

This work is distributed as a Discussion Paper by the

STANFORD INSTITUTE FOR ECONOMIC POLICY RESEARCH

SIEPR Discussion Paper No. 03-23
**Identifying Direct and Indirect Effects.
Estimating the Costs of Motherhood
Using Matching Estimators**

By
Marianne Simonsen
University of Aarhus
And
Lars Skipper
University of Aarhus

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

The Stanford Institute for Economic Policy Research at Stanford University supports research bearing on economic and public policy issues. The SIEPR Discussion Paper Series reports on research and policy analysis conducted by researchers affiliated with the Institute. Working papers in this series reflect the views of the authors and not necessarily those of the Stanford Institute for Economic Policy Research or Stanford University.

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

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.

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 Stefan Bender, Martin Browning, Nabanita Datta Gupta, Michael Lechner, Shelly Lundberg, Helena S. Nielsen, Amalia R. Miller, Michael Rosholm, and Ed Vytlačil 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 Labor economics discussion group at Dept. of Economics, Stanford University and The Empirical Evaluation of Labour Market Programmes Conference 2005. 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 parameter of interest in the treatment evaluation literature is the average effect of treatment on the treated (Heckman, LaLonde & Smith (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 in the literature been designated 'the family gap' or 'the child penalty' (Budig & England (2001), Phipps, Burton & Lethbridge (2001), Korenman & Neumark (1992), Waldfogel (1998*a*), Waldfogel (1998*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 US, Canada, and Great Britain, the before-mentioned studies suggest a penalty of between 10 – 15% lower wages, whereas penalties in Scandinavia are typically smaller, ranging between 0 – 1.5% (Datta Gupta & Smith (2002), Nielsen, Simonsen & Verner (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 & 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 et al. (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 on earnings 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 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 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 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 e.g. 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 subsidised 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 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 *ceteris paribus*, 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 we impose an exclusion restriction to perform regression-adjusted matching recently sug-

gested by Heckman, Ichimura & Todd (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 organised 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 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 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 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 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 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 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 ten 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

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

Hence, the problem becomes that of finding the counterfactual $E[Y_0|C = 1]$ in (1), which is, of course, unobserved. 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, \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 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 & Rubin (1983)⁴

$$Y_0 \perp\!\!\!\perp C | \mathbf{X} \tag{A-1}$$

In other words, we assume that there exists some \mathbf{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 \mathbf{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 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.,

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

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 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 & 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,

$$Y_0 \perp\!\!\!\perp C | P. \tag{A-3}$$

This new conditioning variable, P , changes CIA into (A-3), 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 $E[Y | \mathbf{X}]$ to that of estimating $E[C | \mathbf{X}]$, often estimated by a logit or, as in this paper, a probit. See discussion in Black & 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 & Smith (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 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

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

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

i.e. assignment into C and S is strongly ignorable for $Y_0(s)$. Then $\theta(s)$ is identified and is (on the outset) free of indirect effects. \mathbf{X} is precisely the set of attributes ensuring potential wage outcome in the non-motherhood state be 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 & Rubin (1983).

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

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

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 specify $\Pr(C = 1, S = s | \mathbf{X})$ 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 $S = s$, $\Pr(C = 1 | S = s, \mathbf{X})$.

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 additive separability of the outcome equation in line with Heckman et al. (1998):

$$E \left[Y_0 | C = 1, \tilde{\mathbf{S}}, \mathbf{X}, \mathbf{R} \right] = \mathbf{X} \beta_X + \tilde{\mathbf{S}} \beta_{\tilde{\mathbf{S}}} + E \left[U_0 | C = 1, P(C = 1 | \mathbf{X}, \mathbf{R}) \right] \quad (\mathbf{A}\mathbf{S})$$

Assume furthermore,

$$U_0 \perp\!\!\!\perp C, \tilde{\mathbf{S}} | P(C = 1 | \mathbf{X}, \mathbf{R}), \quad (\mathbf{C}\text{-1})$$

$$\Pr(C = 1 | \mathbf{X}, \mathbf{R}) < 1, \quad (\mathbf{C}\text{-2})$$

where U_0 is the error term in the non-motherhood state. These assumptions mean that assignment into C and $\tilde{\mathbf{S}}$ is strongly ignorable for U_0 . When $E \left[U_0 | C = 1, \tilde{\mathbf{S}}, \mathbf{X} \right]$ is specified as a nonparametric function of (potentially continuous) \mathbf{X} , $(\mathbf{A}\mathbf{S})$ is termed a partial linear

model and identification requires the exclusion restriction, $E[Y_0|C = 1, P(C = 1|\mathbf{X}, \mathbf{R}), \mathbf{R}] = E[Y_0|C = 1, P(C = 1|\mathbf{X}, \mathbf{R})]$, 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:

$$U_0(s) \perp\!\!\!\perp C, S, \tilde{\mathbf{S}} | P(\mathbf{X}, \mathbf{R}), \quad (\mathbf{D-1})$$

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

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. 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, \mathbf{X} , includes 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.⁵ The results are presented in Table 2.

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. 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.

