Mariola Pytliková

Lenka Filipová

Jaromír Gottvald

Jiří Balcar

Petra Vašková

Zuzana Machová

Lenka Janíková

Dagmar Brožová

GENDER WAGE GAP AND DISCRIMINATION IN THE CZECH REPUBLIC

Ostrava, 2012

Mariola Pytliková[1,2], Lenka Filipová[1], Jaromír Gottvald[1], Jiří Balcar[1,3], Petra Vašková[1], Zuzana Machová[1], Lenka Janíková[1], Dagmar Brožová[1,4]

[1] Department of National Economics
 Faculty of Economics
 VŠB-Technical University Ostrava
 Sokolská 33
 701 21 Ostrava, CZ
 Marp@asb.dk, lenka.filipova@vsb.cz, jaromir.gottvald@vsb.cz, balcar@rpic-vip.cz, petra.vaskova@vsb.cz, zuzana.machova@vsb.cz, lenka.janikova@vsb.cz, brozova@isis.vse.cz

[2] Danish Institute of Governmental Research (KORA)
 Købmagergade 22
 1150 Copenhagen K, DK

[3] RPIC-ViP s.r.o.
 Výstavní 2224/8
 709 00 Ostrava, CZ

[4] Department of Economics
 Faculty of Economics
 University of Economics
 W. Churchilla Sq. 4
 130 67 Prague 3, CZ

Reviews

Anders Ekeland, Statistics Norway (SSB), Oslo, Norway
Martin Kahanec, Central European Labour Studies Institute (CELSI), Central European University in Budapest

The text should be cited as follows: Pytliková, M. et al. (2012). *Gender wage gap and discrimination in the Czech Republic*, SAEI, vol. 15. Ostrava: VŠB-TU Ostrava.

Foreword

Gender differentials on the labour market represent one of the most important and most discussed research topics in the field of labour economics. Early research emphasizes the role of human capital and discrimination in explaining gender differentials in the labour market outcomes. Recently researchers have started to bring in a number of other socio-economic, psychological and gender identity factors. The empirical research can be divided into three different types according to the data and methodology used:

   i)   experimental (mostly) laboratory-based evidence,
   ii) studies using longitudinal data and
   iii) surveys.

In majority of the existing empirical research, some type of Oaxaca-Blinder or John-Murphy-Pierce decompositions are used to examine labour market differences between men and women and to quantify levels of discrimination.

The main objective of the book is to give overview of what we know so far about various factors contributing to the gender pay gap via reviewing existing theoretical and empirical literature, and to extend our knowledge on the topic through data analysis. In our analyses we mainly focus on the *new* factors such as the relationship between social and gender identity norms, preferences, and gender wage gap. The book focuses also on the role of soft skills, personality characteristics and workplace practices in explaining gender wage gaps. Several of the chapters are based on data from a survey designed precisely to get data on these *new* factors. The survey was conducted in the Czech Republic in 2011 based on the questionnaire, which we constructed to cover various aspects of work and family life, preferences, personality traits and other characteristics of employees and their jobs. Traditional linked employer-employee data are also used in some chapters.

This book is aimed at social scientists, policy makers and the non-technical reader interested in the questions related to gender-based wage differences.

In the book:
- A review of existing theory and empirical evidence related to gender differentials in the labour market outcomes and discrimination.
- Analyses focusing on socio-psychological factors, such as gender identity, work-family balance, personality traits and soft skills in explaining gender wage gaps.

- A novel survey covering various aspects of work and family life, preferences, personality and other characteristics of employees and their jobs.
- Novel findings based on a rich survey, database of soft skill's characteristics and a comprehensive linked employer-employee dataset.

Acknowledgements

Mariola Pytliková

December 2012, Ostrava

Content

Detailed Content

List of Abbreviations

AEIS	Average Earnings Information System
BMI	Body Mass Index
BS	Business Sector
CAPI	Face-to-face interview using a laptop
CEO	Chief Executive Officer
CR	Czech Republic
CZK	Czech Crown
EFTA	European Free Trade Association
EU	European Union
EU-SILC	The European Union Statistics on Income and Living Conditions
F	Female
FE	Firm Effect
GCE	General Certificate of Education
IDORG	Identification of Organisation
ISCO	International Standard Classification of Occupations
ISPV	Average Earnings Information System
KKOV	Classification of Educational Qualification Types
LEED	Linked Employer-Employee Dataset
M	Male
NACE	Classification of Economic Activities
NBS	Non-business Sector
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organisation for Economic Cooperation and Development
PPP	Purchasing Power Parity
UK	United Kingdom of Great Britain and Northern Ireland
USA	United States of America
USD	U.S. Dollar
USSR	Union of Soviet Socialist Republics

Chapter 1

Introduction

By Mariola Pytliková, Lenka Filipová

The twentieth century saw a dramatic improvement in women's economic status thanks to changes in social norms, reduced gender discrimination, affirmative action, lower occupational segregation and increase in access to higher education for women. Technological change has also made women's lives easier through technological advance at homes. Home appliances such as washing machines, refrigerators, vacuum, dishwasher etc. helped women to break free from the traditional *housewife* roles and they allowed women to join the labour force in greater numbers (Greenwood and Seshadri, 2005). However, much of the progress in terms of the gender labour market differentials slowed down in the recent decades.

