FROM GENDER GAPS IN SKILLS TO GENDER GAPS IN WAGES: EVIDENCE FROM THE PIAAC

Yolanda F. Rebollo-Sanz[1] (**Corresponding Author:** yfrebsan@upo.es)

Universidad Pablo de Olavide

ORCID ID:**0000-0003-1876-709**

Sara De la Rica[2]

Universidad del País Vasco & ISEAK

This version: September 2019

Abstract

Our paper makes the first attempt using a cross-national perspective, to address the impact of skills in the analysis of gender gaps in labor market performance. To that end we use the OECD´s PIAAC dataset, which contains information on OECD and non OECD economies. Firstly, we document that there are gender gaps in cognitive skills for numeracy, which are found to be around 3-4% and to increase with age. Next, we show that gender gaps in math skills are crucial to understanding a substantial proportion of the gender gaps observed in labor market outcomes. In particular, the higher math skills exhibited by males explain 45% of the gender gap in labor force participation for young workers and 29% for workers aged 30-39. Finally, gender differences in math skills help to explain 40% of the gender wage gap observed. That impact increases from entry age to 30-39 (motherhood age), as we find that math skills explain 44% of the gender wage gap observed for young workers and as much as 55% for workers aged 30-39.

JEL Classification: J16, J24, J31

Keywords: gender wage gap, gender gap in labor force, cognitive skills, PIACC

[1] Corresponding author: Yolanda F. Rebollo-Sanz: Departamento de Economía, Métodos Cuantitativos e Historia Económica, Area de Análisis Económico, Universidad Pablo de Olavide, Carretera de Utrera km.1 41013, Seville, Spain. Email: yfrebsan@upo.es. ORCID ID:**0000-0003-1876-709**

[2] Sara de la Rica: Departamento de Economia, Universidad del Pais Vasco, Email: sara.delarica@ehu.es

Financial support from MINECO of Spain through project ECO2013-48884-C3-3-P, and from Junta de Andalucía through project SEJ-1512 is gratefully acknowledged.

1. Introduction

Over the last twenty years, there has been a narrowing of gender gaps in different labor market outcomes such as employment rates, hours worked, and wage rates, among others, in all advanced economies, mainly driven by a large expansion in women's education. However, these gender gaps still persist even after many observed characteristics of workers concerning education such as age, experience, and demographic characteristics are taken into account (Blau, Brinton, and Grusky 2008; England, Gornick, and Shafer 2012; Blau and Kahn 2017). This is referred to as the "unexplained" or "adjusted" gender gap and, as some empirical papers show (Blau and Kahn, 2017; Boll et al, 2016), unexplained factors still accounted for a substantial share of the raw wage gender gap in 2010.[3] In this paper we test whether it is possible to reduce the unexplained component of this gender gap by using more direct measures of human capital endowments. In this vein, the appearance of new datasets that seek to construct such direct measures enable more precise tests to be conducted into whether gender differences in human capital endowments other than years of schooling are behind the unexplained gender gap in labor market outcomes. The the Programme for the International Assessment of Adult Competences (PIAAC) offers these direct measures of human capital endowments, such as cognitive skills for the adult population across a significant number of countries. Our paper deals precisely with the association between the adjusted gender gaps in labor market outcomes and gender gaps in cognitive skills. To measure this association, we suggest that cross-country variation provide clearer evidence than within-country analysis because of the better picture of broad differences in the social and economic environment.

More precisely, our contribution with respect to previous studies is threefold: First, this paper uses precise measures of cognitive skills provided by PIACC to assess whether gender gaps exist in cognitive skills for different ages and if so how big they are. Second, we focus on the relationship between gender gaps in math skills and labor force participation, which to our knowledge, has not previously been addressed. In this, we seek to explore the link between math skills and self-selection into the labor market, particularly for women. Third, related to the second issue, we investigate the link between math skills and wages. We account for the potential endogeneity of math skills for labor force participation and wages given that the former may be affected by different labor market trajectories. We are particularly interested in comparing findings at entry age (24-29) and at the next age group (30-39), where motherhood mostly takes place. This perspective provides a more holistic view of how the link between gender gaps in skills and labor market outcomes evolve through life. Gender gaps in areas such as wages and hours worked change substantially at different stages of the

[3] Blau and Kahn (2017) estimate this unexplained component is around 85% of the raw gender gap. Boll et al (2016) estimate this is around 60%.

life cycle as a result of motherhood. Moreover, educational attainment may still play a role in explaining gender differences for older workers but not so much for younger workers.

In related literature, cognitive skills have been found to be positively associated with the success of individuals in the labor market, participation in society, and economic growth (Oreopoulos & Salvanes, 2011; Hanushek & Woessmann, 2015; Hanushek et al 2015; Hampf, Wiederhold & Woessmann, 2017). However, there is hardly any evidence on the impact of properly measured skills in the labor market on gender gaps observed in different labor market outcomes. So far, only Hanuseck et al (2015) and Fortin (2008) provide some insights on the link between cognitive abilities and gender wage gaps, although their empirical analysis does not seek to measure the role of gender gaps in cognitive skills with a view to explaining gender gaps in labor market performance.

The assessment of gender differences in cognitive skills, particularly in numeracy skills in adulthood, is appealing since results from PISA persistently find that females at age 15 perform consistently around 5% more poorly in numeracy skills than their male counterparts (see Arora & Pawlowski, 2017). Such gender disparities may lead, at least partly, to the documented lower presence of women in the fields of study of science, technology, engineering, and mathematics. Joensen & Nielsen (2014) shows that encouring more students to opt for advanced mathematics has a sizeable positive earnings effects for girls, but no effect for boys at the margin. Moreover, the gender segregation by occupation observed in the labor market – young women tend to study fields such as education, health, and social sciences whereas technical studies are primarily male fields – may also be a consequence of the poorer performance of girls in numeracy skills. Hence, it is necessary to document empirically the size of gender gaps in cognitive skills, particularly in numeracy skills, and then assess the extent to which such gender disparities affect labor market performance in order to examine in depth the determinants of gender gaps in the labor market.

Gender gaps in labor market outcomes are likely to be heterogeneous depending on whether or not there are children in the household. In most developed countries, women must combine employment with home responsibilities to a greater extent than their male partners. This affects their decisions with respect to their labor supply, affects their human capital accumulation, and hence affects their labor-market performance in terms of time employed, type of job, wages, and accumulation of skills.[4] To account for this, in our empirical analysis we test whether the contribution of gender gaps in cognitive skills differs for parents and non-parents.

Data shows that on average gender gaps in literacy skills are negligible even on entry into the labor market and remain so at different ages. However, men display cognitive numeracy skills around 4% higher than those of women (a difference of around 7-8 points on a 500-point scale) and this gap

[4] For instance, previous studies have concluded that for women, being married and having young children reduce labor force participation and the probability of paid employment, whereas for men being married increases labor force participation and the probability of paid work and having young children has no significant impact.

increases substantially from 2.3% at the age of entry into the labor market (24-29) to 4.3% for the 30-40 age group. Our results indicate that these gender differences in skills strongly affect gender gaps observed in labor market participation. In particular, they explain 45% of the gender gap observed in labor force participation for young workers and only 23% for individuals in the next age group. Substantial differences are found when gender gaps are compared for non-parents and parents separately. Lastly, gender differences in math skills are also crucial to understanding the gender wage gap observed. On average, differences in math skills explain 40% of the gender wage gap observed. That impact also increases from entry age to 30-39 (motherhood age), as we find that math skills explain 44% of the gender wage gap observed for young workers and as much as 55% for workers aged 30-39.

The rest of the paper is organized as follows: Section 2 describes the dataset used. Section 3 analyzes gender gaps in literacy and numeracy skills and gives descriptive statistics of gender gaps by age and by country. Section 4 focuses on the relationship between math skills and gender gaps in labor market outcomes. Section 5 concludes.

2. The Programme for the International Assessment of Adult Competencies (PIAAC)

The data source used in the paper is the Programme for the International Assessment of Adult Competencies (PIAAC), developed by the OECD to provide internationally comparable data on the skills of the adult population. The first round of PIAAC data, collected between August 2011 and March 2012, produced data on mostly OECD countries (see OECD, 2013). In a second round[5], PIAAC conducted the same skill survey in nine more countries (including both non-OECD countries and new members of the OECD) between April 2014 and March 2015 extending the usable sample with comparable skill data to 23 countries with information on wages (though the total number of countries is 31). At least 5,000 adults participated in the PIAAC assessment in each country. In each participating country a representative sample of adults between 16 and 65 years of age was interviewed at home in the language of their country of residence.

The PIAAC was designed to measure key cognitive skills needed for individuals to advance at work and participate in society. It hence extends the information obtained from the various PISA waves, which measured those skills at the age of 15. These new skill variables together with socioeconomic and labor market covariates are offered in the PIACC database. The combination of these two sets of information gives greater value to this database. In particular, the survey includes three main sets of information: First, a personal interview comprising a questionnaire about personal background, educational attainment and training, current work status, wages, and work history. Information on

[5] Round-1 countries: Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Ireland, Italy, Japan, Korea, Netherlands, Norway, Poland, Slovak Republic, Spain, United Kingdom. Round-2 countries: Chile, Greece, Israel, Lithuania, New Zealand, Slovenia

family background, linguistic background, health status, and civic participation is also provided. Second, there is an assessment of cognitive skills in three domains – literacy[6] (mainly reading skills), numeracy, and problem-solving skills in technological settings. The assessments are explicitly implemented as international or cross-national assessments designed to provide valid, reliable measures of proficiency across different countries, languages, and cultures. This unique information on skills at individual level, together with standard labor market information such as wages, educational attainment, labor market experience, and type of job makes the PIACC a highly suitable data source for our purpose. Gender gaps in skills can be accounted for at different ages using international data from a harmonized dataset.

