VILNIUS UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

JUSTYNA LAVRYNOVIČ

3rd year, Quantitative economics

Bachelor's thesis

GENDER PAY GAP OVER THE LIFE CYCLE IN LITHUANIA

Supervisor dr. Linas Tarasonis

Vilnius, 2022

Table of Contents

1. Introduction

In the past several decades gender wage gap have played a significant role when talking about gender inequality around the world. This problem attracts more attention than earlier due to the growth in demand of equal opportunities in society. Recent research have revealed the complexity of the pay gap between men and women. This field of study deals with many possible factors that affect the gap. Most of the theories are focused on explaining the drivers that cause the problem of inequality. Gender gap is being widely investigated in recent times. Therefore, nowadays, it still remains the sector of labour economics that is actively researched. The focus on Lithuanian market is chosen due to limited investigations in this country. Researches done in other countries assist for this paper in estimating more detailed results. Carefully selected literature helps to explain the basis of the gender gap. It represents recent findings of the gender pay gap drivers in European Union. Discoveries from a variety of economists across the world gives the opportunity to examine the applicability of these theories for Lithuanian market. Analysis of the pay gap over the life cycle can lead to a breakthrough for new actualities that could help to reduce inequality. Statistics are the key tool for discovering new methods for gender wage gap reduction. Diminishing the wage gap is beneficial to the economy of the country, as it is proven by several recent studies. It also tends to be necessary for civilization, as it profits not only monetary, but also morally. Overcoming the problem of gender pay gap would indicate the growth of the society and equal opportunities for all genders not only in labour market, but in a way wider spectrum. Lithuania is still struggling with post-soviet ideology syndrome, which distinguishes with discrimination and old-fashioned beliefs. However, the modern community is progressing steadily and there is less and less discrimination based on gender. To overcome this problem it is significant to find the key factors of gender discrimination problem in labour market. Removing gender pay gap or at least decreasing it to a minimum should be a priority in Lithuania. Equal career opportunities for women should not be a privilege, but a necessity.

The main idea of this paper is to investigate the evolution of gender wage gap over the life cycle. Particular trends can illustrate the main factors that cause the pay gap. Differences between age groups can lead to outstanding results while looking into this problem. Specific results in particular age categories can be significant for identification of pay gap drivers. For example, child penalty as a cause of gender inequality directly corresponds with age of a woman. Therefore, investigation that is based on age gives an opportunity to understand the impact of this phenomenon. In addition, estimating the size of unadjusted gender pay gap helps to evaluate the problem of gender inequality in Lithuania and compare it with other European union countries. The unadjusted wage gap provides an overall picture of gender pay inequality. This international

measure is used widely, thus analysing it globally is a straightforward approach. The ultimate goal is to investigate possible trends that are proposed by used microdata and to suggest explanations of the results. All estimations are done while using microdata of 2014 Structure of Earnings Survey. 2018 data is also publicly available, although 2014 survey was chosen due to a better fit for this research. However, newer data is used for comparison and assists in observation of the wage gap throughout years. Comparison of 2014 and 2018 unadjusted wage gap in Lithuania helps to indicate the change in the past decade. It could help to prove that there is an improvement and Lithuania is becoming more critical about gender-based discrimination. This research aims in finding possible solutions and suggestions for the government to overcome or at least lessen the detrimental effects of the gender inequality problem. Proving that the biological difference between men and women is one of the main factors that affect the pay gap could help to find possible solutions for diminishing it. Child penalty is a complex issue for women in labour market. Unpaid work a home, childcare tend to affect not only compensation, but also choices of a woman. Maternity leave interrupts women's career and result in slowing human capital accumulation, which effectuate a slower wage growth over the life cycle. However, biological difference is only one element from a variety of possible reasons of the gender gap. The factors that make impact on gender inequality in Lithuanian labour market are complicated. Analysing hourly wage differences over the life cycle helps to investigate those factors.

On average, women in European Union earn 13% less per one hour of work comparing to men in 2020.[1] Over the last decade gender pay gap is mostly reducing, however some countries notice opposite trends. Comparing unadjusted pay gap in Lithuania with other countries and with 2018 data will help to see the overall picture how pay gap changes. Investigating the drivers of the gender gap in different age groups has a potential to be a pre-condition for designing new policies for reduction of the pay gap. Adapting main drivers from the previous research will help to see if the causes of gender pay gap are similar in different countries.

For this study multiple linear, ridge and lasso regressions are chosen to find the best fit for the data. Multiple linear regression is considered to be one of the best ways for such a research and is used for main calculations. Ridge and lasso models are also evaluated to highlight advantages and disadvantages of calculations. Hourly wage performs as a dependent variable, that is measured on a continuous scale. Independent variables (sex, company size and etc.) are mostly nominal, which means that they are split into categories, female and male for example. Logarithmic approach for hourly wage has an advantage of resulting in more precise results. It improves the accuracy of multiple linear regression results by equalizing the hourly wage vector.

---

[1] Based on data presented by Eurostat – Statistical office of European Union.

The limitation of this method is that it is more difficult to present and summarise the results. Log-percent differences are one the same as percent differences, which makes it complicated to compare it with past findings. However, the main goal of this paper is to see the differences between age groups and it has no detrimental effect on this task. At first, multiple regression is done only by adding one control variable – age. This seems to be the best approach for investigating the differences between all given age groups. Therefore, more control variables are being added step by step, to see how the model changes. It is expected to see an improvement of the wage gap as control variables should explain some part of the wage gap.

The plan of the paper is as follows. In section 2 past findings on the gender gap are used to explain the basis of it. It begins with gender inequality as it has been the main cause of compensation differences across the world. Then the possible wage gap drivers are emphasised from different theoretical papers and empirical experiments. In section 3 empirical setting of research is introduced. In the beginning of this section the structure of the used data is briefly explained. In section 4 the methodology of the work is combined with the results to avoid repetitiveness. In section 5 conclusion about the outcomes of this research is added with the following suggestions for the government on how to improve the current policies for diminishing gender inequality in labour market.

2. Related literature

2.1. Gender inequality

Gender inequality is a well-known negative social phenomenon. Different genders experience specific privileges and disadvantages across many areas, such as education, religion, health care, politics and labour market.[2] Every continent and even every country has a different experience with gender inequality. One of the main drivers of different privileges are religion and cultural approach. Such drivers explain the difference of the gender gap in countries that differ by the level of development. Gender inequality in developing countries can be mainly explained by the cultural norms resulting in normalization of the favouritism of men.[3] Some traditions and religions diminish the impact of women to a minimum and leave no space for equality. On the other hand, the majority of developed countries have a whole different approach to women. Cultural norms allow women to be an equal part of society. Females and males have the same possibilities for education, medical care, voting and participation in politics. However, the labour participation and wage gap between genders still exists. One of the ways of understanding the equality of the country on the gender basis is the examination of Gender Equality Index. It is a complex indicator created for measuring and comparing structural gender equality in the European Union.[4] Figure 1 shows the gender equality index in European Union and Lithuania. As indicated

Figure 1

Gender equality index 2021
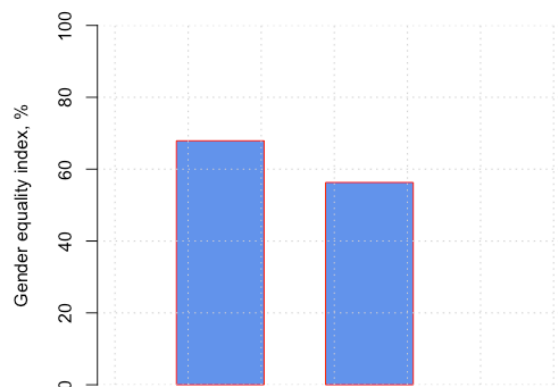
Source: made by author, based on European institute for gender equality data

[2] Lorber, J. (2001). Gender inequality. *Los Angeles, CA: Roxbury*.

[3] Jayachandran, S. (2015). The roots of gender inequality in developing countries. *economics*, 7(1), 63-88.

[4] Bericat, E. (2012). The European gender equality index: Conceptual and analytical issues. *Social Indicators Research*, *108*(1), 1-28.

by the figure, the gender equality index in Lithuania is 11.6% lower than the average EU rate in 2020.

Gender discrimination starts in early age, but going into labour at older age makes the investigation possible as there are more numbers that can be compared. *"The valuation in the market-place of personal characteristics of the worker that are unrelated to worker productivity",-* discrimination defined by Kenneth Arrow[5]. Nowadays, women often face discrimination at work, which is strongly driven by stereotypes. An example of such discrimination would be a common belief that an average male employee is more productive and motivated than a female. Distrust of employers and consumers in women as specialists leads to the discrimination in the labour market.[6] Scepticism and social stereotypes negatively influence the employment and wage gaps.

2.2. Gender participation in labour market

In labour market gender gap is evident not only in compensating, but also in working-time and employment. Globally, women participate less in the labour because of different factors. Cultural background, maternity, house work are the most common drivers of the gender gap. In addition, in paid employment men accordingly work more hours than women. These factors are diminished in developed countries as religion and culture is not an obstacle. However, maternity leave and discrimination in general are still relevant problems. In many countries women face higher likelihood of unemployment which is not related to human capital. Long-established stereotypes is one of the main reason in European Union forming a gender gap in employment and working time. Statistical office of EU reports that in 2018, around 30% of women were working part-time comparing to only 8% of men in European Union. Part-time work can be a reflection of women's personal choices, however, may also be a result of child-care responsibility and unpaid work at home.[7] The prioritisation of household care reflects the social norms that form individual preferences of women.[8] However, in Lithuania part-time job is popular only among students and specifically for less-qualified jobs. Positions that need tertiary degree usually do not offer part-time contracts.

---

[5] Kenneth Joseph Arrow – an American economist, joint winner of Nobel Memorial prize in economic sciences.

