Industriens Utredningsinstitut
THE INDUSTRIAL INSTITUTE FOR ECONOMIC AND SOCIAL RESEARCH

A list of Working Papers on the last pages

No. 414, 1994

PITFALLS IN THE MEASUREMENT OF THE RETURN TO EDUCATION:
An Assessment Using Swedish Data

by

Eugenia Kazamaki Ottersten, Erik Mellander,
Eva M. Meyerson and Jörgen Nilson

April 1994

Postadress
Box 5501
114 85 Stockholm

Gatuadress
Industrihuset
Storgatan 19

Telefon
08-783 80 00
Telefax
08-661 79 69

Bankgiro
446-9995

Postgiro
19 15 92-5

10.92 10.000

Pitfalls in the Measurement of the Return to Education: An Assessment Using Swedish Data

Eugenia Kazamaki Ottersten, Erik Mellander,
Eva M. Meyerson, Jörgen Nilson*

The Industrial Institute for Economic and Social Research (IUI)
Box 5501, S–114 85 Stockholm, Sweden
Fax: + 46-8-6617969, E-mail: iui.@iui.se

April 11, 1994

Abstract

We consider three pitfalls associated with the standard econometric approaches to measuring the return to education, namely; selection bias, omitted variable bias, and random error in the measure of schooling. We start from the traditional wage equation and formulate a model which explains education and the selection into "wage-earners" and "others", as well as earnings. The empirical analysis uses two Swedish data sets for 1984 and 1986. Measurement error in schooling is found to be the most important pitfall. In addition, we conclude that omitted variables in the form of social relations significantly influence earnings, both directly and indirectly, through education and selection.

*We wish to thank seminar participants at IUI in particular Thomas Andersson and Gunnar Eliasson, and seminar participants at SOFI for useful comments.

1 Introduction

In this paper we deal with three pitfalls associated with the standard econometric approaches to measuring the return to education. The starting point is the traditional wage equation originally formulated by Mincer (1974). However, in investigating the relation between earnings and schooling we ultimately treat education as well as the wage as endogenous variables. The three pitfalls we address are *i)* selection bias, *ii)* omitted variable bias, and *iii)* random error in the measure of schooling. As far as we know, this study is the first to account for these problems simultaneously. To assess their empirical importance we make use of two Swedish cross-section data sets, referring to 1984 and 1986 respectively.

In modeling the selection of individuals into "wage earners" and "others" we extend the common practice of considering only individuals in the labor force to include also some individuals outside the labor force, in particular, students.

It has long been recognized that the human capital variables in the wage equation have to be complemented by individual-specific information about innate abilities and family background or, more generally, hereditary and environmental factors. In dealing with the second pitfall we take a comprehensive approach to the problem of accounting for environmental factors by considering variables associated with what the sociological literature has termed social capital. A person's social capital is embodied in his/her relations to other individuals. Thus, in contrast to human capital which is an *intra*-personal concept, social capital is an *inter*-personal concept. By referring to social relationships outside the family, as well as within the family, social capital is a much more general concept than family background.

To take the random error in the measure of schooling into account we explicitly model education as a stochastic variable, the variation in which is explained within the model.

The model that we use can be partitioned into two blocks, each consisting of recursive two-equation systems. In the first block the individual's education is ex-

plained in terms of a set of social capital indicators and a number of standard control variables. The selection into "wage-earners" and "others" is then determined as a function of education, other human capital variables, social capital indicators, and control variables.

In the second block only the "wage-earners" are considered. Their education is determined by adjusting the education equation of the first block for sample selection. The wage is then estimated by means of a Mincer - type wage equation including social capital indicators as well as human capital variables. The wage equation thus allows for direct effects of social capital on the wage, as well as indirect effects through the education and selection variables.

To assess the relative importance of the three pitfalls we estimate the wage equation sequentially. We begin by controlling for selection bias. Next, we include social capital variables to allow for omitted variables. Finally, we report the results corresponding to the extended wage equation which also takes measurement error in education into consideration.

Section 2 assesses previous research in this area and describes the concept of social capital. In Section 3 we specify the econometric approach taken in this paper. Section 4 and 5 turn to the description of data and results. In Section 6, we bring together some of the major conclusions from this study and identify where future research is warranted.

2 Background

Building on a notion of equalizing differentials by Adam Smith in "Book I of the Wealth of Nations", the theory of human capital is predominant in the determination of wages. There have been important contributions in this area by Jacob Mincer (1957, 1958, 1962), Theodore Schultz (1960, 1961), and Gary Becker (1962, 1964).

In its most simplistic form, the traditional human capital equation can be specified as follows (cf., e.g., Berndt (1991, ch. 5))

$$lny_i = lny_o + \beta_1 s_i + \beta_2 x_i + \beta_3 {x_i}^2 + u_i. \quad (1)$$

In this model the wage y is determined by schooling, s, and experience x. That these two types of human capital may well be interdependent is

The human capital equation has been extensively tested. A typical finding in the literature is that the returns to schooling are substantially higher in developing countries than in developed countries. In the 1970's the average return to schooling was about 15 percent in developing countries and 9 percent in high-income countries (Psacharopoulos, 1985).[1] A recent Swedish study shows that the income premium to schooling in Sweden has been declining over the 1968 to 1981 period from 7 percent to about 3 percent. Thereafter there has been a small increase in the return to education to a level about 5-6 percent (Björklund & Kjellström, 1993).[2]

Recent studies have also incorporated the effects of family background on earnings and returns to schooling and discussed the importance of unobservable worker attributes. For example, when examining the differences between developing and developed countries it has been argued that a number of biases and unobservable worker attributes may affect the results. A model of assortive mating developed by Lam & Schoeni (1993) supports the notion that parental characteristics represent unobservable worker attributes.

2.1 Shortcomings of the Human Capital Approach

In the literature there has been a discussion about omitted variable bias in the traditional way to estimate the returns to schooling. If, e.g., family background and innate abilities affect the wage and if these variables are positively correlated with schooling, omitting family background and ability measures and estimating eq (1)

[1] Most of the human capital studies deal with US data (e.g. Freeman 1975, Lydon 1990, Pencavel, 1991, Psacharopoulos 1980, Willis 1986).

[2] For an overview of some empirical evidence from Sweden see also Edin, Fredriksson & Holmlund (1994).

by OLS will bias the estimated returns to schooling, because the s_i will be correlated with the disturbance u_i.

Berndt (op. cit.) refers to several studies in which this problem has been investigated. Griliches and Mason (1972) considered the possible bias arising from the omission of ability measures. Their conclusion was that ability measures constructed by means of intelligence test scores were virtually uncorrelated with schooling. In contrast, Taubman (1976a,b) concludes that ability and schooling may be highly correlated. However, Griliches (1977, 1979) argues that Taubman's result is due to random errors in the measurement of schooling.

