Erasmus School of Economics

THE INFLUENCE OF COMPANY SIZE ON GENDER SALARY INEQUALITY

Analyzing the impact of company size on gender salary inequality by doing an Ordinary Least Squares (OLS) regression and a Blinder-Oaxaca decomposition.

ABSTRACT

In this paper the existence of a relationship between gender salary inequality and company size is investigated. A positive correlation is expected based on the assumptions of the employer size-wage effect and the increasing wage disparity between men and women throughout the wage distribution. However, no positive relationship between the two phenomena is found.

Eri Kooistra
413710
Thesis B.Sc. Economics & Business Economics
Erasmus School of Economics

Supervisor: prof. dr. (Olivier) O.R. Marie

Date of completion: 12/08/2017

Table of Contents

1. Introduction 2
2. Theoretical framework 5
3. Data 9
4. Methodology 15
5. Results 18
6. Conclusion 25
7. Discussion 26
8. Recommendations 28
9. References 29

1. Introduction

Judging from the latest studies, the unadjusted pay gap between men and women in the European Union has stayed the same over the last couple of years. Females in The Netherlands earn 16.1% less than males on average, whereas the difference is even higher between male and female managers, namely 21.4% (Vrooman, 2017).

In regard to the abovementioned pay gap there has to be made a clear distinction. The discrepancy can be found in the definition of 'pay gap'. There are two applications of the phenomenon. The first one is called the unadjusted pay gap and signifies only the difference in average gross hourly income of male and female employees as percentage of the male gross income (Gender pay gap in unadjusted form, n.d.). The second one is the adjusted pay gap and controls for a range of productivity measures and personal characteristics. The purpose of the adjusted pay gap is to find out if discrimination in the workplace is responsible for the difference in income of men and women. Ideally, the adjusted pay gap would be zero, because wage of men and women should be the same for equal jobs (Grimshaw & Rubery, 2002).

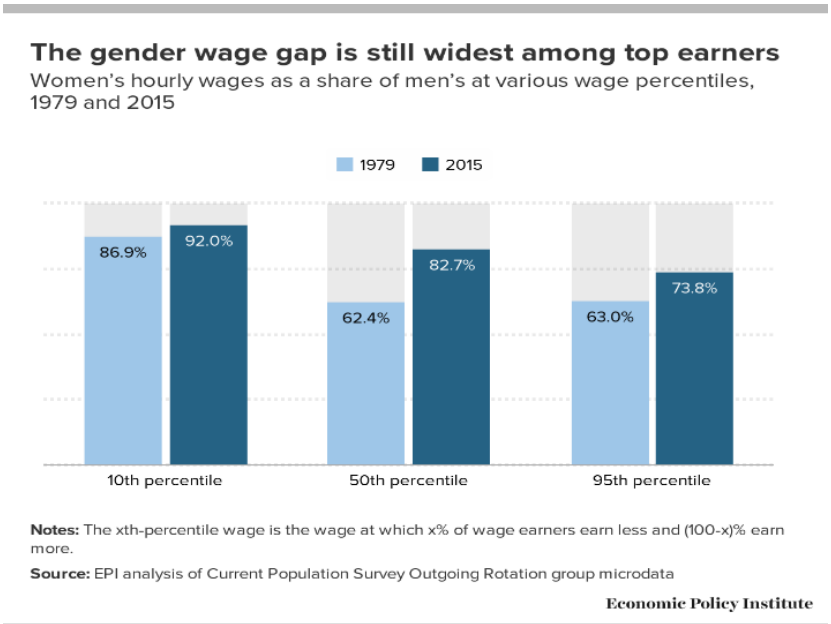
The existence of the gender wage differential, which cannot be attributed to human capital indicators, has led to countless empirical studies about labour market discrimination. One of the latest studies coming from The Netherlands concludes that there are payment differences between men and women practicing science which cannot be contributed to age, job classification or job level (de Goede, van Veelen, & Derks, 2016). Apparently, despite the law of wage equality between men and women for the same working activities, enacted in 1975, there is still an inexplicable pay gap between the two genders.

In addition to the known human capital factors that influence someone's salary like education, on the job training and years of experience, a possible factor needs to be examined that influences salary and can simultaneously have a discriminating effect in order to explain a part of the adjusted gender pay gap.

One of the possible discriminating and wage increasing factors is company size. At least, this seems to fulfil the requirement of positively affecting the salary of an employee. In 1989 Charles Brown

and James Medoff came up with evidence for the existing positive relationship between the size of an employer and its remitted salary to employees and called it the employer size-wage effect (Brown & Medoff, 1989). They used older theories as well as, at that time, nowadays findings to support their presumptions. Medoff and Brown give six plausible explanations causing the effect, namely (1) big corporations bring in higher-quality employees, (2) big corporations provide worse labor circumstances, (3) big corporations use it to forestall the employees to unionize, (4) big corporations are more able to pay higher salaries, (5) big corporations deal with tinier pools of candidates relative to vacancies and (6) big corporations find it more difficult to monitor their employees. These explanations will be thoroughly clarified in the theoretical framework.

Currently, the unadjusted gender pay gap is not evenly distributed. The gap between male and female hourly wages in 2015 was smallest for the worst earners, with 10th percentile females making 92 per cent of male's payments, 50th percentile females making 82.7 per cent of male's payments and the gap is at its greatest at the ceiling of the wage distribution, with hourly wages of females being 73.8 per cent of the hourly wages of men at the 95th percentile. The graph below displays the different wage gaps for the aforementioned percentiles[1]:



[1] CPS is a monthly household survey prepared by the U.S. Census Bureau for the Bureau of Labor Statistics (BLS). This subsample gives the hourly wages of American respondents who are aged 18-64, employed in the public or private sector and have hours worked within the range 1-99 per week.

