« The impact on wages and worked hours of childbirth in France »

Auteurs

Bruno Rodrigues, Vincent Vergnat

Document de Travail n° 2016 – 48

Novembre 2016
The impact on wages and worked hours of childbirth in France
Bruno Rodrigues, Vincent Vergnat∗
November 2, 2016

Abstract
Using French administrative data, we estimate the impact of the birth of a first, second and third child on hourly wages, as well as for hours worked, for both women and men. We compute the impact on these outcome variables, two, four and six years after the birth of the child, and focus on the distinction between highly educated women and women with a high school degree or less. We also take the maternity leave (or paternity leave in case of men) duration into account. Estimation is done with difference-in-differences and we compute bootstrapped confidence intervals. Results show both lower and highly educated women decrease significantly their working hours after the birth of their child. Men are, for the most part, not much impacted by the birth of their children. Maternity leave duration influences the magnitude of the impact of the birth, especially on the hourly wages of educated women.

Keywords: Fertility decisions, Labour Supply, Difference in Differences, Family pay gap

JEL classification: D10, J13

∗Rodrigues (corresponding author): Institut national de la statistique et des études économiques du Grand-Duché du Luxembourg (STATEC), and Agence pour la normalisation et l’économie de la connaissance (ANEC), 19-21, boulevard Royal, L-2449, Luxembourg and Bureau d’Économie Théorique et Appliquée (BETA), Université de Strasbourg, CNRS, 61 avenue de la Forêt Noire. Email: bruno.rodrigues@statec.etat.lu Vergnat: Bureau d’Économie Théorique et Appliquée (BETA), Université de Strasbourg, CNRS, 61 avenue de la Forêt Noire. Email: vincent.vergnat@unistra.fr We would like to thank the participants of the Strasbourg-Konstanz workshop in applied econometrics, in particular Winfried Pohlmeier, as well as Bertrand Krebel. We also thank Hans-Martin von Gaudecker, François Laisney, Mathieu Lefebvre, Marie Blaise and Holger Stichnoth for helpful comments. We would also like to thank the organizers and participants of the WIEM 2016 conference in Warsaw, Poland. The access to the data was carried through the CASD dedicated to researchers authorized by the French Comité du secret statistique. This work was supported by the ANR FamPol project, grant ANR-14-FRAL-0007 of the French Agence Nationale de la Recherche.
1 Introduction

Low fertility is a phenomenon to which a lot of major industrialised countries are confronted. France seems to be an exception to this rule however. In France, fertility rates are high relative to other industrialised countries such as Germany, but still below the replacement rate of 2.1. This low fertility rate in industrialised countries may be linked to the emancipation of women. Gayle and Miller [2006] explain that in the USA, a typical woman had four children at the beginning of the 20th century, but that number had decreased to 1.9 at the end of the last century. This decrease in childbearing has also been accompanied by an increase in female labour supply. Still according to Gayle and Miller [2006], “the participation of all wives increased by 36% over the last 25 years, the rates of mothers with children under the age of three increased by 83%, and by 91% for women with children one year old or younger” (p. 2). Women can now have dynamic careers just like men, but evidence shows that it is still mostly mothers who decrease their labour supply to take care of their children. In France, this is the case as shown by Pailhé and Solaz [2007]. In the survey Families and employers conducted by the INSEE and the INED and exploited by Pailhé and Solaz [2007], after having their first child, 22% of women perform a professional transition linked to a birth against only 5% for men. These transitions can slow down women’s careers and the literature on female labour supply and fertility decisions give a clear picture of this phenomenon. As an example for Australia, Baxter et al. [2008] show that women increase their housework hours greatly after giving birth, which is not the case for men, and men actually increase their labour supply for higher order births.

This negative impact on a woman’s career would probably be more important for educated women, who had invested more in human capital and therefore had a more lucrative career in front of them. For example, Adda et al. [forthcoming] show that in Germany, women with abstract occupations have a much higher skill atrophy rate than women with less abstract occupations (at most 6.9% for women in abstract occupations versus 0.6% for women in routine occupations). These abstract occupations are in general demanded for jobs that require higher degrees. These highly educated women have also, on average, less children. Adda et al. [forthcoming] also estimate the amenity values of different occupations and show that abstract occupations have a very low amenity value when children are present. They show that if abstract occupations had the same amenity value that routine occupations, the number of women that would choose to work in abstract occupations would increase by 5%. Francesconi [2002] studies the fertility and labour supply decisions of young married women. In his simulations, Francesconi [2002] shows that “increased schooling decreases the expected number of children substantially” (p. 367) and that lower wages increase fertility. These findings seem also to indicate that it is mostly highly educated women that have low fertility and that having a child is more costly for them than for lower educated women.

1The INSEE is the national institute of statistics and economic studies, the INED is the national institute for demographic studies
This paper focuses on France, which has a higher fertility rate than most other industrialized countries and which has very generous social policies. In this paragraph, we shall review these policies to give the reader an idea of the current French situation. Already in 1985, the *Allocation Parentale d’Éducation* (APE) allowed parents of 3 or more children to receive a lump sum if they reduced their labour supply to take care of their children: either by completely stopping working, or by working less hours. However, one condition required that the youngest child was under 3 years old. In 1994, the APE had been extended to parents of two children. The impact of this reform has been studied by Lequien [2012]. In 2004, the APE was replaced by the *Complément de libre choix d’activité* (CLCA). The CLCA is also a lump sum, which depends on family resources, given to parents that completely stop, or reduce their labour supply to take care of a child, that must be younger than 3 years old. For their first child, parents received the CLCA during 6 months, and from their second child, they got the CLCA until the third birthday of the last child. It is mostly women who claim this allowance as shown by Boyer and Nicolas [2012]. Boyer and Nicolas [2012] and Joseph et al. [2013] show that the proportion of working women reducing their labour supply after the first birth has increased and the impact of the reform depends on the education level of the mother. To incite fathers to stay at home and increase the labour supply of women, the CLCA was replaced in 2015 with the *Prestation partagée d’éducation de l’enfant* (PreParE). The PreParE works in a different fashion than the CLCA: parents both get a lump sum for 6 months until the first birthday of their child. But one parent alone cannot stay 1 year at home and claim the benefits of his/her partner (or reduce his/her labour supply to part-time work for 1 year, except for lone parents). This incites both parents to either stay at home with their newborn child, or to both reduce their labour supply.