Concerning the effects of omitting family background variables the results are more mixed. Some studies find that family background matters a lot, see, for instance, Leibowitz (1974), while others conclude that its effects on earnings are limited.[3]

Let us think a bit about the above problems. It is a stylized fact that some individuals acquire more education and human capital than others. In general, individuals with higher abilities will accumulate more human capital. It is thus reasonable to expect abilities to be correlated with years of schooling. However, ability is probably correlated with family background as well. For example, genetic intelligence is likely to be captured by variables which are associated with the parent's characteristics.

2.2 Social Capital

Social capital is a broad concept. In this paper three different dimensions of social capital will be considered.

[3]In assessing the effects of family background it is desirable to control for ability (and vice versa). In an analysis of educational attainment based on Swedish data Härnqvist (1993) controlled for ability by selecting a population that was relatively homogenous with respect to talent. He then found strong effects from family background on educational performance. Härnqvist's results indicate that family background at least influences the wage indirectly, through education.

The first dimension relates to the number of relationships that the individual has accumulated at a certain point in time. In general, more relationships are better than fewer. However, at some point it may be efficient to limit the number of relations because the costs of terminating relationships may be high and difficult to predict. In addition, there may also be indirect costs involved because the termination of some relationships can affect other, existing or potential, relations.

The second dimension concerns the type of the relationship. A number of different ways to characterize relations can be found in the sociological literature. Here, we will make a distinction between *strong* and *weak* ties. Strong ties are used to denote relations to, e.g., family members and close friends while weak ties may refer to acquaintances or distant relatives. Strong and weak ties fulfill different purposes. For instance, strong ties have been found to be important for children's educational performance, cf. Loury (1977, 1987) and Coleman (1988, 1990).[4] Weak ties, on the other hand, are by some sociologists claimed to be more instrumental to occupational attainment (see Boxman, De Graaf & Flap (1991), Granovetter (1973), Lin, Ensel & Vaughn (1981), Lin (1982, 1988)).[5] The composition of ties depends on the goals that the social capital is expected to achieve. For example, a mix of strong and weak ties appropriate for the creation of new relationships can be poorly suited to maintaining a specific set of values and attitudes.

The third dimension of a person's social capital has to do with the concept of time. Social relationships change over time. For example, a weak tie in a given

[4]Coleman (1988) refers to an interesting piece of evidence for this effect. When controlling for the parents education surprisingly many Asian immigrant families in the U.S. have high performing children. A closer investigation revealed that families with poorly educated parents purchased two copies of the children's required textbooks. One copy was bought for the mother to study, in order to help her child do well at school. Thus, comparatively poor general human capital education was compensated for by means of extensive social capital.

[5]There is no general consensus on this matter, however. For instance, if there is room for nepotism strong ties will play an important role as well. For a recent empirical investigation about the influence of strong ties on wages, see the study by Lam & Schoeni (op.cit.) where the relations to parents-in-law are considered. (See also Meyerson (1993).)

period can become a strong tie or be disolved in another period. From an analytical point of view this has the important implication that social capital cannot be treated like a "fixed effect". Accordingly, the problem of taking social capital into account cannot be eliminated simply by differencing the data.

Finally, to make the above dimensions of social capital operational we relate them to the sociological concept of closure (Coleman (1990)). The closure of a given social structure - a family, a community, or an organization - is a measure of the extent to which this structure can be separated from other social structures, physically and mentally. For example, for a family the degree of closure may be dependent upon how much time the parents spent with their children when the children grew up. Social structures with a high degree of closure are in general based on strong ties between the members in the group. Likewise, structures where the degree of closure is low are often held together by weak ties. Closure indicators can thus be used to measure the strength or tightness in the relationships associated with different social structures.

3 Econometric Approach

The econometric approach adopted below starts from a traditional Mincer equation allowing for the effects of education, experience, tenure, and a number of control variables on wages. In assessing the traditional model we consider three major problems

1) *selection bias*, this bias arises because in estimating earnings equations we are restricted to individuals receiving positive wages;

2) *omitted variable bias*, if some variables which both affect earnings and are correlated with the education variable are left out of the wage equation this will induce a bias in the estimated return to education;

3) *measurement error*, if there are random errors in the measure of education this will also bias the coefficient for education.

A pedagogical discussion of the relation between the second and the third problem can be found in Lam and Schoeni (op. cit.). They show that the biases caused by omitted variables and measurement error can work in different directions. If an omitted variable is positively related to (true) schooling then the omitted variable bias will be positive while the bias arising from the measurement error in education will be negative. In order to assess the relative importance of the two biases they perform a number of simple simulations which indicate that the measurement error bias strongly dominates the omitted variables bias. Lam and Schoeni do not attempt, however, to explicitly account for measurement error directly in the estimation, which we do in this paper.

To consider the three pitfalls, we use an extended version of Leibowitz (1974) model. Leibowitz postulates a recursive relationship between education and the wage; education is determined by individual and family background characteristics whereas income is a function of education, other human capital variables, and individual and family background characteristics.[6]

This modeling framework makes it possible to address the second and the third pitfall. First, by recognizing that individual and family background characteristics may affect the wage directly, as well as indirectly through education, Leibowitz accounts for the problem of possible omitted variable bias. Second, in contrast to equation (1), education is explained within the model. This means that random error in an endogenous variable, which is relatively unproblematic, can be substituted for random error in an exogenous variable, which is a serious problem. While, apparently, Leibowitz was not concerned about the potential measurement error in education, her formulation clearly points to a way of handling this problem.

[6]Leibowitz considers "home investment" rather than family background, but, in operationalizing home investment she essentially uses family background data.

Leibowitz' framework does not take the problem of selectivity bias into account, however. We remedy this shortcoming by including a selection equation in our analysis. Formally, the model can be partitioned in two blocks.

The first block consists of the following two equations, determining education and the selection of individuals into "wage-earners" and "others"

$$s_i = \boldsymbol{\beta}' \mathbf{q}_i + \epsilon_{1i} = E(s_i) + \epsilon_{1i} \tag{2}$$

$$\mathbf{I}_i = \alpha_1 E(s_i) + \boldsymbol{\alpha}_2' \mathbf{z}_i + \epsilon_{2i} \tag{3}$$

where $\mathbf{I}_i$ is a latent variable determining if the individual is gainfully employed.

In equation (2) individual i's schooling is explained by means of social capital indicators and some control variables, all of which are contained in vector $\mathbf{q}_i$. E denotes the expected value operator and ϵ_{1i} is a normally distributed random error.

The selection into "wage-earners" and "others" is determined by eqation (3). If the latent variable $\mathbf{I}_i > 0$ then i's wage is positive, otherwise zero. The latent variable $\mathbf{I}_i$ is not observed. However, its dichotomous realization $\mathbf{I}_i^*$ is observed. If $\mathbf{I}_i > 0$ then $\mathbf{I}_i^* = 1$, otherwise $\mathbf{I}_i^* = 0$. Note that due to measurement error in schooling it is expected schooling rather than actual schooling that enters the selection equation. In addition, the selection is influenced by social capital and control variables, summarized by the vector $\mathbf{z}_i$. The disturbance term ϵ_{2i} is assumed to be normally distributed and possibly correlated with ϵ_{1i}. The vectors $\mathbf{q}_i$ and $\mathbf{z}_i$ have some, but not all elements in common.

