The Costs of Motherhood: An Analysis Using Matching Estimators.

Marianne Simonsen*
Department of Economics
University of Aarhus

Lars Skipper
Institute for Local
Government Studies

Abstract

In this paper we analyse whether motherhood causes lower wages for women. Importantly, covariates are likely to be affected by motherhood giving rise to indirect effects. We estimate net effects of motherhood on wages implementing matching and find significantly negative impacts. To learn about direct effects we implement two strategies: Firstly, we confine the analysis to consider sector specific treatment effects. We find differences in impacts across sectors, likely to be due to considerable job flexibility in the public sector. Secondly, we impose additive separability on the outcome equation and perform residual based matching and find small direct effects.

JEL Classifications: C13, C14, C35, J13, J31, J45, J71

Keywords: Fertility, family gap, wages, matching.

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 Edward J. Vytlacil, Shelly Lundberg, Nabanita Datta Gupta, Michael Rosholm, Helena Skyt Nielsen, and Amalia R. Miller for their helpful suggestions. We also appreciate comments from participants at the EALE Conference 2003 and the Labor economics discussion group at Dept. of Economics, Stanford University. All errors are our own.

*Corresponding address: Marianne Simonsen, Dept. of Economics, University of Aarhus, Building 326, 8000 Aarhus C, Denmark. Email: msimonsen@econ.au.dk. Phone: +45 8942 1599. Fax: +45 8613 6334.

1 Introduction

The most common counterfactual of interest in the treatment evaluation literature is the mean direct effect of treatment on the treated (Heckman, LaLonde & Smith (1999)). In the case where the subpoulation under study consists of women, the treatment is being a mother, and the outcome is wages this counterfactual has in the literature been designated 'the family gap' or, alternatively, 'the child penalty' (Budig & England (2001), Phipps, Burton & Lethbridge (2001), Waldfogel (1998*a*),Waldfogel (1998*b*)). The naming of the counterfactual is, of course, a direct consequence of the results found in the existing literature: Having a child seems to be costly in terms of wages.[1] 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 & Neumark (1992) and Robins & 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 wages has, by far, received the most attention in the literature (Nielsen, Simonsen & Verner (2004), Phipps et al. (2001), Waldfogel (1998*a*)) yet the recovering of the parameter is not immediate due 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 as emphasized in a recent paper by Black & 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 find 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 childbearing.

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 heterogenous 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 e.g. income, demographics, and education on a yearly basis. Furthermore, the individual event history in terms of periods of employment, unemployment, as well as maternal leave and other types of publicly subsidised leaves are known on a weekly basis. We discuss the assumptions needed for identification of direct and net effects of being a mother when variables are affected by the treatment. If the researcher has sufficient information to make wage outcomes statistically 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 characterised 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 all other things being equal, cf. 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, therefore, we take advantage of an exclusion restriction to perform regression adjusted matching recently suggested by Heckman, Ichimura & Todd (1998). To our knowledge matching analysis has never been applied in this area before and we know of very few examples of the implementation of regression adjusted matching in general (e.g. Smith & Todd (2004)). We find significant family gaps no matter the choice of matching estimator but the results vary between public and private employees. The direct effect is small indicating that most of the difference in wages can be explained by variables that are likely to be affected by having children.

The paper is organised as follows. In the next section we discuss our sample of data. In Section 3 we discuss our parameter 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 all 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 subsidised 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 in the state 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 1 shows descriptive statistics on selected variables 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 have 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 hours 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

1 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]

[**Table 1 around here**]

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 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 (four 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 is extended with ten weeks of leave amounting to a maximum of 28 weeks and fathers are granted two 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 two weeks or less and the mode is two weeks (cf. Statistics Denmark - www).

In 1994, publicly subsidised 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 eight, while sabbatical leave amounts to a maximum of 52 weeks. The length of these two types of leaves is registered from 1995 and onwards.

Furthermore, the public sector is characterized by being 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 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 ten fully funded care days per child per year. Finally, apart from the existence of a smaller qualification bonus wages in the public sector are mechanically determined by seniority. Therefore, mothers in the public sector do not risk loosing experience based increases in wages when taking maternity leave. Table 2 shows the distribution of mothers in the two sectors.

[Table 2 around here]

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 'un-treated' 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

$$\begin{aligned} \theta &\equiv E\left[Y_1 - Y_0 | C = 1\right] \qquad (\mathbf{1}) \\ &= E\left[Y_1 | C = 1\right] - E\left[Y_0 | C = 1\right]. \end{aligned}$$