[Table 2 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 wrong⁶, 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. Moreover, we investigate the specification of our estimated propensity using the nonparametric test suggested by Shaikh, Simonsen, Yildiz & Vytlacil (2005). Formally, we test whether $E[C|P] = P$ and find that the test cannot reject our null at the 1% significance level.⁷

We choose the optimal bandwidth using cross validation as in Black & Smith (2004). As expected, 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. 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.⁸

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 firstly estimate the net effect of motherhood for the group of mothers, that is θ assuming **(A-1)** and **(A-2)**. 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)):

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

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 3 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 & Rubin (1985)) exceed 7%, see Appendix A for details.⁹

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:

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

When specifying the conditioning variables in $\tilde{\mathbf{S}}$, we follow human capital theory.¹⁰ 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 (cf. 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 subsidised 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 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 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 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 these effects, we include an indicator for being a mother and 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 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 \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 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 $\tilde{\mathbf{S}}$, see again Table 3. 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 (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 & 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.

[Table 3 around here]

[Figure 2 around here]

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, $\theta(\textit{public})$ and $\theta(\textit{private})$. We estimate both net effects, assuming **(B-1)** and **(B-2)** along with direct effects, assuming additive separability in addition to **(D-1)** and **(D-2)**

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 (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 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, \mathbf{X} includes the same set of variables as above augmented by the latter variables. Regarding estimation of the direct effects, $\tilde{\mathbf{S}}$ is unchanged except for the exclusion of the sector indicator.

The results from the bivariate probit can be seen in Table 2.¹¹ 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 3 show the results from the matching analysis taking endogeneity of sector into account and demonstrate clear differences between sectors, in particular wrt. 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 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 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 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 sector 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 by Note 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.

References

- 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.
- Datta Gupta, N. & Smith, N. (2002), ‘Children and Career Interruptions: The Family Gap in Denmark’, *Economica* **69**, 609–629.
- 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. A. (1999), The Economics and Econometrics of Active Labor Market Programs, in A. Ashenfelter & D. Card, eds, ‘Handbook of Labor Economics’, 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, in M. Lechner & F. Pfeiffer, eds, ‘Econometric Evaluation of Labour Market Policies’, Physica/Springer, Heidelberg, pp. 43–58.
- Lechner, M. (2002), ‘Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods’, *Journal of Royal Statistical Society* **165**(1), 59–82.
- Munch, J. R. (2003), ‘Do Municipality Taxes and Service Affect Settlement? (Har Kommunal Skat og Service Betydning for Flyttemønstrene)’, *Danish Journal of Economics (Nationaløkonomisk Tidsskrift)* pp. 53–67.

- Nielsen, H. S., Simonsen, M. & Verner, M. (2004), 'Does the Gap in Family-Friendly Policies Drive the Family Gap?', *Scandinavian Journal of Economics* **106**, 721–744.
- 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**, 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**, 656–666.
- Rosenbaum, P. R. & Rubin, D. B. (1983), 'The Central Role of the Propensity Score in Observational Studies for Causal Effects', *Biometrika* **70**, 41–55.
- 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**, 33–38.
- Shaikh, A. M., Simonsen, M., Yildiz, N. & Vytlacil, E. J. (2005), On the Identification of Misspecified Propensity Scores. Mimeo, Stanford University.
- Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall/CRC.
- Waldfogel, J. (1998a), '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. (1998b), 'Understanding the "Family Gap" in Pay for Women with Children', *Journal of Economic Perspectives* **12**, 137–156.

Notes

¹The 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.

²This 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.

³This 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.

⁴Notice 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.

⁵Note 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.

⁶We define number of wrong predictions as $\sum_{i=1}^n (C_i - \mathbf{1}_{(\hat{p}_i \geq .5)})^2$.

⁷Note that this is the sufficient condition for the balancing score property, see Theorem 2 in Rosenbaum & Rubin (1983).

⁸The 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, $p \in (0.45, 0.95)$.

⁹The 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%

¹⁰See Appendix A for a full list of the conditioning variables used in regression adjusted matching along with coefficients from estimating the partial linear regression.

¹¹To 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 probabilities over the 4 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 1.

TABLE 1^a
SELECTED MOMENTS, DESCRIPTIVE STATISTICS

Variables	All	Mothers	Non-mothers
Log wages	4.80 (0.28)	4.81 (0.26)	4.79 0.30
Age (years)	29.72 (5.80)	32.87 (4.20)	25.88 (5.13)
Experience (years)	7.28 (4.87)	9.39 (4.29)	4.70 (4.25)
Length of completed education (years)	12.22 (2.45)	12.20 (2.42)	12.26 (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
Health (0/1)	0.12	0.15	0.07
Pedagogic (0/1)	0.06	0.09	0.04
Humanistic (0/1)	0.03	0.03	0.04
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
Unknown (0/1)	0.07	0.07	0.08
Owner of real estate (0/1)	0.42	0.55	0.26
Married (0/1)	0.43	0.66	0.15
Private sector (0/1)	0.52	0.46	0.60
Top level occupation (0/1)	0.10	0.10	0.10
Higher level occupation (0/1)	0.21	0.23	0.18
Medium level occupation (0/1)	0.52	0.51	0.53
Lower level occupation (0/1)	0.17	0.16	0.19
Total duration of unemployment (weeks)	64.40 (94.60)	87.50 (106.90)	36.30 66.80
Years since last u-spell	4.63 (3.80)	5.22 (4.03)	3.66 (3.15)
Total duration of non-participation (weeks)	176.50 (119.30)	162.00 (119.50)	194.20 116.60
Years since last n-spell	4.54 (3.57)	3.74 (3.25)	5.56 (3.69)
Number of siblings when 15-17 years of age	1.16 (1.02)	1.00 (1.05)	1.35 (0.95)
Number of siblings missing (0/1)	0.19	0.23	0.14
Mother working in private sector	0.68	0.72	0.62
Father working in private sector	0.89	0.90	0.87
Sample size	29,027	15,958	13,069

^aStandard deviations shown in parentheses.