The skills measured in the PIAAC are Cognitive Foundation Skills (CFS) or "key information-processing skills". This paper focuses on literacy and numeracy skills since problem-solving[7] skills are not available for some of the countries in our reference sample. Firstly, "literacy" is defined as *the ability to understand, evaluate, use, and engage with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential*; Secondly, "numeracy" is defined as *the ability to access, use, interpret, and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life.* Each cognitive skill is assessed on a 500-point scale.[8]

An issue worth mentioning is the implication of complex sample design for calculating error variance. The error variance of sample statistics in the PIAAC consists of two components: Sampling variance, which reflects uncertainty due to obtaining a specific sample from the population, and Imputation Variance, which reflects uncertainty due to the random draw of plausible values. The Jackknife Replication Approach was used to calculate replication weights. We use the information provided by the PIACC on sampling variance and imputation variance to estimate the parameters of interest as well as standard errors.

[6] The focus of the PIAAC is on certain aspects of literacy, in particular the understanding and use of texts. Writing skills and the ability to produce or format documents are not assessed. This is not because these skills are not considered as important aspects of literacy in broad terms but largely because of the practical difficulties associated with assessing adults' writing in large-scale international surveys.

[7] "Problem solving in technology-rich environments" defined as the ability to use digital technology, communication tools, and networks to acquire and evaluate information, communicate with others, and perform practical tasks (ICT skills - that is, skills in using information and communications technology).

[8] The objective of the assessments is to describe the level and distribution of the skills of the adult population, not to test the proficiency of individuals. The total number of items used in the assessments is greater than the number answered by any single respondent, each of whom undertakes a subset of the tasks administered.

3. Statistical Evidence of Gender Gaps in Cognitive Competences

Our baseline sample is composed of native workers aged 24-49.[9] In order to maintain a homogeneous sample when linking skills with gender gaps in labor market outcomes, we restrict our analysis to the 23 countries with information on individual wages. We start our empirical analysis by testing whether there is evidence of gender gaps in cognitive abilities and, if so, describing how they change for different age groups. Our reference sample is divided into three age groups: a) Entrants into the labor market (aged 24-29); b) Prime age workers aged 30-39; c) Prime age workers aged 40-49. Particular attention is paid to the age group that we denote as "entry age" (24-29 years) compared with the next age group (30-39), when motherhood plays an important role for women. Empirical evidence (Kleven et al 2018 among others) finds a huge increase in the gender gap in labor market participation and earnings, due to the penalty that maternity imposes for women's careers relative to men's. Paull (2008) finds a substantial movement towards part-time work for women that occurs with the first birth and continues steadily for ten years. Gallen, Lesner and Vejlin (2019) show for Denmark that 30 percent of the gender hours gap can be explained by the sorting of women into lower-hours workplaces. This hours gap is driven by mothers, the group for whom differences in employer, occupation, education, and experience also imply large differences in wages. The PIACC database is purely cross-sectional and it does not permit individual labor market trajectories to be followed. To look at differences in gender gaps in different life cycle phases we treat gender gaps in different age groups separately. We focus on gender differences particularly for these two age groups, as gender differences for older workers may be affected by differences in labor market trajectories that our dataset cannot capture.

Sample statistics shown in Table 1 motivates the analysis by examining how numeracy and literacy skills vary by gender and across ages. This Table 1 depicts mean literacy and numeracy test scores measured on a 500-point scale for all individuals and for current workers, respectively.[10] The first point to note is that women exhibit lower competency levels in numeracy skills. On average, the gap for all individuals is 4.1% (11 points on the 500 scale). As expected, it is somewhat smaller for the sample of workers at 3.5% (10 points on the 500 scale). This difference in gender gaps between the whole sample and the employed population points to some kind of positive sorting into the labor market for women. For the two groups of individuals, the pattern by age is very similar. This reinforces interest in understanding the link between cognitive abilities and gender gaps in different areas of

[9] We restrict our sample to native workers since results could be distorted by using the full sample for two reasons: Firstly, immigrants might face more problems in answering correctly concerning cognitive skills; secondly, measurement of educational attainment levels can be very different between samples of natives and immigrants.

[10] To interpret these statistics appropriately, note that in the PIAAC each area of cognitive skill is a latent variable that is estimated using item-response-theory models (see OECD 2013 for details). The PIAAC provides 10 plausible values rather than only one individual score for each respondent and each skill domain. Using the average of the 10 plausible values provides an unbiased estimate of individual skills in each domain. The sample statistics shown in Table 1 use this average, which uses the weights provided by the PIACC to control for sampling variance that reflects uncertainty due to obtaining a specific sample from the population.

labor market performance by age groups. The gap is -2.16% in the 24-29 age group and reaches a maximum of -4.3% in the 40-49 age group. This evidence is striking when one considers that females either systematically outperform males or have made enormous gains on many educational dimensions.[11]

With respect to literacy, however, the picture is rather different. On average, men and women score very similarly so there is no gender gap on average for the full sample. However, there is a positive gap (in favor of women) for the sample of workers more attached to the labor market. This gender gap in literacy remains constant across all age groups.

Finally, another interesting issue is that from age 40 onwards both men and women obtain lower scores in both literacy and numeracy skills. This is likely to be primarily an age rather than a cohort effect. For instance, Green and Riddell (2013) document a cohort-level fall in literacy after age 45, suggesting that skills deteriorate over the life-cycle. Moreover, a look at gender differences in numeracy and literacy skills from the different PISA waves shows gender gaps to be surprisingly stable across the different waves, and hence across the different cohorts (at least at the age of 15). Unfortunately, we have no evidence from different waves for the PIAAC that could provide additional evidence with regard to the importance of cohort effects at older ages.

Table 1 reveals that gender gaps in cognitive abilities are significant for math skills and negligible for literature abilities. De la Rica and Rebollo-Sanz, (2018) offer a more detailed analysis and show that adjusted gender gaps in numeracy hardly vary (remain around 3.5%) even when controlling for individual observed characteristics such as age and education and labor market characteristics such as type of occupation. Similarly, they find that adjusted gender gaps in literacy are not statistically significant.

Henceforth, though the type of education women receive has changed in many countries toward more mathematics-oriented programs and females , relevant gender differences remain in numeracy skills. So from now on, our analysis is restricted to gender gaps in math skills and their role in explaining gender gaps in labor market performance (employment rates and wages).

[Insert Table 1 here]

4. From Gender Gaps in skills to Gender Gaps in Labor Market Outcomes

4.1. Empirical Methodology

To examine the link between gender gaps in skills and gender gaps in labor market outcomes, we use a standard human capital model. Typically, gender gaps are estimated as equation (1) describes.

[11] For instance, There are more graduating females from four-year colleges than males (Goldin, Katz, and Kuziemko (2006). Additionally, the high school dropout rate tends to be lower for females compared to males.

$$Labor - Market - Outcome(i) = \alpha + gender(i)\beta_0 + H(i)\gamma_0 + \varepsilon(i) \quad (1)$$

For each individual i, labor market outcome (labor market participation or log wages, in our case) is regressed on a set of covariates that proxy for human capital levels (H_i) plus an indicator of gender, which takes a value of one if the respondent, is female and zero if male. The coefficient $\{\beta_0\}$ shows the adjusted gender gap in the corresponding labor market outcome conditional on the same observed human capital levels. We then expand this basic function to include a more direct measure of human capital, in particular, cognitive skills (*skills*) along with the more traditional –indirect-, measures of human capital such as years of schooling and labor market experience (effective, measured by years of employment and/or potential, measured by age).[12]

$$Labor - Market - Outcome(i) = \alpha + gender(i)\beta_1 + H(i)\gamma_0 + skills(i)\gamma_1 + v(i) \quad (2)$$

Now the coefficient β_1 measures the adjusted gender gap for individuals who share not only traditional human capital levels but also similar cognitive levels. A growing literature shows that even within educational levels, there are statistically significant returns on cognitive skills in terms of labor market outcomes $\{\gamma_1>0\}$.[13] Hence, if there are gender gaps in cognitive skills (in favor of men) we expect the estimated gender gap $\{\beta_1\}$ from equation 2 to be smaller than that in equation 1, β_0, as men and women share similar cognitive skills. Such a drop is precisely the contribution of differences in cognitive skills to explaining the adjusted gap in labor market outcomes. We examine two main labor market outcomes: Labor Market Participation –a crucial factor in understanding development in women´s wages-, and wage rate (in logs) [14]. Labor Market Participation is measured by the employment status of the individual and takes value one when she is observed employed and zero otherwise.[15] The wage information collected in the PIAAC refers to gross (pre-tax) earnings. Our definition of wages also considers discretionary bonus payments since the unexplained part of the gender wage gap is typically higher and this is typically related to more qualified jobs or jobs where skills might play a major role.

Numerical skills may be endogenous for both labor market outcomes, as skills and labor market experience are likely to be causally related (see Hampf, Wiederhold & Woessmann, 2017; Jimeno,

[12] Even when measures of skills are included, we expect years of schooling to affect labor market participation because of measurement error (i.e. our skills measures are not likely to be a comprehensive set) and because years of schooling may be valued per se in the labor market, where they are perceived to signal useful information about the individual.

