Cardiff Economics Working Papers

Working Paper No. E2014/23

Gender Wage Gap Trends in Europe: The Role of Occupational Allocation and Skill Prices

Ezgi Kaya

November 2014

Cardiff Business School
Aberconway Building
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

Gender Wage Gap Trends in Europe: The Role of Occupational Allocation and Skill Prices*

Ezgi Kaya †
Cardiff University-Cardiff Business School

AUGUST, 2014

Abstract

In this paper, we explore the recent gender wage gap trends in a sample of European countries with a new approach that uses the direct measures of skill requirements of jobs held by men and women. We find that, during the 1990s and 2000s, the gender wage gap declined in the majority of the European countries. Similar to the U.S. experience, a part of this decline is explained by changes in returns to brain and brawn skills in Austria and in the U.K. However, in contrast to the U.S. experience, the changes in returns to brain and brawn skills had a widening effect on the gender wage gap in Southern European countries and in Ireland. Furthermore, we find that a substantial part of the changes in the gender wage gaps in European countries and in the U.S. cannot be explained by the changes in brain and brawn skill prices. The findings of this study suggest the importance of changes in labor market institutions in explaining the gender wage gap trends.

Keywords: Gender wage gap, brain skills, brawn skills, decomposition.

JEL Classification: J16, J24, J31, J71.

---

*I would like to thank Manuel Arellano, Francine Blau, David Card, Michael Coelli, Juan J. Dolado, Francesco Fasani, Libertad González, Lawrence Kahn, Pedro Mira, Ana I. Moro-Egido, Pau Salvador Pujolás Fons, Virginia Sánchez Marcos and especially to Nezih Guner and Ana Rute Cardoso for many useful comments and suggestions. I also benefited from discussions with seminar participants at the XXVII AIEL National Conference of Labour Economics, IAFFE 2012 Annual Conference, COSME-FEDEA Annual Workshop on Gender Economics, ENTER Jamboree 2012 and Students Workshop at Universitat Autònoma de Barcelona. Financial support by the Spanish Ministry of Science and Innovation through grant "Consolidated Group-C" ECO2008-04756 and FEDER is gratefully acknowledged. All remaining errors are mine.

†Contact email: KayaE@cardiff.ac.uk

1 Introduction

There was a dramatic decline in the gender wage gap in the U.S. during the 1980s. The fact that this happened despite a significant rise in overall wage inequality, shifted the attention in the literature to the relationship between the overall wage structure and the gender wage gap. The key change in the U.S. wage structure in 1980s was the rising returns to education and experience due to an increase in demand for high–skilled labor (Katz and Murphy, 1992; Juhn, Murphy and Pierce, 1993).[1] In their seminal paper, Blau and Kahn (1997) find that the change in the U.S. wage structure should have widened the gender wage gap since women had an initial relative deficit in labor market characteristics such as education and experience. However, women were able to overcome this deficit by improving their labor market characteristics, especially their experience levels and and their occupational allocation.

The existing literature attributes the increase in relative demand for high-skilled labor to the technological change, in particular to the developments in computer technology.[2] The task based approach of skill biased technological change proposed by Autor, Levy, and Murnane (2003) moves beyond traditional measures of labor market characteristics (such as education and experience) and models the relation between skills and technological change through tasks performed at jobs. In this framework, work performed in an occupation is broken down into routine and non–routine tasks, which are substitutes and complements with computers, respectively. Therefore, with the development of computer technologies, a shift in the production technology occurred that favored more skilled workers who perform non-routine cognitive tasks in their jobs.

If occupations are characterized by their skill requirements, one can infer the skill intensities of workers given their occupational allocation.[3] Since there exists gender differences in occupational allocation, we expect changing relative demand for skills to have an impact on the gender wage gap.[4] Focusing on the different aspects of the skills required to perform an

[1]During the same period, wage inequality also rose also within education and experience groups reflecting higher returns to unmeasured labor market skills (Blau and Kahn, 1997.)

[2]See Katz and Autor (1999) for a survey.

[3]This allocative process may result from different choices of individuals, discrimination in the process of recruitment or hiring or differences in comparative advantage of workers as in Roy (1951).

[4]Welch (2000) assumes that women are relatively more intensive in intellectual or brain skills while men being more physical or brawn skill intensive. Hence, an increase in the relative value of brain skills, should actually

occupation (such as cognitive, motor, people skills and physical strength), Bacolod and Blum (2010) study how changes in the prices of various skills affected the gender wage gap in the U.S. Their results show that changes in prices of different types of skills (cognitive, motor, people skills and physical strength) contributed to narrowing the gender gap between 1968 and 1990. During this period, cognitive and people skills became relatively more valuable compared to motor skills and physical strength. Since females held occupations that require more cognitive and people skills relative to males, this narrowed the gender wage gap in the U.S. between 1968 and 1990.

This paper revisits the findings of Bacolod and Blum (2010) for a set of European countries: three Southern Europe countries (Italy, Portugal, Spain), two Anglo-Saxon countries (Ireland and the U.K.), and Austria (as an example of Continental European countries).[5] The skill requirements of occupations are obtained from Occupational Information Network (O*Net) data. First, using the data from O*Net, we characterize occupations by two primary attributes, "brains" and "brawns". Then, we matched the brain and brawn skill requirements of jobs with the individual level data from European Community Household Panel (ECHP) and European Union Statistics on Income and Living Conditions (EU-SILC) given the occupational allocation of workers. As a result, we determine the skill intensities of each individual in the sample and estimate the wage return to each skill.[6] Finally, we quantify the contribution of changes in skill prices to the changes in gender wage gap by decomposing the changes in gender wage gap for each country into its components using the technique developed by Juhn, Murphy and Pierce (1991). In order to explore whether the patterns in the U.S. during the 1990s and 2000s changed compared to 1970s and 1980s, we also analyze the changes in the U.S. gender wage gap for the same time period using data from Current Population Survey (CPS).

We find that, from 1993 to 2008 in the U.S. brain skills became more valuable, while brawn skills became relatively less valuable. The experience of Austria and the U.K. was similar to the U.S. In contrast, the increase in returns to brain skills and decrease to brawn skills was not a common phenomenon for the Southern European countries –Italy, Portugal and Spain–

narrow the gender wage gap.

[5]The sample of countries does not include examples of the Nordic and eastern European countries due to the lack of comparable data for the analysis. See Section 3 for the description of data sources.

[6]See Autor et al. (2003) and Bacolod and Blum (2010) for a similar approach.

and for Ireland. During the period of analysis, in Italy, Portugal, Spain and Ireland, brawn skills became relatively even more valuable. The change in skill prices in these countries are potentially affected by the period of the analysis. Ireland and Spain from the mid-1990s experienced a construction and housing boom which potentially explains the increase in returns to brawn skills.

Additionally, we also show that, from 1993 to 2008, the U.S. gender wage gap declined (0.051 log points) and a part of the convergence in the gender gap can be explained by the change in brain and brawn skill prices, similar to the findings of Bacolod and Blum (2010) for 1970s and 1980s.[7] In particular, 17% of the closing U.S. gender wage gap can be explained by changing returns to brain and brawn skills.[8] During the same period, the gender wage gaps also declined in the European countries in our sample, except Spain.[9] Similar to the U.S. experience, a part of the decline in the gender wage gaps in Austria and in the U.K. can be explained by the changes in returns to brain and brawn skills. In particular, the changes in returns to brain and brawn skills account for around 15.4% of the closing gender wage gap in Austria and around 7.6% in the U.K. On the other hand, in contrast to the U.S. experience, in Southern European countries and in Ireland, the changes in returns to brain and brawn skills had a widening effect on the gender wage gaps. In the absence of changes in skill prices, the gender wage gap would have narrowed even further in Ireland (0.032 log points more), in Italy (0.022 log points more) and in Portugal (0.037 log points more). On the other hand, if skill prices would not have changed, the Spanish gender wage gap would have widen only around 0.025 log points instead of 0.035.

Despite these differences across European countries and the U.S. a striking fact is that, a substantial part of the changes in the gender wage gaps cannot be explained by the changes in observable gender-specific factors (i.e. labor market characteristics or brain and brawn skills) or changes in wage structure (i.e. returns to characteristics, skill prices or residual wage inequality). Of course a natural question is then why the gender wage gaps still declined during 1990s

[7]Bacolod and Blum (2010) show that 20% of the narrowing gender gap in the 1980s in the U.S. is due to change in prices of cognitive, people and motor skills as well as the physical strength.

[8]This result is similar when the decomposition analysis is performed for an earlier period. See Table D.2 of Appendix D for the decomposition results for 1979–1988.

[9]The increase in the gender wage gap in Spain from 1994 to the beginning of 2000s is documented also by Guner, Kaya and Sánchez-Marcos (2014).

and 2000s. Other factors that may have contributed to the convergence of the unexplained gender pay gap include changes in selection to the employment, changes in gender differences in unobservable skills and labor market discrimination, as well as the changes in labor market institutions. To answer this question, we explore the relationship between the gender wage gaps that can not be explained by changes in observable gender-specific factors and wage structure and changes in various measures that captures the labor market institutions and discrimination. We find that the changes in these measures are highly correlated with the unexplained part of the gender wage gap trends. Furthermore, we provide some evidence consistent with the role of changes in the labor market institutions, such as decline in the trade union density and increase in the employment protection of temporary workers, in explaining the gender wage gap trends even if the bias induced by non-random selection to employment is corrected.

The number of studies that focus on the skill requirements of occupations to analyze the gender wage gaps in the European labor markets is rather limited. This paper is intended to fill this gap in the literature. A recent paper that is particularly related to the current study is Black and Spitz-Oener (2010). Using self-reported measures of tasks performed within occupations, Black and Spitz-Oener (2010) employ a task-based approach to study the effect of changing tasks on the gender wage gap trends in Germany. Their results indicate that changes in the relative task and relative prices together explain more than 40 percent of the narrowing of the gender gap in West Germany despite the widening effect of changing task prices. Overall, these results are parallel to the findings of this study. In contrast to Black and Spitz-Oener (2010), this study considers skills to be required to perform an occupation and characterizes occupations by skills rather than self-reported measures of routine or non-routine tasks.

The results of the current study are also related to the findings of Borghans, ter Weel, and Weinberg (2006). Using data for Germany, for the U.S. and for the U.K., they show that occupations that require more computer usage and higher extent of team work require more people skills. Moreover women have relatively higher employment rate in occupations which require people skills. They suggest that the increased importance of people skills by the technological change and innovative work practices have raised women's relative employment in those occupations. This study complements their findings by showing the increasing representation of women in occupations which require brain skills. In addition to that, this study quantifies the role of changes in skill intensities and skill prices on the gender wage gap trends in various countries.

The remainder of the paper is organized as follows. The next section explains the details of the decomposition technique employed. Section 3 describes the data sources and concepts used in the analysis and presents the empirical specification. Section 4 analyses the gender wage gap trends, changes in brain and brawn skill intensities of male and female workers and trends in skill prices in the sample of European countries and the U.S. The main results for the decomposition of the changing gender wage gaps are presented in Section 5. Finally, Sections 6 and 7 discuss the role of labor market institutions and non-random selection to the labor market that might also have an impact on the gender wage gap trends and Section 7 concludes.

2 Analytical Framework

The existing literature classifies the factors affecting the gender wage gap into two groups: (i) gender specific factors and (ii) factors related to wage structure. Gender specific factors capture the relative differences of males and females in labor market characteristics (such as education, experience, brain and brawn skill intensities) as well as the gender differences in unobserved qualifications or discrimination. Returns to labor market characteristics, skill prices or the residual wage inequality are not related specifically to aspects of gender and considered as factors related to wage structure. The method developed by Juhn, Murphy and Pierce (1991), hereafter JMP, enables one to decompose the change in the gender pay gap into changes in gender specific factors and those related to the changes in wage structure. This section briefly explains the JMP decomposition technique that is employed in the analysis to quantify the role of each component on the gender wage gap trends. To this end, let the wage equation for males at time t be given by

$$lnW_t^M = X_t^M\beta_t + S_t^M\gamma_t + \sigma_t\theta_t^M, \quad (1)$$

where lnW_t^M is the logarithm of hourly wages, X_t^M is a matrix of labor market characteristics (including education and experience) with returns vector β_t, S_t^M is the matrix of brain and brawn skill intensities of workers determined by the skill requirements of the jobs that they hold and γ_t is the price vector for brain and brawn skills. θ_t^M is the vector of standardized residuals (with mean zero and variance one) and σ_t is the residual standard deviation of male wages for year t (i.e. unexplained level of male residual wage inequality). Given consistent estimates of Equation 1, the gender wage gap for year t can be decomposed as

$$\Delta lnW_t \equiv lnW_t^M - lnW_t^F \quad = \quad [\Delta X_t \beta_t + \Delta S_t \gamma_t] + \sigma_t \Delta \theta_t, \tag{2}$$

where lnW_t^M and lnW_t^F are the average log hourly wage for males and females, respectively, ΔX_t is the male-female differences in labor market characteristics, ΔS_t is the male–female differences in brain and brawn skill intensities, and $\Delta\theta_t$ is the male–female differences in the average standardized residuals. Hence, the gender wage gap for year t can be decomposed into two components, one component due to male–female differences in average labor market characteristics and in average brain and brawn skills weighted by the male prices for these characteristics and skills ($\Delta X_t \beta_t + \Delta S_t \gamma_t$), and another component due to differences in the average standardized residuals weighted by the male residual wage inequality ($\sigma_t \Delta\theta_t$).[10] Given the gender wage gap in two years, s and t, the change in the gender wage gap from year t to s, can then be decomposed as

$$\begin{aligned} \Delta lnW_s - \Delta lnW_t \quad &= \quad [(\Delta X_s - \Delta X_t)\beta_s + (\Delta S_s - \Delta S_t)\gamma_s] \\ &+ \quad [\Delta X_t(\beta_s - \beta_t) + \Delta S_t(\gamma_s - \gamma_t)] \\ &+ \quad (\Delta\theta_s - \Delta\theta_t)\sigma_s \\ &+ \quad \Delta\theta_t(\sigma_s - \sigma_t). \end{aligned} \tag{3}$$

[10]We follow the parametrization by Blau and Kahn (1997) by formulating the wage gap based on male's wage equation. Alternatively, the formulation could based on the female's wage equation or pooled regression. Using male's wage equation lies in the assumption that the prices from the male regression are equivalent to competitive prices. Since, male-female differences in returns can reflect discrimination, the use of male's equation is employed to simulate the wage equation in a nondiscriminatory labor market.

In this four component decomposition, the first component reflects the contribution of changing gender differences in labor market characteristics as well as the skill intensities from year s to year t and is called "observed X effect". The second component captures the effect of changing returns to characteristics and prices of skills for males and is called "observed β effect". The two components are straightforward to calculate using the estimated coefficients from the male wage equation and the sample means by gender.

The third and the forth components are called "gap effect" and "unobserved price effect", respectively, and they are calculated using the entire male and female residual distributions. In particular, the gap effect is calculated as follows. First, for each women in each year a hypothetical wage residual is computed by estimating what her wage residual would be if her labor market characteristics and skills were rewarded as they would be rewarded for men for that year (i.e. female residuals from male regression). Then, a percentile number is assigned to her corresponding to the position of her hypothetical residual in the male residual wage distribution for that year. Second, given her percentile number in year t and the male residual wage distribution in year s, her imputed wage residual is computed for year t. Similarly, her imputed wage residual for year s is the male residual in year s that corresponds to her percentile number in year s. For males, the imputed wage residual for year t is calculated by using their percentile ranking in year s and their wage residuals for year t. Finally, the gender difference between the average of the imputed wage residuals in time period t and s are used to compute the gap effect. Since both computations use the same year s distribution, this term captures the effect of changing positions of females in the male wage residual distribution. Such a change is considered as either the convergence in unobservable skills of females and males or a decline in the discrimination (Juhn et al., 1991). Analogously, the unobserved price effect is calculated by comparing the same year t individuals and by allowing only male residual wage inequality to change. Provided that $\Delta\theta_t$ is negative (since females typically earn less than the mean), a rise in male residual inequality would lead to an increase the gender wage gap.

Since the first and the third term of Equation 3 captures the changing male–female differential in observed and unobserved qualifications respectively, the sum of these two terms are called "gender-specific factors". On the other hand, the sum of the second and the forth component reflects the changing observed and unobserved prices and is called as "wage structure effect" (Blau and Kahn, 1997; Juhn, et al., 1991). By decomposing the changes in the residual gap into into price and quantity effects, JMP decomposition technique can be used to quantify the relative importance of gender-specific factors and wage structure in the gender pay gap trends.