In the second block of the model, which also consists of two equations, the education and the wage are determined for "wage-earners";

$$(s_i | \mathbf{I}_i^* = 1) = \boldsymbol{\beta}' \mathbf{q}_i + \mu_i + \epsilon_{3i} = E(s_i | \mathbf{I}_i^* = 1) + \epsilon_{3i} \tag{4}$$

$$lnwage_i = \gamma_1 E(s_i | \mathbf{I}_i^* = 1) + \boldsymbol{\gamma}_2' \mathbf{v}_i + \eta_i + \epsilon_{4i} \tag{5}$$

where μ_i is the selection effect with respect to schooling, i.e. $E(s_i | \mathbf{I}_i^* = 1) - E(s_i)$, $\mathbf{v}_i$ is a vector of observed individual and job characteristics that influence i's wage, and η_i is the selection effect with respect to the wage, i.e. the expected value of

$lnwage_i - \gamma_1 E(s_i|\mathbf{I}_i^* = 1) - \boldsymbol{\gamma}_2' \mathbf{v}_i$. The disturbances ϵ_{2i} and ϵ_{4i} are assumed to be joined normally distributed with zero means.

Consistent estimates of the model's parameters are obtained by means of the following four-step procedure. First, equation (2) is estimated by OLS, yielding an estimate of $E(s_i)$. Second, using the estimate of $E(s_i)$ in the selection equation (3) we estimate the parameters α_1 and $\boldsymbol{\alpha}_2$ by probit maximum likelihood.

Next, consider equations (4) and (5). Taking the sample selection into account following Heckman's (1978) procedure, the equations can be written as:

$$(s_i|\mathbf{I}_i^* = 1) = E(s_i) + \sigma_s \lambda_i + \epsilon_{1i} \tag{6}$$

and

$$lnwage_i = \gamma_i E(s_i|\mathbf{I}_i^* = 1) + \boldsymbol{\gamma}_2' \mathbf{v}_i + \sigma_w \lambda_i + \epsilon_{2i} \tag{7}$$

where $\lambda_i \equiv \phi(r_i)/(1 - \Phi(r_i))$ is Heckman's λ, the coefficients of which are $\sigma_s \equiv cov(\mu_i + \epsilon_{3i}, \epsilon_{1i})$ and $\sigma_w \equiv cov(\eta_i + \epsilon_{4i}, \epsilon_{1i})$, respectively, and $r_i = \alpha_1 E(s_i) + \boldsymbol{\alpha}_2' \mathbf{z}_i$. The disturbances, finally, are given by $e_{1i} = \mu_i + \epsilon_{3i} - \sigma_s \lambda_i$ and $e_{2i} = \eta_i + \epsilon_{4i} - \sigma_w \lambda_i$.

Due to the recursive structure of the system (6) - (7) it can be estimated by OLS. In the third stage of the estimation we use the estimates of $E(s_i)$ and λ_i obtained in the first and the second step to form the difference

$$(s_i|\mathbf{I}_i^* = 1) - E(s_i) = \sigma_s \lambda_i + e_{1i}. \tag{8}$$

Estimation of equation (8) by OLS yields a consistent estimate of σ_s. Note that (8) implies that

$$E(s_i|\mathbf{I}_i^* = 1) = E(s_i) + \sigma_s \lambda_i. \tag{9}$$

In the fourth and final step we substitute our estimate of the right hand side of (9) for $E(s_i|\mathbf{I}_i^* = 1)$ in (7) and estimate the resulting equation by OLS.

To be able to compare our approach with the traditional human capital specification we start by estimating the following variant of the traditional wage equation:

$$lnwage_i = \boldsymbol{\gamma}_0' \boldsymbol{\nu}_{0i} + \gamma_1 educ_i + + \gamma_2 noexp_i + \gamma_3 exp_i + \gamma_4 exp_i^2 + \gamma_5 ten_i + \epsilon_i \tag{10}$$

where *lnwage* is the log hourly wage, $\boldsymbol{\nu}_{0i}$ is a vector of control variables, *educ* is education, *noexp* stands for no experience, *exp* for experience, and *ten* for tenure. Education and experience are measured by number of years; *noexp* is a dummy equal to one if the respondent's total working experience is less than three months. Tenure is measured by the number of years at the current employer. The control variables include, e.g., dummy variables for gender, marital status, number of children[7],the region where the respondent lives, and variables measuring part time employment. This equation will be our reference case.

In order to assess the relative importance of the three pitfalls we consider them sequentially. We start by correcting for selection bias. Next, we take omitted (social capital) variables into account. Finally, we report the results for the extended model, as presented above. Hence, after the estimation of the selection equation (3) where Heckman's λ is obtained, we estimate the following equation:

$$lnwage_i = \boldsymbol{\gamma}_0^{'}\boldsymbol{\nu}_{0i} + \gamma_1 educ_i + +\gamma_2 noexp_i + \gamma_3 exp_i + \gamma_4 exp_i^2 + \gamma_5 ten_i + \sigma_w \lambda_i + \epsilon_i \quad (11)$$

Compared to traditional models, our approach to dealing with the selection problem is based on a more broadly defined population. Traditionally, the estimation of the probit equation is based on a sample comprising only people in the labor force, characterized as "wage earners" and "others", or as "employed" and "unemployed". However, the relevant population is determined by the variables that are endogenous in the analysis. When education is treated as an endogenous variable it is natural to enlarge the "others" group and include students as well. We do so in our analysis. In addition, we include individuals working in the home (mostly women) but likely to become students or enter the labor force in the future.[8]

Secondly, we test for possible direct effects of social capital variables on the wage

$$lnwage_i = \boldsymbol{\gamma}_0^{'}\boldsymbol{\nu}_{0i} + \gamma_1 educ_i + +\gamma_2 noexp_i + \gamma_3 exp_i + \gamma_4 exp_i^2 + \gamma_5 ten_i + \sigma_w \lambda_i +$$

[7]In a paper by Lazear & Michael (1980) it has been pointed out that the relationship between real income and family size is important. In this paper family size is controlled for by the number of children.

[8]See further Section 4.

$$\gamma_6' \nu_{6i} + \epsilon_i \tag{12}$$

where ν_{6i} is a vector of social capital variables.[9]

Finally, the extended wage equation can be written

$$\begin{aligned} lnwage_i &= \gamma_0' \nu_{0i} + \beta \widehat{peduc}_i + \gamma_1 noexp_i + \gamma_2 exp_i + \gamma_3 exp_i^2 + \gamma_4 ten_i + \\ &\quad \lambda z + \gamma_5' \nu_{5i} + e_{2i} \end{aligned} \tag{13}$$

where $\widehat{peduc}_i$ denotes the predicted education of the ith "wage-earner". As described above the education of "wage-earners" is estimated in two steps. In the first step the education of both "wage-earners" and "others" is explained by social capital indicators and control variables. This step gives the unconditional predicted education, $\widehat{educ}_i$. In the second step we first estimate the selection effect by regressing the difference $educ_i^* - \widehat{educ}_i^*$ on Heckman's λ, where the superindex denotes that only "wage-earners" are considered. We then construct $\widehat{peduc}_i$ by adding the estimated selection effect to the unconditional estimate.

