep them from shirking or to attract better workers (see for an overview e.g. Katz, 1986; Carmichael, 1990). In the framework of efficiency-wage theory, women's lower wages could be explained by assuming that efficiency-wages for women are different or believed to be different from those of men. This could be the case in two ways: First, firms could assume (wrongly or correctly) that women exhibit high effort/don't shirk already for low wages, i.e., lower wages than for men are sufficient to get the desired behavior from female workers. Second, firms could assume (again wrongly or correctly) that women do not react with more effort to higher wages (e.g. because they are not work/career-oriented anyway), so it might not be worth it to pay them high wages. In that case, firms would either employ women for the minimum possible wage or not employ them at all. Both assumptions can be based on stereotypes about women, describing in general men as more agentic, assertive and work-oriented, and women as more communal, caring, and

family-oriented (for a review on gender stereotypes see e.g. Deaux & LaFrance 1998, Glick et al. 1995). It is quite probable that gender stereotypes play a role in the labor market, as gender (and race) are very salient categories and thus are likely to lead to stereotype activation. In addition, typical sex-roles can have an impact on employers' beliefs about the reliability of women at the workplace as well: As women most of the times are the ones who take care of the children, they are believed to be absent or quitting work more often than men (for a discussion of gender differences in turnover-rates see e.g. Viscusi, 1980; Blau & Kahn, 1981; Sicherman, 1996; Light & Ureta, 1992). If the stereotype-based assumptions about female reactions to incentives are correct, paying women lower wages is rational behavior. However, if the assumptions are not correct, lower wages for women might be economically inefficient, and a more "psychological" explanation for such lower wages might be useful.

In recent times psychological research has put forward an explanation which can be interpreted as similar to an efficiency-wage explanation: the possibility that women are worse in negotiating for their salary and ask for lower wages than men in wage negotiations (see e.g. Pelham & Hetts, 2001, Walters, Stuhlmacher & Meyer, 1998). In an efficiency-wage framework one could interpret this as women credibly promising the effort desired by the firm already for lower wages than men.

The present paper describes an experiment conducted to observe this in a controlled setting. The main reason for the choice of the experimental method lies in the fact that it allows to control for other factors influencing wages offered to or asked by women, which are hard to control for in other kinds of empirical studies. Furthermore, it allows for measuring effort level in a simple way, which is often very difficult in real-world settings.

In an experimental double auction both workers and firms can make wage offers. This makes it possible to see whether women get lower wages than men and whether they primarily ask for these lower wages. Further, it is possible to see whether efficiency-wage arguments in one or the other direction provide an explanation for the wage differential. In a laboratory setting stereotypes about female and male work-attitudes should play a smaller role than in "real life". However, assuming with social psychology that stereotypes are universal and also work subconsciously (e.g. Fiske 1998), one should still expect to find an effect – and if an effect is found in a laboratory setting, the effect in real-life should be even bigger.

Double auction markets are very competitive, and have been used before to test efficiency-wage theory, but also to test other related questions (see e.g. Davis & Holt, 1993). In similar experimental double auctions with effort, it was found that firms don't pay low wages, and that workers reciprocate with high effort when getting high wages (e.g., Charness,

1998; Fehr & Falk, 1999; Hannan, Kagel, & Moser 2002). In a double auction dealing with status differences between traders women – just like subjects who were assigned lower status - made lower profits than men selling “goods” (Ball, Eckel, Grossman & Zame, 2001).

The rest of the paper is organized as follows. First the hypotheses are formulated. In section three, the experimental design is described in detail. Section four reports the results, and section five discusses them and concludes.

2 Hypotheses

In summary, the background of the experiment described here is a combination of efficiency-wage effects with knowledge about common stereotypes of women and men. On this background, the following hypotheses can be formulated:

H1: Gender of employees influences wages, such that women receive lower wages (w) than men in the experimental double auction: $w_f < w_m$

H2: If H1 holds, it holds, because efficiency-wage effects with respect to the effort (e) - wage relationship are different for women and men: $e_f(w) \neq e_m(w)$. Two variants, based on common gender stereotypes, are possible:

H2a): As women are on average more cooperative than men, they exhibit high effort already for low wages: for given (low) w that would imply $e_f > e_m$

H2b): As women are on average not very achievement-oriented, they never exhibit high effort, no matter how high the wage: for given (high) w that would imply $e_f < e_m$

H3: If H1 holds, it holds, because employers expect hypothesis two to hold whereas in fact it is not true (i.e., they discriminate against women as a group): $e_f^e(w) \neq e_m^e(w)$

H4: If H1 holds, it holds, because women ask for lower wages, i.e., credibly promise already for a low wage to exhibit the amount of effort desired by the employer: $w_f^a(e) < w_m^a(e)$

3 The experiment

Participants

190 (94 male and 96 female) students of various faculties (the majority from economics and business) at Universitat Pompeu Fabra in Barcelona participated voluntarily, for performance-based payment, in the experiment.

