Task Profiles and Gender Wage-Gaps Within Occupations

Aspasia Bizopoulou

Preliminary Draft

Abstract

Recent literature for the US, Germany and Australia shows that a significant proportion of the unexplained gender wage-gap can be found within very narrowly defined occupations. I find that this pattern also holds true in European countries. This finding raises the question of why men and women working in very similar jobs are paid substantially different wages. Using a newly available dataset with detailed job-task and occupational information, I investigate whether task segregation by gender within a narrowly defined occupation can account for within-occupational gender wage-gaps. I find that higher levels of task segregation by gender increase the wage-gap within an occupation, in favour of men. I also find that, within occupations, the effect of task segregation on wages is driven by certain tasks that carry a significant wage premium and which are consistently performed by men much more than by women.

Keywords: task approach, gender wage-gap, occupations.

JEL Codes: J16, J24, J31

1 Introduction

The question of why women earn less than similarly qualified men has puzzled economists for many years and has been studied from several angles. One well-studied approach has been to look at the problem from the lens of occupations, in particular the role of occupational segregation by gender (Polachek, 1981; Blau and Kahn, 2000). For many years occupational segregation was considered as a prime determinant of lower female wages, even though there remained uncertainties about the mechanism through which segregation affected female wages: were female occupations truly lower skilled than male occupations or was the wage gap entirely discriminatory? (Baker and Fortin, 2001). The uncertain nature of the mechanism was perhaps best highlighted by the highly controversial and publicised strike of the Ford machinists working in Dagenham in 1968. The machinists, who were all women and whose job was to saw the car's interior seat fittings, went on strike to protest against a substantially lower paycheck compared to their male colleagues in the rest of the Ford production line. The court's decision was that Ford had openly discriminated against its female employees, and the event led to the 1970 Equal Pay Act in the UK. The true significance of occupational segregation for wage gaps was slowly unraveled with time, as more and more women entered the labour force and the ranks of traditionally male occupations. Data from the sixties to today shows that occupational segregation fell in many of the most developed countries (see for example (Blau et al., 2013; Fedorets, 2013)), yet recent research has shown that the declining gender wage-gap that has been observed during the same period can be primarily attributed to the effects of technology rather than declining occupational segregation (Yamaguchi, 2014; Beaudry and Lewis, 2014; Black and Spitz-Oener, 2010).

Nevertheless, despite decreasing occupational segregation and the beneficial effect of technology on diminishing male-female pay differences, the wage-gap is still present in simple Mincer-type earnings models. The recurrent and surprising fact is that of the remaining observed and otherwise unexplained gender wage-gap, the majority is to be found *within* very narrowly defined occupations, rather than *across* occupations (Goldin, 2014; Fedorets, 2013; Cobb-Clark and Tan, 2011). This finding remains robust to controls such as educational attainment, age, race and number of hours worked. It has been observed in the US, Germany and Australia (Goldin, 2014; Fedorets, 2013; Cobb-Clark and Tan, 2011). Two recent studies have put forward explanations: (Goldin, 2014) explores this pattern for the US and finds that a substantial proportion of this inequality can be attributed to a strong premium for working very long and inflexible hours in certain occupations such as lawyers. (Fedorets, 2013) also studies this gap by looking at the evolution of the price of non-routine cognitive tasks over time. I also choose to focus on the question of why women earn less men when working in the same very narrowly defined occupation.

My particular take on the problem is to understand whether within-occupational inequality can be attributed to the way tasks are distributed to men and women within narrowly defined occupations. My motivation to explore the effect of task differences on wages, stems from the recent empirical finding that job tasks in the US vary substantially within broadly defined occupations, and the variation is systematically related to race and gender (Autor and Handel, 2013). I further elaborate this observation by looking at the variation of tasks within the most narrowly defined occupations (4-digit codes rather than Autor & Handel's 1-digit). This way of looking at the problem is interesting in so far as it can allow us to study whether gender-wage gaps within occupations are potentially due to imperfect substitutability between the two genders, in terms of their task profiles.

To further motivate this line of inquiry, I start with an illustrative graph (Figure 1) showing the composition of male and female task profiles within an example occupation, in this case Executive Secretaries. I define a task profile as the distribution of tasks performed by men and women in a certain occupation. The graph below provides an illustration of what the female and male task frequencies look like for Executive Secretaries and, importantly, how they differ. Tasks have been ordered according to the proportion of the women that perform each of task within the occupation, hence the spiral shape of the red line. The blue line provides the proportion of men that do each task within this occupation. The juxtaposition of the two lines clearly shows that men and women have slightly different task profiles when they are Executive Secretaries: men perform more problem solving and people tasks but fewer typing and reading tasks than women.

