

IZA DP No. 4650

## The Family Gap in Wages: What Wombmates Reveal

Marianne Simonsen  
Lars Skipper

December 2009

# The Family Gap in Wages: What Wombmates Reveal

**Marianne Simonsen**

*Aarhus University  
and IZA*

**Lars Skipper**

*Aarhus University*

Discussion Paper No. 4650  
December 2009

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **The Family Gap in Wages: What Wombmates Reveal\***

We shed new light on the effects of having children on hourly wages by exploiting access to data on the entire population of employed same-sex twins in Denmark. Our second contribution is the use of administrative data on absenteeism; the amount of hours off due to holidays and sickness. Our results suggest that childbearing reduces female hourly wages but the principal explanation is in fact mothers' higher levels of absence. We find a positive wage premium for fathers.

JEL Classification: J13, J24, J31, J71

Keywords: fertility, wages, twins

Corresponding author:

Marianne Simonsen  
School of Economics and Management  
Aarhus University  
Building 1326  
Bartolins Alle 10  
8000 Aarhus C  
Denmark  
E-mail: [msimonsen@econ.au.dk](mailto:msimonsen@econ.au.dk)

---

\* We thank the Danish National Center for Social Research for making the data available to us. Financial support from the Danish Social Sciences Research council (Simonsen, grant 275-06-0353) is gratefully acknowledged. We appreciate helpful comments and suggestions from Sandy Black, Paul Devereux, and Helena Skyt Nielsen as well as from participants at the ESPE Conference 2009. The usual disclaimer applies.

# 1 Introduction

The decision to have a child is life-altering, also in the context of the labor market. Not only do one's preferences for non-work time potentially change with parenthood, parents are also likely hit by real productivity shocks in the form of sleep deprivation and an infinite - or so it sometimes may seem - stream of viruses. At the same time, the arrival of a child most often requires additional monetary resources. This paper investigates the effects of parenthood on wages.

The question of whether parenthood has a causal effect on wages has triggered a long line of empirical research. The overall conclusion is that having children is costly for women in terms of hourly wages. See Waldfogel (1998a) for a survey of the literature. These results are often attributed to mothers' different investments in household production resulting in for example career interruptions, different preferences for working conditions such as unplanned overtime, potentially different bargaining power because of stronger geographic ties due to costs of moving children, or possible discrimination.

The results for women are in stark contrast to those of men. Though much less frequently investigated, existing studies suggest that fathers are likely to observe an increase in wages from parenthood and marriage more generally; see Browning (1992), Pencavel (1986), Millimet (2000), and Simonsen and Skipper (2008). This is particularly intriguing given that fathers are also affected by (at least some of the) negative productivity shocks caused by the arrival of a child mentioned above. Clearly, both estimates for mothers and fathers are of crucial importance for understanding individual costs and gains related to childbearing. It is also important to know whether or to what degree wage gaps caused by parenthood contribute to explaining gender wage gaps, see Waldfogel (1998b).

To solve the problem that parenthood is likely endogenous to labor market outcomes, the existing literature mainly relies on conditional-on-observables strategies such as OLS and propensity score matching (e.g. Simonsen and Skipper (2006, 2008) using Danish data) or individual level fixed effects analyses (e.g. Anderson, Binder and Krause (2002) using US data). Instead, we propose to use the natural experiment of twins to overcome the selection problem. In particular, we exploit same-sex twin pairs, where one of the twins is a parent and the other is not (yet).

A similar approach starting with Taubman (1976) has been used in the literature estimating the effects of education on wages. Recent studies also exploit twin pairs to investigate effects of birth weight, see Almond, Chay, and Lee (2005), Black, Devereux, and Salvanes (2007) and Royer (2009) while Antonovicz and Town (2004) investigate the effects

of marriage on male wages. Furthermore, sister-differences more generally have been used to address teenage childbearing; see Geronimus and Korenman (1992), Hoffman, Foster and Furstenberg (1993), Bennet, Bloom, and Miller (1995), Ribar (1994) and Holmlund (2005). Except for the latter study, all of these studies on teenage childbearing suffer from very small and unrepresentative samples.

To the best of our knowledge, our paper is the first to employ the twin approach to uncover the effects of parenthood on wages for the general population, probably because large twin datasets with access to long-term labor market outcomes are rarely available. We exploit register data from 2006 on the entire population of employed Danes. Via person identifiers we are able to link family members in the population. Twins are then identified using exact birth dates from the national fertility register. This gives us the complete population of employed twins, consisting of more than 8,000 individuals. Finally, not only do we consider women, we also add to the scarce literature for men, thus providing new insights on the puzzling existing empirical estimates of the effects of fatherhood on wages.

Using a twin sample that to a very high degree resembles the overall population of employed individuals in terms of observable characteristics, we find that motherhood is indeed costly for the group of women: motherhood decreases hourly wages with around four to five per cent and even more so if the youngest child is aged 2 or below. This result is completely in line with the existing literature. Much of this wage gap is, however, explained by mother's degree of absence. Fatherhood, on the other hand, *does* cause an increase in hourly wages of around four per cent. Therefore, despite likely negative productivity shocks associated with parenthood, fathers are capable of increasing hourly wages. Whether we condition on observable characteristics or not does not affect our estimates at all; once twin fixed effects are accounted for, nothing else seems to drive the results.

The paper develops as follows. In the next section we discuss our empirical framework and identification strategy, and in Section 3 we present the Danish labor market and the institutional setting. Section 4 describes our data, Section 5 presents the results and sensitivity analyses, while Section 6 concludes.

## 2 Empirical Framework: Exploiting Twins

We wish to estimate the effects of having children,  $C = 1$ , as opposed to the 'untreated' state of 'non-parenthood',  $C = 0$ , on a measure of wages,  $Y_C$ . We limit our analysis to individuals who participate in the labor market and the estimated parameters should be

interpreted with this in mind.

Let  $Y_1$  be potential outcome in the presence of children and  $Y_0$  the potential outcome in the absence of children. Clearly, we do not observe the same individual both with and without children at the same point in time. Therefore, instead of uncovering person-specific impacts, attention in the literature usually shifts to that of constructing (conditional) means. Most often, the literature estimates the mean effect of treatment on the treated, defined as

$$\begin{aligned}\theta &\equiv E[Y_1 - Y_0|C = 1] \\ &= E[Y_1|C = 1] - E[Y_0|C = 1].\end{aligned}\tag{1}$$

Hence, the problem becomes that of finding the counterfactual  $E[Y_0|C = 1]$  in (1), which is, of course, unobserved.

In this paper we exploit the natural experiment of twins to identify the parameter of interest. The idea is to find pairs of same-sex twins where one is a parent and the other is not. The wage of the non-parent twin can then be used as the counter-factual outcome. This strategy assumes that comparing twins, perhaps conditional on attributes,  $\mathbf{X}$ , eliminates selective differences between those with and without children. Importantly,  $\mathbf{X}$  should only include variables believed not to be affected by parenthood, see Rosenbaum (1984) on the problems with conditioning on variables likely affected by the treatment. Therefore, the estimated effects include indirect effects stemming from differences between parents and non-parents in terms of experience, labor market interruptions etc. As such, the parameters are informative about wage costs and gains of parenthood including those accruing from individual choices caused by childbearing. Furthermore, because we condition on employment, we implicitly assume that if the parent-twin is observed to work then he would also work in the hypothetical case of non-parenthood. Thus, we assume monotonicity in the sense that parenthood only decreases the propensity to work. In any case, we expect this to be a minor concern, simply because Danish employment rates are high for both genders, see below.

Assuming linearity of the outcome equation but allowing for heterogenous treatment effect along the lines of Björklund and Moffit (1987) we get

$$Y_{ij} = X_{ij}\beta + \theta_j C_{ij} + [t_j + u_{ij}]$$

or alternatively

$$Y_{ij} = X_{ij}\beta + \bar{\theta}C_{ij} + [C_{ij}(\theta_j - \bar{\theta}) + t_j + u_{ij}]$$

where  $i$  indicates the individual,  $j$  indicates the family,  $\bar{\theta}$  is the average treatment effect on the treated,  $t_j$  are unobserved twin specific characteristics and  $u_{ij}$  are unobserved individual specific characteristics.  $\bar{\theta}$  can then be estimated using fixed effects (or by differencing).