[6] Pettinger, T. (2017). The economics of discrimination. Oxford, UK.

[7] Ciminelli, G., Schwellnus, C., & Stadler, B. (2021). Sticky floors or glass ceilings? The role of human capital, working time flexibility and discrimination in the gender wage gap.

[8] Bertrand, M. (2020, May). Gender in the twenty-first century. In *AEA Papers and proceedings* (Vol. 110, pp. 1-24).

According to Services to employees in Lithuania, women take more time to get employed and are more likely to get a terminated job offer. Nevertheless, it does not only depend on employers, but also on employees. Women tend to take lower-paid, safer jobs that require less working hours, which also interacts with the level of compensation. Participation in child-care and household tasks requires flexible working hours, so finding a job that meets expectations becomes a harder task for women. Social norms and stereotypes that it is more efficient to hire a man create a competitive environment for women in labour market. On average, women are considered to be less productive at work than men. Status of a woman also matters, making women less attractive to the employers if they are married or have children. Childcare and possibility of maternity leave penalise the labour participation of women.[9] A few studies have shown that gaps in employment rate or working hours are smaller or there is no gap at all between men and childless women.[10]

Gender employment gap has been investigated by many laboratory-based experiments. Some studies have shown that women are less competitive than men and are less confident even with the same competence level. Females tend to underestimate their abilities, thus affecting the growth of their careers.[11] Other studies show that women are more likely to be risk-averse when it comes to financial risk taking. It is the evidence of differences in psychological traits of genders.[12] However, gender gap is declining, as women labour participation correspondingly

Figure 2

Employment gap 2018
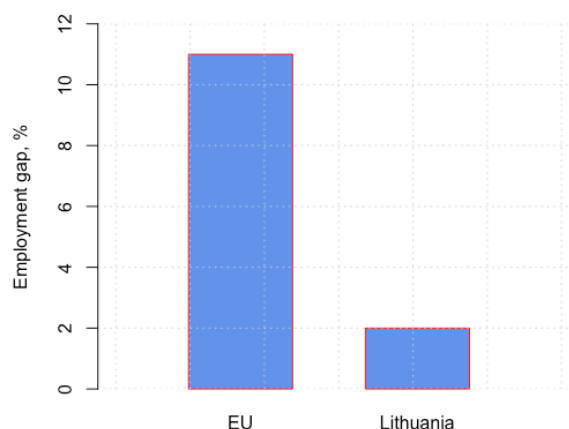
Source: made by author, based on statistical office of European Union 2018 data

---

[9] Meurs, D., & Ponthieux, S. (2005). The gender wage gap in Europe: women, men and the public sector. *WP F0502*.

[10] Angelov, N., Johansson, P. & Lindahl, E. J. (2016). Labor Econ. 34, 545–579

[11] Kay, K., & Shipman, C. (2014). The confidence gap. *The Atlantic*, *14*(1), 1-18.

[12] Schubert, R., Brown, M., Gysler, M., & Brachinger, H. W. (1999). Financial decision-making: are women really more risk-averse?. *American economic review*, *89*(2), 381-385.

increases with human capital.[13] There is a significant progress compared to 20th century, however women still struggle with discrimination. Figure 2 shows the gender employment gap between women and men in European Union and Lithuania as percent of population aged 20 to 64. The gender employment gap is measured as the percentage point difference between the employment rates for men and women. The employment gap is ~9% lower in Lithuania than EU average employment gap.
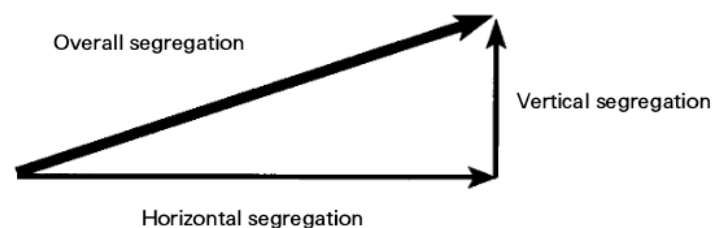
2.3. The gender wage gap
2.3.1.   The basis of the gender pay gap

The gender pay gap is the average gross hourly earnings difference between men and women. Usually, occupation, hours worked, education level and experience are taken into account. The reasons of the gender pay gap are complex and depend on the gender segregation. Firstly, it is caused by horizontal segregation, with more women working in lower paid industries. For instance, OECD (2012) shows that, in advanced economies, 85% of women are employed in services, mainly in education and health. Secondly, vertical segregation influences the gap with less women in better paid positions. It is better known as a "glass ceiling", an invisible barrier that prevents the rising of the career.[14]

Most researchers would agree that there is no single explanation for gender-based segregation. Identified key factors of segregation are biological privileges, under-investment in human capital, income differential, personal preferences, stereotypes and social norms.[15] Figure 3 presents how vertical and horizontal segregations form the overall segregation in labour market.

Figure 3

Overall segregation

Source: Blackburn, R.M., Browne, J., Brooks, B., Jarman, J. (2002). Explaining gender
segregation. The british journal of sociology, 53(4), 513-536.

---

[13] Kunze, A. (2018). The gender wage gap in developed countries. *The Oxford handbook of women and the economy*, 369-394.

[14] Cotter, D. A., Hermsen, J. M., Ovadia, S., & Vanneman, R. (2001). The glass ceiling effect. *Social forces*, *80*(2), 655-681.

[15] Bettio, F., Verashchagina, A., Mairhuber, I., & Kanjuo-Mrčela, A. (2009). *Gender segregation in the labour market: Root causes, implications and policy responses in the EU*. Luxembourg: Publications Office of the European Union.

2.3.2. Gender pay gap explanations

One explanation of the gender wage gap suggests that women tend to have lower wages due to the fact that they accept it. Women choose lower-paid jobs in return for non-financial benefits.[16] Flexible working-schedule or less working-hours give them a possibility for childcare and unpaid work at home. The US evidence suggests that such jobs with more flexibility usually offer lower wages than those with inflexible and long working-hours.[17] In addition, it is worth mentioning that the gender wage gap is lower in occupations where less personal contact is needed and working-hours are shorter.[18] Another example of women preferences is a research in Norway about exporting firms, which suggests that gender wage gap is significantly larger in this industry as it requires traveling, working on non-standard times or night shifts.[19] According to the evidence, women key characteristic about the job is a flexible schedule, even when it results in wage differential.

Previous studies further suggest that gender discrimination has a significant impact on the wage gap. Employers when have a choice between identical applicants where the only visible difference is the gender, usually tend to hire men. Women with the same experience, education and skills have a lower likelihood of getting hired. In addition, if they manage to get the job, they are being offered a smaller compensation than men with identical resumes.[20] Gender-based discrimination can be explained by multiple examples. Consumers and employers distrust in women as employees or the belief of women and men productivity differences. Also, some studies show that women labour supply is more inelastic than men which allows employers to offer lower wages.[21] In addition, some studies have shown that firms tend to negatively react to women's wage negotiation. Women are known to be less risk-averse, less competitive and have less bargaining skills. However, the reaction to bargaining women versus bargaining men also differs. The returns of negotiation are more favourable to men rather than women.

---

[16] Ciminelli, G., Schwellnus, C., & Stadler, B. (2021). Sticky floors or glass ceilings? The role of human capital, working time flexibility and discrimination in the gender wage gap.

[17] Cubas, G., Juhn, C., & Silos, P. (2019). *Coordinated work schedules and the gender wage gap* (No. w26548). National Bureau of Economic Research.

[18] Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, *104*(4), 1091-1119.

[19] Bøler, E. A., Javorcik, B., & Ulltveit-Moe, K. H. (2018). Working across time zones: Exporters and the gender wage gap. *Journal of International Economics*, *111*, 122-133.

[20] Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, *111*(12), 4403-4408.

[21] Hirsch, B., Oberfichtner, M., & Schnabel, C. (2014). The levelling effect of product market competition on gender wage discrimination. *IZA Journal of Labor Economics*, *3*(1), 1-14.

2.3.3. Child penalty

Raising children into socially responsible and self-dependent individuals is a hard task for every family. Giving birth to a child or adopting one is a serious commitment which requires a lot of time and life adjustments. It is a life-changing event in every family's routine. Accordingly, it has a huge impact not only on daily life, but also results career changes, especially for women. An old-fashioned belief that men are obligated to earn money and women's job is to raise children and take care of the household is still relevant nowadays. After having a child, men tend to concentrate on their earnings, while new mothers give up their education or career for motherhood. Even while working full-time women hold full responsibility of the household and childcare. Combining these two time-consuming tasks leads women to child-penalty experience.[22] At first, the break during pregnancy and birth-giving period slows human capital accumulation. Women have to pause their education or take a parental leave at work, which leads to a slower career growth. In addition, the further care of a child is rarely shared equally between mother and father. Gender parenting gap is a significant factor that makes an impact on both wage and employment gaps.