TABLE 2^aCOEFF. EST. AND ASY. STD. ERR., MOTHERHOOD PROBIT AND MOTHERHOOD AND SECTOR BIPOBIT^b

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

Variable	Motherhood		Motherhood		Sector	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Intercept	1.92	0.21	1.72	0.21	-1.03	0.29
Child					-0.37	0.17
20-22 years of age	-2.21	0.17	-2.20	0.17	-0.52	0.22
23-24 years of age	-1.57	0.14	-1.57	0.14	-0.49	0.21
25-26 years of age	-1.25	0.12	-1.25	0.12	-0.25	0.16
27-28 years of age	-0.64	0.10	-0.64	0.10	-0.02	0.13
31-32 years of age	0.10	0.10	0.10	0.10	0.08	0.11
33-34 years of age	0.34	0.10	0.25	0.10	0.02	0.11
35-36 years of age	0.59	0.10	0.68	0.10	0.12	0.11
37-38 years of age	0.46	0.10	0.55	0.10	0.05	0.11
39-40 years of age	0.59	0.12	0.67	0.12	0.15	0.12
Type of highest compl. educ.:						
General (0/1)	0.39	0.26	0.33	0.26	2.75	0.25
Business (0/1)	-1.89	0.25	-1.86	0.26	2.20	0.27
Industry (0/1)	1.10	1.10	1.10	1.12	2.56	0.95
Construction (0/1)	0.47	1.36	0.51	1.37	1.30	1.11
Graphical (0/1)	-1.89	1.34	-1.84	1.33	3.31	1.43
Services (0/1)	-0.13	0.61	-0.15	0.61	0.21	0.56
Food and beverages (0/1)	-0.82	0.43	-0.79	0.43	3.56	0.43
Agricultural (0/1)	-0.74	0.56	-0.69	0.56	2.57	0.54
Pedagogic (0/1)	-0.16	0.53	-0.10	0.54	-2.57	0.96
Humanistic (0/1)	-1.81	0.59	-1.75	0.59	7.98	0.62
Musical (0/1)	-1.09	1.41	-1.13	1.41	3.49	1.38
Social (0/1)	-2.90	0.60	-2.88	0.60	4.91	0.60
Technical (0/1)	-1.13	0.69	-1.09	0.69	2.91	0.69
Unknown (0/1)	-1.16	0.28	-1.39	0.28	2.13	0.28
Length of highest compl. educ.:	-0.11	0.01	-0.10	0.01	0.01	0.01
General (0/1)*length	-0.09	0.02	-0.08	0.02	-0.14	0.02
Business (0/1)*length	0.12	0.02	0.12	0.02	-0.05	0.02
Industry (0/1)*length	-0.13	0.09	-0.13	0.09	-0.05	0.08
Construction (0/1)*length	-0.06	0.11	-0.07	0.11	0.05	0.09
Graphical (0/1)*length	0.12	0.11	0.12	0.11	-0.07	0.12
Services (0/1)*length	-0.05	0.05	-0.04	0.05	0.12	0.05
Food and beverages (0/1)*length	0.03	0.03	0.03	0.03	-0.20	0.03
Agricultural (0/1)*length	0.03	0.04	0.03	0.04	-0.09	0.04
Pedagogic (0/1)*length	-0.02	0.03	-0.02	0.04	0.13	0.06
Humanistic (0/1)*length	0.07	0.04	0.07	0.04	-0.39	0.04
Musical (0/1)*length	0.05	0.08	0.05	0.08	-0.18	0.08
Social (0/1)*length	0.14	0.04	0.14	0.04	-0.22	0.03
Technical (0/1)*length	0.03	0.04	0.03	0.04	-0.07	0.04
Unknown (0/1)*length	0.04	0.02	0.05	0.02	-0.07	0.02
No. siblings			0.07	0.01		
No. siblings missing			0.31	0.03		
Mother employed in private sector					-0.11	0.02
Father employed in private sector					-0.07	0.03

^a 9 regional dummies and interaction terms between age and education dummies are included, reference age group 28-30, educational group health.

^b Estimated correlation between error terms, $\rho=0.10$ (0.10).

TABLE 3
 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}
θ	-0.065^c (0.004)	-0.015^d (0.005)
$\theta(\text{public})$	-0.076^e (0.007)	-0.032^f (0.006)
$\theta(\text{private})$	-0.064^g (0.008)	0.005 ^h (0.006)

^aDensities 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. Std. errors are based on 99 bootstraps with 100% resampling.

^bRegression adjustment includes experience, educational categories and length, occupational categories, regional dummies, and interruptions from the labour market.

^c15,665 mothers and 12,201 non-mothers used,
 $h_{rule-of-thumb} = 0.047, h_{opt} = 0.0002$.

