THE COSTS OF MOTHERHOOD: AN ANALYSIS USING MATCHING ESTIMATORS

MARIANNE SIMONSEN*a AND LARS SKIPPERb

a Department of Economics, University of Aarhus, Aarhus, Denmark
b Institute for Local Government Studies, Copenhagen, Denmark

SUMMARY
We estimate the effect of motherhood on wages using matching. We distinguish between net and direct effects. The net effect includes the total wage costs, whereas the direct represents the causal effect. Since covariates are likely affected by motherhood, the latter effect is not immediately uncovered. We therefore implement two strategies: first, we confine the analysis to consider sector-specific treatment effects; second, we impose additive separability on the outcome equation. We find negative net effects that vary little with sector. The direct effect is small and negative in the public sector and insignificant in the private sector.

1. INTRODUCTION
The most common parameter of interest in the treatment evaluation literature is the average effect of treatment on the treated (Heckman et al., 1999). In the case where the subpopulation under study consists of women, the treatment is being a mother and the outcome is wages, this parameter has been designated ‘the family gap’ or ‘the child penalty’ in the literature (Budig and England, 2001; Phipps et al., 2001; Korenman and Neumark, 1992; Waldfogel, 1998a,b). The naming of the parameter is, of course, a direct consequence of the results found in the literature: having a child seems to be costly in terms of wages.1 For the United States, Canada and Great Britain, the afore-mentioned studies suggest a penalty of between 10% and 15% lower wages, whereas penalties in Scandinavia are typically smaller, ranging between 0% and 1.5% (Datta Gupta and Smith, 2002; Nielsen et al., 2004). In this paper we reconsider this classical problem using recently developed techniques on a high-quality comprehensive register-based data set.

Motherhood may affect wages through several channels and hence there are potentially more parameters of interest. There may be a direct effect of motherhood on wages, i.e. the causal effect of motherhood on wages, and there may be indirect effects running through the effect of motherhood on other covariates (see Korenman and Neumark, 1992, and Robins and Greenland, 1992, for this terminology). For example, mothers may have lower levels of labour market experience due to child-rearing activities. This is likely to affect wages, but the effect is indirect. We define the sum of the direct and the indirect effect as the net effect. The direct effect of motherhood on

1 The family gap may exist for several reasons: mothers may invest differently in household production (resulting, for example, in career interruptions), may have different preferences for working conditions (such as unplanned overtime), and may have different bargaining power (e.g. stronger geographic ties due to costs of moving children). Finally, discrimination may explain a potential wage gap.

Received 23 March 2004
Revised 24 August 2005
wages has, by far, received the most attention in the literature (Nielsen et al., 2004; Phipps et al., 2001; Waldfogel, 1998a), yet the recovering of the parameter is not immediate owing to bias introduced by conditioning on variables that are likely to be affected by the choice of motherhood (see discussion below). Similar issues arise in the literature analysing the effects of college quality on earnings as emphasized in a recent paper by Black and Smith (2004). Here, years of schooling depend partly on college quality, but also have a separate exogenous effect on labour market outcomes. The paper presents results both with and without years of schooling included in the conditioning set and finds large and significant differences in treatment effects.

Both the direct and the net effect are interesting from a policy point of view. The former is needed if we want to make any conclusions on discrimination: if discrimination on the labour market is present it comprises at least part of the direct effect. The latter provides information on the costs of choices that are related to child-bearing.

Common to all the analyses on the estimation of family gaps mentioned above is the assumption of a common treatment effect at least within subpopulations. As noted in Heckman et al. (1999) this is a very strong assumption. Furthermore, all are subject to assumptions on functional form of the equations of interest; for one thing separability of the effects of observables and unobservables is assumed, and the conditional expectations function is assumed to be linear in attributes. Some are also subject to parametric assumptions on correlations of error terms. Alternatively, one may apply matching analysis. Matching is based on the principle that comparing the outcome for an individual from the treatment group, i.e. mothers, with the outcome for an individual from the no-treatment group, i.e. women without children, who in terms of observables is sufficiently similar to the treated individual on average, balances the selection bias arising from self-selection into motherhood. Matching allows for heterogeneous treatment effects, is not subject to parametric assumptions, and does not per se assume separability of the effects of observables and unobservables. It does, however, require extremely rich data sets.

In this paper we apply propensity score matching to estimate the effect of having children on women’s wages. The data at our disposal is a Danish register-based data set that includes very detailed, high-quality information on, for example, income, demographics and education on a yearly basis. Furthermore, the individual event histories in terms of periods of employment, unemployment, as well as maternal leave and other types of publicly subsidized leave schemes are known on a weekly basis. We discuss the assumptions needed for identification of direct and net effects on wages of being a mother when variables are affected by the treatment. If the researcher has sufficient information to make wage outcomes conditionally independent of both motherhood and the variable affected by motherhood, direct effects can be uncovered for specific values of the concomitant variable. Specifically, we consider sector of employment: in Denmark the public sector is characterized by having more flexible working conditions and both duration of and compensation during maternity leave are higher in the public sector than in the private sector. Hence, we expect there to be more mothers working in the public sector ceteris paribus (Nielsen et al., 2004). When considering sector-specific effects of motherhood we think of the effect as the combined treatment of motherhood and sector of employment. If, on the other hand, sufficient information on selection of variables affected by the treatment is unavailable, an alternative solution is to impose additive separability on the outcome equation. To clear out effects of, for example, experience and career interruptions on wages we impose an exclusion restriction to perform regression-adjusted matching, recently suggested by Heckman et al. (1998). To our knowledge matching analysis has never been applied in this area before. We find strong evidence
that mothers select into the public sector and significant, but small family gaps overall along with slightly larger effects in the public sector.

