CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Bachelor's Thesis

2022 **Martina Bartóková**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Decomposition of the gender wage gap before and after COVID-19: did wage determinants across genders change?

Bachelor's Thesis

Author of the Thesis: Martina Bartóková

Study programme: Economics and Finance

Supervisor: Mgr. Barbara Pertold-Gebicka M.A., Ph.D.

Year of the defence: 2022

Declaration

1. I hereby declare that I have compiled this thesis using the listed literature and resources only.
2. I hereby declare that my thesis has not been used to gain any other academic title.
3. I fully agree to my work being used for study and scientific purposes.

In Prague on 31. 7. 2022 Martina Bartóková

References

BARTÓKOVÁ, Martina. *Decomposition of the gender wage gap before and after COVID-19: did wage determinants across genders change?* Praha, 2022. Bachelor's thesis (Bc). Charles University, Faculty of Social Sciences, Institute of Economic Studies. Supervisor Mgr. Barbara Pertold-Gebicka, M.A., Ph.D.

Length of the Thesis: 75 380

Abstract

The topics of inequality, unconditional differences or disproportional rewarding systems are fighting for their place at the top of the interest in public life. In order to fight the inequality present in society, scientists, policymakers, publicly active people and ordinary people are trying to discuss it and find its leading causes. This thesis examines the inequality between men and women in the rewarding system and how the recent pandemic crisis influences it. We analyse the role of many wage gap determinants and try to identify the main ones and how they change with the COVID-19 pandemic. Some of our expectations were confirmed. The pandemic crisis put more burden on women than men. However, the research shows that all the economic sectors suffer from the pandemic; the only difference is by what amount. The results of the decomposition show that change in the determinants of the wage gap is present, but only on the border of the gap. The reason is that even before the pandemic outbreak, the most significant part of the gender wage gap (GWG) was caused by the unexplained gap, i.e., the possible reasons for the GWG are gender discrimination or gender specialisation. After the COVID-19 pandemic began, the unexplained gap strengthened its position.

Keywords

gender, gender wage gap, COVID-19, shecession, change, effect, pandemic, decomposition

Title

Decomposition of the gender wage gap before and after COVID-19: did wage determinants across genders change?

Abstrakt

Témy týkajúce sa nerovnosti, neopodstatnených rozdielov alebo nerovnomerného odmeňovania sa stále viac snažia pretlačiť svoje miesto na vrchol spoločenských tém. Vedci, zákonodarcovia, verejne aktívne osoby ako aj bežní ľudia sa s cieľom bojovať proti všadeprítomnej nerovnosti snažia aktívne diskutovať a hľadať jej príčiny. Predmetom tejto práce je skúmať mzdové rozdiely medzi mužmi a ženami a aký vplyv na tieto rozdiely mala nedávna pandémia koronavírusu. Analyzovali sme vplyv mnohých determinantov na mzdy a snažili sme sa identifikovať tie z nich, s najväčším vplyvom. Ďalej sme sa pozreli ako pandémia ovplyvnila postavenie týchto determinantov. Niektoré z našich očakávaní sa potvrdili. Pandémia zhoršila situáciu žien viac než mužov. Napriek tomu, že výsledky výskumu ukázali, že pandémia ovplyvnila všetky sektory ekonomiky, niektoré z týchto sektorov boli zasiahnuté viac ako iné. Výsledky dekompozície miezd zamestnancov ukázali, že zmena v determinantoch mzdových rozdielov je prítomná len okrajovo. Hlavná časť, ktorá zapríčiňovala mzdové rozdiely medzi mužmi a ženami aj pred pandémiou, je časť, ktorú nevieme vysvetliť pomocou rozdielnych charakteristík mužov a žien. Táto časť môže byť spôsobená diskrimináciou alebo špecializáciou jednotlivých pohlaví. Vďaka pandémii táto časť ešte viac zosilnela.

Klíčová slova

pohlavie, mzdové rodové rozdiely, COVID-19, shecession, zmena, efekt, pandémia, dekompozícia

Název práce

Dekompozícia rodových mzdových rozdielov pred a po COVID-19: zmenili sa mzdové determinanty medzi pohlaviami?

Acknowledgement

I would like to express my gratitude to my supervisor, Mgr. Barbara Pertold-Gebicka M.A., Ph.D., for her professional guidance, valuable suggestions, and patience through the whole process of writing this thesis. I would like to thank my family and friends for their support and trust. Their advice and support help me to finish this thesis. This thesis was supported by the NPO "Systemic Risk Institute" LX22NPO5101.

Table of Contents

Introduction

Over the past several decades, women's role has been heavily discussed. Whether we look at policymakers or average citizens, everyone knows the word "feminism." Women are trying to equalize their roles in public and personal life (Bertrand, 2018). They are fighting for equal rights, equal division of duties, equality in relationships, and equal remuneration in work life. This fight was paralyzed by the uprising of COVID-19 disease, or wasn't it?

With increasing number of COVID-19 cases, many states worldwide considered various pandemic restrictions. Many services had to cancel their work; where possible, jobs were transferred into the online world. Schools, administrative jobs, and many others lost face-to-face systems and moved to the world of computers and internet connections (Alon et al., 2020). In many households, families, who were used to spending time together only in the evenings or weekends, had to change it all at one moment. All of this brings attention to the role of women – they had to be mothers and employees simultaneously, and their primary responsibilities started to overlap. It leads to the question, how has the COVID-19 pandemic changed the role of women in the labour market? (Alon et al., 2020; Fortier, 2020; Bonacini et al., 2021))

It is well known that women were discriminated against in the labour market even before the pandemic started. They earn less than men, they are expected to work shorter hours than men, and their role as mothers (or future mothers) plays an essential role in women's employment (Becker, 1964; Bertrand, 2018; Blau and Kahn, 2017; Fortier, 2020).

The existing literature points to the historically painted role of women as mothers and their need to equalize their position in the labour market with men (Blau and Kahn, 2017). All these things are referred to as gender discrimination (Becker, 1964; Bertrand, 2018). They are causing a different view of men and women, which in the work environment leads to a difference in the rewarding of genders, i.e., it leads to the gender wage gap (Becker, 1964; Blau and Kahn, 2017; Bertrand, 2018). The focus of this thesis will be on the case of the USA, mainly due to their availability of data and a good comparison with previous studies and how other world crises have influenced the gender wage gap there. For a better picture, I will present four cross-section datasets for

years from 2018 to 2021. They should paint a picture of the pre-pandemic situation and what the COVID-19 crisis caused to the labour market and women's role.

The effect of the COVID-19 pandemic on the gender wage gap is negative. The pandemic crisis and the restrictions applied stress the women's role as mothers and caregivers and cause many women to leave the labour market. Due to this, the crisis is referred to as "shecession".

The following pages are organized into four chapters – Literature review, Theoretical framework, Methodology and Data, and Results. In the Literature review, I will provide a brief overview of the current state of research on the gender wage gap and the evolution of women's role in society. Next, the theoretical background will look at the various definitions of concepts discussing the gender wage gap, and together with the theory, it will propose my research questions. The chapter discussing Methodology and Data will provide a reader with the concept of measuring the wage gap across various groups and will introduce my datasets. In the last chapter, I will present the results of my research and give a possible explanation. In the end, I will summarize my findings and the main contribution to the field of research.

1. Literature review

The topic of gender discrimination has been under severe discussion for an extended time nowadays. There can be found many sources for different types of gender discrimination. From all sorts, this thesis will focus on wage discrimination among women and the way how the COVID-19 pandemic has influenced it in the US labour market. The definition of the gender wage gap and what portion of it is due to discrimination has changed over time. When the research began in the middle of the 20th century, the idea of the gender wage gap was a difference in the income of equally productive workers (Becker, 1964). The gender wage gap was presented as wage discrimination primarily by labour economists observing labour market structure (Navarová, 2022). As time passed, the view of what GWG is has changed. Nowadays, the gender wage gap is understood as a difference in the average income of equally endowed people (i.e., people with the same human capital characteristic) caused by externalities (i.e., labour market structure, time preferences of employees, discriminatory point of view of employers, risk aversion of women). When we talk about the raw or unadjusted gender wage gap, we refer to one without any other controlling variables. The GWG estimated using a model with control variables for human capital specification, personal characteristics, or variables describing the working condition is called the adjusted gender wage gap (Babcock and Laschever, 2003; Blau and Kahn, 2017; Bertrand, 2018).

The argument with labour market structure is still present and can be supported by the data, e.g., the data by OECD from 2019, which shows that 82.6% of women work in the service sector, while for men, it is just 60.2%. In comparison, 33.7% of men worked in the production sector in 2019, and, on the other hand, it was only 13.7% of women (OECD, 2019).

However, in the past years, there were significant converging steps made. Not only in human capital characteristics, where women overcome men, but also in sector-specific jobs (Goldin, 2014). Like Goldin specified in her research paper from 2014, there does not need to be government intervention to change the gender pay difference, but what is

needed much more is a change in the mindset of employers toward rewarding and engaging employees.

1.1 Gender wage gap – definition and its main sources

The exact definition of what is the gender wage gap is not united even among researchers. In the 20th century, the gender wage gap was present mainly due to human capital differences (Becker, 1964). Becker, in his work, presented gender wage discrimination as wage difference between workers with the same human capital characteristics. However, at that time, the differences in human capital endowments between men and women were present. Whether we look at schooling, women's labour participation, or other factors, women were scarce in these factors compared to men – the numbers speak for themselves (Goldin, 2014). The average years of schooling in 1981 were 13.3 years for men and 13.2 years for women, but even with the same distribution of schooling, in 1981, only 7.4% of women had an advanced degree, but for men, it was 10% (Blau and Kahn, 2017). At the end of the 20th century, the 1990s in the United States, the conditions in the labour market changed (Goldin, 2014). Women became more educated; the factor of schooling has changed even to favour women against men. This trend has continued to the 21st century. The average years of schooling were 14.3 years for men and 14.5 years for women in 2011, but the distribution of advanced degrees changed in favour of women – 15.7% of women had an advanced degree and only 12.9% of men (Blau and Kahn, 2017). With equalizing role of men and women in terms of human capital characteristics, the gap is closing in time (Blau and Kahn, 2017).

Nevertheless, as the data presented later in this chapter show, some amount of the gap is still present. Even though the problem of uneven human capital distribution is resolved, there is still another remaining problem – the role of woman as a mother (Becker, 1964; Young, 2011; Blau and Kahn, 2017; Alon et al., 2020; Fortier, 2020). The unpaid parenting years take an essential portion of women's working experience and put women into the category of less experienced workers (Young, 2011; Blau and Kahn, 2017).

In the 21st century, we can define the gender wage gap as a difference in pay based on a rewarding system. Firms (or other employers) tend to pay more to employees who work

long and particular hours (Goldin, 2014; Bertrand, 2018). We may see it as discrimination against women because of their "second shift" (Lichard et al., 2021). Lichard, Pertold and Škoda (2021) look at the time use of women in different regions. They observed that the elasticity of substitution of time spent on housework changes with the increasing income of women, but at some point, it reaches a constant value. Their explanation suggests it is some sort of "glass-ceiling" at home put on women. Women with high incomes usually need to work more and longer hours, but their compulsory work at home does not decrease. They are forced to spend the same time in their households, even if their time at work increase. This factor usually implies a trade-off between work and home, which might lead to lower productivity at work (Bertrand, 2018; Lichard et al., 2021).

According to their data, Blau and Kahn have presented another concept generating the pay difference between men and women – female workers tend to be more risk averse than male workers. They presented females as risk-averse employees, which resulted in not being interested in demanding a better-paid, but unstable position (Babcock and Laschever, 2003). It is commonly referred to as a glass ceiling put on women (Kunze, 2008; Blau and Kahn, 2017; Ciminelli et al., 2021).

In the past years, researchers have been referring to the gender wage gap as a concept which occurred due to differences in productivity (Goldin, 2014). We can define *wage* as a summary statistics of an individual's human capital characteristics, prior working experience and future labour force participation. Again, we use the same argument as in the past – the unpaid childcare of mothers and the risk aversion of female workers put women into worse positions, but not for a lower level of education or not being experienced enough. The problem arises in the number of hours worked and the expected productivity (Kunze, 2008; Goldin, 2014).

1.2 Approaches towards quantifying the gender wage gap

1.2.1 Mincer regression

The critical method for measuring the gender wage gap is the Mincer regression (Mincer, 1974). It estimates the mean difference in hourly wage based on the human capital characteristics expanded for variables describing work background (Mincer, 1974; Mincer and Polachek, 1974). This approach might be the most popular among the

researchers. To better understand how the Mincer regression helps estimate the GWG, in 2020, the average female full-time worker earned about 82.3% of what the male worker did. (US Bureau of Labor Statistics, 2022)

Blau and Kahn (2017) used it in their overview of the GWG in the United States across time. The results they presented show the female-to-male unadjusted earnings ratio had kept stable at around 60% up until the 1980s, when it had promptly increased and has kept on this trend. The unadjusted ratio for 2010 has been circa 80%. Using panel data (PSID – Panel Study of Income Dynamic), they estimated three different logarithmic wage ratios from 1980 to 2010. They calculated the unadjusted ratio, the adjusted ratio for human capital specification and the adjusted ratio with full specification of employees. To get human capital specification, they added variables depicting education, experience, and region to the regression where the employees live and the employee's race. In the next step to obtain full specification, they had a regression model with variables describing human capital and added variables for unionisation, industry, and occupation. In 1980, the unadjusted log wage ratio was 62.3%, the log wage ratio adjusted for human capital specification was 71.1%, and the full-specification log wage ratio was 79.4%. Following the same order, the values taken for 2010 are 79.3%, 82.1% and 91.6%[1] (Blau and Kahn, 2017).

Fortin (2008) also used the Mincer regression in her study about the GWG among young adults in the US. She interpreted it as a regression model with a female dummy variable representing the unadjusted gender wage gap. The data used in the study are from the National Longitudinal Study of the High School Class of 1972 (NSL72), where they observed students in their final year of high school in 1972 and continued to interview them in 1973, 1974, 1976, 1979 and finally 1982. The data represent a widely used source to examine the effect of education on gender differences. As another source for data, she used the National Education Longitudinal Study of 1988/1994 (NELS88). The character of the dataset allows Fortin to observe the respondents' noncognitive traits, the attributes that are not observable from labour market characteristics. The data for the wage variable is recorded for respondents in their mid-twenties and their early thirties in both datasets (Fortin, 2008). The raw gap in 1979 (respondents of age 25) for

[1] For graphical representation see Figure 9 in the Appendix

data NSL72 is estimated to be -0.237[2], and if we compare it with the raw gap for the same cohort in 1986 (respondents of age 32), we observe the raw gap equals -0.229[3]. It is the change in the gap across the same group of respondents in different years, but using the data from NELS88, Fortin compares the same age groups but in different years. The unadjusted GWG for NELS88 is estimated to be -0.181 in 2000 (respondents of age 24)[4]. After estimating the raw GWG, Fortin added variables describing noncognitive traits, and as a fully specified model, she added the education describing variables as well as personal characteristics, such as race, marital status, and parental status. These correspond to specific columns in tables 7,8,9 in the Appendix, col. (1) raw GWG, col. (2) only noncognitive traits and (6) full specification holds for all three tables. From the results, we see that without any other specification, just using the noncognitive characteristics, the estimate of the gender wage gap improved. It took -0.22 in 1979; for the same cohort in their thirties, it was -0.187; for the data NELS88, it took -0.173 (Fortin, 2008).

