Direction des Statistiques Démographiques et Sociales

N° F1508

Worker-firm matching and the family pay gap : Evidence from linked employer-employee data

Lionel WILNER

DOCUMENT DE TRAVAIL

Institut National de la Statistique et des Etudes Economiques

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES
Série des Documents de Travail
de la
DIRECTION DES STATISTIQUES DÉMOGRAPHIQUES ET SOCIALES

N°F1508

Worker-firm matching and the family pay gap : Evidence from linked employer-employee data

LIONEL WILNER

(DIVISION SALAIRES ET REVENUS D'ACTIVITÉ)

Document de travail

novembre 2015

SOMMAIRE

Worker-Firm Matching and the Family Pay Gap: Evidence from Linked Employer-Employee Data*

Lionel Wilner†

November 2015

Abstract

The family pay gap is not fully explained by human capital depreciation and unobserved heterogeneity. Endogenous worker-firm matching could also account for such wage differences. This hypothesis is tested thanks to linked employer-employee data on the French private sector between 1995 and 2011. Distinct hourly wage equations are estimated for women and for men, including firm- and worker- fixed effects on top of usual measures of human capital. Though omitted variable biases due to worker-firm matching explain none of the motherhood wage penalty, they play a role in the case of men who do not experience any wage loss after childbirth, but do not enjoy any premium either. In a counterfactual where women do not incur any penalty after childbirth, the gender gap still amounts to 2/3 of the one that currently prevails.

Keywords: High dimensional fixed effects; worker-firm matching; family pay gap; gender inequalities; linked employer-employee data.

JEL Classification: J12; J13; J16; J31; J62; J71.

*I am grateful to Richard Blundell, Élise Coudin, Francis Kramarz, Fabrice Lenglart, Laurent Linnemer, Thierry Magnac, Sophie Ponthieux, Johannes Spinnewijn, Michael Visser, and Andrea Weber for their insightful comments and especially to Dominique Meurs and Sébastien Roux for their discussions. I also thank audiences at Insee (Paris) seminars and at the European Winter Meeting of the Econometric Society 2014 (Madrid). All errors and opinions are mine.

†CREST (INSEE). 18, boulevard Adolphe Pinard, 75 014 Paris, France. Tel: (+33)141175460. Email: lionel.wilner@ensae.fr

1 Introduction

Not only do gender inequalities persist within households (in terms of the share of domestic work or bargaining power; see e.g. Goldin, 2014; Meurs and Ponthieux, 2014), but they also persist within firms, which labor economists and sociologists have documented for decades. The gender pay gap, occupational gender segregation and the glass ceiling are the most striking examples: they are both unfair and inefficient as long as they are not justified by productivity differentials. Although the gender gap has begun to decline slightly, it has not fallen to zero (OECD, 2012). Composition effects tend to locate women at lower positions within the hierarchy, which contrasts with their higher average level of education–at least in most European Union member states. This discrepancy between education and occupation has been designated as a form of gender segregation. In the same vein, the glass ceiling that prevents women from reaching the top positions of the institutional hierarchy (CEOs, national presidencies, etc.) is difficult to break.

However, the most obvious gender inequality is related to childbirth. The family pay gap, which accounts for hourly wage differentials following childbirth, was first documented by Waldfogel (1997). Numerous papers have since determined the magnitude of motherhood penalties. More recent papers have also assessed the existence of wage differences among men between fathers and non-fathers. Four primary theoretical explanations for post-childbirth hourly wage differentials have been proposed, namely (i) human capital depreciation, (ii) individual unobserved heterogeneity (parents would have specific average productivities), (iii) firm matching (parents would match with specific firms), and (iv) discrimination. However, the relevance of each of these explanations remains to be assessed empirically.

The contribution of this paper consists in testing whether the family pay gap stems from selection effects of parents into firms, that is, endogenous worker-firm matching. To the best of my knowledge, this firm matching explanation has never been investigated. However, the typical approach of estimating Mincer equations might suffer from omitted variable biases due to endogenous sorting. The latter biases are important if individuals who plan to become parents move to family-friendly firms, either before or after childbirth. This is particularly likely if parents trade-off money against job amenities. In that case, the estimated wage differentials would spuriously account for endogenous mobility instead of referring to the causal impact of children.

I resort to panel data that are also linked employer-employee data (LEED), that is, the statistical unit of which is a (individual, firm, year) triplet. LEED have proven exceptionally useful in the estimation of wage equations since Abowd, Kramarz, and Margolis (1999) proposed a two-factor specification of Mincer equations, namely firm and worker fixed effects. I estimate an adjusted family pay gap by employing such a two-factor model with high dimensional firm and worker fixed effects, which permits me to control for as much observed and unobserved heterogeneity as possible. Importantly, this econometric specification provides a solution to the previously mentioned omitted variable biases.

My application is based on the comprehensive DADS panel, which contains exhaustive information on French salaried employees' careers in the private sector from 1995 to 2011. This paper is the first attempt to analyze the family pay gap by resorting to LEED. By disentangling the effect of childbirth from spurious correlation between parenthood and other firm-specific wage determinants, this estimation contrasts with previous analyses, which only control for individual unobserved heterogeneity. A *caveat* due to data limitations is the following: work hours have been available only since 1995, which leads me to select relatively young individuals for whom I know the complete path of hourly wages from 1995 to 2011.

After controlling for full-time and part-time experience, as well as for both firm and worker fixed effects, I still find a difference between non-mothers and mothers after childbirth that amounts to an effect of approximately -3% per child on the hourly wage. In the case of women, I reject the firm matching explanation, as well as the hypothesis of unobserved heterogeneity; a possible human capital depreciation seems to provide a more compelling explanation. These results are consistent with the dynamic long-run effects found by Kleven, Landais, and Søgaard (2015). My results are strikingly different with respect to men. My sample of men does not experience any loss after childbirth but also does not enjoy any premium. Moreover, the matching between workers and firms matters here, while unobserved heterogeneity and human capital depreciation play no role. To explain the absence of fatherhood premia, I provide evidence of an erosion of those premia over time, such that one cannot reject the null hypothesis during the period considered. My results are also consistent with previous findings documenting heterogeneity in those premia; overall they indicate gender inequalities with respect to parenthood. Finally, I propose an evaluation of the contribution of the family pay gap to the gender gap by simulating a counterfactual *scenario* in which women and men experience the same childbirth penalty: the former would explain approximately 1/3 of the latter in my sample of young individuals.

The remainder of the paper is organized as follows. Section 2 is devoted to a literature review. Section 3 presents my matched employer-employee database. I describe my econometric specification in Section 4. Section 5 contains the results, namely the test of the firm matching explanation, as well as a measure of the contribution of the family gap to the gender gap, and robustness checks. Section 6 concludes with some policy recommendations.

2 Literature review

The seminal contributions of Waldfogel (1997, 1998) document the existence of a motherhood wage penalty both in the US and in the UK. Relying on data from 1968 to 1988, she estimates Mincer equations on log hourly wages with individual fixed effects and finds a wage loss of approximately -6% per child. Considering a different period, from 1982 to 1993, Budig and England (2001) obtain a loss of -3% at first childbirth, -9% at second childbirth and -12% at third childbirth. Similar figures are found in other European countries: an exhaustive list is given in Davies and Pierre (2005). In Germany, where the gender pay gap is already high (approximately 22%), the motherhood wage penalty is greater than 10% in absolute terms: Buligescu, De Crombrugghe, Menteşoğlu, and Montizaan (2009) find a family pay gap of between -10% and -14%; Beblo, Bender, and Wolf (2009) report estimates higher than 19% in absolute terms; according to Felfe (2012), this gap is approximately -10.7%. An on-going work by Kleven, Landais, and Søgaard (2015) is devoted to the Denmark; it finds dynamic, long-run effects of childbirth on mothers' wages.

The French case has been investigated by Meurs, Pailhé, and Ponthieux (2010). They focus on the effect of child-related time out of the labor market on the gender pay gap (previously documented in Meurs and Ponthieux, 2000). Lequien (2012) and Joseph, Pailhé, Recotillet, and Solaz (2013) specifically analyze the impact of parental leave by exploiting two reforms that enable them to recover causal effects of time out of the workforce on wages. The former examines a reform that occurred in 1994 related to a monthly benefit for parents called the *Allocation Parentale d'Education* (APE); it changed the incentives to take parental leave after the birth of a second child by extending allowance eligibility from third-born to second-born children. The latter is devoted to a second reform that created a childhood benefit, the *Prestation d'Accueil du Jeune Enfant* (PAJE), which replaced the APE; the authors exploit a specific feature of this reform, namely, the allocation of a supplementary benefit for part-time activity.

Some papers nevertheless find no wage differential between mothers and non-mothers. Using US data, Korenman and Neumark (1992) do not find any evidence of a family pay gap. However, their analysis relies on first-differences over a short period, from 1980 to 1982, which could explain this dissonance. More convincing is the paper by Simonsen and Skipper (2006) that finds a "net gap" (unadjusted) but no "direct gap", *i.e.*, no effect of childbirth on wages. The authors explain that most of the gap could stem from indirect channels relating motherhood to other covariates, which may cause spurious correlation. For instance, motherhood may be negatively correlated with experience, as child-rearing activities lead mechanically to lower work experience. In the same vein, one expects to find more mothers working in the public sector or in more family-friendly occupations. These aspects must be controlled for when attempting to identify causal effects. Using a propensity-score matching approach, the authors conclude that in Scandinavia, mothers self-select into the public sector (see also Nielsen, Simonsen, and Verner, 2004) and that there is no causal effect of motherhood on wages in the private sector after controlling for selection.

It has been suggested that to recover the causal effect of childbirth on wages, data on twin sisters should be used, which would provide a natural experiment. Comparing wage trajectories of mothers with respect to those of non-mothers, Lundberg and Rose (2000) find a causal wage gap of -5%, a figure close to that obtained by Simonsen and Skipper (2012) following a similar approach. Another method consists in relying on exogenous childbirths, for instance, by exploiting fertility shocks. The introduction of contraceptive methods and the passage of abortion laws provide researchers with quasi-experimental settings that generate the variation required for identification. Miller (2011) uses biological fertility shocks as instruments for age at first birth to recover a causal impact of the timing of childbirths on wages: delaying birth by one year would raise earnings by up to 9/10% and increase work hours by 6%.