[13] Hanushek et al (2015) use individual information on numeracy cognitive skills from the PIAAC to account more precisely for the size of the returns on skills for wages and conclude that a one-standard-deviation increase in numeracy skills is associated with an 18% increase in wages among prime age workers. Note, however, that their baseline model does not include years of schooling. Hampf, Wiederhold & Woessmann (2017) also use the PIACC to explore several approaches that seek to address potential threats to causal identification of returns on skills in terms of both higher wages and better employment chances.

[14] Using pre-tax earnings has the advantage of capturing how the market rewards certain characteristics before the mediating effect of the tax system is felt. However, it might potentially bias the cross-country comparison of wage dispersion to the extent that different countries differ in the progressivity of their tax systems.

[15] Retired individuals and full time students are omitted from the sample of analysis.

LaCuesta & Villaverde, 2016). Cognitive skills acquired in childhood are likely to affect future labor market paths: cognitive skills enable individuals to understand and perform better. Work experience may vary across similar individuals due to extended periods of unemployment or nonparticipation in the labor market, which, in turn, may affect cognitive skills. This productivity-enhancing effect of skills increases a person's wages (workers obtain better-paid jobs and have more stable labor market paths) or allows him or her to escape unemployment and find a job in the first place.[16] To cater for this, we apply instrumental variable estimation, using as our instruments covariates which might influence cognitive skills levels but that are predetermined to the entry into the labor market (parental educational attainment, whether the individual is a second or third generation immigrant (i.e native-born children of Immigrants) and whether she/he is a native speaker.[17] Our identifying assumption is that they are correlated with math scores but conditional on general capital endowment such as individual education level and labor market experience, are uncorrelated with labor market outcomes.[18]

For the wage equation, we test whether taking into account the labor market participation decision is relevant to understanding the contribution of cognitive skills to the gender wage gap. To that end, we use the Heckman two-step estimator for the wage equation.[19] We use this approach also to address the measurement error due to the unobserved wages of men and women who are unpaid workers. We use marital status, whether or not there are children, health status and the presence of other workers in the household to identify the labor force participation function. We assume that these variables affect wages only through the employment decision, but not directly.

Taking into account motherhood:

In most developed countries, a disproportionate burden is placed on women to provide unpaid care in the home – women provide 3.1 times the care work of men. This affects their decisions with respect to their labor supply, affects their human capital accumulation, and hence affects their labor-market performance in terms of time employed, type of job, wages, and accumulation of skills. For instance, Meurs et al. (2010) conclude that a child has an impact on career interruption and consequently on women's wages. Similarly, Weeden et al. (2016) argue, namely that much of what appears to be a gender wage gap is a gender-specific family gap in pay and that most of it could be explained by factors directly or indirectly related to motherhood. To factor in children, we estimate an expanded

[16] The covariate *years of working experience* correlates with performance in the PIAAC mostly among less educated individuals (see Jimeno, LaCuesta & VIllaverde, 2016).
[17] To obtain a consistent estimator of the impact of math skills on labor market outcomes we must disentangle the math skills variable from the effects of heterogeneous labor market experience across individuals so as to leave those gender differences in cognitive skills, if any, that emerge from differences arising prior to labor market entry.
[18] This endogeneity problem is expected to be much less significant for younger workers than for older ones, whose abilities are measured at about labor market entry wage. Hence, using data on skills at young ages minimizes the problem of work experience contributing to the formation of observed cognitive skills. However, we still face potential endogeneity biases for workers with labor market experience.
[19] We estimate the two-step Heckman selection estimator instead of the full maximum likelihood model to use the sampling weights and replication weights properly to get consistent estimates and adequate standard errors of parameter estimates.

model where the *gender* covariate is interacted with an indicator of whether individuals are parents or not, i.e. we also estimate models of the following kind (Equation 3):

$$\begin{aligned} Labor-Market-Outcome(i) \\ &= F(\alpha + (Women(i) * Non-Parent(i))\beta_{wnp} + (men(i) * Non \\ &- Parent(i))\beta_{mnp} + (Women(i) * Parent(i))\beta_{wp} \\ &+H(i)\gamma_0 + skills(i)\gamma_1 + u(i)\ ) \end{aligned} \tag{3}$$

A comparison of the coefficients of *Women*Non-Parent and Men*Non-Parent,* $\{\beta_{wnp}-\beta_{mnp}\}$, reflects the estimated gender gap in the corresponding labor market outcome of non-mothers versus non-fathers, conditional on human capital variables. Gender gaps between mothers and fathers are defined by the coefficient associated with *Women* Parent* $\{\beta_{wp}\}$. Notice that this specification also enables us to test for the family gap between genders. For instance, the family gap for women is the difference between $\{\beta_{wnp}-\beta_{wp}\}$, whereas that for men is defined by β_{pnp}.

4.2. Results – From GG in Skills to GG in Labor Market Performance

In accordance with the previous empirical literature, our data show substantial gender gaps in the main two outcomes of interests. Table 2 shows average gender gaps in Labor Market Participation –proxy by the probability of being employed-, and hourly wages for the whole sample and for different ages. On average, the two gaps are very similar at around 12-13%, and both increase with age. The increase in the gap in Labor Market Participation from entry age to the 30-39 age group, is higher than for wages.

[Insert Table 2 here]

Gender Gaps in Math Skills and Gender Gaps in Labor Market Participation

Gender differences in experience and labor force attachment have been seen as central to the understanding of the gender wage gap. For that reason, we proceed to test the link between gender gaps in math skills and gender gaps in labor market participation by estimating the probability of being in work (working versus not working) for our sample of 23 OECD countries. We do this for the whole sample (20-49) and for each of the aforementioned three age groups separately: 24-29 (entry age), 30-39 (entry into maternity) and 40-49 (mature or older). We pool observations of all countries and introduce country-specific dummies. In this section, we use the full sample of individuals: employed, unemployed or inactive if the reason for inactivity is different to health issues, retirement or full time studies.

We present different models following the steps described in the Methodology section. The baseline model –Model 1- contains control variables for individual characteristics such as gender, age, education (four dummies), health status (good/bad health) and some household characteristics (having a partner or not; having a child or not; labor situation of the partner). Country fixed effects are included in all estimations so each coefficient must be interpreted as within country estimates. We estimate a second model – Model 2 – that adds to the baseline model 1, the covariate of cognitive ability in numeracy, represented by individual math test scores. Model 3 uses the same covariates as Model 2 but uses instrumental variable estimation to account for the potential endogeneity of math skills. The variables used as instruments are parental educational attainment for mothers and fathers separately, whether the individual is a second or third generation immigrant and whether she/he is a native speaker. All these variables are predetermined to the labor market entry of individuals: basically, we are assuming that family background defined by parental education, immigrant status of parents and native language influences an individual's labor market paths through its effects on his/her human capital endowment such as math scores. The main sample statistics for the additional covariates used in the analysis are presented in Table A.1 in the Appendix.

Tables 3A and 3B show estimated marginal effects from probit estimations of the three models of interests described above. Table 3A display estimated gender gaps in labor market participation (LMP from now on) without disaggregating men and women by their status as regards having children as described in Equation (2) in the methodological section. Table 3B presents the estimated gaps separately for non-parents and parents as described in Equation (3) in the methodological section. For the sake of brevity, in both tables, only the estimated marginal effects associated with the gender covariates are reported. The whole set of coefficients is presented only for our preferred Model 3 in Tables A.3 and A.4 in the Appendix.

Table 3A, where parenthood is not explicitly considered, reveals that gender gaps in labor market participation are substantial. For the pool sample, the conditional average gender gap in LMP in the baseline model is 17.9 pp, i.e. the probability of being in work is 17.9 percentage points lower for women than for men with similar observed characteristics in terms of education level, age, and health and family status. Looking at age groups, we find, as expected, that the gender gap is lowest for young workers (10 pp.) and increases substantially for workers in the next age group (to 22 pp.).

Models 2 and 3 show that math skills affect positively on the probability of being in work and reduce the adjusted gender gap. The estimated coefficient of math skills is positive and statistically significant in almost all cases. We use Model 3 –model using an instrumental variable approach- to measure the extent to which gender gaps in skills explain gender gaps in labor force participation. To that end, we compare adjusted gender gaps from Model 3 with those from the baseline model (Model 1), where math scores are not included in the estimation. That comparison reveals a substantial (and statistically significant) decrease in all gender gaps except for older individuals. In particular, for young workers, the gender gap drops from 10 pp. to 5.5 pp. (a drop of 4.5 pp. in absolute terms or 45% in relative terms), and for the 30-39 age group the labor market participation gap drops from 21.8 pp. to 16.8 pp. (a drop of 5 pp. or 22.9% in relative terms).

[Insert Table 3A here]

A look at Table 3B reveals whether the contribution of skills to gender gaps in LMP differs for parents and non-parents. As in Table 3A, returns to math skills are positive and statistically significant in all samples except for older workers. From the estimated marginal effects shown in Table 3B we compute gender gaps in labor market participation separately for parents and non-parents. These gender gaps are presented in Figure 3 to better illustrate the results. The first striking result when looking at all individuals and all models is the difference between parents and non-parents in gender gaps for participation. This confirms that parenthood is a crucial aspect to account for when looking at gender gaps in LMP. The second result worth highlighting is the role of math skills in explaining these gender gaps for parents and non-parents separately. Comparing the estimated gender gap in the baseline model (Model 1) with that in Model 3 (where math skills are factored in) a reduction of the estimated gender gap can be seen. In both cases, the gap decreases by around 2 pp, but in relative terms the reduction is much larger for non-parents given that the gap in absolute terms is much lower. This indicates that for parents there are likely to be other determinants not captured in the model, which probably explain more of the gender gap in LMP.