Several questions arise from the observation that men and women do not perform exactly the same tasks within an occupation. The first one, which is also the main question I investigate in this thesis, is whether more differentiated male-female task profiles within occupations affect gender wage-gaps. The basic economic intuition is that the more the task-profiles differ within an occupation, the less substitutable the two genders and the more likely is the existence of a gender wage-gap. A further question of interest is whether certain tasks or combinations of tasks systematically 'drive' the gap. Recent research shows that occupations involving mostly cognitive and complex tasks have accrued higher premia

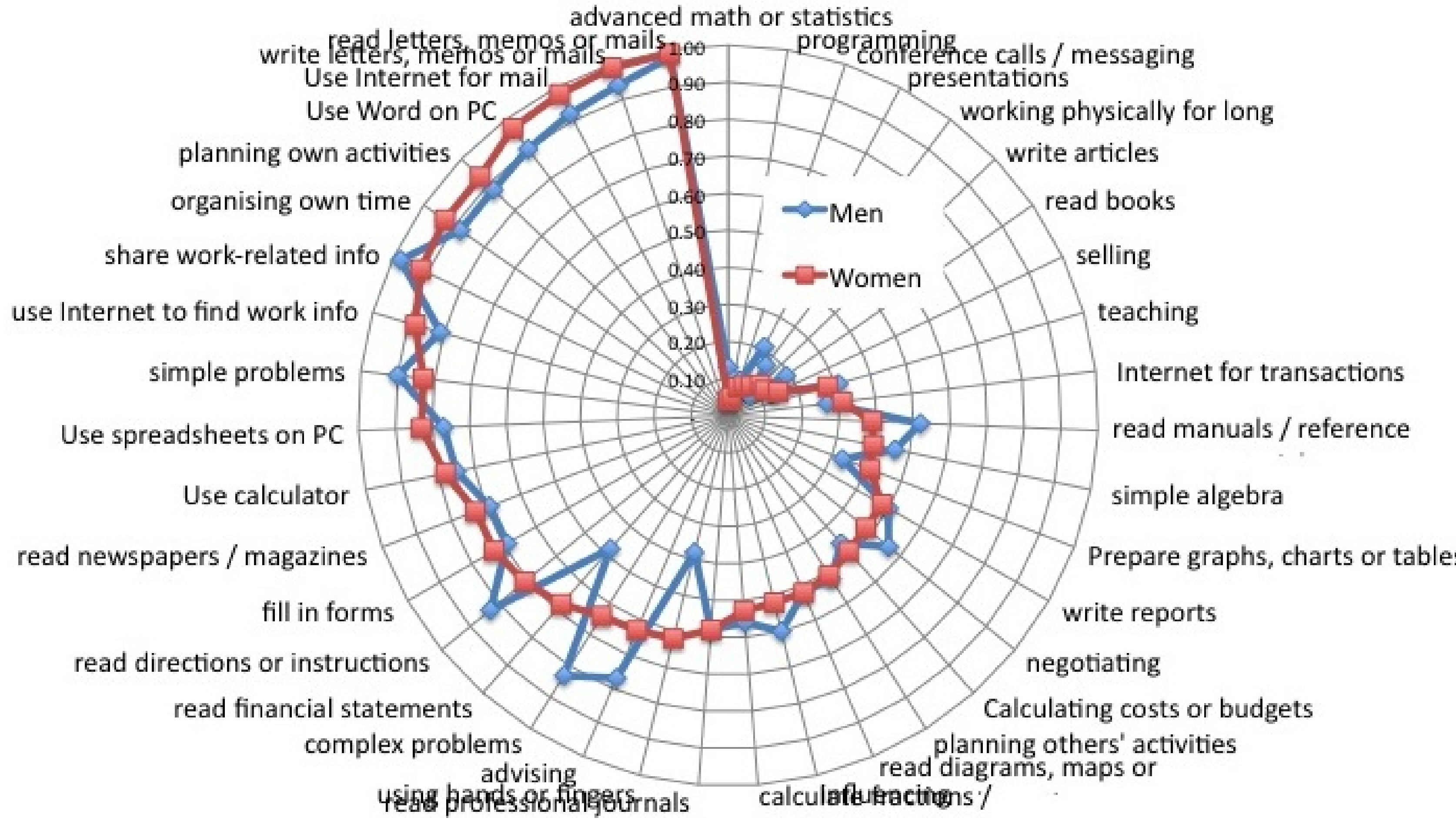

Figure 1: Distribution of female and male tasks among Administrative and Executive Secretaries (48 men, 198 women)

with the introduction of technology in the workplace (Autor et al., 2003; Acemoglu and Autor, 2011), which suggests that the productivity of all tasks within an occupation is not the same.

This paper is also relevant to recent business literature highlighting that women spend more time doing office 'house work' than men, and may be penalised for it through missed promotions or lower earnings. Although in my data I cannot distinguish what constitutes office 'house work', I can observe that men spend significantly more time than women doing tasks that are more valued and carry higher premiums.