A partial explanation for the gap being smallest among the lowest earners is the existence of minimum wage. This sets a wage floor for everyone which results in more equal wages at the bottom of the wage distribution (Gould, Schieder, & Geier, 2016). Interesting to broach is that the wage gap has decreased throughout the whole distribution over the last 40 years. This could be due to improvements of human capital factors like education and experience which have improved a lot with regard to women the last couple of decades (Blau & Kahn, 2016).

A combination of these stated two phenomena, which present the basic concepts of this thesis, would logically lead to the inference saying that larger companies yield more people at the top half of the wage distribution and therefore a bigger gender pay gap should be present there than at smaller companies. To investigate if the proposition is indeed true in The Netherlands, an extended Mincer earnings function in the form of an Ordinary Least Squares (OLS) regression and a Blinder-Oaxaca decomposition will be used. The assumption is made that the adjusted pay gap takes on the same form as the unadjusted pay gap, since the adjusted pay gap will be examined with the OLS. The following research question will be examined:

What is the influence of company size on gender salary inequality in The Netherlands?

This thesis consists of a theoretical framework discussing the employer size-wage effect, the gender wage gap and its corresponding distribution. Furthermore, the used data and its corresponding specifications, the applied methodology and the results will be explained. At the end of the thesis, a conclusion, a discussion of the results, limitations of this paper and recommendations for further research will be given.

2. Theoretical framework

Employer size-wage effect

Recurring to the introduction in which a short summary of the important ideas of Medoff & Brown was given, a more enhanced discussion of the essential factors leading large companies to pay higher wages will be given.

The first reason dates from the neoclassical era and states that large corporations bring in employees with a high productivity because big corporations are more capital (read 'machinery') intensive (Hamermesh, 1980). The substitution between capital and unskilled labor is more elastic than between capital and skilled labor, assuming that an increase in capital results in a higher marginal productivity of a skilled worker but a lower marginal productivity of an unskilled worker, which is called capital-skill complementation (Krusell, Ohanian, Ríos-Rull, & Volante, 2000). Another motive to hire high-skilled employees is because a corporation can diminish its monitoring costs (Oi, 1983).

Secondly, the unsatisfactory working environment at large corporations ensures that employees have to be paid more salary to keep them satisfied and motivates them to perform at the same quality level of labor they would otherwise perform at when they would be working in a pleasant working environment. Examples of discords are more bureaucracy and less responsibilities, own contributions and planning (Masters, 1969; Stafford, 1980), longer commuting (Scherer, 1976) and cold working ambience (Lester, 1967).

There is also an institutional point of view that reasons big nonunion corporations behave like they were unionized to prevent unionization. This behavior consists of paying higher salary, arranging more advantages and offering improved working circumstances. Union salary and profits differ therefore contrarily with size (Curtin, 1970; Foulkes, 1980; Freeman & Medoff, 1984).

A different kind of explanation can be described as the willingness of a firm to share some of the excessive profits with its employees by overpaying them simply because they can due to monopoly power (Weiss, 1966; Mellow, 1982). It is uncertain if the market power relates to large companies itself or to the sector the companies are participated in.

An argument about the recruitment of companies argues that big corporations deal with tinier pools of candidates relative to their vacancies (Weiss & Landau, 1984). Weiss and Landau assume that the increase in labor supply does not rise in the same proportion as the demand for labor does, resulting in a decrease of candidates as proportion of the amount of vacancies leading to payment of higher wages in order to fulfill the increased demand of labor by attracting more employees.

The last reason somehow follows the explanation of Oi about monitoring but it is not explained from a costs cutting perspective. Big corporations find it more difficult to monitor their employees and therefore monitor their employees less than small corporations, resulting in a higher salary on average from a lack of information about the quality of work, such as effort (Stigler, 1962). A related vision states that large corporations' monitoring difficulties contribute to various offered salaries and this is something employees take into consideration when choosing an employer (Garen, 1985).

The article of Brown and Medoff summarizes a lot of earlier studies and complements it with empirical research by testing the previous opinions. Its conclusion is somewhat unsatisfying because there is clear affirmative evidence about the assumption that larger companies pay higher salaries but the findings are not well grounded. They are unable to answer the 'why' of the question.

Gender pay gap and its distribution

One of the biggest hurdles for male and female equality is the still remaining gender pay gap. Although the unadjusted gap started shrinking form the 1980s because of improved human capital characteristics of women, especially working experience. Nowadays, the remaining gap is mostly unexplainable (Blau & Kahn, 2007). The unexplained wage gap can be partially attributed to labor market discrimination. Over time the unexplained gender wage gap has declined. This could be due to diminishing discrimination or an improvement of unmeasured skills of females.

In general, women seem to have worse gender-specific qualifications which implies that, on average, men are better skilled than women and have usually higher paid jobs than women or are working in a higher paid sector. This brings to mind that the structure of wage is a crucial determinant of the gender pay gap. Wage structure is the diversion of wage for different kind of labor-market skills and the rents that are received for employment in specific subdivisions of the economy (Blau & Kahn, 1994).

Blau & Kahn made an interesting and at the same time shocking discovery about the gender pay gap in the USA in comparison to other industrialized countries. The problem was relatively paradoxical. Women in the United States show better competences and hold higher-status occupations on average compared to other countries and the United States has a more active regime to obtain equal pay, but the returns to skill are much higher wherefore it is very unfavorable to be occupied in the lower half of the wage distribution (Blau & Kahn, The Gender Earnings Gap: Learning from International Comparisons, 1992).

Let's take a look at the situation in The Netherlands. Dutch law prohibits unequal wage for equal work. This is embedded in our country in the Wet Gelijke Behandeling van mannen en vrouwen (WGB). A long history of striving towards wage equality between men and women is known. In 1951, Convention no. 100 concerning equal payment was accepted by the International Labour Organisation (ILO). In the Netherlands this would only be endorsed in 1971. Following this convention, the Wet Gelijk Loon voor vrouwen en mannen (WGL) came into force that turned into the Wet Gelijke Behandeling (WGB) in 1980 as we know it right know. Complaints regarding labor

market discrimination can be submitted to the Commissie Gelijke Behandeling (CGB) (Tijdens & Klaveren, 2011).