The identifying strategy in some ways resembles that of matching. In contrast to the standard case of matching, however, twins have the same set of parents and similar genes, reproductive ability, upbringing, skills and appearance. Furthermore, as opposed to siblings in general, twins will most often experience the same *timing* of parental investments. What is important for our purposes is precisely to cancel out such unobservables that explain both the choice to have children and wage outcomes. Apart from 'nature' factors determining skills and reproductive ability that are shared between the twins, there is good reason to think that preferences for family and labor market participation are shaped by one's parents. A recent study that confirms this is Fernandez, Fogli, and Olivetti (2004) who credibly argue that men are more likely to marry a partner who participates in the labor market if their own mothers worked. There is also a large literature within psychology and sociology that investigates parental influence on children's attitudes towards family formation; Axinn and Thornton (1996) for example investigate the effects of parents' marital dissolution on children's attitudes towards premarital sex, cohabitation, marriage, and childbearing. Furthermore, Cunha, Heckman, Lochner and Masterov (2006) demonstrate that the family environment in general is extremely important for determining later (labor market) outcomes. The authors suggest that "strong families promote cognitive, social, and behavioral skills. Weak families do not" (p. 715). Cunha and Heckman (2008) shows that about half of the inequality in the present value of lifetime earnings is due to factors determined by age 18.

Though the idea of exploiting twin differences is intuitively appealing, there are also a number of challenges that needs to be addressed. First of all, twins may be similar with regards to a wide range of characteristics, but they may also differ on other characteristics. Mill et al. (2006) show that even monozygotic twins differ in their epigenetic profiles<sup>1</sup> at age five. Furthermore, their epigenetic differences increase with age and even more so if they have a history of non-shared environments, see Fraga et al. (2005) and Oates et al. (2006). If more able (earnings-wise) individuals become parents, the between-twins estimator will tend to overestimate the effect of parenthood, whereas the opposite will hold if less able individuals become parents. Griliches (1979) and Bound and Solon (1999) discuss this in the context of education and wages and Ribar (1994) in the context of teenage childbearing and educational attainment. Ribar (1994) point out that the within-family estimator is

less biased than the cross-sectional estimator if the unobserved twin specific characteristics,  $t_j$ , has a stronger correlation across siblings than has other unobserved determinants of fertility. Thus if the bulk of the variation in fertility stems from the partner's (fertility preferences or) fecundity, and these are assumed to be independent from the other twin's partner's characteristics, the between-twins estimator will be less biased.<sup>2</sup>

To accommodate the problem that twins are not identical,<sup>3</sup> we condition on a rich set of variables that explain both wages and the decision to enter parenthood and potentially differ between the two twins. Apart from level and type of education and geographic information we are also - for part of the sample - able to introduce pre-parent starting wages and an indicator for early partnership formation, see Section 4 below for details. Starting wages presumably both captures labor market skills (and the effect of grades, see Holmlund (2005)) that are predetermined to parenthood and also informs about the costs of temporarily leaving the labor market to engage in parenthood. Presence of early partner decreases the costs of conception and inform about preferences for family-formation in general.

A second concern with within-family estimators is that the identifying population is potentially very small. Put differently, twins may not vary a lot with regards to outcomes; after all, they are twins. This is particularly pronounced in the case of education. Because presence of a partner and the partner's characteristics (preferences and fecundity) likely affect the couple's fertility, there is significant variation in the choice of parenthood. In our twin sample, 290 twin pairs (out of 1,147 possible) identify the effects of motherhood on wages and 329 twin pairs (out of 979 possible) identify the effects of fatherhood on wages, see Table 4 below.

Finally, one twin's fertility choice may affect the other twin in ways that affect his or her labor market outcomes, see Currie (1995) and Holmlund (2005). If, for example, the pregnancy of one twin sparks the other twin's interest in becoming a parent *and* he or she then as a consequence changes labor market behavior, the results will be biased. More generally, twins comparisons will only be effective if the shared attributes are the only relevant omitted variables.

### 3 The Danish Labor Market and Policies

The Danish labor market is characterised by being as flexible in terms of hiring and firing policies as the US labor market and turnover is extremely high. At the same time, however, individuals are faced with considerable security in terms of degree and length



of compensation during unemployment spells. The degree of compensation is 90 % of previous wages but with an upper limit of DKK 3.335 (about \$ 600) per week in 2006. The average compensation rate is roughly 70 %.

Denmark is also among the countries in the world with the highest female employment rates (0.77 for women and 0.83 for men in the 20-42 age interval in 2006) and family leave is extensive. In 2006, when our analyses are conducted, mothers had the right to 14 weeks of paid maternal leave after birth followed by 32 weeks of paid parental leave that could in principle be shared with the father. Fathers, on the other hand, had the right to two paid weeks during the first 14 weeks following child birth. Compensation depended on the choice of sector and industry but everybody is granted at least the level of unemployment insurance yet others may receive full wage compensation. In any case, compensation rates for mothers almost always exceed that of men. This is simply because men on average earn higher wages than women and therefore more often are affected by the upper limit of UI compensation.

Child care is provided by the municipalities and is heavily subsidized; parents pay a maximum of 30 % of the total costs. At age three, about 80 % of all children were enrolled in publicly provided care.

## 4 Data

Our data stem from several administrative registers all maintained by Statistics Denmark. The wage data hold information on all individuals in Denmark in 2006 who are employed in firms with at least 10 employees. The reason for using this specific population instead of all employed Danes is the availability of exceptional information on not only total labor earnings but also its components (salary, value of fringe benefits, pension contributions etc.) in addition to information on both administrative records on normal hours, overtime, and all absences along with the reasons for absence (sickness, holidays). Thus, *de facto*, we know the amount of hours worked. Via person identifiers, we are able to link parents and their children in the full population. Twins are then identified using exact birth dates from the national fertility register. If siblings are born on the same day, we define them to be twins. Furthermore, the 2006 wage data are augmented with variables describing demographics and education on a yearly basis, individual level employment and unemployment histories along with parental fertility and educational information. We do observe the high quality wage and hours information in the 2002-2006 period. In practice, however, because our twin sample is still relatively small, the panel is too short for us to

employ a differences-in-differences strategy. We do observe cruder wage information as far back as 1997. This is used to construct yearly pre-birth labor income at age 25, which is exploited in our sensitivity analyses below. The information about the individuals' parents is used to establish external validity of our twin based analyses and also exploited in our OLS analyses for the overall population. An unfortunate downside to the data set at hand is that we cannot investigate, for example, the propensity to be employed because we only have the full set of information on employed individuals. Again note that employment propensities are very high for both genders.

We select a subsample of individuals aged 20-42 years with a core attachment to the labor market.<sup>4</sup> In particular, we select individuals who are employed more than 200 hours per year, who are not self-employed, and not undertaking education. The lower age bound is chosen to exclude individuals who are between two types of education, for instance high school and university. The upper age bound is chosen because of an age restriction on the availability of information about the individuals' parents.

We define parenthood based on existence children under the age of 18 in the household in 2006. This gives us the wage effect of the presence of at least one child in the household (where the child is not necessarily biologically linked to the unit of observation). In our sensitivity analyses, we also consider parenthood based on biological link and the conclusions are completely robust to this. This second definition gives us the wage effect of being the biological parent (but where the unit of observation does not necessarily live with or has ever lived with the child). As such, the parameters are informative about the effects of different aspects of parenthood. Due to data limitations, we are unfortunately not able to define fatherhood based on this latter criteria and it is possible that results may differ for men depending on the definition of parenthood. Table 1 shows the share of parents among men and women and the distribution of the age of the youngest child in our twin sample.

**Table 1**  
DISTRIBUTION OF AGE OF YOUNGEST CHILD,  
PARENTS IN TWIN SAMPLE

|  | Mothers | Fathers |
|--|---------|---------|
| <i>Share with youngest child aged:</i> |         |         |
| 0-2 years                              | 0.33    | 0.43    |
| 3-6 years                              | 0.29    | 0.30    |
| 7-9 years                              | 0.17    | 0.15    |
| 10-14 years                            | 0.17    | 0.11    |
| 15-17 years                            | 0.03    | 0.01    |
| Share of parents in sample             | 0.70    | 0.49    |
| Number of parents in sample            | 1,612   | 967     |

The outcome variable used in the analysis is *log hourly labor earnings in 2006*. This wage measure includes all paid components, also pension contributions and fringe benefits. Using hourly earnings allows us to directly address effects of parenthood on compensation for a unit of time investment. We consider hourly earnings both in terms of normal and actual hours on the job. As defined by OECD, normal hours of work are the hours of work fixed by or in pursuance of laws or regulations, collective agreements or arbitral awards, or the number of hours in excess of which any time worked is remunerated at overtime rates. Actual hours on the job are defined as normal hours minus hours absent due to sickness, holidays etc plus overtime. Hourly wages defined by normal hours has the advantage of being extremely precise. In some professions, it is presumably very difficult to correctly assess the actual number of hours. If overtime to a higher degree is underreported for the group of nonparents, estimates using actual hours at the workplace will tend to be biased downwards. Moreover, to the extent that changes in actual hours are part of the costs of parenthood - and we want to measure that - defining hourly wages by normal hours is the relevant measure. On the other hand, hourly wages defined by normal hours are obviously not as informative about actual behavior. Apart from hourly labor earnings, we also separately consider basic hourly wages or salaries, i.e. the amount paid for nominal time of work.