A recent strand of research has focused on the heterogeneity of the motherhood wage penalty, according to either the rank in the wage distribution or the level of education. There is a controversy opposing Budig and Hodges (2010) and Wilde, Batchelder, and Ellwood (2010) concerning the link between education and this wage differential: the former find higher motherhood penalties for women with higher cognitive skill levels, while the latter obtain higher penalties at lower wage levels. England, Bearak, Budig, and Hodges (2013) address this issue and attempt to reconcile the two perspectives by introducing other dimensions of heterogeneity such as race in the US.

Another contemporaneous area of research has begun to focus on men and has investigated the issue of a fatherhood wage gap. Contributions include Lundberg and Rose (2000, 2002) and Glauber (2008). The results instead demonstrate the existence of a fatherhood premium, which contrasts with the motherhood penalty and constitutes a gender-based inequality with respect to childbirth. However, Killewald (2013) finds considerable heterogeneity in those premia: certain groups of fathers, including unmarried residential fathers, nonresidential fathers and stepfathers, experience no significant premium.

The motherhood penalty is a puzzling issue, and several theoretical explanations have been proposed to account for that wage differential. Most of the arguments below also apply to men –except the one concerning maternity leave, although in Scandinavia for instance, welfare programs offer generous paternity leaves, which reduces the gender asymmetry in this respect.

First, motherhood implies some human capital depreciation due to mandatory parental leave. This "human capital deterioration" explanation dates back to Becker (1985). Human capital is a composite concept: it aggregates at least education, experience and training. Women who wish to become mothers would rationally opt for lower educational levels. Their career has mechanically more frequent interruptions (sick-child leaves, in addition to maternity leave), which depreciates their work experience. Furthermore, the time spent out of the labor force is likely to have a negative impact on their training, especially if training is some function of continuous employment. Finally, mothers may work part-time, which further diminishes their work experience.[1] Under the assumption of perfectly com-

[1]One could argue that working part-time constitutes a negative signal that individuals send to their employers by reducing voluntarily their activity. However, this explanation does not belong to "human capital" theory but rather to a competing explanation, the "signaling" theory proposed by Spence (1973).

petitive labor markets, lower hourly wages must reflect a lower productivity caused by career interruptions or a lower training or education level. In Sweden, Albrecht, Edin, Sundström, and Vroman (1999) ask whether this hypothesis alone could be responsible for the family pay gap; they find that human capital depreciation is not the sole explanation for the negative effect of career interruptions on subsequent wages.

Second, individual unobserved heterogeneity has been invoked to explain the wage gap between mothers and non-mothers. The former may choose more family-oriented careers; this self-selection being primarily driven by preferences and/or personal abilities. The negative correlation between labor market outcomes and fertility could then reflect stronger preferences for family, domestic activities, or leisure or lower on-the-job productivity. Women endowed with such preferences and capacities would *ex ante* invest less in education and training, hence acquiring less human capital. The family gap could thus reflect a different willingness to work in a competitive environment. Disentangling spurious correlation between childbirths and wages due to preferences from the causal effect of parenthood therefore requires controlling for unobserved heterogeneity, for instance that due to individual fixed effects–assuming that unobserved heterogeneity does not vary over time. Even after controlling for worker fixed effects in panel data, as is the case in most empirical papers cited previously, a substantial share of wage differences between mothers and non-mothers remains unexplained.

A third explanation (hereinafter "firm matching") claims that mothers would be employed in less productive firms. To reconcile family life and careers, women who plan to be mothers would look for jobs that allow them to spend more time in child-rearing activities. For instance, they would favor jobs with flexible hours, on-site day care, jobs in which personal phone calls are authorized during work, or jobs that do not require overtime work, evening work, work during weekends, etc. In that case, in equilibrium, occupational segregation should emerge in the labor market. As a result, forward-looking women who want children would seek more convenient and convenient more family-oriented jobs. In the same vein, mothers may have higher search costs on the labor market, which restricts their mobility and prevents them from looking for better positions while resulting in poor job matches. Surprisingly, this explanation has received little attention thus far. Budig and England (2001) proposes controlling for as many job characteristics as possible, including part-time employment, to neutralize any "family-friendly" job feature. From the substantial literature devoted to job search, we know that there is considerable heterogeneity in the quality of the employer-employee relationship and that mobility offers potentially large wage gains. Nielsen, Simonsen, and Verner (2004) show that mothers tend to self-select into the public sector. Beblo, Bender, and Wolf (2009) argue that the selection into private establishments has to be taken into account because it could represent up to 7pp of the family pay gap in the German case (-19% after controlling for establishment effects in a matching approach against -26% when such effects are ignored). This explanation is also advanced by Felfe (2012), who suggests that mothers are prepared to trade off earnings against amenities and hence proposes a compensating wage differentials (CWD) explanation. Among mothers, she distinguishes those who remain in the same position after maternity leave from other mothers. She finds that the former experience a significantly smaller wage gap (-9.3% against -24.3%) than the latter, which supports this hypothesis. However, part of the difference stems precisely from an adjustment of work conditions, and after controlling for this adjustment, the family pay gap cannot be solely explained by a CWD explanation.

Fourth and finally, a further explanation for the motherhood wage penalty is the possibility of discrimination against mothers at work. Employers could be reluctant to hire mothers-to-be or women who they expect to become mothers, which would affect mothers' employment. Moreover, employers could also be more rigid in the wage bargaining process and offer prospective mothers fewer opportunities to distinguish themselves (through the provision of overtime work, more risky assignments, etc.), which would result in a less frequent attribution of irregular bonuses. Generally speaking, such discrimination could result from labor reallocation, either within firms or within establishments. The strategy adopted in this paper consists in providing indirect evidence in favor of discrimination by testing for the three previous explanations and by excluding them as possibilities. At least the elimination of alternative explanations, combined with asymmetric results for women and men, suggests that gender biases are likely.

3 Data

3.1 Sources

My analysis is based on the merger of two French administrative datasets commonly known as the DADS-EDP panel collected by Insee.

The first source is the DADS panel, a comprehensive database of salaried employees, and the longitudinal version of the cross-sectional *Déclaration Annuelle de Données Sociales* (DADS). It is mandatory for French firms to fill in annually a DADS for every employee subject to payroll taxes, especially for every salaried employee in the private sector or in government-owned firms. From 1967 to 1975, the panel does not contain any information on firms while from 1976 to 2011, the data are available at the individual-firm level. Every firm – more precisely every establishment – has a unique identifier, the SIRET,[2] a 14-digit number, while individuals are identified by their NIR, a social security number with 13 digits. Before 1976 an observation is made up of a unique (individual, year) pair while after 1976 it is composed of a unique (individual, firm, year) triplet, which features the data as linked employer-employee data (LEED). The DADS panel contains information about individuals born on October of even-numbered years –a representative sample of the French salaried population at rate 1/24. From 2002 onwards though, the panel has been completed with individuals born on October of odd-numbered years, which corresponds to a sampling rate of 1/12; however, the longitudinal depth is mechanically shorter for such individuals in comparison with those born on October of even-numbered years. Since filling in the DADS form is mandatory,[3] and because of the comprehensiveness of the DADS panel with respect to individuals' careers, the data are of exceptional quality and have low measurement error in comparison with survey data. Some years are missing (1981, 1983 and 1990) because there was no data collection by Insee during the 1982 and 1990 censuses. In 1994, 2003, 2004 and 2005, the quality is nevertheless questionable. Since overseas appeared in the panel from 2002 onwards, I restrict my attention to metropolitan France. Finally, these data contain detailed information about gross and net wages, work days, work hours, other job characteristics from 1976 onwards

[2]The SIRET is a concatenation of the SIREN, a firm identifier, and of an establishment identifier.

[3]The absence of a DADS as well as incorrect or missing answers are punished by law with fines.

(like the beginning and the end of an employment's spell, seniority, a dummy for part-time employment), firm characteristics (industry, size, region) and individual characteristics (age, gender).

The second source is the *Échantillon Démographique Permanent* (hereinafter EDP). This longitudinal database covers a representative part of the population born on one of the first four days of October. It contains administrative registers of births and marriages from 1968 to 2011, as well as partial information on education[4] from 1968, 1975, 1982, 1990 and 1999 censuses. However, for half of the sample, namely people born on October 2^{nd} or 3^{rd}, birth registers have not been properly filled in from 1983 to 1997. The information is incomplete from 1983 to 1989 and missing from 1990 to 1997. It is possible to recover part of missing data by exploiting 1990 and 1999 censuses, but at some cost (namely measurement error since the information conveyed by both censuses does not coincide perfectly). I choose the most conservative option: I rely on birth registers and keep therefore individuals born on October 1^{st} or 4^{th} only. This subsample is still representative of the French population[5] at rate 2/365.

The two sources can be merged through their common individual identifier, the NIR. I exclude "wrong" NIRs that are present in the DADS panel for cross-sectional use and for statistical reasons only, as I cannot follow the careers of such individuals.[6] I also exclude some (few) self-employed individuals who appeared in the panel from 2009 onwards. Finally, I eliminate observations that correspond to home-work or to unemployment amenities.

A methodological contribution of this paper consists in computing an accurate measure of salaried experience. I exploit the administrative feature of the dataset and derive experience at the individual level from the sequence of observed working times. I count the actual (past and present) number of work hours and express it in full-time units (FTU). In France, a full-time worker used to work 2028 (1820) hours per year before (after) 2002 –the mandatory working time decreased by 4 hours per week after the adoption of the Aubry laws. However, I face data limitations: work hours have been available only since 1995. Hence, I consider individuals who entered the panel after 1995 to avoid biasing my measure of experience because of missing or incomplete sequences of working time. After merging the two datasets

[4]See Footnote 12.

[5]Yet individuals born abroad are missing from the EDP.