According to Figure 4, which represents employment rate differences between men and women depending on their parental status in European Union (2020), women with children have lower employment rate than women without children. However, employment rate increases for men after having children. It can be assumed, that child is an incentive for men to work and to take financial care of the family, while it works in opposite for women. In addition, it is important

---

[22] Baker, M. (2010, May). Motherhood, employment and the "child penalty". In *Women's Studies International Forum* (Vol. 33, No. 3, pp. 215-224). Pergamon.

to mention that employment rate between women with and without children decreases correspondingly with education. So women with tertiary degree tend to keep working after giving birth more than women with primary or secondary education. This could be evidence that the

Figure 4

Employment rate between individuals with and without children
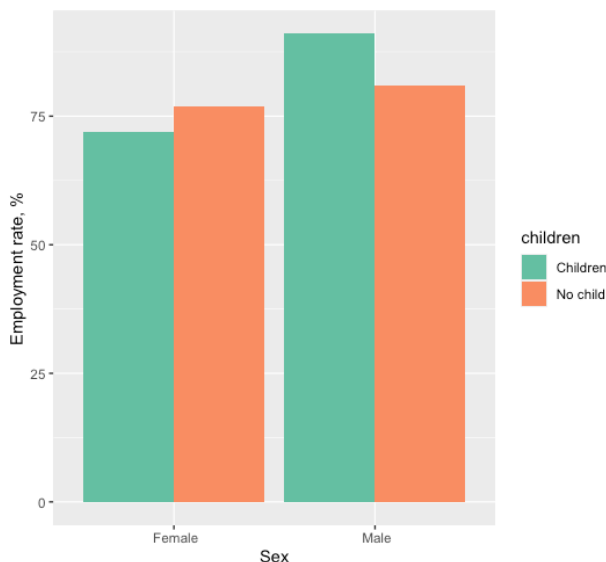
majority of "patriarchal" families consists of less educated women.

Source: made by author, according to Statistical office of European union data 2020

Child penalty's impact is clearly visible when computed wage gap depicts age of workers. Wage gap in Europe is the lowest in the age group of 20-40, which is known as average fertility age period. Figure 5 represents the theoretical explanation of the wage gap over the working time life-cycle. Y-axis shows the evolution of compensation, while x-axis represents the age. At the start of the working time period women earn less than men due to social norms and discrimination.

As it is shown, after the childbirth the wage gap shifts upwards. Slowed human capital accumulation is a result of further childcare and unpaid work at home, which slows down the wage of women.

Figure 5

Wage gap evolution over the life cycle
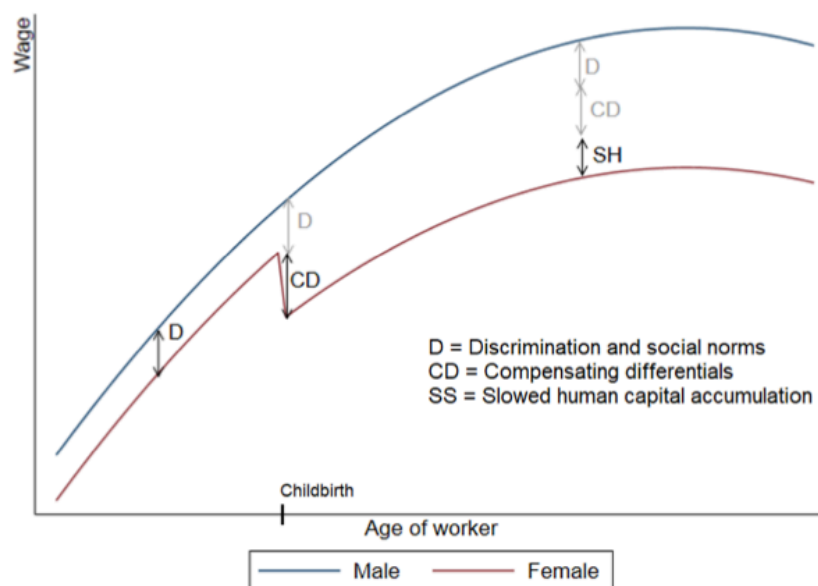
Source: Ciminelli, G., Schwellnus, C., Stadler, B. (2021). Sticky floors or glass ceilings? The role of human capital, working time flexibility and discrimination in the gender gap.

2.3.4. Parental leave in Lithuania

Child penalty is one of the main drivers of gender gap. Parenthood has a negative effect on wage growth and career path. In majority of the families women take parental holidays more often than man. In Scandinavia, it is common to share the parental leave between parents. Government creates a gender friendly environment and equal opportunities for mother and father to share the parenthood. It could ensure a significantly lower percentage of inequality and smaller gender pay gap. However, in Lithuania the situation is different. It is more beneficial for women to take the parental holiday. For example, only every second male took the proposed month off after the child's birth. The reason for this is that the limit of payment is set only for men – the payment cannot excess the average wage of past two months. There is no such limit set for women. In addition, legislation makes it more difficult for men to take the parental leave due to the fact, that the employer is not obligated to give the paid month off at the required time. [23]

---

[23] Pilinkaitė-Sotirovič, V. Kontvainė, V. (2020). Modern men and gender equality: incentives and obstacles of participation in childcare.

Two-years parental leave system in Lithuania is based on financial benefits and does not encourage both genders to involve into childcare. Families often tend to choose an option, where on the second year after child birth man takes the parental leave only on paper and still works full-time. In this case, mother is left to take care of the child because the financial benefit is larger. The system in Lithuania is not flexible and leaves no space for combining work and childcare, thus cannot regulate equal participation in parenthood. Figure 6 represents the distribution of males and females parental leave pay receivers in first and second year after child's birth in 2019.[24] There is a visible rise in second year male data as it was mentioned before.

Figure 6

Parental leave pay receivers in first and second year after child's birth 2019
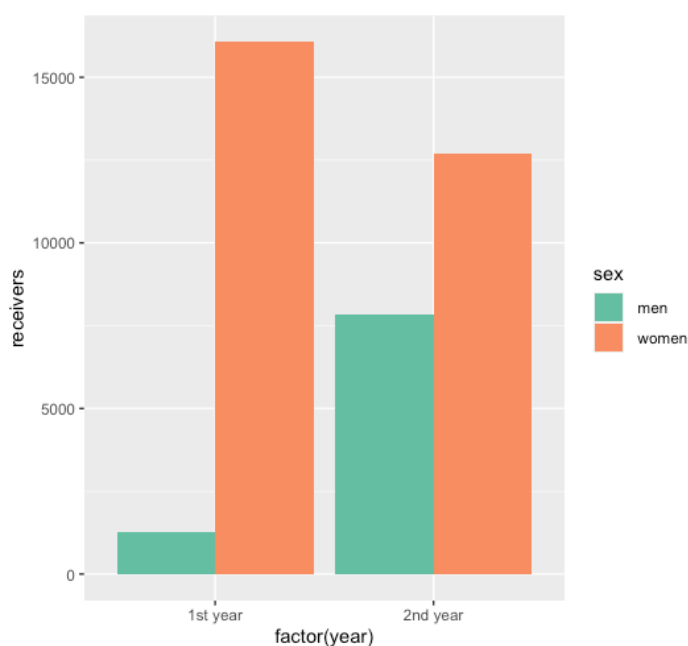
Source: made by author, data used from Pilinkaitė-Sotirovič, V. Kontvainė, V. (2020). Modern men and gender equality: incentives and obstacles of participation in childcare.

Such a system that does not encourage sharing childcare between parents can lead to increasing the wage gap. In comparison, Scandinavian countries have different systems that involves fathers to be take an equal part in parental leave. For instance, in Sweden both parents have 90 days of holiday reserved. As a result, it lowers the gender pay gap to a minimum. However, even when the parental leave opportunities are equal, women's wage growth still falls after giving birth.[25]

---

[24] Data used from official statistical office of Lithuania.

[25] Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of labor economics*, *34*(3), 545-5.

3. Empirical setting

3.1. Data

The source of used data is the Structure of Earning Survey (SES). It is a 4-yearly survey that contains workers information about their annual, monthly and hourly earnings, hours worked, contract type. It also contains characteristics of the worker, such as his sex, age, education level and experience. The research is based on 2014 data, as in 2018 survey results age groups of 14-19 and 20-29 were merged into one making it difficult to see the child penalty impact on gender gap. Also, it would worsen the ability to show the change over the life-cycle. However, this high-quality information allows to find the gender pay gap between similar women and men correspondents in Lithuania. The selection of samples is contains of two stages. Sampled employees are surveyed by specific questionnaires which are made by statistical offices.

The main variable of interest is gross hourly wage of workers. It measures how much workers earn before taxes in one particular month (October in Lithuanian data), accounts for regular and standard payments. It does not count bonuses or other additional payments. SES also provides such information as total hours worked, overtime hours, hours that were paid but not worked (paid sick leave, national holidays and etc.). For confidentiality reasons, age of respondents are divided into groups: 14-19, 20-29, 30-39, 40-49, 50-59, 60+. It was decided to eliminate the age group 60+ from regression, as the wages at this age are no consistent and deform the model. Education data is structured into 4 groups (primary, secondary, bachelor's and master's or PhD). However, to simplify the results of this research education is changed into two groups: primary or secondary and tertiary degree. Furthermore, it was decided to perform regression by using only full-time working respondents. Losing some part of the data is not beneficial, however, it may improve the model as compensations of part-time work are not consistent.

3.2. Descriptive statistics

The attributes of interest:

- Hourly wage: continuous
- Sex: female; male
- Age: 14-19; 20-29; 30-39; 40-49; 50-59
- Education: Primary or secondary degree; Tertiary degree
- Company size: 1-49; 50-249; 250+
- Sector: Private; Public
- Contract: Non-terminated; Terminated

Figure 7 shows that the distribution of logged hourly wage is right-skewed. Y-axis shows the number of workers, while x-axis is logged hourly wage. Positive skewness is a result of few outstandingly high values and explains how data differs from normal distribution. It explains the results presented in Table 1 where mean and median are different, as well as standard deviation is a high number. These features makes it difficult to allocate a typical value, as there is no clear high point of the data. A long-tailed distribution makes the results of multiple linear regression less accurate.

Figure 7

Histogram of log(hourly wage)

Source: made by author, by using Structure of Earnings survey 2014 data

Table 1

Distribution of hourly wage

Target variable	Min	Max	Mean	Median	Std Deviation
hourly wage	5.17	386.63	14.53	12.02	10.68

Source: made by author, by using Structure of Earnings survey 2014 data

Figure 8 illustrates the distribution of the full-time workers by main factor variables (female, age and education) with y-axis being the percentage of respondents. This dataset has approximately 52% male and 48% female observations. The age of workers is similarly distributed except the youngest group of 14-19, which is only 0.65% of observations. This could be explained as people of young age do not usually work, especially full-time. Such a small number of observations can show inconsistent results, however, this age group is important for the research for checking the gap among different age groups, so it is not eliminated. In addition, this data has 8% more observations on employees with primary or secondary education than tertiary degree.