When we move to the 21st century, the research on the GWG is still present. From my perspective, one of the fundamental studies done in the field is by Goldin (2014), which I have already mentioned earlier on. In her "last chapter", she used the Mincer form of regression to analyse the impact of within and between occupation discrimination on different groups of employees. She observed the American Community Survey (ACS) data from 2009 to 2011. For the purposes of research, she divided data into four samples – one carrying only the full-time, full-year workers, the second group for all workers – it does not depend on the type of the worker and their education level, the third and the fourth group are for college graduates (at least BA), who are working full-time and those where their working hours are not crucial. The results are presented in Table 1 below. Whether we look at each of the four samples, we can see a similar pattern there. By adding more explanatory variables, the GWG is improving. For example, the "Full-time, BA" sample has almost identical standard error for all four specifications. While the "Basic" including only female dummy, quadratic age variable, race and year estimates GWG at -0.285 with $R^2 = 0.131$, the improvement, with full specification adding logarithms of the amount of hours work per week, education dummies and the most important dummy for occupations, is significant. The specified sample estimates

[2] Table no. 7 in the Appendix
[3] Table no. 8 in the Appendix
[4] Table no. 9 in the Appendix

GWG for – 0.163 with $R^2 = 0.374$. It leads to a discussion on what exactly are the drivers of the wage differences. (Goldin, 2014). She promotes two explanations – one looks at women's lower bargaining power and risk aversion (Babcock and Laschever, 2003), and the other puts weight on employers and their unwillingness to hire women (Lazear and Rosen, 1990). With these explanations, Goldin (2014) stresses that the majority of the GWG is present *within* the occupation and not *between*. She even points out that changing the distribution of men and women between occupations does not make the change. Another example, looking at the college graduates, who are working full-time, full year, the GWG is 0.323 log points, where 68% of the gap is due to within gap while only 32% corresponds to between occupation gap (Goldin, 2014). As she observed similar results in the different groups (part-time workers, not college graduates and overall), she suggested that the last chapter of gender convergence should focus more on the working conditions within the occupations. The more flexible working conditions (e.g., not strict working hours, home-office allowed), the better it will also be for women and men.

Table 1 Role of occupation for various samples (Goldin, 2014)

TABLE 1—RESIDUAL GENDER DIFFERENCES IN EARNINGS AND THE ROLE OF OCCUPATION

Sample	Variables included	Coefficient on female	Standard error	R^2
Full-time	Basic	−0.248	0.00101	0.112
Full-time	Basic, time	−0.193	0.00100	0.163
Full-time	Basic, time, education	−0.247	0.000905	0.339
Full-time	Basic, time, education, occupation	−0.192	0.00104	0.453
All	Basic	−0.320	0.00105	0.102
All	Basic, time	−0.196	0.000925	0.353
All	Basic, time, education	−0.245	0.000847	0.475
All	Basic, time, education, occupation	−0.191	0.000963	0.563
Full-time, BA	Basic	−0.285	0.00159	0.131
Full-time, BA	Basic, time	−0.230	0.00158	0.177
Full-time, BA	Basic, time, education	−0.233	0.00155	0.216
Full-time, BA	Basic, time, education, occupation	−0.163	0.00158	0.374
All, BA	Basic	−0.384	0.00173	0.119
All, BA	Basic, time	−0.227	0.00151	0.380
All, BA	Basic, time, education	−0.229	0.00148	0.407
All, BA	Basic, time, education, occupation	−0.163	0.00151	0.525

Notes: "Basic" regression is the log of annual earnings regressed on the female dummy, age as a quartic, race, and year. "Time" adds log hours per week and log weeks. "Education" adds dummies for education categories (and those above a BA for the college graduate sample). "Occupation" adds three-digit occupation dummies. "Full-time" is 35 and above hours per week and 40 and above weeks per year. "All" includes workers 25 to 64 years old with positive earnings and positive hours worked during the past year. The "full-time" sample consists of full-time, full-year individuals 25 to 64 years old excluding those in the military using trimmed annual earnings data (exceeding 1,400 hours × 0.5 × 2009 minimum wage). The "BA" sample includes workers with at least a college or university bachelor's degree. The number of observations is 2,603,968 for full-time, 3,291,168 for all, 964,705 for full-time BA or more, and 1,162,638 for all BA or more.

Source: American Community Survey 2009 to 2011.

Source: Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, *104*(4), 1091-1119, p. 1098.

1.2.2 Blinder-Oaxaca Decomposition

The researchers are usually interested not only in the absolute value of the gender wage gap but in the components, it consists of, as well. They decompose the raw wage gap (difference between logarithmic wages) into two parts – the explained part corresponding to the differences in endowments of men and women and the unexplained (or residual) part; its existence is usually interpreted as discrimination (Kunze, 2008). Several forms of decomposition are known, e.g., JMP decomposition (Juhn et al., 1991), which aims to study the gender wage gap across time and different countries. Another known decomposition is the so-called "gender difference at mean " - the Blinder-Oaxaca (B-O) decomposition (Blau and Kahn, 2017) presented by its authors in the 1973 (Blinder, 1973; Oaxaca, 1973).

The idea behind decomposition into two parts is nicely presented by Kunze (2008) in her study "Gender wage gap studies: consistency and decomposition ". She presented reasons why is decomposing wage a helpful tool for research. The decomposition (not only B-O) decomposes the total gap into explained and residual part of the raw wage gap. It navigates research towards the main drivers of the gender wage gap and gives a more consistent view than the Mincer model (Kunze, 2008). As it is mentioned above, the difference in human capital between men and women can be neglected nowadays, so the explained part of the decomposition is getting closer to zero in time, or it might result even in favour of women (Goldin, 2014; Blau and Kahn, 2017). On the other side, the unexplained part or, in other words, the residual part has risen in time. The researchers are asking the question "What does the residual part stands for?". The economic theory delivers a few possible explanations. The classical thought is that the unexplained part of the gender pay gap is due to discrimination. Contrary to this, some think the unexplained part represents women's risk aversion and, therefore, might be found as "wage difference driven by occupation" (Goldin, 2014). To build up, other theories describe the unexplained part as the difference due to lower productivity of women, or like discrimination view, the unexplained part of the wage gap serves for the gender difference in probability of leaving labour force (Goldin, 2014; Blau and Kahn, 2017).

The results of Blinder-Oaxaca decomposition performed by Blau and Kahn (2017) confirmed the findings from previous studies using the decomposition methods and the

findings given by the Mincerian regressions. The explained part of the gender wage gap played crucial role in the 1980s in the US labour market. The values of endowments, which depicted the proficiency of women, have risen rapidly in the late 20th century and cause the explained part of the wage gap even favour women. In explaining the GWG, they calculated the female to male logarithmic wage ratio. The unadjusted ratio is the ratio of logarithmic wages not controlling for covariates of explanatory variables (i.e., human capital characteristics, job specification – in their terms, "full specification"). While as an adjusted ratio, they mean ratio of logarithmic wages controlling for "full specification". The unadjusted value of the female-to-male logarithmic wage ratio was 62.1% in 1980. When controlling for human capital specification, it took value of 71.1% in 1980. The improvement of women's proficiency, we see in the change of the value in 1998 and 2010, respectively. The unadjusted ratio has risen to 77.2% and 79.3%, and for the adjusted ratio with human capital specification it was 82.7% and 82.1%. Nevertheless, the wage gap remains present – the observations conclude it is due to residual part of the gender wage gap (Blau and Kahn, 2017).

1.3 International overview

Results from previously mentioned studies can be used for other countries, not only for the USA. The demography of the USA is evenly distributed. They have different ethnicities, races, and married and single working people. To describe other states in the world, we can use the results from the USA. Olivetti and Petrongolo, in 2008, performed an international study that looked not only at European countries or only at the USA: They look on the gender wage gap in the US and Europe. They worked with PSID data for the USA, same as Blau and Kahn in their studies. To explore gender difference in Europe, they analysed data from 1994 to 2001 gained through European Community Household Panel Survey (ECHPS). This survey questions a few thousand households across European Union.

The results in Table 2 show the USA has a more significant gap than any other European country. Col. 1 reports the median value of gender wage gap based on the hourly wage for 1999 (the base year of the study). Columns 2 and 3 depicts the same as column 1, just that they are adjusted for missing values. Col. 2 is adjusted using the actual value from the nearest wage observation in 2 years window, where col. 3 is adjusted based on the actual value from the nearest wage observation. Comparing

values of the first three columns, the GWG increase with more values included. The authors claimed downward bias due to nonrandom sample selection as the methods to replace missing values might cause a lower chance of including low-wage women (Olivetti and Petrongolo, 2008). Table 2 shows that the sample selection method has a more substantial effect in countries that report high employment gaps as France, Spain, and Greece, where the change from column 1 to column 3 is more than 60%. Compare to the UK or Scandinavian countries, for whom the sample selection method only marginally affected the GWG (Olivetti and Petrongolo, 2008).

Table 2 Median Wage Gaps across European Countries (Olivetti and Petrongolo, 2008)

Table 2
Median Wage Gaps under Alternative Sample Inclusion Rules: Wage Imputation Based on Wage Observations from Adjacent Waves

	1	2	3	4	5
Country:					
United States	.339	.352	.361	.354	.363
United Kingdom	.256	.277	.284	.291	.301
Finland	.160	.194	.197	.203	.206
Denmark	.086	.100	.100	.093	.093
Germany	.191	.223	.214	.234	.234
Netherlands	.178	.193	.199	.195	.199
Belgium	.078	.099	.112	.098	.112
Austria	.192	.224	.234	.220	.229
Ireland	.232	.273	.284	.292	.300
France	.095	.133	.152	.140	.164
Italy	.059	.062	.075	.070	.079
Spain	.097	.153	.168	.143	.157
Portugal	.150	.168	.186	.169	.185
Greece	.111	.148	.184	.148	.185
Correlation	−.329	−.263	−.181	−.269	−.199
Coefficient of variation	.484	.416	.382	.427	.392

SOURCES.—Michigan Panel Study of Income Dynamics and European Community Household Panel Survey.

NOTE.—All wage gaps are significant at the 1% level. Figures in the last two rows display the cross-country correlation between the reported gender wage gap and the gender employment gap and the coefficient of variation of the gender wage gap. Sample description: aged 25–54, excluding the self-employed, the military, and those in full-time education, 1999. Sample inclusion rules by columns: (1) Employed at time of survey in 1999. (2) Wage imputed from other waves when nonemployed (−2, +2 window). (3) Wage imputed from other waves when nonemployed (−5, +2 window). (4) Wage imputed from other waves when nonemployed (−5, +2 window), adjusted for real wage growth by gender and country. (5) Wage imputed from other waves when nonemployed (−5, +2 window), adjusted for real wage growth by gender and country.

Source: Olivetti, C., & Petrongolo, B. (2008). Unequal pay or unequal employment? A cross-country analysis of gender gaps. *Journal of Labor Economics*, *26*(4), 621-654, p. 632

1.4 Gender wage gap and COVID-19 pandemic

This section will look at the most recent studies about the GWG and how it has been affected by the recent pandemic crisis. This thesis is orientated explicitly on how the COVID-19 pandemic changed what factors drive the gender wage gap. Alon et al. (2020) analysed the structure of employment in recent years. They argued that pandemic crisis will have a stronger effect on the women than on men, in opposition to

the financial crisis in 2008-09 (Alon et al., 2020; Bonacini et al., 2021). The data show that women are primarily employed in the countercyclical sectors, e.g., Government, Education and Health Services. Employment in these sectors accounts for 40% of women's employment compared to only 20% of men's employment. It is suggested that it might be due to job preferences of men, who are working in mostly procyclical sectors, e.g., Trade, Manufacturing (Alon et al., 2020). Another way of looking at women's employment structure is that they prefer caregiving jobs more than other jobs (Thomason, 2020). The one thing that caregiving jobs have in common, except being heavily dominated by women, is that they belong to a "lower paid" group of jobs and in the view of the COVID-19 pandemic, the burden of the pandemic restrictions lays on their shoulders (Thomason, 2020; Alon et al., 2020).

The trend the COVID-19 pandemic crisis brings into the gender discrimination field is among economists referred to as "shecession". While the financial crisis in 2008-09 had a more extensive effect on the employment of men and is, therefore, referred to as "mancession", the latest evidence for current crisis suggests otherwise (Alon et al., 2020; Bonacini et al., 2021).

The current crisis is specific to the pandemic restrictions considered by governments. Working sectors, where possible, moved their performance to the online world and the employees were required to work from home (WFH). The sectors affected by WFH condition are education, administrative jobs and others. These are the ones dominated by women. Bonacini et al. (2021) look at how working from home changed the gender wage gap in their working paper. They divided jobs into ones with "low WFH attitude" (e.g., emergency workers) and those with "high WFH attitude" (e.g., teachers). They concluded that GWG is present and significant in both groups but is higher among the occupations in the “high WFH attitude” group, i.e., higher for jobs occupied by women, (Table 3, Bonacini et al., 2021). The severe effect of the pandemic crisis on women led them to introduce the "shecession" term.

Table 3 Estimates and decomposition of the average GWG, (Bonacini et al., 2021)

Table 3 – Estimates and decomposition of the average GWG

Group of employees	Difference	Explained	%	Unexplained	%
Total sample	0.249***	0.055***	22.1	0.194***	77.9
High WFH attitude	0.316***	0.127***	40.2	0.189***	59.8
Low WFH attitude	0.195***	0.037**	19.0	0.159***	81.0

*Notes: *** p<0.01, ** p<0.05, * p<0.1. The table presents coefficients of the variable of interest (i.e. being male) only. Complete estimates for the pooled sample are provided in Tables A.3–A.5. Employees with a high WFH attitude level are defined as those reporting a value of the WFH attitude index over the sample median (i.e. 52.2).*

Source: Bonacini et al., (2021) *Will it be a shecession? The unintended influence of working from home on the gender wage gap related to the COVID-19 pandemic* (No. 771), p.11.

The woman's role as mother plays a crucial part in the final effect of the COVID-19 pandemic on the GWG (Fortier, 2020). The attack of the COVID-19 virus on countries has pushed the governments to take extreme restrictions. They were forced to close schools; many industries must transfer their working life into the online world. The applied quarantine measures placed a restriction on the US labour market and led women to either quit their jobs or switch from the full-time to part-time – the amount of unemployed single-parent mothers has more than tripled in 4 months period in 2020, moving from 4.1% to 15.9%, (Bonacini et al., 2020; Fortier, 2020). Fortier (2020), based on the available data, estimates that switching from full-time to part-time job will results in a 15% decrease of average women's income in the future. The reduction in women's income, together with the risk-aversion of women might be the cause of the "step back" in the equal payment for female and male workers.

2. Theoretical framework

This chapter will continue to align with the previous one, offer economic explanation of the wage and other terminology, and explain the chosen methods and data presented in the following chapter. It will link the theory, the methodology and the chapter presenting research results.