[6]For instance, some of them were not born in October.

and imposing this "entry condition", my sample includes 46,280 individuals. I proceed then to further selection that is described more extensively in Appendix A: I focus on individuals aged 16 to 65 working in the private sector and whose annual wage exceeds 10 euros in 2011 terms. I disregard the years 2003 to 2005 in the analysis for the reasons mentioned above. The working sample contains 41,531 individuals, which represents 212,189 observations at the individual-year level and 301,079 observations at the individual-firm-year level. Although the sample is not representative of the entire French salaried population, the method I propose to address the omitted variable bias in the family gap framework applies to any LEED sample.[7]

3.2 Descriptive analysis

The individuals in the sample are mechanically younger: they are aged 26.9 on average. In addition, 48% of them are women, 18.9% are married, and 33.9% have at least one child. Some of them worked continuously in the private sector from 1995 to 2011. The average potential experience amounts to 6.4 years, where potential experience is defined as the difference between the current year and the year the individual first appears in the panel. On average, full-time experience is 2.3 years in FTU, while part-time experience amounts to 0.8 years. Average seniority is approximately 2.5 years, where seniority is the difference between the current year and the year that an individual first appears with his/her the current firm. The annual job duration amounts to 237 days, or 1037 hours, which reflects composition effects explained by both youth insertion in the labor market and part-time activity. The average net hourly wage amounts to slightly less than 10 euros. As I proceed to trim some very low hourly wages (see Appendix A and Section 5.3), the minimum observed hourly wage is 3.51 euros.

Tables 1 and 2 provide a summary description of the working sample for women and men, respectively. In the labor market, women and men differ with respect to wages: women receive, on average, a lower hourly wage (9.43 euros) than men do (10.42 euros), which yields an unadjusted gender pay gap of 9.6% in this sample of young individuals. Women also tend to have lower full-time experience but higher part-time experience. In FTU, 9.7% of men's activity stems from part-time work,

[7] As time passes, the entry condition will become less restrictive in terms of age–hence the selection will be less drastic in future works relying on the same source.

against 24.7% of women's activity –the average volume of part-time activity is 15.8%.

This issue of part-time work cannot be disregarded in an empirical analysis devoted to childbirth, as argued, for instance, by Budig and England (2001). In contrast to studies relying on full-time workers only, part-time workers are not selected out of the sample. However, the literature has not reached consensus on whether part-time work induces a penalty or a premium on the hourly wage. Using Australian data, Booth and Wood (2008) find that the negative coefficient of part-time work in a Mincerian equation on the hourly wage disappears after controlling for covariates (especially experience) and unobserved heterogeneity. I adopt their methodology and reproduce their Table 2. Hence I specify:

$$\log W_{it} = X'_{it}\beta + \alpha P_{it} + \theta_i + \epsilon_{it}, \tag{1}$$

where W_{it} is the log hourly wage of individual i in year t (computed as the ratio between the sum of her wages and the sum of her work hours), X_{it} a set of covariates, P_{it} a dummy for part-time work, θ_i an individual fixed effect and ϵ_{it} an error term. Table 3 reports the estimates of α under different specifications (with or without fixed effects θ_i, with or without covariates X_{it} including worker characteristics, firm characteristics, experience, etc.). In both pooled OLS and fixed effects approaches, the sign of α is negative in the absence of controls, which means that part-time workers earn unconditionally less than full-time workers. However, $\hat{\alpha}$ becomes positive after controlling for observed and unobserved heterogeneity. This empirical result holds both for women and men.[8]

Another empirical issue worth examining is the relationship between family events, including childbirths and marriages, and experience. Once again, it is crucial to careful distinguish between full-time and part-time experience. I therefore estimate:

$$\text{Exp}_{it} = X'_{it}\beta + \theta_i + \epsilon_{it}, \tag{2}$$

where Exp_{it} refers either to full-time experience (measured in FTU) or part-time experience. I do not include part-time work P_{it} as a covariate here because it would be correlated with the determinants of the dependent variable and hence endogenous in (2). Table 4 displays the results, which exhibit interesting gender

[8]In what follows, I will not distinguish P_{it} from the other covariates in X_{it}.

differences. Women tend to accumulate less full-time experience every year than men do (.41 year as opposed to .5) but twice as much part-time experience (.12 versus .06). While women may lose up to three years of full-time experience after the fourth childbirth, men's full-time experience never diminishes with parenthood: it appears as if men worked more to compensate for the arrival of children in the household, especially at the first and second childbirth. Women should incur a loss of experience following births because of mandatory maternity leave (16 weeks,[9] which amounts to approximately -.3 years of FT experience), but they might compensate for this penalty in the long run by working more. Interestingly, after the first birth, mothers tend to work slightly more than non-mothers, on average by .2 years: in the long run, and accounting for the gender asymmetry of -.3 years, this yields the same level as men (+.5 years at first birth). Yet, the situation changes dramatically from the second childbirth onwards: even in the long run, mothers tend to spend less time on the labor market than non-mothers (-2.6 years at fourth childbirth). The substitution with part-time activity does allow the women to compensate for this difference: part-time experience increases little, by +.3 years, after the third childbirth. Hence, labor supply decisions within households following parenthood may still be strongly biased in favor of men pursuing their labor market activity while women reduce theirs. These results confirm the hypothesis of gender biases and suggest that remaining out of the workforce remains an option for women after the third childbirth. Finally, marriage is associated with positive effects on FT experience–higher for men than for women–but has less clear-cut effects on PT experience.

4 Econometric specification

4.1 Worker fixed effects

The literature devoted to the family pay gap has thus far focused on the estimation of Mincerian wage equations using panel data at the individual-year level. The typical dependent variable is such estimates is the logarithm of the hourly wage, which is regarded as a proxy for productivity if labor markets are perfectly competitive. Explanatory variables include age, experience, seniority, job charac-

[9] As regards a first-born or a second-born child; 26 weeks for third-born and other children. Also, the maternity leave is not mandatory *stricto sensu* and a mother may choose not to have her full leave, but she is not allowed to work as an employee during at least 8 weeks.

teristics (sector, firm size, location), potentially other controls, as well as time and individual fixed effects. In that vein, for individual i observed in year t, I specify first:

$$\log W_{it} = X'_{it}\beta + N'_{jt}\delta + V_t + \theta_i + \epsilon_{it}, \tag{3}$$

where W is the ratio between the sum of wages received by i in year t and the sum of his worked hours, X contains part-time activity, marriage and children and quadratic specifications of age, full-time experience, part-time experience and seniority,[10] N includes firm characteristics (size, industry, *département*), and ϵ_{it} is an idiosyncratic error term, the variance of which is allowed to be individual-specific. The variables of interest are Childbirth_{itk}, $\forall k = 1, \ldots, 5$, which are dummy variables indicating whether individual i had already experienced his k-th childbirth at time t. I introduce the covariate Marriage_{it} to capture the effect of marriage on hourly wages. Moreover, I control for numerous job characteristics that correspond to the characteristics of the individual's main employment (see Appendix A for the definition of main employment) including the firm's size, the *département* where the establishment is located and its sector of activity. Size is coded with 12 categories, while industry is defined by the first two digits of the NACE classification and has 39 categories, including a "missing" one. Further, 95 *département* dummies account for the spatial dispersion of earnings in metropolitan France. V_t is a year fixed effect that captures the contemporaneous effects affecting earnings (business cycle, macro shocks, etc.), while θ_i is an individual fixed effect that encompasses permanent unobserved heterogeneity including talent, employability, cohort effects and initial education. The EDP provides me with a schooling variable that indicates the highest degree obtained by an individual. However, in the presence of individual fixed effects, the coefficient of education is identified provided that this variable is time-varying, which is not the case with initial formation. Finally, in the spirit of Waldfogel (1997, 1998) and Budig and England (2001), I compare a fixed effect estimation with a first-difference estimation of model (3). While the former enables me to recover rather long-run effects, the latter accounts for a short-run effect.

[10] These characteristics are attached to an individual's main employment.

4.2 Worker and firm fixed effects

Despite their qualities, previous models suffer from an omitted variable bias caused by the omission of firm fixed effects, which was first emphasized by Abowd, Kramarz, and Margolis (1999). High-wage workers are expected to match with high-wage firms; more generally, the matching process may allocate specific workers to firms with specific compensation schemes. If the job matching results in important selection effects, previous estimations will suffer from endogeneity bias due to a correlation between explanatory variables such as Marriage_{it} or Childbirth_{itk}, $\forall k = 1, \ldots, 5$, and a firm-specific term ψ_j of the error term in (3) that would write as $\psi_j + \epsilon_{ijt}$. Abowd, Kramarz, and Woodcock (2008) explicitly model this omitted variable bias as a function of the covariance between the matrix X and the design matrix of indicator variables for the employer for which individuals work, conditional on the design matrix of individuals' indicator variables. I choose then to exploit the LEED nature of my dataset without aggregating information at the individual-year level. To the best of my knowledge, this paper is the first attempt to address the family pay gap issue at the individual-firm-year level and hence avoids omitted variable bias due to endogenous matching. I specify a model *à la* Abowd, Kramarz, and Margolis (1999) in which I am able to identify firm fixed effects:

$$\log w_{ijt} = x'_{ijt}\beta + v_t + \theta_i + \psi_{j(i,t)} + \epsilon_{ijt}, \tag{4}$$

where w_{ijt} is the hourly wage earned by individual i working in firm j in year t, x contains part-time activity, marriage and children, as well as quadratic specifications of age, full-time experience, part-time experience and seniority, v_t is a year fixed effect, θ_i is an individual fixed effect, $\psi_{j(i,t)}$ is a firm fixed effect and ϵ_{ijt} an error term, the variance of which is individual-specific. I do no longer control for location, size or industry because these covariates correspond to an aggregation of the pure firm effects $\psi_{j(i,t)}$ (Abowd, Kramarz, and Woodcock, 2008); more precisely, they are an employment-duration weighted average of the firm effects within the *département*/size[11]/industry.

[11]In an abuse of terminology, size refers to all firms with a size belonging to one of the 12 previously mentioned size categories.