Hence, the problem turns to that of finding the counterfactual $E\left[Y_0 | C = 1\right]$ 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 those women who choose to have children, (**1**), and not the mean effect for the population of women. With potential heterogenous impacts these parameters will likely differ.[4] 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, that 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, $\mathbf{X}$, 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 outcome distribution of those having children is the same as the observed outcome distribution of those without. 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 & Rubin (1983)[5]

$$Y_0 \amalg C|\mathbf{X} \qquad (\mathbf{A\text{-}1})$$

In other words, we assume that the outcome for a mother *had she not had children* is the same as the outcome for a non-mother with the same observed characteristics. 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 be able to utilise (**A-1**) it is necessary to make sure that there is a woman without a child analogue to each mother, the 'common support' assumption, i.e.,

$$P \equiv \Pr(C = 1\,|\mathbf{X}) < 1. \qquad (\mathbf{A\text{-}2})$$

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 nonparametric 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 & Rubin (1983). Here the focus is shifted from the set of covariates to the probability of motherhood, $P = \Pr(C = 1\,|\mathbf{X})$. As long as (**A-1**) and (**A-2**) hold,

$$Y_0 \amalg C|P. \qquad (\mathbf{A\text{-}3})$$

This new conditioning variable, P, changes CIA into (**A-3**) which together with $P < 1$ are sufficient conditions required in order to justify propensity score matching to estimate the mean impact on the treated, see Heckman et al. (1998). Clearly, the functional form of P is rarely known and has to be estimated, shifting the high-dimensional estimation problem from that of estimating $E[Y|\mathbf{X}]$ to that of estimating $E[C|\mathbf{X}]$, often estimated by logit or, as in this paper, a probit. See discussion in Black & Smith (2004) on this issue. Moreover, as will become apparent, the adaption 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 determined by the treatment

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

We stress at this point that the conditioning set, $\mathbf{X}$, in (**A1**) *of course* does not include any concomitant variables potentially affected by motherhood. Therefore, interpretation of (**1**) assuming (**A1**) and (**A2**) 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 $\mathbf{S}$. For simplicity, in the following discussion let $\mathbf{S}$ 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 variables results in unbiased estimates of the average treatment effect only when the variables are 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*:

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

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 information not only about selection into motherhood but also on the selection into S such that:

$$Y_0(s) \amalg C, S|\mathbf{X} \ \forall \ s \in S \tag{\textbf{B-1}}$$

$$\Pr\left(C=1, S=s|\mathbf{X}\right) < 1 \ \forall \ s \in S \tag{\textbf{B-2}}$$

i.e. assignment into C and S is strongly ignorable for $Y_0(s)$ then $\theta(s)$ is identified despite the presence of S and is (on the outset) free of indirect effects. $\mathbf{X}$ is precisely the set of attributes ensuring potential wage outcome in non-motherhood be independent of C and S. In fact, by Lemma 4.2, (i) in Dawid (1979), if $Y_0(s) \amalg C|\mathbf{X}$ and S is a function of C

then it follows that $Y_0(s) \amalg S|\mathbf{X}$. This corresponds well with our idea that S is caused by C. Alternatively, C and S can be thought of as a new combined type of treatment.

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

$$\theta = \sum_S \theta(s) \cdot \frac{\Pr(C = 1, S = s)}{\Pr(C = 1)} \tag{2}$$

yet the interpretation of θ is no longer clear. (2) 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 match on the probability of motherhood conditional on S, $\Pr(C = 1|S = s, \mathbf{X})$. We therefore specify $\Pr(C = 1, S = s|\mathbf{X})$ as a bivariate probit, explicity allowing S to be a function of C.

3.2 The Linear Case

Apart from sector choice we expect neither experience nor the interruption variables to be exogenous wrt. 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 $\tilde{\mathbf{S}}$.

To uncover causal relationships, assume that we have access to exclusion restrictions, $\mathbf{R}$. Furthermore, assume additive separability of the outcome equation:

$$E\left[Y_0|C = 1, \tilde{\mathbf{S}}, \mathbf{X}, \mathbf{R}\right] = \mathbf{X}\beta_X + \tilde{\mathbf{S}}\beta_{\tilde{\mathbf{S}}} + E\left[U_0|C = 1, \tilde{\mathbf{S}}, \mathbf{X}, \mathbf{R}\right] \tag{\textbf{2}}$$

along with

$$U_0 \amalg C, \tilde{\mathbf{S}}|\mathbf{X}, \mathbf{R} \tag{\textbf{C-1}}$$

$$\Pr\left(C = 1|\mathbf{X}, \tilde{\mathbf{S}}, \mathbf{R}\right) < 1 \tag{\textbf{C-2}}$$

where U_0 is the error term in the non motherhood state. These assumptions mean that assignment into C and S is strongly ignorable for U_0. When $E[U_0|C = 1, S, \mathbf{X}]$ is specified as a nonparametric function of (potentially continous) $\mathbf{X}$, (**2**) is termed a partial linear model and identification requires exclusion restrictions, $\mathbf{R}$, 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:

$$U_0(s) \amalg C, S, \tilde{\mathbf{S}}|\mathbf{X}, \mathbf{R} \tag{\textbf{D-1}}$$

$$\Pr(C = 1, S = s|\mathbf{X}, \mathbf{R}) < 1 \tag{\textbf{D-2}}$$

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 heterogenous 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. We utilise the parental information as exclusion restrictions. By applying this new 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.[6] We condition on age, type of education e.g. health care, services, technical education etc., length of education given type, place of habitation outside capital area (greater Copenhagen), and the woman's number of siblings.[7] The results are presented in Table 3.

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. 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 1. The education variables presumably capture some effects of preferences for having children. 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 but we acknowledge that unobserved factors may affect both the choice of education and fertility. As long as the choice to become a mother is conditionally independent of outcome in the non motherhood state holds this is not a problem. Habitation outside the capital area also increases the probability of motherhood. This variable is meant to capture the effects of a possibly more traditional view on having children. Finally, the woman's number of siblings significantly increases the probability of having children. This variable will serve as our exclusion restriction when performing residual based matching below.