The paper is organized as follows. In the next section we discuss our sample of data. In Section 3 we discuss our parameters of interest and outline our econometric strategy. Section 4 presents the results from propensity score matching, while Section 5 discusses the results from accounting for sector selection. Finally, Section 6 concludes.

2. DATA

The original data set contains information on a representative sample of 5% of all Danish individuals in the 15–74 age bracket. Information stems from several registers maintained by Statistics Denmark. The registers include variables describing income, demographics and education on a yearly basis. Furthermore, the individual event history in terms of periods of employment, unemployment, maternal leave, as well as other publicly subsidized leave (child-rearing or sabbatical) and the residual category non-participation, is known on a weekly basis.

In the empirical analysis below, we use a 1997 cross-sectional subsample of women aged 20–40 years, 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 parental information used to construct exclusion restrictions applied in the econometric analyses. The analysis is performed using retrospective information on the labour market history.

Table I shows descriptive statistics for the sample used in our analyses along with descriptive statistics of mothers and non-mothers. We classify women as mothers if they have given birth to a child. Thus it is assumed that the presence of biological children is more important than the presence of stepchildren. 2 54.8% of the women in our sample had children in 1996.

The outcome variable of interest in the analysis is log hourly wage. It is calculated from annual earnings and number of working hours. The measure of working hours used in this calculation is very precise in that the information comes from registers on compulsory contributions to supplemental pension payments that are closely linked to the working hours actually paid for by employers. It is seen that average log wage for mothers does not differ from average log wage for non-mothers. Yet it is clear from Table I that mothers differ significantly from non-mothers in terms of observables: mothers are on average 7 years older, are more often employed in the public sector, have twice as much labour market experience, are more likely to have an education directed towards the health care sector or the schooling system, and have on average longer unemployment spells. Furthermore, they are more likely to own real estate and are more often settled in the province. Finally, they are, not surprisingly, more likely to be married and have fewer siblings of their own. 3

The information on interruptions consists of a subset of spells created from accurate event histories known on a weekly basis. Incidences of unemployment and non-participation are

---

2 This may, of course, be a problem if children in the household other than biological children of the woman affect her choices and actions. Most often, though, children from separated homes live with their mother. Moreover, in 1996 the number of adoptions amounted to only 600 in total. In our sample this would amount to approximately 30 adoptions per year using 1996 numbers.

3 This latter variable will be used as our exclusion restriction in what follows. We argue below that number of siblings is correlated with fertility but not with productivity, i.e. labour market outcome.
Table I. Selected Moments, Descriptive Statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>All</th>
<th>Mothers</th>
<th>Non-mothers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log wages</td>
<td>4.80</td>
<td>4.81</td>
<td>4.79</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.26)</td>
<td>0.30</td>
</tr>
<tr>
<td>Age (years)</td>
<td>29.72</td>
<td>32.87</td>
<td>25.88</td>
</tr>
<tr>
<td></td>
<td>(5.80)</td>
<td>(4.20)</td>
<td>(5.13)</td>
</tr>
<tr>
<td>Experience (years)</td>
<td>7.28</td>
<td>9.39</td>
<td>4.70</td>
</tr>
<tr>
<td></td>
<td>(4.87)</td>
<td>(4.29)</td>
<td>(4.25)</td>
</tr>
<tr>
<td>Length of completed education (years)</td>
<td>12.22</td>
<td>12.20</td>
<td>12.26</td>
</tr>
<tr>
<td></td>
<td>(2.45)</td>
<td>(2.42)</td>
<td>(2.48)</td>
</tr>
<tr>
<td>Type of highest completed education:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General (0/1)</td>
<td>0.22</td>
<td>0.19</td>
<td>0.26</td>
</tr>
<tr>
<td>Business (0/1)</td>
<td>0.34</td>
<td>0.33</td>
<td>0.35</td>
</tr>
<tr>
<td>Industry (0/1)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Construction (0/1)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Graphical (0/1)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Services (0/1)</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Food and beverages (0/1)</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Agricultural (0/1)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Health (0/1)</td>
<td>0.12</td>
<td>0.15</td>
<td>0.07</td>
</tr>
<tr>
<td>Pedagogic (0/1)</td>
<td>0.06</td>
<td>0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>Humanistic (0/1)</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>Musical (0/1)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Social (0/1)</td>
<td>0.04</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>Technical (0/1)</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Unknown (0/1)</td>
<td>0.07</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>Owner of real estate (0/1)</td>
<td>0.42</td>
<td>0.55</td>
<td>0.26</td>
</tr>
<tr>
<td>Married (0/1)</td>
<td>0.43</td>
<td>0.66</td>
<td>0.15</td>
</tr>
<tr>
<td>Private sector (0/1)</td>
<td>0.52</td>
<td>0.46</td>
<td>0.60</td>
</tr>
<tr>
<td>Top-level occupation (0/1)</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Higher-level occupation (0/1)</td>
<td>0.21</td>
<td>0.23</td>
<td>0.18</td>
</tr>
<tr>
<td>Medium-level occupation (0/1)</td>
<td>0.52</td>
<td>0.51</td>
<td>0.53</td>
</tr>
<tr>
<td>Lower-level occupation (0/1)</td>
<td>0.17</td>
<td>0.16</td>
<td>0.19</td>
</tr>
<tr>
<td>Total duration of unemployment (weeks)</td>
<td>64.40</td>
<td>87.50</td>
<td>36.30</td>
</tr>
<tr>
<td></td>
<td>(94.60)</td>
<td>(106.90)</td>
<td>66.80</td>
</tr>
<tr>
<td>Years since last u-spell</td>
<td>4.65</td>
<td>5.22</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td>(3.80)</td>
<td>(4.03)</td>
<td>(3.15)</td>
</tr>
<tr>
<td>Total duration of non-participation (weeks)</td>
<td>176.50</td>
<td>162.00</td>
<td>194.20</td>
</tr>
<tr>
<td></td>
<td>(119.30)</td>
<td>(119.50)</td>
<td>116.60</td>
</tr>
<tr>
<td>Years since last n-spell</td>
<td>4.54</td>
<td>3.74</td>
<td>5.56</td>
</tr>
<tr>
<td></td>
<td>(3.57)</td>
<td>(3.25)</td>
<td>(3.69)</td>
</tr>
<tr>
<td>Number of siblings when 15–17 years of age</td>
<td>1.16</td>
<td>1.00</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>(1.02)</td>
<td>(1.05)</td>
<td>(0.95)</td>
</tr>
<tr>
<td>Number of siblings missing (0/1)</td>
<td>0.19</td>
<td>0.23</td>
<td>0.14</td>
</tr>
<tr>
<td>Mother working in private sector</td>
<td>0.68</td>
<td>0.72</td>
<td>0.62</td>
</tr>
<tr>
<td>Father working in private sector</td>
<td>0.89</td>
<td>0.90</td>
<td>0.87</td>
</tr>
<tr>
<td>Sample size</td>
<td>29.027</td>
<td>15.958</td>
<td>13.069</td>
</tr>
</tbody>
</table>