Design

The design was adapted from Fehr & Falk’s (1999) study of underbidding on the labor market. The basic structure is a double auction with effort, with excess supply of workers (four firms and six workers). Each firm can only contract one worker per period. Workers and firms can make wage offers simultaneously, following an improvement rule. This means for

workers that each new wage offer has to be lower than the last offer made before (by any of the workers). For firms, each new wage offer has to be higher than the last offer made before (by any of the firms). All offers can be accepted at any point in time, i.e., ‘old’ offers, which are still standing, can be accepted as well. No one has to accept an offer, i.e., firms do not have to conclude a contract with a worker, and workers do not have to conclude a contract with a firm. After three minutes trading is ended, no matter how many contracts are concluded. In a second step, when all contracts are concluded or the three minutes are over, those workers who concluded a contract select an effort level (a percentage) which is costly for them and which determines their final payoffs and the profits of their firm. Workers can determine the effort level such that firms make losses¹. Firms’ payoffs π_F are calculated as follows:

$\pi_F = 120 * \text{percentage chosen by the worker} - \text{salary (between 20 and 120)}$

Workers’ payoffs π_W are calculated like this:

$\pi_W = \text{salary} - \text{costs of percentage}$
--

If a firm does not conclude a contract with a worker, the firm does not earn any money in that period. If a worker does not get a contract, he or she still gets twenty experimental guilders. Costs for the effort-levels (percentages) are defined as follows, based on the design by Fehr, Kirchsteiger & Riedl (1998):

Percentage	10	20	30	40	50	60	70	80	90	100
Cost	0	1	2	3	4	8	10	12	15	18

The payoffs in experimental guilders were transformed in €-cent by multiplying them by two.

The experiment was run with all possible combinations of sexes of firms and workers, and everybody always knew the sex of workers resp. firms. In the four “experimental” treatments, workers and firms were always homogeneous with respect to sex. Three (control) sessions were mixed, i.e., workers and firms were distributed into their roles randomly, regardless of their sex. This resulted in five different treatments, with four resp. three (in the mixed treatment) sessions of each treatment. Ten periods of the auction with effort were played in each session.

¹ The earnings-function for the firms differed from those used by Fehr & Falk (1999) in that the function used here allows for firms to make losses.

	Sex firm			
	male	female	mixed	
Sex worker	male	4	4	-
	female	4	4	-
	mixed	-	-	3

Table 1: Number of sessions for each constellation.

In the following, the treatments will be referred to mentioning always sex of worker first, and then sex of firm, i.e., worker male – firm female will be called male-female or mf in the graphs, and worker female - firm male female-male resp. fm.

Procedure

The experiment was conducted computerized in the experimental laboratory of Universitat Pompeu Fabra. A program, which Fehr & Falk had used in one of their earlier double-auction experiments, was adapted. The experiment was programmed in z-Tree (Fischbacher 1999).

Upon arrival, students were divided into the two roles of firms and workers. In the sex-homogeneous sessions and the mixed sessions, this was done randomly. All participants gathered in front of the lab and draw lots with either “firm” or “worker” on it. Then, firms and workers entered the lab separately. In sex-heterogeneous sessions, this constituted the basic experimental manipulation: Before starting the experiment, the experimenter said loudly outside of the lab, in front of all participants “The girls (boys) come with me now, you are firms!” When the “firms” were seated, he brought the remaining subjects into the laboratory and said loudly when entering “The boys (girls) are workers”. Then, worker/employer instructions were distributed. Participants in the experiment were seated separately in front of their computers and could not see the other participants’ screens. Subjects read the instructions and completed one page of exercises to test their understanding of the payoff structure. From the moment students were seated in their cubicles, no communication was allowed and questions about the experiment were only answered privately. While participants filled in the exercises the experimenter watched them carefully and pointed them to mistakes, to secure understanding of the payoff structure of the experiment. When everybody had completed the exercises, the auction started. Ten periods of the double-auction with effort were played. At the end of the experiment, subjects filled in a short questionnaire collecting demographical data. They were paid privately, based on what they earned in the experiment plus a 5€ show-up fee. If subjects in the role of firms made losses they had to back it up with their show-up fee.

4 Results

The analysis was done on different levels of aggregation. Most hypotheses were tested using data from period one only, as behavior in later periods is not truly independent from what happened in earlier periods. Some analyses, however, use data over all periods, especially to show differences in trends over time for the treatments. If this is the case it will be indicated. Average wages are calculated only for workers/firms who had concluded a contract. Because of the relatively small number of independent observations, non-parametric Mann-Whitney-U tests were used to compare treatments. Correlational analyses are done using Pearson correlation coefficients, as linear relationships were assumed. However, as especially in the case of effort and wage normal distributions cannot be assumed, Spearman's rank correlations were calculated here as well. For significance testing holds the same as in previous chapters - marginally significant (10%-level) results are reported, but have to be interpreted cautiously. If not otherwise reported, significance testing has been done two-tailed.

In the following, the results are presented by hypothesis, followed by a description of some exploratory results. A final section deals with the mixed sessions separately.

Hypothesis 1:

Hypothesis 1 was fully confirmed: Keeping sex of the firm constant, there is a significant difference in average wage (aggregated on session-level) in period one between female and male workers: Females earn significantly less than males ($Z = -2.21$ as. sig. = .014 (one-tailed)). Figure 1 shows box plots of average wages in period one. The dark line represents the median; the box represents 50% of the data, the lower boundary being the 25th percentile and the upper boundary the 75th percentile. The vertical lines represent values within 1.5 box lengths from the upper or lower edge of the box; values further away are defined as outliers (up to 3 box lengths from the edges) or extreme values. Figure 2 shows the development of the average wage over all periods per constellation.

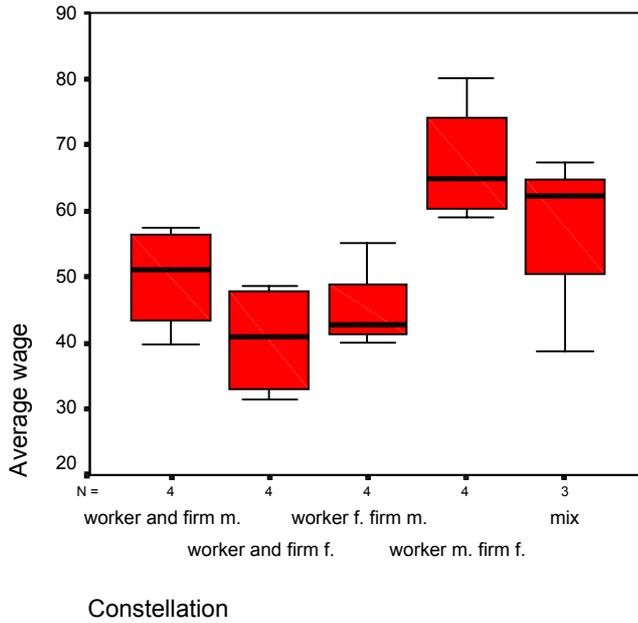


Figure 1: Box plots of average wage per constellation, period one.