There are, however, two potential drawbacks of JMP decomposition (Blau and Kahn, 1997; Kunze, 2007; Suen 1997). First, the inconsistent estimates of β_t and γ_t in Equation 1 may affect the interpretation of each component. Since female employment rates have changed considerably, the selection bias might be one reason that would lead inconsistent estimates (Heckman, 1979). The sign of the bias is ex-ante unpredictable, since the selected group might be positively or negatively selected in terms of their unobserved characteristics (Blau and Beller, 1988; Blau and Kahn, 1997). Selectivity bias correction (Heckman, 1979) is a common approach to overcome this problem. In our benchmark estimates we use male prices to ameliorate the problem due to changes in non-random selection into work since male employment rates are quite stable over time.[11] Moreover, changes in male prices abstracts from the change in male-female differences in returns that may be relate with discrimination. In Section 7, we explore the possible contribution of sample selection to the gender wage gap trends by implementing the correction for selection into work using a two-stage Heckman (1979) selection model.

Second, the residual gender wage gap can be separated to gender-specific factors and wage structure component only if the residual gap does not change over time due to sample composition, measurement error, equation misspecification or a change in the distribution of unobserved characteristics. Since the aim of this paper is to quantify the role of brain and brawn skills on the gender wage gap trends rather than identifying the role of gender specific factors and the factors related to wage structure per se, this is less of a concern for the purpose of this study. Nonetheless, the forces that may affect the residual gender wage gap is discussed in Section 6 with providing descriptive evidence that residual gap attributed to gender–specific factors actually may be changes in discrimination as well as the changes in labor market institutions such as trade union density or employment protection.

[11] See Blau and Kahn (1997) for a similar approach.

3 Data, Concepts and Empirical Specification

3.1 Wage and Employment Data

For European countries, individual level data on wages and labor market characteristics comes from two different sources, European Community Household Panel (hereinafter, ECHP) and European Union Statistics on Income and Living Conditions (hereinafter, EU-SILC) provided by Eurostat. The ECHP is a panel survey of 15 European countries from 1994 to 2001, covering a wide range of topics like income, health, education, housing, demographics and employment characteristics. From 2001 the ECHP was succeeded by the EU-SILC. EU-SILC provides cross-sectional and longitudinal data on income, poverty, social exclusion and living conditions pertaining to individual-level changes over time, observed over a four year period since 2003. As a result, there is no single data source to study the long term dynamics of the wage structure in Europe. However, despite the differences between ECHP and EU-SILC, harmonizing some of the variables of the two datasets is possible.[12]

The key variable for this study is the gross hourly wage. Hence, the analysis are restricted to the countries which provide complete information on hourly wages in both surveys, namely Austria, Ireland, Italy, Portugal, Spain and the U.K. The analysis are based on the data from the initial wave of ECHP and cross-sectional component of the EU-SILC because of their representativeness.[13] Both surveys include information on demographic characteristics and employment of individuals.[14] For the U.S., the data come from Integrated Public Use Microdata Series (IPUMS) of Current Population Survey (hereinafter, CPS) March Supplements. The CPS survey years, 1994 and 2009, were selected to match the sample period of the ECHP and EU-SILC data used.

The country samples are restricted to individuals of working age, between 25 and 54 years old who are working at least 15 hours per week with valid observations on all the variables used

[12] Goos, Manning and Salomons (2009 and 2011) make use of wages from these two surveys to investigate job polarization trends in Europe.

[13] In the first wave of ECHP, in 1994, a sample of nationally represented households were interviewed in Ireland, Italy, Portugal, Spain and the U.K. Austria have joined the project in 1995. Data from EU-SILC is used from the 2009 cross-sectional component for all countries except 2008 for the U.K. due to the differences in income reference period.

[14] See Appendix A.1 and A.2 for the description of variables and procedures followed to construct country samples.

in the wage equations. Wage observations five times greater than the 99th percentile or lower than the half of the 1st percentile of the country wage distributions in each year are excluded from the country samples.

3.2 Data on Skill Requirements of Occupations

Brain and brawn skill requirements of occupations constructed using Occupational Information Network (hereinafter, O*Net) data developed by the U.S. Department of Labor. O*Net database is the most well known source for information on occupations in the U.S. labor market and it is a replacement for the Dictionary of Occupational Titles (DOT) which was extensively used in earlier research.[15] Recently, O*Net has been used to determine occupational skill requirements and task content of occupations also for several European countries.[16]

O*NET database provides detailed information about worker and job characteristics for more than 1110 occupations with a set of measurable descriptors. O*Net descriptors have the importance scale where O*Net rank each descriptor as not important at all (1), somewhat important (2), important (3), very important (4) or extremely important (5) to perform an occupation. A subset of the these descriptors is classified by O*Net as worker abilities and measures cognitive abilities, psycho-motor abilities and physical abilities required to perform at each occupations. [17] To construct brain skills, all the descriptors classified under cognitive abilities and to construct brawn skills all the descriptors classified under psycho-motor and physical abilities are used – twenty one different measures of cognitive ability intensity, ten measures of psycho-motor ability intensity and nine measures of physical ability intensity.[18] Appendix A.3 provides the list and the description of the variables used, organized by brain and brawn skill

[15]See Autor, Levy and Murnane, 2003; Goos and Manning, 2007; Bacolod and Blum, 2010.

[16]We follow the common practice in the literature on matching occupational skill requirements of the U.S. labor market with European datasets. See Amuedo-Dorantes and de la Rica (2011) for Spain; Ortega and Polavieja (2009) for 25 European countries to analyze the task specialization of immigrants and Goos, Manning and Salomons (2009, 2011) for analyzing the job polarization in 16 European countries.

[17]It is common in the literature to reduce the large number of descriptors to a relevant subset using textual definitions. See Amuedo-Dorantes and de la Rica, 2011; Autor, Levy and Murnane, 2003; Bacolod and Blum, 2010; Goos, Manning and Salomons, 2009; Peri and Sparber, 2009.

[18]We checked the consistency of this classification using principal component analysis performed among all set of descriptors under worker abilities title. Informed by this analysis, we categorized cognitive ability measures as brains and psycho-motor abilities together with physical abilities as brawns. The results of the principle component analysis performed among all set of attributes are available upon request. The details of the principle component analysis technique can be found in Appendix C.

type.[19]

3.2.1 Constructing Skill Requirements of Occupations

The occupation information in the individual level data is provided at aggregate level as eighteen occupation categories. To construct skill requirements of these eighteen occupation categories, we proceed as follows. First we convert the occupation codes used in O*Net to the codes used in individual level data.[20] Second, since under each of the eighteen occupation category there are several jobs classified, for each category, we calculate a weighted average of all the descriptor values. For this purpose, we use the percentage of workers employed in the U.S. labor market by 2001 for the jobs classified under the broad category as weights. It is important to note that matching O*Net data with European data relies on the assumption that the occupations in the U.S. and in Europe being examined herein are not different with regards to their skill requirements.

Next, we rescale O*Net descriptors to reflect the relative importance of each skill among all occupations. As pointed out by the earlier research, O*Net descriptors values range from one to five, but the score of each descriptor varies considerably across occupations. Peri and Sparber (2009) and Amuedo-Dorantes and de la Rica (2011) overcome this problem by rescaling the measures. Formally, let s_{kj} be the value of skill descriptor k for occupation j where $j = 1, 2, ...18$; and the maximum and minimum value of the descriptor s_k among occupations be $\overline{s_k}$ and $\underline{s_k}$. Each skill descriptor value is rescaled as the following: $s^*_{kj} = (s_{kj} - \underline{s_k})/(\overline{s_k} - \underline{s_k})$.

Finally, using the rescaled descriptor values, s^*_{kj}, we construct the measures of brain and brawn skills by taking the simple average of corresponding set of descriptors' rescaled values. Table 1 displays the occupations under the broad classification, as well as the brain and brawn skill summary measures for each of the occupations.

As presented in Table 1, occupations at the top of the brain skill measure distribution

[19]The remaining O*NET worker ability descriptors largely pertain to sensory dimension which we do not include in the analysis. We excluded the descriptors measuring sensory abilities mainly for two reasons. First, sensory abilities include descriptors that are not clearly being classified under one of the two sets (brains and brawns) according to their textual definitions. Second, they are related with some measures of cognitive abilities and at the same time with psycho-motor and physical abilities which prevents the clear classification of skills.

[20]See Appendix A.4 for the details of mapping 2010 Standard Occupational Code (SOC) used in the O*NET data to ISCO-88 codes.

Table 1: Brain and brawn skill intensity of occupations

Occupation code	Average of Rescaled Values: Brains	Brawns	Brains/Brawns	Occupation title
1112	0.86	0.33	2.59	Physical, mathematical, engineering, life science, health professionals
1300	0.78	0.1	7.94	Managers of small enterprises
2122	0.76	0.08	9.56	Teaching professionals
2300	0.74	0.16	4.68	Legislators, senior officials, corporate managers
2400	0.71	0.11	6.24	Other professionals
3132	0.65	0.51	1.27	Physical, engineering, life science, health associate professionals
3334	0.52	0.07	7.59	Teaching and other associate professionals
4142	0.51	0.78	0.65	Agricultural, fishery and related laborers
5100	0.49	0.87	0.56	Extraction, building, other craft and related trades workers
5200	0.47	0.78	0.6	Metal, machinery, precision, handicraft, printing and related trades workers
6100	0.47	0.83	0.56	Stationary-plant and related operators, drivers and mobile-plant operators
7174	0.45	0.33	1.36	Models, salespersons and demonstrators
7273	0.42	0.22	1.97	Office and customer services clerks
8183	0.38	0.62	0.6	Personal and protective services workers
8200	0.32	0.64	0.5	Machine operators and assemblers
9100	0.28	0.8	0.35	Skilled agricultural and fishery workers
9200	0.15	0.74	0.2	Laborers in mining, construction,manufacturing and transport
9300	0.02	0.53	0.03	Sales and services elementary occupations
Mean	0.50	0.47	2.63	
Std. dev.	0.23	0.30	3.12	
Pearson correlation coefficient	-0.58			

Note: Occupation codes are based on regrouped (group B) classification of ECHP data. If the occupations are regrouped, the first and the last two digits of the occupation code corresponds to the 2-digit ISCO-88 classification of occupations.

are professionals (physical, mathematical, engineering, life science, health and teaching), and legislators, senior officials and managers (corporate managers and managers of small enterprises). At the bottom of the brain skill distribution, there are laborers (in mining, construction,manufacturing and transport) and elementary occupations (sales and services). If occupations are ranked according to their brain skill requirements, the average difference in brain skill requirement between two consecutive positions in the occupational ranking is 0.05, which is equal to 1/4 standard deviation difference in brain skills (standard deviation of brawn skills is 0.2). On the other hand, occupations at the top of the brawn skill distribution are mainly blue-collar workers (extraction and building workers and stationary-plant operators). Once again, if the occupations are ranked according to their brawn skill requirements, again the average difference in brawn skill requirement between two consecutive positions implies, on average, 0.05 change in brawn skill measure which corresponds to a 1/6 standard deviation change in brawn skill requirement (standard deviation of brawn skills is 0.3).

Constructed skill measures are merged with the individual level data using the occupational allocation of individuals. This allocative process may result from different choices of individuals, discrimination in the process of recruitment or hiring or differences in comparative advantage of workers as in Roy (1951) which is taken as given over the time period of analysis. Moreover, the brain and brawn skill measures do not vary by worker within occupations. On the other hand, since there is no time variation in O*Net, the time variation in brain and brawn skill intensity differences between men and women comes only from the occupational differences. The results of the current analysis are valid only if the skill composition within occupations is constant over time. Throughout a long period, some skills might become idle for certain occupations possible due to change in the task content of occupations by technological progress. However, using DOT (earlier version of O*Net) Goos and Manning (2007) show that most of the overall changes in task composition of occupations in U.S. labor market happened between occupations not within occupations. Autor and Handel (2009) also provide evidence on the dominance of occupation as a predictor for the variation in the task measures using the individual level Princeton Data Improvement Initiative. Given the results of previous studies and considering the relatively recent and short length of our individual level data (from 1993 to 2008), it is reasonable to assume that any kind of progression might affect the distribution of skills and skill prices rather than the skill content of the occupations.

3.3 Empirical Specification

Using the matched data set, we implement the JMP decomposition by estimating the following specification:[21]

$$\begin{aligned} \ln Wage_{ijct} &= \beta_{1ct} + \beta_{2ct} Edu_{2ijct} + \beta_{3ct} Edu_{3ijct} + \beta_{4ct} Exp_{ijct} + \\ &+ \beta_{5ct} Exp^2_{ijct} + \beta_{6ct} Brains_{jct} + \beta_{7ct} Brawns_{jct} + u_{ijct} \end{aligned} \tag{4}$$

where $\ln Wage_{ijct}$ is the logarithm of gross hourly wage of male worker i employed in occupation j in country c at year t. Edu_2 and Edu_3 are dummies for the secondary and higher levels of educational attainment leaving the low level of educational attainment as the omitted

[21]We investigated the possibility of different functional forms using higher order polynomials in brains and brawns (including quadratic and cubic terms). In no case, these terms were statistically significant and had an effect on the ceteris paribus returns to other labor market characteristics.

category. Exp is the proxy for labor market experience. Finally, $Brains_{jct}$ and $Brawns_{jct}$ are the skill requirements of the occupation that the worker holds at time t.

Since brain and brawn skills can not be sold separately there is no market for skills we employ hedonic price model to determine the skill prices separately. Hence, occupations are assumed to be described by their bundle of skills, brains and brawns, and the ordinary least squares estimates for the skill coefficients in Equation 4 are interpreted as the marginal contributions of brains ($\partial \ln Wage/\partial Brains$) and brawns ($\partial \ln Wage/\partial Brawns$) to the logarithm of hourly wages.

4 Descriptive Analysis

4.1 Gender Wage Gap Trends

Table 2 summarize the main characteristics of the variables used in the analysis. As shown in Table 2, female workers on average earned less than males in all the countries in each year indicating the persistence of gender wage gaps. The unadjusted gender wage ratio in the U.S. was around 75% in 1993 and 79% in 2008.[22] The unadjusted female–male wage ratio for the European countries varied between 74% (for the U.K.) and 92% (for Italy) in 1993. During the 1990s and 2000s, the majority of European countries experienced a decline in the gender wage gap similar to the U.S. except Spain. The decline in the gender wage gap in European countries and the U.S. varied from the lowest 0.006 log points (in Portugal) to the highest 0.143 log points (in Ireland). By 2008, the unadjusted female–male wage ratio was lowest for the U.K. and for the U.S. (about 79% for both countries). In Spain, the unadjusted gender wage gap increased from 0.080 log points in 1993 to 0.115 log points in 2008.[23]

One obvious reason for closing of the gender wage gaps might be the improved labor market characteristics of women. In fact, during the period of analysis, women have been catching up with men in their educational attainment levels. By 2008, the share of women at higher

[22] The unadjusted gender wage gap ratio is calculated as $exp(Log(\text{hourly wage})^F - Log(\text{hourly wage})^M) * 100$.

[23] See the report by Eurofound on the increase in the gender wage gap in Spain during the late 1990s: http://www.eurofound.europa.eu/eiro/studies/tn0912018s/es0912019q.htm. Using data from the ECHP and EU-SILC, Guner, Kaya and Sánchez-Marcos (2014) also show that Spanish gender wage gap increased 0.074 log points from 1994 to 2004.