^d15,775 mothers and 12,072 non-mothers used,
 $h_{rule-of-thumb} = 0.047, h_{opt} = 0.049$

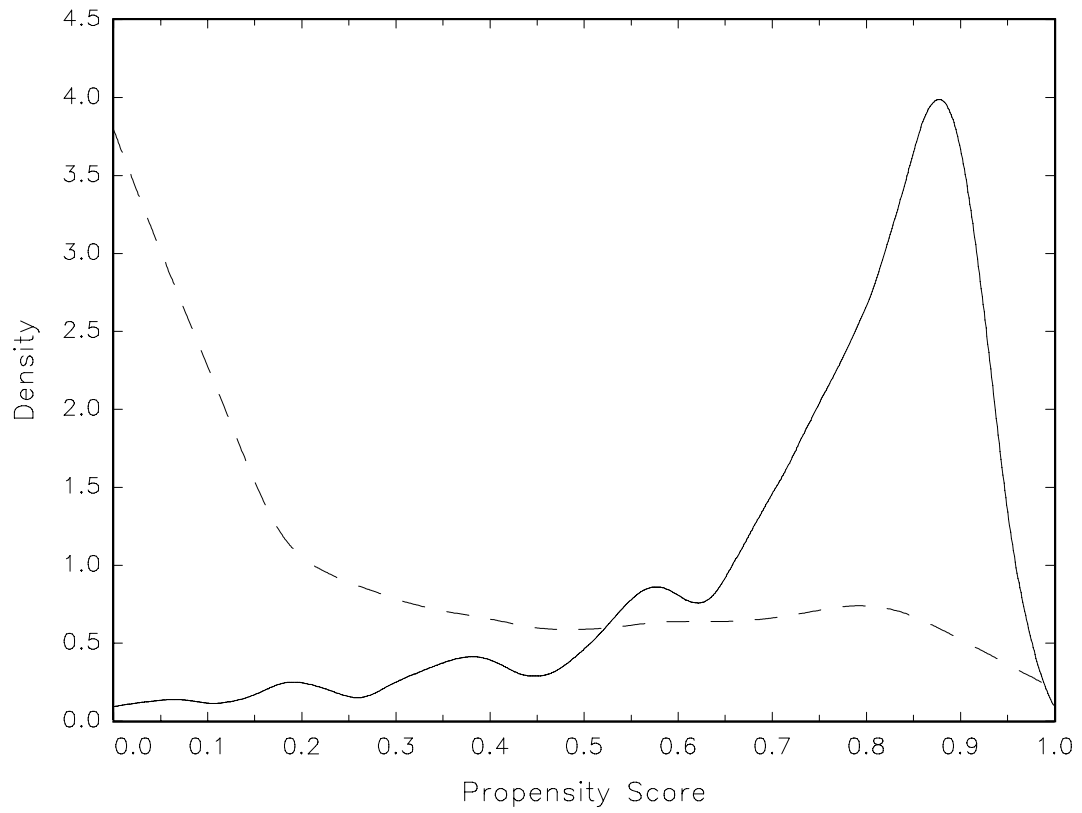
^e8,433 mothers and 4,826 non-mothers used,
 $h_{rule-of-thumb} = 0.057, h_{opt} = 0.005$

^f 8,452 mothers and 4,802 non-mothers used,
 $h_{rule-of-thumb} = 0.058, h_{opt} = 0.047$

^g 7,276 mothers and 7,332 non-mothers used,
 $h_{rule-of-thumb} = 0.052, h_{opt} = 0.010$

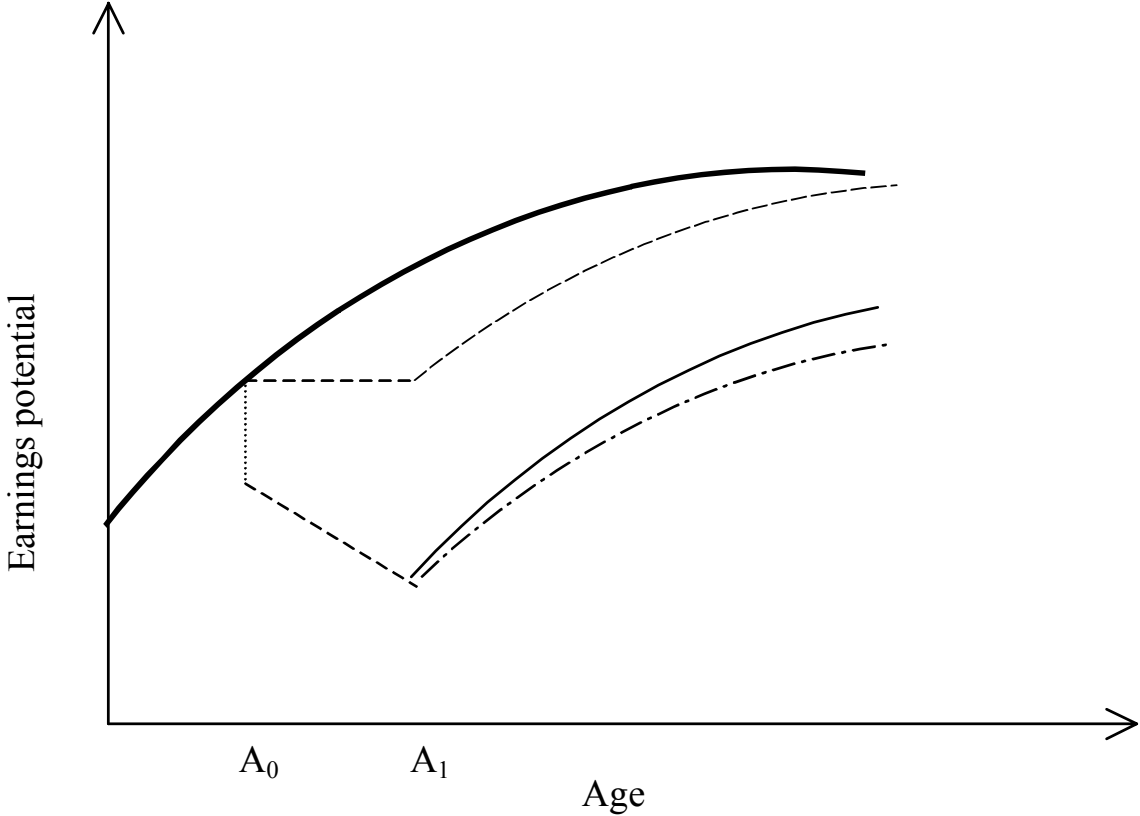
^h 7,264 mothers and 7,359 non-mothers used
 $h_{rule-of-thumb} = 0.053, h_{opt} = 0.045$