[Insert Table 3B here]

These striking differences in gender gaps in LMP for non-parents and parents are observed in all age groups. With respect to the extent to which math skills explain these differences, it is interesting to note that for non-parents aged 24-29 differences in math skills explain the entire observed gender gap (though it is small in absolute terms). For non-parents in the next age group the introduction of math skills reduces the GG in LMP from 9.2 pp to 6.5 pp (2.7 pp, 29% in relative terms), which does not account for the whole gap but is nevertheless substantial. Turning to parents, for young individuals differences in math skills also explain a substantial fraction of the gap in participation (6.8 pp, 22% in relative terms). And for parents aged 30-39 math skills account for a reduction of 5.7 pp in LMP gender gaps, which in relative terms amounts to 18 %.

[Insert Figure 3 here]

Summarizing, though differentials in LMP have decreased in many countries we obtain that these gaps are still important even comparing men and women with similar human capital endowments. Moreover, gender differences in math skills play a determinant role in explaining gender gaps in LMP, particularly for workers less than 40 years old and for non-parents. For parents math skills still play a relevant role for gender gaps in participation, but that role is more limited as gender gaps conditional on math skills are still very substantial. This indicates that in addition to differences in math skills there are other (unobserved) factors which determine the differences in labor force participation between mothers and fathers even at young ages.

Gender Gaps in Math Skills and Gender Wage Gaps

To estimate the extent to which math skills explain the gender wage gap we proceed as before: We run a typical log wage regression for all groups and for each age group separately as a function of the standard human capital variables such as age, educational attainment, and labor market experience (Model 1). Then we run an additional regression where math scores are added as a covariate (Model 2). Thirdly, we use instrumental variables to free estimations from potential endogeneity biases related to labor market trajectory from the observed math skill variable (Model 3). Finally, we add other controls related to job characteristics, such as occupation (nine groups, one digit level), full time job, and firm size (more than 250 employees) to estimate gender wage gaps conditioning men and women who share very similar job conditions (Model 4). The lack of common support and the fact that the results for Model 4 barely change from those of Model 3 lead us to consider Model 3 as our reference model. The main sample statistics for the additional covariates used in the analysis are presented in the Table A.2 in the Appendix.

The main results for the estimated gender gap in wages are presented in Table 4A. Detailed results for our preferred Model 3 are presented in Tables A.5 and A.6 in the Appendix.[20] The first result to highlight is that the adjusted gender wage gap between men and women with similar (observed) human capital levels averages 20%, meaning women earn on average 80% of male earnings. If this gap is examined separately for different ages, we get that the gender pay gap begins early in the working life and a clear increase with age becomes apparent: At entry, it is 10%, but it increases to 18.6% in the 30-39 age group and again to 25% for men and women aged over 45.

[Insert Table 4A here]

To answer the main question posed by the paper, i.e. the extent to which gender gaps in math skills explain the adjusted gender wage gap, the gender gaps from Model 1 must be compared with those from Model 3. The answer is that for the full sample, math skills explain about 8 pp (40% in relative terms) of the adjusted gender wage gap. In other words, if men and women had the same math skills, in addition to other observed characteristics such as education, gender wage gaps would be on average 40% lower. Looking separately at the impacts for different ages, Table 4A reveals that differences in math skills explain 4.2 pp (44% in relative terms) of the whole gender wage gap for the youngest

[20] The typical variables such as experience and education show, as expected, a significant positive effect on wages.

group. That contribution then increases to 10 pp (55%) for the medium age group before dropping to 8.5 pp (34%) for the oldest group. As mentioned above, at older ages the gender wage gap may be due to different labor market trajectories between men and women, and much of it is likely to be unobserved in the data. As such, we focus on the youngest group and on the differences between entry age (24-29) and the onset of motherhood age (30-39). For these two groups we find a very strong impact on entry and an even stronger one in the 30-39 age group for math skills as a determinant of the gender wage gap observed.

Given the increasing extent to which math skills explain the gender wage gap for the age group where the onset of motherhood mainly takes place, we estimate that contribution separately for non-parents and for parents. Table 4B displays the results. As before, to better illustrate the results separately for non-parents and for parents given the presence of different interactions in the estimation, Figure 4 displays the estimated gender wage gap for each model separately for non-parents and for parents. As with participation, the first issue to highlight from Figure 4 is that gender wage gaps are notably higher for parents than for non-parents. As before, this indicates that parenthood is an issue that must be considered when looking into gender wage gaps.

[Insert Table 4B here]

Next, we focus on the whole sample (ages 24-49) and compare adjusted wage gaps between men and women with similar observed characteristics (Model 1 - baseline model). For non-parents, the gender wage gap is 12.7% whereas for parents it is 25.7% (more than double). Second, to learn the extent to which math skills explain this substantial gender wage gap we compare gender gaps from model 1 with those from model 3, as before, where math skills are included as a covariate and instruments are used to get unbiased estimates. Including math skills reduces the gender wage gap by 6.6 pp (48% in relative terms) for non-parents and by 15.7 pp (39%% in relative terms) for parents, which is lower in relative terms, but still very substantial. This indicates that differences in math skills between men and women explain 48% of the gender wage gap for non-parents and 39% for parents. As above, the conclusion is that to reduce the gender wage gap it is essential to try to eradicate the advantage of men over women in math skills.

To learn whether this impact changes at different ages, we present the results separately by age groups. At entry (ages 24-29), gender wage gaps are lower than at higher ages for both non-parents and parents. For men and women with similar observed human capital levels, we find a gender wage gap of 8% for non-parents and 19.4% for parents. The extent to which math skills explain these gaps is large: 4.1 pp (51%) for non-parents and 6.1 pp (31%) for parents. In the next age group gender wage gaps are greater than in the youngest group and the impact of math skills in explaining them is again very substantial at 6.3 pp (48%) for non-parents and 13.1 pp (57%) for parents.

Other interesting issues worth mentioning emerge from the results from Model 3 shown in Table 3B. Related to the impact of math skills on wages rather than gender wage gaps, we find that an increase of 10 points in math scores implies an increase in wages of around 4.7% at entry age, 9.2% at 30-39, and 7.4% for workers aged over 40. These results are in accordance with previous studies such as that of Hanuseck et (2015). They obtain that on average a one-standard-deviation increase in math skills increases wages by 18%. It is known that one standard deviation in math skills corresponds to 40 points, so it can be stated that Hanuseck et al. (2015) find that workers with math scores of about 10 points higher will get a return in the form of wages about 4.5% higher.

It is often argued that occupational segregation by gender is an important driver of the difference in the pay of men and women (see Blau and Kahn, 2000). Model 4 includes job characteristics and occupation in our wage regression but gender gaps remain similar. This result is in line with the idea that gender wage gap is specially driven within narrow defined occupations and not across different occupations (Golding, 2014). And henceforth, gender gaps in skills contribute more to explain the gender gap than segregation in occupations.

Robustness check- Selection into the labor market

We have stated above that math skills affect entry into the labor market, primarily for women. However, in the previous section we compute the impact of math skills in explaining gender wage gaps only for the sample of workers, and thus do not take self-selection into the labor market into account. It is necessary to check the extent to which the results found in the previous section change when that selection is explicitly factored into the estimation.

To do this, we re-run the four previous wage models, but using a two-step Heckman approach. In the first stage, we estimate the probability of being in work, obtain the Inverse Mills Ratio (IMR) (which computes (the inverse of) the probability of being in work), and include it in the second-step wage regression. The results can be seen in Table 5. For identification, we include whether or not there are children, health status and whether the partner is working.

[Insert Table 5]

The most interesting result from Table 5 is that the Inverse Mills Ratio is almost never significantly different from zero in the second-stage wage regression, particularly when math skills are factored into the estimation. This makes sense given the role played by math skills in labor market participation as described above. This reveals that self-selection bias is not an issue for the results found above, particularly if math skills are included as a predictor of wages.

6. Summary and Conclusions

The availability of more direct measures of skills for a cross-section of 23 OECD and non-OECD countries permits closer investigation than previously possible of the link between gender gaps in skills and gender gaps in labor market outcomes, in particular in the decision to work and in wages. To our knowledge, there are no prior empirical studies, which address this issue from an international perspective. In particular, we test whether there are gender gaps in cognitive abilities and seek to understand how those gaps might influence some important labor market outcomes at different ages, which represent different stages along the life cycle. We focus on literacy and numeracy skills since data on problem solving skills is not available for all 23 countries considered in the analysis. To that end, we use data from the OECD's Programme for the International Assessment of Adult Competencies (PIAAC), which offers unique information on skills at individual level together with standard labor market information such as wages, educational attainment, labor market experience, and type of job, making it a highly suitable data source for our purpose.