Table 2: Summary statistics on female and male workers, 1993 and 2008

Panel A. Descriptive statistics for 1993														
	Austria		Ireland		Italy		Portugal		Spain		U.K.		U.S.	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Log (hourly wage)	2.184	2.461	2.107	2.346	2.125	2.200	1.382	1.509	2.066	2.146	2.111	2.408	2.557	2.842
	(0.544)	(0.435)	(0.571)	(0.532)	(0.405)	(0.367)	(0.680)	(0.597)	(0.550)	(0.518)	(0.484)	(0.517)	(0.610)	(0.630)
Primary edu(%)	23.7	14.8	22.4	34.8	38.2	49.6	66.1	76.7	39.5	52.9	43.7	35.4	3.4	5.6
Secondary edu(%)	66.8	76.8	51.1	39.7	48.6	38.2	14.0	11.9	20.8	19.5	25.7	25.9	59.0	57.4
High edu(%)	9.5	8.4	26.3	25.3	12.8	11.3	13.3	9.2	39.7	27.6	30.2	38.0	37.6	37.0
Experience (year)	17.282	16.655	15.683	18.185	14.375	16.609	15.711	19.082	13.930	17.579	18.838	17.774	18.667	18.572
	(9.672)	(9.522)	(8.960)	(9.416)	(9.370)	(9.830)	(9.585)	(9.855)	(9.716)	(10.464)	(10.382)	(9.818)	(8.510)	(8.442)
Brain skills	0.411	0.479	0.502	0.508	0.453	0.446	0.435	0.474	0.455	0.474	0.499	0.547	0.515	0.496
	(0.193)	(0.172)	(0.214)	(0.198)	(0.194)	(0.177)	(0.192)	(0.161)	(0.246)	(0.181)	(0.207)	(0.199)	(0.210)	(0.222)
Brawn skills	0.375	0.550	0.330	0.488	0.372	0.508	0.446	0.570	0.355	0.546	0.329	0.435	0.318	0.464
	(0.231)	(0.293)	(0.213)	(0.300)	(0.265)	(0.293)	(0.280)	(0.289)	(0.243)	(0.291)	(0.216)	(0.290)	(0.229)	(0.300)
Number of obs.	952	1,440	951	1,365	1,597	2,551	1,229	1538	1,276	2,551	3,132	3,357	22,062	23,172
Panel B. Descriptive statistics for 2008														
	Austria		Ireland		Italy		Portugal		Spain		U.K.		U.S.	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Log (hourly wage)	2.324	2.523	2.570	2.666	2.170	2.228	1.614	1.735	2.203	2.318	2.612	2.842	2.665	2.899
	(0.438)	(0.441)	(0.502)	(0.504)	(0.406)	(0.369)	(0.552)	(0.548)	(0.476)	(0.443)	(0.518)	(0.524)	(0.630)	(0.672)
Primary edu(%)	14.3	8.5	12.4	20.0	26.5	42.1	54.6	68.2	26.1	36.3	8.0	7.0	3.4	5.4
Secondary edu(%)	50.3	56.7	24.4	23.2	43.2	39.9	19.3	17.4	24.4	23.4	52.5	52.2	48.0	52.9
High edu(%)	35.4	34.7	61.7	50.5	30.3	17.9	25.2	12.9	49.4	39.9	46.7	47.0	48.6	41.7
Experience (year)	19.828	22.168	16.496	18.226	15.436	17.315	17.967	20.675	15.473	18.451	14.831	15.622	19.828	19.820
	(9.178)	(9.178)	(9.212)	(10.018)	(8.454)	(8.953)	(9.570)	(10.241)	(8.759)	(9.286)	(10.390)	(9.529)	(9.251)	(9.055)
Brain skills	0.438	0.492	0.541	0.523	0.462	0.453	0.417	0.474	0.459	0.482	0.527	0.558	0.538	0.504
	(0.208)	(0.188)	(0.210)	(0.227)	(0.200)	(0.194)	(0.233)	(0.181)	(0.237)	(0.196)	(0.188)	(0.221)	(0.211)	(0.220)
Brawn skills	0.327	0.505	0.314	0.443	0.347	0.372	0.418	0.587	0.346	0.518	0.300	0.401	0.305	0.451
	(0.220)	(0.294)	(0.211)	(0.287)	(0.249)	(0.265)	(0.252)	(0.276)	(0.220)	(0.291)	(0.206)	(0.278)	(0.223)	(0.301)
Number of obs.	1,744	1,960	1,170	1,113	5,172	6,193	1,351	1,377	3,854	4,140	2,232	2,070	31,018	31,907

Data Source: For 1993-1994 sample, European Community Household Panel (for Ireland, Italy, Portugal, Spain, U.K. 1994 and for Austria 1995) and CPS March Supplements (for the U.S. 1994). For 2008 sample, European Union Statistics on Income and Living Conditions (EU-SILC, 2009) and CPS March Supplements 2009. *Notes:* See Appendix A.2 for variable definitions.

educational levels rose considerably as compared to 1993 in all countries in the sample. Although there was an increase in the share of higher educated males during this period, the increase was larger for females than males in all countries, again except for Spain. On the other hand, in 2008 women workers were more experienced than they were in 1993. However, from 1993 to 2008 the experience levels of men also increased. Hence, the male–female difference in experience levels persisted in most of the European countries to the detriment of women.[24] On the other hand, women in the U.S. became on average more experienced than men already at the beginning of 1990s.

Furthermore, from 1993 to 2008, both men and women shifted their occupational allocations to more brain skill and less brawn skill intensive occupations, except for Portugal.[25] In Portugal, women shifted their occupational allocation to less brain skill insensitive jobs and men to more brawn skill insensitive jobs. For all other countries in the sample, women moved into brain skill intensive occupations at a faster pace than men, except Spain. Hence, from 1993 to 2008, the gender gap in brain skill intensities increased favoring females. However, except Austria and the U.S., in all the sample of countries male workers moved out of brawn skill intensive occupations at a faster pace than women.

4.2 Brain and Brawn Skill Prices

How did the skill prices change? To answer this question, Table 3 presents the wage returns of males to brain and brawn skills in the U.S. and in the sample of European countries in 1993 and 2008.[26] In all the countries in common, brain skills were positively and significantly valued throughout the period, while the marginal contribution of brawn skills to the logarithm of hourly wages was negative and small.

To be concrete, as discussed in Section 3.2.1, a change in occupation associated with a 1/4 standard deviation increase in brain skill requirements such as going from having the brain

[24] In particular, from 1993 to 2008 the experience gap between males and females narrowed in Ireland, Italy, Portugal and Spain, while the gap widened in Austria and the U.K.

[25] See Tables B.1 and B.2 of Appendix B for the occupational allocation of males and females in 1993 and in 2008 for the sample of countries.

[26] As discussed in Section 2 Table 3 is based on the wage regression estimates using the male sample to simulate the wage equation in a nondiscriminatory labor market. See Tables D.3 and D.4 of the Appendix D for the estimation results using the males and females pooled sample and using only females, respectively.

skills required to be a protective service worker to be an office or service clerk. In 1993, such a skill premium was associated with the lowest 1.9 percent (for Italy) and with the highest 4.3 percent (for Portugal) rise in wages in the European countries in the sample. In 2008, the same occupational change was associated with the least 2.1 percent (for Italy) and the most 3.6 percent (for the U.S.) higher wage. On the other hand, in 1993, a change in occupation that implied a 1/6 standard deviation increase in brawn skill requirements received a wage penalty, penalty being highest in Portugal (with around 3.7 percent) and the lowest in the U.S. (with around 0.8 percent). By 2008, this penalty was lowest in Ireland (with around 0.7 percent) and highest in Portugal (with around 1.9 percent. From 1993 to 2008, returns to brain skills increased in Austria, Italy, the U.K. and the U.S., while in Ireland, Portugal and Spain brain skills became relatively less valuable.[27] On the other hand, during the same period, brawn skills became significantly less valuable in the U.S., as well as in Austria and in the U.K. In contrast to the U.S. experience, however, the brawn skill penalty declined in Southern European countries and in Ireland.

4.2.1 Robustness Checks

We carried out a range of robustness tests to check whether the estimated returns to brain and brawn skills are affected by the construction of skills discussed in Section 3.2.1 or the empirical specification (Equation 4). For checking whether the construction process of skill measures affects the wage equation estimates, a different technique, Principle Component Analysis is employed to generate brain and brawn skill measures. Principle Component Analysis is a data reduction technique which maximizes the amount of variation of the large number of variables explained by a smaller number of components (Jolliffe, 1986).[28] Using the brain and brawn skill measures constructed via Principle Component Analysis, we determine the skill intensity of jobs held by females and males and re-estimate the empirical model specified in Equation 4. Table D.6 in Appendix D provides the estimate of the male wage regression

[27] For each country, the changes in brain skill prices are statistically significant at 1% significance level.

[28] Principle Component Analysis has been commonly used in the literature to construct measures from DOT or O*Net data (Autor, Levy and Murnane, 2003; Bacolod and Blum, 2010; Goos, Manning and Salomons, 2009; Ortega and Polavieja, 2012). See Appendix C for a brief explanation of the technique and the procedure followed to construct skill measures and summary statistics of brain and brawn skills using this method.

Table 3: Wage Regression Estimates

	Austria		Ireland		Italy		Portugal		Spain		U.K.		U.S.	
	1994	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008
Secondary Education	0.117 ***	0.069 ***	0.244 ***	0.102 *	0.115 ***	0.094 ***	0.142 ***	0.281 ***	0.223 ***	0.142 ***	0.118 ***	0.130 **	0.432 ***	0.315 ***
	(0.039)	(0.031)	(0.039)	(0.051)	(0.022)	(0.022)	(0.037)	(0.051)	(0.030)	(0.027)	(0.015)	(0.059)	(0.047)	(0.030)
Higher Education	0.235 ***	0.249 **	0.395 ***	0.405 ***	0.350 ***	0.243 ***	0.538 ***	0.663 ***	0.349 ***	0.270 ***	0.275 ***	0.301 ***	0.739 ***	0.673 ***
	(0.073)	(0.044)	(0.061)	(0.064)	(0.039)	(0.035)	(0.123)	(0.175)	(0.054)	(0.040)	(0.035)	(0.053)	(0.072)	(0.057)
Experience	0.002	0.022 ***	0.033 ***	0.022 ***	0.007 **	0.027 ***	0.015 ***	0.043 ***	0.014 **	0.029 ***	0.022 ***	0.014 ***	0.043 ***	0.043 ***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.003)	(0.004)	(0.009)	(0.007)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)
Experience^2	0.000	- 0.000 **	- 0.000 ***	- 0.000	- 0.000	- 0.000 ***	- 0.000 ***	-0.001 ***	- 0.000	- 0.000 ***	- 0.000 ***	- 0.000 **	-0.001 ***	-0.001 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Brains	0.595 ***	0.614 ***	0.682 ***	0.570 ***	0.384 ***	0.438 ***	0.866 ***	0.450 **	0.724 ***	0.583 ***	0.649 ***	0.680 ***	0.568 ***	0.715 ***
	(0.137)	(0.059)	(0.150)	(0.127)	(0.101)	(0.106)	(0.216)	(0.208)	(0.192)	(0.103)	(0.132)	(0.070)	(0.136)	(0.123)
Brawns	-0.188 ***	-0.260 ***	-0.307 **	-0.141	-0.314 ***	-0.204 **	-0.731 ***	-0.371 **	-0.405 ***	-0.328 ***	-0.234 *	-0.291 ***	-0.166	-0.170 **
	(0.028)	(0.036)	(0.108)	(0.089)	(0.071)	(0.075)	(0.122)	(0.141)	(0.079)	(0.085)	(0.126)	(0.072)	(0.099)	(0.074)
Constant	2.097 ***	1.918 ***	1.546 ***	1.858 ***	2.004 ***	1.746 ***	1.289 ***	1.123 ***	1.693 ***	1.715 ***	1.790 ***	2.236 ***	1.592 ***	1.688 ***
	(0.121)	(0.089)	(0.110)	(0.115)	(0.056)	(0.101)	(0.133)	(0.139)	(0.121)	(0.126)	(0.116)	(0.060)	(0.153)	(0.141)
VIF(Brains)	1.38	1.43	1.73	1.88	1.28	1.26	1.42	1.45	1.46	1.43	1.63	1.51	1.73	1.86
VIF(Brawns)	1.26	1.31	1.55	1.7	1.46	1.32	1.38	1.55	1.37	1.49	1.48	1.51	1.78	1.87
R^2	0.15	0.26	0.34	0.37	0.32	0.25	0.43	0.36	0.36	0.37	0.25	0.23	0.25	0.27
Number of obs.	1,440	1,960	1,365	1,113	2,551	6,193	1,538	1,377	2,551	4,140	3,357	2,070	23,172	31,907

Notes: i) Occupational level clustered standard errors are in parentheses. (ii) *, ** and *** significant at 1, 5 and 10 % significance level respectively. (iii) The omitted category is taken as low level for education dummies. (iv) Variance inflation factor: VIF= $1/(1-R_i{}^2)$ and $R_i{}^2$ is the coefficient of determination of the regression equation where each explanatory variable regressed on all the other explanatory variables.

specified in Equation 4 using the skill measures constructed via Principle Component Analysis. A comparison of the returns to skills using these new measures of skills with the estimation results discussed in the previous section shows that construction process of skill measures does not alter our results.[29]

The second check for robustness focuses on the empirical specification. The empirical specification that is considered, excluding the brain and brawn skill measures, is simple but fairly standard in the literature (Blau and Kahn, 1997; Willis, 1986). However, the estimation of wage equation including brain and brawn skills simultaneously might exhibit collinearity. As presented in Table 1, brain and brawn skill measures are negatively correlated. The existence of collinearity would inflate the variances of the parameter estimates and can produce parameter estimates of the "incorrect sign" and of implausible magnitude (Greene, 1993). Taking into account this concern, the variance inflation factor (VIF), the collinearity diagnostic statistics is computed and presented in Table 3. VIF is based on the proportion of variance in the each independent variable that is not related to the other independent variables in the model. Conventionally, a variance inflation factor of ten or larger have been used as rule of thumb to indicate serious multicollinearity (Kennedy, 1992; Hair, Black, Babin and Anderson, 1995). As seen in Table 3 the mean variance inflation factor values for brain and brawn skill measures for each regression are much lower than ten indicating no collinearity.

As a third robustness check, we also consider an alternative specification that includes only the ratio of brain to brawn skill measures instead of the brain and brawn skill measures sepa-

[29]The skill measures constructed by the Principle Component Analysis are unit free as the rescaled skill measures, but note that the scale of measurement in both technique is different. Using skill measures constructed with another process produces negligible changes in the estimated coefficients. For example, the coefficient estimate for brain skills using the measures constructed by Principle Component Analysis is around 0.143 for the U.S. in 1993. In this case, the standard deviation of brain skill measure is one by construction (See Table C.2 in Appendix C). Then one standard deviation increase in brain skills is associated with 14.3% increase in hourly wages. Once again, if occupations are ranked according their brain skill requirements, a change in occupation implies on average 0.2 increase in brain skill measure, i.e. 1/5 standard deviation increase in brain skills. This change (1/5 standard deviation increase) is associated with 2.8% ($14.3 \times 1/5$) increase in hourly wages which is the same as the main estimations presented in Table 3 (0.568×0.05). A similar comparison can be done for the rest of the coefficients.

rately:

$$\begin{aligned} \ln Wage_{ijct} &= \beta_{1ct} + \beta_{2ct} Edu_{2ijct} + \beta_{3ct} Edu_{3ijct} + \\ &+ \beta_{4ct} Exp_{ijct} + \beta_{5ct} Exp^2_{ijct} + \beta_{6ct} (Brains/Brawns)_{jct} + u_{ijct}, \end{aligned} \tag{5}$$

where (Brains/Brawns) is the brain to brawn skill ratio of the job j that the individual i in country c held at time t. In this case, the coefficient estimate for brain to brawn ratio, β_6, reflects the marginal contribution of working in an occupation relatively more brain skill intensive than brawn skill. The full set of coefficient estimates from this specification is presented in Table D.5 of Appendix D. Once again, the estimation of this specification give positive and significant coefficient estimates for the brain to brawn ratio implying a positive return of working in a relatively more brain skill intensive occupation. We find that, in line with the results presented in the previous section, returns to brains to brawns ratio increased in the U.S. over time period of analysis. Among the European countries in the sample, in the U.K., return to brains to brawns ratio increased, in Austria and in Ireland did not change significantly, while in Southern European countries, the returns declined from 1993 to 2008.

5 Decomposition of the Changes in the Gender Wage Gap

We turn now to the decomposition results of the changes in the gender wage gaps from 1993 to 2008. The decomposition analysis results are presented in Panel B of Table 4. But before, several interesting descriptive findings regarding the changes in the gender wage gaps are presented in Panel A.

The first four rows of Panel A present the residual standard deviation for males and females (from own wage regressions specified by Equation 4) in 1993 and 2008. A higher residual standard deviation indicates a higher wage inequality within education, experience and brain–brawn skill levels. As shown in Table 4, both male and female residual wage inequalities were much higher in the U.S. than any other European country in 1993 and 2008. Moreover, during this period, the residual wage inequality for males and females increased in the U.S.