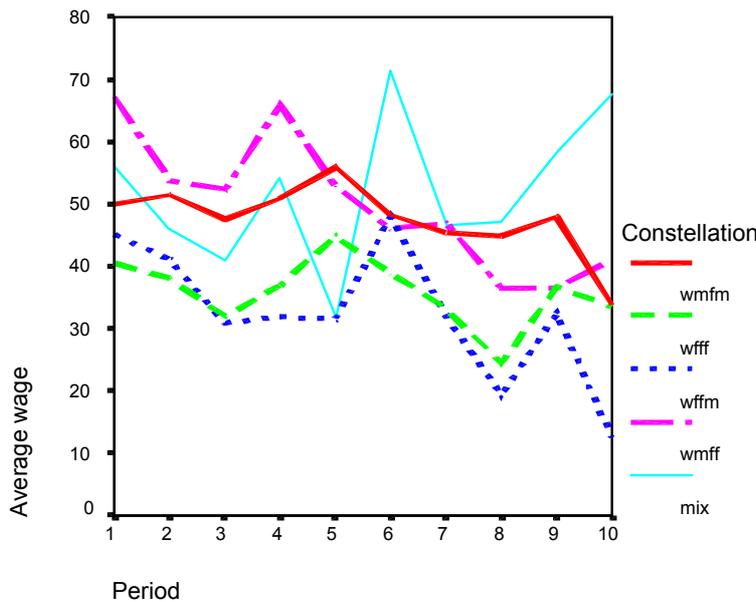


Figure 2: Development of average wage per constellation

Hypotheses 2 and 3:

First, a test for reciprocity in general was done: The correlation of wage with effort, calculated on the individual level (in period one). This correlation is highly significant, indicating that subjects reciprocate – when they get a higher wage, they exhibit more effort (Pearson correlation: $r = .53$, $p = .000$ excluding mixed groups; $r = .49$, $p = .000$ including mixed groups.)² When split by sex of worker, only for men the positive correlation between

² Spearman's rank correlation: $\rho = .45$, $p = .000$ excluding mixed groups, $\rho = .42$, $p = .000$ including mixed groups.

wage and effort level becomes significant at the 5% level, while for women, the correlation is only marginally significant (Pearson correlation for male workers $r = .5$, $p = .002$, for female workers $r = .25$, $p = .081$ (all 1-tailed))³. Thus, in general a positive relationship between wage and effort level exists, but in the first period it becomes clearly significant only for men (see figure 3).

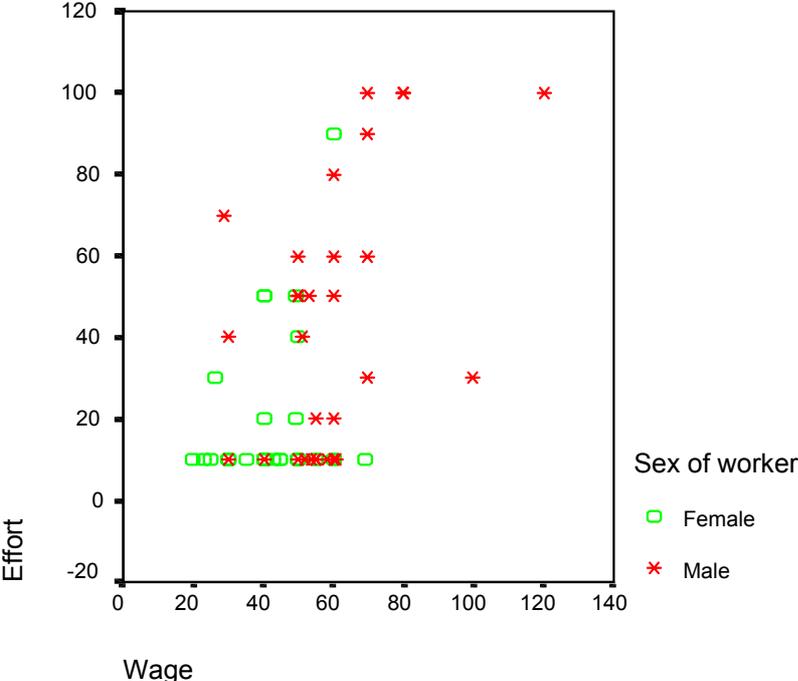


Figure 3: The effort-wage relationship for male and female workers in the first period.

In the following, wages were organized into three categories: Low wages ($20 \leq \text{wage} \leq 45$), a middle category ($46 \leq \text{wage} \leq 65$), and high wages ($\text{wage} \geq 66$). These categories are the same as used by Fehr & Falk (1999), and using them led to approximately the same overall pattern as in their paper, as figure 4 shows: Subjects in the lowest wage category hardly exhibit more than the minimum effort, subjects in the middle-category exhibit slightly higher effort, and only those in the highest wage category really exhibit high effort levels.