The reduction of the labour supply has an impact on the careers of women which can potentially create a wage gap between mothers and women without children: the family wage gap. During the last two decades, increasing attention has been given to the family wage gap. The literature has been concerned with explaining this family wage gap. For example, Waldfogel [1997] is one of the first contributions to this literature. The author shows, using American data, that labour market experience is not the only explanation for the family wage gap. The author tests two other possible explanations: unobserved heterogeneity and part-time employment. Unobserved heterogeneity seems to be not an important factor to explain the family wage gap, unlike part-time employment. Even after controlling for these factors, an important part of family wage gap remains unexplained. The family wage gap can also be observed in other industrialised countries; Gangl and Ziefle [2009] estimate wage equations for the UK, Germany and the United States of America and show that motherhood is associated with wage penalties ranging from 9% to 18%. German women are especially penalized, because they tend to take long childcare breaks, which is not the case for British and American women. The wage penalty for German women is estimated to be around 16% to 18% in hourly wages, 13% for British women and 9% for American women. Gangl and Ziefle [2009] also show that women
with children invested less in education than comparable childless women and that women tend to favour child-friendly occupations. Davies and Pierre [2005] focused on the family gap in European countries. To compare the impact of motherhood on earnings in Europe, the authors used the ECHP to estimate wage equations for 11 countries. They show that the size in penalties in pay are different across countries and depend also on the number of children and the timing of the first birth. They find significant penalties on the wage in many European countries like Germany, the United Kingdom or Denmark. For France, they find a wage penalty of 10% after controlling for selection, but only for mothers of more than 3 children.

The family wage gap could also be applicable for men. This issue was studied, for example, by Lundberg and Rose [2000]. They study the impact of a birth on the wage and worked hours of married men and women. They show that the negative family wage gap for women depends of the duration of maternity leave and that a birth has a positive impact on the hourly wage of men. Other papers confirm this last result (Lundberg and Rose [2002], Glauber [2008], Hodges and Budig [2010]). Killewald [2013] estimates wage equations using the NLSY 1979 and finds that the wage premium for fathers depends on the family context, namely that biological fathers living with the mother of their children gain around 4% in hourly wages, but unmarried fathers, or stepfathers, do not. However, this 4% wage premium decreases to 1.3% for married, residential fathers who are married to women working full-time. In couples where both the husband and the wife are working full time, specialization cannot occur. Thus, the author argues, these fathers have also household responsibilities which makes it difficult for them to commit more to their careers and thus increase their wages.

The studies that have focused on France are Davies and Pierre [2005], discussed above, Meurs et al. [2010], Duguet et al. [2015] and Wilner [2016]. Meurs et al. [2010] use the French “Families and Employers” survey to study the impact of children and duration of maternity leave on the gender wage gap. Their results show that having a child creates no direct pay penalties for women and a bonus for men. Moreover, having a child has an indirect negative impact on the hourly wage of women through the reduction of labour supply (part-time job or time out of labour force to take care of children).

This paper contributes to this literature by studying the impact of a child’s birth on women’s and men’s hourly wages and supplied hours, using difference-in-differences estimations. Previous literature on female labour supply and human capital accumulation shows that the impact of having a child is very different for women with different levels of human capital. Thus, we analyse the effect of a child’s birth for different educational attainments. We also take the maternity leave duration into account. Evidence from the literature suggests that the impact on our variables of interest increases with education; higher educated women (men) suffer from a larger penalty (reward) than lower educated ones. For example, Mincer and Polachek [1974] find that the human capital depreciation rate for women that stay at home after the birth of a child is -

---

ECHP: European Community Household Panel Survey
1.1% for women with elementary schooling or less, -1.4% for women with some high school, and -4.3% for women with at least some college education. Human capital depreciation rate is then translated by a loss in hourly wages.

Unlike other studies on France, which estimate wage equations, Duguet et al. [2015] use difference-in-differences estimates to study the family wage gap of women and men working either in the private sector or in the public sector. Duguet et al. [2015] find that for women in the public sector, the impact on the wage is a loss from about 3.5% to 6.5%, while it is 9.1% for women in the private sector, a result in line with the findings of Lequien [2012]. Three years later, women still earn less, from about 0.9% to 2.4% less than their initial wage. Worked hours also decrease. Duguet et al. [2015] also show that for men, having a child is associated with an increase in wages but a decrease in hours worked. Wilner [2016] uses French administrative data to test whether the self-selection of women that expect to have children into low wage firms could explain the family wage gap. After controlling for this selection as well as for unobserved heterogeneity and human capital, the author finds that mothers had a penalty in hourly wage of approximately -2.2% per child and fathers do not enjoy any loss or premium.

Unlike these papers, we focus on the role of education level and maternity (parenthood) leave duration on family wage gap. The analysis finds that having a child reduce the labour market participation of educated and non-educated women but not of men. The birth of the first child have also a negative impact on hourly wage of high educated women who take a long maternity leave duration. As Wilner [2016], we find no premium in hourly wage for men.

This paper is structured as follows: section 2 presents the data set used, as well as the econometric methodology. Section 5 discusses the results and finally section 6 concludes the paper.