1.1 Women are more likely to not to work; when they do have jobs, they earn less than men

Despite the fact that there has been a lot of attention paid to improving women´s position on the labour market in the recent decades, data shows that there are still significant gender differences on the labour market.

Figure 1–1 shows female to male median wage ratio for a number of OECD countries for years 2000 and 2010. We may observe that women earn less than men in all developed OECD countries, but the wage gap varies across countries. In Hungary, Poland, Spain, New Zealand and Norway women earn about 90% as much as males, whereas in Israel, Japan and Korea women earn 60 - 70% as much as men. The gender median wage gap has narrowed down in almost all countries since 2000, except for Italy, Portugal and France. Statistics show that women continue to earn less than men even after controlling for characteristics that researchers think might determine productivity (OECD, 2001) and that the differences mentioned are even greater for more educated women and mothers of two or more children (OECD, 2001; Waldfogel, 1998).

There are also substantial differences between men and women in the labour force participation rate. According to Table 1–1 the female labour force participation rate is lower than male participation rate in all developed OECD countries.
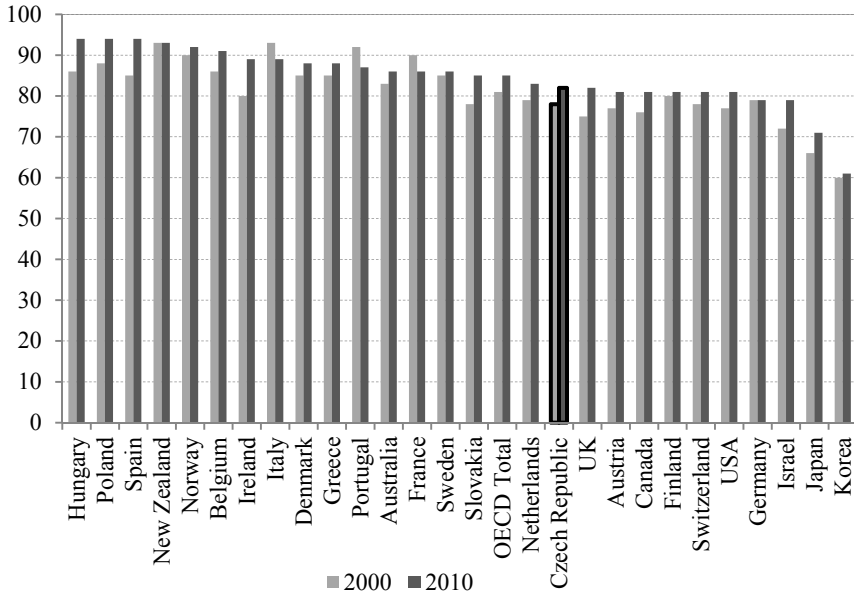
Figure 1–1 Female/male median wage ratio, OECD countries, 2000 and 2010, in %
Source: OECD (2012)

Table 1–1 Gender labour force participation gap, OECD countries, years 2000 and 2011

	2000		2011			2000		2011	
Country	Female	Male	F	M	Country	F	M	F	M
Australia	65	83	71	83	Japan	60	85	63	84
Austria	62	80	70	81	Korea	52	77	55	77
Belgium	57	74	61	72	Netherlands	65	83	73	84
Canada	70	82	74	81	New Zealand	67	83	72	84
Czech Republic	*64*	*79*	*62*	*79*	Norway	76	85	76	80
Denmark	76	84	76	82	Poland	60	72	59	73
Finland	72	78	73	77	Portugal	64	79	70	79
France	62	75	66	75	Slovakia	63	77	61	77
Germany	63	79	72	83	Spain	53	80	68	81
Greece	50	77	58	78	Sweden	76	81	78	83
Hungary	53	67	57	69	Switzerland	72	89	77	89
Ireland	56	80	63	77	United Kingdom	69	84	70	83
Israel	56	67	61	68	United States	71	84	68	79
Italy	46	74	51	73	*OECD Total*	*59*	*81*	*62*	*79*

Source: OECD (2012)
Notes: F = Female, M = Male

Traditionally, Scandinavian countries have the highest female labour participation rates, whereas some Southern European like Greece or Italy, and Central European such as Hungary and Poland, the lowest female labour participation rates among the developed OECD countries. We can as well observe that the female labour force participation increased between 2000 and 2011 for most of the OECD countries with a notable exception of the Czech Republic, Poland and Slovakia. Spain has seen the greatest gains in terms of female labour force participation but from a poor starting point. Further, the gender labour force participation gap has narrowed over time across the OECD (again with the exception of the Czech Republic, Poland and Slovakia).