*Note:* Standard deviations are shown in parentheses.

registered from 1981 and onwards, while maternity leave and parental leave in connection with childbirth can be traced back to 1984 (before 1984, maternity leave is included in the non-participation category). In the period before 1984, mothers are eligible for 18 weeks of maternity leave (4 weeks before expected birth and 14 weeks after), which means that for the oldest women in our sample some maternity leave may be hidden in the residual category of non-participation. In
1984, the maternity leave scheme was extended by 10 weeks of leave, amounting to a maximum of 28 weeks, and fathers were granted 2 weeks of leave during the first 14 weeks after the birth. This is the scheme effective in 1997. The 10 additional weeks can in principle be shared with the father, yet this is not the norm. In fact, in 1997 96.6% of all fathers on leave take 2 weeks or less and the mode is 2 weeks (Statistics Denmark, www).

In 1994, publicly subsidized sabbatical leave and child-rearing leave were introduced. For employed individuals, child-rearing leave amounts to a maximum of 52 weeks per child under the age of 8, while sabbatical leave amounts to a maximum of 52 weeks. The length of these two types of leave is registered from 1995 and onwards.

As mentioned in the Introduction, the public sector in Denmark is more family friendly than the private sector. Specifically, women employed in the public sector receive a full wage compensation during maternity leave, as opposed to the much lower UI benefits received by private sector employees. In addition, public sector employees are allowed to take leave earlier compared to private sector employees and have the right to 10 fully funded care days per child per year. Finally, apart from a small qualification bonus, wages in the public sector are mechanically determined by seniority. Therefore, mothers in the public sector do not risk losing experience-based wage increases when taking maternity leave.

3. THE PARAMETERS OF INTEREST

The goal of the evaluation is to measure the effect or impact of a given treatment, $C$, on an outcome variable, $Y_C$. Here the treatment is having children, $C = 1$, as opposed to the ‘untreated’ state of ‘non-motherhood’, $C = 0$, and the outcome of interest is log hourly wages. Let $Y_1$ be potential outcome in the presence of children and $Y_0$ the potential outcome in the absence of children. We are now faced with what is known as the Fundamental Evaluation Problem in that we do not observe the same woman both with and without children at the same point in time. Moreover, instead of uncovering person-specific impacts, attention in the literature usually shifts to that of constructing (conditional) means. Most often, we estimate the mean effect of treatment on the treated, defined as

$$
\theta = E[Y_1 - Y_0|C = 1] = E[Y_1|C = 1] - E[Y_0|C = 1]
$$  

Hence, the problem becomes that of finding the counterfactual $E[Y_0|C = 1]$ in (1), which is, of course, unobserved; i.e., some assumptions are needed to obtain identification.

We estimate the mean effect of having children for women who choose to have children, (1), and not the mean effect for the population of women. With potential heterogeneous impacts these parameters will likely differ. Our focus is thus whether women who choose to have children are punished in terms of wages, not whether women who do not have children would be punished had they chosen to have children.

In this paper we apply the method of matching, which has recently received much attention in applied econometrics in general and in programme evaluation in particular. Matching is based on the assumption that conditioning on attributes, $X$, eliminates the selective differences between those with and without children. More precisely, the method of matching assumes that the econometrician has access to conditioning variables sufficiently rich such that the counterfactual

Copyright © 2006 John Wiley & Sons, Ltd.  
outcome distribution of mothers is the same as the observed outcome distribution of non-mothers. By conditioning on the covariates at our disposal, we will thus be capable of balancing the bias coming from the self-selection into motherhood.