2 Data and econometric methodology

2.1 The data used

The data set used in this paper is called the "DADS-EDP" data set and is provided by the INSEE. This data set is actually composed of two other sources: the DADS panel and the EDP merged to form the DADS-EDP. The DADS is an administrative data set with information on wages, the type of employment contract, employment sector, the size of the firm the person is working in, the starting and closing dates of the period of paid work, the number of paid hours, and much more. Each year, firms have to make a declaration for each of their employees. Every working person in France is covered by these declarations, except for employees of government bodies, self-employed people and employees of French firms abroad. Civil servants working in public institutions of an

3DADS stands for Déclaration Annuelle des Données Sociales, or Annual Declaration of Social data. EDP stands for Échantillon Démographique Permanent, or Permanant Demographic sample.
industrial and commercial nature are included in the DADS (since 1991 and 1992) as well as publicly-employed hospital staff (since 1984), civil servants of territorial communities (since 1988), unemployment benefits recipients (since 2002) and agricultural workers (since 2003). There exists different versions of the DADS, for instance a version that includes every civil servant. We have access to the panel version of the DADS (from 1976 to 2010), which is a 1/25th sample of the DADS until 2001 (we have employees born in October in an even year). Since 2002, the sample size was doubled.

The EDP is a panel with information on marital status, fertility, degrees obtained and the place of residence. From 1967, to 2004, people born from the 1st to the 4th of October are in the EDP. Since 2004, the data set was enriched with the individuals born from the 2th to the 5th of January, 1st and 4th April and July. The data are gathered from civil registries each year, and also from the census, whenever needed to complete the information from the civil registries. For example, for people born between 1989 to 1997, the information on children comes exclusively from the census. For most people born between 1982 and 1989, this is also mostly the case. Before 2004, only people living in continental France were in the EDP. Now, people living in the French overseas territories are also included.

The merged DADS-EDP data to which we have access through the CASD are composed of individuals born on EDP days that are also in the DADS panel. Therefore, the data set does not include civil servants of national public services, men or women who have never worked and self-employed people. People born abroad are not included in the data.

2.2 Data preparation
Before using the data for analysis, a lot of data preparation and cleaning was done. The raw data is in spell format, not very well suited for analysis. This short subsection explains the operations we performed on the data. We did the same operations for both women and men, so in this subsection we only describe the steps for women.

First of all, we had to order the births of the children. In the description of the data it was written that for each child the variable aeni gave the birth year of the ith child. One would have assumed that aen1 was thus the year of birth of the first child, so the oldest, and aen2 was the year of birth of the second child and so on. However, this turned out to be wrong. Instead, aen1 was always the year of birth of the youngest child, while the year of birth of the oldest child was aen2 for a woman who had two children, but for a woman who had 3 children it was then aen3. Something as simple as computing the mean age of childbearing by cohort for the first child was thus impossible to do with a single line of code. The first step was thus to order the births so that the first

4 Actually, we only know the wedding date of a person. We do not know if that person is single, divorced, in a civil union (such as the French PACS), or is a widower.
5 These days are called EDP days.
child, and the oldest, was in \text{aen1}, the second child to be born to a woman in variable \text{aen2} and so on.

The second step was to remove obvious errors; for example, we had some women that were born after their children. We completely removed such lines. We filled up incomplete data if possible: for example if \text{starting date of contract} and the \text{duration of payment} were both available, but not the \text{ending date of contract}, it was easy to deduct the \text{ending date of contract}.

We then created two variables that gave the age in years of the individuals (current year minus year of birth) and the number of children someone had at a given year. Until then we only had the total number of children someone ever had.

A methodological contribution of this paper consists in creating a variable that counted the number of days a person stayed out of the labour force due to having a child. Since the data is in spell format, and also because the starting and ending days of work contracts are reported, it was possible for us to compute this variable. This variable is useful to compare the impact of the birth of a child for women that had short maternity leaves (less than 6 months) to mothers who took longer maternity leaves.

We removed the parallel spells by defining the main employment for each woman and for each year. If a woman has had more than one job in a given year, we only kept the one where she worked the most days. If there were two jobs (or more) where she had worked an equal number of days in, we kept the job with the highest wage. For each job, we also had the worked hours, the number of days worked as well as wages. We summed up all these variables together and finally removed every other spell. After these operations, a line in our data set gives, for a given year, a woman’s total number of days worked, the total of the earned wages and only her main job.

We also created a variable giving the attained education level in five categories: no information on education, less than high school education, completed high school (or similar degree), 2-3 years of higher education (university or similar institution) and 4-8 years of higher education.\footnote{The original data gave much more detailed information, especially for lower degrees that are not relevant for this study. Later on, for the difference-in-differences analysis, we grouped these 5 categories into 2 coarse categories, “at most a High school degree”, “more than a High school degree”.

Finally we created hourly wages by dividing the annual wage by the number of hours worked in the year.

\subsection{2.3 Econometric methodology}

In order to identify the impact of the birth of a child on hourly wages and the number of hours worked, we use a standard DiD approach. The following description is fairly standard and can be skipped for readers familiar with DiD.

Card and Krueger [1994] use the DiD method to study the impact of the raise of the minimum wage in New Jersey that occurred on the 1st of April 1992. DiD works by considering two groups, a control group and a treated...
In experiments, creating such groups is relatively easy: it suffices to assign the treatment (for example, taking a certain medication) randomly. The people that were randomly selected to get the treatment thus become the treated group.

In economics however, treatment is rarely assigned randomly. One famous example of randomly assigned treatments are the Vietnam lotteries. Angrist [1990] studies the impact of these lotteries on the wages of men and shows that “[...] as much as ten years after their discharge from service, white veterans who served at the close of the Vietnam era earned substantially less than non-veterans” (p. 330).