1.2 Women Matter

Yet, most of the developed (and not only) countries face problems of population aging[1] and increasing female labour participation together with attracting skilled immigrants and encouraging people to work longer is one of the solutions to counteract the population aging. As recent statistics show, women tend to be more educated than men and thus women constitute an important source of skilled labour force, in particular in countries with a relatively low female labour force participation rate. In addition, recent research suggests that women representation on company boards and management is positively correlated with the company's performance, such as profits (Adler, 2001; Adams and Ferreira, 2009). Gender diversity in firms might be also positively correlated to innovation and creativity (Hong and Page, 2001 and 2004; Berliant and Fujita, 2008). Thus, women constitute a valuable part of the work force, and reducing gender inequality should benefit employers, employees and the entire society.

1.3 Factors Explaining Gender Labour Market Differentials

Given the widely observed differences between men and women on the labour market, explaining the differentials has attracted much attention among researchers. In fact, the gender wage gap has been studied extensively since the early 1970s and it became one of the most important topics in the labour economics. The basic research questions regarding the gender wage gap is whether discrimination is an explanatory factor and if so, to what extent.[2] Most of the existing earlier research tries to uncover how much of the gender wage gap remains after adjusting for an individual's productivity enhancing characteristics such as education, experience, occupation and industry. Various decomposition techniques are applied to express the part of wage gap that can be explained by

---

[1] Demographic projections by the United Nations suggest that during the next three decades, populations in Europe might ceteris paribus decline by 12 per cent (Raftery et al., 2012). The main factor responsible for population aging is a large decline in the total fertility rate over the last half century.

[2] The economics of discrimination started with Becker´s seminal study in 1957 and has been important in explaining differences in wages of equally productive males and females.

different characteristics of men and women and the *unexplained part* that may then be ascribed to gender-based wage discrimination. From the methodological point of view, the methods by Oaxaca (1973) and Blinder (1973) or that of Juhn et al. (1993) and their modifications are usually applied. Using this approach most research finds a sizable *unexplained* gender wage gap, which has been usually considered as discrimination. However, many argue that this approach may be biased because it omits important, although unobservable in the data, variables. On the contrary, some of the included variables may reflect discrimination (like occupation segregation or tenure and other gender-specific factors) and discrimination would be underestimated. Thus the residual gap cannot directly be seen as a measure of discrimination (Blau and Kahn, 2000). That is why we are reluctant to use the term *discrimination* and instead will use the concept of *unexplained part of the gender wage gap*.

What do we know about factors explaining the observed gender differentials? So far we know that family situations play an important role in explaining the gender wage gap. Research on this topic can be divided into two groups according to the core of interest which is: (1) marital status[3], or (2) children.[4] The relationship between marital status and pay is different for men and women. Research has shown that there is a *marriage premium* for men, whereas there is a *marriage penalty* for women. As far as impact of children is concerned, the *motherhood penalty* in wages for bearing the childcare duties has been the most investigated topic. In connection to that the effect of family-friendly policies, such as the length of maternity and parental leave, on gender wage gap has been investigated, e.g. in Gupta et al. (2008). Also the need for family-friendly workplaces, i.e. workplaces, which make it possible for employees to balance better their family and work commitments, such as flexible work schedule, access to part-time and work from home, have been discussed recently.

In addition, there has been an increasing focus on psychological and socio-psychological factors as possible explanations for gender differences on the labour market. A comprehensive review of this research is given by Bertrand (2010). As shown in the experimental economics literature women tend to be more risk-averse than men, prefer a less competitive work environment, are more altruistic, have stronger preferences for redistribution and finally are less likely to negotiate on behalf of themselves as men (Bertrand, 2010; Borghans et al., 2009). Further, other socio-psychological factors, such as the effect of gender identity and social norms are put forward as possible determinants of the gender differences in the labour market. In particular the hypothesis that as long as the social norm that *men work in the labour market and women work at home* exists, women will have lower motivation to participate in the labour market than men, with consequences on wages and other labour market outcomes (Akerlof and

---

[3] See for instance Korenman and Neumark (1991), Antonovics and Town (2003), Gupta et al. (2005).
[4] See for instance Mincer and Polachek (1974), Becker (1985), Waldfogel (1998), Joshi and Davies (2002) or Blau et al. (2009).

Kranton, 2000). Further, the theory of preferences points towards different priorities of men and women considering work-family balance (Hakim, 2002 and 2008). Hakim claims that women choose three combinations of lifestyle preferences: home-centered, work-centered, or adaptive; and that these lifestyle preferences are an important determinant of fertility, employment patterns, and job choice. However, empirical research on the role of those psychological and socio-psychological attributes in explaining the labour market outcomes is still in its infancy.

1.4 Women on the Czech Labour Market

In the Czech Republic, similarly to other post-communist countries, gender equality was a highly proclaimed policy goal during the socialist era and evidence shows that the differences in wages between women and men were rather low at that time (Brainerd, 2000). The transition towards market economy has led to an increase in the gender earnings gap similar to those in the developed market economies (Večerník, 2001). Jurajda (2003b) shows that women's hourly wages were up to 30% lower than that of men 1998. According to more recent data the gender median wage gap in the Czech Republic was above the OECD average at around 18% in 2011, see Figure 1–1. The difference in average wages between men and women was around 25% in 2011, see Chapter 5. Jurajda (2005) finds that segregation of women into low-wage jobs is a significant source of wage differences between men and women in the Czech Republic.