4.3 Identification

The identification of the model is discussed in Abowd, Kramarz, and Margolis (1999) and provided in greater detail in Abowd, Creecy, and Kramarz (2002). It proceeds from the connectedness properties of the graph formed by individuals (let designate their number by N) and firms (let designate their number by J). Specifically, the data must be partitioned into G mutually exclusive groups of either individuals or firms such that the members of one group cannot have employed–or have been employed by–any member of another group. These G groups are the maximally connected sub-graphs of the entire graph, the vertices of which correspond to the union of the set of persons and the set of firms, while its edges are pairs of firms and persons. For each group g with N_g persons and J_g firms, $N_g - 1$ individual fixed effects and $J_g - 1$ firm fixed effects can be identified such that $N + J - G$ effects can be identified on the whole. The uniqueness of the effects within a group stems from the elimination of one person effect: it can be achieved by setting the group mean to zero as Abowd, Creecy, and Kramarz (2002) suggest.

4.4 Estimation

Technical details regarding the estimation of two-way high dimensional fixed effects are provided in Abowd, Creecy, and Kramarz (2002); in particular, one practical solution to cope with the inversion of large matrices consists in exploiting their sparsity. Efficient algorithms include the conjugate gradient and the "zigzag" Gauss-Seidel routine.

Among the four explanations for the parenthood pay gap presented in Section 2, this two-factor model enables me to distinguish carefully between individual unobserved heterogeneity and firm matching. First, to document selection effects, I recover the estimated individual fixed effect $\hat{\theta}_i$ and explain it using cohort effects and education dummies[12] in Z as follows:

$$\hat{\theta}_i = \gamma_z Z_i + \eta_i. \tag{5}$$

In other words, Equation (5) seeks to project estimated unobserved heterogeneity on observed heterogeneity.

[12]Including a "missing" category when no information is available on the highest degree obtained, as is the case for 18,277 individuals, namely 44% of the sample.

Second, to assess the existence of an endogenous matching process that would match individuals to specific firms, I compute the correlation between individual and firm fixed effects $cor(\theta, \psi)$ that indicates the extent to which high-wage workers self-select into high-wage firms. Separate correlations for parents and for non-parents shed some light on the differences between the assignment of parents to high-wage firms from the assignment of non-parents to high wage firms. For instance, it has already been documented that the correlation is almost zero in the US while it is negative in France, whatever the parenthood status; my findings are consistent with the latter result.

In practice, I estimate different models for women and for men, which allows me to proceed to separate analyses of both genders with respect to the issue of the parenthood penalty. Relatively little attention has been devoted to men, and *a fortiori* to both genders simultaneously, which is another dimension in which this paper contributes to the literature.

5 Results

5.1 Testing for endogenous matching

The main results are displayed in Table 5, columns 3 and 6, which report the estimates from the model including individual- and firm- fixed effects (2FE) with $G = 4,742$ groups for women and $G = 5,268$ groups for men. For both genders, I estimate three different specifications: first-difference (FD) in columns 1 and 4, individual fixed effects (FE) in columns 2 and 5 and individual and firm fixed effects in columns 3 and 6.

Overall, and in line with previous findings, the estimations suggest the existence of a parenthood wage penalty for French women working in the private sector, of approximately -3% per child. As transitions into public sector, self-employment, unemployment or inactivity are missing from the data, this penalty is likely underestimated because it is estimated conditional on working in the private sector even after childbirth. No significant effect is obtained for French men: in particular, I observe no fatherhood premium. The motherhood penalty exhibits some non-linearity with the rank of birth: -4.7% for the first child, -7.1% for the second child, -7.7% for the third child and -9.8% for the fourth child. Estimates corresponding to the fifth childbirth are much more imprecise due to the low sample size. These results argue for the existence of gender bias in the relationship between children and wages. They are consistent with heterogeneity in childbirth returns (the sample comprises young individuals) and with long-run or dynamic effects of motherhood. The FD approach by contrast tends to estimate a short-run effect; interestingly the short-run motherhood penalty is systematically lower in absolute terms than the long-run loss measured by FE or to 2FE methods, which is consistent with the cumulative penalties and dynamic effects of parenthood found by Kleven, Landais, and Søgaard (2015).

However, neither the FD nor FE estimates correct for the possibility of poorer job matches for parents than for non-parents, which 2FE does, and which may help in explaining part of the observed wage differential. The comparison of column 2 (resp. 5) with column 3 (resp. 6) precisely measures the share of the gap that is explained by endogenous worker-firm matching and by the corresponding omitted variable bias. However, the entire motherhood pay gap remains. By contrast, for men, the FE specification leads to some (small) penalties, while no significant effect of childbirth is obtained in the 2FE specification. As a result, the "firm matching" explanation is rejected in the case of women, while it seems to matter for men. To go further, I report several estimates of the childbirth coefficients in Tables 6 and 7, which correspond to different specifications of Equations (3) and (4). The coefficient of the third childbirth is nearly doubled when one fails to control for experience and for firm fixed effects. More generally, these tables quantify the role played by each of the explanations in the family gap previously mentioned in the literature.

The first explanation–human capital proxied by experience–determines up to 1/3 of the adjusted motherhood pay gap, especially from the third childbirth onwards (comparison of columns 2 and 3c). However, this result does not hold for men. Controlling for potential experience (column 3a) or failing to distinguish full-time from part-time experience (column 3b) also biases the coefficients of interest.

The second explanation–individual unobserved heterogeneity–is not the major reason for lower hourly wages after childbirth, for women or men. Columns 1 and 2 exhibit some differences, but they are are not significant at 5% or their magnitude is economically small.

The third explanation–firm matching–is rejected for women, which constitutes the main result of this paper: comparing estimates from column 3c to those from column 4 does not reveal any significant difference. In other words, women who plan to become mothers would not move to firms which offer particularly lower wages–when they anticipate a pregnancy, immediately after childbirth, or subsequently. On the contrary, there is a small though significant difference for men: failing to control for firm fixed effects yields small penalties, while after including the latter the effect of childbirth is no longer significant at usual levels. Fathers would tend to move to firms that offer slightly lower compensation, and the negative coefficient in the FE specification would then account for such a spurious correlation.

Table 8 sheds some light on worker unobserved heterogeneity by displaying the results of (5), which depicts how estimated unobserved heterogeneity $\hat{\theta}_i$ depends on the covariates. Once cohort effects have been taken into account, the average individual productivity is almost an increasing function of education. This holds for both men and women. Table 9 presents correlations between individual and firm unobserved heterogeneity. Consistently with previous findings, this correlation is negative and amounts in the sample to roughly -.21. It is approximately -.21 for non-mothers and -.25 for mothers but -.22 for non-fathers and -.2 for fathers. Overall, these figures suggest that there are only limited firm matching forces that would trap parents in low wage firms. These results also indicate that firm matching works slightly differently for women and men because the previous correlation is lower for mothers than for non-mothers, while the opposite is the case for men.

My results are also consistent with the literature focusing on the effect of marriage on wages. I find that a "marriage pay premium" amounts to 2.6% for women and to 2.7% for men. Interestingly, most of the corresponding literature focuses on men's marriage premium, while one cannot reject that this premium is as high for women as it is for men.

Finally, I investigate why there is no fatherhood premium in the data. Although a thorough analysis would be beyond the scope of this paper,[13] I provide here an estimation of the evolution of that fatherhood premium over time. I estimate the same model as (4) but allow both marriage and fatherhood premia to vary over time. Based on the same data, I consider a longer period (1976-2011) and hence rely on the daily wage instead of the hourly wage. The corresponding

[13]It is actually the object of an on-going work.

sample of interest is composed of men working full-time in the private sector. Figure 1 displays the results, which are consistent both with the fatherhood premium reported in the literature and with my results on the effect of childbirth on men's wages since 1995. This premium seems to have eroded over time, from roughly 5% at first childbirth until 1998, to almost zero at the end of the 2000s. Understanding why the fatherhood premium has disappeared is a challenging but rewarding task for applied research in this domain; it requires however to disentangle composition effects due to the lower participation of spouses to the labor market, from the childbirth effect.

5.2 How much does the family gap contribute to the gender gap?

To evaluate the contribution of this gender-biased parenthood penalty to the gender pay gap, I simulate a counterfactual *scenario* in which women would experience the same childbirth penalty as men, that is, no penalty at all. Public interventions might well consist in promoting paternity leaves, which could reduce or even eliminate such gender inequality with respect to parenthood. From women's observed wages w_{it}^o, I compute therefore their simulated wages w_{it}^s in the case in which they face no motherhood penalty:

$$\log w_{it}^s = \log w_{it}^o - \sum_{k=1}^{5} \beta_{ck}^{\text{Women}} \text{Childbirth}_{itk}. \tag{6}$$

From pooled cross-sectional data, I estimate annual adjusted gender pay gaps Δ_t^o and Δ_t^s on both observed and simulated wages. Denoting by G_i the gender dummy that is equal to 1 if individual i is a woman, I specify $\forall l \in (o, s)$,:

$$\log w_{it}^l = \Delta_t^l G_i + X_{it}' \beta_t^l + N_{jt}' \delta_t^l + \epsilon_{it}^l. \tag{7}$$

Figure 2 depicts the fraction of women's wages in terms of men's wages in both observed and counterfactual *scenarios*. Figure 2a displays the corresponding patterns over time, while Figure 2b plots these patterns against age. First, the sample is composed of individuals aged 26.9 on average, hence for whom the gender gap is rather low. In 2013, the unadjusted French gender gap is almost zero for individuals aged less than 25. Here, the adjusted gender pay gap varies from

3.5% at the end of the 1990s to 5.5% at the beginning of the 2010s. Second, the increase in the gender gap over time that is apparent from Figure 2a is largely due to a composition effect: in 2011, the sample is mechanically composed of older individuals than in 1995 because of the entry condition. As Figure 2b shows, older individuals experience a higher gender gap. Third, the sample contains few individuals aged older than 50, which makes the estimates of the gender gap very imprecise at those ages. Nevertheless, the *scenario* in which women do not encounter any motherhood penalty is still far from corresponding to equal wages between men and women. Even in 2011, the family pay gap would explain at most 1/3 of the gender pay gap, even though at higher ages the former could represent almost one-half of the latter (observe, for example, the result at age 35 in Figure 2b). Investigating, in greater detail, the contribution of the motherhood penalty to the gender gap at higher ages crucially depends on the availability of appropriate data. To conclude, in this sample one has both gender inequalities with respect to parenthood and remaining gender wage differentials.