³ Spearman's rank correlation (1-tailed) for male workers $\rho = .43$, $p = .007$, for female workers $\rho = .17$, $p = .171$

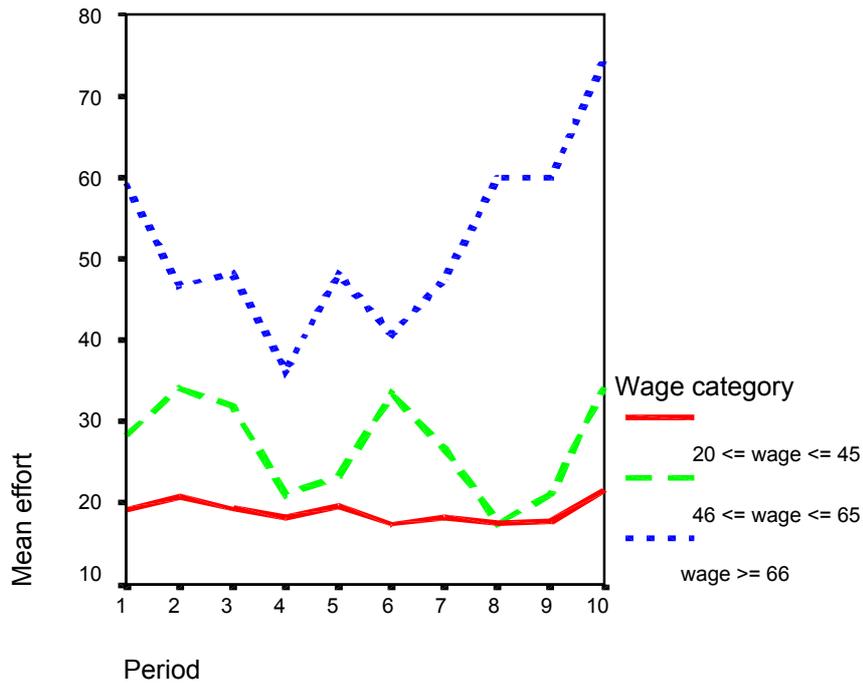


Figure 4: Development of mean effort per wage category (mixed session included).

Women exhibit high effort only very seldom, in all categories. This could, however, be the result of women getting paid less than men in each category. To test for this, U-tests per wage category were calculated.

Wage category	Sex of worker	N	Mean wage	Std. Deviation	Z	As. Sig (2-tailed)
20 <= wage <= 45	Male	4	32.25	5.19	-.70	.484
	Female	18	34.56	7.9		
46 <= wage <= 65	Male	19	54.95	4.25	-1.42	.156
	Female	12	53.17	4.59		
Wage >= 66	Male	8	82.5	18.32	-1.63	.104
	Female	1	69			

Table 2: Mann-Whitney-U test of difference in wages between male and female workers per wage category (period one, individual-level data).

Tables 2 and 3 show that only one woman is in the highest wage category in period one, while most women are in the lowest wage category. This is very different for men – only four men are in the lowest category, most men are in the middle category and still eight are in the highest category. This means that one can only restrictedly interpret an U-test between men and women for the highest and the lowest category, because of a too small N in one of the groups each time. In none of the wage groups, the test finds a significant difference between men and women – neither for wage level nor for effort level.

Wage category	Sex of worker	N	Mean effort	Std. Deviation	Z	As. Sig (2-tailed)
20 <= wage <= 45	Male	4	32.5	28.72	-1.30	.193
	Female	18	16.11	13.35		
46 <= wage <= 65	Male	19	32.11	23.23	-1.32	.188
	Female	12	23.33	24.98		
Wage >= 66	Male	8	76.25	31.59	-1.63	.104
	Female	1	10			

Table 3: Mann-Whitney-U test of difference in effort between male and female workers per wage category (period one, individual-level data).

It seems as if male and female firms expected that male workers would react positively to higher wages, whereas female workers would not. However, only one firm tried to pay a high wage to a female worker, which means, the question whether women do not react to high wages with more effort cannot be answered with these data. As table 4 shows, men also only change effort when moving to wage category 3, and in this category there is only one woman.

Hypothesis 4:

Mann-Whitney-U-tests are calculated for wages asked by women and men, to see whether women ask for lower wages or just get offered lower wages. All offers made in the first period were aggregated for firms and workers separately for this analysis. It was found that offers of firms in the first period are significantly lower when workers are women than when workers are men ($Z = -2.94$, as. sig. = .003). On the other hand, workers of both sexes do not differ in the wages they ask for ($Z = -1.47$, as. sig. = .141). Looking at the development over periods, from period four to six inclusively we also find a difference in the offers of the workers by sex – women obviously learn very quickly that they cannot ask for high wages and adjust their offers. The result is a significantly higher wage asked by male workers than by female workers from period four to six. The offers of firms remain significantly different for male and female workers until period four.

Other results:

For the four homogeneous constellations, over all periods together, a Kruskal-Wallis test (Chi-Square = 33.41, $p = .000$) reveals a significant difference in the number of trades between the constellations. Figure 5 shows clearly that the female-male-constellation resulted in a strongly decreasing number of trades down to zero until the end of the session, while in all other constellations, number of trades varied only between three and four (all possible trades). Testing for sex of firms res. workers, there is a significant difference in average number of trades (over all periods) between male and female workers: when workers are

male, there are on average more trades over all periods than when workers are female ($Z = -3.86, p = .000$) - which can also be seen in figure 5.