Research originating from Sweden in 2003 presents an increasing gender log wage gap throughout the wage distribution and acceleration at the top of the distribution due to a glass ceiling (Albrecht, Björklund , & Vroman, 2003). A glass ceiling can be understood as an intangible barrier within a hierarchy that prevents women or minorities from obtaining upper-level positions (glass ceiling, n.d.). If we assume that only better educated and more articulate women would do something about the detrimental treatment, the anti-discrimination policy is assumed to tackle the existence of a glass ceiling. The comparatively high salaries at the bottom of the wage distribution in Sweden entails that women choose less demanding jobs, because it makes it harder for ambitious women to pay for childcare or housekeeping.

There is also a constant increase of the gender pay gap noticeable throughout the wage distribution in Finland and The Netherlands. In Germany, Denmark and Belgium a wider gap at the top of the wage distribution can be noticed. In Spain and Ireland an opposite course of the pay gap can be seen and in Britain the pay gap remains constant throughout the wage distribution at around 20% (Arulampalam, Booth, & Bryan, 2007).

The gender pay gap in The Netherlands is above the EU average. The unadjusted pay gap for full-time work is higher than for part-time work in The Netherlands and the unadjusted pay gap increases with age. This can be attributed to childbirth and that older women could not gain from equality policies which were set-up during their off time (Gowling, 2014). The Dutch government will start closing the gender wage gap by increasing the labor participation of women by encouraging women to start working and by pointing out the importance of being financial independent.

Research dating from 2006 concludes that one third of the gender pay gap in The Netherlands comes due to the overpayment of male dominated sectors and one sixth by the underrepresentation of females at high level jobs. The other half is a mix of unobserved characteristics of men and women and the presence of discrimination at the workplace (Fransen, Plantenga, & Vlasblom, 2010).

3. Data

Used data

The used data originates from the LISS panel data archive. LISS panel makes use of just about 5000 households, what is equal to around 8000 individuals, which are located all over The Netherlands. The individuals are randomly drawn from the population. Variables for the years 2012 and 2013 from three different studies are combined into one data set.

The first survey from which variables are taken is called *Background Variables*. This is single wave data and consists of the most important general characteristics of the households. The selected variables are *gender, age, amount of children* and *level of education*[2]. *Gender* is a crucial variable in order to determine the salary difference between males and females. *Age* is used instead of work experience. Work experience is expected to have a positive effect on wage growth (Lazear, 1974). Employees with more work experience need less on the job training. The *amount of children* will most likely have a negative effect on the earned salary, because a lot of women who have (young) children start working part-time and stop working full-time (Jackson, 1993). *Level of education* seems to have a positive effect on the salary of employees due to the ability of performing more difficult jobs conditional on the assumption that difficult jobs often pay out high earnings (Pereira & Martins, 2004). The educational levels consist of: (1) primary school, (2) VMBO (intermediate secondary education), (3) HAVO/VWO (higher secondary education/preparatory university education), (4) MBO (intermediate vocational education, (5) HBO (higher vocational education) and (6) WO (university).

The second survey is named *Economic Situation: Income* and contains salary of the individuals earned over the entire year, including holiday allowance, 13th month and profit sharing schemes, at their main employer. This is longitudinal data. A dataset is longitudinal if it contains data about the same subjects at multiple points in time (What are Longitudinal Data, n.d.). Not everyone was transparent about his/her income. Some people preferred not to share their wages and other people did not know their exact wages.

[2] Level of education in CBS (Statistics Netherlands) categories.

The last used survey is *Work and Schooling* and is also longitudinal data. This survey consists of variables that are related to specific job information and educational level of the individuals. The selected variables are *hours worked, profession* and *company size. Hours worked* is the average amount of hours worked per week computed over a whole year. The explanation for the use of hours worked is that employees who work more, given the same job, earn more. *Profession* seems a relevant variable, because some professions bring higher yields on average. The different professions are: (1) higher academic or independent profession (e.g. architect, physician, scholar, academic instructor, engineer), (2) higher supervisory profession (e.g. manager, director, owner of large company, supervisory civil servant), (3) intermediate academic or independent profession (e.g. teacher, artist, nurse, social worker, policy assistant), (4) intermediate supervisory or commercial profession (e.g. head representative, department manager, shopkeeper), (5) other mental work (e.g. administrative assistant, accountant, sales assistant, family carer), (6) skilled and supervisory manual work (e.g. car mechanic, foreman, electrician), (7) semi-skilled manual work (e.g. driver, factory worker), (8) unskilled and trained manual work (e.g. cleaner, packer) and (9) agrarian profession (e.g. farm worker, independent agriculturalist). *Company size* is measured in the amount of people that worked at the branch/location where the individual mainly worked at his/her last job.

Data transformation

Ordinary Least Squares (OLS)

The three different data sets of the years 2012 and 2013 are merged into one data set. In order to perform an ordinary least squares (OLS) regression some changes had to be made to the assembled data. To figure out the effect of company size on the gender pay gap, a regression has to be made that displays the effect of the different variables on the percentage change in salary, given that the other variables remain constant. This can be obtained by changing *salary* into a log variable.

It is important to have a complete data set. This requires removal of individuals that did not fill in all the essential questions of the survey. For instance, data of individuals who did not know their salary or the amount of people they worked with at their job, but do know their profession and their educational level, gives incomplete information and those particular individuals are removed from the dataset.

To create different company sizes the variable is given in three categories. A company is small when it has less than 50 employees (1), medium when it has 50-249 employees (2) and big when it has 250 employees or more (3).

Creating a dummy can make a comparison between males and females. Also an interaction effect between being a female and company size can be measured with this dummy variable. Within the interaction effect, company size three together with being female is taken as the reference base.

Second, a dummy variable for having kids is designed. There is no distinction anymore in the amount of kids the individuals have. The assumption is made that having one child is sufficient for men and especially women to dissociate from their ambition to get everything out of their working career.