In focusing on (1) we make the following conditional independence assumption (CIA; Rosenbaum and Rubin, 1983): \( Y_0 \perp C \mid X \) (2)

In other words, we assume that there exists some \( X \) such that the outcome for a mother had she not had children is the same as the outcome for a non-mother with the same \( X \). In particular, this means that women must not take into account wages in the non-motherhood state when deciding whether to become mothers or not. They may, however, consider wages in the motherhood state. This is consistent with the case where mothers and non-mothers are equally productive in the non-motherhood state conditional on their characteristics but are potentially different in the motherhood state. In order to utilize equation (2) it is necessary to make sure that there is a woman without a child analogue to each mother, the ‘common support’ assumption, i.e.:

\[ P \equiv \Pr(C = 1 \mid X) < 1 \] (3)

At this point we do not want to assume any functional form of the outcome equation as opposed to much of the literature already mentioned in the Introduction. We are therefore potentially faced with the non-parametric curse of dimensionality due to our rich register data. A way to circumvent the curse of dimensionality without imposing arbitrary assumptions on the outcome equation is based on the results in Rosenbaum and Rubin (1983). Here the focus is shifted from the set of covariates to the probability of motherhood, \( P = \Pr(C = 1 \mid X) \). As long as (2) and (3) hold:

\[ Y_0 \perp C \mid P \] (4)

This new conditioning variable, \( P \), changes CIA into (4), which together with \( P < 1 \) are sufficient conditions required to justify propensity score matching to estimate the mean impact on the treated. Clearly, the functional form of \( P \) is rarely known and has to be estimated, shifting the high-dimensional estimation problem from that of estimating \( E[Y \mid X] \) to that of estimating \( E[C \mid X] \), often estimated by a logit or, as in this paper, a probit. See discussion in Black and Smith (2004) on this issue. Moreover, as will become apparent, the adoption of a one-dimensional specification of selection clearly illuminates both the common support considerations as well as the differences in distributions of covariates that would not be addressed by standard OLS.

3.1. Accounting for Variables Affected by the Treatment

A potential problem in our set-up, as in many other applications, e.g. Black and Smith (2004), is endogeneity of variables affecting the main outcome. This will create indirect effects of the treatment (see Robins and Greenland, 1992).

We stress at this point that the conditioning set, \( X \), in (2) of course does not include any concomitant variables potentially affected by motherhood. Therefore, interpretation of (1) assuming

\[^{4}\text{Notice that these assumptions are generally stronger than the mean independence assumption of Heckman et al. (1998). See Lechner (2001) for discussion on this point.}\]
(2) and (3) is that of a net effect of motherhood on wages for the group of mothers. Denote the set of concomitant variables potentially affected by motherhood $S$. For simplicity, in the following discussion let $S$ be one-dimensional. Rosenbaum (1984) examines the consequences for the average treatment effect of including such a variable that has potentially been affected by the treatment in the conditioning set. He finds that, in general, adjustment for such a variable results in unbiased estimates of the average treatment effect only when the variable is in fact not affected by the treatment. This is because the matching estimator integrates over the density of the set of observables conditional on treatment. To estimate correctly the mean outcome for non-mothers this density must not change with treatment.

A way to learn about the effect of $S$ is, however, to consider $s$-specific treatment effects:

$$
\theta(s) = E[Y_1(s)|C = 1, S = s] - E[Y_0(s)|C = 1, S = s]
$$

where $Y_0(s)$ and $Y_1(s)$ are outcomes in the non-motherhood and motherhood states conditional on $S = s$. This is, for example, the effect of motherhood on wages for women observed to be employed in sector $s$. $\theta(s)$ is informative about impacts for a given value of $S$.

Assume that we have sufficient information not only about selection into motherhood but also on the selection into $S$ such that

$$
Y_0(s) \mid C, S \mid \forall s \in S
$$

$$
\Pr(C = 1|S = s, X) < 1 \forall s \in S
$$

i.e., assignment into $C$ and $S$ is strongly ignorable for $Y_0(s)$. Then $\theta(s)$ is identified and is (at the outset) free of indirect effects. $X$ is precisely the set of attributes ensuring potential wage outcome in the non-motherhood state is independent of $C$ and $S$. Here, $C$ and $S$ can be thought of as a new combined type of treatment. Since we only make pairwise comparisons, the proof of the balancing score property follows immediately from Rosenbaum and Rubin (1983).

Note that, in principle, from $\theta(s)$, one can construct a population-weighted average effect of motherhood on wages:

$$
\theta = \sum_s \theta(s) \cdot \frac{\Pr(C = 1, S = s)}{\Pr(C = 1)}
$$

yet the interpretation of $\theta$ is no longer clear. Equation (7) merely represents the average costs of motherhood, but it is not a treatment effect as such.

As above, we still face the curse of dimensionality, now accentuated by the presence of $S$. To estimate $\theta(s)$ we specify $\Pr(C = 1, S = s|X)$ as a bivariate probit, explicitly allowing $S$ to be a function of $C$, select the group of women for whom $S = s$, and match on the probability of motherhood conditional on $S = s$, $\Pr(C = 1|S = s, X)$.

### 3.2. The Linear Case

Apart from sector choice we expect neither experience nor the interruption variables to be exogenous with respect to fertility: presumably, mothers are more likely to interrupt their careers to engage in child-rearing activities, i.e. non-participation or formal child-rearing leave. Yet we
often face the case with insufficient information on the influence of our treatment, \( C \), on these concomitant variables. Denote this extended set of variables potentially affected by the treatment \( \bar{S} \).