Table 4: Decomposition of the change in gender wage gap, 1993 vs 2008

Panel A. Descriptive statistics	Austria	Ireland	Italy	Portugal	Spain	U.K.	U.S.
Male residual SD*							
1993	0.402	0.433	0.303	0.451	0.416	0.448	0.545
2008	0.378	0.399	0.320	0.440	0.352	0.458	0.575
Female residual SD**							
1993	0.488	0.434	0.308	0.416	0.399	0.395	0.533
2008	0.353	0.421	0.332	0.340	0.344	0.449	0.546
Mean female residual from male wage regression							
1993	-0.264	-0.274	-0.125	-0.192	-0.155	-0.272	-0.333
2008	-0.188	-0.143	-0.103	-0.218	-0.146	-0.229	-0.311
Mean female residual percentile***							
1993	31.14	32.62	38.76	35.89	39.08	30.71	32.01
2008	33.72	39.82	40.64	33.61	38.65	34.24	33.46
Panel B. Decomposition of the change in gender wage gap $\Delta\overline{lnW_{2008}} - \Delta\overline{lnW_{1993}}$							
Change in gender wage gap	**-0.078**	**-0.143**	**-0.018**	**-0.006**	**0.035**	**-0.066**	**-0.051**
Gender wage gap-1993 ($\Delta\overline{lnW_{1993}}$)	0.277	0.238	0.075	0.127	0.080	0.297	0.285
Gender wage gap-2008 ($\Delta\overline{lnW_{2008}}$)	0.199	0.096	0.057	0.121	0.115	0.231	0.234
(1) Observed X's	**0.007**	**-0.041**	**-0.042**	**-0.071**	**0.015**	**-0.013**	**-0.023**
Education variables	-0.003	-0.016	-0.030	-0.044	0.010	-0.021	-0.018
Experience variables	0.019	-0.017	-0.002	-0.009	-0.007	0.018	0.003
Brains	-0.009	-0.017	0.007	0.015	0.004	-0.011	-0.008
Brawns	-0.001	0.009	-0.017	-0.033	0.008	0.001	0.000
(2) Observed Prices	**-0.009**	**0.030**	**0.045**	**0.039**	**0.030**	**-0.010**	**-0.006**
Education variables	-0.003	0.001	0.014	-0.018	0.008	0.000	-0.001
Experience variables	0.006	0.006	0.010	0.019	0.011	-0.005	0.000
Brains	0.001	0.002	0.001	-0.024	-0.003	0.001	-0.005
Brawns	-0.013	0.022	0.021	0.061	0.013	-0.006	-0.001
(3) Unobserved Prices	**-0.003**	**-0.018**	**0.004**	**0.006**	**-0.032**	**0.010**	**0.013**
(4) Gap	**-0.073**	**-0.114**	**-0.025**	**0.020**	**0.022**	**-0.053**	**-0.034**
Sum gender-specific $(1+4)$	-0.066	-0.155	-0.067	-0.051	0.037	-0.066	-0.058
Sum wage structure $(2+3)$	-0.012	0.012	0.049	0.045	-0.002	0.000	0.007

Notes: The change in the differential is the change in the male-female log wage differentials between 1993 and 2008. * Estimated using male wage regression. ** Estimated using female wage regression. *** Computed by assigning each women a percentile ranking in the indicated year's residual male wage distribution and calculating the female mean of these percentiles.

Similarly, Ireland, Italy, and the U.K. experienced increases in the male and female residual wage inequality. In contrast to the U.S. experience, however, in Austria, Portugal and Spain residual wage inequality for both genders declined.

The mean female residual from the male wage regression and the mean female residual percentile presented in the following four rows of Panel A. The mean female residual from the male wage regression is generally interpreted as a measure of discrimination but might also capture the omitted productivity differences between males and females (Blau and Kahn, 1997). On the other hand, the mean female residual percentile show the progression of females within education, experience and skill groups similar to the change in the absolute value of the female residuals from male wage regression. The results presented in Table 4 indicates that the mean female residuals from the male wage equation are lower for the U.S. than all the European countries in the sample. In other words, after controlling for education, experience as well as the brain and brawn skills, the gender gap was highest in the U.S. in both years. Moreover, by 2008, the mean female residual percentile was higher compared to 1993 in all countries, except Portugal and Spain. In Portugal, from 1993 to 2008, the mean female residual decreased which resulted in a lower ranking of the mean female residual percentile. On the other hand, in Spain although the the mean female residuals (from male regression) increased from -0.155 in 1993 to -0.146 in 2008, women did not move up within the residual wage distribution of males.

The unadjusted gender wage gaps in 1993 and in 2008 as well as the change in the gap between these two years is presented in Panel B. During 1990s–2000s all European countries experienced a decline in the unadjusted gender wage gap in common, except Spain. Despite the common trend in the gender wage gap in other countries, the rate of convergence varies substantially across countries, from 0.006 log points (in Portugal) to 0.143 log points (in Ireland). How did the changes in skill prices affect gender wage gap trends? In the U.S., from 1993 to 2008, 17% of the closing gender wage gap can be explained by changes in returns to brain and brawn skills. Similar to the U.S. experience, in Austria and in the U.K. a part of the convergence in the gaps was due to changes in skill prices, about 15.4% of the gender wage gap in the former and around 7.8% in the later. However, in contrast to the U.S. experience, the changes in returns to brain and brawn skills had a widening effect on the gender wage gap in Ireland, Italy, Portugal and Spain. The main reason for this is that, in all Southern European countries and in Ireland, brawn skills became more valuable, skills that women had an initial deficit. Hence, if the occupational allocation of men and women had remained constant, this should have widen the gender wage gaps. Despite the decline in brain skill prices in Portugal and in Spain that favored women since women had also initial deficit in brain skills, the change in brawn skill prices reclaimed the potential gains of women from the decline in brain skill prices.

What accounts for the convergence of the gender wage gaps? We start with the contribution of "Observed X's" and "Observed Prices". "Observed X's" heading indicates the contribution of the labor market characteristics and skills to the gender wage gap trends in each country for 1990s-2000s. The "Observed Prices" captures the effect of changing returns to characteristics and prices of skills on the gender wage gap trends. We start with the only exception, Spain that experienced an increase in the unadjusted gender wage gap, around 0.035 log points. As seen in Table 4, the change in gender differences in labor market characteristics and skills as well as the the change in returns to these characteristics and skills were responsible for the widening gender wage gap in Spain. On the other hand, in other European countries, a substantial part of the convergence in the gender wage gaps, cannot be explained by observed factors. The part of the convergence explained by changes in characteristics and skills and returns to these characteristics and skills is only around 2.5% for Austria, 7.7% for Portugal, 18.8% for Ireland and 35% for the U.K. In Italy, however, the changes in these factors cannot explain the closing gender wage gap. However, in the U.S. the changes in "Observed X's" and "Observed Prices" accounted for more than the half (around 57%, $0.029/0.051$) of the convergence in the gender wage gap.

The change in male residual wage inequality measured as "Unobserved Prices" had a widening effect on the gender wage gap in Italy, Portugal, and the U.K. as well as the U.S. Hence female workers in these countries were adversely affected from the increase of male residual inequality (see Male residual SD in Table 4). On contrary, the decline in the wage residual inequality in Austria and Spain favored women. However, as seen in Table 4, the change in male residual wage inequality is unlikely to explain the gender wage gap trends.

Turning the attention to the last term of the decomposition, "Gap effect" reveals the importance of the unexplained factors on gender wage gap trends. In countries where females moved up in the male residual wage distribution (see Mean female residual percentile in Table 4), namely in Austria, Ireland, Italy, the U.K. and the U.S., the gap effect is negative. In other words, womens progression within groups had a narrowing effect on the gender wage gap from 1993 to 2008 in these countries. In Portugal and in Spain, however, the gap effect had a widening effect on the gender wage gaps. What is striking is that, the "Gap effect" accounts for a substantial part of the convergence of the gender wage gap in most of the countries. The part of the convergence gender wage gap due to unobserved factors measured as "Gap effect" is around 94% for Austria, 80% for Ireland and the U.K., and 66% for the U.S. In Spain, "Gap effect" accounts for the 63% of the widening gender wage gap. In Italy, the unexplained part reach a value greater than 100% implying if observed and unobserved prices, i.e. wage structure would not have changed, the gender wage gap would narrow even further. In Portugal, the gap effect is not able to explain the convergence in the gender wage gap.

6 The Role of Labor Market Institutions

One of the results of this study arise from the fact that a substantial part of the changes in the gender wage gaps can not be explained by the changes in the gender gaps in labor market characteristics and skills or changes in the wage structure (measured as gap effect presented in Table 4). It is important to note that, not only the convergence in unobserved skills of males and females, but also the changes in the labor market institutions and/or the changes in discrimination would be captured by the gap effect. Indeed, Gayle and Golan (2012) develop a model of the labor supply, occupational sorting and human capital accumulation in which statistical discrimination and a wage gap arise endogenously. They use this dynamic equilibrium model to quantify the driving forces behind the decline in the gender earnings gap in the U.S. They find that for the period 1967–1997 decline in the statistical discrimination accounts for a large fraction of the decline in the gender wage gap in the U.S.

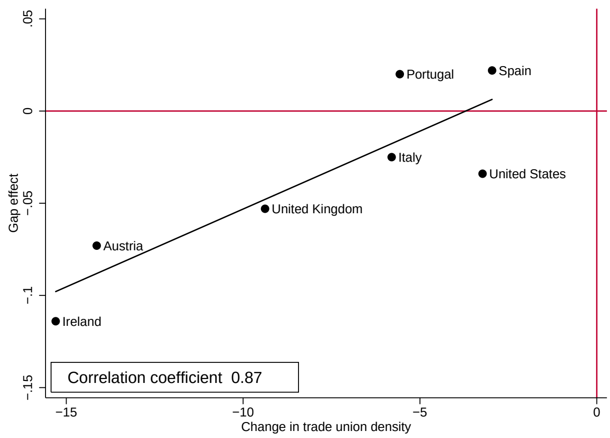
Actually, women in the U.S. advanced in the male residual distribution much more during 1980s compared to 1990s–2000s (measured as the mean female residual percentile).[30] If a part of the gap effect is due to statistical discrimination and such a sharp decline in statistical

[30]See Tables D.1 and D.2 of Appendix D for the wage regression estimates and decomposition of the change in the U.S. gender wage gap from 1979 to 1988

discrimination occurred in the U.S. during the 1980s, this might partially explain the larger contribution of the gap effect to the change in the gender wage gap during the 1980s as compared to the later decades. Indeed, the main bulk in women's improvement in labor market characteristics and skills and the increase in the labor market commitment of women in the U.S. occurred during the 1980s which might contributed to a reduction in statistical discrimination (Blau and Kahn, 2000). On the other hand, as shown by Pissarides, Garibaldi, Olivetti, Petrongolo and Wasmer (2005), more generous institutions that compress the wage distribution will also tend to decrease the gender wage gap. In other words, the changing positions of females in the male wage residual distribution might be not only due to the convergence in unobservable skills of females and males, but also to the decline in the discrimination or changes in the labor market institutions. Unfortunately, the harmonized data used in this study lack variables on either union status or union coverage. However, for further investigating the issue, this section provide descriptive evidence on the role of some of these factors on the changing gender wage gap via gap effect. For this purpose the relationship between the gender wage gaps that can not be explained by the characteristics and skills or returns to these characteristics and skill prices (measured as gap effects presented in Table 4) and the changes in various measures that captures the labor market institutions and discrimination is explored. We find that the changes in these measures are highly correlated with the gap effect. In other words, the change in female's progression in each country's male residual wage distribution might be capturing the effect of changes in some of these measures or a combination of them.

The measures that capture the changes in the labor market flexibility include the change in the trade union density as well as the OECD employment protection measures for temporary and regular workers. The change in trade union density in each country is defined as the change in the percentage of employees who are members of a trade-union from 1993 to 2008. For calculating the change in employment protection of regular and temporary workers in each country, the changes in the value of two OECD indicators from 1993 to 2008 are used. These two OECD indicators are individual dismissal of workers with regular contracts and regulation of temporary contracts.[31]

[31] To find out more about the employment protection measures see www.oecd.org/employment/protection.

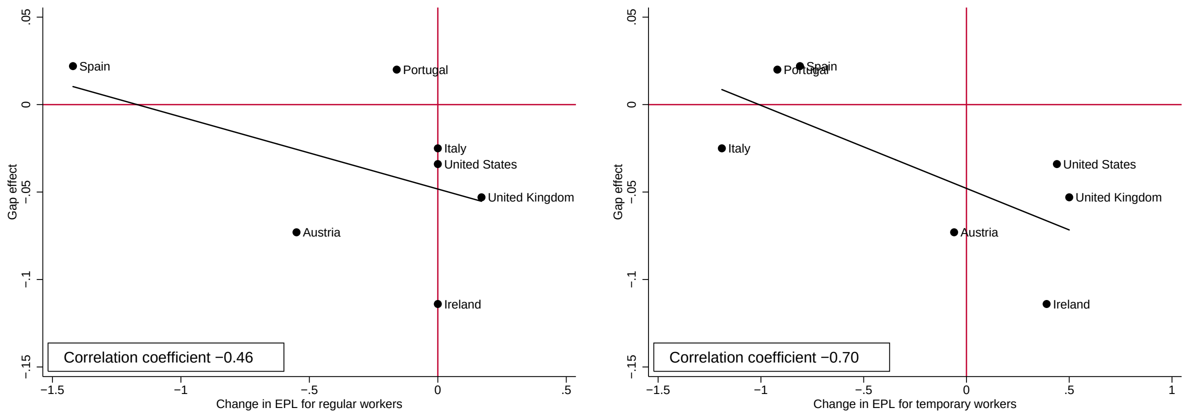


Note: The change in trade union density is defined as the change in the percentage of employees who are members of a trade-union from 1993 to 2008.

Figure 1: The Change in Trade Union Density and The Gap Effect

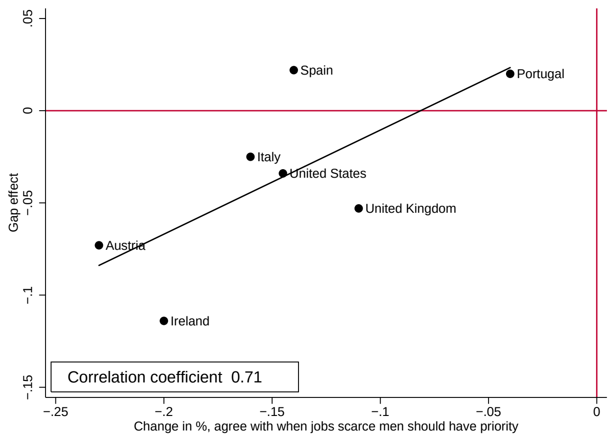
Figure 1 relates the change in trade union density from 1993 to 2008 to the gap effects presented in Table 4. As Figure 1 shows, in all countries there was a decline in trade union density from 1993 to 2008. The higher the decline in the trade union density, the larger the gap effect implying a higher rank for females within the residual wage distribution of males. This might be explained by the fact that the decline in the trade union membership might decrease the unionization gap between men and women given the higher trade union membership among men. Hence, the declining unionization rate had a larger negative impact on male than female workers narrowing the gender wage gap (Blau and Kahn, 1997 and 2000).

Moreover, as presented in Figure 2 the correlation coefficient between the change in employment protection and the gap effect is high and negative. Although the change in employment protection of regular workers do not show a clear pattern, countries with higher employment protection of temporary workers in 2008 compared to 1993, are also countries in which the gap effect tends to decrease the gender wage gap. Indeed, for 25–34 years old age group, Pissarides et al. (2005), find that stricter employment legislation for temporary contracts tend to decrease the gender wage gap, while the stricter employment protection legislation for regular workers tend to increase the gap since it is designed to protect the "insiders in the labor market and have a larger a negative impact on the young. Since there is no measure of discrimination, an indirect measure is generated to capture the attitudes towards gender. Using data from



Note: The change in employment protection of regular workers is defined as the change in the OECD indicator of individual dismissal of workers with regular contracts and the change in employment protection of temporary workers is defined as the change in the OECD indicator of regulation of temporary contracts from 1993 to 2008.

Figure 2: The Change in Employment Protection of Workers and The Gap Effect



Note: The change in attitudes toward gender roles is measured as the change in mean response in World Value Survey to the statement 'When jobs are scarce, men should have more right to a job than women' from 1993 to 2008 (0-1 scale: 0 indicates no agreement with the statement, 1 indicates complete agreement with the statement).

Figure 3: The Change in Attitudes Toward Gender Roles and The Gap Effect

World Values Surveys (WVS), the proportion of people in each country agree with the statement "When jobs are scarce, men should have more right to a job than women is computed. Although this is an imperfect measure for discrimination, the change in the proportion of people that agree with this statement would capture the change in attitudes about gender roles in work. As Figure 3 presents, the correlation coefficient between the change in attitudes toward gender roles from 1993 to 2008 and the gap effect is high and positive. In other words, if in 2008 less people agree with the statement "When jobs are scarce, men should have more right to a job than women" as compared to 1993, the higher the ranking of females in the male residual distribution. This is parallel to the argument that the decline in the statistical discrimination would allow female's progression leading a decline in the gender wage gap.

7 The Role of Selection

A further issue to note is the substantial increase in women's employment rates over time.[32] Despite the use of male wage regression in analysis ameliorates the problem due to changes in non-random selection into work, the estimated gap effects that are attributed to the change in labor market institutions may include the impact of changes in unmeasured selectivity of women participants to the labor market. Earlier studies emphasized the importance of selection in explaining the gender wage gap trends.[33] For instance, Blau and Kahn (2006) study changes in the U.S. gender wage gap between 1979 and 1998 and find that sample selection implies that the 1980s gains in womens relative wages were overstated and that selection may also explain part of the slowdown in convergence between male and female wages in the 1990s. Mulligan and Rubinstein (2008) also argue that in the U.S. between 1975 and 2001, selection into employment shifted from negative to positive for women, and the narrowing of the gender wage gap during this period reflects changes in female workforce composition.