The factor variables *company size, level of education* and *profession* need a reference base to make them comparable with each other. For *company size*, small companies (1) function as the reference category. For *level of education*, primary school (1) functions as the reference category

and for *profession*, unskilled and trained manual work (e.g. cleaner, packer) (8) functions as the reference category.

Blinder-Oaxaca decomposition

Exactly the same variables as with the Ordinary Least Squares (OLS) are used with the Blinder-Oaxaca decomposition. To include the categorical variables another strategy is needed, because factor (categorical) variables and time-series operators are not allowed in oaxaca. Therefore, those variables had to be normalized. Instead of an automatic subdivision of factor variables with their respective coefficients, like with the OLS, sets of dummy variables had to be made to make sure the results of the decomposition were invariant of the choice of the base category.

Descriptive statistics

The objective of this section is to give some insights on the used data and to show if the assumptions that are presumed to be true in this thesis are indeed true. First of all, there is assumed that wages at big firms are higher than at small firms. The following table shows the average salaries of individuals working at the three different companies:

branch/1	Mean	Std. Dev.	Freq.
1	20412.827	15261.082	948
2	24478.208	13073.599	523
3	28303.626	14551.235	431
Total	23318.783	14868.217	1,902

One can see that the bigger the company, the higher the average wage. Therefore, the first assumption can be accepted.

The second assumption that has to be checked is the differences in pay gaps at the different percentiles in the wage distribution. The first table shows the average salary at the different percentiles for females:

.

	Percentiles	Smallest		
1%	600	5		
5%	1600	10		
10%	3000	147	Obs	835
25%	10300	190	Sum of Wgt.	835
50%	17000		Mean	17813.87
		Largest	Std. Dev.	13671.11
75%	23600	65468		
90%	30000	80000	Variance	1.87e+08
95%	35000	190000	Skewness	6.04735
99%	50000	210000	Kurtosis	78.89861

The second table shows the average salary at the different percentiles for men:

.

	Percentiles	Smallest		
1%	900	170		
5%	2400	260		
10%	7500	319	Obs	1,067
25%	21000	400	Sum of Wgt.	1,067
50%	27157		Mean	27626.75
		Largest	Std. Dev.	14341.8
75%	35000	85000		
90%	43000	90522	Variance	2.06e+08
95%	50000	130000	Skewness	1.036407
99%	70000	130000	Kurtosis	8.565611

To calculate if the pay gaps are indeed ascending throughout the wage distribution, the same three percentiles are used as in the introduction to compare the differences in pay gaps, namely the 10^{th}, 50^{th} and 95^{th} percentile.

10^{th} percentile: 3000 / 7500 = 0.40. Women at the 10^{th} percentile earn 60% percent less than their male counterpart.

50^{th} percentile: 17000 / 27157 = 0.63. Women at the 50^{th} percentile earn 37% percent less than their male counterpart.

95^{th} percentile: 3500 / 5000 = 0.7. Women at the 95^{th} percentile earn 30% percent less than their male counterpart.

This is an interesting founding, because the pay gap is decreasing throughout the wage distribution and not increasing as was previously expected and claimed by Arulampalam et al (2007). A reason for this finding could be that the division of company sizes is not well chosen. Changing the partition of company sizes could lead to other outcomes. Another reason could be that there are a few individuals with very low salaries, but no individuals with salaries which are sky-high. One of the two assumptions is not met and this could lead to opposite outcomes. One should note that there could have been made different assumptions beforehand, thus this result does not exclude a positive relation between company size and gender pay inequality yet. That is to be found out by looking at the results in the next section.

4. Methodology

Mincer earnings function

One of the two methods that is used in this paper to estimate if there is any connection between the inequality in salary between men and women and the size of a company, is the Ordinary Least Squares (OLS) regression. The regression looks like a function that found its origin in 1974. In that year, Jacob Mincer devised the human capital earnings function. This function conceptualized schooling and experience as predictors of income (Mincer, 1974). The basic equation looks like this:

(1) $\ln E_t = \ln E_0 + rs + \beta_1 t - \beta_2 t^2$

Where E_0 stands for the earnings of someone without any schooling or education, s is years of schooling and t is years of experience. The parameters r, β_1 and β_2 are the returns to schooling and experience. This empirical breakthrough is often used to calculate the returns on schooling and the impact of schooling and work experience on the wage gap between men and women.

A meaningful working paper issued in 2003 rewrites this equation to a form that is close to the equation that will be used in this thesis to investigate the gender pay gap, but with fewer variables (Heckman, Lochner, & Todd, 2003):

(2) $\ln[w(s, x)] = \alpha_0 + \rho_s s + \beta_0 x + \beta_1 x^2 + \varepsilon$

With $w(s, x)$ wage at schooling level s and working experience x, α_0 is the wage without any years of schooling and years of work experience, ρ_s is the rate of return on schooling (the same for al schooling levels) and ε is a zero mean residual with $E(\varepsilon | s, x) = 0$. In this thesis is worked with a logarithm (log) of wage and not the natural logarithm of wage (ln). This has got no consequences for the eventual outcome. The quadratic part of the function stands for the years of potential work experience. Heckman et al. (2003) gave some criticism on the old Mincer earnings function. Tuition costs and income taxes should be taken into consideration for higher levels of schooling, the effect of schooling on earnings is nonlinear, experience and schooling cannot be separated and uncertainty has to be taken into account.

The regression wherewith the effect of company size on gender pay inequality is computed ultimately looks like this:

(3) $Log\ (salary) = \alpha + \beta_1 * year + \beta_{2,3,4,5,6,7} * education + \beta_{8,9,10,11,12,13,14,15,16} * profession + \beta_{17} * age + \beta_{18} * age^2 + \beta_{19} * (hours\ worked) + \beta_{20} * kids + \beta_{21,22,23} * (company\ size) + \beta_{24} * female + \beta_{25,26,27} * female * (company\ size) + \beta_{28} * female * kids$

An interaction effect between having kids and being female has been added to the regression in order to see if there is a different effect on one's salary for having kids when being a male or being a female. Also age-squared is added to the equation, since the relationship between age and salary can happen to be non-linear.