Overall, the results are consistent with the idea that gender gaps in skills matter to understand gender gaps in labor market outcomes even when comparing individual with similar age and education levels. We show that gender gaps in numeracy cognitive skills for adults exist and play a role in explaining the gender gaps in labor market performance. The gender gap in cognitive numeracy skills is 3.5%-4%. Furthermore, it undergoes a substantial increase from 2.3% to 4.3% from age at entry into the labor market (24-29) up to age 30-39. After that, it decreases slightly. With regard to gender gaps in labor force participation, we find that differences in math skills explain the full gender gap observed (though it is small in absolute terms) for non-parents and at entry age (24-29). For non-parents in the next age group math skills explain 29% of the gender gap observed in participation, which is also a substantial figure. For parents, differences in math skills explain 22% of the total gap for young individuals, and for parents aged 30-39 they account for 18% of the total gender gap observed. In conclusion, gender differences in math skills are essential to understanding a very substantial fraction of the differences observed in labor market participation. Turning to gender gaps in hourly wages, we find that on average math skills explain 40% of the gender wage gap observed. Additionally, their impact increases from entry age to ages 30-39 (motherhood age): We find that math skills explain 44% of the gender wage gap observed for young workers and 55% for workers aged 30-39.

Our findings recommend a gender perspective not only when designing labor market policies but also in the design of other policies in area such as education. Given these results, the conclusion reached in this study is that to reduce gender gaps in labor market performance, in particular in labor force participation and in wages, it is essential to implement measures that make up the disadvantage of women compared to men in numerical skills. That advantage already exists at the age of 15, so to do these measures need to be implemented at school age. In light of this evidence, policy makers should move beyond a narrow focus on boosting female school enrolment and concentrate as well on the contents and quality of teaching. Special efforts should be made to get girls more interested in mathematics and science, and boys more interested in reading.

References

Arora, A. & Pawlowski, E. (2017), “Examining Gender Differences in the Mathematical Literacy of 15-Year_olds and the numeracy skills of the Age cohorts as adults”, Commissioned Paper OECD, Washington, DC.

Autor, D. H. and M. J. Handel (2013). “Putting tasks to the test: Human capital, job tasks, and wages”. *Journal of Labor Economics* 31 (S1), 59-96.

Blau, F., Briton, M. and Grusky, D. (2008). *“The declining Significance of Gender?”* Ed. Francine D. Blau, Mary C. Brinton, David Grusky

Blau, F. D. and L. M. Kahn (2017). “The gender wage gap: Extent, trends, and explanations.” *Journal of Economic Literature* 2017, 55(3), 789–865 https://doi.org/10.1257/jel.20160995

Boll, C., J. Leppin, A. Rossen, and A. Wolf (2016): “Magnitude and impact factors of the gender pay gap in EU countries” *Report prepared for and financed by the European Commission-Directorate-General for Justice, European Union (ed.).*

Christl, M. and Köppl-Turyna, M. (2017): “Gender wage gap and the role of skills: evidence from PIAAC dataset”, GLO Discussion Paper, No. 63

De la Rica, S. and Rebollo-Sanz, Y.F. (2018) "Brechas de género en competencias cognitivas: evidencia internacional" (2018). *Cuadernos Economicos* ICE Nº95 2018

Desjardins, R. & A. Warnke (2012), "Ageing and Skills: A Review and Analysis of Skill Gain and Skill Loss Over the Lifespan and Over Time", *OECD Education Working Papers*, No. 72, OECD Publishing, Paris. http://dx.doi.org/10.1787/5k9csvw87ckh-en

England, P. ; Gornick, J.; Shafer, E.F. (2012). “Women´s Employment, Education and the Gender Gap in 17 countries”, *Monthly Labor Review*.

Gallen, Y., Lesner, R.V, Vejlin, R. (2019) “The labor market gender gap in Denmark: sorting out the past 30 years” *Labour Economics* January Vol 56, pp. 58-67

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

Green D. and C. Riddell (2013) “Ageing and Literacy Skills: A Comparison between Canada, Norway and the United States” *Labour Economics* June Vol 22, pp. 16-29

Hampf, F., Wiederhold, S., & Woessmann, L. (2017). Skills, earnings, and employment: exploring causality in the estimation of returns to skills. *Large-scale Assessments in Education*, 5, 12.

Hanushek, E. A., G. Schwerdt, S. Wiederhold, & L. Woessmann (2015), “Returns to Skills Around the World: Evidence from PIAAC”, *European Economic Review, 73, pp. 103-130*

Hanushek, E. A., & Woessmann, L. (2015). *The knowledge capital of nations*. Cambridge, MA: MIT Press.

Hertzog, C., et al. (2009): “Enrichment effects on adult cognitive development”, *Psychological Science in the Public Interest*, vol. 9, no. 1, pp. 1-65.

Jimeno, J.F., Lacuesta, A., Martínez-Matute and, Villanueva, E. (2016). “Education, Labour Market Experience and Cognitive Skills: Evidence from PIAAC”, Documentos de Trabajo, n.º 1635, Banco de España.

Joensen, J.S & Nielsen, H.S (2014) “Mathematics and Gender: Heterogeneity in Causes and Consequences” *The Economic Journal* Vol 126, Iss 593, pp 1129-1163

Kleven, H., Landais, C. and Søgaard, J.E. (2018) “Children and Gender Inequality: Evidence from Denmark”, NBER Working Paper No. 24219 Issued in January 2018

Lindemann, D. J. (2015) “Gender and Numeracy Skill Use: Cross-National Revelations from PIAAC”. Retrieved [insert date], from [insert website].

Meurs, D., A. Pailh_e, and S. Ponthieux (2010). “Child-related career interruptions and the gender wage gap in France” *Annals of Economics and Statistics*, 15-46.

Ñopo, H. (2008), “Matching as a tool to decompose wage gaps”, *Review of Economics and Statistics*, Vol. 90 No. 2, pp. 290-9.

Olivetti C. & Petrongolo, B. (2008) "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps" *Journal of Labor Economics*, University of Chicago Press, vol. 26(4), pages 621-654, October.

Oreopoulos, P. and Salvanes, K.G. (2011) “Priceless: The nonpecuniary benefits of schooling” *Journal of Economics Perspectives,* Vol25, Nº1, Winter 2011, pp. 159-184

OECD. (2012). *Better skills, better jobs, better lives: A strategic approach to skills policies*. Paris: OECD Publishing. https ://doi.org/10.1787/97892 64177 338-en.

OECD (2013): *OECD skills outlook 2013: First results from the survey of adult skills*, Paris: OECD publishing

Paul, G. (2008) “Children and Women´s Hours of Work” *The Economic Journal* Vol 118, Iss 526

Perry, A., S. Wiederhold, & D. Ackermann-Piek (2014): How can skill mismatch be measured? new approaches with PIAAC. Methods, Data, Analyses: *A Journal for Quantitative Methods and Survey Methodology* 8 (2), 137-174

Main Tables and Figures

Table 1: Gender Gaps in Cognitive Numeracy and Literacy Skills (Overall and By Age)

All Individuals and Workers

			Women		Men		Gender Gap (%)
			mean	sd	mean	sd	
Full Sample	Numeracy	Aged 24-29	272	47	282	46	**-3.71%**
		Aged 30-39	272	44	282	47	**-3.74%**
		Aged 40-49	261	46	273	50	**-4.52%**
		Overall	**267**	**46**	**278**	**50**	**-4.11%**
	Literacy	Aged 24-29	284	42	285	42	**-0.25%**
		Aged 30-39	283	42	284	44	**-0.44%**
		Aged 40-49	272	43	274	46	**-0.60%**
		Overall	279	43	280	44	**-0.53%**
Workers	Numeracy	Aged 24-29	278	44	284	45	**-2.16%**
		Aged 30-39	276	42	285	45	**-3.14%**
		Aged 40-49	266	44	277	**48**	**-4.30%**
		Overall	**272**	**44**	**282**	46	**-3.51%**
	Literacy	Aged 24-29	290	39	287	42	**0.95%**
		Aged 30-39	286	40	286	44	**-0.18%**
		Aged 40-49	275	41	277	50	**-0.80%**
		Overall	**282**	**41**	**283**	**50**	**-0.30%**

Note: PIACC Individual Sample *weights are considered. These math and literacy scores are the mean of the corresponding ten plausible values.*

Table 2: Gender Gaps in Main Labor Market Outcomes

	Women	Men	Gender Gap
Labor Market Participation (% Working)			GG (Women-Men)
All	77.3%	89.4%	-12 pp.
24-29	75.4%	86.0%	-10 pp.
30-39	76.2%	91.0%	-14 pp.
40-49	80.5%	91.1%	-10 pp.
Wages (hourly wages, logs)			Average GGW
All	14.8	17.1	-12.9%
24-29	13.1	14.3	-9.7%
30-39	15.7	17.5	-10.6%
40-49	15.7	19.3	-18.4%

Note: Labor Market Participation refers to the number of individuals observed employed as a proportion of the total labor force. Full time students, Retired and disabled individuals are omitted from the analysis. Gender gaps in Labor Market Participation are measured in percentage points (pp.) whereas gender gaps in wages are average gender gaps in percentages (%).