5.3 Robustness checks

I proceed to three robustness checks of the results. First we examine the sensitivity of the estimates to outliers. I perform several estimations with and without trimming hourly wages. Table 10 displays the corresponding results: column 1 corresponds to no trimming, column 2 corresponds to the elimination of hourly wages below a .8 minimum hourly wage (the base specification), column 3 to the elimination of hourly wages below a 1 minimum hourly wage, and column 4 further imposes a cap at 100 euros following Felfe (2012). Overall, and although eliminating outliers tends to reduce the estimated loss, I find a limited impact on the motherhood penalty, while the absence of trimming at the bottom of the distribution leads to significant and large fatherhood wage penalties; no trimming at the top results in close estimates.

Second, I investigate whether different measures of experience alter the result. As argued above, when seriously investigating the family gap issue, it is important to compute the experience covariate as accurately as possible. Resorting to administrative data is an helpful tool that enables me to provide an almost ideal variable with little measurement error. The definition of experience matters: in addition to counting the amount of time spent on-the-job, carefully distinguishing full-time from part-time experience has an impact of the estimated effect of children on wages. Childbirth coefficients differ slightly according to whether one controls for experience as a whole or for both full-time and part-time experience. Potential experience, which is a poor measure of the actual time spent in the workforce, performs worse. [14]

Third, I check whether the above results are robust to the inclusion of occupational covariates in log hourly wage equations. In general, I am reluctant to control for occupation in the wage equations because it is likely to be correlated with unobserved determinants of wages including talent or productivity, and hence occupation may be regarded as an endogenous variable. I nevertheless assess whether controlling for such covariates dramatically alters the conclusion, as there is no consensus in the literature on that topic. Table 12 displays the corresponding results and shows that not only do the signs and significance of childbirth effects remain once occupation (namely dummies defined by the two-digit PCS-ESE French classification) has been controlled for but also their magnitude. There is hardly any attenuation in the FE specification, but no significant difference is observed in the 2FE specification between columns 3 and 6 of Tables 5 and 12.

As documented in the literature, childbirth returns are heterogeneous across individuals and depend on many factors, including the position in the wage distribution, education, occupation, etc. I choose here not to investigate further these several dimensions of heterogeneity to focus on the omitted variable bias issue. Other dimensions may well affect this family gap, such as the presence of unions, the distance between home and the workplace, etc.

6 Conclusion

This paper has reexamined the family pay gap by employing linked employer-employee data and controlling for three explanatory factors in wage equations: experience as a proxy for human capital depreciation, worker fixed effects and firm fixed effects. It provides a test of the firm matching explanation, according to which endogenous selection of parents into low-wage firms would spuriously

[14] In the presence of age and individual fixed effects, the slope of potential experience is not identified due to collinearity, as potential experience is defined here as the difference between the current year and the year an individual first appears in the panel.

explain the parenthood penalty. I estimate a linear model in the presence of two-way high dimensional fixed effects on a sample of young French individuals working in the private sector from 1995 to 2011. I find a motherhood wage penalty of approximately -3% per child on the hourly wage, the effect being more pronounced at the first childbirth. By contrast, fathers do not experience any significant loss following childbirths, but they also do not enjoy any premium. While I reject the firm matching explanation as the main reason for the gender-biased parenthood penalty affecting women, mobility between firms is likely to play a role in the case of men. Moreover, it appears that the so-called fatherhood premium has eroded over the period from 1976 to 2011 in France.

I also find that the human capital depreciation explanation accounts for a share of gender inequalities in parenthood. The remainder of the inequality could be due to discrimination against mothers at work, which might stem from within-firm labor reallocation: mothers would be less exposed to risky assignments and thus less likely to receive bonuses or even be trapped in low-wage trajectories. Explicitly testing for the presence of discrimination against mothers at work is a task that should be developed in future research. Such a gender inequality is both unfair and inefficient, which legitimates further public intervention, including campaigns against discrimination, the development of on-the-job childcare and the development/extension of paternity leave as is the case in Scandinavian countries. If women interrupt their careers for 16 weeks at childbirth, a paternity leave of the same duration may provide a way to bring down this gender gap. The paternity leave would certainly have a causal, positive impact on a less unequal housework division between spouses (especially as regards childcare, see Pailhé, Solaz, and Tô, 2015).[15]

[15] This might also help explain the erosion of the fatherhood premium in the 2000s. The introduction of a short paternity leave (11 days) in France dates to 2002.

References

ABOWD, J. M., R. H. CREECY, AND F. KRAMARZ (2002): “Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data,” Cornell University Working Paper.

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): “High wage workers and high wage firms,” *Econometrica*, 67(2), 251–333.

ABOWD, J. M., F. KRAMARZ, AND S. WOODCOCK (2008): “Econometric Analyses of Linked Employer-Employee Data,” in *The Econometrics of Panel Data*, ed. by L. Matyas, and P. Sevestre, pp. 727–760. Springer.

ALBRECHT, J. W., P.-A. EDIN, M. SUNDSTRÖM, AND S. B. VROMAN (1999): “Career interruptions and subsequent earnings: A reexamination using Swedish data,” *Journal of Human Resources*, XXXIV(2), 294–311.

BEBLO, M., S. BENDER, AND E. WOLF (2009): “Establishment-level wage effects of entering motherhood,” *Oxford Economic Papers*, 61(suppl 1), i11–i34.

BECKER, G. S. (1985): “Human capital, effort, and the sexual division of labor,” *Journal of Labor Economics*, 3(1), S33–S58.

BOOTH, A. L., AND M. WOOD (2008): “Back-to-Front Down Under? Part-Time/Full-Time Wage Differentials in Australia,” *Industrial Relations*, 47(1), 114–135.

BUDIG, M., AND M. HODGES (2010): “Differences in disadvantage: How the wage penalty for motherhood varies across women’s earnings distribution,” *American Sociological Review*, 75(5), 705–728.

BUDIG, M. J., AND P. ENGLAND (2001): “The wage penalty for motherhood,” *American Sociological Review*, 66(2), 204–225.

BULIGESCU, B., D. DE CROMBRUGGHE, G. MENTEŞOĞLU, AND R. MONTIZAAN (2009): “Panel estimates of the wage penalty for maternal leave,” *Oxford Economic Papers*, 61(suppl 1), i35–i55.

DAVIES, R., AND G. PIERRE (2005): “The family gap in pay in Europe: A cross-country study,” *Labour Economics*, 12(4), 469–486.

England, P., J. Bearak, M. Budig, and M. Hodges (2013): "How the Motherhood Penalty Varies by Wage and Cognitive Skill: A Reassessment," Working paper.

Felfe, C. (2012): "The motherhood wage gap: What about job amenities?," *Labour Economics*, 19(1), 59–67.

Glauber, R. (2008): "Race and Gender in Families and at Work The Fatherhood Wage Premium," *Gender & Society*, 22(1), 8–30.

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

Joseph, O., A. Pailhé, I. Recotillet, and A. Solaz (2013): "The economic impact of taking short parental leave: Evaluation of a French reform," *Labour Economics*, 25, 63–75.

Killewald, A. (2013): "A Reconsideration of the Fatherhood Premium Marriage, Coresidence, Biology, and Fathers' Wages," *American Sociological Review*, 78(1), 96–116.

Kleven, H. J., C. Landais, and J. E. Søgaard (2015): "Parenthood and the Gender Gap: Evidence from Denmark," Working Paper.

Korenman, S., and D. Neumark (1992): "Marriage, Motherhood, and Wages," *Journal of Human Resources*, 27(2), 233–255.

Lequien, L. (2012): "The impact of parental leave duration on later career," WP Crest 2012-04.

Lundberg, S., and E. Rose (2000): "Parenthood and the earnings of married men and women," *Labour Economics*, 7(6), 689–710.

——— (2002): "The effects of sons and daughters on men's labor supply and wages," *Review of Economics and Statistics*, 84(2), 251–268.

Meurs, D., A. Pailhé, and S. Ponthieux (2010): "Child-related career interruptions and the gender wage gap in France," *Annales d'Économie et de Statistique*, 99/100, 15–46.

MEURS, D., AND S. PONTHIEUX (2000): "Une mesure de la discrimination dans l'écart de salaire entre hommes et femmes," *Économie et statistique*, 337(1), 135–158.

MEURS, D., AND S. PONTHIEUX (2014): "Gender Inequality," in *Handbook of Income Distribution SET vols. 2A-2B*, ed. by A. Atkinson, and F. Bourguignon. North Holland.

MILLER, A. R. (2011): "The effects of motherhood timing on career path," *Journal of Population Economics*, 24(3), 1071–1100.

MINCER, J. (1958): "Investment in human capital and personal income distribution," *The Journal of Political Economy*, 66(4), 281–302.

NIELSEN, H. S., M. SIMONSEN, AND M. VERNER (2004): "Does the Gap in Family-friendly Policies Drive the Family Gap?," *The Scandinavian Journal of Economics*, 106(4), 721–744.

OECD (2012): *Closing the Gender Gap: Act Now.* OECD Publishing.

PAILHÉ, A., A. SOLAZ, AND M. TÔ (2015): "The Impact of Paternity Leave on Housework Division Between Spouses," Working Paper.

SIMONSEN, M., AND L. SKIPPER (2006): "The costs of motherhood: An analysis using matching estimators," *Journal of Applied Econometrics*, 21(7), 919–934.

——— (2012): "The family gap in wages: What wombmates reveal," *Labour Economics*, 19(1), 102–112.

SPENCE, M. (1973): "Job market signaling," *The Quarterly Journal of Economics*, 87(3), 355–374.

WALDFOGEL, J. (1997): "The effect of children on women's wages," *American Sociological Review*, 123, 209–217.

——— (1998): "The family gap for young women in the United States and Britain: Can maternity leave make a difference?," *Journal of Labor Economics*, 16(3), 505–545.

Wilde, E. T., L. Batchelder, and D. T. Ellwood (2010): "The mommy track divides: The impact of childbearing on wages of women of differing skill levels," NBER Working paper 16582.