The thesis is initially focused on the wage gap between men and women in the USA and how the COVID-19 pandemic affected it. To explore the wage gap, first I need knowledge what the wage is and how it is determined. In plain language, the hourly wage is a price for an hour worked. From the economic perspective it is the opportunity cost of an hour of leisure the worker gives up. The economists claim the wage is the

result of negotiations between risk-neutral employers and risk-averse employees, and it should be the equilibrium price in the labour market (Bertrand, 2004). According to Bertrand (2004), these negotiations between the employer and the employee are sensitive to the external labour market conditions, but they do not respond to the current labour market situation. She referred to these negotiations as "wage-setting process". These claims raise a new question – through the negotiations, what are determinants of wage?

In the 20th century, when wage gap research emerged, the employee's wage was determined by his human capital endowments (Becker, 1964). The amount of schooling he/she has, the previous job training, and other represent an individual's investment into his or her education level. These factors are easily measured, and their values can be obtained through population surveys. On the other hand, the wage should mirror not only the human capital endowments of employee but also his working experience and productivity. The experience of a worker cannot be measured; the commonly used is an approximation using some form of mathematical transformation of age (Becker, 1964; Mincer, 1974; Willis, 1986). Putting all the mentioned determinants together, we observe the famous Mincer equation (Eq.1, Chapter 3). The equation might differ across economists, one of the reasons for it, might be the perception of human capital. The investments in human capital differ across the employees, contrary to Becker's intuition of homogeneity of human capital (Willis, 1986).

> *Research question no.1: "Does hour wage change in time in the US labour market?"*
>
> *Research question no.2: "Does the gender wage gap change in time in the US labour market?"*

Following the assumption of heterogenous human capital, there is a place to wonder what portion of the wage gap can be described by differences in the endowments. However, differences in the endowments cannot be observed using only Mincer regression and I need to know determinants of the wage gap. It leads me to decompose the wage gap into two parts - the explained and the unexplained (Eq. 7, Chapter 3). The decomposition can be provided using the method of Oaxaca and Blinder (1973), explained in more detail in the next chapter.

Research question no.3: "Does determinants of gender wage gap change in time?"

Research question no.4: "Does gender wage gap change due to difference in endowments?"

Nevertheless, the main goal of this thesis is not to analyse the gender wage gap alone. This thesis aims to find out the effect of the COVID-19 pandemic on the gap and its influence on the individual determinants of wage, including the employee's occupation and industry. The economic theory offers its point of view. The theory suggests that women will suffer from the consequences of the pandemic restrictions more than men (Alon et al., 2020; Bonacini et al., 2021; Fortier, 2020). Fortier (2020) brings attention to the role of women as mothers – with restrictions put on society by pandemics, e.g., closure of schools/day-cares and home office; it forced many women to leave the job or to switch from full-time to part-time in order to maintain work-life balance. Contrary, Bonacini (2021) looks at it from a labour market point of view. The pandemic restriction caused a shrinking of the labour market and preferred not-in-person options. Closing services and highlighting online shopping have resulted in redundant work positions. As they argue with data, women occupy most of these working positions.

Research question no.5: "Does hour wage change in US labour market after the outbreak of COVID-19?"

Research question no.6: "Does the hourly wage of men and women react differently to the COVID-19 pandemic?"

Research question no.7: "Does the determinants of gender wage gap change after the COVID-19 pandemic?"

The literature predicts that pandemic restriction put much more weight on women than men (Bonacini et al., 2021). This thesis looks at the problem similarly. Based on the abovementioned factors (women as mothers, risk aversion of women), I expect my data to support these factors. The Mincer regression will be a tool to explore the development of hourly wage. One way to look at it is that the hourly wage of women should raise as they are mostly working in sectors heavily burden by pandemic, e.g., health sector, (Alon et al, 2020), which causes a need for workers and should serve as a motivation factor. On the other hand, the "second shift" of women might cause many to

leave the labour market to be at home and provide care for their families and households (Lichard et al., 2021; Fortier, 2020).

3. Methodology and Data

Chapter 3 will introduce the research approach and how to answer the research questions. The first three research questions are interested in the gender wage gap and its value. To tackle these questions, we need to estimate the value of the GWG. As presented in literature review, various ways of measuring the gender wage gap exist. Because the interest of this thesis is in its absolute value and its change in time, the previous research led us to use the Mincer regression (Mincer, 1974).

The estimated values of either GWG or the hourly wage are tools for further discussion. They show the development of the GWG in time and the hourly wage over time and even help deduce, what the main drivers of wage differences across genders are. Nevertheless, they do not answer whether the GWG is solely due to differences in human capital between men and women or other reasons. On the other hand, the decomposition method is used to find determinants of the wage gap and what portion of the GWG can be explained due to differences between men and women. The method focuses on the difference between the mean value of mutual group indicator, e.g., gender, race, and work position. It offers a more detailed view of elements behind the GWG and helps us interpret what amount of the wage difference is explained by the chosen bundle of control variables. The types of decomposition might differ; this thesis will proceed with the Blinder-Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973).

The topic of this thesis requires working with the latest data. At the beginning of writing, the latest available data came from the United States. The argument of availability and time accuracy plays a decisive role in the research. Therefore, all arguments mentioned above point to choosing to work with the data from the US. I obtained my data using the online database IPUMS and focused on data gained through the Current Population Survey. These data consist of personal and sensitive information, e.g., the hourly wage, age, and level of education; for approved access by the institution, I have to declare the purposes to use the data and that I will not freely distribute the data. Because of these reasons, I will not provide the transcript of datasets and code in the Appendix.

3.1. Methodology

3.1.1. Mincer regression

The Mincer regression has been widely taken as a benchmark for linear approximation between the earnings and the human capital endowments. Mincer proposed a regression relationship (Equation no. 1) in his book "Schooling, Experience and Earnings " (1974), and since then, it has widely influenced the research in the field of labour economics.

I will use a slightly modified version of his linear regression in my thesis. In the original model, the earnings, in logarithmic form, are regressed on the variables describing human capital characteristics, including gender. The variable depicted person's experience is in the form of the second polynomial function of *age* The model is estimated using Ordinary Least Squares.

$$ln(E_i) = \alpha ln(E_0) + \beta_i \sum_{i=1}^{n} k_i + \gamma_i \sum_{i=1}^{n} l_i + u_i \quad (1)$$

To demonstrate, the left-hand side of my model consists of logarithm of hour wage regressed on various characteristics of a person, as equation no. 2 symbolizes. The latter equation shows more concrete view of the models. In equation no. 3, the natural logarithm of hour wage is regressed on different human-describing variables. The first sum of the vector X_i depicts the personal characteristics of each respondent, such as his gender, age, working experience and his marital status. The other sum of the vector Z_i corresponds to the work characteristics, i.e., education level, occupation and industry. As specified by equation no. 3, the model will be referred to as the base model, but without the vector Z_i, only with personal characteristics X_i.

$$ln(W_i) = \alpha + \beta_i \sum_{i=1}^{n} T_i + u_i \quad (2)$$

$$ln(W_i) = \alpha + \beta_i \sum_{i=1}^{n} X_i + \gamma_i \sum_{i=1}^{n} Z_i + u_i \quad (3)$$

3.1.2. Blinder – Oaxaca decomposition

The method was introduced in 1973, by researchers - Blinder and Oaxaca separately. Because of the separate work of both authors, the decomposition is named Blinder-Oaxaca (B-O). The B-O decomposition is built on the simple regression equation:

$$Y_i = \beta_0 + \sum_{j=1}^{n} \beta_i X_{ij} + u_i,$$

(4)

where Y_i stands for the natural logarithm of the dependent variable; in our case, it is hour wage. The following term X_{ij}, j = 1, …, n indicates explanatory variables for all *i* observations in the dataset used to explain, in our case, the wage differentials between male and female workers (Y_i). The next step in the decomposition is the comparison of the outcomes for two different groups – men and women. We transform the original equation into two separate ones.

$$Y_i^M = \beta_0^M + \sum_{j=1}^{n} \beta_j^M X_{ij}^M + u_i^M$$

(5)

$$Y_i^F = \beta_0^F + \sum_{j=1}^{n} \beta_j^F X_{ij}^F + u_i^F$$

(6)

In the equations above, the upper index M stands for male and the upper index F stands for female. Further, these two equations lead us to the actual decomposition proposed by Blinder and Oaxaca. The decomposition is based on the difference in the mean outcome of the dependent variable, i.e., hourly wage, for the two proposed groups, i.e., men and women. Through the series of algebraic steps, we arrived at the final line, which is, as denoted, composed of two main elements (Equation no. 7). The first element corresponds to the so-called "explained gap". The explained gap is the part of the difference in average hourly wages of the two selected groups, which can be justified by the differences in the endowments and clarified by the explanatory variables, e.g., difference in education. The second part matches the so-called "unexplained gap". After controlling for the explained gat, the unexplained gap remains

from the hourly wage difference. Its non-zero value is a sign of a wage gap between the two experimental group. There are two main logical branches – the one that argue that the unexplained part describes the difference due to unequal productivity of the selected groups and the other, whose argument for presence of an unexplained gap is discrimination, (Jann, 2008; Navarová, 2022).

Secondly, there has been made an inevitable extension of the method. Neumark (1988) proposed one of the main extensions. As the base for his method, he used pooled regression without using the group-specific intercept. On the other hand, he cleared the data and compared only the observation with similar productivity. By clearing the dataset, he argued that discrimination caused by the difference in productivity is no longer present.

$$\overline{Y_i^M} - \overline{Y_i^F} = \widehat{\beta_0^M} + \sum_{j=1}^{n} \widehat{\beta_j^M} \cdot \overline{X_j^M} - \widehat{\beta_0^F} - \sum_{j=1}^{n} \widehat{\beta_j^F} \cdot \overline{X_j^F} =$$

$$= \left(\widehat{\beta_0^M} + \sum_{j=1}^{n} \widehat{\beta_j^M} \cdot \overline{X_j^M} - \left(\widehat{\beta_0^M} + \sum_{j=1}^{n} \widehat{\beta_j^M} \cdot \overline{X_j^F} \right) \right)$$

$$+ \left(\left(\sum_{j=1}^{n} \widehat{\beta_j^M} \cdot \overline{X_j^F} \right) - \widehat{\beta_0^F} - \sum_{j=1}^{n} \widehat{\beta_j^F} \cdot \overline{X_j^F} \right)$$

$$= \underbrace{\left(\sum_{j=1}^{n} \widehat{\beta_j^M} \left(\overline{X_j^M} - \overline{X_j^F} \right) \right)}_{explained} + \underbrace{\left(\left(\widehat{\beta_0^M} - \widehat{\beta_0^F} \right) + \sum_{j=1}^{n} \overline{X_j^F} \cdot \left(\widehat{\beta_j^M} - \widehat{\beta_j^F} \right) \right)}_{unexplained}$$

(7)

3.2. Data

Due to the decision to have a bachelor thesis focused on the recent COVID-19 pandemic, it is necessary to have the latest data possible. This, together with other reasons, decided to work with the data covering situation in the United States of America. The data were obtained from the available public source – the Integrated Public Use Microdata Series, and the observations were collected using the Current Population Survey.

3.2.1. Dataset

The primary dataset covers the period from 2018 to 2021 and consists of four subsets. The data from 2022 are excluded only for consistency and work efficiency. The four subsets are datasets consisting of cross-sectional data, and each corresponds to one year. When working with data, I divide each of the four years into quarters. As there occurs a problem with some variables in the fourth quarter of 2021, I exclude it from the decomposition models. Firstly, as mentioned, each of the subsets is in the form of cross-sectional data with specific variables corresponding to the year and the month. Secondly, the cross-sectional data contains essential demographic characteristics of respondents – US citizens. Lastly, following the research problem I choose to work with hour wage as the marker of the person's working evaluation.

3.2.1.1. The Integrated Public Use Microdata Series

The Integrated Public Use Microdata Series (IPUMS) is the individual-level public population database. The data are organised by different characteristics and the source of the collecting observations. For this thesis, I work with the data collected via the Current Population Survey (CPS). The CPS offers time-accurate information for each month of every year. The Current Population Survey is a monthly US Labor Force survey. The data are conducted jointly by the US Census Bureau and the Bureau of Labor Statistics. The data include various information, e.g., demographic information, employment data, program participation or other supplements such as tobacco use, volunteering, and more. (IPUMS USA | The IPUMS Project: An Update)

In addition, I need information about the income of the respondents. Considering it is sensitive information, there is an Outgoing Rotation Group (ORG) category. Those people, whom the National Burau of Economic Research interviews for 4 months, then have an eight-month break and are interviewed again for another 4 months. After this procedure, they are thrown out from the sample. These people are asked about their economic situation, such as hourly wage, hours worked a week, and similar information. The sample is restricted to consist only of people from the ORG, aged 15 and older, who are employed and taking wage or salary (self-employed people are not considered).

Access to the CPS data is available through the IPUMS online database. The condition for approved access is to declare the purposes of using the data – personally,

confirmation of the student's status at Charles University and that the data will be used for purposes of bachelor thesis research.

3.2.2. Data overview

To depict the situation before the COVID-19 pandemic and the changes after its outbreak. The cross-sectional data are restricted to four years (2018-2021) and divided into four different datasets corresponding to each year. The datasets consist of 156 720, 151 618, 130 669 and 131 338 observations each year. The restrictions put on the data are an age window - from 15 to 70, and availability - consists of only people with an available hourly wage. As the next step, the dataset is composed as follows. Except for the predefined variables, it consists of another 12 variables – AGE, SEX, RACE, MARITAL STATUS, EMPLOYMENT STATUS, LABOUR FORCE PARTICIPATION, OCCUPATION, INDUSTRY, LEVEL OF EDUCATION, EARNINGS WEIGHT, HOURLY WAGE and restrict to only eligible data by variable EARNINGS ELIGIBILITY.

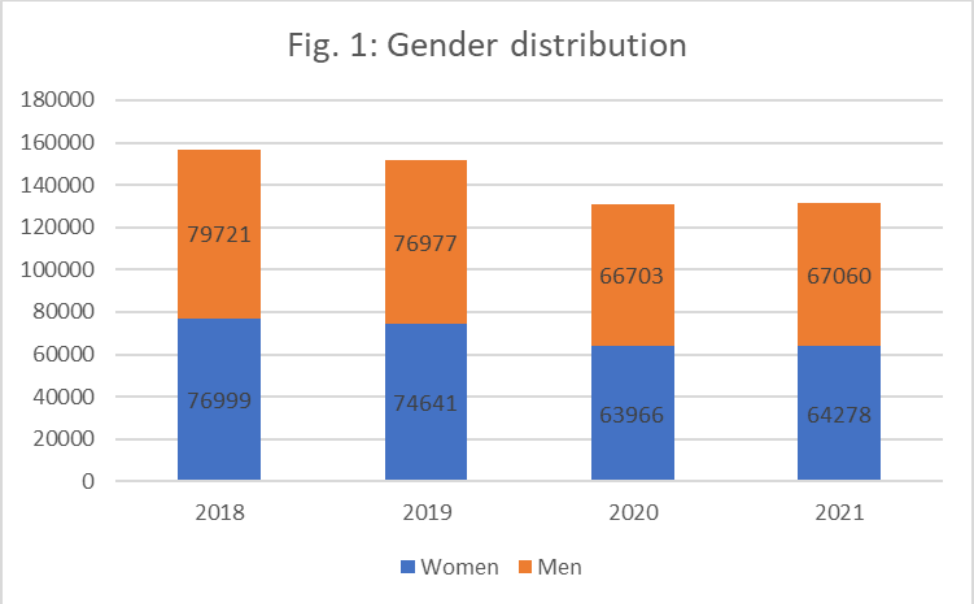
Another step in processing the data was the correct transformation. Some variables were transformed into binary form. These concern variables are SEX, MARITAL STATUS, LABOUR FORCE PARTICIPATION and EMPLOYMENT STATUS. For the SEX variable, the transformation took the form of FEMALE, indicating whether the respondent is a woman or not. The same logic implies to the MARITAL STATUS, LABOUR FORCE PARTICIPATION and EMPLOYMENT STATUS. In the case of the last two mentioned, the redefined variables took value one strictly, as there was a dominant restriction on the dataset to take only those observations with an available hourly wage. It leads us to obtain a dataset with participants who are EMPLOYED and actively participated IN LABOUR FORCE.