Figure 8

Distribution of company size, sector and contract variables
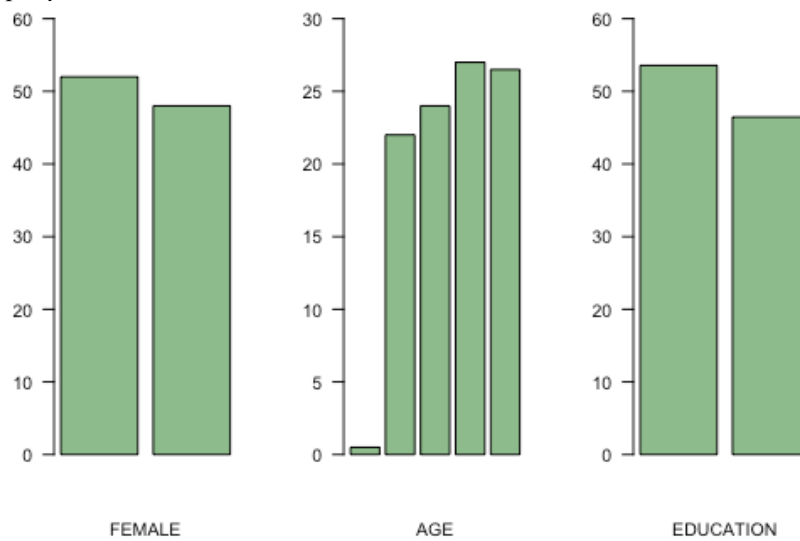
Source: made by author, by using Structure of Earnings survey 2014 data

Figure 9 illustrates the distribution of remaining factor variables (company size, sector and contract). According to the chart, the majority of the respondents work in bigger companies. The workers of small companies that have 1-49 employees form only 20% of the data. Distribution of the sectors is also not equal with only 30% of employees working in public sector and the

Figure 9

Distribution of female, age and education variables
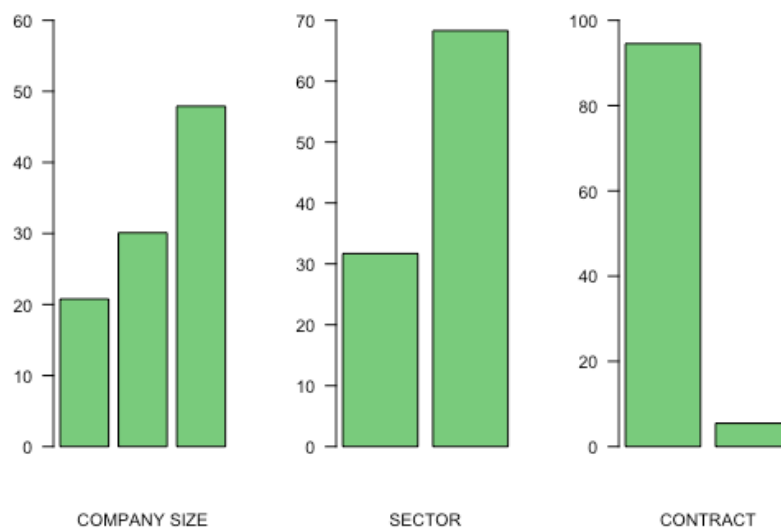
Source: made by author, by using Structure of Earnings survey 2014 data

remaining 70% in private companies. However, it seems reasonable as private companies form the majority of the market. Lastly, only 5% of respondents have terminated contracts. This also can be explained by tendencies in Lithuanian market. Terminated contracts are not valuable among employees as they usually offer lower compensation and less social benefits, so a small percentage of those who have it is reasonable.

3.3. Basis of regression

3.3.1. Multiple linear regression

Multiple linear regression (MLR) is an extension of simple ordinary least-squares (OLS regression). Dependent variable, or variable that is meant to be predicted, is continuous. Independent variables, or predictors, can be either continuous, either factors. Predictor variables used for multiple linear regression are all factors (female, age, education, company size, sector, contract). The output of β coefficients is presented in a table form, which allows to easily explain and compare the results.

3.3.2. Lasso and ridge regression

Lasso regression or the least absolute shrinkage method is also an extension of linear regression. Function is modified to minimise the complexity of the model by setting a limit on the sum of absolute coefficient values. To build a lasso regression model it is necessary to find the optimal (cross-validated) lambda value. Furthermore, to test the performance of lasso regression data is split into two parts: train and test. Evaluation of model on train and test data assists in understanding the overall performance of the model. In addition, lasso regression works by selecting particular coefficients for each predictor variable that minimise the loss function. To avoid over-fitting, regularisation is compulsory for the model.

Ridge regression has a similar approach as lasso regression. The difference is that penalty parameter that is equal to the square of the magnitude of coefficients is added to the model. In addition, the difference between lasso and ridge regressions can be explained by determining the alpha value in the model. Alpha in ridge regression contributes as a parameter which adds a factor of sum of square of coefficients, whereas in lasso it adds sum of absolute value of coefficients. Furthermore, lambda coefficient is added to the model to control the penalty term. Increasing lambda determines values of the coefficients to tend towards zero.

Both lasso and ridge models have some limitations. Ridge regression is easier to perform as it decreases the complexity of the model. However, the number of predictor variables stays the same. Ridge regression minimises coefficients, whereas never leads values to zero. Contrarily,

lasso leads some coefficients to zero, but struggles with data that has a lot of predictor variables. Lasso can determine some coefficients as zero even if they are significant to the model.

4. Results and discussion

4.1. The labour income gap

According to SES data from the 2014 survey wave, on average, women with the same experience and education earn around 15% less than men. However, the gender wage gap is not the only significant measure of gender inequality. Working-time and employment gaps should also be considered when talking about inequality. Together these gaps determine the main differences between men and women in labour market income.

The components of the overall gender gap are illustrated in Figure 10. On average, EU labour income of women is 40% lower than men. According to 2014 data, wage and employment gaps differ by only 1%. Results indicate that average women participation rates in the European Union are low compared to men. However, the labour income gap in Lithuania is almost two times lower than the EU's, by standing above 20%. The wage gap takes the biggest part in forming the labour income differences due to limited working-time and employment gaps. Participation gaps are significantly lower than the average.

Figure 10

Components of the gender gap
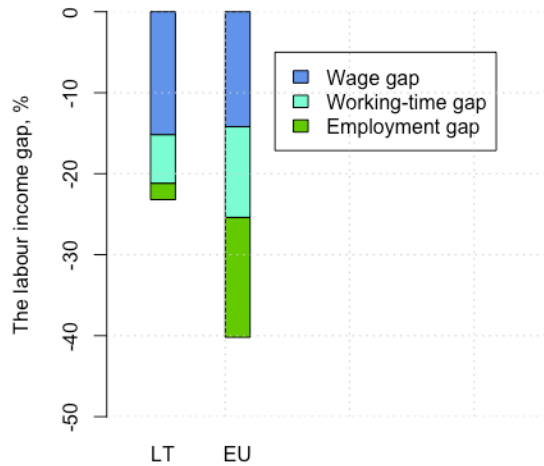
Source: made by author, by using Structure of Earnings survey 2014 data

4.2. Gender wage gap

$$w_i = \beta_{1,i} female_i + \epsilon_i \tag{1}$$

The gender wage gap is calculated by using linear regression. It can be simply explained by Equation 1. By estimating β coefficients the results are summarised in Table 2. Estimates indicate that sex of the worker correlates with shifts of the wage as both predictors are strongly significant. The estimated result of full-time workers wage gap is 16% (calculated possible error is 0.2%).

$$\frac{\beta_{1,0} - \beta_{1,1}}{\beta_{1,0}} \times 100 \tag{2}$$

Error calculation is done by adding and subtracting standard error values from regression multiplied by 1.96 at 5% significance level. The gap is being calculated by eliminating part-time workers from the data, as their wages and working-time are not consistent. Part-time workers 17% of the surveyed individuals, so eliminating such a big part of the data results in losing many respondents. However, it improves the model and makes it more solid.

Measure slightly differs from other completed public calculations. It can be caused by different methods of calculation, utilized data and depends on what factors are taken into account.

Table 2

Linear regression results

Predictors	Estimate	Standard error
female 0	15.74334 ∗ ∗∗	0.079
female 1	13.21376 ∗ ∗∗	0.083

Source: made by author by using Structure of Earnings survey 2014 data

4.3. Gender wage gap with controlling age

$$log(w_{i,j}) = \beta_{1,i} female_i + \beta_{2,j} age_j + \beta_{2,j} age_j \times I(i = 1) + \epsilon_{i,j} \qquad (3)$$

The further estimation of the gender wage gap continues with adding one control variable - age. The estimation of gender wage gap by age is described by Equation 3, where $w_{i,j}$ is the hourly earnings of individual, female is gender dummy variable, and age is a dummy variable for different age groups (14-19, 20-29, 30-39, 40-49 and 50-59).