Tables

Table 1: Sample of women - Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
First year in panel	19932	2003	4.728	1995	2011
Married	19932	0.211	0.408	0	1
One child	19932	0.181	0.385	0	1
Two children	19932	0.132	0.339	0	1
Three children	19932	0.046	0.209	0	1
Four children	19932	0.012	0.107	0	1
Five children	19932	0.006	0.077	0	1
Age	95499	26.899	7.549	16	65
Potential experience	95499	6.253	4.240	1	17
Full-time experience	95499	2.117	2.717	0	16.9
Part-time experience	95499	0.715	1.146	0	12.1
Seniority	95499	2.402	2.586	0.01	17
Nb. of working days	95499	229.4	134.1	1	360
Nb. of working hours	95499	936	706	11	4056
Part time	95499	0.433	0.495	0	1
Part time (FTU)	95499	0.247	0.431	0	1
Net hourly wage	95499	9.43	5.74	3.51	1010
Seniority	135431	2.046	2.403	0.01	17

Sample of 19932 women working in the private sector from 1995 to 2011 (95499 individual×year observations, 135431 individual×firm×year observations). Wages: in 2011 euros. Full-time experience: in full-time units (FTU).

Table 2: Sample of men - Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
First year in panel	21599	2002	4.653	1995	2011
Married	21599	0.168	0.374	0	1
One child	21599	0.151	0.358	0	1
Two children	21599	0.110	0.313	0	1
Three children	21599	0.033	0.178	0	1
Four children	21599	0.009	0.094	0	1
Five children	21599	0.002	0.048	0	1
Age	116690	26.901	7.235	16	65
Potential experience	116690	6.598	4.242	1	17
Full-time experience	116690	3.039	3.179	0	17
Part-time experience	116690	0.372	0.711	0	14
Seniority	116690	2.556	2.629	0.01	17
Nb. of working days	116690	243.6	130.1	1	360
Nb. of working hours	116690	1120	734	11	4400
Part time	116690	0.243	0.429	0	1
Part time (FTU)	116690	0.097	0.295	0	1
Net hourly wage	116690	10.42	8.44	3.51	1760
Seniority	165648	2.174	2.457	0.01	17

Sample of 21599 individuals working in the private sector from 1995 to 2011 (116690 individual×year observations, 165648 individual×firm×year observations). Wages: in 2011 euros. Full-time experience: in full-time units (FTU).

Table 3: Estimates of part-time/full-time log hourly wage differential

	Women		Men	
Specification	Pooled OLS	FE	Pooled OLS	FE
(1)	-0.078***	-0.017***	-0.072***	-0.040***
	(0.002)	(0.003)	(0.002)	(0.003)
(2)	-0.050***	0.030***	-0.003***	0.056***
	(0.002)	(0.002)	(0.002)	(0.003)
(3)	-0.017***	0.034***	0.016***	0.054***
	(0.002)	(0.002)	(0.002)	(0.003)
(4)	-0.011***	0.034***	0.027***	0.054***
	(0.002)	(0.002)	(0.002)	(0.003)
(5a)	-0.003	0.034***	0.033***	0.054***
	(0.002)	(0.002)	(0.002)	(0.003)
(5b)	0.013***	0.040***	0.051***	0.062***
	(0.002)	(0.002)	(0.002)	(0.003)
(5c)	0.045***	0.052***	0.065***	0.067***
	(0.002)	(0.002)	(0.002)	(0.003)
Observations	95499	95499	116690	116690

(1) contains a constant.
(2) adds worker characteristics (quadratic specification in age, dummy if married).
(3) adds firm characteristics (*département*, two-digit industry dummies and establishment size).
(4) adds a quadratic specification in tenure.
(5a) = (4) + quadratic specification in potential experience.
(5b) = (4) + quadratic specification in experience.
(5c) = (4) + quadratic specification in full-time and part-time experience.

Table 4: Impact of marriages and childbirths on experience

	Women		Men	
Dependent	FT exp.	PT exp.	FT exp.	PT exp.
Age	0.409***	0.123***	0.500***	0.057***
	(0.007)	(0.004)	(0.006)	(0.002)
Marriage	0.337***	-0.046	0.593***	-0.043*
	(0.053)	(0.028)	(0.048)	(0.018)
First childbirth	0.183***	0.052*	0.474***	-0.044***
	(0.040)	(0.021)	(0.035)	(0.012)
Second childbirth	-0.237***	0.260***	0.902***	-0.078***
	(0.070)	(0.037)	(0.057)	(0.020)
Third childbirth	-1.362***	0.339***	0.959***	-0.062
	(0.129)	(0.069)	(0.109)	(0.039)
Fourth childbirth	-2.570***	0.158	1.072***	-0.123*
	(0.250)	(0.112)	(0.242)	(0.058)
Fifth childbirth	-3.530***	-0.047	-0.494	-0.027
	(0.572)	(0.155)	(0.485)	(0.137)
Year dummies	Yes	Yes	Yes	Yes
Individual effects	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes
Observations	95499	95499	116690	116690
R^2	0.858	0.779	0.902	0.768

FT exp.: full-time experience, PT exp.: part-time experience.

Table 5: Log hourly wages

	Women			Men		
	(1) FD	(2) FE	(3) 2FE	(4) FD	(5) FE	(6) 2FE
Marriage	0.034***	0.029***	0.026***	0.054***	0.045***	0.027***
	(0.004)	(0.005)	(0.007)	(0.004)	(0.006)	(0.007)
First childbirth	-0.057***	-0.047***	-0.048***	-0.005	-0.006	0.001
	(0.003)	(0.004)	(0.005)	(0.003)	(0.004)	(0.005)
Second childbirth	-0.095***	-0.072***	-0.074***	-0.007	-0.018**	-0.002
	(0.005)	(0.006)	(0.008)	(0.005)	(0.006)	(0.008)
Third childbirth	-0.107***	-0.081***	-0.080***	-0.012	-0.038***	-0.004
	(0.008)	(0.011)	(0.015)	(0.008)	(0.011)	(0.013)
Fourth childbirth	-0.161***	-0.103***	-0.103***	-0.007	-0.013	0.007
	(0.015)	(0.020)	(0.029)	(0.015)	(0.021)	(0.028)
Fifth childbirth	-0.107***	-0.050	-0.071	-0.086*	-0.089*	0.031
	(0.021)	(0.049)	(0.060)	(0.034)	(0.034)	(0.047)
Age	0.048***	0.044***	0.014	0.057***	0.058***	0.040
	(0.002)	(0.002)	(0.101)	(0.002)	(0.002)	(0.114)
Age^2 (1e-3)	-0.408***	-0.356***	-0.302***	-0.531***	-0.580***	-0.461***
	(0.024)	(0.030)	(0.045)	(0.023)	(0.031)	(0.046)
Part-time	0.063***	0.057***	0.060***	0.066***	0.068***	0.077***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)
Seniority	0.018***	0.015***	0.011***	0.025***	0.015***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Seniority^2$ (1e-3)	-0.995***	-1.006***	-0.753***	-1.576***	-1.021***	-0.600***
	(0.112)	(0.116)	(0.136)	(0.109)	(0.135)	(0.159)
FT Experience	0.023***	0.034***	0.032***	0.019***	0.041***	0.039***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
FT $Experience^2$ (1e-3)	-0.533***	-0.976***	-0.912***	-0.547***	-1.347***	-1.317***
	(0.146)	(0.159)	(0.203)	(0.117)	(0.131)	(0.168)
PT Experience	-0.038***	-0.021***	-0.011*	-0.027***	0.001	-0.009
	(0.003)	(0.004)	(0.005)	(0.004)	(0.007)	(0.008)
PT $Experience^2$ (1e-3)	3.499***	2.666***	1.960**	4.320***	-0.368	2.791
	(0.452)	(0.503)	(0.716)	(0.653)	(1.380)	(1.484)
Year dummies	No	Yes	Yes	No	Yes	Yes
Individual effects	No	Yes	Yes	No	Yes	Yes
Industry dummies	Yes	Yes	No	Yes	Yes	No
Regional dummies	Yes	Yes	No	Yes	Yes	No
Firm size controls	Yes	Yes	No	Yes	Yes	No
Firm effects	No	No	Yes	No	No	Yes
Observations	63260	95499	135431	80808	116690	165648
Nb. individuals	15721	19932	19932	18012	21599	21599
Nb. firms	21513	31189	42937	25770	36408	49556
R^2	0.220	0.683	0.817	0.270	0.722	0.840

Clustered standard errors at the individual level in parentheses
Industry dummies: 39 two-digit dummies (NACE)
Firm size controls: 12 dummies

Table 6: Coefficients of childbirth in Mincer equations - Women

Specification	(1)	(2)	(3a)	(3b)	(3c)	(4)
First childbirth	-0.051***	-0.044***	-0.046***	-0.050***	-0.047***	-0.048***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Second childbirth	-0.073***	-0.083***	-0.087***	-0.080***	-0.072***	-0.074***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Third childbirth	-0.129***	-0.118***	-0.123***	-0.089***	-0.081***	-0.080***
	(0.010)	(0.013)	(0.013)	(0.013)	(0.012)	(0.015)
Fourth childbirth	-0.176***	-0.164***	-0.171***	-0.104***	-0.103***	-0.103***
	(0.013)	(0.024)	(0.024)	(0.023)	(0.023)	(0.029)
Fifth childbirth	-0.175***	-0.129**	-0.135**	-0.046	-0.050	-0.071
	(0.020)	(0.063)	(0.063)	(0.054)	(0.055)	(0.060)
Observations	95499	95499	95499	95499	95499	95499
R^2	0.314	0.676	0.676	0.680	0.683	0.817

(1) = Pooled OLS with year and cohort dummies.
(2) = FE with year dummies and a quadratic specification in tenure.
(3a) = (2) + quadratic specification in potential experience.
(3b) = (2) + quadratic specification in experience.
(3c) = (2) + quadratic specification in full-time and part-time experience.
(4) = 2FE.