There was a transformation needed for the HOURWAGE variable, too. The numbers took nominal values; to correctly compare the wages in time, they must be adjusted for inflation. The recommended adjustment by the data provider is using Consumer Price Index. Following their lead, 2019 is the reference year, the last year unaffected by pandemic restrictions. Based on this adjustment, in the following pages, the variable HOUWAGE reports values as 2019 US dollars.

In essence, the description of the data is provided using numbers. Besides the information about how many yearly observations yearly datasets consist of, Figure 1 shows the gender distribution in each year. There is a sign of an attempt to equalize the number of men and women, but men are slightly above each year. The number of observations is the highest in 2018 and lowest in 2020. Figure 1 shows it as well.

Figure 5 and Figure 6 (in the Appendix) picture the distribution of married and single people among both genders. While looking at the graph depicting females, the distribution is generally even. In the case of men, married men dominate the datasets for three years except 2021, where they take 49,27% out of all men.

Figure 1 Gender distribution across datasets



Source: Author

Table 1 presents the summary statistics for HOUR WAGE and AGE variables across all datasets. Other variables are excluded as there are either categorical or dummy variables. Statistics observing the HOURWAGE variable are biased because of substantial values for some respondents. The respondents with considerable hourly wages are depicted as 999.99. For age, many respondents are in their early 40s as the mean and median have similar values across datasets.

Table 4 presents the summary statistics for HOUR WAGE and AGE variables across all datasets. Other variables are excluded as there are either categorical or dummy variables. Statistics observing the HOURWAGE variable are biased because of

substantial values for some respondents. The respondents with considerable hourly wages are depicted as 999.99. For age, many respondents are in their early 40s as the mean and median have similar values across datasets.

Table 4 Descriptive Statistics

Table 4 - Descriptive Statistics

		Hour Wage	Age
2018	Mean	423,77	41,57
	Median	26,15	41
	St. deviation	483,5593	13,75
2019	Mean	429,6	41,68
	Median	38	41
	St. deviation	484,2143	13,83
2020	Mean	456,25	42,03
	Median	33	42
	St. deviation	487,2388	13,76
2021	Mean	451,6	41,69
	Median	32,5	41
	St. deviation	486,429	15,85

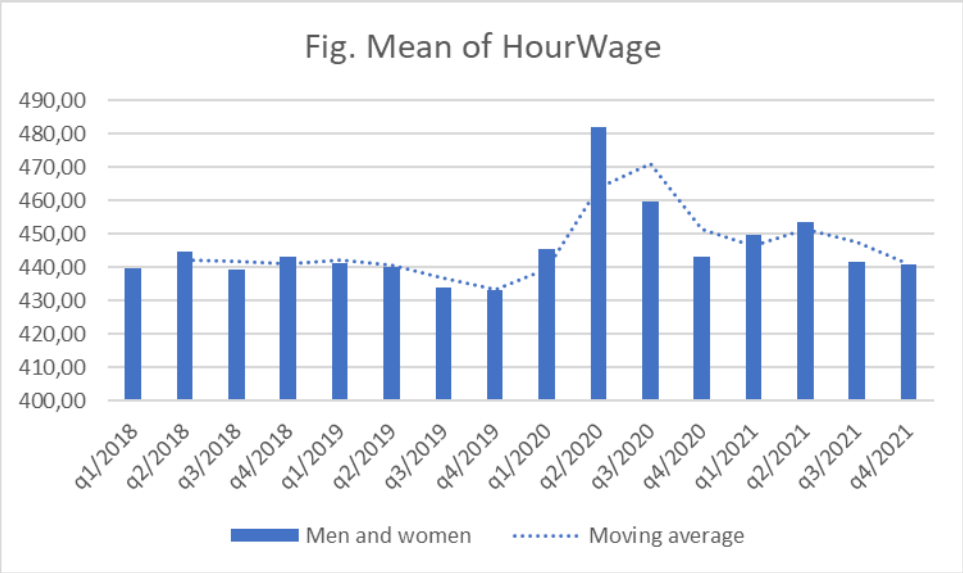
Source: Author

4. Results

Chapter 4 will describe research results and discuss possible reasons. It will answer the research question and draw conclusions based on the possible drivers of the results. To interpret wage differences and situation in the labour market before and after the beginning of the COVID-19 pandemic, I choose as the first period affected by the pandemic crisis the 2nd quarter of 2020. The initial restrictions came into effect during February and March or early April 2020 in most US states, so their consequences are expected to appear in the following periods (Forbes, 2020).

Before I proceed with the analysis of the gender wage gap determinants, I will present the evolution of average hourly wages in the US and the evolution of the gender wage gap as estimated using the Mincerian regression. These will help me to answer some of the research questions. After that, I will continue in reporting the results of the decomposition of gender wage gap. The results answer the rest of the research questions and help to understand whether the COVID-19 pandemic restrictions change determinants of the GWG and conditions in the labour market.

Figure 2 Mean of the hourly wage



Source: Author

Figure 2 reports the development of the average hourly wage for US workers across time. All wages are adjusted to the level of 2019 US dollars, as it is described in the previous chapter. The trendline set the trend based on the Moving Average of order 2. The mean of hourly wage kept stable up until the end of 2019. Starting 2020, it increased rapidly and peaked at 482 USD in the 2nd quarter of 2020. After this steep increase it gets closer to its initial value, and from the trendline we can expect to follow its trend and keep decreasing, until it reaches pre-COVID-19 values about 440 USD per hour. The hour wage reaches its maximum in the period I take as the beginning of the COVID-19 pandemic. It was the period when massive restrictions were put on society – lots of people were forced to leave their job, together with the shrinking of the labour market. The exit of mostly low-wage employees, whose working position might be redundant may cause a steep increase the average hourly wage. It is possible that many women, who leave the labour market due to their unpaid parental work, belong to the

"low wage " group of workers. The shrinkage of the labour market caused the employers to cut their cost; therefore, they kept only the essential employees. As many of those who either lost their job or left the labour market do not belong in the category of essential workers, we might expect they did not belong in the group of high-wage employees also.

The answer to research question no. 1 is "Yes, the hourly wage changes in time." As it does not keep stable in time. Figure 2 answers research questions no.1 and no.5. In the graph above, we see the time development of the average hourly wage for the whole working-age population. The pre-pandemic levels are stable, but then, there is a steep increase followed up by a steady decrease. Figure 7, in the Appendix, shows the average hourly wage for men and women separately. The women's real wage is at a lower level through the whole research period. The wage ratio is stable at about 90%.

To answer the fifth question, "Yes, the hour wage change after the COVID-19 pandemic. " In the chapter Methodology and Data, I declare the 2nd quarter of 2020, as the starting point for the period affected by pandemic restrictions. In Figure 2, we see the 2nd quarter of 2020 is the exact moment of a rapid increase in average hourly wage for the whole working-age population. This proves that pandemic restrictions have affected employees' hourly wages regardless of gender. Figure 7, in Appendix, inspects changes in the hourly wage separately for men and women; the direction and pattern of change are the same. The only difference is the level of women's average hourly wage. To conclude answer to the sixth research question is "No, the hourly wage of men and women does not react differently to the COVID-19 pandemic. " Tables, which depict the gender distribution across various occupation or industry sectors, show us that each category has its representants for men and women. Even though either men or women dominate some categories since there is no category without representants of both genders, their hour wages react in similarly. The thing influenced by gender distribution is the level of an hourly wage; what we can see differs.

A sudden increase in the value of the average hourly wage and then continuous decrease declares the change in the average hourly wage after the outbreak of the COVID-19 disease. The potential reason for such a steep increase in the average hourly wage is a change in the distribution of employees. From the data provided in the Appendix, the numbers of employees drop almost in every sector. The previous research claims that

most women are working in counter-cyclical sectors and men dominate pro-cyclical sectors. The data provides statistics that pandemic restrictions hit almost every working sphere, regardless of whether it is changing with the economy. For example, the gastronomy workers are dominated by women, but not as strongly as teachers or health assistants. The gastronomy service is, without dispute, one of the most burdened sectors by pandemic restrictions. The labour market shrinks by 30% comparing 2019 and 2020. In 2020, 28% fewer men were working as gastronomy workers and same percentage fewer women. However, as women take a dominant share in the gastronomy (about 56%), the 28% of women who leave the labour market accounts for a more significant share in the women's employment than men. Therefore, it hurts women more strongly than men, even though the effect of pandemic restrictions was the same for men and women. Another example provides changes in the sector of Manufacturing. The sector is dominated by men in a ratio of 3:1, men to women. The data in the Appendix shows the labour market was shrinking there too. The number of men working there before the COVID-19 pandemic was about 11 thousand, after the attack of pandemic restrictions it is less than 9,5 thousand. Approximately 15% of men loose or leave their job. It supports the argument, that mostly people with lower hourly wages usually leave the labour market, causing a steep increase in average hourly wage for the 2nd quarter of 2020. Table 5 shows the distribution of employees in the Information sector. In line with expectations, the sector is dominated by men; they took about 58%. The sector belongs to "high-wage" sectors, with the average hourly wage being more than 550USD and in 2020, attacking the value of 600USD. The number of employees decreases in the Information sector only shortly in 2020. By 2021, we see a 5% increase in employment. These arguments support the theory that COVID-19 hits mostly "low-wage " employees.

Eq. 8 refers to the original idea of labour economists, that the primary source of GWG is the difference in productivity between men and women (Becker, 1964). To capture the difference in productivity, I use the Mincer regression formulated as in Eq. 9. At beginning is the *Basic model*, where a second order polynomial of age is used as a proxy variable for working experiences. Next, to expand the *Basic model*, other variables will be added one by one until the final model equals to model formulated as in Eq. 9. To this model I will refer to as a *Fully specified model* or *Full specification*.

For the next purposes, I will look at the estimate of GWG by the *Basic model* as a raw gender wage gap.

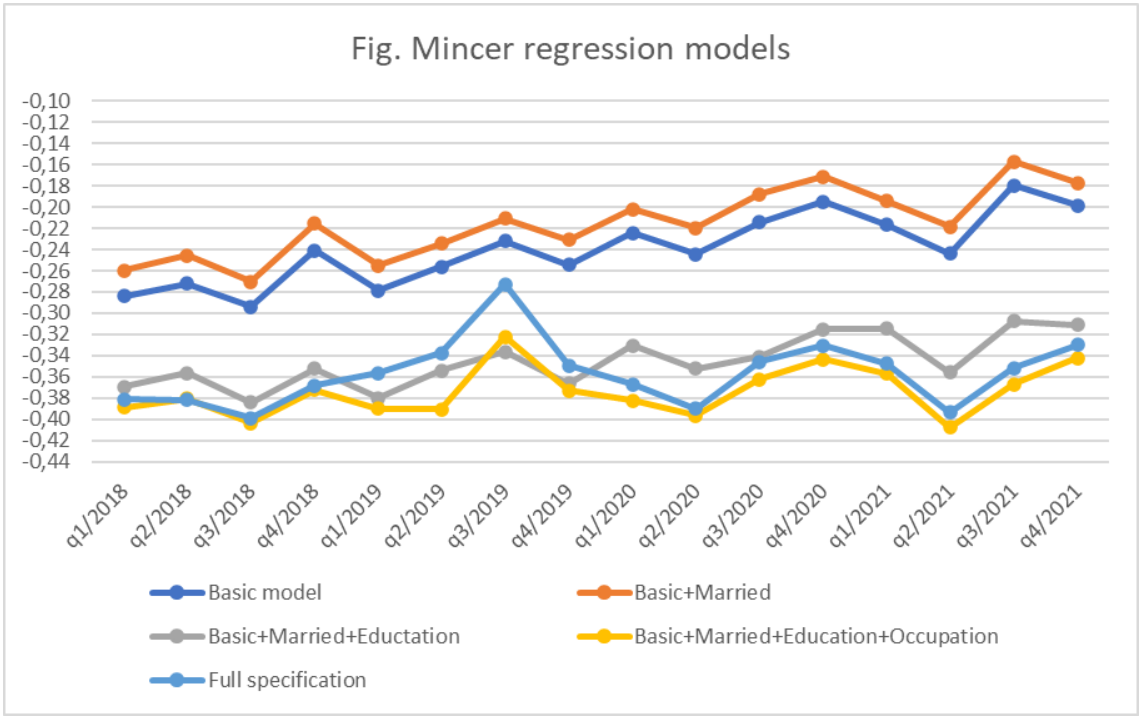
$$log(HourWage) \sim Female + Age + Age^2 \quad (8)$$

$$log(HourWage) \sim Female + Age + Age^2 + Married + High\ School\ diploma + University\ degree + Occupation\ dummies + Industry\ dummies \quad (9)$$

Figure 3 plots the results of the Mincer regression models. In the estimated results, the reference category observes, from education dummy variables, people with less than a high school diploma. Looking at the dummy variables for occupation, in the reference category are included all agricultural professions and industry, it is sector that groups accommodation and gastronomy services. From the plotted values in Fig. 3, we observe a trend in improving the raw GWG, as the coefficient for the *Female* variable in the *Basic model* and the model controlling for marital status is increasing over time. In models with more control variables, the trend is not that clear. To answer the research question no.2, I conclude that the gender wage gap is changing over time. For raw GWG, we see the trend in improving the gender wage difference (i.e., closing the GWG), but for adjusted GWG, it is not changing in a clear direction. We see some sharp increases, as well as a sudden decrease in the value of the coefficient.

In Fig. 3, another aspect of the Mincer models emerges: the raw GWG given by the *Basic model* and even when controlling for marital status is significantly lower than the rest of the models. The coefficient for the 1st quarter of 2018 in the *Basic model* equals -0.2839, while for the *Fully specified model,* it is -0.3761. When I start controlling for the *Education level*, the model immediately jumps to lower values (-0.3695) in the 1st quarter of 2018, and I observe the lowest value (-0,3891) in model, which controls for *Occupation status*, but not for *Industry sectors*. This trend holds through the whole inspected period. Looking at the periods affected by the pandemic crisis, the gender wage gap for the 3rd quarter of 2020 in the *Basic model* is estimated to be -0,2143. When controlling for *Education*, the GWG fall to -0,3431 and in the *Fully specified model,* the GWG is similar (-0,346). Table 15 (in the Appendix) provides an overview of the estimated GWG.