Table 2 shows average labor earnings and its main components and hours of work by gender and parenthood status in our twin sample. On average, parents have more normal hours than non-parents; presumably, it is more costly to be unemployed as a parent than otherwise because also one's children are affected by this.<sup>5</sup> See Figure A1 in Appendix A for

the relationship between age of youngest child and normal hours for the entire population of employed Danes. Here it is clear that fathers work more than non-fathers regardless of the age of the youngest child, whereas mothers catch-up when the child is between one and two years old. Unfortunately, we observe too few twins for us to meaningfully construct an analogous figure for this population. However, we demonstrate below that twins are very similar to the population in general. We therefore expect Figure A1 to be representative for the twins as well. Consider next actual hours worked. Comparing actual hours with normal hours reveals that all groups of individuals on average work less than normal hours. Much of this is due to holidays.<sup>6</sup> Furthermore, this latter pattern is more pronounced among women and particularly so for the group of mothers. See Figure A2 in Appendix A for the relationship between age of youngest child and the difference between normal and actual hours for the entire population of employed Danes. This demonstrates that the parents accumulate more hours of absence irrespective of the age of the youngest child. An important explanation may be the child's sick days, which likely affects both parents. Ekberg, Ericson, and Friebel (2005), for example, find that Swedish fathers' share of care for sick children is around 35 %.

**Table 2**  
WAGE COMPONENTS AND HOURS OF WORK, TWIN SAMPLE<sup>a</sup>  
PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD

|   | Mothers              | Non-mothers          | Fathers              | Non-fathers          |
|---|----------------------|----------------------|----------------------|----------------------|
| Labor earnings 2006, DKK                    | 267,818<br>(124,533) | 258,793<br>(140,775) | 364,295<br>(170,622) | 298,619<br>(159,382) |
| Share of labor earnings                     |                      |                      |                      |                      |
| - basic pay                                 | 0.693                | 0.683                | 0.713                | 0.716                |
| - pension contributions and fringe benefits | 0.117                | 0.115                | 0.114                | 0.107                |
| - other components                          | 0.190                | 0.202                | 0.172                | 0.177                |
| Normal hours                                | 1,512<br>(493)       | 1,443<br>(567)       | 1,590<br>(502)       | 1,456<br>(566)       |
| Actual hours                                | 1,244<br>(431)       | 1,223<br>(489)       | 1,410<br>(449)       | 1,302<br>505         |
| Absence (normal-actual hours)               | 268                  | 220                  | 180                  | 154                  |
| Absence relative to normal hours            | 0.18                 | 0.15                 | 0.11                 | 0.11                 |
| Sample size                                 | 1,612                | 682                  | 967                  | 991                  |

<sup>a</sup>As of May 2009 \$ 1 corresponds to DKK 5.5

Table 3 shows descriptive statistics on selected variables for mothers and non-mothers, fathers and non-fathers in the twin sample. The upper part of the table informs about background variables, while the lower part of the table show the outcome variable and other variables likely affected by parenthood. If there is an effect of parenthood on wages, we do, of course, not expect labor market outcomes in general to balance across the groups of parents and non-parents either. Note that experience is actual experience in the sense that it does not include unemployment spells or parental leave. It does on the other hand include days off due to sickness and holidays.<sup>7</sup>

We see that parents are about six years older than non-parents and they are better educated than non-parents. Mothers are more likely to have a vocational degree and less likely to have high school or less education, while fathers are more likely to have a vocational degree or more education. There are no big differences in neither own number of siblings nor their parents' level of education. Parents and non-parents behave differently on the labor market, however. Parents have higher hourly wages. Mothers are more likely to work in the public sector, less likely to have a top level occupation and have significantly longer unemployment spells and, of course, birth related leave. Fathers, on the other hand, are more likely to work in top and higher level occupations, and have significantly shorter unemployment spells. There are large differences between mothers'

and fathers' level of leave-taking; the average mother has accumulated 342 days of leave, whereas the average father in our sample has 23 leave days. The differences are at least partly due to the institutional set-up that grants mothers longer leave at an most often higher degree of compensation. Note that non-parents may have some birth-related leave. If both biological parents take leave and divorce later on, for example, children will only be registered as living with one of the parents, yet both adults will have positive amounts of leave.

One might worry that our twin sample is very different from the average population and therefore conclude that the results from the formal analyses below are non-representative. Table A1 in Appendix A documents, however, that while many differences in means are statistically significantly different because of the very large sample sizes, the differences themselves are fairly small and might to some extent be driven by the difference in average age between the two samples. Thus despite the fact that twins in general start out in life being significantly different from singletons (Almond, Chay, and Lee (2005), Black, Devereux, and Salvanes (2007)) our twin sample resembles the overall population in terms of the observable characteristics from Table 3.

Table 4 show descriptive statistics for the identifying population, namely the sub-population where one twin is a parent and the other is not (yet). Note that mothers and non-mothers in this population have equal amounts of experience and unemployment. Furthermore, there are no significant differences between obtained levels of education across parents and non-parents. As in Table 2 (and Table A1), mothers are more likely to work in the family-friendly public sector. Fathers have slightly more experience than non-fathers and have a correspondingly lower unemployment duration.

**Table 3**  
 DESCRIPTIVE STATISTICS, SELECTED VARIABLES, TWIN SAMPLE  
 PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD

| Variables  | Mothers         | Non-mothers     | Fathers         | Non-fathers     |
|--|-----------------|-----------------|-----------------|-----------------|
| <i>Background variables :</i>                                |                 |                 |                 |                 |
| Number of children   | 1.93<br>(0.75)  | •               | 1.88<br>(0.74)  | •               |
| Age of youngest child (years)                                | 5.57<br>(4.37)  | •               | 4.32<br>(3.93)  | •               |
| Age (years)  | 36.36<br>(4.25) | 31.20<br>(6.23) | 37.05<br>(4.21) | 32.28<br>(6.00) |
| Length of completed education                                |                 |                 |                 |                 |
| less than high school  | 0.11            | 0.15            | 0.16            | 0.20            |
| high school  | 0.06            | 0.08            | 0.04            | 0.07            |
| vocational degree  | 0.40            | 0.32            | 0.47            | 0.46            |
| short further education                                      | 0.06            | 0.07            | 0.07            | 0.06            |
| medium length further education                              | 0.28            | 0.28            | 0.16            | 0.09            |
| long further education                                       | 0.09            | 0.10            | 0.11            | 0.11            |
| Number of siblings in family                                 | 4.36<br>(1.04)  | 4.24<br>(1.09)  | 4.35<br>(1.11)  | 4.15<br>(0.91)  |
| Mother's length of completed education                       |                 |                 |                 |                 |
| less than high school  | 0.46            | 0.42            | 0.48            | 0.42            |
| high school  | 0.01            | 0.01            | 0.01            | 0.01            |
| vocational degree  | 0.33            | 0.33            | 0.34            | 0.35            |
| short further education                                      | 0.03            | 0.04            | 0.03            | 0.03            |
| medium length further education                              | 0.15            | 0.17            | 0.12            | 0.15            |
| long further education                                       | 0.02            | 0.02            | 0.01            | 0.03            |
| Father's length of completed education                       |                 |                 |                 |                 |
| less than high school  | 0.35            | 0.35            | 0.34            | 0.31            |
| high school  | 0.02            | 0.02            | 0.01            | 0.02            |
| vocational degree  | 0.43            | 0.39            | 0.45            | 0.46            |
| short further education                                      | 0.04            | 0.06            | 0.05            | 0.05            |
| medium length further education                              | 0.11            | 0.11            | 0.09            | 0.11            |
| long further education                                       | 0.06            | 0.07            | 0.06            | 0.05            |
| <i>Labor market outcomes :</i>                               |                 |                 |                 |                 |
| Log wages, normal hours                                      | 5.13<br>(0.29)  | 5.11<br>(0.30)  | 5.38<br>(0.31)  | 5.25<br>(0.32)  |
| Log wages, actual hours                                      | 5.33<br>(0.32)  | 5.27<br>(0.32)  | 5.50<br>(0.31)  | 5.36<br>(0.32)  |
| Experience (years)   | 13.66<br>(5.46) | 9.07<br>(5.94)  | 15.61<br>(5.53) | 10.86<br>(6.19) |
| Private sector (0/1)   | 0.43            | 0.48            | 0.83            | 0.80            |
| Top level occupation (0/1)                                   | 0.13            | 0.15            | 0.17            | 0.12            |
| Higher level occupation (0/1)                                | 0.31            | 0.23            | 0.18            | 0.12            |
| Medium level occupation (0/1)                                | 0.36            | 0.31            | 0.39            | 0.43            |
| Lower level occupation (0/1)                                 | 0.03            | 0.05            | 0.07            | 0.06            |
| Total duration of unemployment, last ten years (days)        | 162<br>(327)    | 152<br>(288)    | 85<br>(206)     | 129<br>(260)    |
| Total duration of birth-related leave, last ten years (days) | 342<br>(298)    | 6<br>(55)       | 24<br>(43)      | 3<br>(23)       |
| Sample size  | 1,612           | 682             | 967             | 991             |
| Share of parents   |                 | 0.70            |                 | 0.49            |