In cases were random treatment is not possible, quasi-experimental methods such as DiD have proved to provide consistent estimates of the parameters of interest. The problem econometricians face is that they do not observe what would have been the outcome variable (for example, hourly wages) for the control group if it were treated, because it might be that people self-selected into the treatment group. Let us illustrate how DiD works. Let $y^C_{it} | D = 0$ be the outcome variable for individual $i$ at time $t$ for the control group. $D$ is a dummy variable that equals 1 if the individual is treated and 0 if not. Consider the following quantities:

$$E[y^C_{it} | D = 0]$$
$$E[y^T_{it} | D = 1]$$

As explained previously, we do not observe $E[y^C_{it} | D = 1]$ for the control group, meaning that we do not observe what the outcome variable $y^C_{it}$ would have been for the control group if its members were treated. We only observe this for the treated group. Furthermore, let us assume that we can decompose the quantities in (1) in a time-invariant and a time-variant component:

$$E[y^C_{it} | D = 0] = \gamma^C_i + \lambda^C_t$$
$$E[y^T_{it} | D = 1] = \gamma^T_i + \lambda^T_t + \beta$$

where $\beta$ is a constant that captures the effect of the treatment on the outcome variable. Thus it is $\beta$ that the econometrician wants to estimate, net of disturbance parameters (the $\lambda$’s). Let us suppose that the econometrician computes the following quantity:

$$E[y^T_{it} | D = 1] - E[y^C_{it} | D = 0] = \gamma^T_i + \lambda^T_t + \beta - (\gamma^C_i + \lambda^C_t)$$

If the econometrician assumes that the individuals composing the control group are exactly the same as the individuals composing the treatment group, the time-variant effects cancel each other:

$$E[y^T_{it} | D = 1] - E[y^C_{it} | D = 0] = \gamma^T_i - \gamma^C_i + \beta$$
This can be seen by plotting the outcome variables for both groups before and after the treatment, and checking if the plotted curves have a common trend before treatment. This common trend hypothesis is important and is what makes DiD work. In the Appendix 7.2 these “common trend graphs” are showed. To get rid of the time-invariant effects, the econometrician needs to observe both groups again: after the treatment has taken place. He can then compute the following quantity:

\[
E[y_{it+1} | D = 1] - E[y_{it+1} | D = 0] - (E[y_{it} | D = 1] - E[y_{it} | D = 0]) (6)
= \gamma_T^i - \gamma_C^i - (\gamma_T^i - \gamma_C^i) + \beta
= \beta,
\]

where \( \beta \) can be either positive, negative or statistically 0, in which case the treatment did not have any effect. Estimation of \( \beta \) can be achieved by replacing the expected values by their empirical counterparts. In this paper, \( \beta \) is estimated using (6) and then we bootstrap the confidence intervals.

### 3 Summary statistics and graphs

There are 409'107 individuals in total in the panel (1976-2010), from which 194'956 are women and 214'151 are men.

Table 1: Age at which women and men are first observed, for older cohorts

<table>
<thead>
<tr>
<th>Date of birth</th>
<th>1934 - 1943</th>
<th>1944 - 1953</th>
<th>1954 - 1963</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>Mean</td>
<td>38.23</td>
<td>37.65</td>
<td>29.58</td>
</tr>
<tr>
<td>1st quartile</td>
<td>36</td>
<td>35</td>
<td>26</td>
</tr>
<tr>
<td>Median</td>
<td>40</td>
<td>40</td>
<td>32</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>38</td>
<td>38</td>
<td>28</td>
</tr>
<tr>
<td>Observations</td>
<td>46711</td>
<td>85205</td>
<td>179592</td>
</tr>
</tbody>
</table>

Table 1 and 2 give the mean age of the first and last observation for older cohorts, for women and men. People in the data set are first observed when they start working and are last observed whenever they leave the country, go into retirement, or die.

Table 3 and 4 give the same information as Tables 1 and 2 but for younger cohorts. The cohorts with the most people are the 1954-1963, 1964-1973 and the 1974-1983 cohorts, which is not surprising as these cohorts are those that would suffer the less from left and right censoring.
Table 2: Age at which women and men are last observed, for older cohorts

<table>
<thead>
<tr>
<th>Date of birth</th>
<th>1934 - 1943</th>
<th>1944 - 1953</th>
<th>1954 - 1963</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>Mean</td>
<td>47.99</td>
<td>48.16</td>
<td>47.68</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.27</td>
<td>2.03</td>
<td>3.67</td>
</tr>
<tr>
<td>1st quartile</td>
<td>48</td>
<td>48</td>
<td>49</td>
</tr>
<tr>
<td>Median</td>
<td>49</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>49</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>Observations</td>
<td>46711</td>
<td>85205</td>
<td>179592</td>
</tr>
</tbody>
</table>

Table 3: Age at which women and men are first observed, for younger cohorts

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>6.67</td>
<td>6.72</td>
<td>3.15</td>
</tr>
<tr>
<td>1st quartile</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Median</td>
<td>31</td>
<td>31</td>
<td>23</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>21</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>Observations</td>
<td>475092</td>
<td>551005</td>
<td>348546</td>
</tr>
</tbody>
</table>

Table 4: Age at which women and men are last observed, for younger cohorts

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>Mean</td>
<td>41.06</td>
<td>41.00</td>
<td>31.27</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.94</td>
<td>3.93</td>
<td>3.26</td>
</tr>
<tr>
<td>1st quartile</td>
<td>39</td>
<td>38</td>
<td>29</td>
</tr>
<tr>
<td>Median</td>
<td>44</td>
<td>44</td>
<td>34</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>42</td>
<td>42</td>
<td>31</td>
</tr>
<tr>
<td>Observations</td>
<td>475092</td>
<td>551005</td>
<td>348546</td>
</tr>
</tbody>
</table>
In figure 1 we see the number of people in the panel for a given year, and by cohort. Figure 1 also shows the expansion of the DADS panel, as well as the years for which the panel is not available, 1981, 1983 and 1990. We also observe two dips, one in 1994 and one in 2003-2004. The first is due to technical problems met by the INSEE in the years 1993 and 1994 and the second to a significant number of DADS identifiers being wrong or unreported.
Figure 2 shows the fertility rate in France. As can be seen, the true fertility rate is underestimated by our data. The dotted curve is the fertility rate estimated by the World Bank, while the solid curve is the fertility rate estimated by the DADS-EDP. The trend, however, is the same for both curves. Also, as mentioned above, for some years, the information on fertility is missing and completed with data from the census. The problem with this, is that this only counts children currently living with their parents. Children that already left home are not included in the census. Another issue is that we do not have all civil servants in the data and women who do not work at all during their life.
Figures 3 and 4 show the age at the first, second, third and fourth birth for women and men respectively. Confidence intervals are represented by the grey

areas. Ignoring the first and last cohorts, which suffer from censoring, we see that the age at which people have their children has increased.