FIGURE 1



Non-parametric estimates of propensity densities (biweight kernel), probit $P(\mathbf{X})$ 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.

A Appendix

TABLE A1
COVARIATE IMBALANCE, NET EFFECT
Means and Standardized Differences in Percentage Points

Variable	Treated means	Matched means	%bias
20-22 years of age	0.010	0.009	0.468
23-24 years of age	0.023	0.023	-0.114
25-26 years of age	0.047	0.046	0.283
27-28 years of age	0.082	0.084	-0.706
31-32 years of age	0.153	0.159	-1.542
33-34 years of age	0.155	0.153	0.604
35-36 years of age	0.159	0.155	1.292
37-38 years of age	0.159	0.156	0.836
39-40 years of age	0.080	0.079	0.381
Unknown (0/1)	0.067	0.074	-2.542
General (0/1)	0.193	0.196	-0.755
Business (0/1)	0.328	0.338	-2.040
Industry (0/1)	0.009	0.009	0.252
Construction (0/1)	0.011	0.009	1.726
Graphical (0/1)	0.009	0.010	-0.937
Services (0/1)	0.022	0.019	1.575
Food and beverages (0/1)	0.043	0.044	-0.622
Agricultural (0/1)	0.009	0.010	-0.910
Pedagoghic (0/1)	0.087	0.084	0.758
Humanistic (0/1)	0.027	0.022	2.885
Musical (0/1)	0.003	0.002	1.592
Social (0/1)	0.027	0.026	1.055
Technical (0/1)	0.013	0.014	-0.406
Length of highest compl. educ.:	0.122	0.121	4.174
Unknown (0/1)*length	0.007	0.007	-1.662
General (0/1)*length	0.020	0.020	-0.573
Business (0/1)*length	0.037	0.039	-2.191
Industry (0/1)*length	0.001	0.001	0.227
Construction (0/1)*length	0.001	0.001	1.658
Graphical (0/1)*length	0.001	0.001	-0.937
Services (0/1)*length	0.002	0.002	1.565
Food and beverages (0/1)*length	0.005	0.005	-0.404
Agricultural (0/1)*length	0.001	0.001	-0.555
Pedagoghic (0/1)*length	0.013	0.012	0.794
Humanistic (0/1)*length	0.004	0.004	2.935
Musical (0/1)*length	0.001	0.000	1.519
Social (0/1)*length	0.005	0.004	1.070
Technical (0/1)*length	0.002	0.002	-0.536
Frederiksborg	0.123	0.124	-0.334
West Sealand	0.109	0.110	-0.300
Fuen	0.085	0.092	-2.430
Southern Jutland	0.097	0.090	2.619
Aarhus and Vejle	0.189	0.198	-2.291
Ringkoebing	0.100	0.093	2.161
Northern Jutland	0.095	0.083	4.280

TABLE A2
COVARIATE IMBALANCE, NET EFFECT, PUBLIC SECTOR
Means and Standardized Differences in Percentage Points