The sign of the bias is ex-ante unpredictable, since the selected group might be positively or negatively selected in terms of their unobserved characteristics (Blau and Beller, 1988; Blau and Kahn, 1997). Moreover the selection process into work may be different for women compared to men and selection rule may have changed with the large changes in employment rates. For instance, if women in 1990s who were employed tend to have relatively high-wage characteristics, an increase in women's employment rates may understate the convergence of the gender wage gap may be understated since there will be more women in the labor market who

[32]See Figures B.1 and B.2 of Appendix B for the trends in employment rates in the European countries and the U.S. respectively.

[33]See Olivetti and Petrongolo (2008) for how nonrandom selection into work may affect international comparisons of gender wage gaps.

tend to have relatively low-wage characteristics. On the other hand, if the market becomes more positively selective over time, the convergence in the gender wage gap will be overstated.

We explore the possible contribution of sample selection to the narrowing of the gender wage gap by re-decomposing the gender wage gap trends in the selection corrected model. Additionally, we explore whether the labor market institutions are related to unexplained part of the gender wage gap trends, i.e. gap effect, even after the selectivity correction. Correction for selection into work is implemented here using a two-stage Heckman (1979) selection model. In particular, we estimate probit participation equations for males and for females in each country for each year and reestimate the wage regressions:

$$lnW_t^M \quad = \quad X_t^M \beta_t + S_t^M \gamma_t + \sigma_t \theta_t^M + \psi_t \lambda_t^M, \tag{6}$$

where λ is the inverse Mills ratio derived from a probit participation equation and is a measure of the selection bias, and ψ is its estimated coefficient in the wage equation that measures the wage effects of selection.[34] Then, we decompose the change in the gender wage gap over time in the selection corrected model as:

$$\begin{aligned} \Delta lnW_s - \Delta lnW_t &= [(\Delta X_s - \Delta X_t)\beta_s + (\Delta S_s - \Delta S_t)\gamma_s] \\ &+ [\Delta X_t(\beta_s - \beta_t) + \Delta S_t(\gamma_s - \gamma_t)] \\ &+ (\Delta\theta_s - \Delta\theta_t)\sigma_s \\ &+ \Delta\theta_t(\sigma_s - \sigma_t) \\ &+ (\Delta\lambda_s - \Delta\lambda_t)\psi_s + \Delta\lambda_t(\psi_s - \psi_t). \end{aligned} \tag{7}$$

The decomposition of the change in gender wage gaps in the selection corrected model in Equation 8 refines the JMP decomposition model (Equation 3) as it includes now two additional components, $(\Delta\lambda_s - \Delta\lambda_t)\psi_s$ and $\Delta\lambda_t(\psi_s - \psi_t)$. Neuman and Oaxaca (1998) show that wage decompositions are sensitive to the way the selection term is interpreted. Since, our interest is

[34] The left-hand side in the probit is whether or not a person is employed and the probit is identified by including brains and brawns in the wage function but not in the probit function, and by including the non-labor family income, number of children and a dummy for the presence of 0-6 years old children in the probit function. Probit results are not reported here but are available on request.

the unexplained part of change in the gender wage gaps after correcting for the nonrandom selection to employment, i.e. "Selectivity-corrected gap effect", we follow the simplest approach suggested by Gupta et. al. (2008), in which gender differences in the selectivity over time are treated as a separate component of the wage decomposition (the last two terms of equation 7).

Table 5: Addressing selection bias: Selectivity-corrected gender wage gaps

	Austria	Ireland	Italy	Portugal	Spain	U.K.	U.S.
Change in gender wage gap	-0.078	-0.143	-0.018	-0.006	0.034	-0.066	-0.051
Gap effect	-0.073	-0.114	-0.025	0.020	0.022	-0.053	-0.034
Selectivity corrected-gap effect	-0.030	-0.578	0.046	-0.079	-0.011	-0.334	0.183

Selectivity-corrected gap effect is based on estimating the selection corrected model using a two-stage Heckman (1979) selection model. See text for details.

The results change somewhat for the selection corrected model, particularly with respect to the gap effect.[35] Table 5 presents the change in the gender wage gaps from 1993 to 2008 for each country once again and the gap effect from the JMP decomposition of the selection corrected model (measured as *selectivity corrected-gap effect* in Table 5). For comparison, Table 5 also includes the gap effect estimated from the model that does not account for selection. First, for Ireland, Portugal, Spain and the U.K., even after selection correction, a substantial part of the changes in the gender wage gaps remains unexplained. Hence, in these countries due to sample selection, the gains in womens relative wages were understated. Moreover, the unexplained part of the gender wage gap trends becomes negative for Portugal and Spain in the selection corrected model, going from 0.020 log points to -0.079 log points for Portugal and from 0.022 log points to -0.011 log points for Spain. Correcting for selection reveals the convergence in the unexplained part of the gender wage gap in Portugal and Spain from 1993 to 2008. On the other hand, for Italy and the U.S. the unexplained part of the gender wage gap trends becomes positive in the selection corrected model indicating that controlling for the inverse Mills ratio, the gender wage gap widened slightly in these countries. In Austria, the convergence in the gender wage is overstated due to selection, however about 38%(-0.030/-

[35]Table E.1 of the Appendix E provides the contribution of each component to the change in gender wage gaps from 1993 to 2008 based on the selection corrected model.

0.078) of the slowdown in the narrowing of the gender wage gap cannot be explained neither by labor market characteristics and skills or changes in the wage structure nor selection.

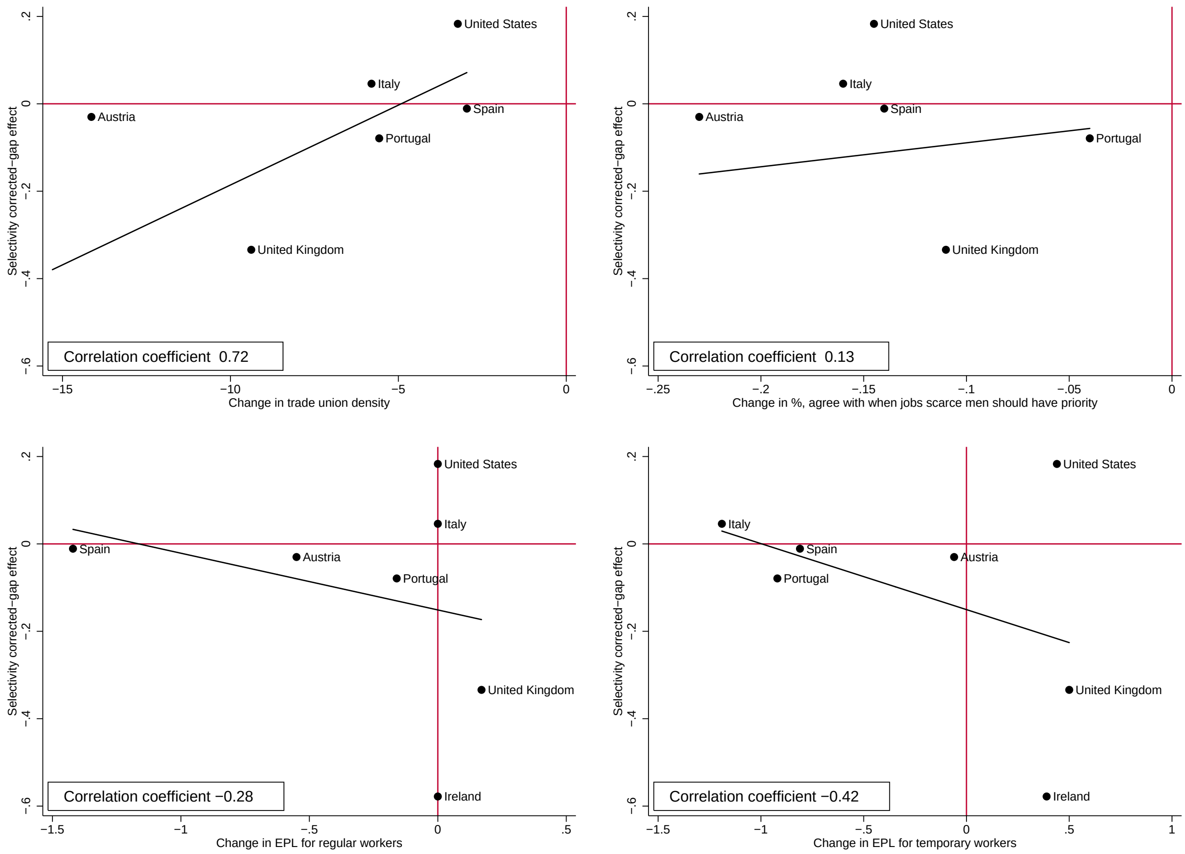


Figure 4: The Change in Labor Market Institutions, Discrimination and The Selectivity-corrected Gap Effect

In Figure 4, we present again the relationship between the unexplained part of the gender wage gap trends, "Selectivity-corrected gap effect and the change in labor market institutions. The correlation coefficient between the gap effect and changes in measures that captures the labor market institutions and discrimination decrease when selection bias is corrected. In particular the correlation between the gap effect and the change in proportion of people that agree with the statement "when jobs are scarce, men should have more right to a job than women" decreases from 0.62 to 0.13. This is not surprising, considering that selection into work and the changes in attitudes towards gender roles may be altered together. However, despite the decline in correlation coefficients, the change in the trade union density and the change in employment protection of regular workers are still strongly correlated with the selectivity corrected-gap effect. This evidence confirms that, to the extent that labor market institutions are an important component in explaining the degree of overall gender wage gap trends.

8 Concluding Remarks

The recent literature focusing on the U.S. emphasizes the role of various skills required by occupations and changing prices of those skills on the closing gender wage gap. In this paper, we explore the recent gender wage gap trends in various European countries as well as in the U.S. using the direct measures of skill requirements of jobs.

Our findings reveal that, although in Austria and in the U.K., similar to the U.S. experience, a part of the closing gender wage gap can be explained by the changes in brain and brawn skill prices, the increase in returns to brain skills and decrease to brawn skills was not a common phenomenon for the Southern European countries Italy, Portugal and Spain and for Ireland. In contrast to the U.S. experience, in Southern European countries and in Ireland, the changes in returns to brain and brawn skills had a widening effect on the gender wage gaps. Nevertheless, from 1993 to 2008 the gender wage gaps declined in the sample of European countries, except Spain. However, a substantial part of the changes in the gender wage gaps cannot be explained by the changes in observable gender-specific factors (i.e. labor market characteristics or brain and brawn skills) or changes in wage structure (i.e. returns to characteristics, skill prices or residual wage inequality).

Other factors that may have contributed to the convergence of the unexplained gender pay gap include changes in selection to the employment, changes in gender differences in unobservable skills and labor market discrimination, as well as the changes in labor market institutions. The results of this study reveal the relation between the changing attitudes toward gender and/or the labor market flexibility and the unobservable gender specific factors that contribute to closing gender wage gap even after the non-random selection to employment is corrected.

References

[1] Amuedo–Dorantes, Catalina, and Sara de la Rica. 2011. "Complements or substitutes? Task specialization by gender and nativity in Spain." *Labour Economics* 18(5):697–707.

[2] Autor, David H., Frank Levy, and Richard Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118(4):1279–1333.

[3] Autor, David H., and Michael J. Handel. 2009. "Putting Tasks to The Test: Human Capital, Job Tasks and Wages." *NBER Working Paper* 15116.

[4] Bacolod, Marigee P., and Bernardo S. Blum. 2010. "Two Sides of the Same Coin: U.S. Residual Inequality and the Gender Gap." *Journal of Human Resources* 45(1):197–242.

[5] Black, Sandra E., and Alexandra Spitz-Oener. 2010. "Explaining Womens Success: Technological Change and the Skill Content of Womens Work." *The Review of Economics and Statistics* 92(1):187–94.

[6] Blau, Francine D., and Andrea H. Beller. 1988. "Trends in earnings differentials by gender, 1971-1981." *Industrial and Labor Relations Review* 41(4):513–29.

[7] Blau, Francine D., and Lawrence M. Kahn. 1997. "Swimming Upstream: Trends in the Gender Wage Differential in the 1980s." *Journal of Labor Economics* 15(1):1–42.

[8] Blau, Francine D., and Lawrence M. Kahn. 2000. "Gender Differences in Pay." Journal of Economic Perspectives 14(4):75–99.

[9] Blau, Francine D., and Lawrence M. Kahn. 2006. "The us gender pay gap in the 1990s: Slowing convergence." *Industrial and Labor Relations Review* 60(1):45–66.

[10] Borghans, Lex, Bas ter Weel, and Bruce Weinberg. 2006. "People People: Social Capital and the Labor-Market Outcomes of Underrepresented Groups." NBER Working Paper 11985.

[11] CEC: Commission of the European Communities. 2003. "Gender pay gaps in European Labor markets." *Commission Staff Working Paper (SEC)* 93.

[12] Datta Gupta, Nabanita and Oaxaca, Ronald L. and Smith, Nina. 2001. "Swimming Upstream, Floating Downstream: Trends in the U.S. and Danish Gender Wage Gaps". *CLS Working Paper* No. 01–06.

[13] Dolado, Juan J., Florentino Felgueroso Fernndez and Juan Jimeno-Serrano. 2000. "Explaining youth labor market problems in Spain: crowding-out, institutions, or technology shifts?." *IZA Discussion Paper* 142.

[14] Gayle, George-Levi, and Limor Golan. 2012. "Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968–1997." *Review of Economic Studies* 79(1): 227–67.

[15] Goos, Maarteen, and Alan Manning. 2007. "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain." *Review of Economics and Statistics* 89(1):118–33.

[16] Goos, Maarteen, Alan Manning, and Anna Salomons. 2009. "Job polarization in Europe." *American Economic Review Papers and Proceedings* 99(2):58–63.

[17] Goos, Maarteen, Alan Manning, and Anna Salomons. 2011. "Explaining Job Polarization: The Roles of Technology, Offshoring and Institutions." Center for Economic Studies, Discussion Paper 11.34.

[18] Greene, William H. 1993. Econometric Analysis, 2nd edn. New York: Macmillan.

[19] Guner, N., Kaya, E. & Sánchez-Marcos, V. (2014). Gender Gaps in Spain: Policies and Outcomes over the Last Three Decades. *SERIEs (the Journal of Spanish Economic Association)*, 5(1): 61–103.

[20] Hair, F. Hair Jr., William C. Black, Barry J. Babin, and Rolph E. Anderson. 1995. Multivariate Data Analysis, 3rd ed. New York: Macmillan.

[21] Heckman, James J. 1979. "Sample selection bias as a specification error." *Econometrica* 47(1):153–63.

[22] Jolliffe, Ian T. 1986. Principal Component Analysis. Springer, New York.

[23] Juhn, Chinhui, Kevin Murphy, and Brooks Pierce. 1991. "Accounting for the Slowdown in Black-White Wage Convergence." In *Workers and Their Wages*, pp. 107–43, ed. Marvin Kosters. Washington, D.C.: American Enterprise Institute.

[24] Juhn, Chinhui, Kevin Murphy, and Brooks Pierce. 1993. "Inequality and the Rise in Returns to Skill." *Journal of Political Economy* 101(3):410–42.

[25] Katz, Lawrence and David Autor. 1999. "Changes in the Wage Structure and Earnings Inequality." In *Handbook of Labor Economics, Vol. 3A*, eds. Orley Ashenfelter and David Card, 1463–558. North-Holland Elsevier Science.

[26] Katz, Lawrence and Kevin Murphy. 1992. "Changes in Relative Wages, 19631987: Supply and Demand Factors." *Quarterly Journal of Economics* 107(1):35–78.

[27] Kennedy, Peter. 1992. A Guide to Econometrics. Oxford: Blackwell.

[28] Kunze, Astrid. 2008. "Gender wage gap studies: consistency and decomposition." *Empirical Economics* 35(1): 63–76.

[29] Manski, Charles F. 1989. "Anatomy of the selection problem." *Journal of Human Resources* 24(3):343–60.

[30] Mulligan, Casey B., and Yona Rubinstein 2008. "Selection, investment and women's relative wage over time." *Quarterly Journal of Economics* 123(3): 1061-1110.

[31] Neuman, Shoshana, and Ronald L. Oaxaca. 2004. "Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note." *Journal of Economic Inequality* 2(1):3–10.

[32] Olivetti, Claudia, and Barbara Petrongolo. 2008. "Unequal pay or unequal employment? A cross-country analysis of gender gaps." *Journal of Labor Economics* 26(4): 621-654.

[33] Ortega, Francesc, and Javier G. Polavieja. 2012. "Labor–market exposure as a determinant of attitudes toward immigration." *Labour Economics* 19(3):298–311.

[34] Peracchi, Franco. 2002. "The European Community Household Panel: A Review." *Empirical Economics* 27(1):63–70.