4 Data

The data used in this paper are from the Swedish Household Market and Non Market Activities (HUS) study, see Klevmarken and Olovson (1993). We have used 1984 and 1986 data in this study.[10]

The 1984 data set is based on a sample of over 2100 Swedish households, identified with the help of a stratified sample of Swedish speaking individuals aged 18-74. In each household with couples either married or living together, both partners were interviewed. The response rate was approximately 74 percent, resulting in 1564 households and 2795 individuals. Information was gathered on, i.a., conditions during childhood, education, labor market experience, wages, and family composition.

[9] For a discussion of the social capital variables, see Section 4.

[10] The HUS data base contains data for 1988 and 1991 as well and by the end of 1994 there will be data for 1993, too. We plan to make use of this additional information in later work.

The data obtained were then augmented with register information. In addition, in 1984 a time-use survey was conducted, based on telephone interviews.

The 1986 data was obtained partly by updating the information from the personal interviews conducted in 1984 and partly by means of a new supplementary sample, directed to over 800 households. The individuals included in the supplement were asked approximately the same questions as in the 1984 personal interview. No time-use survey was conducted in 1986. For the 1986 panel the response rate was about 80 percent and for the supplement 60 percent, yielding total numbers of households and individuals of 1574 and 2724, respectively.

In this study we have excluded all individuals that fulfill any of the following criteria, *i)* age $\geq$ 65, *ii)* unable to work or retired, *iii)* self-employed, farmers, fishermen, and hunters. Remaining individuals in the samples are both members of the labor force and persons outside the labor force. In the HUS survey, the former category comprises those that are either gainfully employed or unemployed/laid off individuals that have looked for work within two months of the date of the interview. The latter category consists of students and individuals working in the home, mainly house-wives. Individuals working in the home but unlikely to enter the labor force or become students in the future are excluded from our analysis.[11]

For 1984 we constructed two samples, one without the information from the time-use survey (1984 A) and one which takes it into account (1984 B). The former sample has the advantage that it is comparable with the 1986 data whereas the 1984 B sample is more informative. The latter adds four variables to the 1984 A sample; three social capital variables and one control variable. The latter is an attitude variable measuring the respondent's assessment of his/her work. In terms of observations, the 1984 B sample is somewhat smaller than the 1984 A sample,

[11]Those with earlier experience have been asked if they want to work in the future. We have retained only individuals who have responded in the affirmative to this question. Concerning individuals without labor market experience, we have retained those who were students the previous year.

due to a number of non-responses in connection with the time-use survey.

According to Table 1 the 1984 share of wage-earners is 89.9 percent in the A sample and 89.3 percent in the B sample whereas the 1986 share is somewhat higher, 91.6 percent. In Table 2 the "others" category has been broken down into unemployed, students, and persons working in the home. The distribution over the three subcategories differs between 1984 and 1986. In 1984 the distribution is perfectly even across subcategories; each represents one third. In the 1986 sample the share of unemployed and laid off is down to about 26 percent while the students share is almost 40 percent. The share of "house-wives" remains roughly unchanged. By means of both Table 1 and Table 2 it can be seen that the shares of the unemployed are 3.4, 3.6, and 2.2 percent in the 1984 A, 1984 B, and 1986 sample respectively. These shares are close to the 1984 and 1986 Swedish unemployment figures which were 3.1 and 2.2 percent, respectively.

Finally, a few words about the social capital variables that we use. As discussed in Section 2.2 the structure of an individual's social capital varies over his/her life cycle. To capture this variation we have tried to distinguish the social capital associated with childhood and youth from the social capital generated in the individual's adult life. In addition, we want to make a distinction between closures (relational structures) characterized mainly by strong and weak ties, respectively.

For the respondent's childhood and youth we include indicators of family closure (strong ties) and community closure (weak ties). Among the family closure variables we consider the extent to which the mother stayed at home when the children were young, measures of the occupation of the respondent's father, and the father's and mother's education. Community closure will, e.g., relate to the place where the respondent was brought up and to information that indicates the respondent's citizenship status.

The social capital generated in the individual's adult life is partitioned correspondingly. Adult family closure is measured by marital status, number of children etc. The indicators of community closure are measures of the time used by the

individual in activities that differ with respect to the amount of external contacts (relations outside the family) that they involve. On the one extreme there are activities that are characterized by no external contacts at all, like reading books. On the other extreme there are activities directly focused on creating and maintaining external contacts, like entertaining quests at home and visiting friends.[12]

[12]Unfortunately, data on adult life community closure are available only for the 1984 cross section.

	"Wage-earners"		"Others"		Total	
	#	%	#	%	#	%
1984 A	1523	89.9	172	11.1	1695	100
1984 B	1406	89.3	168	11.7	1574	100
1986	1473	91.6	135	8.4	1608	100

Table 1: The data sets partitioned into "wage-earners" and "others"

	Unemployed		Students		"House-wives"		Total	
	#	%	#	%	#	%	#	%
1984 A	57	33.1	58	33.8	57	33.1	172	100
1984 B	56	33.3	56	33.3	56	33.3	168	100
1986	36	26.7	53	39.3	46	34.0	135	100

Table 2: Break-down of the individuals in the "others" category

5 Results

As mentioned in the previous section, we have two data sets for 1984 and one for 1986. This yields three sets of estimated equations. Each of these sets contain seven regression equations. First, there is the reference equation (10). Next, there are the education and probit equations (2) and (3), and the wage equation (11) corrected for sample selection. These are followed by the wage equation which also takes omitted variables into account (12). Finally, we have the education equation for wage earners (6) and the extended wage equation (13) which also acounts for measurement error in education. We will comment on these regressions in turn and then conclude the section by summarizing the effects of the different specifications on the estimated return to education.

5.1 The Reference Wage Equation

The results obtained in the estimation of the reference equation (10) are given in Table 4. All the human capital variables, with the exception of the dummy variable for no experience, are significant. The return to education is about 3.7-3.8 percent in 1984 and slightly below 4.2 percent in 1986. The difference between the 1984 and 1986 estimates is significant at the 5 percent level. Experience (at least equal to one year) has a positive effect for most individuals.[13]

The estimated experience profiles are quite different in the three samples. The maximum experience effects occur at 26, 23.5, and 42 years in the 1984 A, 1984 B, and the 1986 samples respectively. This corresponds to marginal returns equal to 16.5, 14.3, and 30.0 percent. In contrast to experience, the marginal effect of tenure is about the same in 1984 and 1986, in both years it is somewhere between 0.4 and 0.5 percent.