Figure 11 explains the results of multiple linear regression and shows the evolution of gender wage gap over the life-cycle. The graph shows log-percent differences between women and men hourly wages by age, and is obtained by estimating β coefficients from Equation 3. The graph is obtained by Equation 4. The shaded area is two standard errors confidence band.

$$\frac{\beta_{1,1,j} - \beta_{1,0,j}}{\beta_{1,1,j}} \qquad (4)$$

The 14-19 age group has the smallest wage gap. It can be explained by the focus young people put on their education and career, while not needing to take care of children or housework. It is visible that the wage difference is the biggest at the age of 30-39. In the European Union, the majority of first births are given by women aged 20-39. The mean age of first time mothers is

Figure 11

Evolution of the pay gap over the life cycle
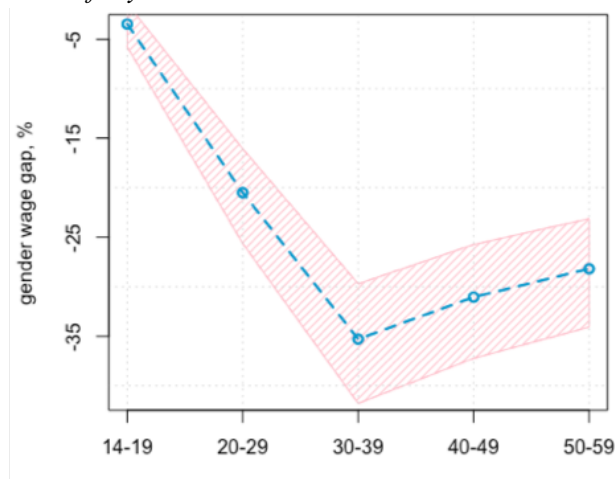
Source: made by author by using Structure of Earnings survey 2014 data

Table 3

Multiple linear regression results 1

Predictors	Estimate	Standard error
Female 0	2.139 ***	0.042
Female 1	2.067 ***	0.063
age 20-29	0.332 ***	0.042
age 30-39	0.560 ***	0.042
age 40-49	0.454 ***	0.042
age 50-59	0.421 ***	0.042
female:age 20-29	−0.016	0.077
female:age 30-39	−0.073	0.077
female:age 40-49	−0.088	0.077
female:age 50-59	−0.070	0.077
multiple R^2	0.9583	
adjusted R^2	0.9583	

Source: made by author by using Structure of Earnings survey 2014 data

increasing, for example the age went from 26.2 years in 2014 to 28.9 years in 2018.[26] It indicates that the model in graph 4 could be the evidence of the child penalty impact. After reaching the maximum, the wage gap stabilises within the age groups of 40-49 and 50-59. Table 3 shows

---

[26] Data taken from Eurostat – official statistical office of EU.

estimated β coefficients and standard errors, where gender dummy is called female (0 is for male, 1 is for female). As mentioned above, the smallest gap is in the youngest age group due to lack of experience and minimal education. Jobs worked by youth are often physical and do not require any specific sets of skills. Furthermore, the difference between men and women earnings increases by age until 30-39 and stabilises after that.

4.4. Gender wage gap with controlling age and education

$$log(w_{i,j,k}) = \beta_{1,i} female_i + \beta_{2,j} age_j + \beta_{3,k} education_k +$$
$$\beta_{2,j} age_j \times I(i=1) + \beta_{3,k} education_k \times I(i=1) + \epsilon_{i,j,k} \quad (5)$$

Secondly, in this part control variables of age and education are included. Model is explained by Equation 5. Education is split into two parts: "primary or secondary education" and "tertiary degree" to evaluate how gender wage gap variates over the life-cycle between people

Figure 12

Evolution of the gender wage gap by age and education
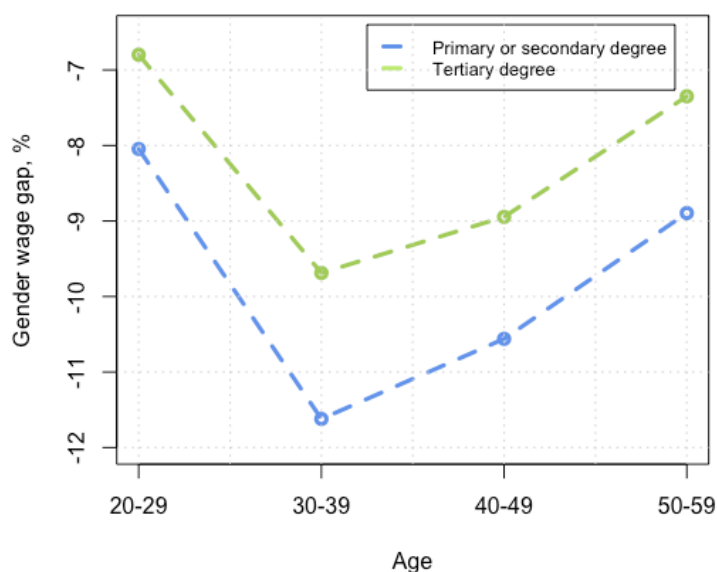
Source: made by author by using Structure of Earnings survey 2014 data

with different education levels. Figure 12 shows the comparison of highly educated population against low educated population. Log-percent differences show that the lowest point of the gap is within the same age groups 30-39. However, wage gap is smaller for higher educated individuals by more than 1%. Receiving a tertiary degree at the ages of 14-19 is not possible in Lithuania. As a result, these estimations for are not shown in Figure 12.

Table 4 of multiple linear regression shows the estimated coefficients for logged hourly earnings for females and males. By comparing multiple and adjusted R^2 from table 2 and table 3,

the overall model improves as the control variable of education is added. Table 4 shows similar trends as the previous Table 3 only showing the control variable of age.

In the new model, the smallest wage gap is also within the youngest 14-19 group age. The wage gap with controlling age and education variable changes in a similar pattern as the previous model, with 30-39 age group being the largest. However, compared to the first multiple linear regression model, these results indicate that wage gap stabilises more rapidly after the fertility age. There is a bigger decrease of the gap in the age group 40-49. In addition, within the 50-59 age group, the wage gap decreases and becomes similar to 20-29 age group. However, the discrepancy still remains lower.

Table 4

Multiple linear regression results 2

Predictors	Estimate	Standard error
female 0	$2.131 * **$	0.039
female 1	$2.049 * **$	0.058
age 20-29	$0.205 * **$	0.040
age 30-39	$0.290 * **$	0.040
age 40-49	$0.304 * **$	0.040
age 50-59	$0.282 * **$	0.040
education 2	$0.298'$	0.172
female 1 : education 2	$-0.018'$	0.011
female 1 : age 20-29	-0.091	0.070
female 1 : age 30-39	-0.170	0.070
female 1 : age 40-49	-0.150	0.070
female 1 : age 50-59	-0.115	0.070
age 20-29 : education 2	0.020	0.172
age 30-39 : education 2	$0.220*$	0.172
age 40-49 : education 2	$0.184*$	0.172
age 50-59 : education 2	0.186	0.172
multiple R^2	0.9659	
adjusted R^2	0.9659	

Source: made by author by using Structure of Earnings survey 2014 data

4.5. Gender wage gap controlling age, education and other factors

$$log(w_{i,j,k,z,r,p}) = \beta_{1,i} female_i + \beta_{2,j} age_j + \beta_{3,k} education_k + \beta_{4,z} companysize_z$$
$$+ \beta_{5,r} sector_r + \beta_{6,p} contract_p + \beta_{2,j} age_j \times I(i=1) + \beta_{3,k} education_j \times I(i=1) +$$
$$\beta_{4,z} companysize_z \times I(i=1) + \beta_{5,r} sector_r \times I(i=1) + \beta_{6,p} contract_p \times I(i=1) + \epsilon_{i,j,k,z,r,p}$$
$$(6)$$

Further regression analysis evaluates the gender wage gap with controlling for age, education, company size, industry sector and contract type. Model is explained by Equation 6. Company size variable is split into three groups: 0, 1 and 2 which represent the number of workers

within a company of 1-49, 50-249, and 250+ employees respectively. Sector variable contains two groups: public (A) and private (B). Lastly, contract type describes the category of employment agreement between the worker and the firm. This variable has three groups: not-terminated, terminated and apprenticeship. The last group was eliminated from the regression analysis as there were no apprenticeship contracts among the surveyed workers. Sector variable is described as not-terminated (A) and terminated (B).

Figure 13 illustrates β coefficients of multiple linear regression. The x-axis represents utilized variables whereas the y-axis shows the corresponding β coefficients. Predictor variable is logged hourly wage. The first two coefficients show a low wage gap of 2 log points between men and women within the 14-19 age group which received primary or secondary education, work in a small company in a public sector. As illustrated in Figure 13, the wage gap increases with age. Values "female1:age" indicate how the age of women directly impact on their wages. Seeing as they are all negative, it means that age has a negative effect on earnings. As shown previously, the

Figure 13

Multiple linear regression estimates
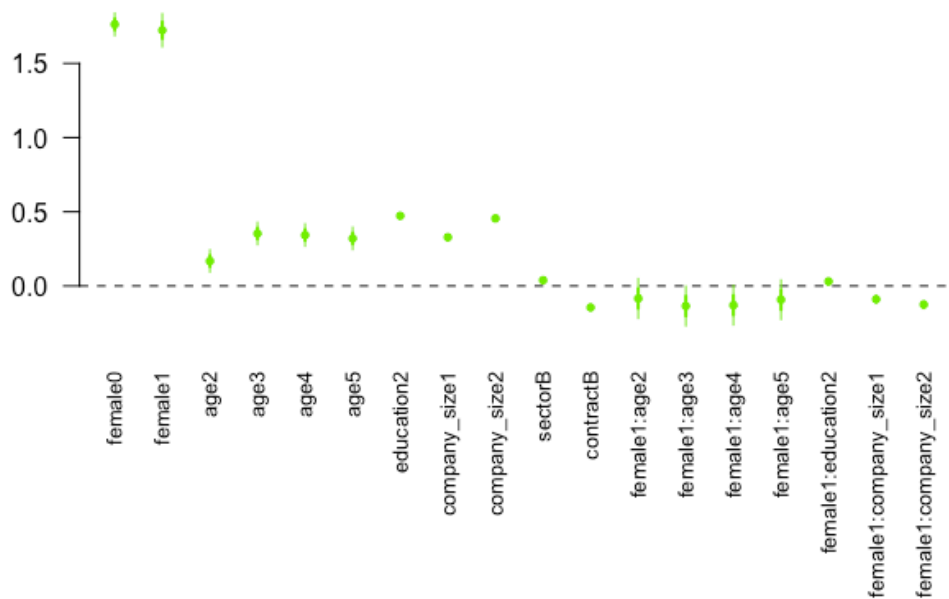
Source: made by author by using Structure of Earnings survey 2014 data

pay gap is the biggest within the age of 30-39. However, the pay gap minimises to less than 9% (log-percentage) meaning that the majority of the wage gap can be explained by education, company size, and employment agreement type. Furthermore, the data shows that the hourly wage increases sharply depending on the size of the company and also shows slight increases within the private sector. However, it is reported that terminated job agreements have a negative effect on earnings.