To uncover causal relationships, assume additive separability of the outcome equation in line with Heckman et al. (1998):

\[
E[Y_0|C = 1, \bar{S}, X, R] = X\beta_X + \bar{S}\beta_{\bar{S}} + E[U_0|C = 1, P(C = 1|X, R)]
\]

Assume furthermore:

\[
U_0 \perp C, \bar{S}|P(C = 1|X, R) \quad (9)
\]

\[
Pr(C = 1|X, R) < 1 \quad (10)
\]

where \( U_0 \) is the error term in the non-motherhood state. These assumptions mean that assignment into \( C \) and \( \bar{S} \) is strongly ignorable for \( U_0 \). When \( E[U_0|C = 1, \bar{S}, X] \) is specified as a non-parametric function of (potentially continuos) \( X \), (AS) is termed a partial linear model and identification requires the exclusion restriction, \( E[Y_0|C = 1, P(C = 1|X, R), R] = E[Y_0|C = 1, P(C = 1|X, R)] \) (see Robinson, 1988). Invoking these new stronger assumptions we can, of course, now interpret (1) as the direct effect of motherhood.

Similarly, \( \theta(s) \), our \( s \)-specific treatment effects from before, may be suffering from the presence of other variables affected by motherhood. Assume additive separability of the \( s \)-specific wage functions along with

\[
U_0(s) \perp C, S, \bar{S}|P(X, R) \quad (11)
\]

\[
Pr(C = 1|S = s, X, R) < 1 \quad (12)
\]

where \( U_0(s) \) is the error term in the non-motherhood state in sector \( s \). Then the \( s \)-specific direct effects are identified.

As above we still want to allow for heterogeneous treatment effects. We therefore estimate the partial linear model and perform regression-adjusted matching as recently suggested in Heckman et al. (1998) as opposed to OLS. By applying this method we clear out the effects of potential returns in the non-mother state to observable characteristics from \( Y \) and perform matching on the residuals.

4. PROPENSITY SCORE MATCHING

The first step in the empirical analysis is to estimate the probability of being a mother. We model the propensity score by a standard probit. The conditioning set, \( X \), includes age, type of education, e.g. health care, services, technical education, length of education given type, place of habitation, and the woman’s number of siblings.\(^5\) The results are presented in Table II.

We find that age significantly increases the probability of being a mother in 1996 with a decreasing effect, reaching its maximum at the age of 44, which is outside the range of our data.

\(^5\) Note that this variable is our exclusion restriction. It is not included when estimating the propensity used to estimate the net effect of motherhood. The remaining coefficients are not sensitive to the exclusion of the information on number of siblings. Moreover, predictive power and estimated densities do not change.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Motherhood Coeff.</th>
<th>Motherhood SE</th>
<th>Sector Coeff.</th>
<th>Sector SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.92</td>
<td>0.21</td>
<td>-1.03</td>
<td>0.29</td>
</tr>
<tr>
<td>Child</td>
<td>-2.21</td>
<td>0.17</td>
<td>-2.20</td>
<td>0.17</td>
</tr>
<tr>
<td>20–22 years of age</td>
<td>-1.57</td>
<td>0.14</td>
<td>-1.57</td>
<td>0.14</td>
</tr>
<tr>
<td>23–24 years of age</td>
<td>-1.25</td>
<td>0.12</td>
<td>-1.25</td>
<td>0.12</td>
</tr>
<tr>
<td>27–28 years of age</td>
<td>-0.64</td>
<td>0.10</td>
<td>-0.64</td>
<td>0.10</td>
</tr>
<tr>
<td>31–32 years of age</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>33–34 years of age</td>
<td>0.34</td>
<td>0.10</td>
<td>0.25</td>
<td>0.10</td>
</tr>
<tr>
<td>35–36 years of age</td>
<td>0.59</td>
<td>0.10</td>
<td>0.68</td>
<td>0.10</td>
</tr>
<tr>
<td>37–38 years of age</td>
<td>0.46</td>
<td>0.10</td>
<td>0.55</td>
<td>0.10</td>
</tr>
<tr>
<td>39–40 years of age</td>
<td>0.59</td>
<td>0.12</td>
<td>0.67</td>
<td>0.12</td>
</tr>
</tbody>
</table>

**Type of highest completed education:**
- General (0/1)
- Business (0/1)
- Industry (0/1)
- Construction (0/1)
- Graphical (0/1)
- Services (0/1)
- Food and beverages (0/1)
- Agricultural (0/1)
- Pedagogic (0/1)
- Humanistic (0/1)
- Musical (0/1)
- Social (0/1)
- Technical (0/1)
- Unknown (0/1)

**Length of highest completed education:**
- General (0/1) × length
- Business (0/1) × length
- Industry (0/1) × length
- Construction (0/1) × length
- Graphical (0/1) × length
- Services (0/1) × length
- Food and beverages (0/1) × length
- Agricultural (0/1) × length
- Pedagogic (0/1) × length
- Humanistic (0/1) × length
- Musical (0/1) × length
- Social (0/1) × length
- Technical (0/1) × length
- Unknown (0/1) × length

**Note:** Nine regional dummies and interaction terms between age and education dummies are included; reference age group 28–30, educational group health. Bold estimates significant at the 5% level.