<sup>a</sup>Standard deviations shown in parentheses.

**Table 4**

DESCRIPTIVE STATISTICS, SELECTED VARIABLES, IDENTIFYING TWIN POPULATION  
PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD

| Variables  | Mothers         | Non-mothers     | Fathers         | Non-fathers     |
|--|-----------------|-----------------|-----------------|-----------------|
| <i>Background variables:</i>                                 |                 |                 |                 |                 |
| Number of children   | 1.70<br>(0.66)  | •               | 1.76<br>(0.68)  | •               |
| Age of youngest child (years)                                | 4.72<br>(4.46)  | •               | 3.73<br>(3.98)  | •               |
| Age (years)  |                 | 34.70<br>(4.92) |                 | 35.02<br>(4.58) |
| Length of completed education                                |                 |                 |                 |                 |
| less than high school  | 0.09            | 0.12            | 0.16            | 0.20            |
| high school  | 0.06            | 0.06            | 0.04            | 0.04            |
| vocational degree  | 0.41            | 0.40            | 0.50            | 0.49            |
| short further education                                      | 0.07            | 0.06            | 0.05            | 0.06            |
| medium length further education                              | 0.32            | 0.28            | 0.14            | 0.10            |
| long further education                                       | 0.06            | 0.09            | 0.11            | 0.12            |
| Number of siblings in family                                 |                 | 4.35<br>(1.16)  |                 | 4.22<br>(0.90)  |
| Mother's length of completed education                       |                 |                 |                 |                 |
| less than high school  |                 | 0.40            |                 | 0.46            |
| high school  |                 | 0.02            |                 | 0.02            |
| vocational degree  |                 | 0.35            |                 | 0.32            |
| short further education                                      |                 | 0.02            |                 | 0.03            |
| medium length further education                              |                 | 0.19            |                 | 0.14            |
| long further education                                       |                 | 0.02            |                 | 0.02            |
| Father's length of completed education                       |                 |                 |                 |                 |
| less than high school  |                 | 0.34            |                 | 0.29            |
| high school  |                 | 0.03            |                 | 0.02            |
| vocational degree  |                 | 0.41            |                 | 0.49            |
| short further education                                      |                 | 0.07            |                 | 0.05            |
| medium length further education                              |                 | 0.11            |                 | 0.10            |
| long further education                                       |                 | 0.05            |                 | 0.05            |
| <i>Labor market outcomes:</i>                                |                 |                 |                 |                 |
| Log wages, normal hours                                      | 5.11<br>(0.27)  | 5.17<br>(0.29)  | 5.35<br>(0.33)  | 5.32<br>(0.30)  |
| Log wages, actual hours                                      | 5.32<br>(0.31)  | 5.34<br>(0.32)  | 5.47<br>(0.33)  | 5.43<br>(0.38)  |
| Experience (years)   | 12.04<br>(5.49) | 12.03<br>(5.48) | 14.45<br>(5.83) | 13.93<br>(6.13) |
| Private sector (0/1)   | 0.39            | 0.49            | 0.82            | 0.83            |
| Top level occupation (0/1)                                   | 0.12            | 0.16            | 0.16            | 0.14            |
| Higher level occupation (0/1)                                | 0.28            | 0.22            | 0.18            | 0.14            |
| Medium level occupation (0/1)                                | 0.35            | 0.37            | 0.42            | 0.44            |
| Lower level occupation (0/1)                                 | 0.03            | 0.06            | 0.08            | 0.08            |
| Total duration of unemployment, last ten years (days)        | 191<br>(373)    | 182<br>(305)    | 103<br>(239)    | 146<br>(300)    |
| Total duration of birth-related leave, last ten years (days) | 318<br>(273)    | 11<br>(72)      | 22<br>(41)      | 5<br>(37)       |
| Sample size  | 290             | 290             | 329             | 329             |
| Share of parents   |                 | 0.50            |                 | 0.50            |

<sup>a</sup>Standard deviations shown in parentheses.



## 5 Estimated Family Wage Gaps

This section presents our estimation results. We first present our main results for hourly labor earnings after which we investigate the effects on the different wage components and perform a range of robustness checks.

Table 5 shows the estimated effects of parenthood. We use both OLS on the overall population, OLS on the twin sample, twin fixed effects without explanatory variables, and twin fixed effects conditioning on information that varies within twin pairs. Our conditioning set for the full sample consists of age dummies, variables describing type and length of own education, dummies for place of habitation, number of siblings, and parental educational attainment. See Table A2 in Appendix A for details. Again, we do not want to condition on variables that are believed to be affected by parenthood. We therefore leave out variables such as sector of employment, experience and occupational choice from the conditioning set.

We define parenthood based on presence of children living in the household under the age of 18. We are thus estimating the effects of being responsible for a child (where the child is not necessarily biologically linked to the unit of observation). Consider first the effects on hourly labor earnings defined by normal hours. These results are shown in the upper section of Table 5. Here parenthood decreases hourly labor earnings for women in the order of five percent. Men, on the other hand, gain around four percent in hourly labor earnings from fatherhood. Surprisingly, both sets of results are very robust to the choice of estimation method. For both genders are OLS estimates - for the overall population as well as for the twin population - fairly close to the results using twin fixed effects.<sup>8</sup> Note also that once twins fixed effects are accounted for, nothing else seems to drive the results. The formal tests support this. The estimates for women lie in the lower end of the estimates from the international literature (e.g. Anderson, Binder, and Krause (2002), Budig and England (2001) and Waldfogel (1998a, 1998b)) but are in line with previous Danish findings (e.g. Datta Gupta and Smith (2002), Simonsen and Skipper (2006, 2008)) and also consistent with results from the more recent studies from the US, see for example Millimet (2000). As indicated above, the few existing papers find that men gain from parenthood and marriage more generally. Our findings support this.

The results using actual hours in place of normal hours are shown in the lower part of Table 5. As expected given the differences in hours between mothers and non-mothers shown in Table 2, the estimated effects are now much smaller in size and not significantly different from zero when employing twin fixed effects. The lack of significance could be

due to small sample sizes. It seems, nonetheless, that the majority of the wage penalty for women observed above can be explained when we account for absence. One way to interpret this result is that employers just acknowledge that mothers are likely to be more absent than non-mothers (given the same number of normal hours) and they reward the two groups alike based on the number of hours they are actually present at the workplace. Since fathers are also more absent than non-fathers, the results for men are larger in size compared to the results using normal hours and the conclusion still holds; fatherhood causes a significant increase in hourly labor earnings.