Figure 3 The gender wage gap estimated using Mincer regression models



Source: Author

The Mincer regressions suggest the level of education plays an important role in the gender wage gap problem. When considering men and women with the same schooling, the GWG increases by almost 10 logarithmic points. It points out a possible problem with not rewarding according to someone's qualification. Gender specialization can stand behind this difference, too. In the models, I control only for the level of education; the fields of education are not the main point of interest. Assuming that most of the

employees work in the field of their studies, then the education focus will be captured by *Occupation* and *Industry* variables. It explains not such a significant difference between a *Fully specified model* and a model observing only *Marital status* and *Education*. It causes situation like a woman, who has a university degree, but she is working in "low-wage" positions, e.g., teachers, will earn a lower hourly wage than a man with a university degree, who is working in some managerial positions or as IT specialists. Tables showing the gender distribution across various occupations and industries are provided in the Appendix. The gender distribution within occupation as Teachers, IT & Data Scientist or Doctors is dominated by one gender. In case of teachers more than 75% of positions are occupied by women. These facts could help policymakers to be orientated in the right direction.

In the models, which controls for personal and work characteristics, the pattern of development of the gender wage gap is like in the models, which control only for personal characteristics. The only difference is that the absolute value of GWG is lower in the models with more control variables. It does not give the exact reason for the wage difference, but we can draw some conclusions from the existing literature. Even though there are no sector without representation of both genders; the share of the genders differs. The authors have already proved that women prefer different sectors or occupations than men (Babcock and Laschever, 2003; Goldin, 2014; Alon et al., 2020). However, as the estimates of GWG have the same behaviour only on lower level, I may argue it might be due to other reasons, e.g., gender discrimination or gender specialization. The Mincer regression is not enough to decide, what leads to a difference in wages between men and women; there is a need for the determinants GWG.

As a method to observe determinants of the GWG, I choose to work with the Blinder-Oaxaca decomposition. At first, I proceed with model regressed only on *age* and age^2, in the form specified in Eq.10. The *Gender* variable after the vertical bar represent an indicator variable, for which I want to estimate the difference (Hlaváč, 2022). Similarly, as with Mincer regression models, I started with model as in Eq.10 and I continued with adding one by one other explanatory variables, until I obtain the *Fully specified model* as in Eq.11.

$$log(HourWage) \sim Age + Age^2 | Gender$$

(10)

$$log(HourWage) \sim Age + Age^2 + Married + High\ School\ diploma + University\ degree + Occupation\ dummies + Industry\ dummies | Gender$$

(11)

This gradual adjustment of the decomposed model is shown in table 10 to table 14 in the Appendix. In the 1st model, I try to explain the gender wage difference only by the squared function of age of employees as the approximation of their work experience; the answer is expected that most of the gap cannot be explained. The explained part takes negative values except for the one model. Those negative values imply that women's endowments in human capital should favour women against men. It agrees with the theory of women being more endowed than men. Nevertheless, there is a model, specified in eq. 12, where the endowments cause the GWG to rise (in positive values). When we control only for working experience and marital status, women are penalized for being married. There are many possible interpretations. One is that single women are expected to be more productive due to fewer personal commitments. The explanation accounts for married women, there is assumption put on them to become mothers in a short time, which implies a decrease in productivity and a rise in risk aversion, with new responsibilities towards their child.

$$log(HourWage) \sim Age + Age^2 + Married | Gender$$

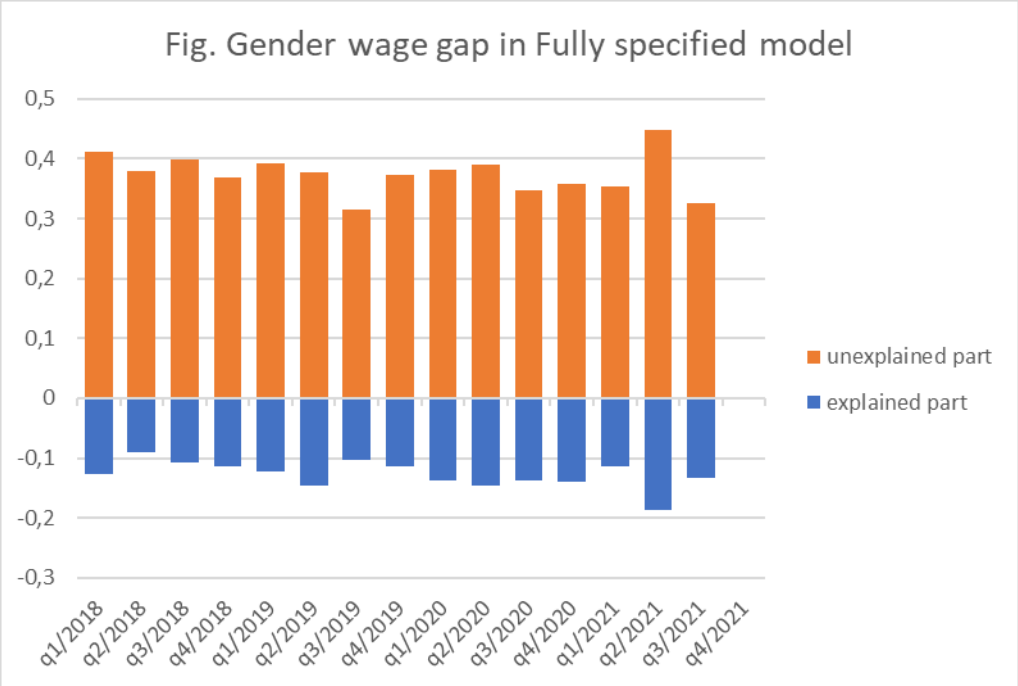
(12)

Fig. 4 provides a bar chart showing the GWG for the *Fully specified* model as a sum of what can be explained by the difference in endowments and the unexplained part. The explained part in the whole observed period has a negative sign. It confirms, what has been already observed, women are better endowed in human capital than men. The whole gap is the sum of explained and unexplained parts. By comparing the bars' height, we can state that the unexplained part plays a much more significant role in the total wage gap than the explained part. In the 3rd quarter of 2019, the explained part represents -48% of the total gap, and the unexplained part stands for 148% of the total gap. Looking at another example, the maximum absolute value of the explained part of

the GWG I measured in the 2nd quarter of 2021 was -0,18612 (-71% of the gap); also, the unexplained part was large (0,44737), which takes for 171% of the gap.

The values gathered from the B-O decomposition can be compared with the estimates from Mincer regression models. The gender wage gap estimated using B-O decomposition is lower than the estimates given by the Mincer regression model, but it needs to be stressed that the unexplained part has values similar to the *Fully specified* Mincer type model (Fig. 8, in the Appendix). Fig. 8 shows the highest present unexplained part of the GWG for the 2nd quarter of 2021. It is the period with the strictest pandemic restrictions, with many people losing their jobs, switching from full-time to part-time or due to their "second-shift" leaving the labour market. It, again, would be primarily low-income employees and single mothers, which causes the GWG to fall into low-lying values.

Figure 4 The decomposition of the gender wage gap in the Fully specified model



Source: Author

To provide empirical support for these arguments, we can look at how the coefficients of variables evolve in time for a model with *full specification*. I have already explained that even though the total GWG is getting smaller over time and, at the same time, the explained part rises, the unexplained part of the GWG plays a dominant role in the total GWG (it makes about 2/3 of the whole GWG). To evaluate the "key players" from all

the variables describing human capital, I calculated ratios of coefficient for the unexplained part to total GWG produced by the variable for each variable in all 16 time periods, table 16 to table 19 in the Appendix.

In the overview of the GWG and how the explained and unexplained parts were evolving, the possible penalization of women for being married is present as it is the only model where the explained part took strictly positive values. From the overview of all variables, I claim the role of marital status does not change over time, even after the pandemic restrictions are applied. This may be important for the mindset of employers about the prejudices put on married (or single) women.

The comparison of the Mincer regression models shows the importance of education level when determining the GWG. The B-O decomposition can evaluate what role has the education level. It tells us that until the COVID-19 pandemic, the education level was one of the main determinants of the gap, but with the arrival of the pandemic crisis, education has lost its power. In the case of the people with a university degree, the ratio has kept stable. However, the situation is different for the "High school diploma" category. The ratio was moving around one even before the pandemic, but it changed rapidly. The ratio takes more than a hundredfold from the 2nd to the 3rd quarter of 2020. Moving to the occupation dummy variables might help to paint a bigger picture about how COVID-19 change the determinants of the gender wage gap. The unexplained gap still accounts for more than half of the gender wage gap, but the COVID-19 pandemic crisis does not cause the fatal change.

Initially, I talked a lot about my expectation of what the COVID-19 pandemic crisis would bring us. I expected that low-income employees would be affected far more than those with an hourly wage above the average. It is one of my explanations for developing the average hourly wage. The Blinder-Oaxaca decomposition confirms the assumptions. In tables 16-19, with ratios of the unexplained part to the total gap for each variable, the values for low-paid occupations are changing heavily. For many people working manually, e.g., Manual workers, Maids and janitors or Mechanics and repairers, the change is glaringly apparent. Even before the pandemic, the unexplained part took a large share of the GWG, but after the restrictions were applied, the ratios were far too big, for some periods even more than 100.

The other aspect of the gap ratio is that the ratio inside each occupation's category is not similarly changing in all occupations. There are low-paid categories, who carry the burden of the pandemic crisis, and high-paid working positions, e.g., lawyers and financial specialists. As Goldin (2014) stresses in her work, discrimination is present within the occupation. My data confirms this statement. We can examine the financial specialist's category when looking for empirical evidence. The gender distribution there is stable in the researched time period and accounts for the ratio of men to women 2:3. Women slightly dominate but are not that strong. Although as for GWG, the ratio of unexplained to total gap was negative throughout the whole year 2018, saying that women are positively discriminated against men, with the pandemic crisis coming, it changes to positive 2,45. Suddenly, the unexplained part of the wage gap moved women into lower-paid groups.

The previous discussion allows me to answer the rest of the proposed researched questions. Research question no. 3 referred to whether determinants of GWG are changing in time. The answer is "Yes, the determinants of the gender wage gap are changing in time". We see those changes in tables 16-19 in the Appendix, and the previous example with education levels maps the situation too. Another example of change in determinants of wage is the improvement of coefficients in time and different models. While in the model, which controls only for age, working experience and marital status, the explained gap is strictly positive and causes women to be penalized for being married. In the *fully specified model*, the married coefficients are stable over time, accounting for about 10-20% of the total gap.

On the other hand, as we see, not only the occupation dummy variables are changing their importance and share of the total gap, but for the industry category, it holds as well. The unexplained gap gets even to favour women for those working in the public administration sector. The administrative workers move from the category of the sectors with heavy discrimination against women to the exact opposite. What might be behind this is a change in the functional requirements. Before the pandemic began, public administration was the sector with strict working hours and the requirement of physical presence at work. It caused many women to work only part-time with a lower income scheme, as they also had to manage their households as their "second shift" (Lichard et al., 2021). The pandemic crisis brings to this sector flexibility, availability of home-

office working or flexible work time. Many women become able to work full-time, and it increases their hourly wage. These are arguments to support my answer to the seventh research question. The answer is: "Yes, the determinants of the GWG change after the COVID-19 pandemic." The direction of change is not united. The positive change is in the example above. On the other hand, sectors with any flexibility available moved in the opposite direction. These sectors are science and manufacturing. In the case of science, many natural science experiments require the physical presence of the researcher and lots of his time, as women had to take care of their families and maintain personal life with close schools/day-cares, they were forced to leave their jobs or switch to part-time.

The remaining question is whether a change in the GWG was due to a difference in endowments or other factors. The answer is "No, the GWG does not change due to the difference in endowments of women." Even though Fig. 4 depicts a rise in the explained part of the GWG, the unexplained part is responsible for a more significant piece of the total gap. I have already mentioned changes that the pandemic restrictions put on the labour market. Though few of them work in favour of women (e.g., flexible working time), many make working even harder for women. The pandemic restrictions denying free movement, requesting shorter opening hours or closure of schools caused the labour market to shrink and require women to prioritize their time between work and their role as mothers. All the mention has the effect of change of the overall gender wage gap and its determinants and leads many authors to refer to this period as "shecession" as women are those who will mainly carry the burden of the COVID-19 pandemic (Alon et al., 2020; Bonacini et al., 2021; Fortier, 2020).

Conclusion

This thesis aimed to analyse the impact of the COVID-19 pandemic on the gender wage gap and whether the determinants of the GWG changed after the pandemic outbreak.

To achieve this thesis's goal, it was necessary to have the latest data possible. Data's availability and time relevance encouraged me to work with data capturing the situation in the USA. The data are spread over four cross-sectional datasets, separate for each year in the period from 2018 to 2021. To better represent the pandemic effect and put a more evident border for the pandemic events, I divided every year into four quarters.

To analyse the influence of the COVID-19 restrictions on the GWG, a detailed understanding of the process of how the GWG emerged, what it is and how it is measured was needed. For the evaluation of people's earnings, I chose to work with an hourly wage, which I adjusted for inflation and transformed to the value of US dollars in 2019.

Before focusing on the determinants of the GWG, I have to know how the GWG is changing in time. For this purpose, I use the famous Mincer regression, where the logarithmic transformation of the income variable (i.e., in my case, hourly wage) is regressed on various groups of variables depicting the human capital characteristics with one control variable for the group we want to compare, for the purposes of this thesis it is gender variable – female. I find out the average hourly wage is rising over time, and the GWG is closing.

As the next step, I pursued my research to another level, where I examined the determinants of the GWG in time and how its role changed after the COVID-19 outbreak. To explore the determinants of GWG, I used the Blinder-Oaxaca decomposition – method, which uses the difference at mean to decompose the wage into explained and unexplained parts. The results show that the gap is closing in time. It is due to the increase in human capital endowments of women and not the effect of improving the labour market condition. The unexplained gap is still present and even accounts for a more significant part of the total GWG.

To answer whether determinants of the GWG changed, the result of my analysis suggests that determinants have changed. While in pre-pandemic years, the education level took a significant role in explaining the GWG, after the COVID-19 outbreak, it lost its dominant role. The explained part is more defined by the differences within the occupation and different industries. Changes in the labour market have culminated in more flexible working conditions in some industries. On the other hand, it forced women to prioritise their time between family and work.

The effect of the COVID-19 pandemic on the gender wage gap appears to be highly heterogeneous concerning specific industries or occupations, whether we look at the low-income employees or not and existing research provided on specific groups of employees. Therefore, the contribution of this thesis lies in the broad spectrum of

observations across different groups of people and its time relevance with the analysis of the main drivers after the pandemic began.

Furthermore, this thesis accounts for the data capturing the latest period as possible up until the end of 2021, while many already published studies are focused mostly on the 2nd part of 2020 or the 1st quarter of 2021. It also provides the effect of a vast and highly specified field of occupations and industries, which might help address future policies to equalise the role of men and women in society, even in the Czech Republic.

The possible issue of this thesis is the lack of information about the parental status of employees and whether they are living with their partners in a shared household or not. Due to the lack of this information, all the explanations relating to the role of women as mothers and their "second shift" at home are only suggestions based on the results of previous studies. It would help to connect the drivers of the GWG within a specific group of employees. It may be an interesting suggestion for possible future extension of the research.