Variable	Treated means	Matched means	%bias
20-22 years of age	0.006	0.005	1.815
23-24 years of age	0.016	0.016	-0.683
25-26 years of age	0.041	0.036	2.880
27-28 years of age	0.073	0.072	0.311
31-32 years of age	0.144	0.160	-4.463
33-34 years of age	0.166	0.162	0.980
35-36 years of age	0.169	0.158	2.879
37-38 years of age	0.175	0.173	0.542
39-40 years of age	0.093	0.089	1.464
Unknown (0/1)	0.054	0.063	-3.938
General (0/1)	0.172	0.150	5.791
Business (0/1)	0.228	0.181	11.752
Industry (0/1)	0.005	0.002	4.963
Construction (0/1)	0.006	0.005	1.893
Graphical (0/1)	0.005	0.004	2.148
Services (0/1)	0.014	0.007	6.294
Food and beverages (0/1)	0.047	0.043	1.778
Agricultural (0/1)	0.007	0.009	-2.693
Pedagoghic (0/1)	0.158	0.203	-11.766
Humanistic (0/1)	0.019	0.019	0.571
Musical (0/1)	0.004	0.005	-2.234
Social (0/1)	0.026	0.020	4.129
Technical (0/1)	0.007	0.007	0.753
Length of highest compl. educ.:	0.128	0.130	-6.911
Unknown (0/1)*length	0.006	0.006	-3.128
General (0/1)*length	0.018	0.015	6.418
Business (0/1)*length	0.026	0.021	11.650
Industry (0/1)*length	0.001	0.000	5.423
Construction (0/1)*length	0.001	0.001	1.614
Graphical (0/1)*length	0.001	0.000	2.020
Services (0/1)*length	0.002	0.001	6.633
Food and beverages (0/1)*length	0.006	0.005	1.498
Agricultural (0/1)*length	0.001	0.001	-2.026
Pedagoghic (0/1)*length	0.023	0.030	-11.437
Humanistic (0/1)*length	0.003	0.003	0.739
Musical (0/1)*length	0.001	0.001	-2.515
Social (0/1)*length	0.004	0.003	4.046
Technical (0/1)*length	0.001	0.001	0.673
Frederiksborg	0.113	0.116	-0.876
West Sealand	0.122	0.116	1.680
Fuen	0.088	0.100	-4.294
Southern Jutland	0.089	0.097	-2.588
Aarhus and Vejle	0.190	0.196	-1.565
Ringkoebing	0.096	0.080	5.367
Northern Jutland	0.097	0.092	1.438

TABLE A3
COVARIATE IMBALANCE, NET EFFECT, PRIVATE SECTOR
Means and Standardized Differences in Percentage Points

Variable	Treated means	Matched means	%bias
20-22 years of age	0.014	0.011	2.954
23-24 years of age	0.031	0.024	4.425
25-26 years of age	0.054	0.054	-0.040
27-28 years of age	0.093	0.090	0.895
31-32 years of age	0.164	0.157	1.822
33-34 years of age	0.143	0.157	-4.040
35-36 years of age	0.148	0.156	-2.173
37-38 years of age	0.139	0.143	-1.040
39-40 years of age	0.065	0.065	-0.025
Unknown (0/1)	0.083	0.094	-3.823
General (0/1)	0.218	0.221	-0.713
Business (0/1)	0.445	0.409	7.368
Industry (0/1)	0.015	0.016	-1.129
Construction (0/1)	0.015	0.014	1.237
Graphical (0/1)	0.014	0.017	-2.156
Services (0/1)	0.031	0.035	-2.095
Food and beverages (0/1)	0.039	0.048	-4.676
Agricultural (0/1)	0.012	0.012	-0.164
Pedagoghic (0/1)	0.004	0.009	-6.957
Humanistic (0/1)	0.035	0.028	3.814
Musical (0/1)	0.003	0.002	1.286
Social (0/1)	0.028	0.028	0.081
Technical (0/1)	0.020	0.018	1.643
Length of highest compl. educ.:	0.115	0.116	-1.480
Unknown (0/1)*length	0.008	0.009	-3.297
General (0/1)*length	0.022	0.022	-1.530
Business (0/1)*length	0.051	0.046	7.362
Industry (0/1)*length	0.002	0.002	-1.188
Construction (0/1)*length	0.002	0.002	1.532
Graphical (0/1)*length	0.002	0.002	-2.071
Services (0/1)*length	0.003	0.004	-1.929
Food and beverages (0/1)*length	0.005	0.006	-4.698
Agricultural (0/1)*length	0.001	0.002	-0.479
Pedagoghic (0/1)*length	0.001	0.001	-7.069
Humanistic (0/1)*length	0.005	0.004	3.636
Musical (0/1)*length	0.000	0.000	1.554
Social (0/1)*length	0.005	0.005	-0.053
Technical (0/1)*length	0.003	0.003	1.659
Frederiksborg	0.134	0.143	-2.636
West Sealand	0.095	0.104	-2.978
Fuen	0.082	0.079	1.201
Southern Jutland	0.107	0.089	6.051
Aarhus and Vejle	0.188	0.182	1.498
Ringkoebing	0.104	0.099	1.774
Northern Jutland	0.093	0.085	2.597

TABLE A4
COVARIATE IMBALANCE, DIRECT EFFECT
Means and Standardized Differences in Percentage Points