*Estimated correlation between error terms, \( \rho = 0.10 \) (0.10).
Hence, the effect is increasing over the relevant range. The variable is, of course, deterministic and thus by no means endogenous to the decision of motherhood. Furthermore, relative to health care-oriented types of education most types of education decrease the probability of being a mother for a given length of education, though some coefficients are insignificant. Note that some may be insignificant due to a small number of individuals in a particular category (see Table I). The analysis also shows that higher level of education of a given type reduces the probability of motherhood. The choice of education is in the vast majority of cases predetermined to the decision of motherhood. We include a set of nine regional dummies to account for effects of local labour markets as well as possible differences in views on having children. According to Munch (2003), Danes mainly settle in connection with their choice of educational institution. After this initial settlement, Danes rarely move between municipalities and hardly ever between regions. For this reason, settlement is expected to be predetermined to fertility. Finally, the woman’s number of siblings significantly increases the probability of motherhood. We assume that this variable satisfies the necessary exclusion restriction for estimating the direct effect (see above).

The model predicts relatively well: 5867 out of 29 210 predictions or 20.1% of all predictions are wrong,\(^6\) and Efron’s \(R^2\) equals 0.42. Figure 1 shows the smoothed densities of the propensity scores for both mothers and non-mothers. It is seen that mothers have more probability mass concentrated around high values of the propensity score compared to non-mothers, who have more mass concentrated around low values of the propensity score. Hence, mothers are likely to differ significantly from non-mothers in terms of observables, meaning that there is a potential gain from matching. Note that the densities seem to have common support: it is possible to find a

\[ \sum_{i=1}^{n} (C_i - 1)_{(\hat{\phi}_i \geq 0.5})^2. \]

Copyright © 2006 John Wiley & Sons, Ltd.
match for a mother among non-mothers, even for mothers with the highest level of the propensity score. Thus the model does not predict too well. Moreover, we investigate the specification of our estimated propensity using the non-parametric test suggested by Shaikh et al. (2005). Formally, we test whether $E[C|P] = P$ and find that the test cannot reject our null at the 1% significance level.\footnote{Note that $E[C|P] = P$ is the sufficient condition for the balancing score property; see Theorem 2 in Rosenbaum and Rubin (1983).}

We choose the optimal bandwidth using cross-validation as in Black and Smith (2004). As expected, owing to our large sample our estimates are not sensitive to the choice of bandwidth or kernel and thus using Silverman’s rule of thumb to choose the bandwidth (Silverman, 1986) instead of cross-validation does not matter. For the same reason, performing nearest-neighbour matching both with and without replacement gives results that are exactly identical to the kernel-based matching analysis. We conduct inference based on bootstrap methods using 99 bootstraps, resampling from the pool of first step predicted propensities, trimming away 2% of the observations with the lowest densities.\footnote{The results are not at all sensitive to reasonable changes in trimming levels, 2–10%. As can be seen from Figure 1, this is due to the existence of ‘enough’ non-mothers in the region with the mass of mothers, $p \in (0.45, 0.95)$.} Therefore, our approximation ignores the fact that the kernel-based matching procedure is based on estimated propensities, and that our enforcement of the common support requirements is data-driven. Incorporating all steps of the estimation in the bootstrap procedure, including re-estimating the propensity along with the corresponding densities, proved to be too computationally expensive. Due to our large sample, we do not expect the uncertainty from the first estimation to inflate the variance to any degree. See also Lechner (2002) on this issue when using nearest-neighbour matching.

We first estimate the net effect of motherhood for the group of mothers, i.e. $\theta$, assuming (2) and (3). We compare log wages for mothers with log wages for non-mothers that are similar in terms of the propensity without making any assumptions on the functional form of the outcome equation. This corresponds to the following expected attribute-adjusted treatment difference (Rosenbaum, 1984):

$$E_X[E[Y_1|C = 1, X] - E[Y_0|C = 0, X]|C = 1]$$

which gives us an estimate of the net effect of being a mother on wages including all effects stemming from concomitant variables potentially affected by the treatment. We see from the first row in Table III that this estimator results in a 6.5% lower wage for mothers. We also obtain reasonably balanced covariates after matching on our estimated propensity score. In no case do the standardized differences in means for covariates (Rosenbaum and Rubin, 1985) exceed 7% (see Simonsen and Skipper, 2004), for details.\footnote{The presented results are robust to changes in the conditioning set $X$. Specifically, we performed the analysis without regional dummies. This resulted in a $-7.9\%$ net effect. Furthermore, changes in the specification of age did not change the results at all. Finally, estimating the model based on the older cohort with completed fertility, 39–40 years old, resulted in a similar net effect of $-7.3\%$.}

We then go on to estimate the direct effect of motherhood, i.e. $\theta$, assuming additive separability along with (9) and (10). This estimator is based on the following expected attribute-adjusted treatment difference:

$$E_{X,\tilde{S},R}[E[Y_1|C = 1, X, \tilde{S},R] - E[Y_0|C = 0, X, \tilde{S},R]|C = 1]$$
Table III. Estimated Treatment Effects and Bootstrapped Standard Errors. Dependent variable: log hourly wage rate in 1997; full sample, 16 012 mothers and 13 198 non-mothers

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Net effect</th>
<th>Direct effect, regression-adjusted matching&lt;sup&gt;a,b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta )</td>
<td>(-0.065^c)</td>
<td>(-0.015^d)</td>
</tr>
<tr>
<td>( \theta(\text{public}) )</td>
<td>(-0.076^e)</td>
<td>(-0.032^f)</td>
</tr>
<tr>
<td>( \theta(\text{private}) )</td>
<td>(-0.064^g)</td>
<td>0.005&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Densities were estimated using a Gaussian kernel and a bandwidth chosen via cross-validation, \( h_{opt} \). The overlapping support region was determined using a 2% trimming rule. Standard errors are based on 99 bootstraps with 100% resampling. Bold estimates significant at the 5% level.