**Table 5**  
ESTIMATION RESULTS<sup>a</sup>  
PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD

|  | Women                                     |               |            | Men  |              |            |
|--|---|---------------|------------|--|--------------|------------|
|  | # obs                                     | Effect        | Std. Error | # obs                                      | Effect       | Std. Error |
| <i>Hourly labor earnings, normal hours:</i>        |   |               |            |  |              |            |
| OLS - population                                   | 372,377                                   | <b>-0.046</b> | 0.001      | 336,633                                    | <b>0.037</b> | 0.001      |
| OLS - twins  | 2,294                                     | <b>-0.052</b> | 0.013      | 1,958                                      | <b>0.057</b> | 0.017      |
| Twin fixed effect w/o explanatory variables        | 2,294                                     | <b>-0.057</b> | 0.020      | 1,958                                      | 0.028        | 0.019      |
| F-test of joint significance of twin fixed effects | F(1146,1146)=2.09, P-value=0.000          |               |            | F(978,978)=2.23, P-value=0.000             |              |            |
| Hausman test of FE vs. RE                          | Chi <sup>2</sup> (1)=17.09, P-value=0.000 |               |            | Chi <sup>2</sup> (1)=29.30, P-value=0.000  |              |            |
| Twin fixed effect w. explanatory variables         | 2,294                                     | <b>-0.051</b> | 0.019      | 1,958                                      | 0.027        | 0.019      |
| F-test of joint significance of twin fixed effects | F(1146,1118)=1.57, P-value=0.000          |               |            | F(978,949)=1.79, P-value=0.000             |              |            |
| Hausman test of FE vs. RE                          | Chi <sup>2</sup> (1)=63.62, P-value=0.002 |               |            | Chi <sup>2</sup> (1)=91.96, P-value=0.000  |              |            |
| <i>Hourly labor earnings, actual hours:</i>        |   |               |            |  |              |            |
| OLS - population                                   | 372,377                                   | <b>-0.020</b> | 0.002      | 336,633                                    | <b>0.056</b> | 0.002      |
| OLS - twins  | 2,294                                     | -0.002        | 0.015      | 1,958                                      | <b>0.067</b> | 0.014      |
| Twin fixed effect w/o explanatory variables        | 2,294                                     | -0.026        | 0.022      | 1,958                                      | <b>0.041</b> | 0.018      |
| F-test of joint significance of twin fixed effects | F(1146,1146)=1.82, P-value=0.000          |               |            | F(978,978)=2.48, P-value=0.000             |              |            |
| Hausman test of FE vs. RE                          | Chi <sup>2</sup> (1)=15.29, P-value=0.001 |               |            | Chi <sup>2</sup> (1)=28.33, P-value=0.000  |              |            |
| Twin fixed effect w. explanatory variables         | 2,294                                     | -0.021        | 0.022      | 1,958                                      | <b>0.041</b> | 0.080      |
| F-test of joint significance of twin fixed effects | F(1146,1118)=1.34, P-value=0.000          |               |            | F(978,949)=1.85, P-value=0.000             |              |            |
| Hausman test of FE vs. RE                          | Chi <sup>2</sup> (1)=58.68, P-value=0.001 |               |            | Chi <sup>2</sup> (1)=110.87, P-value=0.000 |              |            |

<sup>a</sup>Bold coefficients are significant at the 5 % level, italic indicates significance at the 10 % level.

Given the discrepancy between mothers' and fathers' actual level of leave-taking, it is maybe not surprising that the estimated parameters are not the same either. What may be surprising is that fathers actually gain from parenthood since they are also likely to be affected by negative productivity shocks caused by the arrival of a child. They are, for example, also more absent than non-fathers given the same number of normal hours. Of course, it might just be the case that fathers compensate for the loss in labor income from mothers by changing their labor market behavior. In other words, we might suspect the difference between the gender-specific estimates to stem from family responsibilities and subsequent differential investments in the labor market; women clearly engage the most in caring for children. Using time-use data on 6,624 adult Danes in 2001, Bonke (2002) finds that women with children below the age of seven spend close to six hours per day on home production. The corresponding number of hours for men is four. Single men or men in relationships without children or children older than seven years report to be spending between two and three hours per day. Women with children older than seven report spending  $4\frac{1}{2}$  hours per day whereas singles or women living in relationships without children report spending between two and four hours per day. Only for lone parents and singles below 45 are there no statistically differences between the hours spent on home production between the genders. Furthermore, Bonke, Datta Gupta and Smith (2003) show that women in particular are penalized in terms of wages for doing housework and even more so if the housework is timed immediately before or after their job. Also, Simonson and Skipper (2008) show that Danish women temporarily reduce their participation in (post secondary) labor market training during their childbearing period. This is not the case for men.

## 5.1 Sensitivity Analyses

We lastly proceed to our sensitivity analyses. We include a richer conditioning set that is only available for part of our sample, consider estimates of the effects on hourly salary instead of hourly labor earnings, we analyse the group with small children,<sup>9</sup> and explore an alternative definition of parenthood.

The most important sensitivity analysis is consists of adding pre-parent information. Here, in addition to the variables described in Table A2 in the Appendix, we condition on pre-parent starting wages proxied by labor income at age 25, a dummy for nonparticipation at age 25, and an indicator for early partnership formation (also at age 25). To make sure that this information is actually measured before parenthood we only include

individuals who become parents after the age of 26; about 65 % of the parents in our sample enter parenthood after this age. The estimates should be interpreted as such. As described above, our information about labor income goes back to 1997. This means that we only have the early labor income information for the part of the sample aged 26-34 in 2006. 78 per cent of these participate at age 25. To avoid that the sample becomes too small we include also individuals aged above 34 and condition on dummies for missing information on pre-parent starting wages and early partnership. Table 6 presents the corresponding means. While early partnership formation is not associated with future parenthood status regardless of gender, starting wages seem to be positively correlated with future fatherhood.

**Table 6**  
 DESCRIPTIVE STATISTICS, PRE-PARENT VARIABLES,  
 IDENTIFYING TWIN POPULATION  
 EXCLUDE PARENTS BEFORE AGE 26  
 PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD

| Variables                         | Mothers        | Non-mothers    | Fathers        | Non-fathers           |
|-----------------------------------|----------------|----------------|----------------|-----------------------|
| Starting wages (in DKK 100,000)   | 2.49<br>(1.81) | 2.27<br>(1.56) | 2.71<br>(1.69) | <b>2.10</b><br>(1.46) |
| Early partnership formation (0/1) | 0.63           | 0.64           | 0.72           | 0.72                  |
| Sample size                       | 110            | 110            | 99             | 99                    |

<sup>a</sup>Standard deviations shown in parentheses. Bold indicates that non-parent mean is statistically significantly different from parent mean (5 % level).

Table 7 shows the estimation results. We first replicate the analyses from Table 5 using the restricted sample and then add the additional covariates to the conditioning set. Despite the differences in initial characteristics, this exercise does not change the conclusions from above and the results are completely robust to the inclusion of pre-parent information.

**Table 7**  
ESTIMATION RESULTS: INCLUDING STARTING  
WAGE AND EARLY PARTNERSHIP INFORMATION  
EXCLUDE PARENTS BEFORE AGE 26  
PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD

|   | # obs | Women         |            | # obs | Men          |            |
|---|-------|---------------|------------|-------|--------------|------------|
|   |       | Effect        | Std. error |       | Effect       | Std. error |
| <i>Hourly labor earnings, contracted hours:</i> |       |               |            |       |              |            |
| Twin fixed effect w. explanatory variables      | 1,572 | <b>-0.044</b> | 0.023      | 1,616 | 0.018        | 0.021      |
| including wage at age 25 and early partnerdummy | 1,572 | <b>-0.043</b> | 0.023      | 1,616 | 0.017        | 0.021      |
| <i>Hourly labor earnings, actual hours:</i>     |       |               |            |       |              |            |
| Twin fixed effect w. explanatory variables      | 1,572 | 0.005         | 0.025      | 1,616 | <i>0.033</i> | 0.020      |
| including wage at age 25 and early partnerdummy | 1,572 | 0.006         | 0.025      | 1,616 | <i>0.034</i> | 0.020      |

<sup>a</sup>Bold coefficients are significant at the 5 % level, italic indicates significance at the 10 % level.

An important issue is whether parenthood affects the most liquid part of earnings, namely basic pay, more than total earnings. This is analysed in Table 8. Comparing with the results from Table 5 documents, however, that the effects on basic pay are almost identical to the effects on total labor earnings. Here, the effect of motherhood on hourly basic pay defined by actual hours is significant at the 10 % level.

**Table 8**  
ESTIMATION RESULTS: BASIC PAY<sup>a</sup>  
PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD  
TWIN FIXED EFFECTS WITH EXPLANATORY VARIABLES

|  | # obs | Women         |            | # obs | Men          |            |
|--|-------|---------------|------------|-------|--------------|------------|
|  |       | Effect        | Std. error |       | Effect       | Std. error |
| <i>Hourly basic pay, normal hours:</i> | 2,294 | <b>-0.062</b> | 0.023      | 1,958 | 0.017        | 0.019      |
| <i>Hourly basic pay, actual hours:</i> | 2,294 | <i>-0.032</i> | 0.018      | 1,958 | <i>0.031</i> | 0.017      |

<sup>a</sup>Bold coefficients are significant at the 5 % level, italic indicates significance at the 10 % level.