There are also some recent studies suggesting that the unexplained part of wage differences between Czech men and women may be caused by their different preferences related to work-family balance, e.g. Křížková et al. (2006), Mysíková (2007), Filipová and Machová (2011). Transition towards market economy brought also increase in the gender gap in labour force participation. As observed from the Table 1–1 the gender gap in the labour force participation in the Czech Republic got larger over time and it is currently around the OECD average at 17%, see Table 1–1. The female labour force participation is in general rather low and even decreased over time, from 64% in 2000 to 62% in year 2011, whereas the male labour force participation rate remained unchanged. Thus, there is surely a great potential for increasing labour force through raising the female labour force participation and in this way solving (at least partly) the increasing fiscal burden in connection with the population aging.[5]

---

[5] The Czech Republic has one of the lowest fertility rates in the world at around 1.4 children per woman (2011), much below the *replacement* rate of 2.1. Consequently the Czech population is aging dramatically, placing a huge fiscal burden for the future generations.

1.5   Organization of the Book

Chapter 2 written by *Lenka Filipová and Dagmar Brožová* gives an overview of the existing existing theoretical and empirical literature regarding the factors determining differentials on labour market by gender, in particular it reviews literature on gender wage gap. Further, we contribute with our own research based on two different datasets from the Czech Republic. First, we use a linked employer-employee micro- dataset (LEED) based on the Average Earnings Information System (AEIS)[6], which is an official data source collected for purposes of the Czech Statistical Office and the Ministry of Labour and Social Affairs. AEIS encompasses monthly gross wages, hourly wages and working hours as well as personal and job characteristics of individuals. It enables us to investigate wage and gender wage differences in time and account for firm specific fixed effects. However, the set of wage determinants that it is possible to create from this data source is quite limited. It covers traditional wage determinants such as education, tenure and other demographic characteristics, job and firm characteristics. On the other hand, the AEIS does not contain the wage determinants focused in the most recent research literature, i.e. data related to family-career balance, preferences, gender identity or socio-psychological characteristics. Due to these data limitations we have developed our own questionnaire, which covers various aspects of work and family life, preferences, personality and other characteristics of employees and their jobs. A survey using this questionnaire (64 questions in total) was carried out on a representative sample of employees in the Czech Republic in 2011. This unique survey helps us to shed some more light on the factors explaining gender wage gap and to overcome the problem of omitted variable bias mentioned above. Relative to the existing literature, we are able to develop a more comprehensive model of wage determinants and consequently to reduce the unexplained part of the gender based wage gap. *Chapter 3* written by *Jiří Balcar, Lenka Filipová, Petra Vašková and Zuzana Machová* describes the data used in our research and gives a first descriptive analysis of these data. *Chapter 4* written by *Jaromír Gottvald, Petra Vašková and Lenka Janíková* uses the traditional data source (LEED based on AEIS) for evaluation of wage determinants in the Czech Republic and discusses the limitations of this data source. Here, we compare also the unexplained residual estimated using AEIS with the residual estimated on the basis of our survey. *Chapter 5* written by *Lenka Filipová, Mariola Pytliková, Jiří Balcar and Jaromír Gottvald* provides results of analysis of the gender wage gap and discrimination in the Czech Republic using our own survey. This survey is then analysed in more detail in *Chapters 6* and *7*, specifically: *Chapter 6* written by *Lenka Filipová, Lenka Janíková and Jaromír Gottvald* focuses on family factors, family-career balance, preferences and gender identity factors, whereas *Chapter 7* written by *Jiří Balcar* concentrates on soft skills. *Conclusion* part, written by *Mariola Pytliková,* summarizes and provides some policy recommendations.

---

[6] In Czech it is called *Informační systém o průměrném výdělku* (ISPV).

Chapter 2

Theory and Empirics of Gender Wage Gap and Discrimination

By Lenka Filipová, Dagmar Brožová

The first part of this chapter presents a brief overview of the gender wage gap literature. The second part describes decomposition techniques usually applied to express gender-based wage discrimination.

2.1 Gender Wage Gap

This chapter gives a brief overview of the gender wage gap literature. A more detailed literature review related to the specific topics of gender wage gap and discrimination will be presented in the chapters focused on those specific topics. Gender wage gap has been systematically studied since the early 1970s when Mincer (1974) expressed the relation between human capital (education, age and/or experience) and wages as an equation which started to serve as a methodological base for the research of wage determinants. Later some other authors, e.g. Dickens and Katz (1987), Krueger and Summers (1987) and (1988) or Allen (2001) modified Mincer's wage regression by adding a labour status description (the position in one's employment, working hours, type of employment contract, and qualification requirements), employer's characteristics (company size, commercial and non-commercial sectors, and industries), institutional factors (legal protection of employment and the minimum wage), and regional factors. The economics of discrimination which started with Becker´s seminal study in 1957 has been an important approach when explaining differences in wages of equally productive males and females. Various decomposition techniques are mostly applied to express the part of wage difference that can be explained by objective characteristics of men and women and the unexplained part that may then be ascribed to gender-based wage discrimination. From the methodological point of view, the methods by Oaxaca (1973) and Blinder (1973) or that of Juhn et al. (1993) and their modifications are usually applied. However, as emphasized in the introduction, research studies on this topic have a challenge like all regression analysis with omitted variable bias. If some relevant variables were omitted – and if for instance men were

more highly endowed with respect to these omitted variables – then discrimination would be overestimated. On the contrary, some of the included variables may reflect discrimination (like occupation segregation or tenure and other gender-specific factors) and discrimination would be underestimated. Thus the residual gap cannot directly be seen as a measure of discrimination (Blau and Kahn, 2000).