[35] Peri, Giovanni, and Chad Sparber. 2009. "Task Specialization, Immigration, and Wages." *Applied Economics* 1(3):135–69.

[36] Pissarides, Christopher, Pietro Garibaldi, Claudia Olivetti, Barbara Petrongolo, and Etienne Wasmer. 2005. "Women in the labour force: How well is Europe doing?" In *Women at Work: An Economic Perspective*, eds. Tito Boeri, Daniela Del Boca and Christopher Pissarides, Oxford: Oxford University Press.

[37] Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers* 3(2):235-246.

[38] Suen, Wing. 1997. "Decomposing wage residuals: Unmeasured skill or statistical artifact?." *Journal of Labor Economics* 15(3): 555–66.

[39] Welch, Finis. 2000. "Growth in Womens Relative Wages and in Inequality among Men: One Phenomenon or Two?" *American Economic Review Papers and Proceedings* 90(2):444-49.

[40] Willis, Robert J. 1986. "Wage determinants: A survey and reinterpretation of human capital earnings functions." In *Handbook of Labor Economics Vol.1*, eds. Orley Ashenfelter and Richard Layard, 525–602. North-Holland Elsevier Science.

Appendices

Appendix A. Data Appendix

A.1. Variables

Gross hourly wage: Gross hourly wages from ECHP and EU-SILC are constructed by dividing gross monthly wages for the previous year by monthly hours worked in the main job in that year. The income reference period in the ECHP and in EU-SILC is the calendar year preceding the year of data collection in all countries except the U.K. and Ireland. In Ireland it is the 12 months prior to the interview, and in the United Kingdom it refers to the period around the date of interview. Hence ECHP data on wages refers to the year 1993 except for Austria to the year 1994 since Austria joined the survey in 1995. EU-SILC data refers to the year 2008. Wages are converted in 2005 PPP units using the purchasing power parity (PPP) exchange rates and then deflated by using the harmonized consumer price index (HCPI=2005). Both surveys include supplementary information on PPP exchange rates and HCPI is extracted from OECD Main Indicators database. In March CPS the annual earnings are top-codded. Following Katz and Murphy (1992) and Blau and Kahn (1997) procedure, the top-codded values are multiplied by 1,45. Then, the hourly wage from March CPS is constructed using the annual gross wages for the previous year divided by the product of weeks worked and average hours worked per week in that year. Wages are deflated by using the consumer price index (CPI=2005).

Occupation: The occupation information in ECHP and EU-SILC is defined using the International Standard Classification of Occupations (ISCO-88) and codded at the two-digit level. Occupation variable from March CPS is reclassified based on European classification using the ISCO-SOC crosswalk made available by the Center for Longitudinal Studies in the U.K. at http://www.cls.ioe.ac.uk/text.asp?section=000100010005001600002.

Education: The education variable from ECHP and EU-SILC is harmonized by using the International Standard Classification of Education (hereinafter, ISCED) categories. High educational qualifications are defined as ISCED categories 5-7, and include recognized third level education. Secondary education is defined by ISCED categories 3 and 4, and includes all second stage of secondary level education. Low education is defined as having no qualifications or only qualifications below the secondary education level, and corresponds to ISCED categories 0-2. Educational attainment variable from March CPS is reclassified based on the ISCED categories using the mapping provided byUNESCO Institute for Statistics (UIS) http://www.uis.unesco.org/education/ISCEDmappings/Pages/default.aspx.

Labor Market Experience: EU-SILC, provides the exact number of years spent in paid work with two exceptions; Ireland and the U.K. For Ireland and the U.K. the missing information on experience in EU-SILC is proxied using the years passed after the highest level of education was attained. On the other hand, ECHP lacks the information on actual labor market experience. However it provides information about the age of individuals at the highest level of education completed and at the beginning of the working life as well as the number of continuous months of unemployment before current job. Using these variables we generate a proxy for labor market experience. To proceed more formally, let y_t denote the year of the survey, y_s the year when the individual attained the highest education level, y_w the year when the individual began working life and m_u the number of continuous months of unemployment before current job ($y_u = m_u/12$ in years). The measure for labor market experience for individuals who completed their education earlier than starting to the working life (if $y_s >= y_w$) is computed as $exp = y_t - y_w$ and for the ones who started the working life before completing their highest education degree (if $y_s > y_w$) as $exp = y_t - y_s$. Then, the measure for labor market experience is partially corrected by subtracting the continuous months of unemployment before current job ($exp^* = exp - y_u$). Since March CPS does not provide information about the actual labor market experience, we use potential labor market experience variable that is age–years of schooling–6. Values for years of schooling is imputed using the educational attainment levels suggested by Jaeger (1997).

A.2. Sample

ECHP samples come from the initial wave of each country which is representative for the corresponding year. Although the ECHP aims at being both cross-sectionally and longitudinally representative, due to non random attrition and demographic changes arising from the arrival of new waves of immigrants, its cross-sectional representativeness tends to fade away over time (See Peracchi (2002) for an overview of ECHP data). On the other hand, EU-SILC is a four-year rotated panel and provides two types of annual data: cross-sectional data and longitudinal data observed periodically over a four-year period. The analysis are based on the cross-sectional component of EU-SILC which is representative for the corresponding year.

The ECHP and EU-SILC samples are restricted to individuals of working age, between 25 and 54 years old. In ECHP age is top-coded at 85 years in wave 1, 86 years in wave 2, and so on, for all countries, whilst age at first job is top coded for all countries and waves at 60 years. As we are mostly concerned with working age population, these top-coding rules are relatively unimportant. The sample is further restricted to individuals who are working at least 15 hours per week with valid observations on all the variables used in the wage equations. As suggested by Commission of the European Communities, gender wage gap "ought to be based on data covering the whole economy, including all sectors and firm sizes, including possibly also those working less than 15 hours per week" (CEC, 2003). However, the restriction of working at least 15 hours per week is necessary because of the nature of ECHP, since ECHP does not distinguish individuals regularly working less than 15 hours from those out-of the labor force in the first two waves. Finally, wage observations five times greater than the 99th percentile of the country wage distributions in each year or lower than 1$ per hour are excluded from the sample.

The U.S. samples are constructed using the same rules as the ECHP and EU-SILC samples. In particular, sample is restricted to workers 25-54 aged, not living in group quarters, not in school, not working without pay, working at least 15 hours per week and with positive number of years of potential labor market experience with non-missing variable responses. Finally, wage outliers five times greater than 99th percentile of the U.S. wage distributions in each year or lower than 1$ per hour are dropped.

A.3.Descriptors comprising the skill measures

Variables comprising BRAIN SKILLS measure

O*Net Descriptor	Description
oral comprehension	listening and understanding information and ideas presented through spoken words and sentences.
written comprehension	reading and understanding information and ideas presented in writing.
oral expression	communicating information and ideas in speaking so others will understand.
written expression	communicating information and ideas in writing so others will understand.
fluency of ideas	coming up with a number of ideas about a topic.
originality	coming up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.
problem sensitivity	telling when something is wrong or is likely to go wrong.
deductive reasoning	applying general rules to specific problems to produce answers that make sense.
inductive reasoning	combining pieces of information to form general rules or conclusions.
information ordering	arranging things or actions in a certain order or pattern according to a specific rule or set of rules.
category flexibility	generating or using different sets of rules for combining or grouping things in different ways.
mathematical reasoning	choosing the right mathematical methods or formulas to solve a problem.
number facility	adding, subtracting, multiplying, or dividing quickly and correctly.
memorization	remembering information such as words, numbers, pictures, and procedures.
speed of closure	quickly making sense of, combining, and organizing information into meaningful patterns.
flexibility of closure	identifying or detecting a known pattern that is hidden in other distracting material.
perceptual speed	quickly and accurately comparing similarities and differences among sets of letters, numbers, objects, pictures, or patterns.
spatial orientation	knowing the location in relation to the environment.
visualization	imagining how something will look after it is moved around or when its parts are moved or rearranged.
selective attention	concentrating on a task over a period of time without being distracted.
time sharing	shifting back and forth between two or more activities or sources of information.

Variables comprising BRAWN SKILL measure

O*Net Descriptor	Description
arm-hand steadiness	keeping hand and arm steady while moving arm or while holding arm and hand in one position.
manual dexterity	quickly moving hand, hand together with arm, or two hands to grasp, manipulate, or assemble objects.
finger dexterity	making precisely coordinated movements of the fingers of one or both hands to grasp, manipulate, or assemble very small objects.
control precision	quickly and repeatedly adjusting the controls of a machine or a vehicle to exact positions.
multi limb coordination	coordinating two or more limbs while sitting, standing, or lying down.
response orientation	choosing quickly between two or more movements in response to two or more different signals.
rate control	timing movements or the movement of a piece of equipment in anticipation of changes in the speed and/or direction of a moving object or scene.
reaction time	quickly responding to a signal when it appears.
wrist-finger speed	making fast, simple, repeated movements of the fingers, hands, and wrists.
speed of limb movement	quickly moving the arms and legs
static strength	exerting maximum muscle force to lift, push, pull, or carry objects.
explosive strength	using short bursts of muscle force to propel oneself, or to throw an object.
dynamic strength	exerting muscle force repeatedly or continuously over time.
trunk strength	using abdominal and lower back muscles to support part of the body repeatedly or continuously over time without 'giving out' or fatiguing.
stamina	exerting physically over long periods of time without getting winded or out of breath.
extent flexibility	bending, stretching, twisting, or reaching with body, arms, and/or legs.
dynamic flexibility	quickly and repeatedly bending, stretching, twisting, or reaching out with body, arms, and/or legs.
gross body coordination	coordinating the movement of arms, legs, and torso together when the whole body is in motion.
gross body equilibrium	keeping or regaining body balance or stay upright when in an unstable position.

A.4. Mapping of O*Net-SOC Occupational Codes to ISCO Codes

15th edition of O*Net occupational coding is based on SOC 2010, but there exist differences between two occupation codes. O*Net splits up several SOC 2010 occupations into multiple separate occupations. O*Net includes 1110 occupations with detailed information in the database for 974 of them, while SOC 2010 includes 840 detailed occupations. 667 occupations in 15th edition of O*Net are at SOC level which we have the ability requirements and employment shares of these occupations in the 2001 U.S. labor market. However, 37 SOC level occupations are divided into multiple categories at O*Net level. For instance, SOC 2010 code 11-3031 is Financial Managers, which O*Net provides information on ability requirements, but divides up this category into further two categories 11-3031.01, Treasurers and Controllers; and 11-3031.02 Financial Managers, Branch or Department. For these two categories, we have their ability requirements separately but we do not have their employment shares separately. We have dealt with these O*NET categories by simply taking the descriptor values for the main 37 occupation titles (for this example the values for 11-3031, Financial Managers are taken into account). For 269 occupations in 15th edition O*Net do not exist in SOC 2010 separately. For instance, SOC code 13-2011 is Accountants and Auditors. O*NET divides it into 13-2011.01 Accountants; and 13-2011.02 Auditors and provides the ability requirements of detailed categories (for 13-2011.01 and 13-2011.02) but does not provide the ability requirements of the main category (13-2011). Since we do not have the employment shares of detailed categories, we deal with these categories by taking a simple mean of the descriptor values to determine the skill requirement of the main title (for this example we took the simple average of descriptor values for occupations 13-2011.01 and 13-2011.02 to determine the skill requirement of 13-2011 Accountants and Auditors). There is one exceptional case in O*Net classification 19-1020.01 Biologists which does not exist in SOC classification, which we excluded from the analysis. Information on abilities is collected for 854 occupations among those 1100.

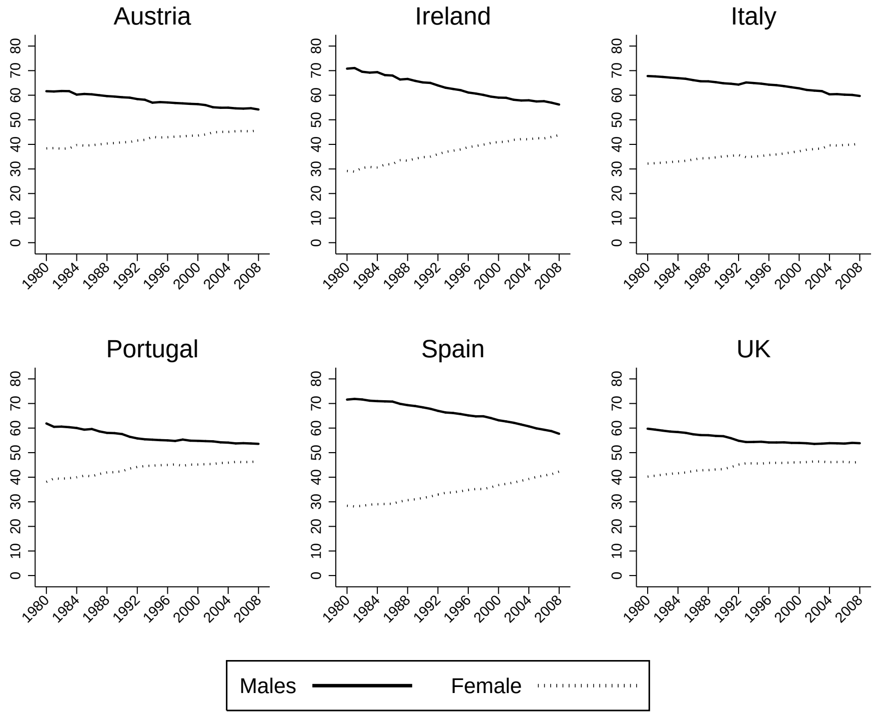
After determining the descriptor values of all SOC 2010 level occupations we proceed as follows: First, we matched ISCO codes with SOC 2000 codes using the ISCO-SOC 2000 made available by the Center for Longitudinal Studies in the U.K. at

http://www.cls.ioe.ac.uk/text.asp?section=000100010005001600002. Then, using SOC 2000-

SOC 2010 crosswalk provided by Integrated Public Use Microdata Series (IPUMS-USA) we matched ISCO codes with SOC 2010 codes. O*Net codes are matched with ISCO codes using these two crosswalks. Finally, using the employment shares of SOC 2010 occupations for 2001 derived from Occupational Employment Statistics Survey 2010 by Bureau of Labor Statistics, we determine the descriptor values of broader occupational titles. In total 849 O*Net occupations are classified under broad categories of ISCO level occupations with total employment share 97% in 2001 in the U.S. labor market.

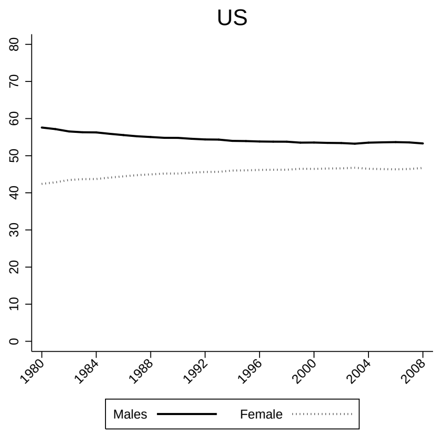
For Portugal, EU-SILC does not differentiate two occupational categories: 1112, Legislators, senior officials and corporate managers and 1300, Managers of small enterprises. Only for Portugal, these two occupations are aggregated while determining the descriptor values of broad level occupations.

Appendix B. Selected Labor Market Statistics



Source: OECD Employment Statistics, 2011.

Figure B.1: Employment rates in the European countries, % of labor force ages 15-64: 1980-2008



Source: OECD Employment Statistics, 2011.