Figures 5 and 6 show the timing of births for different cohorts. The same conclusions can be drawn from these figures and from figure 3 and 4: younger cohorts are having their children later.

Figure 5: Timing of births for women

![Figure 5: Timing of births for women](image)

Figure 6: Timing of births for men

![Figure 6: Timing of births for men](image)
Figures 7 and 8 are both made using the "central cohorts"; due to censoring and low number of observations, we removed the two oldest cohorts and the youngest cohort. These figures suggest that women time childbearing conditionally on their education level, which does not seem to be the case for men.
Figures 9 and 10 show the yearly net wage by education level, again for both women and men. We can clearly see that higher educated women are paid higher wages. The same can be said for men, but men earn much higher wages than women.
Figures 11 and 12 show the yearly net wage by number of children. The more children a woman has, the less she seems to earn. There seems to be a difference between having no, one or two children for women younger than 40,
but then these women catch up to their peers with less children. Starting with three children however, women never seem to catch up. Interestingly, before the women reach the age of 25, there does not seem to be much selection effect. For men, having children seems to be correlated with higher yearly wages as can be seen on figure 12. The reader may wonder why the graphs show that there are 15 year olds with 4 to 8 years of university education. This is because we only have the final education level in our data. For the purposes of the estimations, we removed people that were too young to have such high degrees.

Figure 13: Share of part-time workers by number of children

Figure 13 shows the share of part-time workers by number of children. This graph needs some explanation to be understood; the numbers are the number of children women (dashed curve) and men (solid curve) have. Throughout the years this share has remained somewhat stable for both men and women, but as can be seen, the more children women have, the more likely it is that they work part-time. In recent years especially, we see that even after only one child, women seem to decrease their labour supply. For men however, the number of children does not seem to have any impact on part-time work, apart for 0 children. Maybe this is because men that do not have children tend to be young and thus work part-time to pay for their studies for example. This is just an hypothesis; we do not pretend that we are able to “read” such an explanation solely from looking at this graph.
Figures 14 and 15 show the number of hours worked in a year for women and men by education level. Higher educated people seem to work more hours than lower educated ones.
Figures 16 and 17 show the number of hours worked in a year for women and men by number of children. These last two graphs are very interesting; indeed,
having more children seems to be correlated with less worked hours for women have, but this is reversed for men.

What do all these graphs and summary statistics tell us? All these graphs do seem to go in the same direction as the evidence from the literature. For women, higher education is correlated with later childbearing, having more children is correlated with lower yearly wages. For men, having children does not seem to carry a penalty, quite the contrary. The graphs suggest that having children might be associated with higher yearly wages. As for hours worked, women reduce their labour supply the more children they have, while labour supply seems to increase for men. In the next section, we show how we estimate the impact of the birth of one, two and three children on the following outcome variables: hourly wages and hours worked, but we discriminate by education level and by maternity or paternity leave duration. As stated in the introduction, the literature suggests that higher educated women suffer from a higher penalty in wages. We hypothesise that highly educated women have their children later because they study longer, and thus suffer from a penalty in hourly wages. This paper does not try to give an answer as to why there could be a penalty in hourly wages, but the literature on female labour supply tells us that it may be due to a higher depreciation rate of human capital for highly educated women. In future work, we plan on testing this hypothesis for France using our data and a structural life-cycle model of labour supply.

4 Control and treatment groups

To analyse the impact of the birth of a child on hours worked and on the hourly wage of a woman or a man, we chose to run a difference-in-differences estimation. This allows us to not have to specify a functional form for the wage equation and also deals with unobserved heterogeneity. It is thus important to discuss how we constructed the control and treatment groups. We only kept individuals that had a birth in either 2002 or 2003. We selected both these years in order to increase the number of observations for our analysis. Then, we observed what the hourly wages and worked hours were for these individuals two years before treatment. This counts as our first observation. Then we also observed them again two, four and six years after treatment. The reader might wonder why we chose the years 2002 and 2003 as the treatment years. The first reason is that we wanted to use the most recent data at our disposal. The second reason is the requirement that nothing but the treatment could influence the variables of interest. So if we focus on the 2000’s, our options are actually very limited. In the year 2000, the socialist government at the time in France introduced the 35-hour workweek. Before this reform, the normal workweek in France was 39 hours long, but after the reform it decreased to 35 hours. We had to be careful not to select individuals before and after this reform, or else we would have attributed decrease in hours due to this reform to the birth of a child wrongly.

Another reform was implemented in 2004, the CLCA reform described in the introduction. Here again, we had to be careful, as this reform could have

21
changed the behaviour of people. Since we also wanted to see the labour supply
and hourly wages of treated individuals six years after treatment, we could not
have chosen the year 2005 (or later) as a treatment year, as our data only
goes until 2010. Using data from the 2000’s is also a way to insure that the
information of birth is fairly accurate. Indeed, from 1989 to 1997, and for people
born between 1982 and 1989, the information on the birth of their children is
limited as it comes only from the census and not from civil registries.

To create our control and treatment groups we separated the individuals by
education levels, but our five categories that we used in the previous sections
were too fine-grained so we did not have many individuals in the groups. We
decided to pool individuals together into two categories: having at most a high
school degree (or similar) or having more than a high school degree (such as
any university degree or other types of degrees such as the French BTS). We
also made sure to only keep women and men of a certain age, as the education
variable is highest education level achieved. Thus we removed people that were,
for example, 17 years old in the year 2000 with eight years of college education.
We also made sure that people in groups would not have another child before
2011. We also separated the treatment and control groups further into women
(men) that stayed out of the labour force for less than 6 months, for more than
6 months but less than 2 and a half years and women (men) that stayed
out of the labour force for more than 6 months but less than 6 years. This
allows us to distinguish the impact for people that stay out of the labour force
for different lengths of time.