Using a newly released and still unexplored dataset with very detailed information about the task profiles of workers in 9 European countries, I study the role of tasks in gender wage-gaps within occupations. I first look at whether adding occupations to a traditional decomposition exercise helps explain a larger proportion of the overall gender wage gap, and I find that adding tasks decreases the unexplained part of the gap by 8 percentage points. I then study whether task segregation by gender within very narrowly defined occupations can explain the observed gender wage gaps within said occupations. I use a new application of the Dissimilarity Index as a measure of task segregation within occupations and I find that higher values of the index - which correspond to more task segregation by gender within an occupation - leads to higher gender wage-gaps within occupations in favour of men. I find that controlling for the level of task segregation within an occupation can explain up to 100% of the gender wage gap within the most populous occupations. Furtermore, in a wage accounting exercise task segregation by gender is shown to be positively correlated with men's wages and negatively correlated with women's wages. The result holds for occupations at the most narrowly defined, 4-digit, level. Finally, I disentangle the effect of task segregation on wages: do women do more low paying tasks than the men or do men spend more time on high-paying tasks compared to women? I find over-whelming evidence for the latter: men are consistently observed to be spending more time on high-paying tasks than women and not the other way round.

The rest of the paper is organised as follows: in section 2 I describe the data and provide summary statistics, in section 3 I outline the empirical approach of the analysis, in section 4 I present results and in section 5 I conclude.

2 Data and Summary Statistics

The data is taken from the Programme for the International Assessment of Adult Competencies (PIAAC) for the year 2011, which became available to researchers in October 2013. The survey was conducted in 21 OECD countries, however I chose to focus on 9 European countries (Belgium, Czech Republic, Denmark, Spain, France, Italy, Netherlands, Poland, Slovakia) for data availability reasons[1]. The survey was fielded by the Organisation for Economic Cooperation and Development (OECD) and the next wave is to be expected in 2017. Three characteristics of this data are useful for the purposes of this analysis: 1) for each individual I have detailed information on the type of tasks and activities they perform at work, including but not limited to: types of reading, types of writing, types of numeracy, computer use, types of influences exerted on others, learning activities, organisation and physical activities; 2) the information on tasks is available for 9 European countries; 3) all occupations are characterised by a finite set of 39 broad tasks, thus allowing the identification of high and low premium tasks across all occupations.

2.1 Summary Statistics

Table 1 provides an overview of the 39 tasks that characterise occupations. Each task can be performed at five different levels of intensity: i) never, ii) less than once a month, iii) less than once a week but at least once a month, iv) at least once a week but not every day and v) every day. In Table 2, I provide some basic summary statistics for each of the countries separately. The sample consists of 35,139 employed individuals aged 16-65 and each country takes up 10-15% of the sample. There are slightly fewer women than men in the sample, with the exception of the Czech Republic, which corresponds to the fact that female labour participation is lower for women than for men. Average working hours for full-time individuals fluctuate around 41 hours per week, with the exception of Spain, which is the only country with an average below 40 hours. In terms of demographic distribution, it is notable that 45 percent of the Polish sample consists of 16-25 year olds, while for all other countries the 16-25s are around the 10 percent mark. The most prevalent level of educational attainment is Upper Secondary education, the notable exception being Spain, where the most common level of education is Lower Secondary. Italy also stands out for the proportion of people that have a Master's degree, with only 2,7%, compared to

[1]Unfortunately, several of the participating countries did not provide information on earnings and/or occupational classification.

at least 9% in all other countries.

3 Empirical Approach

In this section I outline the way I test for the main hypothesis of this paper: that task segregation by gender within occupations will lead to gender wage-gaps due to the lower substitutability of the two sexes within the occupation. The basic economic intuition is that in those occupations with higher male-female substitutability wage gaps will be expected to be lower. While it may first sound as a relatively simple hypothesis to test, the richness of the data makes the task quite difficult. As mentioned earlier, for each individual in each occupation I have information on their complete task set: I know which tasks they do, which they do not do and how often they do them. In section 3.1 I explain how I combine the rich information on task use to test for task segregation. In section 3.2 I explain the econometric framework I use to test the task segregation hypothesis.

3.1 Operationalisation of the D-index for measuring task segregation within occupations

An advantage of this sample is that all 483 4-digit occupations are characterised by a maximum set of 39 tasks, something that makes within-occupational task differentiation possible [2]. While each task is performed at a certain intensity (never, less than once a month, more than once a month but less than once a week, more than once a week, everyday), I choose to collapse the tasks into 0/1 dummies, where 1 corresponds to doing the task at least once a month and up to everyday.