Overall, adding more control variables to the regression analysis lowers the gender pay gap. This is an important finding in helping understand the different drivers involved in pay gap. By comparing the unadjusted and adjusted pay gap, similar trend over the life cycle is determined. Therefore, adding more control variables minimises the pay gap and partially explains the main drivers. This model allows to find the wage gap between any desirable group of individuals and compare it.

4.6. Comparison with 2018 data

The data from the 2018 Structure of earnings survey has it's own drawbacks. Merging the 14-19 and 20-29 age groups into one makes evaluation of wage gap over the life-cycle more complex. The difference between these two age groups is significant within 2014 SES data, therefore evaluation of this particular data set could be advantageous for the research. However, it is still possible to observe the change over the life cycle.

The wage gap among full-time workers was calculated using the same methods used for the 2014 dataset and showed a 13.9% (calculated possible error is 0.2%) difference. The result is 2.1% lower compared to 2018, implying that the wage gap could be decreasing throughout the years.

When adding a control variable of age, the results indicate that pay gap evolution over age groups shows a similar trend to that of the 2014 data. Figure 14 illustrates log-percent differences of pay gap based on age, while the. shaded area shows the standard error band. Overall, the graph looks similar, however it is important to note that the 14-19 and 20-29 age groups were merged

Figure 14

Evolution of the wage gap over the life cycle, 2018 data
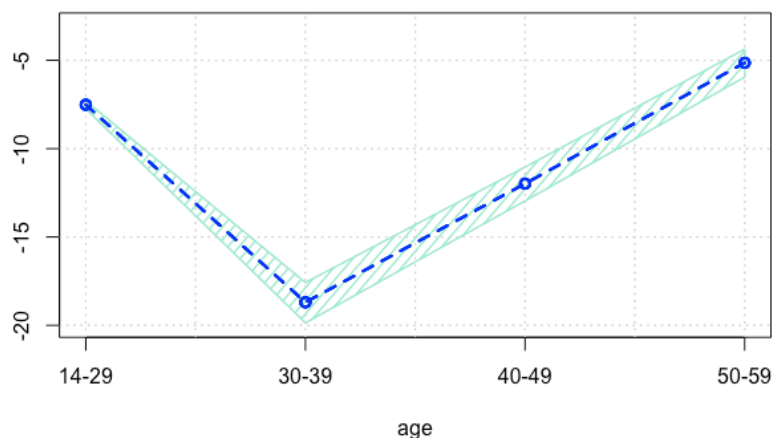
Source: made by author by using Structure of Earnings survey 2014 data

together. This could be the cause of higher log-percent differences in 14-29 age group. When regression with 2014 data was performed, the smallest wage gap was among the 14-19 age group, therefore combining the data with 20-29 group could produce a higher difference.

Additionally, statistical comparisons reveals that the standard errors of 2018 multiple linear regression are significantly lower compared to 2014. The shaded area is significantly larger in the 2014 wage gap graph, which highlights the precision of the 2018 data.

Results in Figure 14 are more accurate than Figure 11. However, the analyses of the results cannot be done due to inconsistencies in age factor. There was a hypothesis that it is because of the greater number of survey observations, but it was rejected as 2014 data has even more observations than 2018 data. A possible explanation is that 2018 data is more consistent and has less "NaN" (missing) values.

4.7. Occupational segregation

Horizontal segregation is one of the main aspects of gender differentials in labour force. It implies that labour market is divided into two groups – men and women – that work in different types of occupation. Table 5 is constructed by filtering the data of full-time workers and investigating most frequent occupations among both genders with primary or secondary education. Results indicate, that professions chosen by men and women are completely different. For example, among the youngest group where workers have no education and experience, men tend to work jobs that require physical force. Packers, storage or construction workers are common

Table 5

Most frequent occupations of respondents with primary or secondary education

Age group	Sex	Occupation
14-19		Manufactoring labourer
20-29		Building frame
30-39	Male	Heavy truck and bus driver
40-49		Heavy truck and bus driver
50-59		Heavy truck and bus driver
14-19		Shop salesperson
20-29		Shop salesperson
30-39	Female	Shop salesperson
40-49		Nursing or midwifery
50-59		Cleaners

Source: made by author by using Structure of Earnings survey 2014 data

occupations for young men. For women most common occupations at young age are shop salespersons or food related workers.

Furthermore, most frequent occupations change with age and depend on education. Men with no tertiary degree tend to choose professions such as building frame workers, truck drivers and protective services. However, women without education choose less risky and not requiring force jobs, such as salesperson, tailor or food factoring, hand packers. On average, these professions are lower-paid in comparison with all kind of drivers, house building and maintenance workers. Physical advantage of men gives them an opportunity to work higher paid jobs without requiring education while women of young age select to work in customer service. Another thing worth mentioning from the Table 5 that older women tend to work in domestic, hotel or office cleaning as well as nursing and elderly care. This could be the effect of the demand for especially young workers in customer service.

Table 6 compares occupations of men and women with tertiary degree. Males mostly work in IT or engineering sectors, which are known for relatively large compensations. In addition, women with tertiary degree choose completely opposite jobs than men. For example, accounting, administrative services, sales and marketing. The results imply that most common professions are ascribed to either feminine or masculine side. However, occupations that require tertiary degree have no need in physical force. Therefore, it suggests the intellectual and psychological difference between genders. As Table 6 illustrates, women tend to be more interested in social science, while men lean to mathematics or technologies. Another explanations for such radical differences is consumer prejudice. Some women doing "masculine" jobs can be seen as unprofessional and too weak to do required tasks properly. The same theory works for "feminine" jobs such as sewing, nursing or cleaning.

Table 6

Most frequent occupation of respondents with tertiary degree

Age group	Sex	Occupation
20-29		Software developers
30-39		Engineering and manufacturing managers
40-49	Male	Managing directors and chief executives
50-59		Managing directors and chief executives
20-29		Administration professionals
30-39		Administration professionals
40-49	Female	Finance professionals
50-59		Administration professionals

Source: made by author by using Structure of Earnings survey 2014 data

Tendency is clear, women without and with tertiary degree select less paid jobs due to a variety of possible reasons. As a results, it makes a huge impact on the wage gap. Generally, women are badly rewarded than men. However, focus on occupation shows that it partly depends on intellectual and psychological characteristics accompanied by personal choices and customer prejudice. The occupation impact on gender wage gap in Lithuania is undeniable.

4.8. Working-time

$$w_{i,f} = \beta_{1,i} female_i + \beta_{2,f} fulltime + \beta_{2,f} fulltime \times I(i = 1) + \epsilon_{i,f} \qquad (7)$$

As it was presented with past findings, working time has a significant impact on the wage gap. Performing the analysis among part-time workers shows some interesting results. As it was expected, women form a bigger part of part-time workers than men. 60% out of 6684 part-time workers are female. Equation 7 explains the multiple linear regression with controlling full-time variable. The results show that there is a 16% gap between full-time workers and respectively small wage gap of 7% between part-time workers. However, this confirms the decision of eliminating the part-time workers from the general model due to inconsistency. Part-time jobs can

Table 7

Multiple linear regression estimates

Predictor	Estimate	Standard error
female 0	15.74334	0.08578
female 1	13.21376	0.08945
full time 2	−2.30229	0.23807
female 1: fulltime 2	1.58711	0.31220

Source: made by author by using Structure of Earnings survey 2014 data

offer higher compensation while people will still earn less in a month due to limited working hours. Consequently, part-time workers accept the wage differential, since advantage of having more time for housework or childcare outweighs it. Usually in Lithuanian market part-time job is chosen by personal preferences, not because of failure in finding a full-time job. Furthermore, the most common occupations among part-time workers are cleaners and sales agents, which are known for being the source of additional money or a short-term option.

4.9. Pay gap drivers

Particular trends of wage gap over the life cycle explain the origin of inequality in the specific age group. At the age of 14-19 the wage gap is more than 7 times smaller in comparison

with 30-39 age group. This indicates that gender based discrimination is not as relevant as it is in other European countries. The wage gap between young people without tertiary degree or experience can be explained by biological advantages and disadvantages. Occupational segregation explains this part of the trend as young women choose jobs which are safer and do not require physical strength. While men take advantage of their physical power and choose better paid occupations. Furthermore, following age groups experience more inequality as the wages of women do not grow at the same rate as the wages of men.

The biggest wage gap is determined in group of 30-39 years. Such a big difference is explained by child penalty. Wage gap drops rapidly after crossing the average fertility age of a woman. Child penalty has the biggest effect on inequality. Slow human capital accumulation and interrupted career continue the slow the growth of female wage. This penalty is valid for up to 20 years after the birth of the first child. Child penalty in Lithuania explains a large part of the remaining gender inequality. Because of the interruption of the child birth, women may not develop specific skills and professional networks that could determine the wage growth. Productivity reduction after having a child results in lower professional performance. A larger share of house work and childcare impedes women to be more successful and hard-working.