[Table 3 around here]

[Figure 1 around here]

The model predicts relatively well: 5,867 out of 29,210 predictions or 20.1% of all predictions are wrong[8] 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 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.

We choose the optimal bandwith using cross validation as in Black & Smith (2004). Due to our large sample our estimates are not sensitive to the choice of bandwidth or kernel and thus using Silverman rule of thumb to choose the bandwidth (Silverman (1986)) instead of cross validation does not matter. Also, performing nearest neighbour matching both with and without replacement give results that are exactly identical to the local linear matching analysis. This is remarkable given that the estimates are based on very different subsamples but obviously because of the large estimation sample.

We firstly estimate the net effect of motherhood for the group of mothers, that is θ assuming (**A1**) and (**A2**). 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)):

$$E_{\mathbf{X}}\left[E\left[Y_1|C=1,\mathbf{X}\right]-E\left[Y_0|C=0,\mathbf{X}\right]\right],$$

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 4 that this estimator results in a 6.9% lower wage for mothers; a result in the lower end of the estimates in the international literature. We also conclude that we obtain reasonably balanced covariates after matching on our estimated propensity score. In no case do the standardized differences in means for covariates (Rosenbaum & Rubin (1985)) exceed 7%, see Simonsen & Skipper (2004) for details.

We then go on to estimate the direct effect of motherhood, i.e. θ assuming additive separability along with (**C1**) and (**C2**). This estimator is based on the following expected attribute adjusted treatment difference:

$$E_{\mathbf{X},\tilde{\mathbf{S}},R}\left[E\left[Y_1|C=1,\mathbf{X},\tilde{\mathbf{S}},R\right]-E\left[Y_0|C=0,\mathbf{X},\tilde{\mathbf{S}},R\right]\right].$$

When specifying the conditioning variables, $\mathbf{X}$ and $\tilde{\mathbf{S}}$, we follow traditional human capital theory.[9] Apart from the variables included in standard wage equations such as actual work experience and experience squared, very specific types of education, length of these types of education, choice of sector, and occupation categories we include yearly indicators for the timing of the end of the last unemployment and nonparticipation spells, thereby allowing the effect of interruptions to decrease in time (cf. Nielsen et al. (2004)). Since regression adjusted matching clears out the effects of potential returns in the non-mother state we are forced to pool the non-participation category with maternity leave and other publicly subsidised types of leave to be able to account for the effects of these types of interruption We do not expect this assumption to be too restrictive; there are no ex ante reason 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 emphasises 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 A0 and age A1. 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. Human capital depreciation while interrupting may cause a gradual decline in wages with the effect possibly depending on the duration of the interruption. We allow for linear depreciation as illustrated in the figure and condition on the total duration of the interruption spells of a given type. Finally, there may be a catching up effect: Women may regain part of (or all) the effect of lost experience and depreciation when they return to work. In the literature, the period of catching up has been labelled "the recovery phase". In our model this is allowed for by including dummies for the year of the latest interruption.

As our exclusion restriction, R, we use 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 putcomes conditional on the $\mathbf{X}$ and $\tilde{\mathbf{S}}$ 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 & Case (1994) and Ermisch & Francesconi (2001) both argue that number of siblings affect wage determination only through the individual's level of education, which supports our use of the variable as an 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 this additional information, see again Table 4. 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 $\mathbf{X}$ and $\tilde{\mathbf{S}}$ even though the distribution of experience may differ among the two groups. Note also that these assumptions - *along* with an assumption of constant treatment effect - are commonly made in the existing literature.

In Appendix A we discuss sensitivity of our results and check the internal consistency of the model.

[Table 4 around here]

[Figure 2 around here]

5 Propensity Score Matching Analysis: Accounting for Sector choice

With sufficiently rich information to make outcome in the nonmotherhood 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, $\theta(public)$ and $\theta(private)$.

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 simultaneous and 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.

[Table 5 around here]

When modelling the sector selection 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. The information on number of siblings again serve as exclusion restriction. Importantly, the exclusion restriction remain significant and the coeffecient seems robust to the change of model. We argue that number of siblings is uncorrelated with both wage outcome (cf. discussion above) and sector choice. This argument relies on conditioning on $\mathbf{X}$. 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 residual based matching.

The results from the bivariate probit can be seen in Table 5.[10] It is obvious from Table 3 and 5 that the coefficients in the motherhood equation have not changed substantially by allowing the error term to be correlated with sector choice. The coefficients from the sector choice equation indicate that mothers are more likely to choose to work in the public sector in line with with our hypothesis, most types of education increase the probability of working in the private sector compared to health care related types of education, whereas the effect of length of education vary.

Rows 2-3 in Table 4 show the results from the matching analysis taking endogeneity of sector into account and demonstrates clear differences between sectors, in particular wrt. the direct effects. 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 may to a large degree reflect that mothers in the public sector make different choices than 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 given 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 could also be caused by job flexibility, a nonpecuniary benefit, within the public sector: Besides having extra care days targeted towards childrens needs mothers may be reallocated to less stressful or timeconsuming jobs, for example jobs involving only standard hours.