What would be the best way to capture the heterogeneous task-profiles of different people within an occupation? In theory, many measures could be used to do this, but in my case I am looking for one that will capture the role of gender. I choose to use a measure that has been commonly used in geographical studies of segregation, namely the Index of Dissimilarity (henceforth D-index). Originally formulated by (Jahn et al., 1947) in the American Sociological Review for the purposes of measuring the level of racial segregation in different areas, it has since been widely used. Fortin & Huberman (2002) use it in a context that is closest to this paper, namely for measuring the extent of gender segregation within an occupation. In the context of this paper the measure is modified to fit the present

[2] Table 1 provides an overview of all the tasks available in the sample.

purpose: rather than measuring segregation of people within an occupation, as has been done by (Watts, 1998; Zveglich and van der Meulen Rodgers, 1999), I use it to measure segregation of tasks by gender within an occupation. An intuitive explanation of how the measure works can be illustrated by going back to Graph 1: noting that for each task the triangle (red) point shows the proportion of women doing this task in that occupation and the blue point the proportion of men, we take the absolute difference between each red and blue point and then aggregate it, so as to get an idea of the overall level of heterogeneity in task profiles. This is exactly how the D-index operates and it can be formalised as follows:

$$D_i = \frac{1}{2}\sum_{t=1}^{39}\left|\left(\frac{f_{it}}{F_i} - \frac{m_{it}}{M_i}\right)\right|$$

The index i refers to an occupation and the index t to tasks. F_i is the number of women in occupation i and $f_{it}$is the number of women in occupation i doing task t. Thus $\frac{f_{it}}{F_i}$ is the proportion of women doing task t in occupation i and $\frac{m_{it}}{M_i}$ is the proportion of men doing task t in occupation i. The measure is bounded between 0 and 39, where 0 corresponds to the situation where task bundles are distributed identically between men and women or, in terms of the graph, the blue and red points are exactly on the same spots for each task. 39 corresponds to the situation of perfect segregation where none of the men do any of the tasks that the women do and vice versa.

3.2 Econometric Approach

To begin the analysis, I closely follow the approach of Fedorets (2014) and Goldin (2014) and investigate whether gender-gaps persist when we account for occupations at the 4-digit level. In terms of econometric framework, I estimate a wage equation that is very similar to Goldin (2014) and Fedorets (2014):

$$\ln w_i = X_i\beta + \alpha_k OCCUP_k + u_i \tag{1}$$

X_i are the characteristics that vary by individual, for which summary statistics by country are provided in Table 1, $OCCUP_k$ are occupation dummies (there are 483 occupations) and u_i is an individual error term. The aim of this part of the analysis is to highlight that for the current sample of European countries, it is true that the large proportion of the wage gap is to be found within occupations, much like in US, German and Australian data.

The next step in the analysis is to understand whether task segregation within occupations can account for the within occupational wage gap. In equation (2), the coefficient a_k captures the effect of the various characteristics of occupation k on the wage, conditional on individual characteristics X_i. We note that we are particularly interested in one particular characteristic of occupations, i.e. the male-female task-profile heterogeneity. Thus, a_k can be decomposed as follows:

$$\alpha_k = \delta + \gamma D_k + \omega_k, \tag{2}$$

where D_k is the index of dissimilarity, which measures the level of task segregation. A straightforward way to estimate (1) would be to plug (2) into (1):

$$\ln w_i = \delta + X_i\beta + D_k + (\omega_k + u_i) \tag{3}$$

Baker and Fortin (2001) note that in this type of econometric model we are estimating the effect of a variable aggregated at the occupational-level on a variable that is at the individual-level, thus giving rise to omitted variable bias at the occupational level. For example, by omitting other characteristics of the occupation, such as the working conditions and average human capital that may be correlated with individual wages, we may overestimate the effect of the particular occupational-level variable that we want to include in the estimation. A similar problem arises in the estimations of gravity equations, where correlating variables at different levels of aggregation without controlling for omitted variables at different levels of aggregation can lead to biased estimates (**?**). The solution that Baker and Fortin (2001) propose is to introduce control variables at all levels of aggregation of the estimation. In our case this means including average human capital characteristics at the occupational level as well as general characteristics with respect to the occupation in question. In this case we control for the potential unobserved correlation between ω_k and X_i. Thus the equation to be estimated will be as follows:

$$\ln w_i = \delta + X_i\beta + \gamma D_k + G_k\lambda + C_k k + \psi_k + u_i \tag{4}$$

G_k is a measure of average human capital characteristics within an occupation and C_k is a measure of occupation-specific characteristics unrelated to human capital such as working conditions. The error term ψ_k is at the occupational level while u_i is at the individual level. The advantage of this approach means that the presence of the occupation-specific human capital variable G causes $\hat{\beta}$ to be estimated as if there were k-specific (i.e. occupation-specific) fixed effects.