The addition of more control variables reduces the wage gap and explains some parts of it. Inequality as well depends on education, company size. Public and private sectors have different trends of the wage gap. In addition, terminated job contracts determine a lower compensation.

4.10.   Ridge and lasso regressions

To evaluate the performance of multiple linear regression, model with controlling age, education and other factors (company size, sector and contract) is performed by using ridge and lasso regression. These models also allow simple calculations of the wage gap between particular groups of individuals by using β coefficients. β coefficients are estimated by using the same Equation 6 as it was done in multiple linear regression method.

$$
\begin{aligned}
log(w_{i,j,k,z,r,p}) &= \beta_{1,i}female_i + \beta_{2,j}age_j + \beta_{3,k}education_k + \beta_{4,z}companysize_z \\
&+ \beta_{5,r}sector_r + \beta_{6,p}contract_p + \beta_{2,j}age_j \times I(i=1) + \beta_{3,k}education_j \times I(i=1) + \\
&\beta_{4,z}companysize_z \times I(i=1) + \beta_{5,r}sector_r \times I(i=1) + \beta_{6,p}contract_p \times I(i=1) + \epsilon_{i,j,k,z,r,p}
\end{aligned}
$$
$$(6)$$

In comparison, multiple linear regression results have shown that wage gap between the youngest age group of 14-19 (with lower education level, working in private sector, small company and not-terminated work contract) is 2% (log-percentage). However, according to ridge regression estimates, wage gap between the same group of men and women is 1.2% (log-

percentage). Lasso regression results imply that the same wage gap is even smaller than 1% (log-percentage). Overall, the results from these three different regression models are not the same, however, they not differ by much. Further ridge and lasso regression results are presented in Appendix 1.

This example explains why wage gap calculations differ by every research. Chosen method and model has a large impact on the results. The presence of many alternative ways of measuring the wage gap form a point of view that calculations of it are unreliable. However, the illustration of evolution of the gap is independent from the way it is measured.

In this case, the main difference between lasso and ridge regressions is that ridge shrinks close to zero so that all predictor variables can be kept in the model. Whereas lasso shrinks coefficients to exactly zero, which means that some of the predictor variables are eliminated from the model. For example, according to the Table 1 in Appendix 1, coefficients female 0 and female 1 are equal to zero. Complete reduction of unnecessary parameters is an advantage.

To find out which model is the best approach for determining the wage gap, root mean square error (RMSE) and R^2 should be calculated for all models. The lower the standard deviation of the residuals, the more accurate model results are. R-squared explains the proportion of the model explained by independent variables. With increasing value of R^2 greater part of the model is explained. To calculate RMSE and R^2 of MLR, lasso and ridge regressions, the data is split into train and test parts. Table 8 represents the results from the evaluation of models. RMSE of ridge and lasso regressions is equivalent in both train and test data. Root mean square error of multiple linear regression is more than three times higher than RMSE of ridge and lasso. This implies that residuals of MLR are more spread out than residuals of lasso and ridge regressions. In addition, results show that R^2 is higher for multiple regression, which means that predictors in multiple regression explain a larger part of the model than predictors in lasso and ridge regression. Substantially, ridge and lasso regression models have more accurate results as R^2 is similar to MLR. However, the difference in root mean square error is large. The spread of prediction errors is a significant factor for predicting the dependent variable (hourly wage). In conclusion, multiple regression is not the best approach for the determination of the gender wage gap. Results suggest that lasso and ridge models are equally accurate.

Therefore, lasso regression is a better approach for evaluation of the wage gap. The advantage of elimination of the predictors that are less significant leads to less difficult calculations. As used dataset has less predictors than observations, all relevant predictors are used in the model. Cross-validation method for choosing the penalty factors assures the success of the model on future data samples. Lasso performs parameter shrinkage by shrinking coefficients to zero and automatically selects variables. To conclude, every discussed method has it's own

advantages and disadvantages. Although, this study claims that the most accurate way to evaluate the gender pay gap is lasso regression. Multiple linear regression has an easier approach for calculating the wage gap, but not evaluating it.

Table 8

RMSE and R² comparison

Type of regression	Data	RMSE	R^2
MLR	Train	1.69	0.3
	Test	1.69	0.3
Ridge	Train	0.45	0.29
	Test	0.45	0.28
Lasso	Train	0.45	0.29
	Test	0.45	0.28

Source: made by author by using Structure of Earnings survey 2014 data

5. Conclusion

The findings of this study confirm the significance of the wage gap. The recent growth in demand of equal opportunities for men and women requires more specific explanations of the wage gap. Lithuanian market lacks the statistics and researches on gender gap. Therefore, analysing it and evaluating possible reasons is the key for overcoming the problem. Multiple linear regression analysis is a common way for estimating and examining the wage gap and it's drivers. Using microdata from Structure of Earning survey leads to successful results.

Calculated unadjusted wage gap in 2014 is 16%. It is approximately 2% higher than the EU average. The result is similar to official calculations, but differs by approximately 1%. It can be caused as different approaches of calculations are used and also depends on the analysed data. Furthermore, this study claims that the traditional explanations continue to influence the wage gap. Although, some factors' impact have increased or decreased. One of the findings is that discrimination forms only a small part of gender gap. At the ages of 14-19 estimated gap shows the low impact of gender based discrimination. The wage gap between young people without tertiary degree or experience can be explained by biological advantages and disadvantages. Occupational segregation explains this part of the gender gap. Psychological and intellectual differences between men and women is also one of the wage gap indicators. Furthermore, slow human capital accumulation after giving birth is observed. Wage gap drops after the average age of fertility. This phenomenon is presented as child penalty on women and forms the majority of wage gap in Lithuania. Interrupted education and career determine slow human capital accumulation. Furthermore, added control variables reduce the wage gap and partly explain some of it. Inequality is partly driven by education, company size, sector and type of contract. The comparison with 2018 data makes it clear that the wage gap remains to have the same trend throughout the life cycle. However, after the fertility age it stabilises more rapidly and reaches a smaller wage gap in years of 50-59, and has a smaller wage gap than the youngest age group. The main difference is that overall wage gap is smaller by 2% than in 2014 meaning that inequality in Lithuanian labour market is reducing. In addition, evaluation of lasso and ridge regression approach for this problem assists in finding the best fit for the problem. Results imply that multiple linear regression is a better approach in calculating the pay gap, whereas lasso regression is better for evaluating it.

The results imply that more statistical analysis and research should be done on child penalty. The absence of literature about child penalty phenomenon in Lithuania has a negative effect on the economy. Ensuring equal pay for both men and women has to start with collecting and analysing the data. Indicating the main drivers of the pay gap would lead to a better decisions in introducing new policies. Some suggestions for Lithuanian government and Statistics

Lithuania: 1) To continue estimation and evaluation of gender pay gap and it's evolution over the life cycle; 2) To improve data collection and analysis. For example, data with information about first child birth, household details and hourly or monthly wages would lead to more accurate and detailed results; 3) To improve the parental compensation system. Current policy is focused on monetary benefits and does not encourage both parents to involve. To ensure equal approach the same conditions should be applied for maternal and paternal leave; 4) To make statistics and evaluations available and properly presented to the public.

SUMMARY

In the past several decades gender inequality in labour market have played a significant role. This problem attracts a lot of attention due to recent demand growth in equal opportunities for all genders. Recent research have proven the complexity of the wage gap. Active research on the gender pay gap results in significant discoveries. Selected literature illustrates the main drivers that affect the differentials. Analysis of the wage gap over the life cycle evaluates the impact of well-known drivers. Particular trends illustrate the main pay gap causes. Overcoming gender inequality problem in labour market requires statistical analysis to find the best policies for the government. Equal career opportunities for women are not a privilege, but a necessity. The factors that make impact on gender inequality in Lithuanian labour market are complex. Investigating hourly wage differences of men and women by using multiple linear regression leads to a better understanding of the situation in Lithuania. In addition, evaluation of lasso and ridge regression approach for this problem assists in finding the best fit for the problem. Results imply that multiple linear regression is a better approach in calculating the pay gap, whereas lasso regression is better for evaluating it. Calculated unadjusted wage gap in 2014 is 16%. It is approximately 2% higher than the EU average. The result is similar to calculations of Statistical office of European Union. Different methods, calculations and selected data are the main causes of the difference in results. This study claims that traditional explanations resume the influence on the wage gap. Therefore, findings confirm the differences in wage gap drivers by age groups. The youngest age group of 14-19 confirms the occupational segregation. Wage gap between young women and men without education and experience is driven by physical gender advantages and disadvantages. Results illustrate that the largest pay gap is within the 30-39 age group and is driven by child penalty. Child penalty explains the biggest part of the wage gap in Lithuania. In conclusion, this research confirms that traditional pay gap drivers remain the same. However, evaluation of the gender gap by age illustrates the distribution of the main pay gap drivers.