Another potentially interesting question is whether the effects vary with the age of the children. Table 9 investigates this. Here, we consider the effects of the youngest child being 0-2 years old compared to not having children and the effects of the youngest child being 3-6 years compared with 0-2 years. Note that in the latter case we are comparing twins who are both parents; what varies is the timing of parenthood. Consider first the effects of the youngest child being 0-2 years old relative to not having children. It is clear that this is more expensive in terms of wages for women; both estimates are much more negative than the average effects in Table 5. When children are 0-2 years old, mothers have just returned to the labor market after extensive leave periods and at the same time children are being introduced to child care, which may be emotionally draining for both the child and the parents. Children are also much more likely to catch diseases when exposed to non-parental care, see e.g. Gordon, Kaestner, and Korenman (2007). This may increase work absence among parents of young children. Still, only the estimate for hourly wages defined by normal hours is significantly different from zero when evaluated at the 5 % level. The effects of fatherhood on measures of hourly wages for the group with young children are slightly larger in size compared to those for the overall population. The results for having a child aged 3-6 relative to a child aged 0-2 show that there is a considerable degree of catching-up among mothers. The estimates for men show that the fatherhood premium continues to rise with the age of the youngest child.

**Table 9**

ESTIMATION RESULTS: SMALL CHILDREN  
PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD  
TWIN FIXED EFFECTS WITH EXPLANATORY VARIABLES

|   |       | Women         |            |       | Men          |            |
|---|-------|---------------|------------|-------|--------------|------------|
|   | # obs | Effect        | Std. Error | # obs | Effect       | Std. Error |
| <i>Hourly labor earnings, normal hours:</i> |       |               |            |       |              |            |
| Youngest child 0-2 years vs. no children    | 902   | <b>-0.107</b> | 0.030      | 1,148 | <i>0.044</i> | 0.027      |
| <i>Hourly labor earnings, actual hours:</i> |       |               |            |       |              |            |
| Youngest child 0-2 years vs. no children    | 902   | -0.054        | 0.035      | 1,148 | <b>0.058</b> | 0.025      |
| <i>Hourly labor earnings, normal hours:</i> |       |               |            |       |              |            |
| Youngest child 3-6 years vs. 0-2 years      | 618   | <b>0.078</b>  | 0.035      | 358   | <i>0.082</i> | 0.047      |
| <i>Hourly labor earnings, actual hours:</i> |       |               |            |       |              |            |
| Youngest child 3-6 years vs. 0-2 years      | 618   | 0.007         | 0.042      | 358   | <i>0.074</i> | 0.044      |

<sup>a</sup>Bold coefficients are significant at the 5 % level, italic indicates significance at the 10 % level.

Table 10 shows the results for motherhood defined by biological link. As explained above, this cannot be constructed for fathers in our data. These results should be interpreted as the effects of being a biological mother (but where the unit of observation does not necessarily live with or has ever lived with her child). There is great overlap between the two groups of mothers, however. All of the mothers defined by presence of children in the household are also mothers when exploiting biological link. Conversely, 90 per cent of the biological mothers are also mothers when defined by presence of children in the household. The results are also highly similar.

**Table 10**  
ESTIMATION RESULTS  
PARENTHOOD BASED ON BIOLOGICAL LINK  
TWIN FIXED EFFECTS WITH EXPLANATORY VARIABLES

|   |       | Women         |            |
|---|-------|---------------|------------|
|   | # obs | Effect        | Std. Error |
| <i>Hourly labor earnings, normal hours:</i> | 2,294 | <b>-0.051</b> | 0.010      |
| <i>Hourly labor earnings, actual hours:</i> | 2,294 | -0.013        | 0.023      |

<sup>a</sup>Bold coefficients are significant at the 5 % level, italic indicates significance at the 10 % level.

## 6 Conclusion

We estimate total effects of parenthood on hourly wages exploiting access to the entire population of employed Danish twins. Sibling based analyses more generally have been used to evaluate the effects of teenage childbearing but we are not aware of any other papers that employ this strategy to consider the effects of parenthood for the general population.

We identify the effect of parenthood using same-sex twin pairs where one twin is a parent and the other is not (yet). We are therefore able to clear out unobservables that are shared between the twins because they have the same set of parents and similar genes, reproductive ability, upbringing, skills and appearance. This is important to the extent that these variables explain both wage outcomes and child bearing. We acknowledge, of course, that twins are not completely alike. To address potential residual selection

we therefore condition on a rich set of explanatory variables including preparent starting wages and information about early partnership formation.

In contrast to much of the existing literature, we not only consider the effects of motherhood, we also estimate effects for men. The twin based analyses reveal that mothers face a significant gap in hourly wages: we find that women earn on average four to five per cent less per hour as a consequence of motherhood, and the effects are larger for women with young children. Most of this wage gap, however, seems to be explained by mother's when we account for absence and use actual instead of normal hours. Given the same number of actual hours, mothers do not earn less than non-mothers. We conjecture, however, that mothers' higher degree of absence is chiefly driven by motherhood. Thus motherhood constitutes a real cost to women; mothers could have earned more on the job had they not been absent. From the point of view of the employers, however, mothers and non-mothers who work the same number of actual hours are treated equally. Men gain around four per cent from fatherhood when considering hourly wages. This roughly corresponds to the value of one year's experience.

It is maybe not surprising the childbearing reduces women's hourly wages. The obvious question is why this does not seem to be the case for men. Fathers in our identifying sample have about six months more experience than non-fathers and correspondingly lower unemployment duration. This is unlikely to explain the entire gap, however. Other explanations may be the amount of household work or take-up of labor market training. Descriptive evidence shows that men and women change the number of hours spent in household production when they become parents. Where they before entering parenthood were more or less equal partners in the production of household goods, women clearly take over after childbirth, likely leaving women less energy to invest in the labor market. An interesting question is, however, to what extent parenthood changes men's (and women's) wage bargaining position.