4 Results

4.1 Accounting for the female wage gap: the effect of occupations

In Table 3, I highlight the importance of within-occupational differences in earnings for the gender wage-gap in these 9 European countries, as has previously been done for the US, Germany and Australia (Goldin, 2014; Fedorets, 2013; Cobb-Clark and Tan, 2011). I closely follow the example of (Goldin, 2014) and observe the change in the coefficient on female of a log earnings regression to which I gradually add control variables. I provide those results for all education groups, for university graduates (BA) and for High-School leavers (HS). For each regression, there is one for all workers and one only for those working full-time. Following (Goldin, 2014), for each regression I include an age quartic, a dummy for native speaker status and 8 country dummies, the reference being Spain. I subsequently add hours worked and education dummies. In the most complete specification I include both occupation dummies (4-digit) and the 39 tasks.

In all specifications, adding occupational dummies decreases the coefficient on female by no more than a third which is similar to the result that (Goldin, 2014) finds for the US. Furthermore, the coefficient on female remains above 10% in all specifications. The main takeaway from the addition of occupational dummies to the regression is that the gender wage-gap displays the same patterns in these nine European countries as it does in previously studied countries like the US, Germany and Australia.

I then take Goldin's approach one step further and add 39 task variables. A decrease in the female coefficient can be observed, yet the unexplained gap within occupations is not reduced by more than a quarter. While at first glance the addition of the 39 tasks seems to play a minor role in the within-occupational wage-gap, we have to keep in mind that the task variables are not mutually exclusive in the way occupational dummies are. If individuals performed only one task at a time, then adding task variables would eliminate the within-occupational wage inequality due to performing different tasks. However, given that we cannot control for each individual's task profile within an occupation, we cannot tell how differences in task sets of men and women affect wage-gaps within occupations in with this specification.

One outcome that is worth noticing is that the coefficient on female is much larger among the low- rather than the high-skilled (i.e. last two sets of estimates, compared to the middle two sets of estimates in Table 3). In light of the recent studies showing that the closing of the gender wage-gap over time has been largely driven by the lower-skilled jobs - while the gap in high-skilled jobs has actually been increasing - it may come as surprise to see that at the absolute levels the gap is lower among high-skilled individuals and higher among lower-skilled individuals (Black and Spitz-Oener, 2010; Acemoglu and Autor, 2011; Fedorets, 2013; Goldin, 2014). (Goldin, 2014) looks at at the absolute level gaps and finds that it is lower for low-skilled than for high-skilled American women. Unfortunately we do not have the corresponding PIAAC data either for the US, which is where (Goldin, 2014) and (Acemoglu and Autor, 2011) base their analysis, or for Germany, which is the source for (Black and Spitz-Oener, 2010) and (Fedorets, 2013), because income values are missing for both countries.

The main takeaway from these summary statistics is that a substantial proportion of the wage-gap remains within narrowly defined occupations. Thus, if we are looking to explain the remaining of the wage-gap we need to understand what is happening within those occupations.

4.2 Occupational Task Segregation and the Gender Wage Gap

I then proceed to control for the level of within occupational segregation on the gender gap. Following (Goldin, 2014), I only include occupations with at least 20 people, and at least 10 men and 10 women, in order to avoid bias from very small occupations. The downside of excluding small occupations is that many of the 483 occupations are completely excluded from the analysis, namely 295 of them. Nevertheless, the remaining sample that is studied is still large, with 20,599 observations (out of the 35,139).

Columns (1)-(4) of Table 5 show the results of adding the D-index in a Mincer-type equation. The coefficient on the D-index is positive and significant in all specifications, while the coefficient on the multiplicative variable D-index*Female is negative and significant. The coefficient on Female also remains negative and significant throughout the 4 specifications and it is notably much lower than all estimations of equation (1) in Table 3. The interpretation of these three variables together is as follows: task segregation is positively and significantly correlated with men's wages and it is negatively and significantly correlated with women's wages. In other words, for a woman to work in a occupation that has higher task segregation by gender is equivalent to losing out on up to 2,5% from her earnings. In columun (3), I include a set of potentially omitted variables at the occupational level, for the reasons explained in section 3. The omitted variables I include are Average Human Capital (HC), i.e. average educational attainment and experience within the occupation; Occupation Characteristics, i.e. average level of freedom to perform job, average level of micro-management by superiors, average level of freedom in work speed and average level of working hours within the occupation; and the Rate of Femaleness of an occupation.

In columns (5)-(7) I limit the estimation sample to very large occupations, so as to check that the results are not driven by extremes in smaller occupations. I limit the sample to individuals working in occupations with at least 100 people. In those occupations, I find that the effect of task segregation is in fact stronger: the negative effect on women's wages is up to 9,2% and the positive effect for men's wages is up to 8,2%. Furthermore, I observe that for this sample the negative coefficient on female entirely disappears once I include the multiplicative dummy Female*D-index. In other words, for large and popular occupations, the gender pay gap within those occupations can be explained by the different activities of men and women. In the corrected estimates - columns (4)-(6) - the coefficient on the D-index is not identical in all three estimations, but it is much more similar between the 1-step OLS and the 2-step FGLS, i.e. 0.007 and 0.003 respectively.