Table 3A: Labor Market Participation Equation: Gender Gaps in Labor Market Participation and Cognitive Skills (Employment Probability estimated using discrete choice models)

	Full Sample	By Age Cohort 24-49		
	24-49	**24-29**	**30-39**	**40-49**
	(1)	**(2)**	**(3)**	**(4)**
Model 1: Baseline Model				
Math Scores	-	-	-	-
Women	-0.179***	-0.100***	-0.218***	-0.178***
	(0.01)	(0.01)	(0.01)	(0.01)
Model 2: +observed math scores				
Math Scores	0.010***	0.011***	0.011***	0.008***
	(0.00)	(0.00)	(0.00)	(0.00)
Women	-0.166***	-0.0858***	-0.204***	-0.168***
	(0.01)	(0.01)	(0.01)	(0.01)
Model 3: + IV for math scores				
Math Scores	0.019***	0.032***	0.032***	-0.010
	(0.00)	(0.01)	(0.01)	(0.01)
Women	-0.152***	-0.055***	-0.168***	-0.190***
	(0.01)	(0.02)	(0.02)	(0.01)
Sample	63,829	15,995	23,603	24,231

Note: Dependent Variable: Binary indicator of whether the individual is employed (=1) and 0 otherwise. Estimation takes into account sample PIACC weights and PIACC replication weights. Numeracy scores are divided by 10. Marginal effects are displayed with their robust standard errors in brackets. All specifications include country fixed effects and other individual characteristics such as age, educational attainment level, health status, and households characteristics such as the presence of children and partner's attachment to the labor market. The instrumental variables used for model 3 are the educational attainment level of parents (mothers and fathers separately), native-born-speaker and first or second generation immigrants (native-born children of Immigrants). * p<0.10, ** p<0.05, *** p<0.01

Table 3B: Labor Market Participation Equation: Gender Gaps in Labor Market Participation (Separately for parents and non-parents)

	Full Sample	By Age Cohort 24-49		
	24-49	**24-29**	**30-39**	**40-49**
	(1)	**(2)**	**(3)**	**(4)**
Model 1: Baseline Model				
Math Scores	-	-	-	-
Men* No Children	-0.062***	-0.015	-0.076***	-0.082***
	(0.01)	(0.03)	(0.02)	(0.02)
Women* Children	-0.260***	-0.303***	-0.306***	-0.206***
	(0.01)	(0.02)	(0.01)	(0.01)
Women* No Children	-0.126***	-0.0433*	-0.168***	-0.160***
	(0.01)	(0.03)	(0.02)	(0.02)
Model 2: +math scores				
Math Scores	0.009***	0.010***	0.011***	0.008***
	(0.00)	(0.00)	(0.00)	(0.00)
Men* No Child.	-0.062***	-0.019	-0.075***	-0.078***
	(0.01)	(0.03)	(0.02)	(0.02)
Women* Children	-0.246***	-0.284***	-0.290***	-0.196***
	(0.01)	(0.02)	(0.01)	(0.01)
Women* No Children	-0.116***	-0.0355	-0.159***	-0.150***
	(0.01)	(0.03)	(0.02)	(0.02)
Model 3: + (IV) math scores				
Math Scores	0.018**	0.030**	0.029**	-0.001
	(0.00)	(0.01)	(0.01)	(0.01)
Men* No Child.	-0.064***	-0.017	-0.076***	-0.086***
	(0.01)	(0.03)	(0.02)	(0.02)
Women* Children	-0.249***	-0.284***	-0.291***	-0.200***
	(0.01)	(0.03)	(0.02)	(0.01)
Women* No Children	-0.119***	-0.038	-0.159***	-0.153***
	(0.01)	(0.03)	(0.02)	(0.02)
Sample	63,829	15,995	23,603	24,231

Note: Estimation takes into account sample PIACC weights and PIACC replication weights. Dependent Variable: Binary indicator of whether the individual is employed (=1) and 0 otherwise. Numeracy scores (observed and predicted) are divided by 10. Marginal effects are displayed and their robust standard errors in brackets are on line 2. Marginal effects and predicted probabilities shown represent the mean of individual effects. For interactions of gender with children, the omitted option is "Men with no children". Detailed estimation results for Model 3 are shown in Appendix Tables A.3 and A.4. All specifications include country fixed effects and other individual characteristics such as age, educational attainment level, health status, and partner´s attachment to the labor market. The instrumental variables used for model 3 are the educational attainments level of parents (mothers and fathers separately), native-born-speaker, first or second generation immigrants. * p=0.10, ** p=0.05, *** p=0.01

Figure 3: Adjusted Gender Gaps in Labor Force Participation (Women-Men) by Age & Family Type.

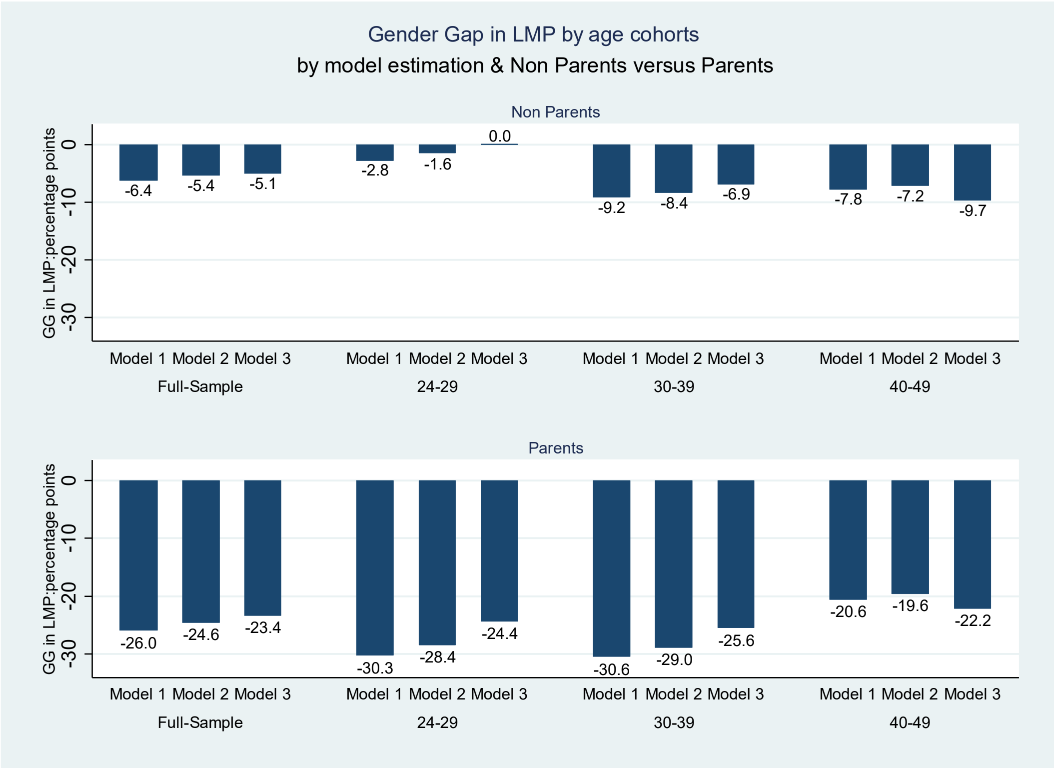

Table 4A: Wage Equation (log hourly wages): Gender Gaps in Wages - Main Results

	Full Sample	Age Cohort 24-49		
	24-49	24-29	30-39	40-49
	(1)	(2)	(3)	(4)
Model 1: Baseline Model				
Math Scores	-	-	-	-
Women	-0.199***	-0.104***	-0.186***	-0.246***
	(0.01)	(0.02)	(0.01)	(0.01)
Model 2: + math scores				
Math Scores	0.024***	0.020***	0.022***	0.028***
	(0.00)	(0.00)	(0.00)	(0.00)
Women	-0.172***	-0.084***	-0.162***	-0.214***
	(0.01)	(0.02)	(0.01)	(0.01)
Model 3: + math scores (IV)				
Math Scores	0.074***	0.048***	0.093***	0.075***
	(0.01)	(0.01)	(0.02)	(0.01)
Women	-0.117***	-0.058***	-0.082***	-0.161***
	(0.01)	(0.02)	(0.03)	(0.02)
Model 4: + math scores (IV) & Labor				
Math Scores	0.057***	0.036***	0.073***	0.059***
	(0.01)	(0.01)	(0.02)	(0.01)
Women	-0.116***	-0.055**	-0.086***	-0.154***
	(0.01)	(0.02)	(0.03)	(0.02)
Sample	41,716	10,579	15,445	15,692

Note: Least squares regressions weighted by sampling weights and replication weights are used to compute robust standard errors (in brackets). Dependent Variable: log gross hourly wage. Sample: employees aged 24-49. Numeracy scores are divided by 10. All specifications include country fixed effects and other individual characteristics such as age, educational attainment level, and labor market experience. Labor characteristics include tenure (years), occupation (nine groups), contract type (fixed-term contract), and firm size (three dummies). * p<0.10, ** p<0.05, *** p<0.01

Table 4B: Wage Equation (log hourly wages): Gender Gaps in Wages & Family: Main Results

	Full Sample	Age Cohort 24-49		
	24-49	**24-29**	**30-39**	**40-49**
	(1)	(2)	(3)	(4)
Model 1: Baseline Model (OLS)				
Math Scores	-	-	-	-
Men* No Children	-0.113***	-0.0521*	-0.0886***	-0.114***
	(0.01)	(0.03)	(0.02)	(0.03)
Women* No Children	-0.240***	-0.132***	-0.219***	-0.274***
	(0.01)	(0.03)	(0.02)	(0.03)
Women* Children	-0.257***	-0.194***	-0.230***	-0.270***
	(0.01)	(0.03)	(0.02)	(0.01)
Model 2: + math scores				
Math Scores	0.024***	0.019***	0.0214***	0.028***
	(0.00)	(0.00)	(0.00)	(0.00)
Men* No Children	-0.104***	-0.058**	-0.079***	-0.095***
	(0.01)	(0.03)	(0.02)	(0.02)
Women* No Children	-0.211***	-0.121***	-0.195***	-0.234***
	(0.01)	(0.03)	(0.02)	(0.02)
Women* Children	-0.224***	-0.168***	-0.200***	-0.236***
	(0.01)	(0.03)	(0.02)	(0.01)
Model 3: + math scores (IV)				
Math Scores	0.073***	0.047***	0.092***	0.074***
	(0.01)	(0.01)	(0.02)	(0.01)
Men* No Child.	-0.0861***	-0.067**	-0.047**	-0.064**
	(0.01)	(0.03)	(0.02)	(0.03)
Women* No Children	-0.152***	-0.106***	-0.114***	-0.168***
	(0.02)	(0.03)	(0.03)	(0.03)
Women* Children	-0.157***	-0.133***	-0.099***	-0.178***
	(0.02)	(0.03)	(0.03)	(0.02)
Model 4: + math scores & Labor (IV)				
Math Scores	0.056***	0.036***	0.072***	0.058***
	(0.01)	(0.01)	(0.02)	(0.01)
Men* No Children	-0.078***	-0.054**	-0.049**	-0.055**
	(0.01)	(0.03)	(0.02)	(0.02)
Women* No Children	-0.141***	-0.092***	-0.099***	-0.161***
	(0.02)	(0.03)	(0.03)	(0.03)
Women* Children	-0.158***	-0.129***	-0.115***	-0.169***
	(0.02)	(0.03)	(0.03)	(0.02)
Sample	41,716	10,579	15,445	15,692