It is noteworthy that the direct effect of motherhood is positive yet insignificant in the private sector. One explanation for the size of the direct wage effect in the private sector may be that wages in the private sector are much more flexible such that wage outcome may better reflect differences in observables. Though insignificant, the fact that we find a wage *premium* in the private sector may be somewhat surprising. For us to be able to interpret the parameter as a direct effect it must be the case that becoming a mother in the private sector *causes* the woman to make career choices that increase her productivity. This could happen if the woman wishes to stay in her job and expects that becoming a mother may negatively affect her chances of continued employment and promotion in the future. To counter this she increases her productivity upon returning after giving birth. Another explanation could be compensation due to lack of job flexibility in the private sector: It is supposedly more costly for mothers than for nonmothers to hold stressful positions leading some mothers to select into the public sector as indicated by the mother indicator in the biprobit. Mothers staying in the private sector seem to be succesful in receiving compensation for the increase in disutility of work.

Nielsen et al. (2004) also consider the effect of motherhood in the public and private sector and find results that seem to contradict the results from this analysis: They find that the effect of motherhood is positive in the public sector and negative in the private sector. The method differs from the one employed in this paper along many dimensions but the most important difference is that the estimated treatment effects are average treatment effects and not average treatment effects for the group of mothers. Hence the interpretation of the estimates from Nielsen et al. (2004) would be the effect of motherhood in a given sector for a woman drawn randomly from the population of women. Furthermore, when allowing for incomplete selection and conditioning on sector choice for a standardized woman they find that the most persistent log wage gap is seen in the counterfactual case; women who are actually public sector workers would, in particular, have been penalized for having children if they had been employed in the private sector. This corresponds well with the conclusion from this paper.

6 Discussion

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. This may include both discrimination, statistical discrimination, and bargaining differences. The net effect includes the direct effect as well as indirect effects resulting from choices causes 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 of being a mother on wages.

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 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 changes the conclusion regarding costs of motherhood.

References

Behrman, J. R., Cheng, Y. & Todd, P. E. (2004), 'Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach', *The Review of Economics and Statistics* **86**(1), 108–132.

Black, D. A. & Smith, J. A. (2004), 'How Robust is the Evidence on the Effects of College Quality? Evidence from Matching', *Journal of Econometrics* **121**, 99–124.

Budig, M. J. & England, P. (2001), 'The Wage Penalty for Motherhood', *American Sociological Review* **66**, 204–225.

Butcher, K. & Case, A. (1994), 'The Effect of Sibling Composition on Women's Education and Earnings', *Quarterly Journal of Economics* **109**, 531–563.

Dawid, A. P. (1979), 'Conditional independence in statistical theory', *Journal of the Royal Statistical Society B* **41**, 1–31.

Ermisch, J. & Francesconi, M. (2001), 'Family Matters: The Impact of Family Background on Educational Attainments', *Economica* **68**, 137–156.

Heckman, J., Ichimura, H. & Todd, P. (1998), 'Matching as an Econometric Evaluation Estimator', *Review of Economic Studies* **65**, 261–294.

Heckman, J., LaLonde, R. & Smith, J. (1999), The Economics and Econometrics of Active Labor Market Programs, *in* A. Ashenfelter & D. Card, eds, 'Handbook of Labor Economics, Vol.', Vol. 3, Elsevier Science, Amsterdam, pp. 1865–2097.

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

Lechner, M. (2001), *Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption*, Physica/Springer, chapter 3.

Nielsen, H. S., Simonsen, M. & Verner, M. (2004), 'Does the Gap in Family-Friendly Policies Drive the Family Gap?', *Scandinavian Journal of Economics* **106(4)**, 721–744.

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.

Robins, J. M. & Greenland, S. (1992), 'Identifiability and Exchangeability for Direct and Indirect Effects', *Epidemiology* **3**(2), 143–155.

Robinson, P. M. (1988), 'Root-N-Consistent Semiparametric Regression', *Econometrica* **56**, 931–954.

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**(5), 656–666.

Rosenbaum, P. R. & Rubin, D. B. (1985), 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score', *The American Statistician* **39**(1), 33–38.

Rosenbaum, P. & Rubin, D. (1983), 'The Central Role of the Propensity Score in Observational Studies for Causal Effects', *Biometrika* **70**(1), 41–55.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall/CRC.

Simonsen, M. & Skipper, L. (2004), Identifying Direct and Indirect Effects. Estimating the Costs of Motherhood Using Matching Estimators. SIEPR working paper, 3-23.

Smith, J. A. & Todd, P. (2004), 'Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?', *Journal of Econometrics* **125**, 305–353.

Waldfogel, J. (1998*a*), '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.

Waldfogel, J. (1998*b*), 'Understanding the "Family Gap" in Pay for Women with Children', *Journal of Economic Perspectives* **12**, 137–156.

A Sensitivity Analysis

It is possible that the treatment effect varies with age. One possible test is to calculate ATET conditional on completed fertility. We perform the matching analysis for the part of the sample aged 39-40. Women aged 39 or older give birth to 2.7% of all children born in Denmark in 1997. Some of these children are obviously not the women's first child.

Hence, choosing 39 years of age as the cut-off point does not seem too unreasonable. We do not find significant changes in the estimated family gaps.