Women's greater responsibility for children is considered as an important factor in explaining why women earn less than men. Research on this topic can be divided into two groups according to the core of interest which is: (1) marital status, e.g. Hill (1979), Korenman and Neumark (1991), Gray (1997), or (2) children, e.g. Mincer and Polachek (1974), Becker (1985), Joshi and Davies (2002), Fuchs (1988), Blau et al. (2009), Waldfogel (1998). According to Waldfogel and Sigle-Rushton (2007) there can be at least four types of explanations of the gender and especially motherhood gap or penalty in pay. The first, human capital theory (Mincer and Polachek, 1974; Becker, 1985; Polachek, 1995) emphasizes that women generally have lower wages than men because they have lower levels of human capital such as education or training; work experience; and job tenure. Women with children tend to earn even lower wages than other women because they spend more time at home when they have children, are more likely to work part-time, and change employers. Since they anticipate shorter and more discontinuous work lives, they have lower incentives to invest in market-oriented formal education and on-the-job training. Several recent studies prove that career interruptions can explain a sizeable proportion of the gender wage gap (Bertrand et al., 2009) which is commonly explained by the above-mentioned theory of accumulating less human capital. A second explanation for the lower wages of women with children is connected with the trade-off women make between wages and flexibility (part-time jobs, location close to home etc.). Third, the lower wages of mothers are also explained by their lower real or perceived productivity and effort. It is supposed that longer hours that women spend on housework and taking care of children may also decrease the effort they put into their market jobs compared to men or women without children. Fourth, the lower wages of mothers can be explained by a selection. It is supposed that women who anticipate having children and difficulties in combining work and childcare, invest less in human capital and thus have lower wages. On the contrary, women who expect to be higher earners may choose not to have children.

There is also a growing number of studies pointing out the unexpected detrimental effect of family-friendly policies on women's economic attainments. They emphasize that although family-friendly policies enhance women's economic independence by facilitating their participation in the paid economy (OECD, 2001), other forms of gender inequality begin to appear. One of them is a gender-based occupational segregation which is more evident in the women-friendly Scandinavian labour markets than in the liberal market economies like USA and Canada (Jacobs and Gerson, 2004; Jacobs and Lim, 1992; Chang, 2000; Wright et al., 1995). This gender-based occupational segregation is

explained by the increased demand for female labour associated with welfare state expansion which is characterized by the growth of the social service sector. State-sponsored family services such as child care facilities, educational institutions, and care homes for the elderly are female dominated occupations which on the one hand bring job opportunities for women, on the other they are connected with low earnings (Esping-Andersen, 1990; Kolberg, 1991; Gornick and Meyers, 2003).

Over the last ten years psychological and socio-psychological factors have been increasingly considered as possible explanations for gender differences in the labour market. Bertrand (2010) presents a comprehensive review of this research. In respect to psychological attributes, gender differences in risk taking, attitudes towards competition, social preferences and attitudes towards negotiation have been evaluated thoroughly. A lot of experimental literature on this topic has proved systematic differences in the attributes mentioned between men and women. I.e. women are more risk-averse than men, prefer a less competitive work environment, are more altruistic with stronger preferences for redistribution and don´t negotiate for themselves as well as men. Other personality traits that are most frequently researched is the so-called Big Five model, i.e. extroversion, agreeableness, conscientiousness, neuroticism and openness to experience. The research of empirical implications of those psychological attributes for labour market outcomes is in its infancy, but promising, see Duckworth and Quinn (2009), Almlund et al. (2011), Borghans et al. (2009).

The research of socio-psychological factors is methodologically based on the identity model imported into economics from social psychology. Akerlof and Kranton (2000) suggested the utility model according to which one´s identity can influence economic outcomes because behaving differently from what is expected for one´s social category could decrease the person´s utility. Regarding gender identity the impact of social norms about what is appropriate for men to do and what is appropriate for women to do is researched. The identity model could then explain labour force participation or occupational segregation by gender with the impact on wages. As long as the social norm that *men work in the labour market and women work at home* exists, women will have less motivation to participate in the labour market than men. Considering occupational segregation Goldin (2002) assumes that men do not want women in certain jobs because men are afraid of losing the prestige they get from working in those jobs and accepting women as co-workers would threaten their own gender identity. Using data from the World Values Surveys Fortin (2005) analyzed the influence of women´s view of themselves regarding their labour force participation and relative earnings in a sample of 25 OECD countries over a 10-year period. The results show that representation of women as homemakers is quite stable across cohorts and over time, which may have an impact on their labour market outcomes. On the basis of the gender identity model the changes in women´s labour market outcomes are dependent on the changes in meaning of the male and female social categories.