Then, to analyse the impact of having a first child, we proceeded in a similar
manner to Duguet et al. [2015]: the control group is composed of women without
child. For the impact of a second child, the control group is thus women with
one children and so on.

After create control and treatment groups, we also have to make sure that
those in the control group only differ from those in the treatment group by
the fact of not being treated. This means that the individuals composing these
groups should be as homogeneous as possible. We thus made sure that the
age structure between these groups was similar. We wanted to avoid to have
actually older people in our treatment group than in our control group. So we
created hundreds of control and treatment groups by selecting individuals that
were of a certain age in the year of the first observation. For each pair of group
created, we compared the distribution of age using a Kolmogorov-Smirnov test
and a simple t-test to compare the means of ages in both groups and only kept
the groups were at least one of these tests passed. In most cases, however, both
tests passed.

After all these steps we ran our DiD analysis for each subgroup. Our vari-
ables of interest were number of hours worked in a year and hourly wages. We
bootstrap confidence intervals using Efron’s bootstrap percentile t method as
described in Chernick [2007]. Efron’s bootstrap percentile t method is simple
to implement and is second-order accurate.
5 Results and discussion

In this section, we present the results of our study. Let us summarise the previous sections before discussing the results.

To assess the impact of the birth of one, two and three children on worked yearly hours as well as on hourly wages, we use a DiD approach on administrative data from France. To avoid several issues, we chose the year 2002 as the treatment year. The issues avoided are two-fold: in the year 2000, France introduced the 35-hour workweek and in 2004 there was a maternity leave reform. So we had to be careful not to choose the treatment year in a year where the first observation would lie before the reform of the 35-hour workweek, while the second observation would lie after it, lest we would have attributed a decrease in worked hours to childbearing while in reality it would have been due to the reform. The second issue we avoid is the 2004 CLCA reform described in the introduction. This reform would have also tainted our results.

Then, to have more observations, we also selected individuals that had a birth in 2003, and not in 2002. We pooled the individuals that were treated in 2002 and 2003 together. We assume that having a child in 2002 and 2003 is very similar, and thus that pooling the individuals is not a problem.

We also made sure that the distribution of ages was the same between the treatment and control groups. Our very preliminary results showed significant differences in worked hours between the treatment and the control group, but this was due to a different age distribution between groups. The treatment groups systematically contained older individuals. To correct this, we created hundred of groups with different individuals and then used the Kolmogorov-Smirnov test on the empirical distributions of age for both groups as well as a simple t-test to compare the means of age in both groups. Finally, we only kept control and treatment groups were at least one of the tests passed, but in most cases, both tests passed. This means that we do not have many observations in each group, but we are fairly confident in the quality of the groups. In the tables below, a single star (*) next to a result means that the result is significant at the 10% level, and two stars (**) next to a result means that the result is significant to the 5% level.
Table 5: Estimation of the impact of the birth of a child on the hourly wages of women. 5% and 10% bootstrapped confidence intervals have been computed using 1000 replications, and are shown below the estimates. “Short”, “Intermediate” and “Long” maternity leaves stand for “less than 6 months out of the labour force after giving birth”, “more than 6 months but less than 2 and a half years” and “more than 6 months but less than 6 years” respectively.

Tables 5 and 6 show the impact of the birth of a child on hourly wages as well as yearly worked hours. The impact is shown in 2004, 2006 and 2008, so two, four and six years after the birth of the child. We further distinguished between women that went on maternity leave for less than 6 months, more than 6 months but less than 2 years and a half and finally for more than 6 months but less than 6 years. The results are in 2007€. For women with at most a high school degree, the results indicate no penalty in hourly wage. For women with more than a high school degree, we see a negative impact, but only for women that stayed more than 6 months out of the labour force. Women that stayed less than 6 months do not suffer a penalty. This might be evidence of human capital depreciation that only occurs if women stay out of the labour force long enough.

Hours decrease significantly for both educated and non-educated women in the short and in the long term. The impact on labour supply of birth does not seem transitory.
<table>
<thead>
<tr>
<th>Year of 2nd observation</th>
<th>Maternity Leave Duration</th>
<th>Less than High School Estimate</th>
<th>More than High School Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Short</td>
<td>$-151.80^{**}$</td>
<td>$-173.78^{**}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-239.82, -67.02]$</td>
<td>$[-276.38, -73.85]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-227.32, -80.91]$</td>
<td>$[-264.43, -88.00]$</td>
</tr>
<tr>
<td>2006</td>
<td>Short</td>
<td>$-191.22^{**}$</td>
<td>$-221.86^{**}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-284.62, -101.74]$</td>
<td>$[-326.78, -117.08]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-269.22, -114.90]$</td>
<td>$[-313.14, -133.32]$</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>$-201.74^{**}$</td>
<td>$-193.11$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-333.10, -60.64]$</td>
<td>$[-437.57, 59.18]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-315.21, -82.11]$</td>
<td>$[-399.55, 20.85]$</td>
</tr>
<tr>
<td>2008</td>
<td>Short</td>
<td>$-197.41^{**}$</td>
<td>$-229.64^{**}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-304.48, -94.21]$</td>
<td>$[-332.39, -131.57]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-288.35, -108.12]$</td>
<td>$[-319.09, -146.19]$</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>$-161.37^{*}$</td>
<td>$-177.16^{*}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-318.09, 1.42]$</td>
<td>$[-361.76, 8.88]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-299.07, -20.56]$</td>
<td>$[-337.09, -23.48]$</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>$-162.63^{**}$</td>
<td>$-284.13^{**}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-282.15, -25.58]$</td>
<td>$[-487.74, -94.37]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[-267.45, -49.33]$</td>
<td>$[-456.23, -129.07]$</td>
</tr>
</tbody>
</table>