<sup>b</sup> Regression adjustment includes experience, educational categories and length, occupational categories, regional dummies and interruptions from the labour market.

<sup>c</sup> 15 665 mothers and 12 201 non-mothers used: \( h_{rule-of-thumb} = 0.047, h_{opt} = 0.0002 \).

<sup>d</sup> 15 775 mothers and 12 072 non-mothers used: \( h_{rule-of-thumb} = 0.047, h_{opt} = 0.049 \).

<sup>e</sup> 8 433 mothers and 4 826 non-mothers used: \( h_{rule-of-thumb} = 0.057, h_{opt} = 0.005 \).

<sup>f</sup> 8 452 mothers and 4 802 non-mothers used: \( h_{rule-of-thumb} = 0.058, h_{opt} = 0.047 \).

<sup>g</sup> 7 276 mothers and 7 332 non-mothers used: \( h_{rule-of-thumb} = 0.052, h_{opt} = 0.010 \).

<sup>h</sup> 7 264 mothers and 7 359 non-mothers used: \( h_{rule-of-thumb} = 0.053, h_{opt} = 0.045 \).

When specifying the conditioning variables in \( \mathbf{S} \), we follow human capital theory.<sup>10</sup> In addition to the variables in \( \mathbf{X} \) from above we therefore include actual work experience and experience squared, choice of sector, occupation categories, yearly indicators for the timing of the end of the last unemployment and non-participation spells, thereby allowing the effect of interruptions to decrease in time (see Nielsen et al., 2004). Since regression-adjusted matching is meant to clear out the effects of potential returns in the non-mother state, we are forced to pool maternity leave and other publicly subsidized types of leave with the non-participation category to be able to account for the effects of these types of interruptions. We do not expect this assumption to be too restrictive; there are no \textit{ex ante} reasons why depreciation of human capital during non-participation should differ from that of maternity leave.

Figure 2 illustrates the potential effects of interruptions on the earnings potential and emphasizes the link between the theoretical effects and our explanatory variables. We consider a woman who interrupts her career to have a child, engaging in child-rearing activities between age \( A_0 \) and age \( A_1 \). During the interruption, the woman fails to accumulate experience. This simply corresponds to a horizontal shift in the earnings profile. This effect is caught in our model by using actual experience as the conditioning variable. In addition, earnings profiles may shift downwards for two reasons: firstly, the existence of a significant child penalty causes a vertical shift in the earnings profile; and secondly human capital may depreciate while interrupting, shifting the earnings potential further downwards depending on the duration of the interruption. To account for the latter, we include the duration of the latest interruption spell, thus allowing for linear depreciation as illustrated in the figure. Finally, there may be a catching up effect: Women may regain part

<sup>10</sup> See Simonsen and Skipper (2004) for a full list of the conditioning variables used in regression-adjusted matching along with coefficients from estimating the partial linear regression.
of (or all) the effect of lost experience and depreciation when they return to work: a ‘recovery phase’. In our model this is allowed for by including dummies for the year of the latest interruption.

$R$ is the woman’s own number of siblings. It may on all reasonable grounds be left out of the wage function; we do not expect number of siblings to affect labour market outcomes conditional on the $X$ and $S$, and it is definitely predetermined to wage determination. Furthermore, the variable is correlated with the choice of having children since it enters the selection equation with a significant coefficient. Butcher and Case (1994) and Ermisch and Francesconi (2001) both argue that number of siblings affects wage determination only through the individual’s level of education, which supports our exclusion restriction: we include the woman’s type and length of education in the wage function and, conditional on that information, the woman’s own number of siblings should not explain wages.

We find a much smaller penalty of 1.5% when conditioning on the additional information in $S$ (see again Table III). In fact, this corresponds to a reduction of almost 80% compared to the net effect. Remember, though, that the identifying assumptions needed to recover the direct effect of being a mother are strong! It must, for example, be the case that the expected unobserved characteristic for a mother is the same as for a non-mother conditional on them having the same $X$ and $S$, even though the distribution of experience may differ among the two groups. Note also that these assumptions—among with an assumption of constant treatment effect—are commonly (more or less implicitly) invoked in the existing literature.

Both sets of results are small in an international comparison, which is possibly due to institutional factors. The direct effect is in line with the Scandinavian findings from Datta Gupta and Smith (2002) and Nielsen et al. (2004). This is remarkable given the distributions of observable characteristics among mothers and non-mothers, see Figure 1 and discussion above.

Figure 2. Earnings potential. —— no interruptions, no family gap; - - - - experience foregone, no family gap; experience foregone, family gap, depreciation, catching up; - - - experience foregone, family gap, depreciation, no catching up
5. SECTOR-SPECIFIC EFFECTS

With sufficiently rich information to make outcome in the non-motherhood state conditionally independent of both motherhood and $S$, we can identify direct $s$-specific without assumptions on the functional form of the outcome equation. In this section, we consider treatment effects in the public and private sector, $\theta(\text{public})$ and $\theta(\text{private})$. We estimate both net effects, assuming equations (5) and (6) along with direct effects, assuming additive separability in addition to equations (11) and (12).