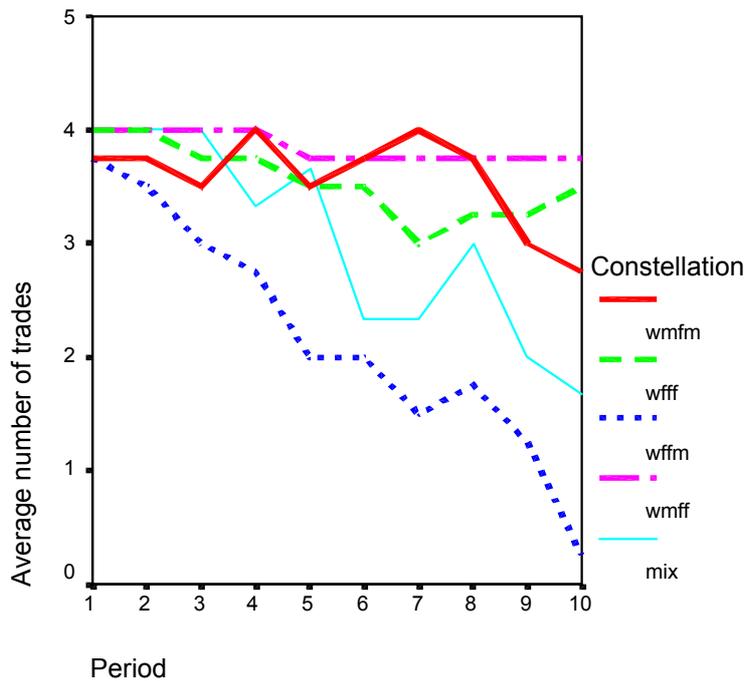


Figure 5: Development of number of trades per constellation.

Male and female firms differed significantly with respect to number of trades as well: Female firms traded more ($Z = -3.97, p = .000$). In fact, male firms often stopped trading at all when they made losses, whereas some female firms went on trading until the last period.

The pattern in the mixed group lies somewhere between the other groups – trades go down as well, not as completely as for male firms and female workers, but stronger than in all other conditions (see below for an explanation). Basically in all those constellations where firms made losses, trading went down over the periods.

A result that could provide an explanation for women’s lower effort is that the difference between offers of firms and wages asked by the workers is higher for female workers than for male workers ($Z = -2.21, \text{as. sig.} = .027$). A t-test to see whether this discrepancy differed significantly from zero was calculated for all constellations (see table 4). It showed that in the male-male-condition, in the male-female-condition, and in the mixed condition discrepancy was not significantly different from zero in period one. In both conditions with female workers, discrepancy was significantly different from zero. This means, only when workers were male or there were at least males among the workers, this discrepancy did not become significant (see also figure 6).

Variable	Constellation	t	df	Sig. (2-tailed)	Mean Difference
Discrepancy	mm	1.02	3	.384	4.74
	ff	5.23	3	.014	18.21
	fm	7.38	3	.005	18.98
	mf	2.32	3	.103	10.27
	mix	.94	2	.447	7.57

Table 4: T-test of difference from zero (test-value = 0) for discrepancy, for all constellations separately (period one per session).

Table 5 shows that men also ask more from female firms than from male firms (and get offered more from female firms than from male firms), but these effects are not significant. There are also no significant differences depending on sex of the firm in neither firm offers, the wages asked by workers nor discrepancy between wages asked and offered (see also table 5).

Constellation		N	Minimum	Maximum	Mean	Std. Deviation
mm	Offers by firms	4	37.67	57.50	48.22	8.20
	Bids by workers	4	31.58	68.63	52.96	16.86
	Discrepancy	4	-6.08	13.80	4.74	9.33
ff	Offers by firms	4	27.00	45.00	35.58	9.24
	Bids by workers	4	44.75	66.30	53.79	9.18
	Discrepancy	4	9.13	24.30	18.21	6.97
fm	Offers by firms	4	22.92	43.33	35.31	8.95
	Bids by workers	4	48.13	60.00	54.29	5.03
	Discrepancy	4	12.79	25.21	18.98	5.14
mf	Offers by firms	4	48.17	75.00	57.94	11.76
	Bids by workers	4	62.89	72.86	68.20	4.71
	Discrepancy	4	-2.14	18.07	10.27	8.85
mix	Offers by firms	3	41.25	68.75	51.56	14.99
	Bids by workers	3	49.00	66.13	59.13	8.98
	Discrepancy	3	-6.50	21.46	7.57	13.98

Table 5: Descriptive statistics for offers by firms, bids by workers and discrepancy between them for all constellations separately (period one, per session).

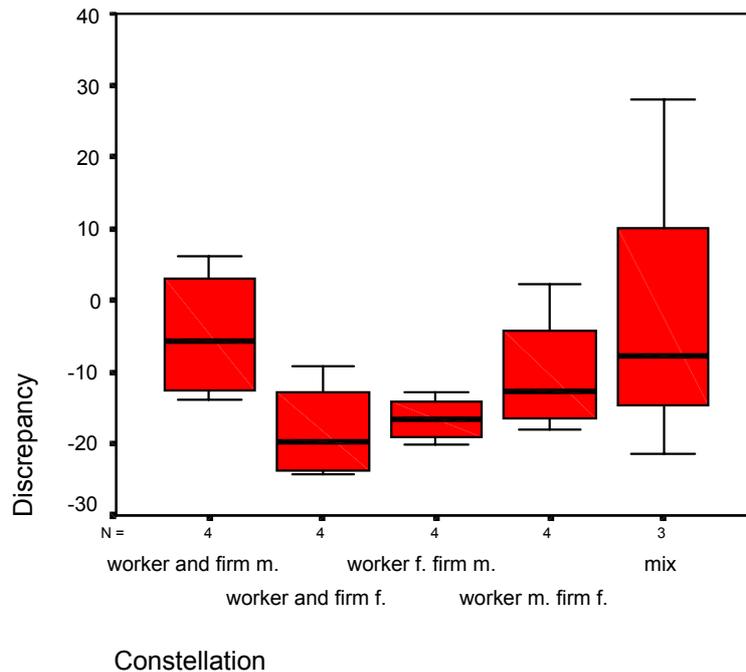


Figure 6: Box plots of discrepancy between wage asked for and offered per constellation.