Concerning the control variables, it is worth noting that the attitude variable *work* (defined in Appendix B), which is available only in the 1984 B sample, is significant and increases the wage by 0.78 percent on the margin. The variables common to the 1984 and 1986 samples show large differences across years. For instance, in 1984 the marginal age effect is about 1 percent for men and 0.5 percent for women. In 1986 the age effect is insignificantly different from zero for men, and significantly negative, -0.3 percent, for women. Likewise, the marginal gender effects seem to change between 1984 and 1986. Considering significant variables only, the gender effect can in 1984 be taken to be equal to (-0.005 $\times$ *age*) for women in the private sector and (-0.005 $\times$ *age* + 0.1) for women in the public sector. At the mean age, 40 years in 1984, this yields effects equal to -20 and -10 percent, respectively. In 1986, the corresponding expressions are (-0.14 - 0.003 $\times$ *age*) and (-0.14 - 0.003 $\times$ *age* + 0.13), respectively. Evaluated at the mean age, 41 years, the marginal effects are -26 and -13 percent.

[13]The only exception are individuals with experience equal to or greater than 47 years in the 1984 B sample.

With respect to public sector employment it is worth noting the differential effects for males and females. The males working in the public sector 1984 had wages that were about 7 percent lower than males in the private sector. In 1986 this difference was over 9 percent. For the women the difference was of the opposite sign; on average the difference between public and private sector wages were + 10 percent in 1984 and + 13 percent in 1986.

Considerable care has been taken to account for part-time and overtime work. As expected, the hourly wage is somewhat higher for those working part-time, compared to individuals working full-time.[14] Regarding working overtime one finding is that those who work overtime without receiving overtime compensation have implicit wage premia equal to 10-11 percent in 1984 and about 5.5 percent in 1986, compared to individuals not working overtime.

In summary, the reference equation fits the data quite well; it explains almost 40 percent of the variation in (the log of) hourly wages in all of the three samples. The parameter estimates differ surprisingly much between 1984 and 1986, however. In Section 5.4 it will be seen that these differences are indeed artifical in the sense that they almost disappear when the random variation in the measure of education is taken into account.

5.2 Correcting for Sample Selection Bias

To correct for sample selection bias we first esimate the education equation (2) and the probit equation (3). In the estimation of the education equation, only a subset of

[14]In 1984 the difference is about 6 percent. In 1986 the premium is about 5 percent for individuals working 20-34 hours/week (*part-time2*), but over 30 percent for those working up to 19 hours/week (*part-time1*). The latter estimate is clearly unreasonable and seems to have been caused by a coding error. A number of part-time working individuals stated their wage income on a monthly basis and in so doing reported the monthly pay corresponding to full-time work, rather than their actual part-time employment. When we introduced a dummy for the 1986 part-timers reporting their wage income in terms of monthly pay the coefficient for *part-time1* became insignificant. However, as the other coefficients were not affected we have not pursued this matter further.

the social capital variables have been used, namely, those related to the respondent's childhood and youth. Indicators of strong (family) ties are: the time that the mother spent home with the children, the parent's education, and the father's occupation. There are two dummy variables for weak (cummunity) ties *i)* whether the respondent grew up in a city environment (*city-childhood*), *ii)* whether Swedish was spoken in his/her home (*Swedish-home*).[15]

The results are presented in Table 5. Concerning the social capital variables indicating strong ties, all but one are significant. The only exception is a dummy variable that is supposed to capture interaction effects between the mother's and the father's education.[16] Furthermore, of the two variables indicating community closure (weak ties) only the *city-childhood* dummy turned out to be significant. Thus, earlier findings on the importance of strong ties for children's educational performance (cf. Section 2.2) are confirmed, both in absolute and relative terms, i.e. compared to weak ties.

Our results correspond quite well with other Swedish studies to the extent that the parent's education and occupation explain a large part of the variance in education. However, our finding that it does not matter for the respondent's education whether Swedish was spoken in his/her home contradicts earlier results which indicate that children of immigrants in Sweden do relatively better in school than children with Swedish parents; cf. SOU 1993:85.

The estimates are also quite interesting with respect to the control variables. The variable *school-reform* makes it possible to carry out a simple consistency check. This variable is equal to one for individuals that were at most 14 years old at the time of the enactment of the 9-year mandatory school in 1971. Since these individuals had to spend nine years in school under all circumstances, the sum of the intercept and

[15]The latter variable is interpreted as a community closure indicator rather than as a family closure indicator because the native language affects the extent to which a family is integrated in the community.

[16]In the 1984B sample the interaction variable *age* x *father-bluecollar* is insignificant, too.

the coefficient for *school-reform* should be close to 9. This condition is well satisfied in 1984: the sum equals 9.1 and 9.4 in the 1984 A and 1984 B samples, respectively. In 1986 the sum is somewhat higher: 10.6 years. A result that holds in all of the three regressions is that the gender effect is quite small. On average, the education of females is less than half a year shorter than the education of males.

In summary, the education equation is quite satisfactory both from a theoretical and a statistical standpoint. The relevant sociological theory is supported and the equation fits the data well - R^2 is close to 0.4.

We next turn to the probit equation. Our choice of human capital variables in this equation is quite natural. In addition to expected schooling we have included the experience variables *noexp*, *exp*, exp^2, but excluded *tenure* because the latter variable is applicable only to wage earners. In choosing social capital indicators we have focused on the social capital generated in the individual's adult life, operationalized as described in Secion 4. Concerning the social capital indicators associated with the respondent's childhood and youth we have tried to pick those that are relatively close - in terms of time distance - to the respondent's labor market entry. For example, we have included measures of the father's occupation but not data about the father's education.[17]

The estimated probit equations are reported in Table 5. Somewhat surprisingly, expected education has a significant effect on the probability of gainful employment only in the 1986 sample. In constrast, the experience variables, i.e. *noexp, exp, and* exp^2, are all significant in both the 1984 and 1986 samples, (with the exception of the 1986 coefficient for *noexp*). The parameter estimates are also very similar in magnitude across the two years.

In 1984, some of the social capital variables reflecting strong ties generated in the respondent's childhood and youth are significant, e.g., the education of the respondent's mother (*mother-educ1*) and the occupation of the respondent's father

[17]We have included measures of the mother's education, however, because there is no information about the mother's occupation.

(*father- bluecol* and *age* × *father-bluecol*). Marital status (*married*), which is an indicator of adult family closure, is significant, too. The time-use variables (*external-contacts, introvert-activities*, and *extrovert-activities*), which are supposed to capture the individual's adult life weak tie relationships, all turn out to be insignificant. Hence, the sociological hypothesis that weak ties are instrumental with respect to occupational attainment is not supported by the probit analysis.

In the 1986 sample, two of the variables measuring the education of the respondent's mother are significant.

Concerning the explanatory power of the probit equation, it can be seen that the 1986 regression is the most successful one in terms of the total share of correct predictions (91.9 percent). In this context the share of correctly predicted in the "others" category is probably a more useful goodness-of-fit criterion. Judged by this latter criterion, 1984 B regression is the best one with a share of correct predictions equal to 22.0 percent while the 1986 regression is the least successful with only 9.6 percent correct predictions.