Oaxaca decomposition

The second method that is going to be used is the Oaxaca decomposition. Around the same time as Jacob Mincer, Ronald Oaxaca came up with a quantitative solution to examine the differences in wages between males and females. Oaxaca concluded that part of the wage differential could be attributed to discrimination. The part that is said to be discrimination was calculated as the residual after controlling for differences in several characteristics (Oaxaca, 1973).

The method points out the difference in means of a dependent variable between two groups. The difference is divided in a part that is due to inequality between groups in the size of the explanatory variables of the outcome (characteristics effect) and a part that is due to the inequality between groups in the effects of the explanatory variables (coefficient effect). The coefficient effect can raise questions about labor market discrimination (Wagstaff, O'Donnell, Doorslaer, & Lindelow, 2008). Consider two wage equations for individuals in groups A and B:

(1) $\ln(wages_{Ai}) = X_{Ai}\beta_A + \mu_{Ai}$

(2) $\ln(wages_{Bi}) = X_{Bi}\beta_B + \mu_{Bi}$

X are vectors of explanatory variables like schooling, work experience, sector, kind of job, β_A and β_B are a vector of coefficients and μ_{Ai} and μ_{Bi} are error terms. As regression estimates of β_A and β_B are taken b_A and b_B respectively. On average, error terms are zero which results in:

(3) $\text{mean}(\ln(wages_A)) - \text{mean}(\ln(wages_B))$

$= b_A * \text{mean}(X_A) - b_B * \text{mean}(X_B)$

$= b_A * (\text{mean}(X_A) - \text{mean}(X_B)) + \text{mean}(X_B) * (b_A - b_B)$

$b_A * (\text{mean}(X_A) - \text{mean}(X_B))$ is the part that explains the differences between groups in the independent variable (characteristics effect). $\text{Mean}(X_B) * (b_A - b_B)$ is the part that is not explained by those differences and can be linked to the presence of labor market discrimination.

Oaxaca (1973) pointed out some difficulties that arise with his method of residuals. He asks himself the question if it is possible that the wage structure for men and women vary without the attendance of discrimination. Another obstacle is that it does not pay attention to the possibility that the residual could express the adaptation of women to the prejudices of the labor market.

5. Results

Ordinary Least Squares (OLS) regression

.

log_salary	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
year						
2013	.0121267	.0332031	0.37	0.715	-.0529922	.0772455
opleiding						
vmbo (in..)	.3236757	.0956085	3.39	0.001	.1361655	.5111859
havo/vwo ..	.396199	.0971019	4.08	0.000	.20576	.586638
mbo (int..)	.3203141	.09221	3.47	0.001	.1394691	.501159
hbo (hig..)	.5386218	.0927997	5.80	0.000	.3566203	.7206232
wo (univ..)	.6435026	.1030467	6.24	0.000	.4414045	.8456007
profession						
higher ac..	.8237869	.1072201	7.68	0.000	.6135038	1.03407
higher su..	.9882154	.1068413	9.25	0.000	.7786752	1.197756
intermedi..	.8012633	.0918378	8.72	0.000	.6211484	.9813782
intermedi..	.90302	.0948897	9.52	0.000	.7169195	1.08912
other men..	.6966718	.0883364	7.89	0.000	.5234239	.8699196
skilled ..)	.7741907	.1072818	7.22	0.000	.5637865	.9845949
semi-ski..)	.5455762	.1083124	5.04	0.000	.3331508	.7580017
agrarian..)	-.0780979	.2173316	-0.36	0.719	-.5043348	.348139
leeftijd	.1224716	.0095583	12.81	0.000	.1037255	.1412177
age2	-.0012601	.0001104	-11.42	0.000	-.0014766	-.0010437
hours_worked	.0148679	.001321	11.25	0.000	.0122771	.0174587
kids	-.0752585	.0518159	-1.45	0.147	-.1768813	.0263643
c_group						
2	.0289493	.0519681	0.56	0.578	-.0729721	.1308707
3	.1514959	.0572149	2.65	0.008	.0392844	.2637075
female	-.1957408	.0800744	-2.44	0.015	-.352785	-.0386965
female#kids						
0 1	.0703269	.0674678	1.04	0.297	-.0619928	.2026467
1 0	0	(omitted)				
1 1	0	(omitted)				
female#						
c_group						
1 1	-.0967854	.0855965	-1.13	0.258	-.2646597	.0710889
1 2	.0249979	.0969702	0.26	0.797	-.1651828	.2151786
1 3	0	(omitted)				
_cons	5.442164	.1810069	30.07	0.000	5.087168	5.79716

.

First, the results of the Ordinary Least Squares (OLS) regression are shown and its coefficients plus their interpretation will be discussed. Interpreting a regression with more than one independent variable needs an important caveat. The value of the coefficients can be described as the effect of a variable when all the other variables are held constant. All the results considered 'significant' are significant at a 5% significance level.

This table shows the response of the independent variable *salary* in percentages when one of the independent variables changes with one unit. 1902 observations form the years 2012 and 2013 have been used to perform the regression. In this section all the results will be reviewed and discussed to get a better understanding of the model that is used.

First thing that comes to notice is that there has been tested for a significant change in salary amongst the two years. Because of the low significance it is concluded that there is no trend detected between the two years and therefore the variable *year* is not a reliable forecaster to predict someone's salary.

The *level of education* on the other hand seems to be an excellent predictor of salaries. Relative to someone's salary when they only finished primary school, the salary of people who finished a higher degree of education is improving. It is worthwhile to notice that individuals who finished MBO do not necessarily earn more than individuals who only finished VMBO. Even more interesting to notice is the outcome of individuals who only completed HAVO or VWO. They earn significantly more on average than individuals who just finished VMBO or MBO. After that, the outcomes are, as one would expect it to be. Individuals with an HBO degree earn much more than the ones we mentioned before and individuals with a university degree get paid the most, almost 65% more on average than people who only finished primary school.