Over all constellations, effort and discrepancy are negatively correlated ($r = -.60$, $p = .003$ (1-tailed; for period one only))⁴, i.e., the higher the discrepancy, the lower the effort (excluding the mixed category, it remains $r = -.58$, $p = .01$)⁵. A regression analysis over all constellations for period one, using wage and discrepancy as explanatory variables for effort, confirms the importance of discrepancy between wage asked by the worker and offered by the firm for the effort level chosen. The complete linear regression model becomes significant ($R^2 = .51$, adjusted $R^2 = .44$, sig. = .004), and both explanatory variables become significant as well, but ‘discrepancy’ is the more important one, as the beta-weight shows. Excluding the mixed sessions, the complete model remains significant (adj. $R^2 = .41$, sig. = .014), but only the beta-weight for ‘discrepancy’ remains significant. Discrepancy seems to be even more important than wage level in the determination of effort level.

		Std. Error	Beta	t	Sig.
Including mixed sessions	(Constant)	14.04		1.04	.315
	Wage	.24	.40	2.18	.045
	Discrepancy	.31	-.49	-2.68	.016
Without mixed sessions	(Constant)	16.54		.91	.379
	Wage	.28	.40	1.95	.073
	Discrepancy	.39	-.47	-2.25	.042

Table 6: Regression of wage and discrepancy on effort (period one).

⁴ Spearman’s rank correlation rho = -.50, p = .014.

⁵ Spearman’s rank correlation rho = -.53, p = .017.

The mixed sessions

In general, tendencies in the mixed sessions taken together are somewhere in-between the homogeneous sessions – as one should expect. However, the mixed sessions varied in their composition, as table 7 shows - and this did not remain without an effect.

Session	Male Firms	Female Firms	Male Workers	Female Workers
17	2	2	1	5
18	3	1	5	1
19	1	3	2	4

Table 7: Sex-composition of the mixed sessions.

One could assume that differences between these sessions were similar to the differences between the homogeneous sessions. More specifically, session 18 could be expected to be similar to the male-male sessions, while session 17 and 19 should be more similar to sessions with only female workers. It is of course not possible to compare the three sessions statistically, as the number of independent observations is too small. Still, an exploratory analysis, looking at descriptive statistics, was done to get an impression whether tendencies are as expected (see also table 8). The assumptions are partly confirmed – session 18 is more similar to an all-male session, while the other two sessions tend more towards male-female or all female sessions. This holds especially with respect to firm offers. However, session 17 seems to be strange in that the average wage is quite high, but still effort is low. Looking at the offers firms made, it becomes clear that in session 17 the discrepancy between wages offered by firms and wages asked by workers was high in the beginning – even if wages contracted in the end were not that low.

Session		N	Minimum	Maximum	Mean	Std. Deviation
17	Offers by firms	3	25	70	44.67	23.03
	Offers by workers	8	39	100	66.13	22.81
	Wage	4	39	80	62.25	17.52
	Effort	4	10	50	20.00	20.00
18	Offers by firms	4	60	75	68.75	6.29
	Offers by workers	4	55	69	62.25	6.08
	Wage	4	60	70	67.25	4.86
19	Offers by firms	4	10	100	50.00	42.43
	Offers by firms	4	25	70	41.25	20.16
	Offers by workers	6	29	110	49.00	30.85
	Wage	4	30	50	38.75	8.54
	Effort	4	10	30	20.00	11.55

Table 8: Descriptive statistics for offers by firms, offers by workers, wage , and effort for the mixed sessions separately (period one).

Discrepancy is lowest for session 18 (-6.5) – here workers even asked less than the firms offered them – and as a consequence, effort is highest. Discrepancy is highest in session

17 (21.46 compared to 7.5 in session 19), but there is no difference in average effort between session 17 and session 19 – however, wages are very low in session 19 from the beginning on. Thus, the comparison of the three mixed groups shows the same tendencies as the analyses of the homogeneous data. Even if no statistical analysis is possible, this can carefully be taken as a confirmation of the results from the homogeneous sessions. It also shows that it might not even be necessary to make gender salient to get such effects.

5 Discussion

To summarize, it was possible to show in a double auction in the laboratory that women earn lower wages than men. Secondly, the findings show that women's behavior might have a partial influence on their wages. Women, however, are not too modest, but might be a bit less inclined to reciprocate higher wages with higher effort than men. However, this result has to be interpreted carefully, because firms determined wages, and firms did not make offers to women in the same wage range as to men. Furthermore, the experiment also points to a new explanatory factor for differences in effort next to the absolute wage: The difference between the wage asked by the worker in the beginning and the wage offered by the firm.

Men and women pay lower wages to women in the experiment, and they do this independently from what men and women ask for wages in the role of the worker. Male and female firms basically behave the same. However, from the second period on women indeed ask for lower wages. One experience of high discrepancy between the wage asked for and the wage offered by the firm seems to be sufficient for women to adapt their offers downwards. But, both men and women reciprocate, i.e., men and women react to low wages with low effort levels and to higher wages with higher effort levels – which made firms lose money in the experiment. Women did not get the chance to reciprocate higher wages with higher effort.

Only in the male-male-constellations firms do not make losses. In all other constellations on first sight the pattern looks like a striking lack of fairness concerns: Workers go up with their effort levels when they get higher wages, but most of the time they stay low enough to let firms make losses. Taking a closer look at the differences between the four conditions, one possible explanatory variable is recognized: Only in conditions with male workers discrepancy between the average wages firms offered and workers asked for was low, i.e., not significantly different from zero. This indicates that in all other conditions, where workers got offered lower wages than they thought they should get, this was basically seen as an unfair action in the beginning. This evoked a reaction that hurt the firms. Even in the condition with female firms and male workers, discrepancy is very close to significantly

different from zero, because male workers ask very high wages from female firms. This could explain why even here firms made losses. Firms, which made losses, rationally responded with stopping to trade – even if adjustment took some time, as figure 5 showed. This confirms the economic intuition that firms do not make losses if they are able to avoid it.