Another approach to evaluating social factors uses the theory of preferences to look at the differences in priorities for men and women considering work-family balance, e.g. Hakim (2002 and 2008). As mentioned in the introduction, Hakim claims that women choose three combinations of lifestyle preferences: home-centered, work-centered, or adaptive and that these lifestyle preferences are an important determinant of fertility, employment patterns, and job choice.

2.2 Gender-Based Wage Discrimination

Discrimination started to be methodologically studied after the publication of Becker´s seminal work (Becker, 1957). After just a short while economists presented an equation to measure discrimination. Discrimination is a multidimensional phenomenon, which can be examined (due to the complexity of social relations) only on the basis of a multidimensional approach. One of the biggest contributions of economics to this research is undoubtedly the method of measuring discrimination, at least approximately. The American economist R. Oaxaca elaborated on this technique and it became known as Oaxaca´s decomposition of wage differential between two groups of workers, which has been a vital part of *labour economics* since the 1970´s. This approach was further developed by other authors, e.g. A. Blinder.

2.2.1 Regression Model

Regression analysis is the technique that is often used to measure earnings differences according to gender. Regression equation describes the relations between a dependent variable (Y) and explanatory variables (X_i).

A regression equation (2.1) has the general form:

$$Y = \alpha + \beta_1 X i_1 + \beta_2 X i_2 + \beta_3 X i_3 + \cdots \mu i, \qquad (2.1)$$

where Y is the dependent variable, X_{is} are the independent variables, those that are determining the result, Y. μ is the error term and represents the influence of factors that affect Y but are unobserved and aren´t included in the model. α and β are the regression coefficients, α is the constant term and β measures the effect on Y of a one unit change in X, if the value of all the other X variables are constant (ceteris paribus). Regression analysis uses the statistical information on the values of Y and the Xs to estimate the values of coefficients α and β for each separate X variable.

In the regression analysis gender wage differentials are dependent variables of the earnings of two groups of workers, men and women, so called gender pay gap. Explanatory variables are the factors that affect these earnings.[7]

---

[7] For more detailed analysis see Hoffman and Averett (2010), chap. 10.

2.2.2 The Oaxaca-Blinder Decomposition[8]

The principal idea of the Oaxaca Blinder decomposition of wage differential is based on the fact that every wage differential between two groups of people, defined for example by gender, can be divided into two parts. An explained part, that is related to differences in factors determining the earnings, and the residual, unexplained part. This part can be seen as a quantitative indicator of the degree of discrimination.

The first step consists in identification of the factors – determinants of differences in earnings of men and women. The earnings depend on investments in human capital, on number of years of education and on number of years of work experience. These investments depend on innate skills and endowment, and on personal characteristics: responsibility, carefulness, purposefulness, persistence, willpower.

The persons with greater endowments of these characteristics acquire new knowledge more easily and with less effort than those less endowed. They acquire skills and competences with lower costs, so they study and invest in their human capital more often and much more than the less endowed ones. The number of years of work experience also belongs to human capital, because these increase its value. Other determinant characteristics of earnings are, for example, the age of the worker, professional market (or branch), the size of the firm, where one works, the size of the town or village, where one lives, or the region where the firm is, and so on. These factors determining earnings are then written as explaining variables into the earnings functions of men and women and their influence on the level of their earnings is identified, respectively the influence of wage differential between men and women. The part of the wage differential that remains unexplained by these observed determinants is then ascribed to discrimination. That is why this methodological approach is described as *residual*, because it describes the unaccounted-for remainder from the difference between earnings of men and women, by the factors observed, and then elaborated into the model.

The earnings equation (2.2) can be written:

$$Earnings = \alpha + \beta_1 x \ years \ of \ education + \beta_2 x \ years \ of \ work \ experience$$
$$+ \ \beta a \ x \ age + \beta g \ x \ gender + \ldots \mu, {}^{9} \qquad (2.2)$$

where *years of education* and *years of work experience* are continuous variables, easily interpreted numerical values, 13 or 18 years of education for example. Constant term α represents the basic earning without investment in human capital and in other characteristics affected productivity of labour and without investment in work experience. Coefficients β in equation represents the impact

---

[8] This classical procedure see Oaxaca (1973), and further development and enrichment see Blinder (1973), Oaxaca and Ransom (1994), Juhn et al. (1993), Oaxaca and Ransom (2010).
[9] Hoffman and Averett (2010), p. 338, adjusted.

on earnings of having one more year of education or one more year of work experience. *Gender* do not have numerical values, variables like these are included in a regression by using a *dummy* variable: let X = 0 if a person is male and let X = 1 if a person is female. Regression coefficient for the dummy variable gender measure the impact on earnings of being female (rather than male). For example, if the estimated regression coefficient on a female dummy variable is negative, it means that women have lower earnings than men, ceteris paribus.