Note: Least squares regressions weighted by sampling weights and replication weights are used to compute robust standard errors (in brackets). Dependent Variable: log gross hourly wage. Sample: employees aged 24-49. Numeracy scores are divided by 10. The constant term refers to "Men with children". Detailed estimation results for Model 3 are shown in Appendix Tables A.5 and A.6. All specifications include country fixed effects and other individual characteristics such as age, educational attainment level, and labor market experience. Labor characteristics include tenure (years), occupation (nine groups), contract type (fixed-term contract), and firm size (three dummies). * p<0.10, ** p<0.05, *** p<0.01

Figure 4: Adjusted Gender Gaps in Wages (Women-Men) by Age & Family Type

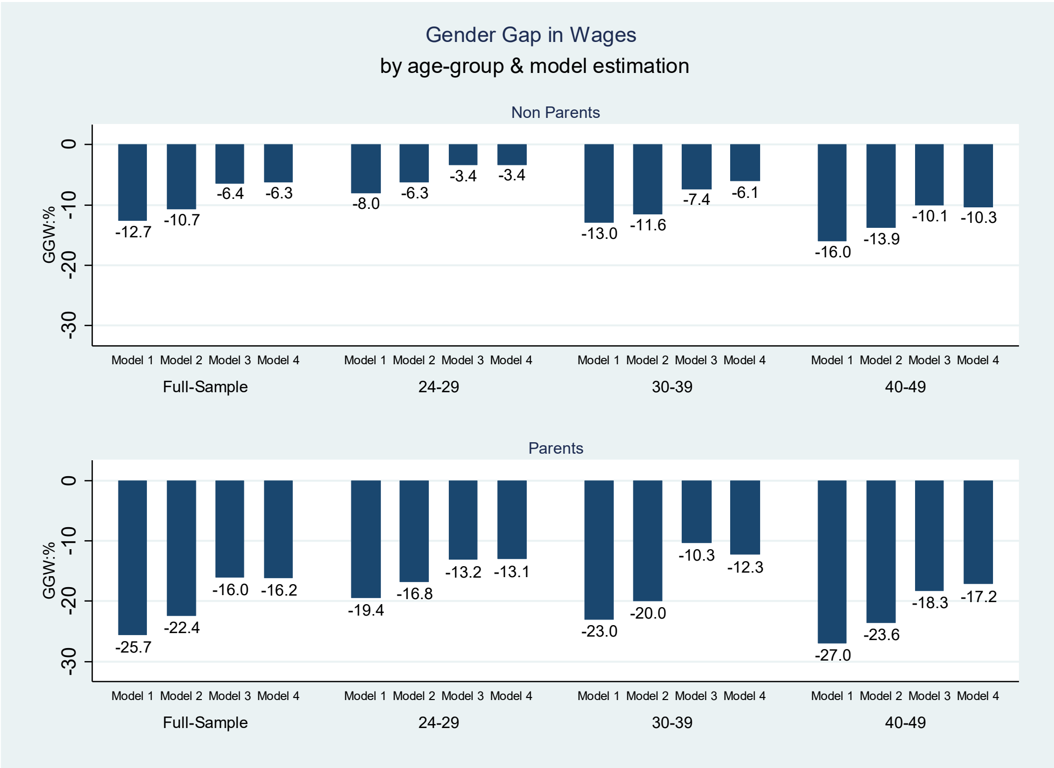

Table 5 : The Wage Equation and The Heckman Selection Model: Gender Gaps in Wages - Main Results.

	Full Sample	By Age Cohort 24-49		
	24-49	**24-29**	**30-39**	**40-49**
	(1)	**(2)**	**(3)**	**(4)**
Model 1: Basic Human Capital				
Math Scores	-	-	-	-
Women	-0.214***	-0.079***	-0.157***	-0.203***
	(0.02)	(0.02)	(0.03)	(0.03)
Inverse Mills Ratio	0.0528	-0.136**	-0.0839	-0.171
	(0.07)	(0.07)	(0.10)	(0.12)
Model 2: + math scores				
Math Scores	0.024***	0.020***	0.022***	0.028***
	(0.00)	(0.00)	(0.00)	(0.00)
Women	-0.180***	-0.067***	-0.122***	-0.166***
	(0.02)	(0.02)	(0.03)	(0.03)
Inverse Mills Ratio	0.0285	-0.0990	-0.116	-0.190
	(0.07)	(0.07)	(0.10)	(0.12)
Model 3: + math scores (IV)				
Math Scores	0.0735***	0.044***	0.086***	0.076***
	(0.01)	(0.01)	(0.02)	(0.01)
Women	-0.118***	-0.0524**	-0.0483	-0.135***
	(0.02)	(0.02)	(0.04)	(0.04)
Inverse Mills Ratio	0.00551	-0.0467	-0.122	-0.102
	(0.07)	(0.08)	(0.11)	(0.12)
Sample	41,716	10,579	15,445	15,692

Note: Least squares regressions weighted by sampling weights and replication weights are used to compute standard errors. Dependent Variable: log gross hourly wage. Sample: employees aged 24-49. Numeracy scores are divided by 10. Robust standard errors are in brackets. All specifications include country fixed effects and other individual characteristics such as age, educational attainment level, and labor market experience. Data Source: PIACC. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Appendix

Table A.1: Main Sample Characteristics used in LMP equation (24-49)

	Women mean	Men mean
Having at least a Child	56.1%	66.1%
Age	41	41
Individual education: Less than secondary	6.2%	6.9%
Individual education: Secondary	62.4%	59.5%
Individual education: Tertiary (not university)	9.7%	14.4%
Individual education: University	21.7%	19.2%
Labor Market Experience (years)	18.77	15.37
Partner in work	35.6%	48.4%
Partner not in work	3.3%	8.0%
Poor Health	22.8%	26.4%

Table A.2 Main Sample Characteristics used in the Wage Equation

	Men	Women
Having at least one child	62%	66%
Age	41	41
Individual education: Primary	4%	4%
Individual education: Secondary	59%	54%
Individual ecucation: Tertiary (not university)	11%	17%
Individual education: University	26%	25%
Labor Market Experience (years)	18.9	17.4
Occupations		
Legislators, senior officials and managers	7%	10%
Professionals	9%	4%
Technicians and associated professionals	14%	19%
Clerks	15%	14%
Service workers and shop and market sellers	7%	17%
Skilled agricultural and fishery workers	13%	28%
Craft and related trades workers	4%	2%
Plant and machine operators and assemblers	18%	3%
Elementary occupations	12%	3%
Large Firm	17%	14%
Part-time	5%	17%
Private Firm	85%	75%
Fixed-term Contract	38%	39%

Note: Occupational classification of respondent's job at 1-digit level (ISCO, 2008)

Table A.3 : Labor Market Participation: Detailed Estimation Results from Preferred Model 3 (Instrumental Variable Estimation):

Sample Estimation	24-49	24-29	30-39	40-49
Men*Non Parent	-0.277***	-0.149	-0.311***	-0.373***
	(0.04)	(0.11)	(0.07)	(0.08)
Women*Non Parent	-0.501***	-0.149	-0.612***	-0.801***
	(0.05)	(0.11)	(0.10)	(0.08)
Women*Parent	-1.022***	-1.007***	-1.119***	-0.975***
	(0.05)	(0.13)	(0.11)	(0.04)
Partner Works (=1)	0.0467	-0.139	0.0846	0.0877
	(0.04)	(0.09)	(0.06)	(0.05)
Poor health (=1)	-0.116***	-0.148*	-0.0669	-0.172***
	(0.03)	(0.08)	(0.05)	(0.05)
Has a partner (=1)	-0.0951**	0.123	-0.145**	-0.0816
	(0.04)	(0.09)	(0.07)	(0.07)
Age	0.0223***	0.0158	0.0234***	0.0129*
	(0.00)	(0.01)	(0.01)	(0.01)
Individual's Education: Less than secondary	-0.437**	0.250	-0.0396	-1.515***
	(0.19)	(0.24)	(0.33)	(0.26)
Individual's Education : Lower secondary	-0.0779	0.230**	0.108	-0.634***
	(0.09)	(0.11)	(0.16)	(0.14)
Individual's Education : Upper secondary	-0.0599	0.322***	0.00980	-0.389***
	(0.06)	(0.09)	(0.10)	(0.09)
Math Skills	0.0755***	0.126***	0.123***	-0.0509
	(0.02)	(0.02)	(0.03)	(0.04)
Constant	-0.650	-2.535***	-1.972*	3.611***
	(0.71)	(0.84)	(1.18)	(1.28)
Observations	60,675	15,162	22,492	23,021

Note: These detailed results correspond to Model 3 in Table 3B (Instrumental Variable Estimation for the Probability of Labor Market Participation). All regressions include country fixed effects. The constant term comprises Men with children with no partner with university studies and in good health. Estimation takes into account PIACC sample weights and PIACC replication weights. Dependent Variable: Binary indicator of whether the individual is employed (=1) and 0 otherwise. Numeracy scores (observed and predicted) are divided by 10.