Table 7: Coefficients of childbirth in Mincer equations - Men

Specification	(1)	(2)	(3a)	(3b)	(3c)	(4)
First childbirth	-0.005	0.004	0.004	-0.006	-0.006	0.001
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Second childbirth	0.010	-0.004	-0.005	-0.018**	-0.018**	-0.002
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Third childbirth	-0.029*	-0.026**	-0.027**	-0.038***	-0.038***	-0.004
	(0.017)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)
Fourth childbirth	0.008	-0.003	-0.004	-0.011	-0.013	0.007
	(0.032)	(0.024)	(0.024)	(0.023)	(0.023)	(0.028)
Fifth childbirth	-0.142***	-0.109**	-0.109**	-0.085**	-0.089**	0.031
	(0.047)	(0.045)	(0.045)	(0.039)	(0.038)	(0.047)
Observations	116690	116690	116690	116690	116690	116690
R^2	0.343	0.718	0.718	0.722	0.722	0.840

(1) = Pooled OLS with year and cohort dummies.
(2) = FE with year dummies and a quadratic specification in tenure.
(3a) = (2) + quadratic specification in potential experience.
(3b) = (2) + quadratic specification in experience.
(3c) = (2) + quadratic specification in full-time and part-time experience.
(4) = 2FE.

Table 8: Individual unobserved heterogeneity - log hourly wages

	Women	Men
Degree is missing	ref	ref
No degree	-0.060***	-0.113***
	(0.006)	(0.006)
Elementary education	-0.035	-0.136***
	(0.023)	(0.025)
Junior high school	-0.020*	-0.078***
	(0.009)	(0.009)
Basic vocational	-0.023***	-0.092***
	(0.005)	(0.005)
Advanced vocational	0.007	-0.061***
	(0.006)	(0.006)
High school degree	0.010	-0.026**
	(0.008)	(0.009)
Some college	0.068***	0.006
	(0.005)	(0.006)
College degree	0.171***	0.149***
	(0.006)	(0.007)
Age effects	Yes	Yes
Observations	19932	21599
R^2	0.184	0.083

Note. The dependent variable is $\hat{\theta}_i$ (see Equation 5).

Table 9: Correlation between individual and firm unobserved heterogeneity

	Women	Men
No child	-0.214	-0.224
One child or more	-0.249	-0.209

Figures

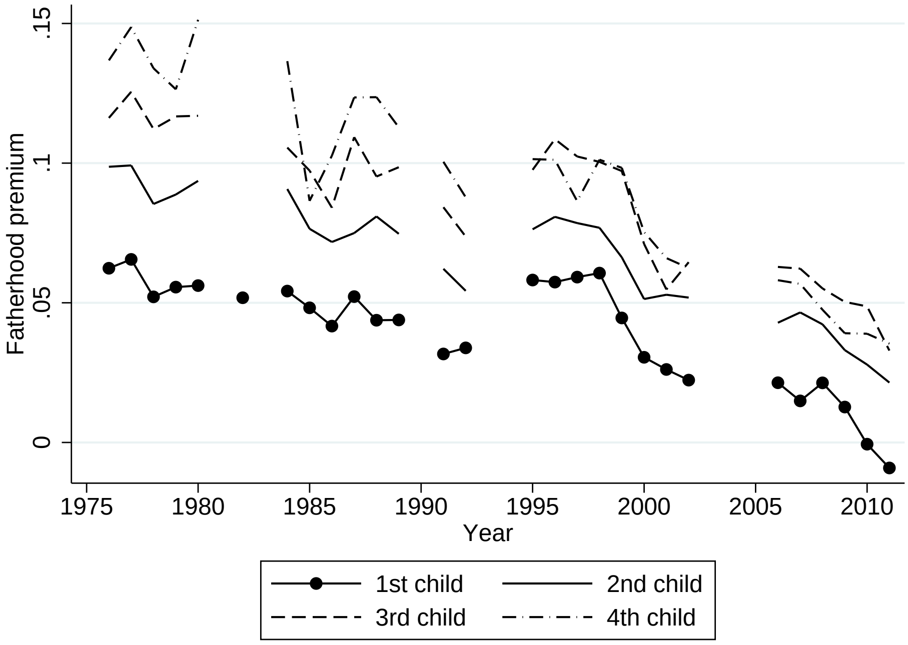
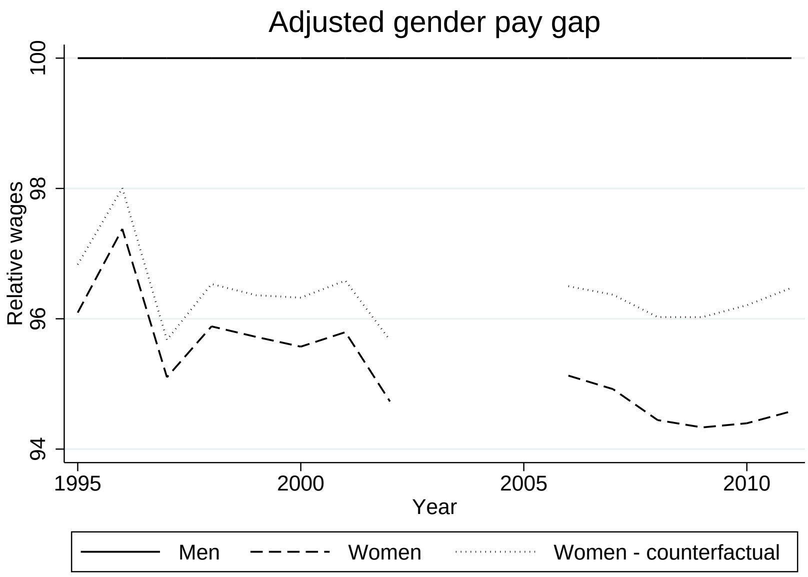
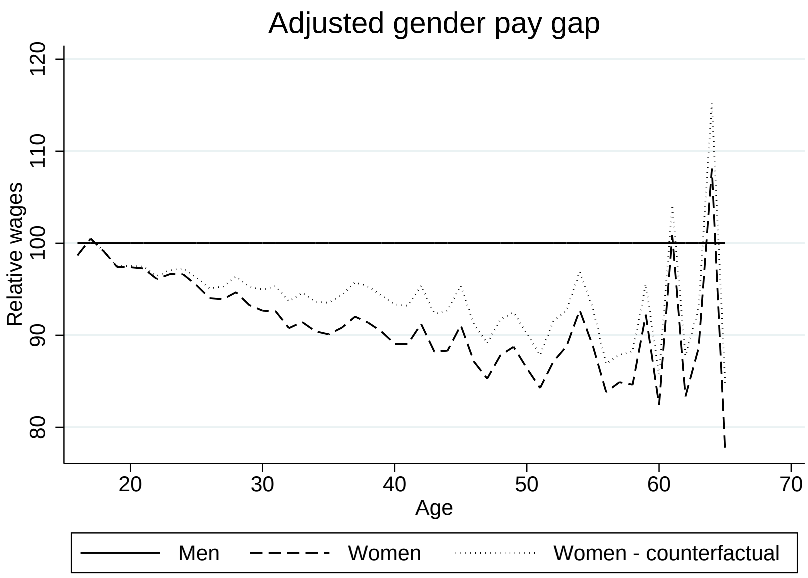


Figure 1: The erosion of the fatherhood premium (daily wage – men, full-time workers, 1976-2011)



(a) Over time



(b) By age

Figure 2: The counterfactual gender pay gap: what if women experienced the same penalty as men regarding childbirth?

Robustness checks

Table 10: Sensitivity to the trimming of outliers - log hourly wages

	Women					Men		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE	FE	FE	FE	FE	FE	FE	FE
Marriage	0.014*	0.029***	0.028***	0.028***	0.015*	0.045***	0.046***	0.045***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.007)	(0.006)	(0.006)	(0.006)
First childbirth	-0.064***	-0.047***	-0.038***	-0.038***	-0.028***	-0.006	-0.005	-0.006
	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
Second childbirth	-0.106***	-0.072***	-0.058***	-0.058***	-0.054***	-0.018**	-0.012	-0.012*
	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
Third childbirth	-0.129***	-0.081***	-0.063***	-0.063***	-0.079***	-0.038***	-0.030**	-0.030**
	(0.013)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Fourth childbirth	-0.156***	-0.103***	-0.083***	-0.082***	-0.063*	-0.013	0.000	0.002
	(0.025)	(0.020)	(0.020)	(0.020)	(0.025)	(0.021)	(0.021)	(0.021)
Fifth childbirth	-0.149*	-0.050	-0.041	-0.040	-0.071	-0.089*	-0.072*	-0.064*
	(0.065)	(0.049)	(0.052)	(0.052)	(0.050)	(0.034)	(0.035)	(0.032)
Age	0.078***	0.044***	0.039***	0.039***	0.113***	0.058***	0.052***	0.051***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age^2 (1e-3)	-0.806***	-0.356***	-0.290***	-0.291***	-1.395***	-0.580***	-0.486***	-0.477***
	(0.036)	(0.030)	(0.029)	(0.029)	(0.039)	(0.031)	(0.030)	(0.029)
Part-time	0.112***	0.057***	0.041***	0.040***	0.131***	0.068***	0.053***	0.052***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Seniority	0.013***	0.015***	0.016***	0.016***	0.011***	0.015***	0.017***	0.017***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Seniority^2 (1e-3)	-0.969***	-1.006***	-1.095***	-1.095***	-0.885***	-1.021***	-1.086***	-1.121***
	(0.133)	(0.116)	(0.113)	(0.113)	(0.154)	(0.135)	(0.131)	(0.126)
FT Experience	0.046***	0.034***	0.029***	0.029***	0.063***	0.041***	0.035***	0.035***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
FT Experience^2 (1e-3)	-1.943***	-0.976***	-0.771***	-0.750***	-2.587***	-1.347***	-1.150***	-1.157***
	(0.177)	(0.159)	(0.155)	(0.153)	(0.150)	(0.131)	(0.128)	(0.126)
PT Experience	-0.032***	-0.021***	-0.018***	-0.019***	-0.006	0.001	0.004	0.005
	(0.005)	(0.004)	(0.004)	(0.004)	(0.008)	(0.007)	(0.007)	(0.007)
PT Experience^2 (1e-3)	3.463***	2.666***	2.313***	2.331***	0.103	-0.368	-0.605	-0.716
	(0.630)	(0.503)	(0.492)	(0.492)	(1.495)	(1.380)	(1.378)	(1.372)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm size controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	99967	95499	91825	91813	124977	116690	112560	112498
Nb. individuals	20300	19932	19665	19664	22213	21599	21288	21286
R^2	0.675	0.683	0.698	0.703	0.724	0.722	0.733	0.733

Same legend as Table 5.
Columns (1) and (5): no trimming.
Columns (2) and (6): base specification (hourly wage>.8 minimum wage).
Columns (3) and (7): hourly wage≥minimum hourly wage.
Columns (4) and (8): hourly wage∈ [minimum hourly wage; 100].