To check for internal consistency of our model, we estimate the average treatment effect on the untreated (TUT): The effect of having children in terms of wages for the group of women who do not have children. We find that our estimates of TUT do not differ from the estimates of ATET. I.e. it is not the case that women who do not have children would benefit in terms of wages for having so indicating a serious problem with the procedure for accounting for unobserved characteristics. All results are available from the author's on request.

Notes

[1]The family gap may exist for several reasons: mothers may invest differently in household production (resulting in e.g. 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.

[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 in 1996.

[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.

[4]All former studies, see Section 1, evaluating the family gap, have assumed a common treatment effect.

[5]Notice that these assumptions are generally stronger than the the mean independence assumption of Heckman et al. (1998). See Lechner (2001) for discussion on this point.

[6]Note that strictly speaking we do not need the probit to be the exact propensity score. Only do we need it to balance the distributions of our attributes - hence we need it to be a balancing score.

[7]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 as well as estimated densities do not change.

[8]We define number of wrong predictions as $\sum_{i=1}^{n}\left(C_i - \mathbf{1}_{(\widehat{p}_i \geq .5)}\right)^2$.

[9]See Simonsen & Skipper (2004) for a full list of the conditioning variables used in regression adjusted matching along with coefficients from estimating the partial linear regression.

[10]To evaluate the predictive power of the bivariate probit we calculate the number of right predictions by comparing the predicted state, characterised by the maximum of the predicted probabilites over the 4 states, to the realized state. In 54.4% of the times the model predicts the real state. This should be compared to approximately 25% had we had no model, see Table 2.

TABLE 1[a]

SELECTED MOMENTS, DESCRIPTIVE STATISTICS

Variables[b]	All	Mothers	Non-mothers
Log wages	4.80	4.81	4.79
	(0.28)	(0.26)	(0.30)
Age (years)	30.71	33.87	26.88
	(5.80)	(4.20)	(5.13)
Experience (years)	7.27	9.39	4.70
	(4.87)	(4.29)	(4.25)
Length of completed education (years)	12.23	12.20	12.26
	(2.45)	(2.42)	(2.48)
Type of highest completed education:			
General (0/1)	0.22	0.19	0.26
Business (0/1)	0.34	0.33	0.35
Industry (0/1)	0.01	0.01	0.01
Construction (0/1)	0.01	0.01	0.01
Graphical (0/1)	0.01	0.01	0.01
Services (0/1)	0.02	0.02	0.02
Food and beverages (0/1)	0.04	0.04	0.03
Agricultural (0/1)	0.01	0.01	0.01
Transportation (0/1)	0.00	0.00	0.00
Pedagoghic (0/1)	0.06	0.09	0.04
Humanistic (0/1)	0.03	0.03	0.03
Musical (0/1)	0.00	0.00	0.00
Social (0/1)	0.04	0.03	0.05
Technical (0/1)	0.02	0.01	0.02
Public security (0/1)	0.00	0.00	0.00
Unknown (0/1)	0.08	0.07	0.09
Owner of real estate (0/1)	0.42	0.55	0.26
Married (0/1)	0.43	0.66	0.15
Province (0/1)	0.63	0.68	0.57
Private sector (0/1)	0.52	0.46	0.60
High level occupation (0/1)	0.10	0.10	0.10
Medium level occupation (0/1)	0.21	0.23	0.18
Low level occupation (0/1)	0.52	0.51	0.53
Total duration of unemployment (weeks)	64.20	87.40	36.20
	(94.40)	(106.90)	(66.60)
Total duration of nonparticipation (weeks)	176.90	162.10	195.00
	(119.80)	(119.60)	(117.60)
Number of siblings when 15-17 years of age	1.16	1.00	1.35
	(1.02)	(1.05)	(0.95)
Number of siblings missing (0/1)	0.19	0.23	0.14
Sample size	29210	16012	13198

[a] Standard deviations shown in parenteses.

[b] Omitted educational type is 'health care', omitted occupation category is lowest level.

TABLE 1 CTD.[a]