List of figures

List of tables

Acronyms

ACS	American Community Survey
BA	Bachelor of Arts
B-O	Blinder-Oaxaca
CPI	Consumer Price Index
CPS	Current Population Survey
ECHPS	European Community Household Panel Survey
EU	European Union
GWG	gender wage gap
IPUMS	Integrated Public Use Microdata Series
JMP	Juhn, Murphy and Pierce
NELS88	National Education Longitudinal Study of 1988/1994
NSL72	National Longitudinal Study of the High School Class of 1972
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
ORG	Outgoing Rotation Group
PSID	Panel Study for Income Dynamic
USA	United States of America
USD	United States Dollar
WFH	working from home

List of References

Alon, T., Doepke, M., Olmstead-Rumsey, J., & Tertilt, M. (2020). *The impact of COVID-19 on gender equality* (No. w26947). National Bureau of economic research.

Babcock, L., Laschever, S., & Don't Ask, W. (2003). Negotiation and the Gender Divide. *Princeton, NJ: Princeton University*.

Babcock, L., & Laschever, S. (2009). Women don't ask. In *Women Don't Ask*. Princeton University Press.

Becker, G. (1964). Gary S. Becker (1964, 1993,). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.

Bertrand, M. (2004). From the invisible handshake to the invisible hand? How import competition changes the employment relationship. *Journal of Labor Economics*, *22*(4), 723-765.

Bertrand, M. (2018). Coase lecture–the glass ceiling. *Economica*, *85*(338), 205-231.

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

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436-455.

Bonacini, L., Gallo, G., & Scicchitano, S. (2021). *Will it be a shecession? The unintended influence of working from home on the gender wage gap related to the COVID-19 pandemic* (No. 771). GLO Discussion Paper.

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.

Forbes. (2020). 45 U.S. States Shut Down And Counting: State-By-State Travel Restrictions. https://www.forbes.com/sites/suzannerowankelleher/2020/03/28/23-states-shut-down-and-counting-state-by-state-travel-restrictions/?sh=332de43758f4

Fortier, N. (2020). COVID-19, gender inequality, and the responsibility of the state. *International Journal of Wellbeing*, *10*(3).

Fortin, N. M. (2008). The gender wage gap among young adults in the united states the importance of money versus people. *Journal of Human Resources*, *43*(4), 884-918.

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

Hlavac, Marek (2022). oaxaca: Blinder-Oaxaca Decomposition in R. R package version 0.1.5. https://CRAN.R-project.org/package=oaxaca

IPUMS - USA. USA IPUMS Current Population Survey https://usa.ipums.org/usa/voliii/1_intro_ms5.shtml

Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, *8*(4), 453-479.

Juhn, C., Murphy, K. M., Topel, R. H., Yellen, J. L., & Baily, M. N. (1991). Why has the natural rate of unemployment increased over time?. *Brookings Papers on Economic Activity*, *1991*(2), 75-142.

Kunze, A. (2008). Gender wage gap studies: consistency and decomposition. *Empirical Economics*, *35*(1), 63-76.

Lazear, E. P., & Rosen, S. (1990). Male-female wage differentials in job ladders. *Journal of Labor Economics*, *8*(1, Part 2), S106-S123.

Lichard, T., Pertold, F., & Škoda, S. (2021). Do women face a glass ceiling at home? The division of household labor among dual-earner couples. *Review of Economics of the Household*, *19*(4), 1209-1243.

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

Mincer, J., & Polachek, S. (1974). Family investments in human capital: Earnings of women. *Journal of political Economy*, *82*(2, Part 2), S76-S108.

Navarová, T. (2022). International gender wage gap: A Meta-analysis.

Neumark, D. 1988. Employers' discriminatory behavior and the estimation of wage discrimination. Journal of Human Resources 23: 279–295.

Olivetti, C., & Petrongolo, B. (2008). Unequal pay or unequal employment? A cross-country analysis of gender gaps. *Journal of Labor Economics*, *26*(4), 621-654.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693-709.

OECD. (2019). Employment : *Distribution of employment by aggregate sectors, by sex*. https://stats.oecd.org/index.aspx?queryid=54755

Tertilt, M., Doepke, M., Alon, T., & Olmstead-Rumsey, J. (2020). This time it's different: The role of women's employment in a pandemic recession.

Thomason, B., & Macias-Alonso, I. (2020). COVID-19 and raising the value of care. *Gender, Work & Organization*, *27*(5), 705-708.

U. S. Bureau of Labor Statistics. (2022). Labor Force Statistics from the Current Population Survey; Earnings. https://www.bls.gov/cps/earnings.htm

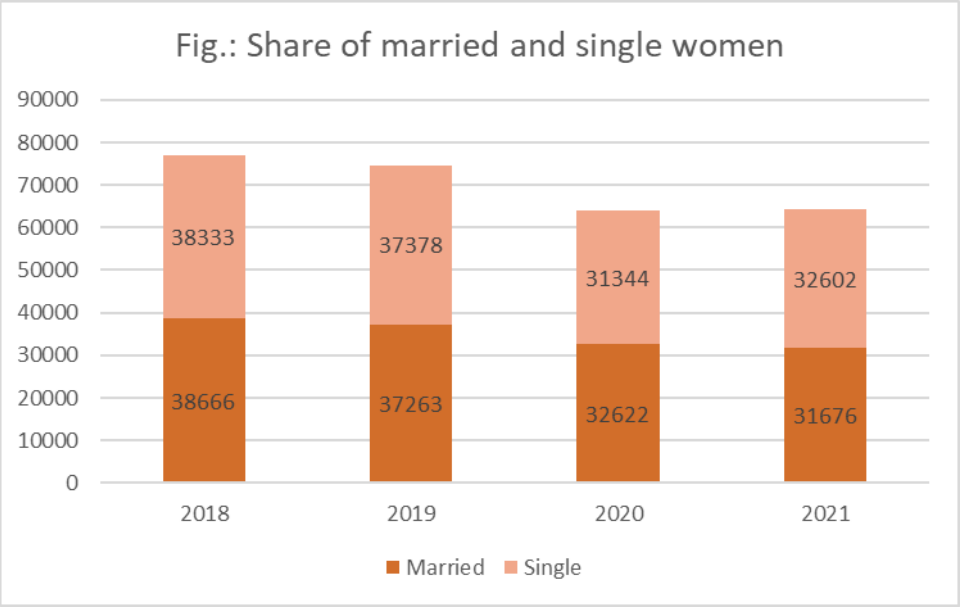
Willis, R. J. (1986). Wage determinants: A survey and reinterpretation of human capital earnings functions. *Handbook of labor economics*, *1*, 525-602.

Young, I. (1990). 1 990 Justice and the politics of difference Princeton University Press. *Princeton NJ*.

Young, I.M. (2011). Responsibility for justice. New York, USA: Oxford University Press.

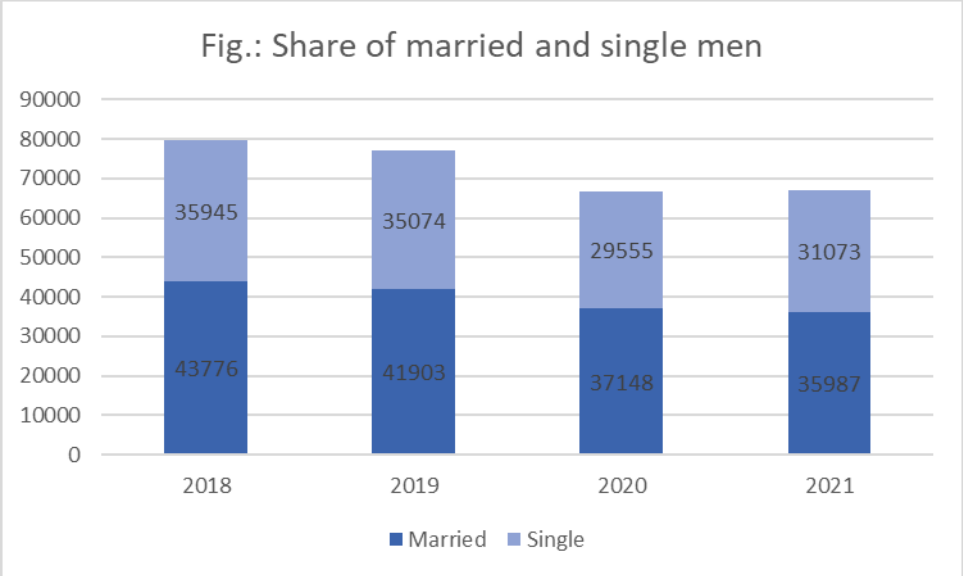
Appendix

Figure 5 Share of married and single women across datasets



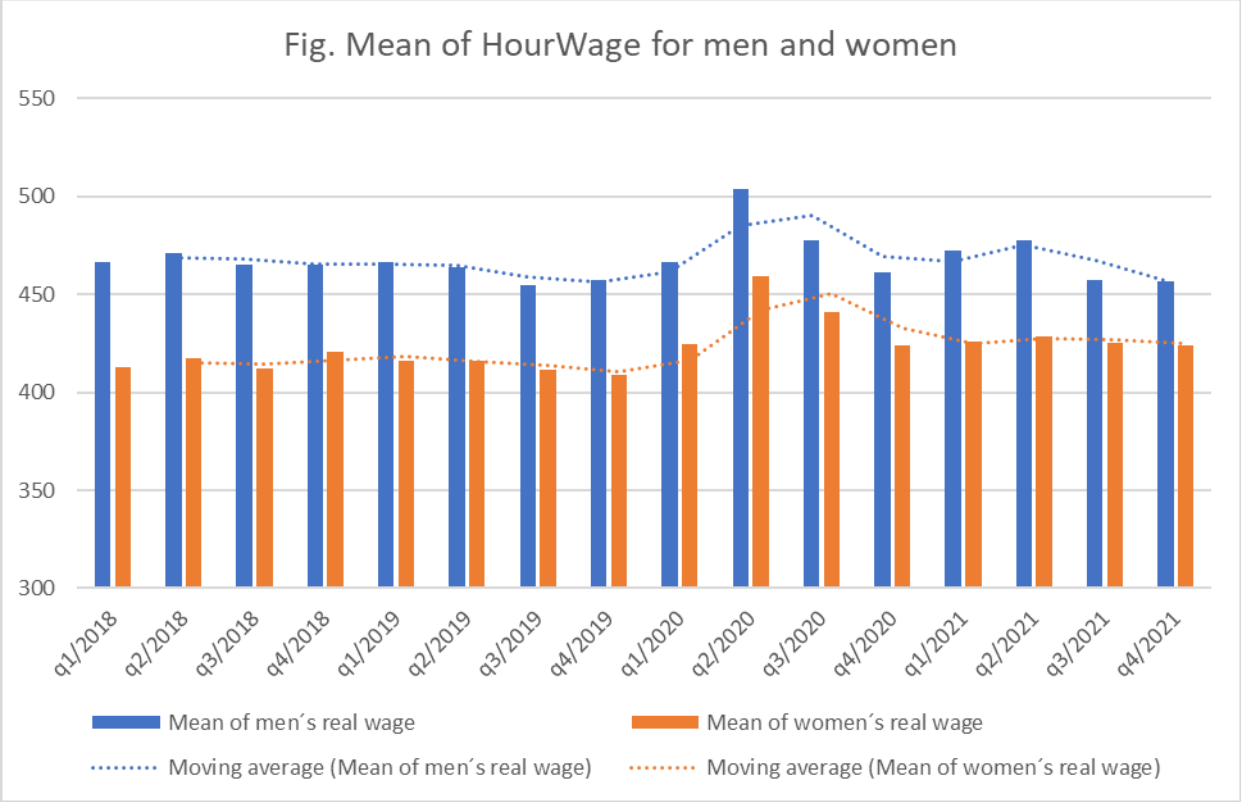
Source: Author

Figure 6 Share of married and single men across datasets



Source: Author

Figure 7 Mean of the hourly wage of men ad women over time



Source: Author

Table 5 Gender distribution across various industries over time

Table 5 - Gender distribution across various industries over time

	2018		2019		2020		2021	
	Men	Women	Men	Women	Men	Women	Men	Women
Education	4782	11045	4584	10799	3950	9470	3899	9251
Health Care	4826	18249	4841	17920	4126	15463	4166	15238
Real Estate and Leasing	270	122	236	111	1019	1024	1041	1060
Finance and Insurance	4101	5591	4025	5520	2847	3945	2796	3799
Manufacturing	11457	4662	11138	4531	9406	3957	9364	3901
Accomodation and Food	5080	6332	5029	6139	3703	4361	3975	4787
Professional and Scientific Services	5984	4782	5987	4759	5629	4250	5578	4354
Information	1759	1206	1524	1113	1287	957	1305	952

Source: Author

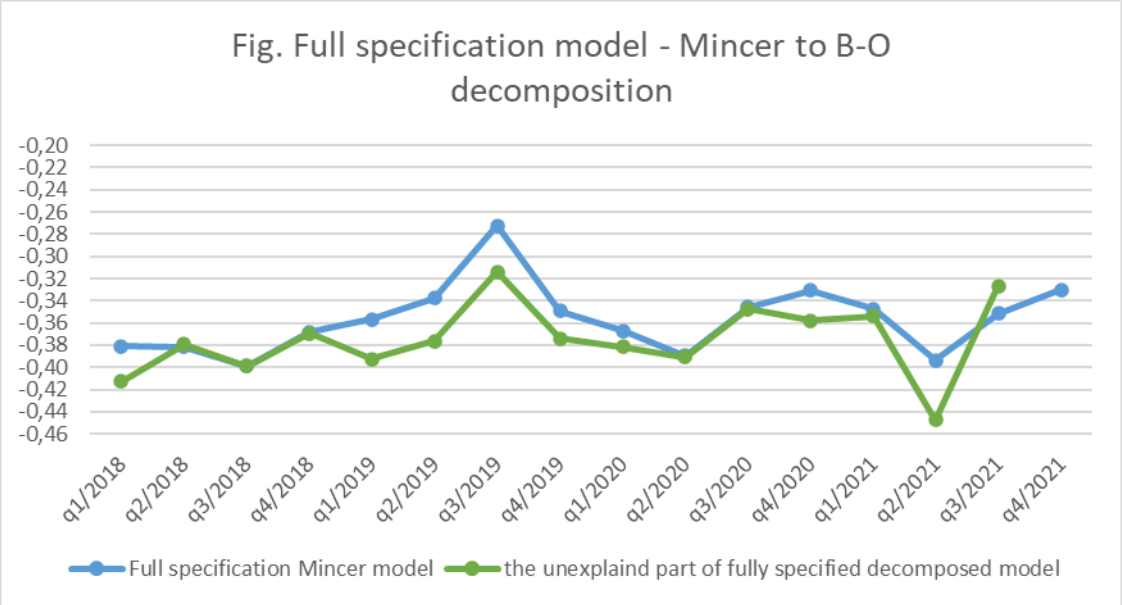
Table 6 Gender distribution across various occupations over time

Table 6 - Gender distribution across various occupations over time

	2018		2019		2020		2021	
	Men	Women	Men	Women	Men	Women	Men	Women
Teachers	1855	7144	1849	6898	2266	6671	2197	6460
Doctors	2391	9780	2431	9707	2103	7071	2058	6952
Gastronomy workers	3827	5104	3816	4954	2713	3557	2993	3832
Managers	9111	7362	8894	7219	8015	6575	7864	6552
Financial specialists	3299	4382	3280	4300	3273	4267	3362	4403
IT&Data scientists	3963	1437	3887	1442	3619	1279	3588	1334
Counselors	708	1325	700	1280	786	1812	819	1742
People in Sales	7242	7741	6838	7140	5825	5943	5676	6007
Maids&janitors	3210	2519	3011	2445	2358	1663	2381	1849