Table 6: Estimation of the impact of the birth of a child on the hourly wages of women. 5% and 10% boostrapped confidence intervals have been computed using 1000 replications, and are shown below the estimates. “Short”, “Intermediate” and “Long” maternity leaves stand for ”less than 6 months out of the labour force after giving birth”, ”more than 6 months but less than 2 and a half years” and ”more than 6 months but less than 6 years” respectively.
Table 7: Estimation of the impact of the birth of a second child on the hourly wages of women. 5% and 10% boostrapped confidence intervals have been computed using 1000 replications, and are shown below the estimates. “Short”, “Intermediate” and “Long” maternity leaves stand for “less than 6 months out of the labour force after giving birth”, “more than 6 months but less than 2 and a half years” and “more than 6 months but less than 6 years” respectively.

Tables 7 and 8 show the impact of having a second child, relative to women with one child already. Hourly wages seem to be mostly unaffected. There is a negative impact but only for lower educated women six years after treatment, and for highly educated women four years after treatment, and only for those that took a longer than six months maternity leave break. Two, four and six years after treatment, hours have decreased again slightly, but maybe surprisingly, only for women that took a short (less than six months) maternity leave break.
Table 8: Estimation of the impact of the birth of a second child on the hourly wages of women. 5% and 10% boostrapped confidence intervals have been computed using 1000 replications, and are shown below the estimates. “Short”, “Intermediate” and “Long” maternity leaves stand for ”less than 6 months out of the labour force after giving birth”, ”more than 6 months but less than 2 and a half years” and ”more than 6 months but less than 6 years” respectively.
Finally, tables 9 and 10 show the impact of having a third child, but only for women who took a short maternity leave break (there were very little women that had three children and took a long maternity leave). Hourly wages are unaffected, but supplied hours decrease yet again particularly two years after birth. The reduction in labour supply is also greater among less educated women, whatever the time horizon.

For the fathers, not many results were statistically significant. So instead of showing uninteresting tables, we simply discuss the few results that were significant.

A first birth significantly decreases the hours of work of fathers (by $-105.07^{**}$ hours, $[-210.67, -6.43]$) six years after birth and only for fathers who took no or a small paternity leave. For the second child, an effect appears only for men with low education who took a small paternity leave; hours of work increase by $131.52^{**}$, $([32.00, 233.84])$ two years after birth and by $96.47^{**}$, $([8.49, 191.55])$ four years after birth. Hourly wages remain unaffected. One explanation for this result is that low educated fathers can only increase their wages by increasing their labour supply.

It seems that there is still in the 2000’s a sharp specialization of tasks between husbands and wives. In the next paragraph, we propose a number of possible explanations for these results.

As pointed out by different studies, such as Meurs et al. [2010] or Wilner [2016], mothers have a greater probability to interrupt their careers or reduce their labour supply than childless women. This could also lead employers to choose to hire or promote men, which also penalizes women that do not want to have children. Therefore, mothers accumulate less human capital and if
<table>
<thead>
<tr>
<th>Year of observation</th>
<th>Maternity Leave Duration</th>
<th>Estimate</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Short</td>
<td>-561.66**</td>
<td>-426.58**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[-803.11, -312.87]</td>
<td>[-641.84, -294.00]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[-749.90, -367.61]</td>
<td>[-620.43, -246.82]</td>
</tr>
<tr>
<td>2006</td>
<td>Short</td>
<td>-353.17**</td>
<td>-195.49**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[-507.96, -205.69]</td>
<td>[-381.57, -7.91]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[-485.22, -221.48]</td>
<td>[-358.84, 39.00]</td>
</tr>
<tr>
<td>2008</td>
<td>Short</td>
<td>-443.14**</td>
<td>-156.55*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[-616.77, -285.21]</td>
<td>[-323.45, 6.12]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[-580.61, -309.15]</td>
<td>[-299.09, -23.48]</td>
</tr>
</tbody>
</table>

Table 10: Estimation of the impact of the birth of a third child on the hourly wages of women. 5% and 10% bootstrapped confidence intervals have been computed using 1000 replications, and are shown below the estimates. “Short”, “Intermediate” and “Long” maternity leaves stand for “less than 6 months out of the labour force after giving birth”, “more than 6 months but less than 2 and a half years” and “more than 6 months but less than 6 years” respectively.

they take a long maternity leave, the human capital they already accumulated depreciates. This “human capital” explanation could explain a great part of our results. Mothers could also choose a job with less constraints but with a smaller hourly wage and/or less professional opportunities. These family-friendly firms are in general low wage firms, as pointed out by Wilner [2016]. Mothers could also be less productive because they spend a lot of time and energy to take care of children.
6 Conclusion

Industrialised countries want to solve three seemingly contradicting problems; increase the labour market participation of women, as well as increase the fertility rates of families but also reduce the family pay gap. In this article, we sought to test the hypothesis that a birth influences negatively the career of a woman and positively the career of a man. We distinguish the impact by different education levels. We focus on France, a country with a long history of generous social policies and high fertility rates. Our results show that it is especially women with a higher education that decrease their supplied hours the most. Highly educated mothers also have a lower hourly wage compared to similar childless women. This could be evidence for human capital depreciation, but this is outside the scope of this paper. In future work, we plan on investigating human capital depreciation of highly educated women in France. Another important aspect that is novel in our paper is that we distinguish between women that take short maternity leaves from women that take longer maternity leaves. Our empirical results confirm that the loss in wages after the birth of a child is proportional to the education level of the mothers. This result raises the question of whether the compensation should be higher for this group of women (as is the case in Germany). These women suffer a greater loss in wages, and thus have less children, in most cases only one. This was not an important issue some decades ago, as highly educated women were not numerous, but nowadays, more and more women have university degrees.