4.3 High-Paying Tasks and Gender

In the previous section, I established that task segregation within occupations does exist and is negatively correlated with women's pay relative to men's. In this section, I want to isolate the mechanism, if any, through which task segregation pushes female wages down within occupations. There are three possibilities:

i) Do women do more of the low-paying tasks compared to the men?
ii) Do men do more of the high-paying tasks compared to the women?
iii) Is the gap driven by a combination of the above?

I disentangle these three questions as follows: first, I perform a simple regression of gender on all tasks, controlling for occupations and country, to identify which tasks are dominated by men or women within occupations. In other words, which tasks are observed to be performed significantly more by men or women? Then, I regress income on all tasks, controlling again for occupations and country, and take note of those tasks with positive or negative and significant coefficients. The question here is which tasks have a significant wage premium/penalty? The two regressions are placed side by side on Table 6. The first column shows the regression of the gender dummy on all tasks and occupations. Positive coefficients mean that women do the task more than men within occupations, while negative coefficients mean the opposite. The second column shows the task coefficient of a regression of wages on tasks. A positive coefficient means that doing more of said task has a positive premium on the wage, controlling for occupations and all other tasks.

My next step is to cross-correlate the dominant male/female tasks with those tasks that are have the highest/lowest and most significant premia on wage. In this way, I can observe which tasks are done by which gender and which tasks have the highest premia. The aim of this cross-correlation is to isolate the direction of the effect of task-segregation on women's within occupations: are women doing more of the tasks that have a wage penalty, are men doing more of the tasks that have a wage premium or both?

For most tasks, there is no cross-correlation effect: either they are done more by women or men without it affecting wages, or they have a significant premium on wages but they are performed equally by men and women. [As has been observed in other studies, we observe that manual labour has a significant wage penalty and is also done by men much more than by women.] In total there are nine tasks that appear to negatively affect women's wages. For 8 out 9 of them the direction of the effect is the same: women are observed to perform much less of said task than men and the task has a significant premium. For 1 of the nine tasks, the effect is the opposite: the task is done much more by women than men and the task is observed to carry a wage penalty.

Given the above observation, the main takeaway of this cross-correlation is that at the level of task segregation, women's lower wages within occupations are partially driven by men doing more high-paying tasks than women, not by women doing more low-paying tasks than men.

5 Conclusion

A large proportion of the unexplained gender wage-gap remains within narrowly defined occupations in several countries. I study the role of task segregation by gender within occupations in explaining this gap. Controlling for a number of observable demographic, educational and work-related characteristics, I find a persistent correlation between higher task segregation within occupations by gender and a higher gender wage-gap in favour of men. I investigate whether the effect of task segregation is driven by women performing more low-premium tasks or by men performing high-premium tasks. I find evidence for the latter: tasks that have a higher relative premium are consistently performed by men much more than by women.

As a continuation of this project it will be of interest to understand how men and women choose what tasks they do - why don't women spend more of their time on high-paying tasks? It would be interesting to disentangle whether women's and men's observed task choices are a result of self-selection or of assignment from higher management.

References

Acemoglu, D. and Autor, D. H. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics*, 4:1043–1171.

Autor, D. H. and Handel, M. J. (2013). Putting Tasks to the Test: Human Capital, Job Tasks, and Wages. *Journal of Labor Economics*, 31(2):S59–S96.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Baker, M. and Fortin, N. M. (2001). Occupational gender composition and wages in Canada , 1987-1988. *the Canadian Journal of Economics/ Revue Canadienne d'Economique*, 34(2):345–376.

Beaudry, P. and Lewis, E. (2014). Do Male-Female Wage Differentials Reflect Differences in the Return to Skill? Cross-City Evidence from 1980 to 2000. *American Economic Journal: Applied Economics*, 6(2):178–194.

Black, S. E. and Spitz-Oener, A. (2010). Explaining women's success: technological change and the skill content of women's work. *The Review of Economics and Statistics*, 92(1):187–194.

Blau, F. D., Brummund, P., and Liu, A. Y.-H. (2013). Trends in occupational segregation by gender 1970 to 2009: Adjusting for the impact of changes in the occupational coding system. *NBER Working Paper 17993.*

Blau, F. D. and Kahn, L. M. (2000). Gender Differences in Pay. *The Journal of Economic Perspectives*, 14(4):75–99.

Cobb-Clark, D. A. and Tan, M. (2011). Noncognitive skills, occupational attainment, and relative wages. *Labour Economics*, 18(1):1–13.

Fedorets, A. (2013). Closing the Gender Pay Gap and Individual Task Profiles: Women's Advantages from Technological Progress. *Working Paper*, pages 1–33.

Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119.

Jahn, J., Schmidt, C. F., and Schrag, C. (1947). The measurement of Ecological Segregation. *American Sociological Review*, 12(3):293–303.

Polachek, S. (1981). Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure. *The Review of Economics and Statistics*, 63(1):60–69.

Watts, M. (1998). Occupational Gender Segregation : Index Measurement and Econometric Modeling. *Demography*, 35(4):489–496.

Yamaguchi, S. (2014). Changes in Returns to Task-Specific Skills and Gender Wage Gap.

Zveglich, J. E. J. and van der Meulen Rodgers, Y. (1999). Occupational Segregation and the Gender Earnings Gap. Technical report, Asian Development Bank.

6 Appendix A

TABLE 1	
People tasks	**Numeracy tasks**
sharing work-related info	calculating costs or budgets
teaching people	use or calculate fractions or percentages
presentations	use a calculator
selling	prepare graphs, charts or tables
advising people	use simple algebra or formulas
influencing people	use advanced math or statistics
negotiating with people	
	ICT tasks
Problem solving	use Internet for mail
simple problems	use Internet to find work-related info
complex problems	use Internet to conduct transactions
	use computer to work with spreadsheets
Literacy tasks	use computer to work with Word
Reading	use computer for programming
read directions or instructions	use computer for conference calls
read letters, memos or mails	
read newspapers or magazines	**Management tasks**
read professional journals or publications	planning own activities
read books	planning others' activities
read manuals or reference materials	organising own time
read financial statements	
read diagrams, maps or schematics	**Manual tasks**
Writing	working physically for long
write letters, memos or mails	using hands or fingers
write articles	
write reports	
fill in forms	

TABLE 2

	Belgium	Czech Republic	Denmark	France	Italy	The Netherlands	Poland	Slovakia	Spain
% in Sample	0.095	0.103	0.150	0.128	0.081	0.111	0.145	0.092	0.095
Female	0.479	0.501	0.488	0.480	0.439	0.480	0.427	0.474	0.467
Log of montly	8.013	7.116	8.057	7.668	7.686	7.76	7.012	7.061	7.564
earnings (PPP adjusted)	(0.630)	(0.812)	(0.326)	(0.690)	(0.725)	(1.004)	(0.723)	(0.751)	(0.711)
Full-Time	0.765	0.902	0.818	0.823	0.840				
Age 16-25	0.115	0.163	0.123	0.099	0.066	0.169	0.453	0.111	0.098
Age 26-35	0.235	0.266	0.157	0.215	0.211	0.176	0.233	0.253	0.251
Age 36-45	0.268	0.235	0.225	0.267	0.348	0.241	0.128	0.258	0.230
Age 46-55	0.282	0.194	0.241	0.281	0.256	0.231	0.120	0.262	0.243
Age 56-65	0.100	0.144	0.254	0.139	0.120	0.157	0.067	0.116	0.109
Hours worked	43.300	43.861	40.56	40.604	41.563	41.42	43.295	43.424	38.102
(Full-Time)	(10.608)	(9.793)	(8.099)	(9.172)	(10.389)	(9.205)	(10.877)	(9.302)	(12.490)
Native Speaker	0.916	0.971	0.815	0.897	0.884	0.903	0.983	0.953	0.943
Live with Spouse	0.821	0.681	0.839	0.818	0.668	0.793	0.519	0.713	0.731
Spouse Employed	0.691	0.709	0.685	0.664	0.576	0.495	0.688	0.740	0.592
Primary School	0.027	0.000	0.008	0.045	0.037	0.055	0.009	0.002	0.157
Lower Secondary	0.093	0.075	0.169	0.137	0.253	0.202	0.071	0.085	0.242
Upper Secondary	0.409	0.640	0.352	0.443	0.481	0.398	0.545	0.668	0.196
Professional Degree	0.292	0.066	0.236	0.137	0.015	0.042	0.055	0.006	0.121
Bachelor	0.018	0.041	0.082	0.115	0.189	0.206	0.104	0.044	0.127
Master/PhD	0.161	0.177	0.153	0.122	0.027	0.098	0.217	0.195	0.159