The current profession individuals exercise, do have a significant effect on salary. It does not matter if you are working in the agrarian sector or doing unskilled and trained manual work. They both pay the least of all the nine possible sectors one can work in. A higher supervisory professional is paid the most and earn almost two times the salary of an unskilled manual worker or agrarian worker.

Although an accurate measure of work experience was not accessible, using age looks like an adequate measurement to predict someone's salary. Becoming older results in a higher salary. The negative coefficient of the age-squared is also significant and can be explained as the decreasing rate of which older people's salary increases and eventually their salary will not increase anymore, because they become less healthy and are not able to work as before. This function is inverted U-shaped.

The hours worked variable shows a significant positive coefficient, which indicates that working more hours in a week contributes to a higher salary. It is unsure if the individuals get paid by the actual amount of hours worked or if they get paid what is stated in their contract regardless of the actual hours worked. Nevertheless, it is assumed that individuals who are working more hours, or have more working hours stated in their contract, compared to someone else working less hours, doing the same job, get paid more.

The effect of having kids is insignificant and therefore, based on this research, there cannot be asserted that having kids downgrades your salary in comparison to an individual who does not have any children. In addition to this variable, an interaction effect between being female and having kids is created to see if women react differently to having children than men. This result also does not give us any significant results.

Company size does seem to have a significant effect on wage, but from a certain level. When company size 2 (5-249) is compared to company 1 (0-49) there is no significant difference in salary but individuals who work at a company with size 3 (250+) earn 15.1% more on average than individuals working in a company with size 1.

The last two results are the ones that should define if there is an adjusted wage gap between men and women. First, there is a significant outcome stating that females earn 19.6% less salary when all of the other variables are held constant. However, the results of the interaction effect between being female at different company sizes are not significant. Therefore, this model does not prove any differences in salary changes between being a female and moving from a small to medium or large firm, a medium to large firm or vice-versa or being a male and moving from a small to medium or large firm, a medium to large firm or vice-versa.

The last value that has to be discussed is the constant term. The constant term in the logarithmic regression is the base value. The logarithm of base value *a* of a number *x* is the power for which one has to raise base value *a* to get *x* as a solution. The base value here is 5.44.

Source	SS	df	MS	Number of obs	=	1,902
				F(24, 1877)	=	61.82
Model	767.462519	24	31.977605	Prob > F	=	0.0000
Residual	970.960006	1,877	.517293557	R-squared	=	0.4415
				Adj R-squared	=	0.4343
Total	1738.42252	1,901	.914477919	Root MSE	=	.71923

The R-squared of the regression is 0.4415. R-squared is the fraction of the dependent variable variance that can be forecasted from the independent variables. As a rule of thumb, the higher the R-squared the higher the trustworthiness of the model, but exceptions of this rule are not uncommon. Almost half of the data's variation is explained around its mean regarding this outcome.

Blinder-Oaxaca decomposition

For the second part of the results, a Blinder-Oaxaca decomposition is run for the three different company sizes and the differences in means for the groups male and female are pooled in an explained part, thus the part that is explained as a difference in characteristics of men and women, and an unexplained part which can be seen as the part in which discrimination can be one of the reasons for the difference in wage. The reason that there is chosen to take a look at the explained and unexplained part in addition to the difference in pay only is because recommendations about closing the gender pay gap can be given, based on the results. In order to make the results interpretable and comparable between the three different companies, the unexplained part is to be fully conceived as a result of discrimination. The 5% significance is also chosen for the Blinder-Oaxaca decomposition.

The following table shows the results for company size 1 (0-49 employees):

```
Blinder-Oaxaca decomposition                    Number of obs     =        948
                                                  Model           =     linear
Group 1: female = 0                               N of obs 1      =        474
Group 2: female = 1                               N of obs 2      =        474
```

log_salary	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
overall						
group_1	9.888686	.0400815	246.71	0.000	9.810128	9.967245
group_2	9.314945	.0488501	190.68	0.000	9.219201	9.41069
difference	.5737408	.063189	9.08	0.000	.4498926	.697589
explained	.2796694	.052456	5.33	0.000	.1768575	.3824812
unexplained	.2940715	.0568967	5.17	0.000	.1825559	.405587

There is a significant difference of 0.57 in salary between men and women who are working at a company of size 1, because the P-value of ‘difference’ and ‘unexplained’ are below 0.05. The percentage that is due to differences in characteristics between males and females is (0.2796694 / 0.5737408) * 100% ≈ 49%. 51 percent of the difference in salary is therefore a result of discrimination.

```
Blinder-Oaxaca decomposition                    Number of obs     =        523
                                                  Model           =     linear
Group 1: female = 0                               N of obs 1      =        336
Group 2: female = 1                               N of obs 2      =        187
```

log_salary	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
overall						
group_1	9.990046	.0501461	199.22	0.000	9.891761	10.08833
group_2	9.586262	.0625602	153.23	0.000	9.463646	9.708878
difference	.4037837	.0801773	5.04	0.000	.246639	.5609284
explained	.1688444	.0667097	2.53	0.011	.0380958	.2995931
unexplained	.2349393	.0776497	3.03	0.002	.0827486	.3871299

At the companies size 2 (50-249 employees) there is even a bigger matter of discrimination, namely (0.2349393 / 0.4037837) * 100% ≈ 58%. However, the difference in pay between men and women is less than at companies of size 1. The results are all significant, since the P-values are below 0.05.

```
Blinder-Oaxaca decomposition                    Number of obs     =        431
                                                  Model           =     linear
Group 1: female = 0                               N of obs 1      =        257
Group 2: female = 1                               N of obs 2      =        174
```

log_salary	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
overall						
group_1	10.21379	.0449901	227.02	0.000	10.12562	10.30197
group_2	9.806886	.0607985	161.30	0.000	9.687723	9.926049
difference	.4069082	.0756344	5.38	0.000	.2586674	.555149
explained	.1665595	.0496708	3.35	0.001	.0692065	.2639125
unexplained	.2403488	.0690366	3.48	0.000	.1050395	.375658

The results of the decomposition of company size 3 (250+ employees) look like the results of the decomposition of company size 2. There is a small difference in the presence of discrimination, namely 1%: (0.2403488 / 0.4069082) * 100% ≈ 59%.