SELECTED MOMENTS, DESCRIPTIVE STATISTICS

Variables[b]	All	Mothers	Non-mothers
Last unemployment spell in 1981 (0/1)	0.01	0.01	0.00
Last unemployment spell in 1982 (0/1)	0.01	0.01	0.00
Last unemployment spell in 1983 (0/1)	0.01	0.02	0.00
Last unemployment spell in 1984 (0/1)	0.01	0.02	0.01
Last unemployment spell in 1985 (0/1)	0.02	0.02	0.01
Last unemployment spell in 1986 (0/1)	0.02	0.03	0.01
Last unemployment spell in 1987 (0/1)	0.02	0.04	0.01
Last unemployment spell in 1988 (0/1)	0.03	0.04	0.01
Last unemployment spell in 1989 (0/1)	0.03	0.05	0.02
Last unemployment spell in 1990 (0/1)	0.05	0.07	0.03
Last unemployment spell in 1991 (0/1)	0.04	0.04	0.03
Last unemployment spell in 1992 (0/1)	0.05	0.05	0.05
Last unemployment spell in 1993 (0/1)	0.07	0.07	0.06
Last unemployment spell in 1994 (0/1)	0.07	0.07	0.08
Last unemployment spell in 1995 (0/1)	0.09	0.08	0.10
Last unemployment spell in 1996 (0/1)	0.20	0.21	0.20
Last nonparticipation spell in 1981 (0/1)	0.00	0.00	0.01
Last nonparticipation spell in 1982 (0/1)	0.01	0.00	0.01
Last nonparticipation spell in 1983 (0/1)	0.01	0.00	0.02
Last nonparticipation spell in 1984 (0/1)	0.02	0.01	0.02
Last nonparticipation spell in 1985 (0/1)	0.02	0.02	0.03
Last nonparticipation spell in 1986 (0/1)	0.03	0.02	0.03
Last nonparticipation spell in 1987 (0/1)	0.03	0.02	0.03
Last nonparticipation spell in 1988 (0/1)	0.03	0.03	0.04
Last nonparticipation spell in 1989 (0/1)	0.04	0.04	0.05
Last nonparticipation spell in 1990 (0/1)	0.05	0.04	0.06
Last nonparticipation spell in 1991 (0/1)	0.06	0.05	0.07
Last nonparticipation spell in 1992 (0/1)	0.07	0.06	0.08
Last nonparticipation spell in 1993 (0/1)	0.09	0.08	0.11
Last nonparticipation spell in 1994 (0/1)	0.10	0.08	0.13
Last nonparticipation spell in 1995 (0/1)	0.20	0.21	0.20
Last nonparticipation spell in 1996 (0/1)	0.18	0.31	0.03

TABLE 2

NUMBER OF MOTHERS AND NON-MOTHERS IN EACH SECTOR

	Mothers, $C = 1$	Non-Mothers, $C = 0$
Public Sector, $S = 0$	8602 (62,0%)	5280 (38,0%)
Private Sector, $S = 1$	7410 (48,3%)	7918 (51,7%)

TABLE 3

Coefficient Estimates and Asy. Std. Err. from Motherhood Probit

Dep. variable: 1 for Mothers, 0 for Non-Mothers

Full Sample, 16,012 Mothers and 13,198 Non-Mothers

Variables[a]	Coeff	Asy Std Error
Intercept	**−12.59**	0.40
Age (years)	**0.76**	0.02
Age squared	**−0.01**	0.00
General (0/1)	0.37	0.24
Business (0/1)	**−1.83**	0.23
Industry (0/1)	1.19	1.00
Construction (0/1)	0.36	1.16
Graphical (0/1)	−0.56	1.41
Services (0/1)	0.18	0.59
Food and beverages (0/1)	−0.47	0.41
Agricultural (0/1)	−0.53	0.46
Transportation (0/1)	−3.26	2.69
Teaching (0/1)	0.27	0.50
Humanities (0/1)	**−1.22**	0.51
Musical (0/1)	−0.94	1.19
Social (0/1)	**−2.78**	0.52
Technical (0/1)	−0.91	0.63
Public security (0/1)	**−0.70**	1.56
Unknown (0/1)	**−0.97**	0.24
Length of education	**−0.11**	0.01
General (0/1)*length	**−0.08**	0.02
Business (0/1)*length	**0.13**	0.02
Industry (0/1)*length	−0.13	0.08
Construction (0/1)*length	−0.05	0.10
Graphical (0/1)*length	0.11	0.12
Services (0/1)*length	−0.04	0.05
Food and beverages (0/1)*length	0.02	0.03
Agricultural (0/1)*length	0.02	0.03
Transportation (0/1)*length	0.18	0.21
Teaching (0/1)*length	−0.03	0.03
Humanities (0/1)*length	0.05	0.03
Musical (0/1)*length	0.04	0.07
Social (0/1)*length	**0.14**	0.03
Technical (0/1)*length	−0.03	0.04
Public security (0/1)*length	0.45	0.25
Unknown (0/1)*length	−0.03	0.02
Province	**0.27**	0.02
No. siblings	**0.07**	0.01
No. siblings missing	**0.24**	0.03

[a]The omitted type of education is health care.

TABLE 4

ESTIMATED TREATMENT EFFECTS

Dep. Variable: Log Hourly Wage Rate in 1997

Full Sample, 16,012 Mothers and 13,198 Non-Mothers

Parameter	Net Effect	Direct Effect Regression Adjusted Matching[a,b]
Motherhood:		
θ	$\mathbf{-0.069}^{c}$	$\mathbf{-0.015}^{d}$
	(0.005)	(0.005)
$\theta(public)$	$\mathbf{-0.073}^{e}$	$\mathbf{-0.036}^{f}$
	(0.007)	(0.007)
$\theta(private)$	$\mathbf{-0.067}^{g}$	0.008^{h}
	(0.008)	(0.006)

[a]Densities were estimated using a gaussian kernel and a bandwidth chosen via cross validation. The overlapping support region was determined using a 2 % trimming rule. Std. errors are based on 99 bootstraps with 100% resampling.

[b]Regression adjustment includes experience, educational categories and length, occupational categories, and interruptions from the labour market.

[c]Kernel based matching, 15,922 mothers and 12,121 non-mothers used, optimal bandwidth 0.0004.

[d]Kernel based matching, 15,925 mothers and 12,119 non-mothers used, optimal bandwidth 0.0004.