Variable	Treated means	Matched means	%bias
20-22 years of age	0.010	0.012	-1.791
23-24 years of age	0.023	0.021	1.536
25-26 years of age	0.047	0.045	0.854
27-28 years of age	0.083	0.085	-0.816
31-32 years of age	0.155	0.174	-5.319
33-34 years of age	0.157	0.165	-2.395
35-36 years of age	0.157	0.149	2.180
37-38 years of age	0.157	0.136	6.053
39-40 years of age	0.079	0.066	4.700
Unknown (0/1)	0.068	0.084	-5.940
General (0/1)	0.192	0.177	3.875
Business (0/1)	0.331	0.334	-0.561
Industry (0/1)	0.009	0.010	-0.444
Construction (0/1)	0.011	0.010	0.190
Graphical (0/1)	0.009	0.011	-1.536
Services (0/1)	0.022	0.024	-1.270
Food and beverages (0/1)	0.043	0.048	-2.422
Agricultural (0/1)	0.009	0.011	-2.269
Pedagoghic (0/1)	0.087	0.088	-0.153
Humanistic (0/1)	0.026	0.026	0.362
Musical (0/1)	0.003	0.004	-0.783
Social (0/1)	0.027	0.025	1.510
Technical (0/1)	0.013	0.014	-0.617
Length of highest compl. educ.:	0.122	0.122	1.178
Unknown (0/1)*length	0.007	0.008	-5.264
General (0/1)*length	0.019	0.018	3.429
Business (0/1)*length	0.038	0.038	-0.261
Industry (0/1)*length	0.001	0.001	-0.380
Construction (0/1)*length	0.001	0.001	0.239
Graphical (0/1)*length	0.001	0.001	-1.505
Services (0/1)*length	0.002	0.003	-1.183
Food and beverages (0/1)*length	0.005	0.006	-2.359
Agricultural (0/1)*length	0.001	0.001	-2.055
Pedagoghic (0/1)*length	0.013	0.013	-0.127
Humanistic (0/1)*length	0.004	0.004	0.487
Musical (0/1)*length	0.001	0.001	-0.844
Social (0/1)*length	0.005	0.004	1.468
Technical (0/1)*length	0.002	0.002	-0.601
Frederiksborg	0.123	0.123	-0.099
West Sealand	0.109	0.108	0.225
Fuen	0.085	0.091	-1.866
Southern Jutland	0.095	0.087	2.865
Aarhus and Vejle	0.191	0.195	-1.078
Ringkoebing	0.099	0.089	3.379
Northern Jutland	0.095	0.084	3.572
No. siblings	0.996	1.012	-1.466
No. siblings missing	0.229	0.229	0.005

TABLE A5
COVARIATE IMBALANCE, DIRECT EFFECT, PUBLIC SECTOR
Means and Standardized Differences in Percentage Points

Variable	Treated means	Matched means	%bias
20-22 years of age	0.006	0.008	-1.789
23-24 years of age	0.016	0.016	-0.447
25-26 years of age	0.042	0.036	3.042
27-28 years of age	0.074	0.072	0.575
31-32 years of age	0.146	0.173	-7.405
33-34 years of age	0.168	0.171	-0.762
35-36 years of age	0.170	0.158	3.398
37-38 years of age	0.171	0.155	4.498
39-40 years of age	0.088	0.080	2.907
Unknown (0/1)	0.055	0.066	-4.714
General (0/1)	0.170	0.143	7.354
Business (0/1)	0.231	0.193	9.392
Industry (0/1)	0.005	0.002	4.965
Construction (0/1)	0.006	0.005	1.784
Graphical (0/1)	0.006	0.004	2.034
Services (0/1)	0.014	0.008	5.378
Food and beverages (0/1)	0.047	0.046	0.478
Agricultural (0/1)	0.007	0.010	-3.249
Pedagoghic (0/1)	0.160	0.205	-11.549
Humanistic (0/1)	0.020	0.019	0.663
Musical (0/1)	0.004	0.005	-2.065
Social (0/1)	0.027	0.020	4.364
Technical (0/1)	0.008	0.007	0.771
Length of highest compl. educ.:	0.128	0.130	-8.646
Unknown (0/1)*length	0.006	0.007	-3.862
General (0/1)*length	0.018	0.015	7.468
Business (0/1)*length	0.027	0.022	9.344
Industry (0/1)*length	0.001	0.000	5.421
Construction (0/1)*length	0.001	0.001	1.504
Graphical (0/1)*length	0.001	0.001	1.894
Services (0/1)*length	0.002	0.001	5.805
Food and beverages (0/1)*length	0.006	0.006	0.117
Agricultural (0/1)*length	0.001	0.001	-2.499
Pedagoghic (0/1)*length	0.024	0.030	-11.389
Humanistic (0/1)*length	0.003	0.003	0.785
Musical (0/1)*length	0.001	0.001	-2.396
Social (0/1)*length	0.005	0.003	4.298
Technical (0/1)*length	0.001	0.001	0.685
Frederiksborg	0.114	0.111	1.016
West Sealand	0.119	0.116	1.008
Fuen	0.089	0.104	-5.170
Southern Jutland	0.088	0.092	-1.250
Aarhus and Vejle	0.191	0.203	-2.925
Ringkoebing	0.095	0.079	5.764
Northern Jutland	0.095	0.089	2.321
No. siblings	0.957	0.979	-2.062
No. siblings missing	0.224	0.233	-2.042

TABLE A6
COVARIATE IMBALANCE, DIRECT EFFECT, PRIVATE SECTOR
Means and Standardized Differences in Percentage Points