Remarkable is the bigger gap in salary at company size 1, but at the same time lowest presence of labor market discrimination. There has to be expound that, in reality, not the whole unexplained part of the wage difference should be attributed to labor market discrimination so we cannot know

for sure that it is the case, but it can be assumed that a lower unexplained component attends with a lower attendance of discrimination.

Another thing that has to be mentioned is that, when the Blinder-Oaxaca decomposition was run on Stata, at company size 2 and 3 the notification ‘zero variance coefficients’ popped up. This means that the coefficients of some dummy variables that were normalized are constant and thus have a unique value. Despite this flaw, the decomposition can still be used. A function at Stata called ‘relax’ causes oaxaca to continue its computation even if some coefficients have zero variances.

6. Conclusion

The objective of this thesis was to prove if there are any signs of a visible relationship between the size of a company and gender salary inequality going on in The Netherlands. Prior to the research, there was reason to believe that a positive relationship between these two phenomena would be noticeable. There are two reasons that feed the belief of a positive correlation between these matters of fact, namely the Employer Size-Wage Effect (ESWE) and the increasing wage gap throughout the wage distribution. In order to find out if there is indeed a connection between the two topics, an Ordinary Least Squares (OLS) regression and a Blinder-Oaxaca decomposition are performed. Unfortunately, both results were not in line with each other.

The OLS regression has not given any significant outcomes that support the earlier presumption. On the other hand, significant results are found by using the Blinder-Oaxaca decomposition, although the results do not support the hypothesis. On the contrary, a bigger pay gap is noticed at smaller firms. It can therefore be rejected that the size of a company has got a positive effect on the gender pay inequality. There cannot be concluded with certainty that there is a negative effect of company size on the gender pay gap, because the difference in these gaps between companies of size 2 and companies of size 3 is almost non-existent. To not be skating on thin ice, the conclusion can be that the gender pay gap is considerably biggest at small firms in comparison to bigger firms at most. It also cannot be excluded that there is no relationship at all between the two matters.

7. Discussion

The internal validity of the used research design can be in danger when the experimental subjects are not randomly chosen. This is called selection bias. In this research people are used from LISS data panel and are randomly drawn from the Dutch population. However, the two years that are used (2012 and 2013) contained a lot of the same individuals. This can give rise to the existence of heteroscedasticity, which means that some individuals have different variabilities from others. The variance of the residuals can be dependent on the value of the dependent variable.

In the conclusion it is said that the gender pay gap at small firms is bigger than at larger firms. More than half of the pay gap can be attributed to differences in characteristics. A lot of reasons for this gap can be pointed out. Women at small companies maybe tend to work more often part-time than females who are working at bigger companies, because women at bigger companies might have children less often as a consequence of focusing more on their career instead of focusing on having a family. The type of profession can also be a crucial factor in the gender pay gap. Maybe females fulfill the less paying jobs more often at small companies in comparison to females at bigger companies. This can be due to the possibility that women at smaller companies often have finished a lower level of education and are therefore less often suited for higher paying occupation.

There are a few limitations that have to be put forward. First of all, data of two years is gathered whereof 1902 observations remained. This is more than enough to run a decent regression, but more observations of more years could always improve validity results. Second, the measure that is used to indicate the size of a company is limited. To get a good indication of company size, more features can be taken into consideration. Total assets, total sales and market capitalization of a company for instance, give a financial measurement for the size of a company (Dang & Li, 2015). A third restriction is the usage of age instead of work experience. Unfortunately, data of work experience was not available at the LISS data panel site. Age does not give an accurate representation of the effect that work experience would have had on the size of the salary. A fourth restriction is the decreasing gender pay gap throughout the wage distribution that was noticed at the descriptive statistics. It is not uncommon in The Netherlands that women are out

of work for some years when the children are young. This can influence a gender effect in work experience for people of the same age. Lastly, the unexplained part of the Blinder-Oaxaca decomposition makes it hard to point a finger at the cause of the gender pay gap and makes it especially worrying when this part occupies much of the gender pay gap compared to the explained part.

8. Recommendations

The difference in wage between males and females could be less if work experience would be used instead of age. The years of experience have a bigger effect on the net hourly wage for women than for men (Ruijter, 2002). The same is concluded by other Dutch statisticians (Bakker, Tijdens, & Winkels, 1999).

In 2015 a student of the Erasmus University Rotterdam examined the effect of firm size on gender wage discrimination and found significant results, saying that wage discrimination was highest at the smallest firms, using an extended Mincer function (Engels, 2015). In this research a distinction was made between working in the public sector and working in the private sector and six years of data were used. For future research can be suggested to include the distinction of current working sector and the adding of more years of data to get the biggest chance of significant results.

The limitation of company size measurement can only be solved when the additional data about financial size of a company will be available in the future. Till then, researchers are restricted to the data that offers information about the amount of employees working at a firm.

The results of the Blinder-Oaxaca decomposition showed that at every company more than half of the gender pay gap cannot be explained. This implicates that discrimination on the labor market still plays a huge role in the salary inequality between men and women nowadays. It looks like law alone is not enough to prevent labor market discrimination. If the presence of labor market discrimination can ever be fully averted is questionable, but that there has to be done more to tackle and reduce the existence of this problem. A new law could make it compulsory for companies to be transparent about their employees' salary.

The explained part can also be partially straightened out. Because women take more often and more months off when they are having a baby, more companies should provide paid leave or subsidize childcare. Offering such utilities is not yet the norm for most companies. Changing that can decrease the gender pay gap. Providing childcare and providing paid leave enable women to start working again earlier and will fade out the gender difference in years of working experience.