Finally, we add the selection variable obtained from the probit equation to the wage equation. This produces only negligible effects on the estimated returns to education. See Table 7. The only substantial change concerns the experience variables. Having been strongly significant in the reference equation the experience variables become completely insignificant. Since *exp* and exp^2 are present in the probit equation this can be interpreted as showing that when sample selection is taken into account these variables only affect the wage indirectly, through the selection mechanism. On the other hand lack of experience, as measured by *noexp*, has both direct and indirect effects on the wage, at least in 1984. Presumably, the positive coefficient of *noexp* merely captures the fact that many individuals without experience, e.g. students, work odd hours and thus are extra compensated to a larger extent than individuals with regular jobs.

Heckman's lambda has the expected, negative, sign in both 1984 and 1986. It is strongly significant in the former year and at the 10 percent level in 1986. The

estimated effects may seem large; e.g., in 1984 a marginal increase in λ increases the wage about 35 percent. However, the selection effect is more appropriately evaluated in terms of the corresponding elasticity, which is obtained by multiplying the coefficient by λ. At the 1984 mean values of λ, this elasticity is about 0.055, i.e. a 1 percent increase in λ increases the wage by 0.055 percent, which is quite low.

5.3 Correcting for Sample Selection and Omitted Variables

Next, we also allow for direct effects of the social capital variables on the wage, in addition to the correction for sample selection. In the first place we have added social capital indicators not included in the probit equation, such as information about the time the mother stayed home with the children, the father's education, and an indicator of childhood community closure, a dummy equal to one if the respondent grew up in a city environment. Adult life family closure is considered indirectly, through some of the control variables indicating marital status and family size. As mentioned above, the time-use variables for adult life community closure are available only in the 1984 B sample.

The results are given in Table 8. Only in 1986 does the introduction of the social capital variables result in a non-negligible change in the estimated return to education. The relative effect on the wage decreases from 4.1 percent to 3.8 percent, but the decrease is not significantly different from zero. Some of the social capital variables are significant in the 1984 samples, however. In the 1984 A sample *mother-home* and *age* $\times$ *mother-home* significantly influence the wage - positively for respondents over the mean age and negatively for those below the mean age. One of the measures of adult life community closure, *introvert-activities*, is significant with the expected, negative, sign in the 1984 B sample. The hypothesis that weak ties are instrumental for job attainment thus receives some support in that time spent on activities *not* involving the creation and maintenance of weak ties lowers

the wage. For 1986 none of the social capital indicators are significant.

The parameters of the control variables are almost unaffected by the introduction of the social capital variables. The same holds for the adjusted R^2's.

5.4 Correcting for Sample Selection, Omitted Variables, and Measurement Error

To correct for measurement error in the schooling variable we first have to construct estimates for the education of "wage-earners". As outlined in Section 3, we do this by correcting the original education equation for sample selection. The estimated selection effects are given in the lower part of Table 5. It can be seen that the parameters are positive and significant in each of the three samples.[18] On average, the education of "wage-earners" is longer than the education of those that are not gainfully employed, which is as expected. Comparing the R^2's at the very bottom of Table 5 with the corresponding measures in the upper part of this table we see that the goodness-of-fit is very close to that of the education equations estimated for the whole sample.

The most noticeable effect of the substitution of *peduc* for *educ* is that for the first time the estimated return to education in 1984 changes substantially, see Table 9. Compared to the regressions correcting for selection bias and omitted variables, the estimated returns increase from 3.69 percent to 4.44 percent and from 3.64 percent to 4.19 percent in the 1984 A and 1984 B samples, respectively. Compare with Table 8. In contrast, the estimated return for 1986 decreases slightly from 3.88 percent to 3.80 percent.

The estimated effects of social capital change, too. In the 1984 A sample, the number of significant social capital is augmented by measures of the father's and the mother's education, in addition to the time that the (respondent's) mother spent

[18] In the 1984 A and 1986 samples the levels of significance are well below 5 percent while in the 1986 B sample the significance level is around 7 percent.

at home with the children. It should be noted that the negative coefficients for the parent's education cannot be taken to imply that having highly educated parents reduce one's wage. Since the parent's education also affect the wage indirectly, through Heckman's λ and the instrument for the schooling variable, all we can say is that the direct effects of these variables on the wage are negative. As can be seen in Table 9, the indirect effects working through education are strongly positive.

In the 1984 B sample one of the measures of the mother's education becomes significant, in addition to the time-use variable *introvert-activities*. In 1986 no direct effects of social capital on the wage can be discerned, however.

Concerning the control variables, it can first be noted that in contrast to the other regressions Heckman's λ is significant in all three samples. Moreover, the coefficients are larger than before and more similar in magnitude between 1984 and 1986.

The finding that the estimates in 1984 and 1986 are more similar than in earlier regressions also applies to the other control variables. For instance, the marginal age effects in the comparable 1984 A and 1986 samples are slightly above 1 percent for men and about 0.5 percent for women in both years. This contrasts sharply with, e.g., the reference wage equation where the estimated male/female effects were 0 percent/-0.3 percent; cf. Section 5.1. For private sector employees the gender effect is also about the same; considering significant variables only, it can be computed as (- 0.0061 $\times$ *age*) and (- 0.0057 $\times$ *age*) for the 1984 A and 1986 samples, respectively. Evaluated at the mean ages, 40 and 41 years, respectively, this yields female wages that are about 24 percent and 23 percent lower than the male wages in the private sector, ceteris paribus.

Compared to the earlier regressions, there are two interesting changes concerning the public/private sector wage differential and the implicit wage premium of those who work uncompensated overtime. Whereas in the previous regressions, males in the public sector were found to have wages 7 - 9 percent below their private sector counterparts the sectoral difference is here not significant in 1984 and significant but

small, less than 5 percent, in 1986. Secondly the estimated implicit overtime premia are consistently larger when the measurement error is taken into account: 16 - 17 percent in 1984 and 14 percent in 1986, compared to 10 - 11 percent and 5 percent, respectivly, in the other wage equations.

Finally, it should be noted that the share of the total variance in the (log of) wages that is explained by the wage equation decreases quite substantially. This indicates that in trying to eliminate the random error in education we have also taken away some systematic variation which influences the wage. Presumably this systematic variation is related to innate abilities. In Section 2.1 we conjectured that genetic intelligence is likely to be captured by information about the parent's characteristics. While our results here do not contradict this conjecture we may draw the tentative conclusion that parental characteristics seem to be insufficient indicators of innate abilities.[19]

5.5 Summarizing the Results

Since the main purpose of the reported regressions is to assess how the estimated return to education changes when the wage equation is corrected for sample selection bias, omitted variables, and measurement error in the schooling variable, it is appropriate to summarize the results in this respect. An attempt in this direction is found in Table 3.[20]

In Table 3 it appears that measurement error in education is by far the most important problem. Whereas the correction for sample selection bias barely changes the return to education and the inclusion of omitted variables only matters in the 1986 sample, accounting for the measurement error results in appreciable changes in the estimated returns, at least in the 1984 samples. Some qualifications are

[19]Of course, genetically transmitted properties may be captured by information about the parents only in the case of biological parents. In our samples, about 90 percent of the respondents grew up with their biological mother and father.

[20]This Table has been put together by means of Tables 4, 7, 8, and 9.