Source: Author

Figure 8 Comparison of the unexplained part of the fully specified decomposed model with the Female coefficient of the fully specified Mincer regression model



Source: Author

Table 7 Overview of the models (Fortin, 2008)

Table 3a
Impact of Noncognitive Traits on 1979 Log Wage (Age 25)—NLS72

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variables						
Female	−0.237**	−0.220**	−0.211**	−0.203**	−0.206**	−0.189**
	(0.010)	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)
Noncognitive traits						
Self-esteem		0.100**	0.080**		0.054**	0.043**
		(0.019)	(0.019)		(0.019)	(0.019)
External locus		−0.099**	−0.063**		−0.025	−0.036
of control		(0.024)	(0.024)		(0.024)	(0.024)
Money/work		0.221**	0.232**			0.209**
very important		(0.030)	(0.030)			(0.030)
People/family		−0.151**	−0.087**			−0.083**
very important		(0.022)	(0.028)			(0.028)
Cognitive skills						
High school math			0.051**	0.035**	0.033**	0.035**
Standardized score			(0.006)	(0.006)	(0.006)	(0.006)
Education (high school omitted)						
Less then high school				−0.090	−0.094	−0.085
				(0.081)	(0.081)	(0.080)
Trade or vocational				0.051**	0.053*	0.048**
				(0.019)	(0.019)	(0.019)
Some college				0.052**	0.051**	0.044**
				(0.014)	(0.014)	(0.014)
College				0.144**	0.140**	0.128**
				(0.017)	(0.018)	(0.018)
Postgraduate				0.202**	0.196**	0.179**
				(0.022)	(0.029)	(0.029)
Experience						
Years				0.039*	0.037**	0.036**
				(0.004)	(0.004)	(0.004)
Tenure				0.021**	0.021**	0.021**
				(0.003)	(0.003)	(0.003)
Personal characteristics						
Part-time				−0.030	−0.033	−0.029
				(0.020)	(0.020)	(0.020)
Black				−0.036	−0.034	−0.041
				(0.021)	(0.021)	(0.021)
Married				0.038**	0.036**	0.039**
				(0.011)	(0.011)	(0.011)
Parental experience				−0.003	−0.003	−0.003
				(0.003)	(0.003)	(0.003)
Child				0.006	0.007	0.009
				(0.017)	(0.017)	(0.017)
Adjusted R-2	0.076	0.093	0.104	0.129	0.131	0.139
Number of observations	6,184	6,184	6,184	6,184	6,184	6,184

Note: Standard errors are in parentheses. Significance at 5 percent level denoted by **; significance at 10 percent level denoted by *.

Source: Fortin, N. M. (2008). The gender wage gap among young adults in the united states the importance of money versus people. Journal of Human Resources, 43(4), 884-918.

Table 8 Overview of the models (Fortin, 2008)

Table 3b
Impact of Noncognitive Traits on 1986 Log Wage (Age 32)—NLS72

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variables						
Female	−0.229**	−0.197**	−0.187**	−0.188**	−0.189**	−0.172**
	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)
Noncognitive traits						
Self-esteem		0.200**	0.168**		0.125**	0.118**
		(0.054)	(0.054)		(0.054)	(0.054)
External locus of		−0.162**	−0.118**		−0.058	−0.070
control		(0.044)	(0.044)		(0.044)	(0.044)
Money/work		0.328**	0.334**			0.258**
very important		(0.047)	(0.047)			(0.047)
People/family		−0.151**	−0.066			−0.025
very important		(0.022)	(0.009)			(0.043)
Cognitive skills						
High school math			0.060**	0.011	0.008	0.014
Standardized score			(0.009)	(0.006)	(0.006)	(0.006)
Education (high school omitted)						
Less then high school				−0.160	−0.160	−0.138
				(0.155)	(0.155)	(0.155)
Trade or vocational				0.031	0.028	0.023
				(0.027)	(0.027)	(0.027)
Some college				0.094**	0.089**	0.081**
				(0.023)	(0.023)	(0.023)
College				0.210**	0.203**	0.189**
				(0.027)	(0.027)	(0.027)
Postgraduate				0.316**	0.306**	0.289**
				(0.034)	(0.034)	(0.034)
Experience						
Years				0.008**	0.007*	0.006*
				(0.003)	(0.003)	(0.003)
Tenure				0.019**	0.019**	0.019**
				(0.002)	(0.002)	(0.002)
Personal characteristics						
Part-time				0.039	0.036	0.048
				(0.028)	(0.028)	(0.028)
Black				0.018	0.017	0.011
				(0.031)	(0.031)	(0.032)
Married				0.019	0.013	0.012
				(0.019)	(0.019)	(0.019)
Parental experience				−0.010**	−0.010**	−0.010**
				(0.002)	(0.002)	(0.002)
Child				0.047	0.046	0.048*
				(0.024)	(0.024)	(0.024)
Adjusted R-2	0.031	0.043	0.050	0.075	0.079	0.081
Number of observations	6,522	6,522	6,522	6,522	6,522	6,522

Note: Standard errors are in parentheses. Significance at 5 percent level denoted by **; significance at 10 percent level denoted by *.

Source: Fortin, N. M. (2008). The gender wage gap among young adults in the united states the importance of money versus people. Journal of Human Resources, 43(4), 884-918.

Table 9 Overview of the models (Fortin, 2008)

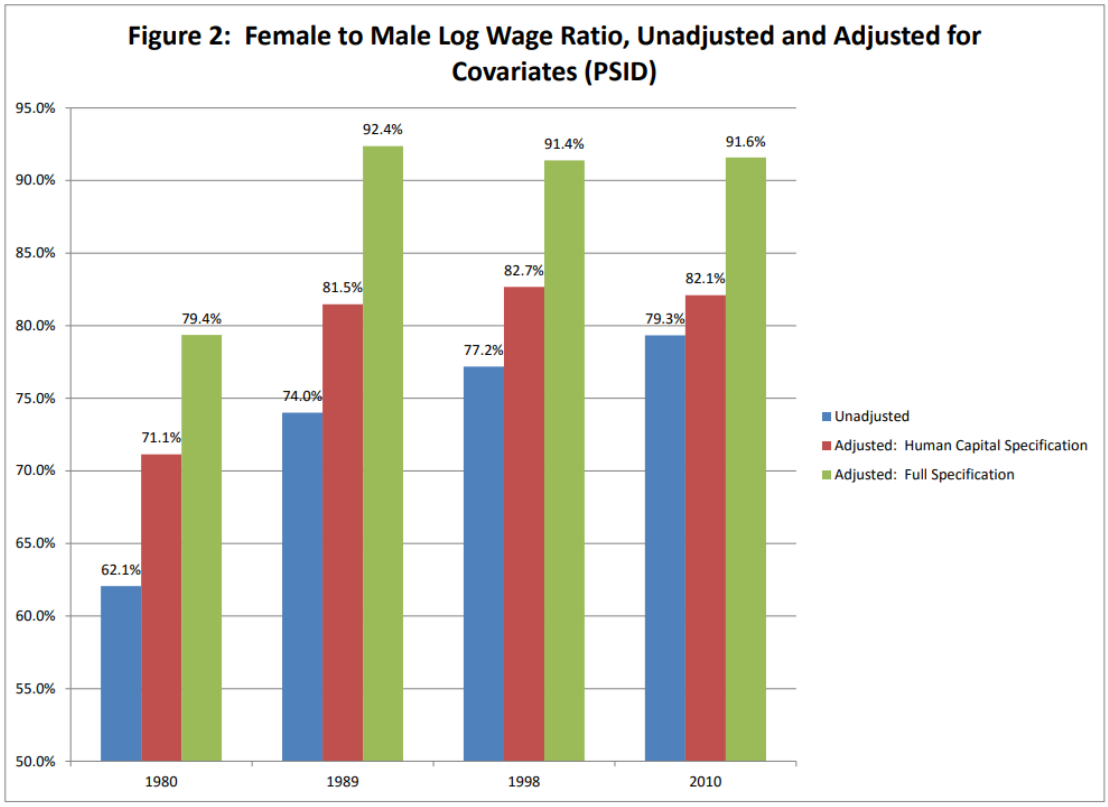
Table 3c
Impact of Noncognitive Traits on 2000 Log Wage (Age 24)—NELS88

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variables						
Female	−0.181**	−0.173**	−0.169**	−0.188**	−0.186**	−0.178**
	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.0171)
Noncognitive traits						
Self-esteem		0.054	0.069*		0.084**	0.082**
		(0.038)	(0.036)		(0.035)	(0.035)
External locus of		−0.207**	−0.081**		−0.035	−0.032
control		(0.054)	(0.054)		(0.024)	(0.024)
Money/work		0.151**	0.190**			0.165**
very important		(0.029)	(0.029)			(0.028)
People/family		−0.082**	−0.016			−0.019
very important		(0.023)	(0.023)			(0.022)
Cognitive skills						
High school math			0.120**	0.063**	0.060**	0.064**
Standardized score			(0.006)	(0.007)	(0.007)	(0.007)
Education (high school omitted)						
Less then high school				0.069	0.066	0.064
				(0.071)	(0.071)	(0.071)
Trade or vocational				0.084**	0.082**	0.086**
				(0.022)	(0.023)	(0.027)
Some college				0.072**	0.073**	0.068**
				(0.017)	(0.017)	(0.017)
College				0.248**	0.246**	0.238**
				(0.021)	(0.021)	(0.021)
Postgraduate				0.447**	0.443**	0.432**
				(0.036)	(0.036)	(0.036)
Experience						
Years				0.001	0.002	0.001
				(0.003)	(0.003)	(0.003)
Tenure				0.015**	0.015**	0.015**
				(0.003)	(0.003)	(0.003)
Personal characteristics						
Part-time				−0.207**	−0.206**	−0.202**
				(0.018)	(0.018)	(0.018)
Black				−0.036**	−0.035*	0.011*
				(0.018)	(0.024)	(0.032)
Married				0.055**	0.054**	0.058**
				(0.011)	(0.012)	(0.012)
Parental experience				0.002	0.002	0.001
				(0.003)	(0.003)	(0.003)
Child				−0.052**	−0.052**	−0.052**
				(0.017)	(0.017)	(0.017)
Adjusted R-2	0.036	0.053	0.111	0.167	0.168	0.172
Number of observations	6,476	6,476	6,476	6,476	6,476	6,476

Note: Standard errors are in parentheses. Significance at 5 percent level denoted by **; significance at 10 percent level denoted by *.

Source: Fortin, N. M. (2008). The gender wage gap among young adults in the united states the importance of money versus people. Journal of Human Resources, 43(4), 884-918.

Figure 9 Female to Male log wage ratio (Blau and Kahn, 2017)



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

Table 10 The Blinder-Oaxaca decomposition of the Basic model

Table 10 – B-O decomposition of Basic model

	the explained part	the unexplained part	GWG
q1/2018	-0,00938	0,29427	0,28489
q2/2018	0,00312	0,28376	0,28688
q3/2018	-0,00617	0,28982	0,28813
q4/2018	-0,00339	0,25995	0,25656
q1/2019	-0,0045	0,28966	0,28517
q2/2019	-0,00267	0,27013	0,26727
q3/2019	-0,00755	0,2525	0,24495
q4/2019	-0,00316	0,2766	0,27344
q1/2020	-0,00443	0,24811	0,24368
q2/2020	0,00019	0,24573	0,24592
q3/2020	-0,00837	0,21903	0,21066
q4/2020	-0,006	0,22584	0,2198
q1/2021	0,0069	0,24085	0,24261
q2/2021	0,00224	0,26412	0,26636
q3/2021	0,00946	0,18253	0,19199
q4/2021	-0,01057	0,20929	0,19872

Source: Author

Table 11 The Blinder-Oaxaca decomposition of the Basic model with addition variables

Table 11 - B-O decomposition of Basic model + Marital status

	the explained part	the unexplained part	GWG
q1/2018	0,01655	0,26834	0,28489
q2/2018	0,03329	0,25359	0,28688
q3/2018	0,02463	0,2635	0,28813
q4/2018	0,02486	0,2317	0,25656
q1/2019	0,02	0,26517	0,28517
q2/2019	0,0223	0,24516	0,26747
q3/2019	0,016	0,22894	0,24495
q4/2019	0,02385	0,24959	0,27343
q1/2020	0,02032	0,22336	0,24368
q2/2020	0,02814	0,21778	0,24592
q3/2020	0,02122	0,18944	0,21066
q4/2020	0,02197	0,19787	0,21984
q1/2021	0,02984	0,21277	0,24261
q2/2021	0,02805	0,23831	0,26636
q3/2021	0,03474	0,15724	0,192
q4/2021	0,01365	0,18507	0,19872

Source: Author

Table 12 The Blinder-Oaxaca decompostion of the Basic model with additional variables

Table 12 - B-O decomposition of Basic model+Marital status+Education

	the explained part	the unexplained part	GWG
q1/2018	-0,09175	0,37664	0,28389
q2/2018	-0,06801	0,35489	0,28688
q3/2018	-0,0842	0,37233	0,28813
q4/2018	-0,10211	0,35866	0,25656
q1/2019	-0,09986	0,38502	0,28517
q2/2019	-0,09459	0,36205	0,26747
q3/2019	-0,10466	0,34962	0,24495
q4/2019	-0,09993	0,37337	0,27344
q1/2020	-0,10644	0,35012	0,24368
q2/2020	-0,09806	0,34398	0,24592
q3/2020	-0,12321	0,33387	0,21066
q4/2020	-0,11176	0,3316	0,21984
q1/2021	-0,08897	0,33158	0,24261
q2/2021	-0,09629	0,36265	0,26636
q3/2021	-0,11329	0,30527	0,19199
q4/2021	-0,12165	0,32037	0,19872

Source: Author

Table 13 The Blinder-Oaxaca decomposition of Basic model with additional variables

Table 13 - B-O decomposition of Basic model+Marital status+Education+Occupation

	the explained part	the unexplained part	GWG
q1/2018	-0,13548	0,42037	0,28489
q2/2018	-0,09379	0,38067	0,28688
q3/2018	-0,11061	0,39874	0,28813
q4/2018	-0,12882	0,38538	0,25656
q1/2019	-0,12936	0,41452	0,28517
q2/2019	-0,13514	0,40261	0,26747
q3/2019	-0,1011	0,34605	0,24495
q4/2019	-0,11138	0,38482	0,27343
q1/2020	-0,14901	0,39269	0,24368
q2/2020	-0,16109	0,407	0,24592
q3/2020	-0,15859	0,36924	0,21066
q4/2020	-0,1494	0,36923	0,21983
q1/2021	-0,12137	0,36398	0,24261
q2/2021	-0,187	0,45336	0,26636
q3/2021	-0,14794	0,33993	0,19199
q4/2021	-0,15667	0,35539	0,19872