As in the section above, we start out with estimation of the propensities. We use a bivariate probit model to estimate the probabilities of being a mother and working in the private sector, allowing the error terms in the two equations to be correlated. From this model we can get the predicted conditional propensities we need to identify the parameters of interest. The selection into motherhood is modelled using the same variables as before, including the information on number of siblings when estimating sector-specific direct effects. We argue that number of siblings is uncorrelated with both wage outcome (see discussion above) and sector choice. This argument relies on conditioning on $X$. The inclusion serves two purposes: firstly, we avoid that identification in the motherhood-sector choice model relies too heavily on the joint normality of the error terms; and secondly, it allows us to perform regression-adjusted matching. In the sector selection equation we condition on motherhood indicator, type of education, length of education given type and place of habitation outside capital area (greater Copenhagen). Furthermore, we include parental information on sector choice. That is, $X$ includes the same set of variables as above, augmented by the latter variables. Regarding estimation of the direct effects, $S$ is unchanged except for the exclusion of the sector indicator.

The results from the bivariate probit can be seen in Table II.\textsuperscript{11} It is obvious that the coefficients in the motherhood equation have not changed substantially by allowing the error term to be correlated with sector choice. In line with our hypothesis, the coefficients from the sector choice equation indicate that mothers are more likely to choose to work in the public sector. In addition, most types of education increase the probability of working in the private sector compared to health care-related types of education, yet the effects of length of education vary.

Rows 2 and 3 in Table III show the results from the matching analysis taking endogeneity of sector into account and demonstrate clear differences between sectors, in particular with respect to the direct effects: the net effect amounts to 7.6% lower wages in the public sector compared to 6.4% in the private sector, whereas the direct effect corresponds to a 3.2% penalty in the public sector and no penalty in the private sector. Note that we cannot use the estimates of the net effects to conclude that being a mother is more expensive in the public sector. The difference in net effects may reflect that mothers in the public sector make different choices from mothers in the private sector in terms of non-participation, part time, choice of occupation etc. In principle, in the public sector, wages for mothers that have been on maternity or child-rearing leave should not differ from non-mothers’ wages, everything else being equal, since wages are highly correlated with seniority in particular in the period under consideration. However, being on leave may affect a mother’s chances of getting a qualification bonus in the public sector. Along with promotion opportunities this is the main incentive scheme in the public sector. Since it is not related to level of experience, non-participation etc., we must expect it to show up in the direct effect. The penalty

\textsuperscript{11} To evaluate the predictive power of the bivariate probit we calculate the number of right predictions by comparing the predicted state, characterized by the maximum of the predicted probabilities over the four states, to the realized state. In 54.4% of the times the model predicts the real state, compared to approximately 25% had we had no model (see Table I).
could also be caused by job flexibility, a non-pecuniary benefit, within the public sector: besides having extra care days targeted towards children, mothers may be reallocated to less stressful or time-consuming jobs, for example jobs involving only standard hours. It is noteworthy that the direct effect is insignificant in the private sector. One explanation for this may be that wages in the private sector are much more flexible such that wage outcome may better reflect differences in observables.

Nielsen et al. (2004) model the effect of motherhood in the public and private sectors in Denmark using a switching regime model and find results that at first glance seem to contradict the results from this analysis: the wage effect of motherhood is positive in the public sector (3.0%) and negative in the private sector (−6.0%). Importantly, the (untestable) identifying assumptions underlying the selection model employed in Nielsen et al. (2004) are not comparable to those of this paper. In addition, the estimated effects in Nielsen et al. (2004) are assumed to be constant over the population. The interpretation of their estimates are therefore the effect of motherhood in a given sector for a woman drawn randomly from the population.

6. CONCLUSION

We contribute to the existing literature on the effect of motherhood on wages by using high-quality data and by implementing propensity score matching that in many respects is less restrictive than what has been used up until now. We estimate both net and direct effects of motherhood on wages as well as sector-specific effects. We interpret the direct effect as the causal effect of motherhood on wages. As pointed out in footnote 1, this effect is potentially made up by several components. The net effect includes the direct effect as well as indirect effects resulting from choices caused by motherhood such as mothers’ potentially lower labour market experience, higher level of non-participation, and greater probability of working in the public sector, which provides better working conditions for families, to mention a few examples. The net effect is an estimate of the total wage cost of having children whereas the direct effect is the causal effect on wages of being a mother.

We find significantly negative impacts on wages for mothers, yet the direct effect is small. Hence, most of the difference can be explained by accounting for covariates that are likely to be affected by having children. The results are small in an international comparison, which is possibly due to institutional factors, but in line with findings from Scandinavia. We conclude that mothers self-select into the public sector where non-pecuniary benefits related to motherhood are larger. The direct effect of motherhood is negative in the public sector and positive but insignificant in the private sector. We argue that it is likely to be due to considerable job flexibility in the public sector.

To date, most empirical work has focused little on endogeneity issues when trying to identify direct effects. Furthermore, whether endogenous variables have any substantial impact on the parameter of interest is an empirical question. In our case, allowing for sector-specific treatment effects clearly informs on costs of motherhood.

ACKNOWLEDGEMENTS

We thank the Danish Social Science Research Council (the FREJA grant) for financial support and the National Centre for Register-Based Research for support to access data from Statistics Denmark. We wish to thank Stefan Bender, Martin Browning, Nabanita Datta Gupta, Michael

Lechner, Shelly Lundberg, Helena S. Nielsen, Amalia R. Miller, Michael Rosholm and Ed Vytlacil, along with Herman van Dijk and two anonymous referees, for their helpful comments and suggestions. We also appreciate comments from participants at the EALE Conference 2003, the Labour Economics Discussion Group at the Department of Economics, Stanford University, and the Empirical Evaluation of Labour Market Programmes Conference 2005. All errors are our own.

REFERENCES