Table 11: Sensitivity to the specification of experience - log hourly wages

	Women				Men			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE	FE	FE	FE	FE	FE	FE	FE
Marriage	0.038***	0.037***	0.030***	0.029***	0.057***	0.056***	0.044***	0.045***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
First childbirth	-0.044***	-0.046***	-0.050***	-0.047***	0.004	0.004	-0.006	-0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Second childbirth	-0.083***	-0.087***	-0.080***	-0.072***	-0.004	-0.005	-0.018**	-0.018**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Third childbirth	-0.118***	-0.123***	-0.089***	-0.081***	-0.026*	-0.027*	-0.038***	-0.038***
	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Fourth childbirth	-0.164***	-0.171***	-0.104***	-0.103***	-0.003	-0.004	-0.011	-0.013
	(0.021)	(0.021)	(0.020)	(0.020)	(0.021)	(0.022)	(0.021)	(0.021)
Fifth childbirth	-0.129*	-0.135*	-0.046	-0.050	-0.109**	-0.109**	-0.085*	-0.089*
	(0.056)	(0.056)	(0.048)	(0.049)	(0.040)	(0.040)	(0.035)	(0.034)
Age	0.054***	0.051***	0.043***	0.044***	0.071***	0.070***	0.056***	0.058***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age^2 (1e-3)	-0.390***	-0.420***	-0.405***	-0.356***	-0.582***	-0.589***	-0.588***	-0.580***
	(0.029)	(0.031)	(0.029)	(0.030)	(0.030)	(0.031)	(0.031)	(0.031)
Part-time	0.040***	0.040***	0.045***	0.057***	0.055***	0.055***	0.063***	0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Seniority	0.018***	0.020***	0.014***	0.015***	0.024***	0.024***	0.014***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Seniority^2$ (1e-3)	-0.941***	-1.086***	-1.116***	-1.006***	-1.322***	-1.359***	-0.985***	-1.021***
	(0.105)	(0.113)	(0.118)	(0.116)	(0.114)	(0.120)	(0.137)	(0.135)
Potential experience		.				.		
		(.)				(.)		
Pot. $exp.^2$ (1e-3)		0.303***				0.078		
		(0.076)				(0.073)		
Experience			0.029***				0.044***	
			(0.002)				(0.002)	
$Experience^2$ (1e-3)			-0.207				-1.300***	
			(0.148)				(0.131)	
FT Experience				0.034***				0.041***
				(0.002)				(0.002)
FT $Experience^2$ (1e-3)				-0.976***				-1.347***
				(0.159)				(0.131)
PT Experience				-0.021***				0.001
				(0.004)				(0.007)
PT $Experience^2$ (1e-3)				2.666***				-0.368
				(0.503)				(1.380)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm size controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	95499	95499	95499	95499	116690	116690	116690	116690
Nb. individuals	19932	19932	19932	19932	21599	21599	21599	21599
R^2	0.676	0.676	0.680	0.683	0.718	0.718	0.722	0.722

Same legend as Table 5.

Table 12: Sensitivity to the inclusion of occupational covariates - log hourly wages

	Women			Men		
	(1)	(2)	(3)	(4)	(5)	(6)
	FD	FE	2FE	FD	FE	2FE
Marriage	0.023***	0.017***	0.018**	0.034***	0.026***	0.019**
	(0.004)	(0.005)	(0.006)	(0.004)	(0.005)	(0.007)
First childbirth	-0.046***	-0.039***	-0.044***	0.000	0.001	0.005
	(0.003)	(0.004)	(0.005)	(0.003)	(0.003)	(0.005)
Second childbirth	-0.071***	-0.055***	-0.067***	0.001	-0.003	0.005
	(0.005)	(0.005)	(0.008)	(0.005)	(0.006)	(0.007)
Third childbirth	-0.075***	-0.063***	-0.073***	0.009	-0.007	0.012
	(0.007)	(0.010)	(0.014)	(0.008)	(0.010)	(0.012)
Fourth childbirth	-0.117***	-0.076***	-0.097***	0.014	0.020	0.030
	(0.015)	(0.019)	(0.028)	(0.014)	(0.019)	(0.027)
Fifth childbirth	-0.069***	-0.008	-0.046	-0.056	-0.041	0.064
	(0.020)	(0.043)	(0.059)	(0.031)	(0.035)	(0.044)
Age	0.040***	0.038***	0.010	0.046***	0.048***	0.047
	(0.001)	(0.002)	(0.092)	(0.001)	(0.002)	(0.088)
Age^2 (1e-3)	-0.308***	-0.267***	-0.256***	-0.400***	-0.441***	-0.378***
	(0.022)	(0.027)	(0.043)	(0.022)	(0.028)	(0.043)
Part-time	0.060***	0.054***	0.058***	0.062***	0.065***	0.075***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)
Seniority	0.017***	0.014***	0.010***	0.023***	0.016***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Seniority^2 (1e-3)	-0.879***	-0.895***	-0.600***	-1.366***	-1.014***	-0.494***
	(0.105)	(0.104)	(0.128)	(0.101)	(0.119)	(0.144)
FT Experience	0.010***	0.018***	0.020***	0.009***	0.025***	0.027***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)
FT Experience^2 (1e-3)	0.075	-0.200	-0.414*	-0.026	-0.613***	-0.855***
	(0.138)	(0.140)	(0.190)	(0.108)	(0.112)	(0.152)
PT Experience	-0.031***	-0.020***	-0.014**	-0.028***	-0.006	-0.017*
	(0.003)	(0.003)	(0.005)	(0.003)	(0.006)	(0.008)
PT Experience^2 (1e-3)	2.910***	2.551***	2.029**	3.889***	0.539	3.500*
	(0.424)	(0.455)	(0.680)	(0.602)	(1.318)	(1.483)
Year dummies	No	Yes	Yes	No	Yes	Yes
Occupational dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual effects	No	Yes	Yes	No	Yes	Yes
Industry dummies	Yes	Yes	No	Yes	Yes	No
Regional dummies	Yes	Yes	No	Yes	Yes	No
Firm size controls	Yes	Yes	No	Yes	Yes	No
Firm effects	No	No	Yes	No	No	Yes
Observations	63260	95499	135431	80808	116690	165648
Nb. individuals	15721	19932	19932	18012	21599	21599
Nb. firms	21513	31189	42937	25770	36408	49556
R^2	0.294	0.721	0.829	0.368	0.764	0.854

Clustered standard errors at the individual level in parentheses

Industry dummies: 39 two-digit dummies (NACE)

Firm size controls: 12 dummies

Occupational dummies: 38 two-digit dummies (PCS-ESE)

A Appendix: data

Cleaning

I proceed to some cleaning of the DADS panel. First I recode the age variable as the difference between the current year and the year of birth. The former age variable exhibits some errors due to scan problems before the numerical DADS was introduced. Second, *département* codes are sometimes one-digit instead of being two-digit; other *département* or region codes are missing. In that case I rely on other observations in the whole database in order to recover that information.

In the EDP database, I eliminate observations for which days or months of marriage or birth are equal either to 00 or 99, as well as observations for which the year of birth is 0000.

Selection

I restrict my attention to individuals born on October of even-numbered years: careers of individuals born on October of odd-numbered years is unknown before 2002. The most important selection is dictated by the necessity of measuring experience properly (see *infra*): I focus on individuals who entered the panel after 1995, which leaves me with 46,280 individuals (338,879 observations at the individual-year level and 489,852 observations at the individual-firm-year level). We eliminate further individuals whose net annual earnings are missing or less than 10 euros in 2011 terms. I also restrict my sample to individuals aged 16 to 65, working at least 10 hours a year, whose job duration is consistent with worked hours (for instance, the ratio of the latter over the former must be less than 24), which leaves me with 45,483 individuals (317,476 individual-year observations). After trimming observations with a hourly wage that is smaller than 80% of the legal minimum wage,[16] and after dropping years 2003 to 2005, my estimation sample is composed of 41,531 individuals (212,189 individual-year observations and 301,079 individual-firm-year observations). Among those individuals, 19,932 are women while 21,599 are men. Last but not least, I define time-varying variables for marriage (parenthood) as the fact of being married (experiencing a childbirth) before time t for individual i.

[16] I proceed to robustness checks with respect to the 80% threshold in Section 5.3.

Definition of main employment

Aggregating data at the individual-year level requires to define for each individual her main employment in the year. I select the employment with (in successive order) the highest number of working days, the highest wage, a full-time position (if any) and the highest number of worked hours. If there are still ties after applying those criteria, I choose the job with the last SIREN in lexicographical order –to keep the code deterministic. Finally, if several observations resisted to the last iteration, I would consider them as authentic doubles and eliminate them –which does not happen here. We define job characteristics (private/public sector, industry, geographic location, firm's size, full-time/part-time, but also seniority) at the individual-year level as being related to the main employment. I sum wages and working hours, and define working days as the minimum of 360 (the annual number of working days in the DADS by convention) and the sum of working days over the whole year.

Computation of experience

Mincer (1958) demonstrated how important it is to control properly for experience and seniority in wage equations. I devote much attention to compute these variables as precisely as possible. Seniority is defined as the difference between the current date and the first appearance of a pair (individual, firm). Thanks to the comprehensive nature of the DADS panel, it is possible to reconstitute the whole salaried career of an individual, hence to compute his experience from observed working times. Experience will thus be defined as closely as possible as the amount of salaried time spent on the labor market. Since worked hours have been available from 1995 onwards only, I restrict my attention to individuals who entered the panel after 1995. I consider that workers increase their full-time/part-time experience variable every year by their share of working hours expressed in full-time units (FTU).