	Reference equation	correction for		
		i)	i) + ii)	i)+ ii) + iii)
1984 A	0.0376	0.0369	0.0369	0.0444
1984 B	0.0374	0.0367	0.0364	0.0414
1986	0.0416	0.0411	0.0388	0.0380

Table 3: Estimates of γ_1 under different specifications, i) correcting for selection bias, ii) correcting for selection bias and omitted variable bias, and iii) correcting for selection bias, omitted variable bias, and measurement error.

appropriate, however.

First, sample selection bias is probably more important than it appears from Table 3, because the table only shows the direct effect of correcting for sample selection. However, Heckman's λ proved to be an important explanatory variable in the education equation as well as in the wage equation. Accordingly, there is an indirect effect of the correction for sample selection which is not considered in the table. Second, the entries in the table are point estimates. By means of the standard errors given in Table 9 it is easy to establish that the estimates in the last column are not significantly different from the estimates in the first column.

6 Conclusions

Our analysis shows that the pitfalls in the measurement of the return to education may be important. In particular this is the case with respect to the existence of random errors in the measure of education, a problem which has not received much attention in the literature. However, our empirical results show that the various biases work in different directions and may cancel out to some extent. Our conclusion

regarding the average return to education in Sweden 1984 and 1986 illustrates this possibility. If we use the traditional wage equation we obtain marginal returns to education equal to 3.79 percent in 1984 and 4.16 percent in 1986. After correcting for sample selection, omitted variable bias, and measurement error, net returns of 4.44 percent and 3.80 percent are obtained in 1984 and 1986 respectively (for the comparable samples). Thus, while our analysis reversed the order of magnitude of the estimated returns generated by the traditional model, the aritmetic mean over 1984 and 1986 changed only slightly, from 3.96 to 4.12 percent.

This points to the importance of more empirical research to examine whether the biases in the estimated return to education tend to "cancel out" over longer time periods as well. This question also has a theoretical counterpart which deserves further research, namely how the biases associated with the different pitfalls are connected and under what conditions they reinforce or weaken each other.

Concerning the issue of omitted variable bias, we can draw two conclusions. First, our results show that social relations significantly influence earnings, both directly and indirectly, through education. To take these relations into account by extending the notion of family background to the sociological concept of social capital as we have done in this study apparently is a fruitful direction of research. Second, the empirics indirectly support the finding of Taubman (1976a, b) and others that schooling may be highly correlated with (innate) ability. We draw this conclusion because less than 40 percent of the variation in schooling is explained by means of social capital indicators and standard control variables and because, more importantly, the fit of the wage equation decreases substantially when we substitute predicted education for the education actually observed. The latter result indicates that our predicted education lacks a systematic component, probably an ability measure. Since our data do not contain any such measures we have had no choice but to try to do without them. However, as the education equation is the driving force in our model it is clear that further work to refine it and increase its explanatory power is warranted.

References

Becker, G.S. (1962), "Investment in Human Capital: A Theoretical Analysis", *Journal of Political Economy*, Vol 70 Supplement, Part 2, pp. S9-549.

Becker, G.S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, New York: National Bureau of Economic Research.

Berndt, E.R. (1991), *The Practice of Econometrics: Classic and Contemporary*, Reading , Massachusetts: Addison-Wesley.

Björklund, A. and C. Kjellström (1993), "Avkastningen på utbildning i Sverige 1968 till 1991", ("The Return to Education in Sweden between 1968 and 1991"), mimeo, Institutet för social forskning.

Boxmanm E.A.W, De Graaf P.M., and H.D. Flap (1991), "The Impact of Social and Human Capital on Income Attainment of Dutch Managers", *Social Act Works*, Vol 13, pp. 51-73.

Coleman, J.S. (1988), "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, Vol. 94, Suppl pp. 95-120.

Coleman, J.S. (1990), *Foundations of Social Theory*, Cambridge, The Belknap Press of Harvard University Press.

Edin, P.A., Fredriksson, P. and B. Holmlund (1994), "Utbildningsnivå och utbildningsavkastning i Sverige", ("The level of Education and the Return to Education in Sweden."), mimeo, Uppsala Universitet.

Freeman, R.B. (1975), "Overinvestment in College Training", *Journal of Human Resources*, Vol.10, pp. 287-311.

Granovetter, M.S. (1973), "The Strength of Weak Ties", *American Journal of Sociology*, Vol. 78, pp. 1360- 1380.

Griliches, Z. (1977), "Estimating the Returns to Schooling: Some Econometric Problems", *Econometrica*, Vol. 45, pp. 1-22.

Griliches, Z. (1979), "Sibling Models and Data in Economics: Beginnings of a Survey", *Journal of Political Economy*, Vol. 87, Part 2, pp. S37-64.

Griliches, Z. and W.M. Mason (1972), "Education, Income, and Ability", *Journal of Political Economy*, Vol. 80, pp. S74-103.

Härnqvist, K. (1993), "Skolprestation, utbildningsval och social selektion till gymnasium och högskola", ("School performance, choice of education, and social selection to high school and college"), in R. Eriksson and J.O. Jonsson (eds.): *Skola och sortering: studier i snedrekrrytering och utbildningskonsekvenser*, Stockholm, Carlssons förlag.

Klevmarken, A. and P. Olovsson (1993), Household Market and Nonmarket Activities: Procedures and Codes 1984D1991, Vol. I and Vol. II, Stockholm: The Industrial Institute for Economic and Social Research and Almqvist & Wiksell International.

Lam, D. and R.F. Schoeni (1993), "Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil", *Journal of Political Economy*, Vol. 101, pp. 710-740.

Lazear, E.P. and R.T. Michael (1980), "Family Size and the Distribution of Real Per Capita Income", *The American Economic Review*, Vol. 70, pp. 91-107.

Leibowitz, A. (1974), "Home Investments in Children", *Journal of Political Economy*, Vol. 82, Part 2, pp. 5115-5131.

Lin, N. (1982), "Social Resources and Instrumental Action", pp. 131-145 in P. Marsden and N. Lin (eds.) *Social Structure and Network Analysis*, Beverly Hills, California: Sage Publications.

Lin, N. (1988), "Social Resources and Social Mobility: A Structural Theory of Status Attainment", in R.L. Breiger (ed.) *Social Mobility and Social Structure*, Cambridge: Cambridge University Press.

Lin, N. , W. M. Ensel, and J. C. Vaughn (1981), "Social Resources and Strength of Ties: Structural Factors in Occupational Status Attainment", *American Sociological Review*, Vol. 46, pp. 393-405.

Loury, G. (1977), "A Dynamic Theory of Racial Income Differences", Ch. 8 in P.A. Wallace and A. Le Mund (eds.) *Women, Minorities, and Employment*

Discrimination, Lexington, Massachusetts: Lexington Books.

Loury, G. (1987), "Why Should We Care About Group Inequality?", *Social Philosophy and Policy*, Vol. 5, pp. 249-271.

Lydon, M. (1990), Movements in the Earnings/Schooling Relationship 1940-86, unpublished Ph.D. dissertation, Stanford University.