Our results do not show that men receive a bonus after the birth of a child. These results are in line with those of Wilner [2016]. In some cases though, lower educated men do increase their supplied hours. This is probably to compensate the loss in wages of their wives.
Acknowledgments We would like to thank the participants of the Strasbourg-Konstanz workshop in applied econometrics, in particular Winfried Pohlmeier, as well as Bertrand Kœbel. We also thank Hans-Martin von Gaudecker, François Laisney, Mathieu Lefebvre, Marie Blaise and Holger Stichnoth for helpful comments. We would also like to thank the organizers and participants of the WIEM 2016 conference in Warsaw, Poland.

Compliance with ethical standards The access to the data was carried through the CASD dedicated to researchers authorized by the French Comité du secret statistique. This work was supported by the ANR FamPol project, grant ANR-14-FRAL-0007 of the French Agence Nationale de la Recherche.

Conflict of interest The authors declare that they have no conflict of interest.

7 Appendix

7.1 Appendix 1: data cleaning

Before selecting people for our study, we cleaned the DADS-EDP data (2010). The following list details exactly what we did:

- Creation of a variable: total number of children.
- Creation of a variable: cumulated number of children for each year.
- Pooling the education variable into 5 categories for the descriptive statistics:
  1. Missing information
  2. Low education (Lower than High school diploma)
  3. Middle education (High school diploma)
  4. Low university degree (2 or 3 years after High school)
  5. High university degree (4 years or more after High school)
- Pooling the education variable into two categories for the DiD:
  1. Having at most a high school degree
  2. Having more a high school degree
- Deletion of individuals born after their children.
- Parallel spells: each line of the data set corresponds to a person, a year and a job. Thus, if a person has multiple jobs in the same year, it will have multiple lines for this year. To keep only one line per individual and per year we computed the following variables for each individuals:

\footnote{Data munging was made possible thanks to the dplyr package [Wickham and Francois, 2015]. Tables were made thanks to [Hlavac, 2015].}
1. Number of hours worked in the year. (Whatever the number of jobs)
2. Number of days of work
3. Sum of the gross wage
4. Sum of the net wage (after social contributions but before income tax)
5. Sum of the net wage with fringe benefits
6. Sum of the fringe benefits

- After computing these variables, we wanted to keep one line per individual and per year. In case there are multiple lines per person and per year we have followed the following rules:
  1. Keep the line with the most days of work
  2. If number of working days are equal to several positions in the same year, we keep the job with the highest wage
  3. If the number of days of work and the wage are equal for various positions in the same year, we randomly keep one line per individual and per year

- We created hourly wages by dividing the annual wage by hours of work
- We deleted observations with extremely high hourly wages or extremely high worked hours

### 7.2 Appendix 2: Common trend graphs

The following graphs show the pre and post-treatment trends for our outcome variables of interest. Gray and dark gray areas are 95% confidence intervals around the mean for the treatment and control groups respectively. Lower educated women (men) are on the left, and higher educated women (men) are on the right.

#### 7.2.1 Impact of the first child

Figure 18: Hourly wages for women, less than 6 months out of labour force after giving birth
Figure 19: Hourly wages for women, more than 6 months but less than 2 and a half years out of labour force after giving birth

Figure 20: Hourly wages for women, more than 6 months but less than 6 years out of the labour force

Figure 21: Hourly wages for men, less than 6 months out of labour force after birth
Figure 22: Hourly wages for men, more than 6 months but less than 2 and a half years out of labour force after birth

Figure 23: Hourly wages for men, more than 6 months but less than 6 years out of the labour force

Figure 24: Worked hours for women, less than 6 months out of labour force after giving birth
Figure 25: Worked hours for women, more than 6 months but less than 2 and a half years out of labour force after giving birth

Figure 26: Worked hours for women, more than 6 months but less than 6 years out of the labour force

Figure 27: Worked hours for men, less than 6 months out of labour force after birth
Figure 28: Worked hours for men, more than 6 months but less than 2 and a half years out of labour force after birth

Figure 29: Worked hours for men, more than 6 months but less than 6 years after birth

7.2.2 Impact of the second child

Figure 30: Hourly wages for women, less than 6 months out of labour force after giving birth
Figure 31: Hourly wages for women, more than 6 months but less than 2 and a half years out of labour force after giving birth

Figure 32: Hourly wages for women, more than 6 months but less than 6 years out of the labour force

Figure 33: Hourly wages for men, less than 6 months out of labour force after birth
Figure 34: Hourly wages for men, more than 6 months but less than 2 and a half years out of labour force after birth

Figure 35: Hourly wages for men, more than 6 months but less than 6 years out of the labour force after birth

Figure 36: Worked hours for women, less than 6 months out of labour force after giving birth
Figure 37: Worked hours for women, more than 6 months but less than 2 and a half years out of labour force after giving birth

Figure 38: Worked hours for women, more than 6 months but less than 6 years out of the labour force

Figure 39: Worked hours for men, less than 6 months out of labour force after birth
Figure 40: Worked hours for men, more than 6 months but less than 2 and a half years out of labour force after birth

Figure 41: Worked hours for men, more than 6 months but less than 6 years out of the labour force after birth

7.2.3 Impact of the third child

Figure 42: Hourly wages for women, less than 6 months out of labour force after giving birth
Figure 43: Hourly wages for men, less than 6 months out of labour force after birth

Figure 44: Hourly wages for men, more than 6 months but less than 2 and a half years out of labour force after birth (only for lower educated men)

Figure 45: Hourly wages for men, more than 6 months but less than 6 years out of the labour force after birth
Figure 46: Worked hours for women, less than 6 months out of labour force after giving birth

Figure 47: Worked hours for men, less than 6 months out of labour force after birth

Figure 48: Worked hours for men, more than 6 months but less than 2 and a half years out of labour force after birth (only for lower educated men)
Figure 49: Worked hours for men, more than 6 months but less than 6 years out of the labour force after birth
References