In the simplest version we suppose only one factor determining earnings – education (denoted by *s*). The average male wage can be indicated w_m and the average female wage can be indicated w_f, while the difference in mean wages we can written $\Delta w = w_m - w_f$.

The earnings functions for men (2.3) and women (2.4) describe the dependence of earnings on the number of years of education for each of the two groups, which can be written as:

$$w_m = \alpha_m + \beta_m s_m \text{ for men,} \qquad (2.3)$$
$$w_f = \alpha_f + \beta_f s_f \text{ for women,} \qquad (2.4)$$

where α_m and α_f describe earnings of men and women for their skills with zero years of education, if employers valued the skills of men and women who have zero years of education equally: $\alpha_m = \alpha_f$. β_m and β_f are coefficients, expressing by how much a man´s and woman´s wage increases if he or she gets one more year of education, if employers valued the education acquired by women as much as they value the education acquired by men , these coefficients would be equal: $\beta_m = \beta_f$.

Raw wage differential (2.5) can be written:

$$\Delta_w = w_m - w_f = \alpha_m + \beta_m s_m - \alpha_f - \beta_f s_f, \qquad (2.5)$$

where s_m a s_f is the mean quantity of education of men and the mean quantity of education of women (in the number of years).

It is possible to modify the wage differentials (Borjas, 2010): to add and subtract the term $\beta_m s_f$ to the right – hand side of equation, and the terms in the equation can be rearranged and rewritten (2.6) in the form:

$$\Delta w = \left(\alpha_m - \alpha_f\right) + \left(\beta_m - \beta_f\right)s_f + \beta_m\left(s_m - s_f\right), \qquad (2.6)$$

where part $\left(\alpha_m - \alpha_f\right) + \left(\beta_m - \beta_f\right)s_f$ expresses the discrimination and $\beta_m\left(s_m - s_f\right)$ expresses the difference in skills originating from education.

The equation breaks down the raw wage differential in two parts. The second term on the right – hand side of equation describes the difference arising from different quantum of education. If men and women have the same average quantum of education, the second term would be zero ($s_m - s_f = 0$). If the average quantum of education for men and women differs, it is clear that part of the wage differential arises because the two groups differ in their skills originating from education.
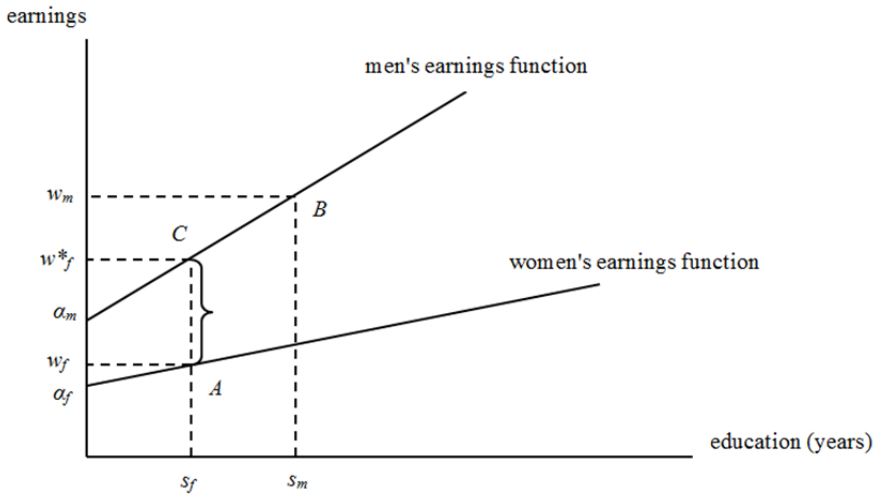
Figure 2–1 The impact of discrimination on wage
Source: Borjas (2010)

The first term represents this unexplained part from education, residuum-part, which has been seen as an indicator of discrimination. It has a positive value if employers appreciate a man´s education more than they value a woman´s education. ($\beta_m > \beta_f$) or if employers pay men more than women for any level of education. The intercept of the earnings function with axis y is higher for men than for women ($\alpha_m > \alpha_f$). The wage gap that arises from this differential treatment of men and women is defined as discrimination.

Figure 2–1 describes the graphical interpretation of the Oaxaca decomposition. The relation between earnings and education attendance is illustrated by the earnings functions of men and women. The steeper the earnings function is and the higher the point in which intercepts the axis y, the higher the evaluation of men in comparison with women by the firms.

The starting position of men is more advantageous. Their abilities and skills are better awarded by higher payment than women even if the two groups have no education attendance. Men have also higher income resulting from additional years of education than women. The wage differential between men and women in the figure is presented by the vertical difference. $w_m - w_f$ and it arises with the growing number of education years. Women have in average s years of schooling and their average earning is w (point A). Men have in average s_m years of education and earn w_m (point B). That part of the wage differential arises from the fact that men have in average more years of education than women. The difference ($w_m - w^*_f$) is the part of differential that is the consequence of different education (line segment BC on the earnings function). If a given woman is paid the same as a given man with the same level of education, they

acquire $w*_f$ (point *C*). The discrimination is then measured by the difference ($w*_f$ – w_f, respectively line segment AC).