One could argue that the wages asked for and offered in part are strategically determined, but this does not counteract the argument that these asks and offers were lower when workers were female than when they were male. However, there is a problem of path-dependency, as even if one looks only at period one, the improvement rule implies that the first offer made determines all other offers. Again, this does not change the finding that these paths led to lower average offers made to women than to men.

What does this tell us for real-life situations and for further research? First, it shows that there is some kind of wage discrimination against women – either individual discrimination or group discrimination. Unfortunately, this experiment could not test whether this is based on wrong stereotypical perceptions of women or on correct estimates of their behavior, as women were not equally represented in all wage categories. If it were based on correct estimates of female behavior, it would of course be rational behavior by firms and thus ‘only’ constitute individual (statistical) discrimination. If it were based on wrong perceptions, it would constitute group discrimination as well.

Secondly, if employers think that they can make higher profits by paying women lower wages, they are wrong – at least in the experiment it did not work out for them: It led to lower effort by women and very low or even negative profits by firms. This seems to be a vicious cycle: Women are paid lower wages, as a consequence they exhibit lower effort, employers see their stereotypes confirmed – and, as happened in the experiment, they rationally stop to employ women at all. This is also evidence against the mechanism Becker (1957) proposes for reducing discrimination, because it is assumed to work via lower wages for women in initial periods – not considering that this might lead to ‘negative’ reactions by the women.

Thirdly, the experiment shows that it is not mainly the absolute wage level, which determines effort levels, but the discrepancy between starting offers of firms and workers. It is not the case in this experiment that women ask for lower wages, they just get them – and therefore, they are in a situation of high discrepancy between their own idea, what wage they would like to get, and the wages firms want to pay them. Men, on the other hand, ask women to pay them more than they ask from other men, also leading to a discrepancy between wage offers. Thus, seemingly expectations – whether right or wrong - both sexes have about each

other lead to discrepancies in wage offers, and to bad outcomes for firms and workers. Women in this experiment quickly adapt their expectations – but they also adapt their behavior to the low wages. It does not seem very unrealistic that women with a good education in the ‘real world’ are confronted with a discrepancy between their own wage expectations and what firms are willing to pay them. Meng (2002) found that female economists in their first job (in male occupations) earn lower wages than their male counterparts with the same education. This does not change with time; instead, the gap rather grows. Although Meng’s study does not provide data about expectations of men and women for their first wage, a study by Blau and Ferber (1991) indicates that expectations women and men have about their starting salaries are quite similar – at least in the US.

Further research should test how women do react to high wages, i.e., whether it is a wrong or a correct expectation that it does not make sense to pay them higher wages, because they don’t react with higher effort (see, e.g., Major, McFarlin, & Gagnon (1984) for evidence in the latter direction). Therefore, experiments should be done were women are confronted with high wages as well, for example using pre-programmed wage offers by firms, or the strategy method (a method which asks about reactions to all possible behaviors of the other party).

It would also be interesting to see what happens when employers can choose between women and men as workers in one auction, and whether they would stop trading with women and continue only with men if possible. In the mixed sessions reported here it was not possible to make an offer explicitly to female or male workers only. Unfortunately, it is difficult to do such an experiment, because if subjects are explicitly asked to choose between male and female workers, one runs the danger of getting only socially desirable behavior.

There might be another explanation for the results, different from an efficiency-wage effect: It could also be that emotional reactions, i.e., certain ‘gut feelings’ influence the wages offered to women - which is also proposed by the integrated model described in chapter 3. This, together with the question what expectations employers have, could be analyzed in future research. Again, this is not that easy, because asking about expectations first might lead to socially desirable behavior, as people become aware of the topic at hand. However, if in a context of an experimental market subjects are asked about their expectations with respect to the effort-choice, and gender is introduced as has been done here, such an effect can probably be avoided. Asking about ‘gut feelings’ might be more difficult, but there are measures like the “Implicit Association Test” (Greenwald et al. 1998), which have proven quite useful in measuring, e.g., prejudice against blacks.

To conclude, the experiment gives an indication why discriminatory outcomes for women on the labor market might continue to exist even – or *especially* if it is possible to pay them lower wages than men, and how the perception that women are less productive than men can be confirmed by women’s reactions to firm’s behavior

Acknowledgements. I would like to thank Armin Falk for allowing me to use his program, David Rodriguez for his great help in preparing and executing the experiment and Jordi Brandts and Marc Vorsatz for translating the adapted, originally German, instructions into Spanish. Further, I would like to thank Matthias Sutter, Martin Strobel, Maarten Vendrik, Peter de Gijssel, participants at the ESA-meeting 2002 in Strasburg, participants at a MiMaMa Seminar at Universiteit Maastricht and participants of a seminar at Universitat Pompeu Fabra, Barcelona, February 2003, for many helpful comments.