Figure B.2: Employment rate in the U.S., % of labor force ages 15-64: 1980-2008

Table B.1: Female share and concentration in occupations, 1993

	Austria		Ireland		Italy		Portugal		Spain		UK		US	
	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F
Legislators, senior officials and corporate managers	0.14	1.46	0.12	1.22	0.15	0.83	0.27	1.03	0.09	0.60	0.30	7.11	0.51	0.27
Managers of small enterprises	0.21	1.08	0.20	1.88	0.00	0.00	0.15	0.18	0.26	0.21	0.34	2.15	0.40	9.19
Physical, mathematical, engineering, life science and health professionals	0.40	2.19	0.51	10.52	0.29	1.91	0.31	2.08	0.38	6.24	0.34	5.30	0.51	6.02
Teaching professionals	0.52	2.31	0.57	9.95	0.69	9.46	0.67	5.81	0.63	13.26	0.62	7.45	0.70	7.33
Other professionals	0.42	0.62	0.32	3.16	0.33	0.82	0.44	2.16	0.46	3.43	0.57	4.72	0.54	8.45
Physical, engineering, life science and health associate professionals	0.25	4.82	0.34	3.65	0.48	5.64	0.39	15.44	0.26	2.94	0.57	6.31	0.62	5.64
Teaching and other associate professionals	0.46	11.60	0.37	6.95	0.44	9.19	0.57	11.88	0.42	11.39	0.45	5.84	0.50	5.02
Office and customer services clerks	0.69	29.46	0.67	28.33	0.50	36.20	0.61	18.88	0.49	18.22	0.71	27.97	0.74	30.14
Personal and protective services workers	0.57	12.11	0.44	10.04	0.39	3.73	0.70	14.20	0.45	11.02	0.65	11.61	0.61	11.92
Models, salespersons and demonstrators	0.79	10.41	0.76	7.23	0.41	2.89	0.38	3.71	0.43	6.37	0.83	6.59	0.42	3.17
Skilled agricultural and fishery workers	0.10	0.27	0.00	0.00	0.23	0.81	0.41	2.10	0.06	0.21	0.13	0.16	0.18	0.49
Extraction, building, other craft and related trades workers	0.11	2.90	0.05	0.76	0.26	7.15	0.38	13.63	0.12	3.91	0.22	1.53	0.06	0.38
Metal, machinery, precision, handicraft, printing and related trades workers	0.08	1.79	0.07	1.00	0.18	4.65	0.11	1.71	0.04	1.08	0.05	0.75	0.06	0.68
Stationary-plant and related operators, drivers and mobile-plant operators	0.05	0.77	0.02	0.35	0.07	0.83	0.03	0.47	0.01	0.10	0.04	0.37	0.11	0.94
Machine operators and assemblers	0.36	1.88	0.47	7.36	0.08	0.26	0.43	4.50	0.24	2.05	0.41	3.99	0.46	0.50
Sales and services elementary occupations	0.72	11.83	0.56	5.40	0.48	9.56	0.73	10.46	0.65	15.59	0.68	6.18	0.46	2.49
Agricultural, fishery and related laborers	0.34	0.21	0.07	0.24	0.52	3.61	0.54	1.89	0.25	1.36	0.23	0.20	0.00	0.00
Laborers in mining, construction,manufacturing and transport	0.30	4.28	0.16	1.94	0.22	2.44	0.21	1.12	0.15	2.00	0.35	1.77	0.35	7.37
Dissimilarity Index	47.64		41.86		29.30		34.79		42.28		39.93		34.06	

Note: F_i/T_i = female employees in occupation i as percentage of total employees in occupation i. F_i/F = female employees in occupation i as a percentage of female employees. The dissimilarity index (ID) is calculated as follows: $ID = [1/2 \sum[F_i/F - M_i/M]] * 100$. The ID has a minimum value 0 when there is same percentage of female and male in each occupation and a maximum value of 100 when each occupation is completely male or female.

Table B.2: Female share and concentration in occupations, 2008

	Austria		Ireland		Italy		Portugal		Spain		UK		US	
	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F	F_i/T_i	F_i/F
Legislators, senior officials and corporate managers	0.26	2.18	0.42	12.17	0.32	0.90			0.24	0.80	0.35	10.31	0.45	1.00
Managers of small enterprises	0.30	0.83	0.52	2.52	0.38	0.80	0.26	1.63	0.29	0.32	0.36	1.60	0.41	9.45
Physical, mathematical, engineering, life science and health professionals	0.36	2.40	0.50	8.14	0.39	3.16	0.49	4.91	0.48	7.61	0.23	3.38	0.58	6.13
Teaching professionals	0.56	6.03	0.62	8.08	0.64	5.29	0.69	4.34	0.61	9.08	0.67	8.14	0.72	9.41
Other professionals	0.55	3.64	0.59	8.52	0.54	3.36	0.72	5.86	0.52	4.38	0.64	6.18	0.52	11.40
Physical, engineering, life science and health associate professionals	0.23	4.13	0.43	3.10	0.39	7.85	0.22	2.15	0.39	4.04	0.62	9.08	0.64	5.74
Teaching and other associate professionals	0.57	14.75	0.46	4.08	0.65	20.52	0.69	9.99	0.47	8.44	0.54	9.01	0.49	3.38
Office and customer services clerks	0.70	26.16	0.73	19.64	0.56	19.41	0.63	15.62	0.67	24.96	0.79	25.74	0.72	26.60
Personal and protective services workers	0.64	11.99	0.62	16.27	0.54	8.87	0.70	15.89	0.53	10.97	0.72	14.51	0.60	13.23
Models, salespersons and demonstrators	0.73	9.57	0.76	8.75	0.67	6.48	0.76	6.32	0.68	8.31	0.74	4.99	0.42	3.31
Skilled agricultural and fishery workers	0.50	0.61	0.10	0.05	0.15	0.39	0.23	0.80	0.13	0.29	0.04	0.06	0.12	0.30
Extraction, building, other craft and related trades workers	0.05	0.78	0.03	0.26	0.19	4.09	0.29	8.47	0.12	2.19	0.06	0.32	0.09	0.66
Metal, machinery, precision, handicraft, printing and related trades workers	0.03	0.36	0.04	0.33	0.13	1.74	0.04	0.51	0.04	0.46	0.04	0.26	0.04	0.37
Stationary-plant and related operators, drivers and mobile-plant operators	0.02	0.29	0.03	0.18	0.07	0.98	0.04	0.53	0.05	0.54	0.06	0.34	0.13	1.01
Machine operators and assemblers	0.38	1.23	0.43	1.15	0.37	3.88	0.46	4.26	0.26	1.47	0.28	1.30	0.53	0.24
Sales and services elementary occupations	0.73	11.71	0.47	3.17	0.64	10.60	0.73	17.33	0.71	13.45	0.46	4.01	0.55	3.37
Agricultural, fishery and related laborers	0.36	0.26	0.13	0.09	0.41	1.29	0.34	0.15	0.37	0.69	0.31	0.21	0.12	0.00
Laborers in mining, construction,manufacturing and transport	0.25	3.07	0.24	3.53	0.09	0.38	0.38	1.25	0.23	2.02	0.12	0.55	0.26	4.40
Dissimilarity Index	46.57		32.35		36.28		44.50		37.97		42.18		34.06	

Note: F_i/T_i = female employees in occupation i as percentage of total employees in occupation i. F_i/F = female employees in occupation i as a percentage of female employees. The dissimilarity index (ID) is calculated as follows: $ID = [1/2 \sum[F_i/F - M_i/M]] * 100$. The ID has a minimum value 0 when there is same percentage of female and male in each occupation and a maximum value of 100 when each occupation is completely male or female.

Appendix C. Principal Component Analysis (PCA)

C.1. PCA Technique

Principle Component Analysis (PCA) is a variable reduction technique which maximizes the amount of variance accounted for in the observed variables by a smaller group of variables called components. The components are not latent factors. PCA is not a model based technique and involves no hypothesis about the substantive meaning of or relationships between latent factors. Technically, let the random vector $\mathbf{X}' = [X_1, X_2, ..., X_p]$ be our observable measures with the covariance matrix $\sum$ with eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant ...\lambda_p \geqslant 0$. The linear combinations:

$$
\begin{aligned}
Y_1 &= \mathbf{a}_1'\mathbf{X} = \mathbf{a}_{11}X_1 + a_{12}X_2 + ... + a_{1p}X_p \\
Y_2 &= \mathbf{a}_2'\mathbf{X} = \mathbf{a}_{21}X_1 + a_{22}X_2 + ... + a_{2p}X_p \\
&... \\
Y_p &= \mathbf{a}_p'\mathbf{X} = \mathbf{a}_{p1}X_1 + a_{p2}X_2 + ... + a_{pp}X_p
\end{aligned}
$$

with $Var(Y_i) = \mathbf{a}_i' \sum \mathbf{a}_i$ and $Cov(Y_i, Y_k) = \mathbf{a}_i' \sum \mathbf{a}_k, i, k = 1, 2, ..., p$ are the principle components i.e. components are uncorrelated linear combinations $Y_1, Y_2, ..., Y_p$ whose variances are as large as possible. Principle components are then defined by:

$$
\begin{aligned}
\text{First principle component} &= \text{linear combination } \mathbf{a}_1'\mathbf{X} \\
\text{that maximizes } Var(\mathbf{a}_1'\mathbf{X}) \text{ st. } \mathbf{a}_1'\mathbf{a}_1 = 1 & \\
\text{Second principle component} &= \text{linear combination } \mathbf{a}_2'\mathbf{X} \\
\text{that maximizes } Var(\mathbf{a}_2'\mathbf{X}) \text{ st. } \mathbf{a}_2'\mathbf{a}_2 = 1 & \quad \text{and } Cov(\mathbf{a}_1'\mathbf{X}, \mathbf{a}_2'\mathbf{X}) = \mathbf{0} \\
& \quad ... \\
i^{th} \text{ principle component} &= \text{linear combination } \mathbf{a}_i'\mathbf{X} \\
\text{that maximizes } Var(\mathbf{a}_i'\mathbf{X}) \text{ st. } \mathbf{a}_i'\mathbf{a}_i = 1 & \quad \text{and } Cov(\mathbf{a}_i'\mathbf{X}, \mathbf{a}_k'X) = 0 \\
\text{for } k \neq i &
\end{aligned}
$$

PCA can be also performed based on the correlation matrix instead of variance-covariance matrix. If the correlation matrix is used, the variables are standardized and the total variance will equal the number of variables used in the analysis since each standardized variable has a variance equal to one. The use of correlation-matrix is necessary when the variables have different scales of measurement or not measured in a natural scale. The number of principle components is decided based on the cumulative variance explained by the components. As a rule of thumb, the first components that explains at least 50% - 70% of the cumulative variance are taken. Kaiser criterion also suggests not to keep components with an eigenvalue of less than 1, since these components account for less variance than the original variable does. Scree plots can be used to represent the ability of principle components in explaining the variation in data by showing the eigenvalues, and hence the variance explained by each component. Moreover, component loadings of each variable involved in the analysis help to interpret the constructed components since they show the weight of each variable in forming the components' scores.

C.2. Constructing Skill Requirement of Occupations by PCA

Principle Component Analysis (PCA) based on the correlation matrix has been widely used in the early research to construct task or skill measures from the various descriptors of DOT or O*Net data (Autor, Levy and Murnane, 2003; Bacolod and Blum, 2010; Goos, Manning and Salomons, 2009; Ortega and Polavieja, 2012). A common approach is performing separate PCAs for different sets of selected standardized descriptors (mean zero and variance one) and using the first principle component of each analysis as the summary measure for that particular set of descriptors (Autor, Levy and Murnane, 2003; Bacolod and Blum, 2010; Goos, Manning and Salomons, 2009). An alternative way is constructing the measures using a joint PCA performed using all standardized descriptors and selecting the components that explain the substantial part of the variance. However, in the case of jointly performed PCA, by construction, principle components will be orthogonal to each other. Building skill measures using a principle component analysis of the all O*Net descriptors will be ruling out the possible complementary or substitution, i.e. correlation between skill measures a priori.

Following the earlier literature, we also construct alternative skill measures via PCA. First, we performed two separate PCAs using the O*Net ability descriptors of O*Net occupations (using the 849 O*Net occupations, those are matched with ISCO level occupations as explained in Appendix A). One PCA is performed among the cognitive ability descriptors and the other among psycho-motor ability descriptors together with physical ability descriptors. The first component of the first PCA explains around 50% of the variation among the cognitive ability descriptors, while most of the variation among the psycho-motor and physical ability descriptors are explained by first principal component (around 72% of the variation). Figure C.1 visually presents the ability of first principle components of each analysis to explain the variation in corresponding descriptor values. Principal components based on transformation of correlation matrix to eigen-basis coordinates are unit free. If the loadings of all descriptors related to the same skill in the corresponding component is positive, then a higher component score implies a higher intensity in that skill. Table C.1 presents the component loadings of each descriptor involved in the analysis. All the cognitive ability descriptors have positive weights on the first principle component of the former PCA analysis (with one exception: Spatial Orientation), while all psycho-motor and physical ability descriptors have positive loadings on the first component of the later PCA without any exception. Hence we call the first and the second components as "brains" and "brawns", respectively.

Then again, to determine the brain skill and brawn skill requirement of broad classification of occupations we make use of the employment shares of SOC 2010 codded occupations for 2001 are derived from Occupational Employment Statistics Survey 2010 by Bureau of Labor Statistics. Basically, we took the weighted average of the component scores of occupations under the broad title where the weights are employment shares. Finally, we standardized the skill measures (mean 0, standard deviation 1). Table C.2 presents the summary statistics of brain and brawn skill measures constructed by this procedure.

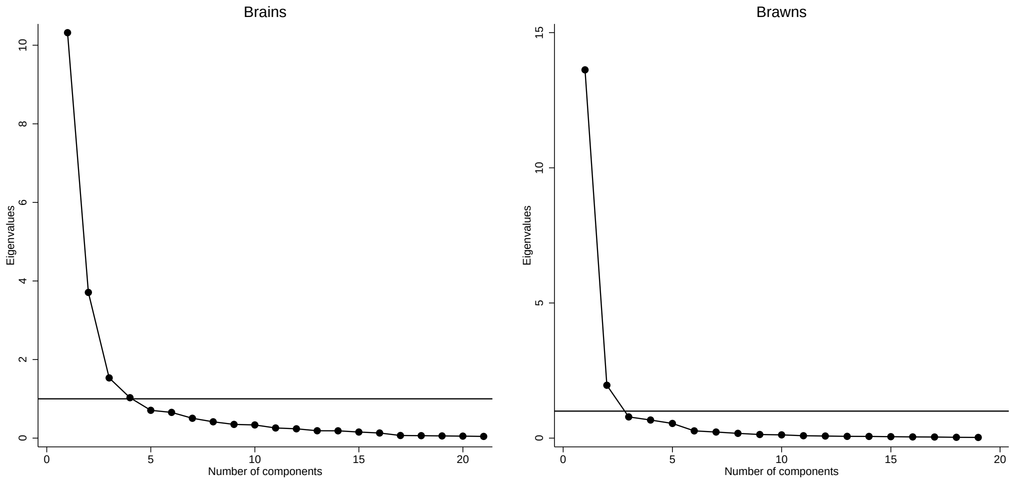


Figure C.1: Scree plot of eigenvalues after separate PCA

Table C.1: Principal component loadings

Brains		Brawns	
Descriptor	Component Loading	Descriptor	Component Loading
Oral Comprehension	0.219	Arm-Hand Steadiness	0.241
Written Comprehension	0.245	Manual Dexterity	0.241
Oral Expression	0.202	Finger Dexterity	0.177
Written Expression	0.240	Control Precision	0.233
Fluency of Ideas	0.258	Multilimb Coordination	0.258
Originality	0.243	Response Orientation	0.238
Problem Sensitivity	0.242	Rate Control	0.235
Deductive Reasoning	0.278	Reaction Time	0.241
Inductive Reasoning	0.270	Wrist-Finger Speed	0.214
Information Ordering	0.249	Speed of Limb Movement	0.241
Category Flexibility	0.251	Static Strength	0.257
Mathematical Reasoning	0.213	Explosive Strength	0.123
Number Facility	0.198	Dynamic Strength	0.250
Memorization	0.235	Trunk Strength	0.240
Speed of Closure	0.229	Stamina	0.246
Flexibility of Closure	0.216	Extent Flexibility	0.252
Perceptual Speed	0.143	Dynamic Flexibility	0.140
Spatial Orientation	-0.050	Gross Body Coordination	0.244
Visualization	0.092	Gross Body Equilibrium	0.234
Selective Attention	0.183		
Time Sharing	0.173		

Table C.2: Brain and brawn skill intensity of occupations, using skill measures constructed by PCA

Occupation code	Principle Component Values: Brains	Brawns	Occupation title
1112	1.11	-1.05	Legislators, senior officials and corporate managers
1300	1.14	-1.27	Managers of small enterprises
2122	1.62	-0.57	Physical, mathematical, engineering, life science and health professionals
2300	1.29	-1.28	Teaching professionals
2400	1.06	-1.22	Other professionals
3132	0.64	0.08	Physical, engineering, life science and health associate professionals
3334	0.27	-1.35	Teaching and other associate professionals
4142	-0.19	-0.86	Office and customer services clerks
5100	-0.52	0.46	Personal and protective services workers
5200	-0.13	-0.44	Models, salespersons and demonstrators
6100	-1.04	1.10	Skilled agricultural and fishery workers
7174	-0.19	1.22	Extraction, building, other craft and related trades workers
7273	-0.22	0.99	Metal, machinery, precision, handicraft, printing and related trades workers
8183	-0.44	1.29	Stationary-plant and related operators, drivers and mobile-plant operators
8200	-0.81	0.61	Machine operators and assemblers
9100	-1.97	0.24	Sales and services elementary occupations
9200	-0.11	1.11	Agricultural, fishery and related laborers
9300	-1.52	0.93	Laborers in mining, construction,manufacturing and transport
Mean	0	0	
Std. dev.	1	1	
Pearson correlation coefficient	-0.68		

Note: Occupation codes are based on regrouped (group B) classification of ECHP data. If the occupations are regrouped, the first and the last two digits of the occupation code corresponds to the 2-digit ISCO-88 classification of occupations.