Table A.4: Labor Market Participation: Detailed Estimation Results from Preferred Model 3 (First Stage of the Instrumental Variable Estimation)

Sample Estimation	24-49	24-29	30-39	40-49
Men*Non Parent	-0.0707	0.403*	-0.156	-0.309*
	(0.09)	(0.23)	(0.15)	(0.17)
Women*Non Parent	-1.096***	-0.685***	-0.975***	-1.329***
	(0.10)	(0.24)	(0.15)	(0.18)
Women*Parent	-1.456***	-1.587***	-1.484***	-1.386***
	(0.07)	(0.24)	(0.11)	(0.09)
Partner Works (=1)	0.301***	0.496**	0.329**	0.196
	(0.09)	(0.22)	(0.13)	(0.13)
Good-health (=1)	-0.435***	-0.522***	-0.337***	-0.435***
	(0.08)	(0.17)	(0.13)	(0.11)
Has a partner (=1)	0.459***	0.225	0.341**	0.585***
	(0.10)	(0.22)	(0.15)	(0.16)
Age				-
	-0.0116***	0.163***	0.0164	0.0536***
	(0.00)	(0.03)	(0.01)	(0.01)
Individual's Education : Less than secondary	-7.448***	-6.005***	-7.619***	-7.564***
	(0.17)	(0.48)	(0.29)	(0.23)
Individual's Education: Lower secondary	-3.527***	-2.997***	-3.665***	-3.529***
	(0.06)	(0.13)	(0.10)	(0.10)
Individual's Education: Upper secondary	-1.846***	-1.787***	-1.979***	-1.733***
	(0.08)	(0.18)	(0.12)	(0.13)
Instruments:				
Parent's Education (Mother): Less than Secondary	-0.833***	-1.640***	-0.906***	-0.186
	(0.10)	(0.19)	(0.16)	(0.19)
Parent's Education (Mother): Secondary	-0.354***	-0.682***	-0.489***	0.0977
	(0.09)	(0.16)	(0.14)	(0.19)
Parent's Education (Father): Less than Secondary	-1.017***	-1.192***	-0.632***	-1.440***
	(0.10)	(0.20)	(0.14)	(0.15)
Parent's Education (Father): Secondary	-0.438***	-0.558***	-0.201	-0.731***
	(0.09)	(0.15)	(0.13)	(0.15)
Native Speaker	0.428**	0.842**	-0.0438	0.581**
	(0.17)	(0.38)	(0.26)	(0.25)
First Generation Immigrant	-0.110	-0.760	0.295	-0.168
	(0.43)	(0.77)	(0.77)	(0.57)
Second Generation Immigrant	1.177***	0.673	1.815**	0.621
	(0.40)	(0.67)	(0.73)	(0.51)
Constant Term	32.04***	27.33***	31.08***	33.99***
	(0.41)	(1.01)	(0.91)	(0.82)
Endogeneity Test (Hausman)	χ^2=132.35	χ^2=112.66	χ^2=93.52	χ^2=50.91
	(p=0.00)	(p=0.00)	(p=0.00)	(p=0.02)

Note: These detailed results correspond to the first stage estimation for the Instrumental Variable Estimation for the Probability of Labor Market Participation presented in Model 3, Table 3B. All regressions include country fixed effects. The constant term contains Men with children with no partner with university studies and in good health

Table A.5: Wage Equation: Detailed Estimation Results from Preferred Model 3 (Instrumental Variable Estimation): (Model 3 Table 4B, IV)

Sample Estimation	24-49	24-29	30-39	40-49
Math Skills	0.0735***	0.0477***	0.0926***	0.0744***
	(0.01)	(0.01)	(0.02)	(0.01)
Men*Non Parent	-0.0861***	-0.0670**	-0.0470**	-0.0643**
	(0.01)	(0.03)	(0.02)	(0.03)
Women*Non Parent	-0.152***	-0.106***	-0.114***	-0.168***
	(0.02)	(0.03)	(0.03)	(0.03)
Women*Parent	-0.157***	-0.133***	-0.0995***	-0.178***
	(0.02)	(0.03)	(0.03)	(0.02)
Age	0.00106	0.0150***	0.000382	-0.00295
	(0.00)	(0.01)	(0.00)	(0.00)
Individual's Education: Less than secondary	-0.0814	-0.0348	0.103	-0.198*
	(0.07)	(0.09)	(0.14)	(0.12)
Individual's Education: Lower secondary	-0.166***	-0.102***	-0.0947	-0.256***
	(0.03)	(0.04)	(0.07)	(0.06)
Individual's Education: Upper secondary	-0.127***	-0.0579*	-0.0694*	-0.232***
	(0.02)	(0.03)	(0.04)	(0.03)
Labor Market Experience	0.0247***	0.0212***	0.0234***	0.00898
	(0.00)	(0.01)	(0.01)	(0.01)
Labor Market Experience^2	-0.0027***	-0.00781**	-0.00224	0.00151
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.708***	1.085***	0.0767	1.076**
	(0.26)	(0.28)	(0.52)	(0.44)
Observations	39,852	10,578	15,445	15,692
R-squared	0.312	0.255	0.182	0.378

Note: These detailed estimation results correspond to Model 3 in Table 4B. Least squares regressions weighted by sampling weights and replication weights are used to compute robust standard errors (in brackets). Dependent Variable: log gross hourly wage. Sample: employees aged 24-49. Numeracy scores are divided by 10. The constant term refers to "Men with children". All specifications include country fixed effects. * p<0.10, ** p<0.05, *** p<0.01

Table A.6 : Wage Equation: Detailed Estimation Results – First Stage Regression : Gender gaps in Wages (IV)

Sample	24-49	24-29	30-39	40-49
Men*Non Parent	-0.384***	0.225	-0.463***	-0.669***
	(0.10)	(0.25)	(0.15)	(0.20)
Women*Non Parent	-1.219***	-0.644**	-1.118***	-1.455***
	(0.10)	(0.26)	(0.16)	(0.20)
Women*Parent	-1.326***	-1.258***	-1.377***	-1.230***
	(0.08)	(0.29)	(0.12)	(0.12)
Age	-0.0387***	0.199***	0.0130	-0.0489**
	(0.01)	(0.04)	(0.02)	(0.02)
Individual's Education: Less than secondary	-7.684***	-6.372***	-7.837***	-7.554***
	(0.29)	(0.93)	(0.45)	(0.37)
Individual's Education: Lower secondary	-3.890***	-3.215***	-3.806***	-3.928***
	(0.08)	(0.17)	(0.12)	(0.13)
Individual's Education: Upper secondary	-2.091***	-2.010***	-2.184***	-1.887***
	(0.09)	(0.21)	(0.14)	(0.15)
Labor Market Experience	0.136***	0.0977	0.161***	0.236***
	(0.02)	(0.06)	(0.04)	(0.04)
Labor Market Experience^2	-0.0347***	-0.0886***	-0.0577***	-0.0540***
	(0.00)	(0.03)	(0.02)	(0.01)
Instruments:				
Parent's Education (Mother): Less than secondary	-0.400***	-0.987***	-0.555***	0.217
	(0.12)	(0.25)	(0.19)	(0.22)
Parent's Education (Mother): Secondary	-0.0631	-0.396**	-0.193	0.417*
	(0.11)	(0.19)	(0.17)	(0.21)
Parent's Education (Father): Less than secondary	-0.798***	-0.922***	-0.651***	-0.913***
	(0.11)	(0.25)	(0.17)	(0.18)
Parent's Education (Father): Secondary	-0.289***	-0.378**	-0.254*	-0.319*
	(0.10)	(0.19)	(0.15)	(0.18)
Native Speaker	0.294	1.076**	-0.0466	0.282
	(0.23)	(0.50)	(0.36)	(0.36)
First Generation Immigrant	-0.0515	-1.059	0.432	-0.0206
	(0.49)	(1.13)	(0.72)	(0.71)
Second Generation Immigrant	1.149**	0.369	1.559**	0.970
	(0.45)	(1.00)	(0.67)	(0.66)
Constant	32.47***	26.29***	30.87***	31.26***
	(0.48)	(1.39)	(0.94)	(1.14)
Endogeneity Test (Wu-Hausman)	F=168.53	F=38.15	F=63.44	F=38.65
	p=0.00	p=0.00	p=0.00	p=0.00
R-squared	0.346	0.324	0.333	0.373

Note: These detailed estimation results correspond to the first stage results from model 3 –instrumental variable model estimation- in Table 4B. Least squares regressions weighted by sampling weights and replication weights are used to compute robust standard errors (in brackets). Dependent Variable: log gross hourly wage. Sample: employees aged 24-49. Numeracy scores are divided by 10. The constant term refers to "Men with children". All specifications include country fixed effects.
* p<0.10, ** p<0.05, *** p<0.01