[e]Kernel based matching, 7,409 mothers and 7,322 non-mothers used, optimal bandwidth 0.008.

[f]Kernel based matching, 7,412 mothers and 7,325 non-mothers used, optimal bandwidth 0.008.

[g]Kernel based matching, 8,587 mothers and 4,748 non-mothers from public sector, 7,413 mothers and 7,315 non-mothers from private sector, optimal bandwidth 0.0004.

[h]Kernel based matching, 8,585 mothers and 4,743 non-mothers from public sector, 7,410 mothers and 7,319 non-mothers from private sector, optimal bandwidth 0.0004.

TABLE 5

Coefficient Estimates and Asy. Std. Err.

Motherhood and Sector Bivariate Probit

Dep. variables: 1 for Mothers, 0 for Non-Mothers, 1 for Private Sector, 0 for Public Sector

Full Sample, 16,012 Mothers, 13,198 Non-Mothers, 15,328 in Private, and 13,882 in Public Sector

-

	Motherhood		Sector Choice	
Variables[a]	Coeff.	Asy. Std. Error	Coef.	Asy. Std. Error
Intercept	**−12.51**	0.38	**−1.17**	0.20
Motherhood			**−0.44**	0.03
Age (years)	**0.76**	0.02		
Age squared	**−0.01**	0.00		
General (0/1)	0.34	0.23	**3.05**	0.22
Business (0/1)	**−1.84**	0.23	**2.35**	0.22
Industry (0/1)	1.29	1.08	**2.54**	0.90
Construction (0/1)	0.39	1.24	1.42	1.06
Graphical (0/1)	−1.54	1.31	**2.89**	1.31
Services (0/1)	0.19	0.57	0.55	0.51
Food and beverages (0/1)	−0.48	0.40	**3.88**	0.40
Agricultural (0/1)	−0.52	0.46	**3.01**	0.45
Transportation (0/1)	−3.24	2.11	0.83	2.43
Teaching (0/1)	0.27	0.51	**−2.13**	0.92
Humanities (0/1)	**−1.22**	0.54	7.56	0.56
Musical (0/1)	−0.92	1.22	**3.79**	1.15
Social (0/1)	**−2.73**	0.55	**5.34**	0.54
Technical (0/1)	−0.90	0.63	**3.17**	0.62
Public security (0/1)	−7.67	4.78	**0.92**	4.78
Unknown (0/1)	**−0.98**	0.23	**2.39**	0.23
Length of education	**−0.11**	0.01	**0.03**	0.01
General (0/1)*length	**−0.07**	0.02	**−0.17**	0.02
Business (0/1)*length	**0.13**	0.02	**−0.07**	0.02
Industry (0/1)*length	−0.14	0.09	−0.06	0.08
Construction (0/1)*length	−0.05	0.10	0.03	0.09
Graphical (0/1)*length	0.10	0.11	−0.09	0.11
Services (0/1)*length	−0.05	0.05	**0.10**	0.04
Food and beverages (0/1)*length	0.02	0.03	**−0.22**	0.03
Agricultural (0/1)*length	0.02	0.03	**−0.12**	0.03
Transportation (0/1)*length	0.17	0.17	0.09	0.20
Teaching (0/1)*length	−0.03	0.03	0.09	0.06
Humanities (0/1)*length	0.05	0.03	**−0.38**	0.03
Musical (0/1)*length	0.04	0.08	**−0.19**	0.07
Social (0/1)*length	**0.14**	0.03	**−0.25**	0.03
Technical (0/1)*length	−0.03	0.04	**−0.10**	0.04
Public security (0/1)*length	0.50	0.34		
Unknown (0/1)*length	−0.03	0.02	**−0.09**	0.02
Province	**0.27**	0.02	−0.02	0.02

TABLE 5, CONTINUED

Coefficient Estimates and Asy. Std. Err.

Motherhood and Sector Bivariate Probit

Dep. variables: 1 for Mothers, 0 for Non-Mothers, 1 for Private Sector, 0 for Public Sector

Full Sample, 16,012 Mothers, 13,198 Non-Mothers, 15,328 in Private, and 13,882 in Public Sector

	Motherhood		Sector Choice	
Variables[a]	Coef.	Asy. Std. Error	Coef.	Asy. Std. Error
No. siblings	**0.07**	0.01		
No. siblings missing	**0.25**	0.03		
Father empl. in public sector			**−0.07**	0.03
Mother empl. in public sector			**−0.11**	0.02