9. References

Albrecht, J., Björklund , A., & Vroman, S. (2003). Is There a Glass Ceiling in Sweden. *Journal of Labor Economics* , 145-177.

Arulampalam, W., Booth, A., & Bryan, M. (2007). Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution. *ILR Review*, 163-186.

Bakker, B., Tijdens, K., & Winkels, J. (1999). Gender, occupational segregation and wages in the Netherlands. *Netherlands Official Statistics*, 36-41.

Blau, F., & Kahn, L. (1992). The Gender Earnings Gap: Learning from International Comparisons. *The American Economics Review* , 533-538.

Blau, F., & Kahn, L. (1994). Rising Wage Inequality and the U.S. Gender Gap. *The American Economic Review* , 23-28.

Blau, F., & Kahn, L. (2007). The Gender Pay Gap. *The Economists' Voice*.

Blau, F., & Kahn, L. (2016). *The Gender Wage Gap: Extent, Trends, and Explanations.*

Brown, C., & Medoff, J. (1989). The Employer size-wage Effect. *Journal of Political Economy*, 1027-1059.

Curtin, E. (1970). *White-Collar Unionization.*

Dang, C., & Li, F. (2015). *Measuring Firm Size in Empirical Corporate Finance.*

de Goede, M., van Veelen, R., & Derks, B. (2016). *Financiële Beloning van Mannen en Vrouwen in de Wetenschap.* Utrecht.

Engels, S. (2015). *The Effect of Firm Size on Gender Wage Discrimination in the Netherlands.*

Foulkes, F. (1980). *Personnel Policies in Large Nonunion Companies.*

Fransen, E., Plantenga, J., & Vlasblom, J. (2010). *Why do women still earn less than man? Decomposing the Dutch gender pay gap, 1996-2006.*

Freeman, R., & Medoff, J. (1984). *What Do Unions Do?*

Garen, J. (1985). Worker Heterogeneity, Job Screening, and Firm Size. *Journal of Political Economy* , 715-739.

Gender pay gap in unadjusted form. (n.d.). Retrieved from eurostat: ec.europa.eu/eurostat/web/products-datasets/-/tsdsc340

Gibson, J., & Stillman, S. (2009, February). Why do big firms pay higher wages? Evidence from an international database. *The Review of Economics and Statistics*, 213-218.

glass ceiling. (n.d.). Retrieved from https://www.merriam-webster.com/dictionary/glass%20ceiling

Gould, E., Schieder, J., & Geier, K. (2016). *What is the gender pay gap and is it real?*

Gowling, A. (2014, february 26). *Gender pay gap in netherlands higher than EU average*. Retrieved from iamexpat: https://www.iamexpat.nl/expat-info/dutch-expat-news/gender-pay-gap-netherlands-higher-eu-average

Grimshaw, D., & Rubery, J. (2002). *The adjusted gender pay gap: a critical appraisal of standard decomposition techniques.* Manchester.

Hamermesh, D. (1980). Commentary. *The Economics of Firm Size, Maket Structure, and Social Performance* .

Heckman, J., Lochner, L., & Todd, P. (2003). *Fifty Years of Mincer Earnings Regressions.*

Jackson, S. (1993). Women's Studies: Essential Readings. New York University Press.

Krusell, P., Ohanian, L., Ríos-Rull, J.-V., & Volante, G. (2000). Capital-Skill Complementary and Inequality: A Macroeconomic Analysis. *Econometrica*, 1029-1053.

Lazear, E. P. (1974, August). *Age, Experience and Wage Growth.* doi:10.3386/w0051

Lester, R. (1967). Pay Differentials by Size of Establishment . *Industrial Relations*, 57-67.

Masters, S. (1969). An Interindustry Analysis of Wages and Plant Size . *The Review of Economics and Statistics*, 341-345.

Mellow, W. (1982). Employer Size and Wages. *The Review of Economics and Statistics* , 495-501.

Mincer, J. (1974). *Schooling, Experience, and Earning. Human Behavior & Social Institutions No. 2.*

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review* , 693-709.

Oi, W. (1983). The Fixed Emmployment Costs of Specialized Labor. *The Measurement of Labor Cost* , 63-122.

Pereira, P., & Martins, P. (2004). Does education reduce wage inequality? Quantile regression evidence from 16 countries. *Labour Economics, 11*(3), 355-371.

Ruijter, j. d. (2002). *Occupational Wage differences: A gender Approach.* Rotterdam: Erasmus University Rotterdam, dissertation.

Scherer, F. (1976). Industrial Structure, Scale Economies, and Worker Alienation. In R. Masson, *Essays on Industrial Organization in Honor of Joe S. Bain* (p. 111).

Stafford, F. (1980). Firm Size, Workplace Public Goods, and Worker Welfare. *The Economics of Firm Size, Market Structure, and Social Performance* .

Stigler, G. (1962). Information in the Labor Market. *Investment in Human Beings*, 94-105.

Tijdens, K., & Klaveren, M. (2011). *De loonkloof tussen mannen en vrouwen: Een review van het onderzoek in Nederland.* Amsterdan: University of Amsterdam.

Vaccaro, G. (2015). *How to Reduce the Unexplained Gender Wage Gap? Evidence from a Regression Discontinuity Design.*

Vrooman, A. (2017, march 7). Loonverschil tussen man en vrouw in de EU gelijk gebleven. *Het Financieele Dagblad* .

Wagstaff, A., O'Donnell, O., Doorslaer, E. v., & Lindelow, M. (2008). Explaining Differences between Groups: Oaxaca Decomposition. In *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation.* (pp. 147-158). World Bank Publications.

Weiss, A., & Landau, H. (1984). Wages, Hiring Standards, and Firm Size. *Journal of Labor Economics* , 477-499.

Weiss, L. (1966). Concentration and Labor Earnings. *The American Economic Review*, 96-117.

What are Longitudinal Data. (n.d.). Retrieved from www.caldercenter.org/what-are-longitudinal-data