SANTRAUKA

Pastaraisiais dešimtmečiais lyčių nelygybė darbo rinkoje atlieka reikšmingą vaidmenį. Ši problema atkreipia ekonomistų dėmesį dėl augančio visuomenės susidomėjimo moterų teisėmis. Įvykdyti tyrimai patvirtina, kad darbo užmokesčio problema yra sudėtinga ir kompleksiška. Aktyvūs tyrimai lemia reikšmingus atradimus darbo ekonomikoje. Įdėmiai atrinkta literatūra iliustruoja pagrindinius darbo užmokesčio skirtumo veiksnius. Darbo užmokesčio skirtumo pokyčių per gyvenimo ciklą analizė įvertina kokią įtaką daro jau žinomi veiksniai. Tam tikra tendencija parodo, kas daro įtaką užmokesčio skirtumui skirtingose amžiaus grupėse. Statistinė analizė yra būtina norint įveikti šią problemą ir pagerinti taikomas strategijas. Lygios galimybės moterims darbo rinkoje yra ne privilegija, o būtinybė. Faktoriai, lemiantys lyčių nelygybę, yra sudaryti iš daugybės veiksnių. Valandinio darbo užmokesčio skirtumo tyrimas naudojant daugialypę tiesinę regresiją padeda įvertinti situaciją Lietuvoje. Įvertinimas kuo skiriasi lasso ir ridge regresija asistuoja ieškant geriausio tyrimo būdo pagrindinei problemai. Remiantis rezultatais, daugialype tiesinė regresija yra vienas geriausias būdas skaičiuoti darbo užmokesčio skirtumą, o lasso regresija – įvertinti ir palyginti. Apskaičiuotas nekoreguotas darbo užmokesčio skirtumas 2014 metais yra 16%. Tai yra 2% aukščiau Europos Sąjungos vidurkio. Rezultatas sutampa su Europos Sąjungos statistikos departamento skaičiavimais. Skirtingi metodai, skaičiavimai ir pasirinkti duomenys sudaro rezultatų skirtumą tarp tyrimų. Rezultatai rodo, jog seniau žinomi, tradiciniai faktoriai tęsia savo įtaką užmokesčio skirtumui. Tačiau, šie faktoriai skirtingai veikia tarp amžiaus grupių. Jauniausioji amžiaus grupė patvirtina profesinę segregaciją. Darbo užmokesčio skirtumas tarp jaunų moterų ir vyrų, kurie yra nepatyrę ir neturi išsilavinimo, yra minimalus. Jį sudaro fizinės būklės pranašumai ir trūkumai. Remiantis tolimesniais rezultatais, didžiausias darbo užmokesčio skirtumas yra 30-39 amžiaus grupėje. Moterys, kurios susilaukia vaikų, dažniausiai uždirba mažiau. Motinystė paaiškina didžiąją dalį lyčių nelygybės Lietuvos darbo rinkoje. Apibendrinant galima teigti, kad tradiciniai darbo užmokesčio veiksniai išlieka tie patys, tačiau skirtingai pasiskirsto ir paaiškina skirtumą tarp amžiaus grupių.

References

Angelov, N., Johansson, P. & Lindahl, E. J. (2016). Labor Econ. 34, 545–579

Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of labor economics*, *34*(3), 545-579.

Anker, R. (1997). Theories of occupational segregation by sex: An overview. *Int'l Lab. Rev.*, *136*, 315.

Baker, M. (2010, May). Motherhood, employment and the "child penalty". In *Women's Studies International Forum* (Vol. 33, No. 3, pp. 215-224). Pergamon.

Bericat, E. (2012). The European gender equality index: Conceptual and analytical issues. *Social Indicators Research*, *108*(1), 1-28.

Bertrand, M. (2020, May). Gender in the twenty-first century. In *AEA Papers and proceedings* (Vol. 110, pp. 1-24).

Bettio, F., Verashchagina, A., Mairhuber, I., & Kanjuo-Mrčela, A. (2009). *Gender segregation in the labour market: Root causes, implications and policy responses in the EU*. Luxembourg: Publications Office of the European Union.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, *55*(3), 789-865.

Böheim, R., & Stöllinger, P. (2021). Decomposition of the gender wage gap using the LASSO estimator. *Applied Economics Letters*, *28*(10), 817-828.

Bøler, E. A., Javorcik, B., & Ulltveit-Moe, K. H. (2018). Working across time zones: Exporters and the gender wage gap. *Journal of International Economics*, *111*, 122-133.

Boll, C., Rossen, A., & Wolf, A. (2017). The EU gender earnings gap: Job segregation and working time as driving factors. *Jahrbücher für Nationalökonomie und Statistik*, *237*(5), 407-452.

Ciminelli, G., Schwellnus, C., & Stadler, B. (2021). Sticky floors or glass ceilings? The role of human capital, working time flexibility and discrimination in the gender wage gap.

Ciminelli, G., Schwellnus, C., & Stadler, B. (2021). Sticky floors or glass ceilings? The role of human capital, working time flexibility and discrimination in the gender wage gap.

Cotter, D. A., Hermsen, J. M., Ovadia, S., & Vanneman, R. (2001). The glass ceiling effect. *Social forces*, *80*(2), 655-681.

Cubas, G., Juhn, C., & Silos, P. (2019). *Coordinated work schedules and the gender wage gap* (No. w26548). National Bureau of Economic Research.

De Quinto, A., Hospido, L., & Sanz, C. (2020). The child penalty in Spain. *Banco de Espana Occasional Paper*, (2017).

Gallego-Granados, P. (2019). The part-time wage gap across the wage distribution.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, *104*(4), 1091-1119.

Hirsch, B., Oberfichtner, M., & Schnabel, C. (2014). The levelling effect of product market competition on gender wage discrimination. *IZA Journal of Labor Economics*, *3*(1), 1-14.

Hughes, J., & Maurer-Fazio, M. (2002). Effects of marriage, education and occupation on the female/male wage gap in China. *Pacific Economic Review*, *7*(1), 137-156.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *economics*, *7*(1), 63-88.

Kay, K., & Shipman, C. (2014). The confidence gap. *The Atlantic*, *14*(1), 1-18.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimuller, J. (2019, May). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings* (Vol. 109, pp. 122-26).

Kunze, A. (2018). The gender wage gap in developed countries. *The Oxford handbook of women and the economy*, 369-394.

Lorber, J. (2001). Gender inequality. *Los Angeles, CA: Roxbury*.

Madalozzo, R. (2010). Occupational segregation and the gender wage gap in Brazil: an empirical analysis. *Economia aplicada*, *14*(2), 147-168.

Meurs, D., & Ponthieux, S. (2005). The gender wage gap in Europe: women, men and the public sector. *WP F0502*.

Neumark, D., R. Bank & K. Van Nort (1996). "Sex Discrimination in Restaurant Hiring: An Audit Study", The Quarterly Journal of Economics, Vol. 111/3, pp. 915-941.

Nyhus, E. and E. Pons (2012), "Personality and the gender wage gap", Applied Economics, Vol. 44/1, pp. 105-118.

Petersen, T., & Morgan, L. A. (1995). Separate and unequal: Occupation-establishment sex segregation and the gender wage gap. *American Journal of Sociology*, *101*(2), 329-365.

Pettinger, T. (2017). The economics of discrimination. Oxford, UK.

Pilinkaitė-Sotirovič, V. Kontvainė, V. (2020). Modern men and gender equality: incentives and obstacles of participation in childcare.

Polachek, S. W. (1987). Occupational segregation and the gender wage gap. *Population Research and Policy Review*, *6*(1), 47-67.

Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, *111*(12), 4403-4408.

Russo, G., & Hassink, W. (2008). The part-time wage gap: A career perspective. *De Economist*, *156*(2), 145-174.

Schubert, R., Brown, M., Gysler, M., & Brachinger, H. W. (1999). Financial decision-making: are women really more risk-averse?. *American economic review*, *89*(2), 381-385.

Semykina, A. and S. Linz (2007), "Gender differences in personality and earnings: Evidence from Russia", Journal of Economic Psychology, Vol. 28/3, pp. 387-410.

Vaccaro, G. (2017), Can Policy Directly Reduce the Unexplained Gender Wage Gap? Evidence from Switzerland.

Weichselbaumer, D., & Winter-Ebmer, R. (2005). A meta-analysis of the international gender wage gap. *Journal of economic surveys*, *19*(3), 479-511.

Appendix 1

Table 1

Lasso regression results

(Intercept)	2.668240e+00
female.0	.
female.1	.
age.1	-2.138440e-01
age.2	-1.243936e-01
age.3	.
age.4	.
age.5	9.864715e-03
education.1	-3.704695e-01
education.2	1.373416e-11
company_size.0	-3.303707e-01
company_size.1	.
company_size.2	.
sector.A	-2.793384e-02
sector.B	.
contract.A	1.350357e-01
contract.B	.
female0:age1	-5.046695e-02
female1:age1	.
female0:age2	-4.955381e-03
female1:age2	.
female0:age3	5.787260e-02
female1:age3	.
female0:age4	4.698168e-02
female1:age4	-1.384438e-03
female0:age5	.
female1:age5	.
female0:education1	.
female1:education1	-1.257209e-01
female0:education2	9.496741e-02
female1:education2	.
female0:company_size0	-3.878308e-03
female1:company_size0	.
female0:company_size1	.
female1:company_size1	-9.165802e-02
female0:company_size2	1.214821e-01
female1:company_size2	.

Source: Made by author

Table 2

Ridge regression results

(Intercept)	2.4569831319
female.0	0.0300253926
female.1	-0.0301236662
age.1	-0.1045585304
age.2	-0.0468840404
age.3	0.0260731221
age.4	0.0137205965
age.5	0.0065933124
education.1	-0.1211913895
education.2	0.1214895879
company_size.0	-0.1131782656
company_size.1	0.0126506779
company_size.2	0.0641058229
sector.A	0.0009363117
sector.B	-0.0009800100
contract.A	0.0565267958
contract.B	-0.0564983924
female0:age1	-0.1095473863
female1:age1	-0.0913669139
female0:age2	-0.0371929164
female1:age2	-0.0464342560
female0:age3	0.0500815251
female1:age3	-0.0097391646
female0:age4	0.0408073143
female1:age4	-0.0161863397
female0:age5	0.0169394824
female1:age5	-0.0048689503
female0:education1	-0.0491007606
female1:education1	-0.1138629840
female0:education2	0.1103595045
female1:education2	0.0629126196
female0:company_size0	-0.1108925546
female1:company_size0	-0.0871287862
female0:company_size1	0.0330428828
female1:company_size1	-0.0158233048
female0:company_size2	0.0793025057
female1:company_size2	0.0085062951

Source: Made by author