Appendix D. Wage Regression Estimates, Decomposition Results and Robustness Checks

Table D.1: Wage Regression Estimates, U.S.

	1979	1988	1993	2008
Secondary Education	0.319 ***	0.351 ***	0.432 ***	0.315 ***
	(0.034)	(0.044)	(0.047)	(0.030)
Higher Education	0.478 ***	0.586 ***	0.739 ***	0.673 ***
	(0.061)	(0.060)	(0.072)	(0.057)
Experience	0.043 ***	0.049 ***	0.043 ***	0.043 ***
	(0.006)	(0.003)	(0.005)	(0.004)
Experience^2	-0.001 ***	-0.001 ***	-0.001 ***	-0.001 ***
	(0.000)	(0.000)	(0.000)	(0.000)
Brains	0.471 ***	0.599 ***	0.568 ***	0.715 ***
	(0.126)	(0.129)	(0.136)	(0.123)
Brawns	-0.013	-0.110	-0.166	-0.170 **
	(0.083)	(0.100)	(0.099)	(0.074)
Constant	1.941 ***	1.683 ***	1.592 ***	1.688 ***
	(0.142)	(0.141)	(0.153)	(0.141)
VIF(Brains)	1.73	1.77	1.73	1.86
VIF(Brawns)	1.82	1.86	1.78	1.87
R^2	0.17	0.23	0.25	0.27
Number of obs.	22,691	20,446	23,172	31,907

Notes: i) Occupational level clustered standard errors are in parentheses. (ii)*, ** and *** significant at 1, 5 and 10 % significance level respectively. (iii) The omitted category is taken as low level for education dummies. (iv) Variance inflation factor: VIF= $1/(1 - {R_i}^2)$ and ${R_i}^2$ is the coefficient of determination of the regression equation where each explanatory variable regressed on all the other explanatory variables.

Table D.2: Decomposition of the changes in gender wage gap, U.S.

Panel A. Descriptive statistics	1979 vs 1988	1993 vs 2008
Male residual SD*		
year t	0.496	0.545
year s	0.533	0.575
Female residual SD**		
year t	0.488	0.533
year s	0.521	0.546
Mean female residual from male wage regression		
year t	-0.499	-0.333
year s	-0.385	-0.311
Mean female residual percentile***		
year t	22.01	32.01
year s	29.03	33.46
Panel B. Decomposition of the change in the gender wage gap		
Change in gender wage gap	-0.138	-0.051
(1) Observed X's	-0.012	-0.023
Education variables	-0.001	-0.018
Experience variables	-0.001	0.003
Brains	-0.010	-0.008
Brawns	0.000	0.000
(2) Observed Prices	-0.012	-0.006
Education variables	0.002	-0.001
Experience variables	0.000	0.000
Brains	-0.001	-0.005
Brawns	-0.013	-0.001
(3) Unobserved Prices	0.026	0.013
(4) Gap effect	-0.140	-0.034
Sum gender-specific (1+4)	-0.152	-0.058
Sum wage structure (2+3)	0.014	0.007

Notes: Year t and year s refer to the years 1979 and 1988 for the second column, and 1993 and 2008 for the third column, respectively. The change in the differential is the change in the male-female log wage differentials between two corresponding years.* Estimated using male wage regression. ** Estimated using female wage regression. *** Computed by assigning each women a percentile ranking in the indicated year's residual male wage distribution and calculating the female mean of these percentiles.

Table D.3: Wage regression estimates 1993-2008, pooled

	Austria		Ireland		Italy		Portugal		Spain		U.K.		U.S.	
	1994	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008
Secondary Education	0.125 *	0.124 ***	0.195 ***	0.075	0.129 ***	0.098 ***	0.149 *	0.244 ***	0.200 ***	0.128 ***	0.129 ***	0.132 *	0.354 ***	0.267 ***
	(0.045)	(0.022)	(0.041)	(0.056)	(0.025)	(0.019)	(0.056)	(0.030)	(0.025)	(0.027)	(0.023)	(0.053)	(0.032)	(0.020)
Higher Education	0.297 ***	0.331 ***	0.480 ***	0.321 ***	0.360 ***	0.241 ***	0.457 ***	0.656 ***	0.347 ***	0.306 ***	0.300 ***	0.347 ***	0.702 ***	0.630 ***
	(0.061)	(0.038)	(0.076)	(0.051)	(0.040)	(0.029)	(0.114)	(0.091)	(0.052)	(0.039)	(0.035)	(0.055)	(0.052)	(0.037)
Experience	0.005	0.018 ***	0.030 ***	0.020 **	0.007 **	0.026 ***	0.019 **	0.039 ***	0.017 ***	0.026 ***	0.017 ***	0.014 **	0.032 ***	0.031 ***
	(0.003)	(0.004)	(0.005)	(0.006)	(0.002)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Experience^2	0.000	- 0.000	- 0.000 **	- 0.000	- 0.000	- 0.000 ***	- 0.000 *	-0.001 ***	- 0.000	- 0.000 ***	- 0.000 ***	- 0.000 **	-0.001 ***	-0.001 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Brains	0.787 ***	0.743 ***	0.832 ***	0.802 ***	0.437 ***	0.530 ***	1.059 ***	0.630 ***	0.797 ***	0.751 ***	0.823 ***	0.859 ***	0.763 ***	0.830 ***
	(0.085)	(0.052)	(0.148)	(0.130)	(0.106)	(0.085)	(0.225)	(0.135)	(0.156)	(0.103)	(0.135)	(0.134)	(0.139)	(0.149)
Brawns	-0.094	-0.114 *	-0.127	-0.009	-0.321 ***	-0.187 **	-0.886 ***	-0.318 *	-0.338 ***	-0.197 *	-0.105	-0.122	0.075	0.040
	(0.070)	(0.042)	(0.125)	(0.107)	(0.075)	(0.061)	(0.147)	(0.128)	(0.070)	(0.081)	(0.120)	(0.096)	(0.127)	(0.105)
Constant	1.822 ***	1.675 ***	1.335 ***	1.703 ***	1.926 ***	1.653 ***	1.146 ***	0.938 ***	1.558 ***	1.511 ***	1.601 ***	1.955 ***	1.435 ***	1.576 ***
	(0.087)	(0.073)	(0.117)	(0.120)	(0.069)	(0.082)	(0.117)	(0.124)	(0.100)	(0.112)	(0.101)	(0.100)	(0.126)	(0.113)
R^2	0.15	0.30	0.32	0.32	0.33	0.27	0.50	0.45	0.38	0.41	0.27	0.23	0.22	0.23
Number of obs.	2,392	3,704	2,316	2,283	4,148	11,365	2,767	2,728	3,827	7,994	6,489	4,302	45,234	62,925

Notes: i) Occupational level clustered standard errors are in parentheses. (ii) *, ** and *** significant at 1, 5 and 10 % significance level respectively. (iii) The omitted category is taken as low level for education dummies.

Table D.4: Wage regression estimates 1993-2008, females

	Austria		Ireland		Italy		Portugal		Spain		U.K.		U.S.	
	1994	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008
Secondary Education	0.008	0.114***	0.164*	0.070	0.127***	0.109***	0.095	0.206***	0.118*	0.113*	0.087**	0.108	0.282***	0.215***
	(0.081)	(0.026)	(0.060)	(0.070)	(0.024)	(0.020)	(0.098)	(0.024)	(0.052)	(0.044)	(0.030)	(0.076)	(0.044)	(0.053)
Higher Education	0.262**	0.353***	0.568***	0.268***	0.337***	0.263***	0.375*	0.712***	0.348***	0.342***	0.276***	0.361***	0.612***	0.576***
	(0.090)	(0.056)	(0.119)	(0.044)	(0.038)	(0.024)	(0.158)	(0.054)	(0.067)	(0.052)	(0.047)	(0.089)	(0.053)	(0.046)
Experience	0.014	0.013**	0.035***	0.023*	0.007*	0.024***	0.015**	0.039***	0.023***	0.019***	0.006	0.005	0.021***	0.017***
	(0.007)	(0.004)	(0.008)	(0.009)	(0.003)	(0.005)	(0.005)	(0.007)	(0.003)	(0.005)	(0.004)	(0.005)	(0.003)	(0.003)
Experience^2	-0.000	-0.000	-0.001***	-0.000	-0.000	-0.000**	-0.000	-0.001***	-0.000***	-0.000	-0.000*	-0.000	-0.000***	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Brains	0.479**	0.564***	0.698***	0.910***	0.457*	0.485***	0.954**	0.423**	0.669**	0.741***	0.592***	0.756***	0.713***	0.739***
	(0.122)	(0.089)	(0.129)	(0.189)	(0.158)	(0.102)	(0.317)	(0.118)	(0.176)	(0.150)	(0.115)	(0.136)	(0.156)	(0.151)
Brawns	-0.660***	-0.321**	-0.456*	-0.046	-0.518***	-0.355***	-1.310***	-0.575***	-0.541***	-0.278**	-0.443***	-0.308**	-0.168	-0.160
	(0.125)	(0.090)	(0.191)	(0.166)	(0.101)	(0.065)	(0.177)	(0.118)	(0.115)	(0.092)	(0.078)	(0.098)	(0.178)	(0.166)
Constant	2.030***	1.788***	1.410***	1.635***	1.926***	1.682***	1.336***	1.036***	1.593***	1.518***	1.854***	2.032***	1.600***	1.732***
	(0.104)	(0.083)	(0.126)	(0.183)	(0.108)	(0.100)	(0.171)	(0.120)	(0.137)	(0.145)	(0.079)	(0.111)	(0.112)	(0.122)
R^2	0.195	0.351	0.423	0.294	0.420	0.332	0.626	0.622	0.473	0.479	0.333	0.247	0.236	0.249
Number of obs.	952	1,744	951	1,170	1,597	5,172	1,229	1,351	1,276	3,854	3,132	2,232	22,062	31,018

Notes: i) Occupational level clustered standard errors are in parentheses. (ii) *, ** and *** significant at 1, 5 and 10 % significance level respectively. (iii) The omitted category is taken as low level for education dummies.

Table D.5: Robustness Check: Wage Regression Estimates, Returns to Brains/Brawns, 1993-2008

	Austria		Ireland		Italy		Portugal		Spain		U.K.		U.S.	
	1994	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008
Secondary Education	0.175**	0.138**	0.300***	0.126*	0.160***	0.125***	0.273***	0.344***	0.275***	0.176***	0.155***	0.169*	0.479***	0.363***
	(0.048)	(0.043)	(0.048)	(0.047)	(0.018)	(0.030)	(0.051)	(0.055)	(0.045)	(0.035)	(0.021)	(0.063)	(0.057)	(0.037)
Higher Education	0.367**	0.394***	0.547***	0.475***	0.421***	0.328***	0.785***	0.775***	0.500***	0.392***	0.382***	0.412***	0.843***	0.781***
	(0.099)	(0.065)	(0.099)	(0.060)	(0.052)	(0.061)	(0.114)	(0.172)	(0.100)	(0.075)	(0.046)	(0.065)	(0.091)	(0.082)
Experience	0.001	0.022***	0.034***	0.022**	0.006*	0.028***	0.011	0.041***	0.014*	0.028***	0.022***	0.016**	0.043***	0.043***
	(0.006)	(0.005)	(0.006)	(0.007)	(0.003)	(0.004)	(0.010)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Experience^2	0.000	-0.000*	-0.000**	-0.000	-0.000	-0.000***	-0.000	-0.001***	-0.000	-0.000**	-0.000**	-0.000*	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Brains/Brawns	0.035**	0.035**	0.044**	0.043***	0.045***	0.029**	0.080***	0.050*	0.052***	0.042**	0.038*	0.048*	0.037**	0.046***
	(0.011)	(0.010)	(0.015)	(0.011)	(0.011)	(0.008)	(0.013)	(0.017)	(0.012)	(0.012)	(0.016)	(0.018)	(0.009)	(0.011)
Constant	2.166***	1.926***	1.555***	1.928***	1.912***	1.750***	1.136***	1.033***	1.650***	1.684***	1.891***	2.267***	1.630***	1.764***
	(0.069)	(0.059)	(0.081)	(0.079)	(0.033)	(0.049)	(0.098)	(0.073)	(0.059)	(0.055)	(0.054)	(0.057)	(0.086)	(0.070)
R^2	0.109	0.203	0.298	0.345	0.309	0.220	0.361	0.342	0.308	0.324	0.197	0.159	0.230	0.243
Number of obs.	1,440	1,960	1,365	1,113	2,551	6,193	1,538	1,377	2,551	4,140	3,357	2,070	23,172	31,907

Notes: i) Occupational level clustered standard errors are in parentheses. (ii) *, ** and *** significant at 1, 5 and 10 % significance level respectively. (iii) The omitted category is taken as low level for education dummies.

Table D.6: Robustness Check: Returns to Skills Using Measures Constructed by PCA, 1993-2008

	Austria		Ireland		Italy		Portugal		Spain		U.K.		U.S.	
	1994	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008	1993	2008
Secondary Education	0.114**	0.065	0.234***	0.104	0.120***	0.095***	0.141**	0.260***	0.222***	0.143***	0.112***	0.124*	0.432***	0.316***
	(0.039)	(0.032)	(0.038)	(0.051)	(0.024)	(0.021)	(0.037)	(0.042)	(0.030)	(0.026)	(0.016)	(0.058)	(0.046)	(0.029)
Higher Education	0.230**	0.242***	0.375***	0.401***	0.358***	0.243***	0.533***	0.608***	0.345***	0.269***	0.264***	0.293***	0.733***	0.669***
	(0.072)	(0.046)	(0.056)	(0.064)	(0.042)	(0.032)	(0.122)	(0.147)	(0.054)	(0.038)	(0.033)	(0.053)	(0.070)	(0.056)
Experience	0.002	0.022**	0.033***	0.022**	0.007*	0.027***	0.015	0.043***	0.014*	0.029***	0.022***	0.014**	0.043***	0.043***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.003)	(0.004)	(0.009)	(0.007)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)
Experience^2	0.000	-0.000*	-0.000**	-0.000	-0.000	-0.000***	-0.000	-0.001***	-0.000	-0.000**	-0.000***	-0.000*	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Brains	0.146***	0.150***	0.177***	0.142***	0.086**	0.105***	0.206***	0.128**	0.174**	0.140***	0.165***	0.164***	0.143***	0.178***
	(0.033)	(0.014)	(0.033)	(0.031)	(0.024)	(0.025)	(0.051)	(0.041)	(0.046)	(0.025)	(0.030)	(0.018)	(0.033)	(0.032)
Brawns	-0.037**	-0.058***	-0.067	-0.022	-0.081**	-0.045	-0.192***	-0.124**	-0.099**	-0.079*	-0.047	-0.067**	-0.032	-0.030
	(0.011)	(0.010)	(0.035)	(0.029)	(0.025)	(0.026)	(0.040)	(0.032)	(0.029)	(0.027)	(0.038)	(0.022)	(0.031)	(0.024)
Constant	2.305***	2.102***	1.750***	2.074***	2.041***	1.864***	1.372***	1.190***	1.861***	1.846***	2.005***	2.441***	1.798***	1.962***
	(0.067)	(0.076)	(0.090)	(0.100)	(0.036)	(0.050)	(0.080)	(0.108)	(0.050)	(0.068)	(0.045)	(0.044)	(0.096)	(0.083)
R^2	0.147	0.266	0.344	0.373	0.313	0.247	0.429	0.365	0.356	0.367	0.253	0.235	0.254	0.269
Number of obs.	1,440	1,960	1,365	1,113	2,551	6,193	1,538	1,377	2,551	4,140	3,357	2,070	23,172	31,907

Notes: i) Occupational level clustered standard errors are in parentheses. (ii) *, ** and *** significant at 1, 5 and 10 % significance level respectively. (iii) The omitted category is taken as low level for education dummies.

Appendix E. Wage Regression Estimates, Decomposition Results and Robustness Checks

Table E.1: Addressing selection bias: Selectivity-corrected gender wage gaps

Decomposition results	Austria	Ireland	Italy	Portugal	Spain	U.K.	U.S.
Change in gender wage gap	-0.078	-0.143	-0.018	-0.006	0.034	-0.066	-0.051
(1) Observed X's	0.012	-0.062	-0.039	-0.078	0.015	-0.022	-0.003
(2) Observed prices	-0.027	0.091	0.044	0.046	0.029	-0.021	-0.011
(3) Unobserved prices	-0.005	-0.013	0.002	0.008	-0.022	0.007	0.012
(4) Gap effect	-0.030	-0.578	0.046	-0.079	-0.011	-0.334	0.183
(5) Selection	-0.028	0.420	-0.071	0.096	0.023	0.303	-0.232

Selectivity-corrected gap effect is based on estimating the selection corrected model using a two-stage Heckman (1979) selection model. See text for details.