# Appendix A

**Table A1**

DESCRIPTIVE STATISTICS, SELECTED VARIABLES, EMPLOYED DANES  
PARENTHOOD BASED ON CHILDREN LIVING IN THE HOUSEHOLD

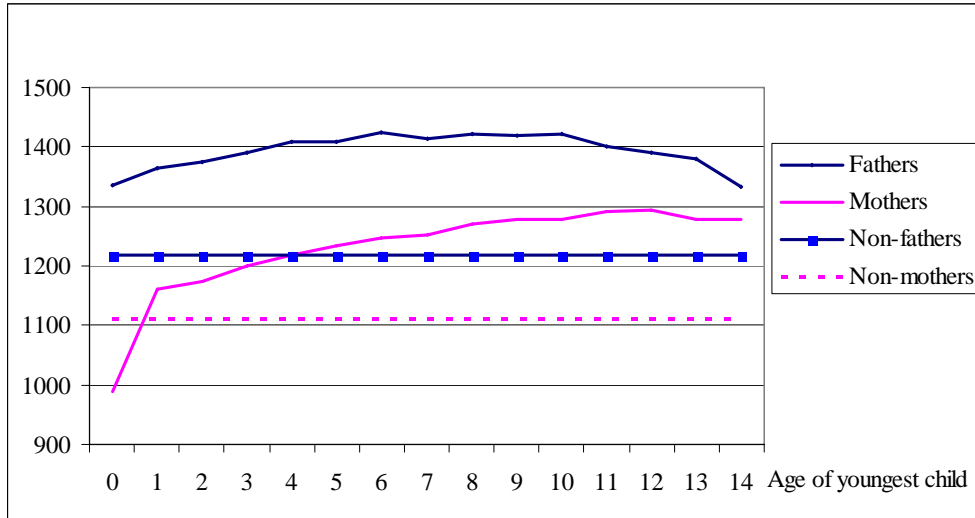
| Variables  | Mothers                | Non-mothers            | Fathers                | Non-fathers            |
|--|------------------------|------------------------|------------------------|------------------------|
| <i>Background variables :</i>                                |                        |                        |                        |                        |
| Number of children   | <b>1.88</b><br>(0.75)  | •                      | 1.87<br>(0.76)         | •                      |
| Age of youngest child (years)                                | 5.42<br>(4.40)         | •                      | <b>4.58</b><br>(4.24)  | •                      |
| Age (years)  | <b>35.31</b><br>(4.95) | <b>28.95</b><br>(6.49) | <b>35.47</b><br>(5.10) | <b>30.01</b><br>(6.62) |
| Length of completed education                                |                        |                        |                        |                        |
| less than high school  | <b>0.14</b>            | <b>0.19</b>            | <b>0.19</b>            | <b>0.27</b>            |
| high school  | 0.06                   | <b>0.14</b>            | 0.05                   | <b>0.10</b>            |
| vocational degree  | 0.39                   | <b>0.29</b>            | <b>0.44</b>            | <b>0.39</b>            |
| short further education                                      | 0.05                   | <b>0.06</b>            | 0.08                   | 0.06                   |
| medium length further education                              | 0.27                   | <b>0.23</b>            | 0.14                   | 0.10                   |
| long further education                                       | 0.08                   | 0.09                   | 0.11                   | <b>0.08</b>            |
| Number of siblings in family                                 | <b>2.49</b><br>(1.18)  | <b>2.37</b><br>(1.11)  | <b>2.48</b><br>(1.18)  | <b>2.44</b><br>(1.14)  |
| Mother's length of completed education                       |                        |                        |                        |                        |
| less than high school  | 0.47                   | <b>0.37</b>            | 0.47                   | 0.41                   |
| high school  | 0.01                   | <b>0.02</b>            | 0.01                   | 0.02                   |
| vocational degree  | 0.35                   | <b>0.37</b>            | 0.35                   | 0.36                   |
| short further education                                      | 0.03                   | 0.04                   | 0.03                   | 0.03                   |
| medium length further education                              | <b>0.13</b>            | 0.18                   | 0.13                   | 0.15                   |
| long further education                                       | 0.02                   | 0.03                   | 0.02                   | 0.02                   |
| Father's length of completed education                       |                        |                        |                        |                        |
| less than high school  | 0.36                   | <b>0.30</b>            | 0.34                   | 0.32                   |
| high school  | 0.01                   | 0.02                   | <b>0.01</b>            | 0.02                   |
| vocational degree  | 0.43                   | <b>0.45</b>            | 0.44                   | 0.45                   |
| short further education                                      | 0.04                   | 0.05                   | 0.04                   | 0.05                   |
| medium length further education                              | 0.10                   | 0.12                   | 0.11                   | 0.11                   |
| long further education                                       | 0.05                   | 0.07                   | 0.05                   | 0.06                   |
| 61   |                        |                        |                        |                        |
| <i>Labor market outcomes :</i>                               |                        |                        |                        |                        |
| Log wages, normal hours                                      | <b>5.10</b><br>(0.31)  | <b>5.04</b><br>(0.34)  | <b>5.36</b><br>(0.35)  | <b>5.20</b><br>(0.35)  |
| Log wages, actual hours                                      | <b>5.30</b><br>(0.36)  | <b>5.32</b><br>(0.35)  | <b>5.46</b><br>(0.35)  | <b>5.42</b><br>(0.38)  |
| Experience (years)   | <b>12.10</b><br>(5.76) | <b>7.24</b><br>(5.61)  | <b>14.53</b><br>(6.21) | <b>9.32</b><br>(6.30)  |
| Private sector (0/1)   | 0.42                   | <b>0.52</b>            | 0.81                   | 0.82                   |
| Top level occupation (0/1)                                   | 0.13                   | <b>0.12</b>            | 0.16                   | <b>0.10</b>            |
| Higher level occupation (0/1)                                | 0.27                   | 0.22                   | 0.18                   | 0.12                   |
| Medium level occupation (0/1)                                | <b>0.33</b>            | <b>0.27</b>            | <b>0.35</b>            | <b>0.35</b>            |
| Lower level occupation (0/1)                                 | 0.04                   | <b>0.04</b>            | 0.08                   | <b>0.08</b>            |
| Total duration of unemployment, last ten years (days)        | <b>209</b><br>(375)    | <b>133</b><br>(287)    | <b>111</b><br>(268)    | <b>153</b><br>(311)    |
| Total duration of birth-related leave, last ten years (days) | 335<br>(299)           | <b>3</b><br>(34)       | 24<br>(46)             | <b>2</b><br>(13)       |
| Sample size  | 244876                 | 127501                 | 165384                 | 171249                 |
| Share of parents   |                        | 0.66                   |                        | 0.49                   |

<sup>a</sup> Standard deviations shown in parentheses. The sample is chosen using the same selection criteria as used to construct the twin sample. Bold coefficients indicate that mean is statistically significantly different from the corresponding mean in the twin sample (5 % level).

**Table A2**  
CONDITIONING SET

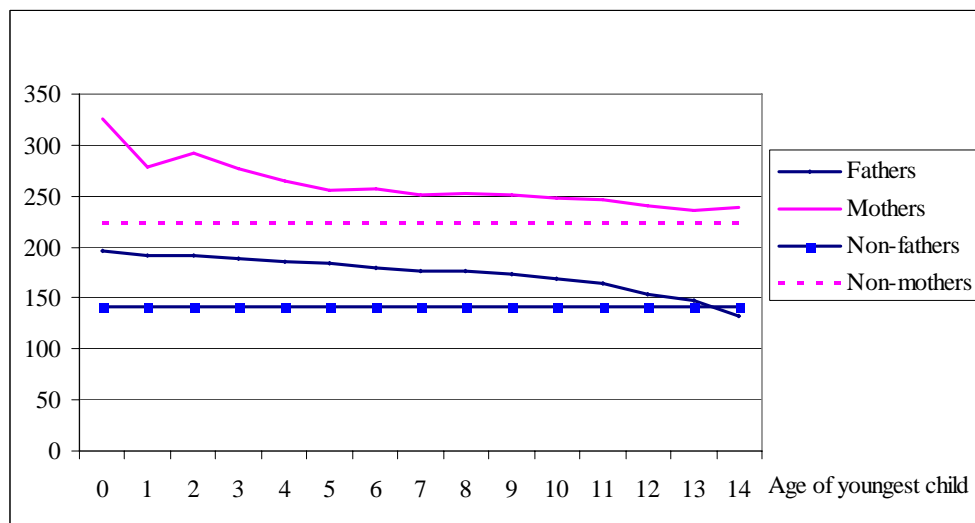
|   |   |
|---|---|
| <b>Length of highest completed education:</b> | Region 2 (0/1) Frederiksborg and Roskilde counties    |
| High school (0/1)                             | Region 3 (0/1) Western Sealand and Storstrøm counties |
| Vocational degree (0/1)                       | Region 4 (0/1) Fuen county                            |
| Short tertiary degree (0/1)                   | Region 5 (0/1) Southem Jutland and Ribe counties      |
| Medium lenght tertiary degree (0/1)           | Region 6 (0/1) Vejle and Ringkøbing counties          |
| Long tertiary degree (0/1)                    | Region 7 (0/1) Aarhus and Viborg counties             |
| <b>Type of highest completed education:</b>   | Region 8 (0/1) Northem Jutland county                 |
| General (0/1)                                 | <b>Age information:</b>                               |
| Humanistic (0/1)                              | Age 20-24 (0/1)                                       |
| Musical (0/1)                                 | Age 25-29 (0/1)                                       |
| Natural sciences (0/1)                        | Age 30-34 (0/1)                                       |
| Social (0/1)                                  | Age 35-39 (0/1)                                       |
| Construction (0/1)                            | Age 40-44 (0/1)                                       |
| Industry (0/1)                                | <b>Family background:</b>                             |
| Graphical (0/1)                               | Number of siblings                                    |
| Technical (0/1)                               | Mother w. high school degree (0/1)                    |
| Services (0/1)                                | Mother w. vocational degree (0/1)                     |
| Food and beverages (0/1)                      | Mother w. short tertiary degree (0/1)                 |
| Agricultural (0/1)                            | Mother w. medium lenght tertiary degree (0/1)         |
| Transport (0/1)                               | Mother w. long tertiary degree (0/1)                  |
| Health sciences (0/1)                         | Father w. high school degree (0/1)                    |
| Public safety (0/1)                           | Father w. vocational degree (0/1)                     |
| Pedagogic (0/1)                               | Father w. short tertiary degree (0/1)                 |
| <b>Regional information:</b>                  | Father w. medium lenght tertiary degree (0/1)         |
| Region 1 (0/1) Copenhagen county              | Father w. long tertiary degree (0/1)                  |

**Figure A1**  
 NORMAL HOURS AND AGE OF YOUNGEST CHILD  
 POPULATION OF EMPLOYED DANES, 2006<sup>a</sup>



<sup>a</sup>The data are chosen using the same selection criteria as for the twin population.