It is obvious that the validity of this method of measuring discrimination depends on the extent to which we are successful in defining, describing and including in the model all the factors determining different earnings of men and women: number of years of work experience, age, marital status, the size of the firm, the size of the residence, region etc. If some of them are not included in the regression model, its evidence about the extent of influence of discrimination on the earnings would be appreciably devalued. In reality, it is impossible to include all the relevant factors determining the level of earnings. For instance, the difference in quality of education acquired by both groups cannot be included. When men and women attend educational institutions of varying quality, the Oaxaca decomposition generates a difference that is attributed to discrimination, when, in fact, the skills acquired differ. Another factor that is difficult to include in the model is individual effort or motivation, which varies from case to case, and can stem from purely personal incentives or preferences. The unexplained part of wage differential, the so-called residual part, comprises also the part that cannot be explained by the model due to lack of information about characteristics of men and women and omission of further relevant characteristics influencing differences in earnings. Therefore this unexplained part of wage differential should be interpreted not as a quota attributed to discrimination, but *as the maximum possible upper limit of wage discrimination*.

This procedure can be extended. Other explanatory variables can be incorporated into the model and the equation can be converted into a logarithmic form that facilitates the interpretation. The equation (2.7) can be expressed by the natural logarithm of wages (Stupnytsky, 2007):

$$lnw_i = \alpha + \beta_1 s_i + \beta_2 a_i + \beta_3 b_i + \beta_4 X_i + e_i, \qquad (2.7)$$

where *s expresses* number of years of education, *a* is number of years of work experience, *b* is age, X_i are the other explanatory variables (e.g. branch, company size, size of settlement, region, proportion of women in the company), e_i expresses residual part.

Earnings functions for men (2.8) and women (2.9) have the following form:

$$lnw_{mi} = X_{mi}\beta_{mi} + \varepsilon_i \text{ for men,} \qquad (2.8)$$
$$lnw_{fi} = X_{fi}\beta_{fi} + e_i \text{ for women,} \qquad (2.9)$$

where lnw_{*i} is the natural logarithm of wages and *X* is vector the other explanatory variables.

The wage differential can be written (2.10) in the form (Stupnytsky, 2007):

$$lnW_m - lnW_f = (X_m - X_f)\beta_m + (\beta_m - \beta_f)X_f + \varepsilon - e, \qquad (2.10)$$

where the first part of the term on the right – hand side of equation represents the explained part (from different characteristic men and women), the second remains unexplained and represents discrimination.

Beblo et al. (2003) break down the gender wage gap in accordance with the general approach by Oaxaca-Blinder into two parts. The first part of decomposed gender wage gap is explained by differences in observable characteristics of human capital, endowment, and the other different, for the job issues relevant, characteristics between the two groups. This part is called endowment effect. It actually reflects differences in work productivity between men and women. The second part of wage differential represents differences in evaluation which is attributed to male and female characteristics, which is therefore the price, respectively the remuneration for endowment and talent, which is afforded these by society – it is called remuneration effect. It is this second part that is often interpreted as wage discrimination.

Provided that $lnw^J = X^J \beta^J$ (for all $J = M, F$), decomposition of wage differential (2.11) by Beblo et al. (2003):

$$\underbrace{(lnw^M - lnw^F)} = (lnw^M - lnw^{*F}) + (lnw^{*f} - lnw^F), \qquad (2.11)$$

raw wage differential.

The term in the first parentheses on the right side of the equation denotes the hypothetical wage differential if women had the same wage structure as men. The term in the second parentheses indicates the disparity between the hypothetical wage rate for women and their actual mean wage, provided that we consider male prices (possibly a weighted price vector as reference). When using female prices as reference, the wage decomposition might offer diverse results. These alternatives and their practical outcomes are presented in particular in the paper of Oaxaca and Ransom (1994).

Decomposition (2.12) of the gender wage gap into *endowment effect* and *remuneration effect:*

$$(lnw^M - lnw^F) = \underbrace{(X^M - X^F)\beta^M}_{endowment\ effect} + \underbrace{X^F(\beta^M - \beta^F)}_{remuneration\ effect}, \qquad (2.12)$$

where vector β^M represents non-discriminatory, i.e. male wage structure, vector β^F represents female wage structure and $lnw^{*F} = X^F \beta^M$.

These authors also present their opinions on the two aspects of the classic approach of Oaxaca and Blinder. The first is concerned with the reference group. Which structure of earnings should be considered as non-discriminatory? The classical approach of Oaxaca (1973) and Blinder (1973) assumed both possibilities, to consider as non-discriminatory either the earnings structure of men or the earnings structure of women. The use of one earnings structure as the referential one, as shown by Beblo et al. (2003), brings potential assymetrics in discovered effects. Other authors consider as non-discriminatory the structure obtained as an average of earnings structures of men and women, while the importance is determined by the relative size of the respective group (Cotton, 1988), which substantially eliminates the above mentioned problem. Neumark