Variable	Treated means	Matched means	%bias
20-22 years of age	0.014	0.014	-0.002
23-24 years of age	0.032	0.023	5.155
25-26 years of age	0.054	0.053	0.421
27-28 years of age	0.094	0.092	0.834
31-32 years of age	0.167	0.169	-0.550
33-34 years of age	0.145	0.163	-4.989
35-36 years of age	0.149	0.149	0.049
37-38 years of age	0.134	0.128	1.706
39-40 years of age	0.059	0.060	-0.272
Unknown (0/1)	0.084	0.095	-3.898
General (0/1)	0.210	0.208	0.367
Business (0/1)	0.452	0.425	5.448
Industry (0/1)	0.015	0.016	-0.775
Construction (0/1)	0.016	0.014	1.022
Graphical (0/1)	0.014	0.016	-1.587
Services (0/1)	0.031	0.035	-1.956
Food and beverages (0/1)	0.040	0.049	-4.440
Agricultural (0/1)	0.012	0.012	-0.169
Pedagoghic (0/1)	0.003	0.009	-6.879
Humanistic (0/1)	0.035	0.029	3.701
Musical (0/1)	0.003	0.002	0.920
Social (0/1)	0.029	0.028	0.277
Technical (0/1)	0.020	0.018	1.560
Length of highest compl. educ.:	0.115	0.116	-1.604
Unknown (0/1)*length	0.009	0.010	-3.518
General (0/1)*length	0.021	0.021	-0.633
Business (0/1)*length	0.051	0.048	5.652
Industry (0/1)*length	0.002	0.002	-0.874
Construction (0/1)*length	0.002	0.002	1.300
Graphical (0/1)*length	0.002	0.002	-1.488
Services (0/1)*length	0.004	0.004	-1.947
Food and beverages (0/1)*length	0.005	0.006	-4.463
Agricultural (0/1)*length	0.001	0.002	-0.519
Pedagoghic (0/1)*length	0.000	0.001	-7.010
Humanistic (0/1)*length	0.005	0.005	3.538
Musical (0/1)*length	0.000	0.000	1.215
Social (0/1)*length	0.005	0.005	0.124
Technical (0/1)*length	0.003	0.003	1.594
Frederiksborg	0.132	0.137	-1.311
West Sealand	0.095	0.103	-2.789
Fuen	0.083	0.080	1.020
Southern Jutland	0.106	0.086	6.671
Aarhus and Vejle	0.190	0.188	0.528
Ringkoebing	0.103	0.098	1.854
Northern Jutland	0.092	0.082	3.516
No. siblings	1.061	1.029	3.055
No. siblings missing	0.223	0.227	-0.899

TABLE A7^aCOEFF. EST. AND ASY. STD. ERR., PARTIAL LINEAR MODEL^b

Variable	Coeff.	Std. Err.
Experience (years)	0.007	0.046
Experience sq.	0.000	0.002
Private sector	0.042	0.066
Type of highest compl. educ.:		
General (0/1)	0.358	0.869
Business (0/1)	0.440	1.059
Industry (0/1)	0.452	2.304
Construction (0/1)	0.635	3.461
Graphical (0/1)	0.896	5.456
Services (0/1)	0.227	2.509
Food and beverages (0/1)	0.250	1.556
Agricultural (0/1)	0.507	2.070
Pedagogic (0/1)	0.380	2.302
Humanistic (0/1)	0.240	2.567
Musical (0/1)	1.539	7.175
Social (0/1)	0.685	1.800
Technical (0/1)	0.421	1.899
Unknown (0/1)	0.387	1.063
Length of highest compl. educ.:	0.052	0.060
General (0/1)*length	-0.023	0.065
Business (0/1)*length	-0.035	0.079
Industry (0/1)*length	-0.036	0.183
Construction (0/1)*length	-0.052	0.286
Graphical (0/1)*length	-0.065	0.454
Services (0/1)*length	-0.018	0.213
Food and beverages (0/1)*length	-0.020	0.120
Agricultural (0/1)*length	-0.037	0.156
Pedagogic (0/1)*length	-0.030	0.154
Humanistic (0/1)*length	-0.019	0.159
Musical (0/1)*length	-0.097	0.414
Social (0/1)*length	-0.042	0.108
Technical (0/1)*length	-0.026	0.124
Unknown (0/1)*length	-0.031	0.080
Owner of real estate	0.018	0.057
Married	-0.007	0.080
Higher level occupation (0/1)	0.097	0.162
Medium level occupation (0/1)	-0.051	0.138
Lower level occupation (0/1)	-0.065	0.085

^a 9 regional dummies included, reference age group 28-30, educational group health.

TABLE A7^a CONTD.COEFF. EST. AND ASY. STD. ERR., PARTIAL LINEAR MODEL^b

U81	-0.014	0.435
U82	0.032	0.434
U83	0.045	0.354
U84	0.005	0.258
U85	-0.005	0.300
U86	0.010	0.246
U87	0.017	0.240
U88	0.008	0.190
U89	0.000	0.155
U90	-0.008	0.158
U91	-0.010	0.167
U92	-0.017	0.102
U93	-0.007	0.116
U94	-0.009	0.097
U95	-0.022	0.089
U96	-0.033	0.083
NON81	-0.013	0.274
NON82	0.002	0.238
NON83	-0.006	0.188
NON84	0.000	0.219
NON85	-0.003	0.198
NON86	-0.015	0.185
NON87	0.007	0.204
NON88	-0.010	0.203
NON89	-0.029	0.196
NON90	-0.025	0.198
NON91	-0.025	0.185
NON92	-0.052	0.173
NON93	-0.056	0.179
NON94	-0.083	0.192
NON95	-0.113	0.179
NON96	-0.094	0.232
Total duration of non-participation (year)	0.005	0.032
Total duration of unemployment (year)	-0.012	0.033

^a 9 regional dummies included, reference age group 28-30, educational group health.