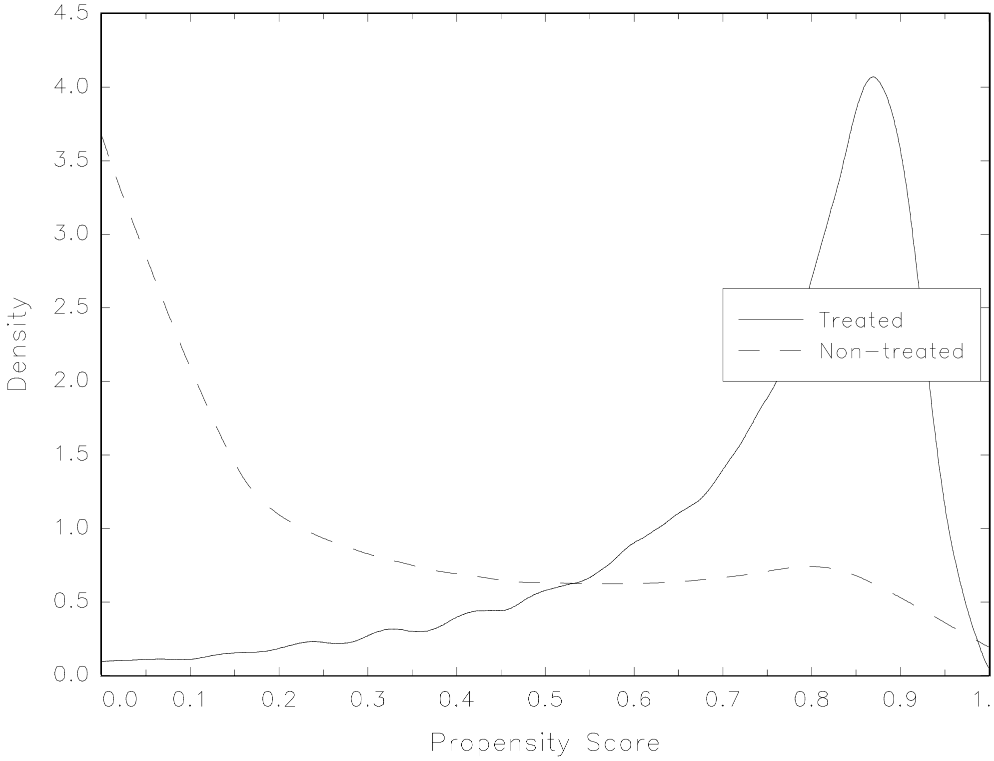
	Coef.	Asy. Std. Error
Correlation between error terms, ρ	**0.15**	0.02

[a]The omitted type of education is health care.

TABLE 6

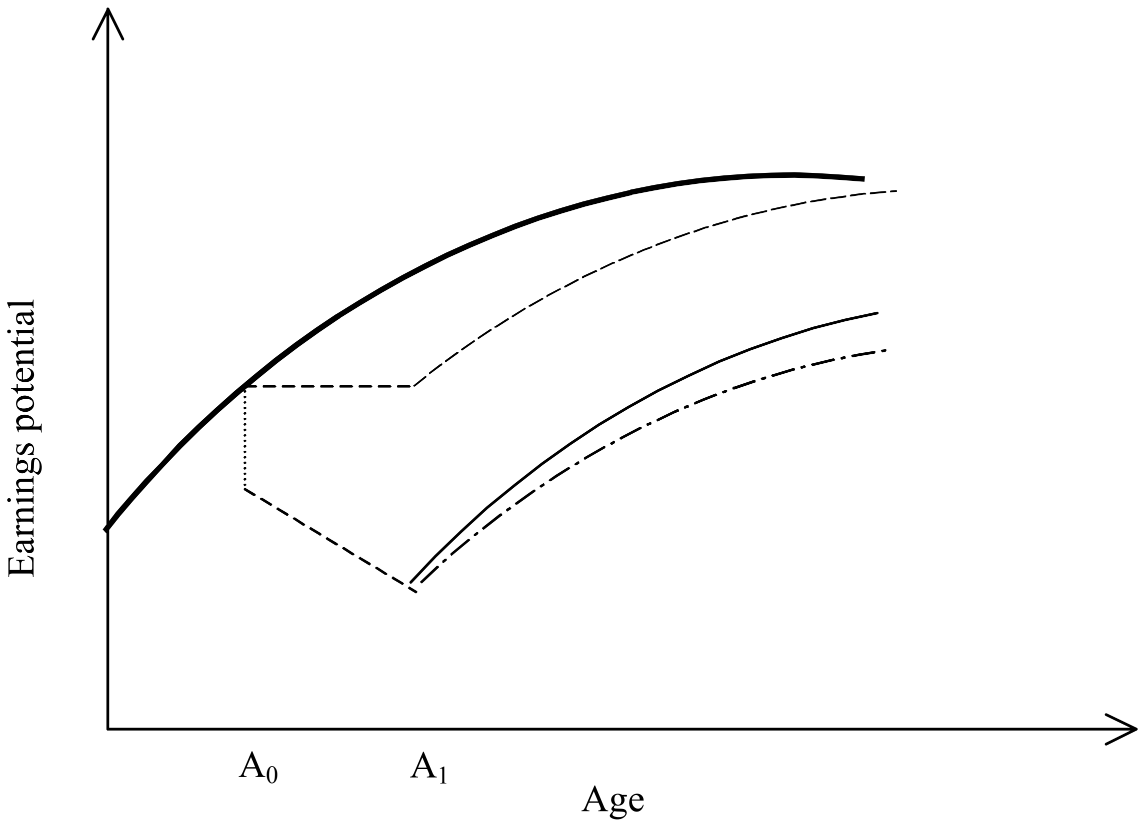
Comparison of Estimated

FIGURE 1



Nonparametric estimates of propensity densities (biweight kernel), probit $P(\mathbf{Z})$ model; bandwidth using Silverman (1986) rule of thumb. —— Mothers, - - - Non-mothers

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.