**Figure A2**  
 DIFFERENCE BETWEEN NORMAL AND ACTUAL HOURS  
 AND AGE OF YOUNGEST CHILD  
 POPULATION OF EMPLOYED DANES, 2006<sup>a</sup>



<sup>a</sup>The data are chosen using the same selection criteria as for the twin population.

## Literature

Anderson, D. J., Binder, M., Krause, K., 2002. The motherhood wage penalty: which mothers pay it and why? *The American Economic Review (Papers and Proceedings)* 92, 354-358.

Antonovicz, K., Town, R., 2004. Are all the good men married? Uncovering the sources of the marital wage premium. *The American Economic Review (Papers and Proceedings)* 94, 317-321.

Almond, D., Chay, K. Y., Lee, D. S., 2005. The costs of low birth weight. *Quarterly Journal of Economics* 120, 1031-1083.

Axinn, W. G., Thornton, A., 1996. The influence of parents' marital dissolutions on children's attitudes toward family formation. *Demography* 33, 66-81.

Bennet, N. G., Bloom, D. E., Miller, C. K., 1995. The influence of nonmarital childbearing on the formation of first marriages. *Demography* 32, 47-62.

Björklund, A., Moffit, R., 1987. The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics* 69, 42-49.

Black, S. A., Devereux, P., Salvanes, K., 2007. From the cradle to the labor market? The effect of birth weight on adult outcomes. *Quarterly Journal of Economics* 122, 409-439.

Bonke, J., 2002. Tid og velfærd [Time and Welfare (in Danish)]. Working Paper No. 26, Danish Social Research Institute.

Bonke, J., Datta Gupta, N., Smith, N., 2003. The timing and flexibility of housework and men's and women's wages. In *Contributions to Economic Analysis*, D. Hamermesh and G. Pfann (eds.), Amsterdam, Elsevier Press, 43-77.

Bound, J., Solon, G., 1999. Double trouble: on the value of twins-based estimation of the return to schooling. *Economics of education review* 18, 169-182.

Browning, M., 1992. Children and household economic behavior. *Journal of Economic Literature* 30, 1434-1475.

Budig, M. J., England, P., 2001. The wage penalty for motherhood. *American Sociological Review* 66, 204-225.

- Cunha, F. and J. J. Heckman 2008. A new framework for the analysis of inequality. *Macroeconomic Dynamics* 12, 315-354.
- Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov 2006. Interpreting the evidence on life cycle skill formation. In E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education* 1, 697-812.
- Currie, J., Thomas, D., 1995. Does head start make a difference? *The American Economic Review* 85, 341-364.
- Datta Gupta, N., Smith, N., 2002. Children and career interruptions. *Economica* 69. 609-629.
- Ekberg, J., Eriksson, R., Friebel, G., 2005. Parental leave - an evaluation of the Swedish "Daddy Month" reform. IZA DP No. 1617.
- Fernandez, R., Fogli, A., Olivetti, C., 2004. Mothers and sons: preference formation and female labor force dynamics. *The Quarterly Journal of Economics* 119, 1249-1299.
- Fraga, M. F., E. Ballestar, M. F. Paz, S. Ropero, F. Setien et al. (2005). Epigenetic Differences Arise During the Lifetime of Monozygotic Twins. *Proceedings of the National Academy of Sciences of the United States of America* 102, 10604-10609.
- Geronimus, A. T., Korenman, S., 1992. The socioeconomic consequences of teen child-bearing reconsidered. *The Quarterly Journal of Economics* 107, 1187-1214.
- Griliches, Z., 1979. Sibling models and data in economics: beginnings of a survey. *Journal of Political Economy* 87, S37-S64.
- Gordon, R. A., Kaestner, R., Korenman, S., 2007. The effects of maternal employment on child injuries and infectious disease. *Demography* 44, 307-333.
- Hoffman, S. D., Foster, E. M., Furstenberg Jr., F. F., 1993. Reevaluating the costs of teenage childbearing. *Demography* 30, 1-13.
- Holmlund, H., 2005. Estimating long-term consequences of teenage childbearing. *Journal of Human Resources* 40, 716-743.
- Mill, J., Dempster, E., Caspi, A., Williams, B., Moffitt, T., Craig, I., 2006. Evidence for Monozygotic twin discordance in Methylation Level at two CpG sites in the promoter

region of the Catechol-O-methyltransferase gene. *American Journal of Medical Genetics* 141, 421-425.

Millimet, D. L., 2000. The impact of children on wages, job tenure, and the division of household labour. *Economic Journal* 110, 139–157.

Nielsen, H. S., 2009. Causes and consequences of fathers' child leave: evidence from a reform of a public sector leave scheme. Mimeo, University of Aarhus.

Nielsen, H. S., Simonsen, M., Verner, M., 2004. Does the gap in family-friendly policies drive the family gap? *Scandinavian Journal of Economics* 106, 721-744.

Oates, N. A., van Vliet, J., Duffy, D. L., Kroes, H. Y., Martin, N. G., Boomsma, D. I., 2006. Increased DNA Methylation at the AXIN1 gene in a monozygotic twin from a pari discordan for a caudal duplication anomaly. *American Journal of Human Genetics* 79, 155-162.

Pencavel, J., 1986. Labor Supply of Men: a Survey, in: Ashenfelter, O., Layard, R. (Eds.), *Handbook of Labor Economics*, Vol. 1, Elsevier Science, New York, pp. 3–102.

Phipps, S., Burton, P., Lethbridge, L., 2001. In and out of the labour market: long-term income consequences of child-related interruptions to women's paid work. *Canadian Journal of Economics* 34, 411–429.

Ribar, D., 1994. Teenage fertility and high school completion. *Review of economics and statistics* 76, 413-424.

Rosenbaum, P. R., 1984. The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society Series A* 147, 656–666.

Royer, H., 2009. Separated at Girth: US Twin Estimates of the Effects of Birth Weight. *American Economic Journal: Applied Economics* 1, 49–85.

Simonsen, M., Skipper, L., 2006. The costs of motherhood: an analysis using matching estimators. *Journal of Applied Econometrics* 21. 919-934

Simonsen, M. Skipper, L., 2008. An empirical assessment of effects of parenthood on wages. *Advances in Econometrics* 21, 359-380.

Simonsen, M. Skipper, L., 2008b. The incidence and intensity of formal lifelong learning. Economics WP No. 2008-07, University of Aarhus.

Taubman, P., 1976. Earnings, education, genetics, and environment. *Journal of Human Resources* 11, 447-461.

Waldfogel, J., 1998a. Understanding the "family gap" in pay for women with children. *Journal of Economic Perspectives* 12, 137–156.

Waldfogel, J., 1998b. The family gap for young women in the United States and Britain: can maternity leave make a difference? *Journal of Labor Economics* 16, 505-545.

## Notes

<sup>1</sup>Epigenetic is defined as something that affects a cell, organ or individual without directly affecting its DNA. An epigenetic change may indirectly influence the expression of the genome.

<sup>2</sup>Griliches (1979) and Bound and Solon (1999) also discuss the effect of measurement errors when applying the between-twins estimator. With register based data we are less concerned about such errors in the definition of parenthood.

<sup>3</sup>See discussion in for example Geronimus and Korenman (1992) and Holmlund (2005). With our data we are, for example, not able to distinguish between monozygotic and dizygotic twins. See however Black, Devereux, and Salveness (2007) who find that pooling monozygotic and dizygotic twins is innocuous when investigating the effects of birth weight on labor market outcomes.

<sup>4</sup>Employment rates are high for both men and women in Denmark: in 2006, 78 % of all Danish men and 72 % of all Danish women were in employment.

<sup>5</sup>This pattern also holds when we consider the identifying population.

<sup>6</sup>A full time employed individual has the right to 25 vacation days per year, see The Law of Holidays ('Ferieloven').

<sup>7</sup>We only have access to information about absence in recent years and therefore would not be able to purge our measure of experience from days off due to sickness and holidays

in general.

<sup>8</sup>This is also found by Black, Devereux, and Salvanes (2007) when investigating the effects of birth weight on adult outcomes.

<sup>9</sup>Unfortunately, despite having access to all employed twins in 2006, our sample size is still relatively small. This means that we can only reasonably investigate a limited number of subgroups.