TABLE 3

Sample	**Variables**	**Coefficient on female**	**Standard Error**	R^2
Full-time	Basic	-.211177	.007657	.4099
Full-time	Basic,hours	-.182233	.007732	.4203
Full-time	Basic, hours, education	-.232798	.007343	.4893
Full-time	Basic, hours, education, occupation	-.179904	.008519	
Full-time	Basic, hours, education, occupation, tasks	-.138170	.009718	
All	Basic	-.345860	.007750	.3892
All	Basic,hours	-.172822	.007384	.4940
All	Basic, hours, education	-.212330	.007064	.5445
All	Basic, hours, education, occupation	-.152139	.008163	
All	Basic, hours, education, occupation, tasks	-.121119	.009350	
Full-time,BA	Basic	-.230887	.014699	.4392
Full-time,BA	Basic,hours	-.176040	.014688	.4669
Full-time,BA	Basic, hours, education	-.174857	.014526	.4795
Full-time,BA	Basic, hours, education, occupation	-.122870	.015486	
Full-time, BA	Basic, hours, education, occupation, tasks	-.105558	.015794	
All,BA	Basic	-.303490	.015417	.3823
All,BA	Basic,hours	-.148457	.014389	.4986
All,BA	Basic, hours, education	-.145786	.014255	.5092
All,BA	Basic, hours, education, occupation	-.101672	.015132	
All,BA	Basic, hours, education, occupation, tasks	-.088828	.015478	
Full-time, HS	Basic	-.271620	.009555	.4143
Full-time, HS	Basic,hours	-.253442	.009633	.4202
Full-time, HS	Basic, hours, education	-.258493	.009589	.4268
Full-time, HS	Basic, hours, education, occupation	-.218627	.011955	
Full-time, HS	Basic, hours, education, occupation, tasks	-.177314	.015549	
All, HS	Basic	-.408911	.009612	.3937
All, HS	Basic,hours	-.234921	.009134	.5007
All, HS	Basic, hours, education	-.240689	.009064	.5091
All, HS	Basic, hours, education, occupation	-.179922	.011164	
All, HS	Basic, hours, education, occupation, tasks	-.154198	.014512	

TABLE 4

ISCO 08	Occupation	Number employed	D index
5223	Shop Sales Assistants	1273	4.514
4110	General Office Clerks	787	2.119
9112	Cleaners and Helpers in Offices, Hotels and Other	748	2.973
2341	Primary School Teachers	720	2.939
2330	Secondary Education Teachers	526	2.782
2221	Nursing Professionals	463	3.227
3322	Commercial Sales Representatives	448	2.530
4311	Accounting and Bookkeepin Clerks	440	3.601
4321	Stock Clerks	395	2.527
5321	Health Care Assistants	394	2.293
9111	Domestic Cleaners and Helpers	371	3.326
5322	Home-Based Personal Care Workers	341	1.983
3313	Accounting Associate Professionals	337	3.570
5131	Waiters	337	1.934
5230	Cashiers and Ticket Clerks	333	2.310

TABLE 5

	(1)	(2)	(3)	(4)
Sample: all occupations with at least 20 people, 10 men & 10 women				
D-index	0.055***	0.049**	0.029***	0.026**
	(0.010)	(0.010)	(0.010)	(0.011)
Female	-0.110***	-0.069***	-0.067***	-0.073***
	(0.023)	(0.023)	(0.022)	(0.022)
D-index*Female	-0.027**	-0.027*	-0.028**	-0.024*
	(0.013)	(0.013)	(0.013)	(0.013)
Female Rate	-	-0.167***	-0.132***	-0.106***
& -	(0.015)	(0.016)	(0.017)	
Occupational				
Characteristics	-	-	YES	YES
1-digit OCC	-	-	-	YES
Country	YES	YES	YES	YES
	(4)	(5)	(6)	(7)
Sample: occupations with 100+ people				
D-index	0.077***	0.082***	0.073***	-0.074***
	(0.014)	(0.014)	(0.013)	(0.014)
Female	-0.018	0.022	0.001	-0.033
	(0.028)	(0.028)	(0.028)	(0.028)
D-index*Female	-0.090***	-0.092***	-0.079***	-0.057***
	(0.017)	(0.017)	(0.017)	(0.017)
Female Rate	-	-0.176***	-0.143***	-0.150***
Occupational				
Characteristics	-	-	YES	YES
1-digit OCC	-	-	-	YES
Country	YES	YES	YES	YES

TABLE 6

Regressions with 4-digit occupational dummies	**Dependent Variables**	
	Female	Wage
Sharing work-info	.011***	.053***
Presentations	-.015***	.031***
Read diagrams, maps or schematics	-.035***	.037***
Use or calculate fractions and percentages	-.024***	.036***
Use advanced maths or statistics	-.017***	.022***
Use spreadsheets	-.012***	.015***
Occupation dummies 4-digit	YES	YES
Remaining tasks	YES	YES
Observations	23,475	20,603

Regressions with 3-digit occupational dummies

	Female	Wage
Sharing work info	.014***	.056***
Presentations	-.016***	.029 ***
Read diagrams, maps or schematics	-.038***	.040***
Use or calculate percentages or fractions	-.023***	.038***
Use advanced mathematics and statistics	-.019***	.024***
Use spreadsheets	-.013***	.017***
Programming	-.017***	-.011**
Occupation dummies 3-digit	YES	YES
Remaining tasks	YES	YES
Observations	23,475	20,603