6 References

- Anderson, D. & Shapiro, D (1996), “Racial Differences in Access to High-Paying Jobs and the Wage Gap Between Black and White Women,” *Industrial and Labor Relations Review*, 49, 273-86.
- Arrow, K. J. (1973), “The Theory of Discrimination,” In O. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets*, Princeton: Princeton University Press, 3-33.
- Ashraf, J. (1996), “Is Gender Pay Discrimination on the Wane? Evidence from Panel Data, 1968-1989,” *Industrial and Labor Relations Review*, 49, 537–546.
- Becker, G. (1957, 1971), *The Economics of Discrimination*, 1st and 2nd. eds. Chicago: The University of Chicago Press.
- Ball, S., Eckel, C., Grossman, P. J. & Zame, W. (2001), “Status in Markets,” *The Quarterly Journal of Economics*, 161 – 188.
- Blau, F. D. & Ferber, M. A. (1991), "Career Plans and Expectations of Young Women and Men: The Earnings Gap and Labor Force Participation," *Journal of Human Resources*, 26, 581-607.
- Blau, F. D. & Kahn, L. M. (1994), “Rising Wage Inequality and the U.S. Gender Gap,” *American Economic Review*, 84, 23–28.
- Blau, F. D. & Kahn, L. M. (1997), “Swimming Upstream: Trends in the Gender Wage Differential in the 1980,” *Journal of Labor Economics*, 15, 1-42.
- Blau, F. D. & Kahn, L. M. (2000), *Gender differences in pay*, NBER working paper 7732.
- Blau, F. D., & Kahn, L. M. (1981), “Race and Sex Differences in Quits by Young Workers,” *Industrial and Labor Relations Review*, 34, 563-77.
- Carmichael, H. L. (1990), “Efficiency Wage Models of Unemployment – One View,” *Economic Inquiry*, 28, 269–295

- Charness, G. (1998), "Attribution and Reciprocity in a Simulated Labor Market: An Experimental Investigation," *Mimeo, University of California, Santa Barbara*.
- Datta Gupta, N., Oaxaca, R. L. & Smith, N. (1999), "Changes in the Gender Wage Gap: A Comparison Between the U.S. and Denmark," *presentation at the 11th EALE Annual Conference*.
- Davis, D. D. & Holt, C. A. (1993), *Experimental Economics*, Princeton, New Jersey: Princeton University Press.
- Deaux K. & LaFrance, M. (1998), "Gender," In Fiske, S. T., Gilbert, D. T., and Lindsay, G. (eds.), *The Handbook of Social Psychology*, 2, New York: McGraw Hill, 788–827.
- Fehr, E. & Falk, A. (1999), "Wage Rigidity in a Competitive Incomplete Contract Market," *Journal of Political Economy*, 107, 106–134.
- Fehr, E., Kirchsteiger, G., & Riedl, A. (1998), "Gift Exchange and Reciprocity in Competitive Experimental Markets," *European Economic Review*, 42, 1-34
- Fischbacher, U. (1999), "z-Tree: Zurich Toolbox for Readymade Economic Experiments. Instructions for Experimenters," *Mimeo, University of Zurich*.
- Fiske, S. T. (1998), "Stereotyping, Prejudice, and Discrimination," In Fiske, S. T., Gilbert, D. T., and Lindsay, G. (eds.), *The Handbook of Social Psychology*, 2, New York: McGraw Hill, 357–414.
- Fortin, N. M. & Lemieux, T. (1998), "Rank Regressions, Wage Distributions, and the Gender Gap," *The Journal of Human Resources*, XXXIII, 610–643.
- Glick, P., Wilk, K. & Perrault, M. (1995), "Images of occupations: Components of gender and status in occupational stereotypes," *Sex roles*, 32, 565-582.
- Haagsma, R. (1993), "Is Statistical Discrimination Socially Efficient?" *Information Economics and Policy*, 5, 31-50.
- Haagsma, R. (1995), "Microeconomic models of gender discrimination in labour markets with asymmetric information," *Unpublished Ph.D-thesis*, University of Amsterdam.
- Hannan, R. L., Kagel, J. H., & Moser, D. V. (2002), "Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences, and Effort Requests on Behavior," *Journal of Labor Economics* 20, 923-951.
- Katz, L. F. (1986), "Efficiency Wage Theories: A Partial Evaluation," In Fischer, S. (ed.) *NBER Macroeconomics Annual*, Cambridge, Mass.: MIT Press, 235–276.
- Light, A. & Ureta, M. (1992), "Panel Estimates of Male and Female Job Turnover Behavior: Can Female Nonquitters Be Identified?" *Journal of Labor Economics*, 10, 156-181.

- Major, B., McFarlin, D. B., & Gagnon, D. (1984), "Overworked and Under Paid: On the nature of gender differences in personal entitlement," *Journal of Personality and Social Psychology*, 47, 1399–1412.
- Mavromaras, K. G. & Rudolph, H. (1999), "Occupational Segregation and the Male Female Wage Gap in Germany Between 1985 and 1995," *presentation at the 11th EALE Annual Conference*.
- Meng, Christoph (2002), "(fe)male jobs and (fe)male wages – disentangling the effect of personal and job characteristics on wages by measuring stereotypes," *Paper submitted for publication*.
- Pelham, B. W. & Hetts, J. J. (2001), "Underworked and Overpaid: Elevated Entitlement in Men's Self-Pay," *Journal of Experimental Social Psychology*, 37, 93–103.
- Phelps, E. S. (1972). "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62, 659-661
- Sicherman, N. (1996), "Gender Differences in Departures from a Large Firm," *Industrial and Labor Relations Review*, 49, 484-505.
- Stanley, T. D. & Jarell, S. B. (1998), "Gender Wage Discrimination Bias? A Meta-Regression Analysis," *The Journal of Human Resources*, XXXIII, 947–73.
- Viscusi, W. K. (1980), "Sex Differences in Worker Quitting," *Review of Economics and Statistics*, 62, 388-398.
- Walters, A. E., Stuhlmacher, A. F., & Meyer, L. L. (1998), "Gender and Negotiator Competitiveness: A Meta-analysis," *Organizational Behavior and Human Decision Processes*, 76, 1–29.