Source: Author

Table 14 The Blinder-Oaxaca decomposition for the Full specification model

Table 14 - Full specification model

	the explained part	the unexplained part	GWG
q1/2018	-0,12732	0,41252	0,2852
q2/2018	-0,09082	0,37941	0,28859
q3/2018	-0,10816	0,39862	0,29046
q4/2018	-0,11327	0,36954	0,25627
q1/2019	-0,12234	0,39272	0,27038
q2/2019	-0,14577	0,37662	0,23085
q3/2019	-0,10303	0,31419	0,21116
q4/2019	-0,11379	0,37393	0,26014
q1/2020	-0,13765	0,38171	0,24406
q2/2020	-0,14702	0,39085	0,24383
q3/2020	-0,13753	0,34768	0,21015
q4/2020	-0,13888	0,358	0,21912
q1/2021	-0,114	0,35403	0,24003
q2/2021	-0,18612	0,44737	0,26125
q3/2021	-0,13409	0,32648	0,19239
q4/2021			

Source: Author

Table 15 Overview of the GWG in Mincer regression models

Table 10 - Overview of the GWG using Mincer regression

	1)	2)	3)	4)	5)
1st quarter of 2018	-0,28390	-0,2598	-0,3695	-0,3891	-0,3812
2nd quarter of 2018	-0,27206	-0,2454	-0,3565	-0,3809	-0,3816
3rd quarter of 2018	-0,29380	-0,2702	-0,3841	-0,4037	-0,3989
4th quarter of 2018	-0,24100	-0,2157	-0,3521	-0,3721	-0,368
1st quarter of 2019	-0,27885	-0,2553	-0,3802	-0,39	-0,3566
2nd quarter of 2019	-0,25620	-0,2341	-0,3542	-0,3909	-0,3378
3rd quarter of 2019	-0,23190	-0,2108	-0,3362	-0,3223	-0,2725
4th quarter of 2019	-0,25440	-0,2311	-0,3667	-0,3731	-0,3492
1st quarter of 2020	-0,22450	-0,2022	-0,3304	-0,3824	-0,3669
2nd quarter of 2020	-0,24430	-0,2194	-0,3524	-0,3967	-0,3898
3rd quarter of 2020	-0,21430	-0,188	-0,3413	-0,3621	-0,346
4th quarter of 2020	-0,19470	-0,1713	-0,315	-0,3434	-0,3305
1st quarter of 2021	-0,21660	-0,1944	-0,3146	-0,3571	-0,3477
2nd quarter of 2021	-0,24400	-0,2185	-0,3559	-0,408	-0,3938
3rd quarter of 2021	-0,17940	-0,1572	-0,3076	-0,3673	-0,3518
4th quarter of 2021	-0,19840	-0,1774	-0,3111	-0,3425	-0,3301

Note: col. 1) refers to the *Basic model*, where the logarithmic wage is regressed on the quadratic function of *age.* Col. 2) is the *Basic model* controlling for *marital status*. Col. 3) is column 2) with addition of education dummy variables. Col 4) is column 3) with the occupation dummy variables. Col. 5) is *Ful specification model*. It controls for quadratic function of age, marital status, education level, occupation category and industry sector.

Source: Author

Table 16 The ratio of the unexplained gap to the total gap obtained through the Blinder-Oaxaca decomposition for the fully specified model for 2018

The ratio of the unexplained gap to total gap for each variable in 2018

	q1/2018	q2/2018	q3/2018	q4/2018
age	1,056366	1,036936	1,021576	1,014331
age^2	1,113824	1,074874	1,042490	1,031722
Married	0,875139	0,860414	0,878502	0,858267
High school diploma	0,647537	1,223912	2,248605	0,805693
Uni degree	0,506971	0,725730	0,618201	0,179657
Artist/Athletes/Musicians/Writers/Journalists	0,215294	1,263314	0,580816	0,792838
Clerks	1,713918	1,566885	1,243172	3,224418
Counselours	-8,907261	0,199768	-0,037931	-9,151000
Doctors	-0,522221	0,572654	-0,546810	44,809264
Emergency employees	0,558522	0,514062	-1,406824	0,858191
Engineers	-0,074923	-0,635772	0,014964	-0,014217
Financial specialists	-0,449246	0,714891	-0,114043	-0,841751
Food producers/Welders/Machine workers	-0,040917	0,312449	0,239997	0,067489
Gastronomy workers	2,210825	1,533195	5,780610	0,183246
Health assistents	-0,914432	2,330612	3,340870	9,332403
IT&Data scientist	-0,144891	-0,512935	-0,171521	0,044069
Lawyers	0,000833	0,617936	-0,111727	4,883742
Maid&janitors	-1,898169	0,769903	0,757792	0,289934
Managers	-0,043107	1,629804	-0,099668	-1,517840
Manual workers	-0,005045	0,081176	0,196491	0,039072
Mechanics&repairers	0,116586	0,056680	0,054708	0,017662
People in sales	1,215247	0,820813	1,209554	1,162304
Scientists	0,184753	0,889589	0,887127	0,586014
Services	1,960699	1,786643	0,075653	-1,744311
Teachers	-0,313907	0,365878	-0,298450	-0,836918
Transport	0,037349	0,331909	-0,071147	0,059278
Administrative support	3,182014	1,079329	1,118053	1,920899
Agriculture	0,078611	1,902331	0,034066	0,244713
Art and Entertainment	0,965439	0,960971	1,025082	0,770650
Construction	-0,105287	-0,089637	-0,232254	-0,087167
Education	0,557972	0,762138	0,655233	0,702108
Finance and Insurance	-0,549812	0,450018	0,513645	0,442664
Health Care	-18,711005	1,151269	0,520066	0,673106
Information	3,626837	1,734334	-0,118504	-2,888134
Management Companies	1,063894	0,954273	1,063472	1,324799
Manufacturing	-1,998361	-15,862185	-11,645780	-0,931882
Mining	0,674172	0,625743	0,621466	0,488559
Other Services	0,680567	0,567460	0,394250	0,456035
Professional and Scientific Services	-0,405378	-1,324906	1,663916	1,637008
Public Administration	1,445985	1,218984	1,621980	1,165479
Real Estate and Leasing	1,196557	0,044994	0,830274	1,133461
Utilities	0,000000	0,000000	0,000000	0,000000
Wholesale Trade	0,122170	2,656470	-8,144000	-0,466390

Source: Author

Table 17 The ratio of the unexplained gap to the total gap obtained through the Blinder-Oaxaca decomposition for the fully specified model for 2019

The ratio of the unexplained gap to total gap for each variable in 2019

	q1/2019	q2/2019	q3/2019	q4/2019
age	1,064901	1,099682	1,057409	1,003842
age^2	1,152262	1,234252	1,116393	1,018869
Married	0,889823	0,877287	0,851243	0,878702
High school diploma	0,920534	0,860092	0,731153	0,898184
Uni degree	0,346447	0,571849	0,492620	0,209341
Artist/Athletes/Musicians/Writers/Journalists	-0,318752	0,941245	1,174499	1,495633
Clerks	1,031903	1,073650	0,994658	0,867515
Counselours	-1,508564	0,213257	0,045822	-0,146690
Doctors	-0,101134	0,508038	0,565826	-0,251308
Emergency employees	3,009133	0,427544	1,496348	-1,288226
Engineers	-0,088100	-0,063412	-0,164200	0,035061
Financial specialists	0,578485	0,576680	0,647913	-0,350692
Food producers/Welders/Machine workers	0,267522	0,306482	0,483857	-0,074325
Gastronomy workers	1,212395	1,175497	1,132484	1,327887
Health assistents	0,985862	0,988260	0,905155	1,924362
IT&Data scientist	-0,161014	-0,394000	-0,491205	-0,151418
Lawyers	-0,032997	0,809504	0,709785	3,270314
Maid&janitors	0,916120	0,797946	0,932149	0,327363
Managers	-2,543817	9,086767	2,822544	-0,154647
Manual workers	0,027415	0,094620	0,257867	0,106726
Mechanics&repairers	0,046347	0,041773	0,278930	0,081706
People in sales	0,546341	1,964673	1,445487	0,497521
Scientists	0,715926	0,762152	0,879314	0,779347
Services	0,687786	1,468248	1,055102	0,155855
Teachers	-0,281833	0,255610	0,340097	-0,246008
Transport	0,164562	0,178068	0,452686	0,035600
Administrative support	2,092371	1,198101	-1,215568	1,492381
Agriculture	-0,098767	6,423529	1,615354	-8,070523
Art and Entertainment	1,127486	0,704730	0,999015	1,299831
Construction	-0,119252	-0,134251	-0,112313	-0,208226
Education	0,566379	0,541061	0,565580	0,687424
Finance and Insurance	-0,449753	0,437518	0,405242	0,721095
Health Care	-0,455080	-0,169037	0,000290	1,182441
Information	0,254859	4,956006	-2,826980	1,380046
Management Companies	0,565483	1,132117	0,999300	0,990454
Manufacturing	-0,054395	-1,457151	-0,605918	-14,595840
Mining	0,134508	-0,124022	-0,256339	-0,862343
Other Services	0,540225	0,663859	0,162251	0,857616
Professional and Scientific Services	-0,303478	-11,565012	-9,095503	1,581260
Public Administration	1,354866	1,184730	2,672911	1,123051
Real Estate and Leasing	-4,925143	3,771502	0,624857	0,712751
Utilities	0,000000	0,000000	0,000000	0,000000
Wholesale Trade	-0,031890	-0,963630	-0,091640	-1,248710

Source: Author

Table 18 The ratio of the unexplained gap to the total gap obtained through the Blinder-Oaxaca decomposition for the fully specified model for 2020

The ratio of the unexplained gap to total gap for each variable in 2020

	q1/2020	q2/2020	q3/2020	q4/2020
age	1,042706	1,014443	1,038486	1,231677
age^2	1,113995	1,028092	1,075524	-1,335311
Married	0,881499	0,860828	0,878973	0,881758
High school diploma	0,886598	0,297923	12,124126	1,246152
Uni degree	0,143718	0,656389	0,524960	0,538879
Artist/Athletes/Musicians/Writers/Journalists	0,630304	1,510807	0,030982	0,825199
Clerks	0,389612	1,004186	0,688371	0,492032
Counselours	-0,634592	2,548987	-0,402286	-0,432694
Doctors	-0,333655	2,295185	-0,287356	0,009871
Emergency employees	0,021059	0,866561	0,765402	-0,031949
Engineers	0,047470	0,125497	-0,087037	-0,061203
Financial specialists	-12,862678	2,449550	0,114240	0,211242
Food producers/Welders/Machine workers	-0,585997	-0,996048	0,041091	0,188640
Gastronomy workers	0,871652	0,915489	0,727484	2,250179
Health assistents	-12,587570	0,554601	-5,291626	10,927249
IT&Data scientist	-0,028940	0,217267	-0,109776	-0,086388
Lawyers	15,717794	1,436106	-0,381181	-0,527095
Maid&janitors	0,782131	1,911480	0,428052	0,601327
Managers	0,017584	0,677257	-0,453554	-0,936845
Manual workers	0,183720	0,025546	-0,058929	0,024945
Mechanics&repairers	-0,007660	0,003397	0,404427	0,082275
People in sales	1,085110	1,061778	1,374752	1,185399
Scientists	0,448664	0,963738	4,714841	1,115148
Services	1,031254	-1,661347	0,529722	0,520644
Teachers	-0,770248	43,288800	-0,078261	-0,314042
Transport	0,172266	-0,620483	0,225018	0,214632
Administrative support	1,122631	0,273516	0,135781	-1,092402
Agriculture	0,173967	0,054428	0,130019	-0,245082
Art and Entertainment	0,856432	1,502638	0,930115	0,928084
Construction	-0,110684	-0,010325	-0,028191	-0,150619
Education	0,634181	0,515787	0,163703	0,599110
Finance and Insurance	0,406440	-0,062683	13,661569	0,348399
Health Care	0,153178	0,014981	-7,416090	0,438994
Information	1,464029	-0,416524	0,393355	4,923748
Management Companies	2,070879	0,345918	1,844895	1,032638
Manufacturing	-1,394500	-0,316817	0,132663	-0,324390
Mining	-0,078594	-0,067062	-0,496105	-0,478962
Other Services	-0,114098	0,177950	-4,243889	0,471288
Professional and Scientific Services	-9,903295	-0,938655	0,379233	-12,019217
Public Administration	1,264634	1,208495	0,365481	1,113014
Real Estate and Leasing	0,959050	0,950224	0,363185	1,137339
Utilities	0,675556	0,694180	0,727924	0,567007
Wholesale Trade	-0,720905	-0,391152	0,222832	-0,277844

Source: Author

Table 19 The ratio of the unexplained gap to the total gap obtained through the Blinder-Oaxaca decomposition for the fully specified model for 2020

The ratio of the unexplained gap to total gap for each variable in 2021

	q1/2021	q2/2021	q3/2021
Age	1,087347	1,033558	0,997163
age^2	0,595440	1,073242	-0,004170
Married	0,886982	0,847715	0,154853
High school diploma	-1,197292	0,569260	-0,954294
Uni degree	0,455970	0,609180	0,647656
Artist/Athletes/Musicians/Writers/Journalists	0,871897	1,192951	-0,505361
Clerks	0,612124	0,675564	92,530177
Counselours	3,767412	-1,038375	-1,789686
Doctors	3,530203	0,198132	-0,939122
Emergency employees	3,206427	0,586558	-1,090988
Engineers	0,068283	-0,090159	192,150686
Financial specialists	4,316393	0,347908	-0,958548
Food producers/Welders/Machine workers	-0,511776	0,171878	-8,436573
Gastronomy workers	0,568829	2,013393	0,379486
Health assistents	-0,404779	-2,753502	-6,935755
IT&Data scientist	0,122384	-0,053216	5,718223
Lawyers	-1,603901	-1,047652	-0,899282
Maid&janitors	1,844581	0,399863	-0,832512
Managers	0,540353	-3,873911	0,367368
Manual workers	-0,044565	0,073242	15,709799
Mechanics&repairers	0,023841	0,068340	382,627044
People in sales	1,054337	0,042735	-0,072903
Scientists	0,137329	0,995366	0,826687
Services	0,240558	0,703944	-0,730115
Teachers	-3,476470	-0,094683	-2,038693
Transport	-0,760912	0,251986	-7,802496
Administrative support	0,194630	-4,892948	0,842032
Agriculture	0,394124	0,078775	3,476943
Art and Entertainment	0,856429	0,904669	0,099370
Construction	-0,104334	-0,032676	28,660439
Education	0,597430	0,262711	-22,035284
Finance and Insurance	-0,020118	-0,130168	-1,148533
Health Care	-0,558264	-2,311905	-0,758672
Information	8,036554	0,551546	1,031090
Management Companies	0,393380	1,447562	0,035726
Manufacturing	-0,095960	-0,515708	12,899999
Mining	-0,186956	-0,008319	-19,128557
Other Services	0,078654	-1,115393	-0,569547
Professional and Scientific Services	-0,066305	5,537511	0,590196
Public Administration	0,156034	0,118144	-0,093167
Real Estate and Leasing	0,699686	1,018651	0,144654
Utilities	0,629431	0,756099	0,255500
Wholesale Trade	-0,330808	0,094603	5,324104

Source: Author

Appendix no. 1: Comparison of the voter‘s base (graph)

Appendix no. 2: Success rate of political parties in 1990-2010 (table)