Meyerson, E. M. (1993), "Human Capital, Social Capital and Compensation: The Relative Contribution of Social Contacts to Managers' Incomes.", Working Paper No 395, IUI, Stockholm.

Mincer, J. (1957), "A Study of Personal Income Distribution", unpublished Ph.D. dissertation, New York: Columbia University, Department of Economics.

Mincer, J. (1958), " Investment in Human Capital and Personal Income Distribution", *Journal of Political Economy*, Vol. 66, pp. 281-302.

Mincer, J. (1962), "On the-Job Training: Costs, Returns, and Some Implications", *Journal of Political Economy*, Vol. 70, Part 2, pp. S50-579.

Mincer, J. (1974), *Schooling, Experience and Earnings*, New York: Columbia University Press for the National Bureau of Economic Research.

Pencavel, J. (1991), "Higher Education, Productivity, and Earnings: A Review", *Journal of Economic Education*, Fall, pp. 331-359.

Psacharopoulos, G. (1985), "Returns to Education: A Further International Update and Implications, *Journal of Human Resources*, Vol. 20, pp. 583-597.

Schultz, T.W. (1960), "Capital Formation by Education", *Journal of Political Economy*, Vol. 68, pp. 571-583.

Schultz, T.W. (1961), "Investment in Human Capital", *American Economic Review*, Vol. 51, pp.1-17.

SOU 1993:85 (1993), *Ursprung och utbildning: social snedrekrytering till högre studier, ("Origin and Education: Socially Biased Recruitment to Higher Education")*, Stockholm, Utbildningsdepartmentet.

Taubman, P. (1976a), "Earnings, Education, Genetics and Environment", *Journal of Human Resources*, Vol.11, pp. 447-461.

Taubman, P. (1976b), "The Determinants of Earnings: Genetics, Family and Other Environment - A Study of Male Twins", *American Economic Review*, Vol. 66, pp. 858-870.

Willis, R.J. (1986), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions ", pp. 525-602 in O. Aschenfelter and R. Layard (eds.) *Handbook of Labor Economics, Vol. l*, Amsterdam: North-Holland Elsevier Science Publishers.

Appendix A: Estimation Results

In the tables, the following conventions have been used. Below the estimates, in parenthesis, asymptotic T-ratios are given. Estimates that are significant on at least the 5 percent level of significance are marked be "**". Estimates that are significant at the 10 percent level of significance are marked by "*".

Note that Table 5 is partitioned. The upper part of the table gives the results for the education equations estimated for both "wage earners" and "others". The lower part of the table gives the results exlusive for the group of "wage-earners".

Equation	(10)	(10)	(10)
Sample	1984 A	1984 B	1986
Dep. var.: log hourly wage			
Human Capital Variables			
educyear	0.0376**	0.0374**	0.0416**
	(17.24)	(16.42)	(18.87)
noexp	-0.0062	0.0034	-0.0053
	(-0.081)	(0.042)	(-0.048)
exp	0.0126**	0.0122**	0.014**
	(4.42)	(4.08)	(5.00)
exp^2	-0.00024**	-0.00026**	-0.00017**
	(-4.53)	(-4.44)	(-3.21)
tenure	0.0046**	0.0047**	0.0042**
	(4.86)	(4.67)	(4.61)
Control Variables			
constant	2.9245**	2.8439**	3.2505**
	(54.00)	(45.94)	(60.08)
age	0.0092**	0.0102**	0.0020
	(5.57)	(5.82)	(1.23)
female	-0.0449	-0.0305	-0.1370**
	(-0.79)	(-0.51)	(-2.43)
female × age	-0.0047**	-0.0054**	-0.0031**
	(-3.69)	(-3.96)	(-2.47)
female × married	-0.0011	-0.0012	0.0315
	(-0.04)	(-0.04)	(1.25)
few-children	0.0128	0.0133	0.0208
	(0.86)	(0.85)	(1.46)
many-children	-0.0813	-0.0676	-0.0421
	(-1.34)	(-1.07)	(-0.72)
city-counties	0.0236	0.0273*	0.0332**
	(1.58)	(1.74)	(2.27)
forest-counties	-0.0372	-0.0370	-0.0095
	(-1.59)	(-1.52)	(-0.42)
female × forest-counties	0.0485	0.0477	-0.0079
	(1.56)	(1.46)	(-0.26)
part-time1	0.0588*	0.0614	0.3155**
	(1.64)	(1.60)	(7.71)
part-time2	0.0647**	0.0690**	0.0487**
	(3.65)	(3.70)	(2.83)
not-paid-overtime	0.1118**	0.1059**	0.0543*
	(3.52)	(3.24)	(1.68)
paid-overtime	-0.0750**	-0.0786**	-0.1042**
	(-3.99)	(-4.017)	(-5.63)
public-sector	-0.0683**	-0.0762**	-0.0920**
	(-3.22)	(-3.45)	(-4.46)
female-public-sector	0.0991**	0.1017**	0.1304**
	(3.58)	(3.51)	(4.83)
work	-	0.0078**	-
	-	(2.13)	-
R^2	0.387	0.392	0.387
$\bar{R}^2$	0.379	0.383	0.378
# observations	1522	1405	1472

Table 4: Wage equations, reference specifications

Equation	(2)	(2)	(2)
Sample	1984 A	1984 B	1986
Dep. var.: educyear			
Social Capital Variables			
mother-home	0.1520**	0.1310**	0.0994*
	(3.02)	(2.49)	(1.85)
age × mother-home	-0.0045**	-0.0039**	-0.0037**
	(-3.55)	(-2.90)	(-2.71)
father-educ0	1.1193**	1.0793**	1.0457**
	(4.32)	(3.99)	(4.27)
father-educ1	1.4582**	1.4903**	1.4841**
	(4.57)	(4.42)	(4.27)
father-educ2	3.2715**	3.2523**	2.9714**
	(8.65)	(8.26)	(7.97)
father-educ3	3.2987**	3.1678**	3.0376**
	(7.79)	(7.18)	(7.09)
father-educ4	3.4638**	3.3208**	3.1237**
	(8.04)	(7.40)	(7.52)
mother-educ0	0.7972**	0.7553**	0.6417**
	(2.91)	(2.62)	(2.43)
mother-educ1	1.8387**	1.7803**	1.1639**
	(4.80)	(4.49)	(3.15)
mother-educ2	1.6818**	1.6531**	1.5350**
	(4.66)	(4.35)	(4.25)
mother-educ3	2.2110**	2.2702**	1.4442**
	(4.68)	(4.52)	(3.01)
mother-educ4	3.0879**	3.2569**	2.7332**
	(5.75)	(5.63)	(5.18)
f & m-educ2-4	0.2179	0.1092	-0.0215
	(0.54)	(0.26)	(-0.05)
Swedish-at-home	0.0686	0.0365	0.0464
	(0.30)	(0.15)	(0.19)
father-whitecol	-1.7599**	-1.5179**	-1.2691*
	(-2.75)	(-2.26)	(-1.91)
father-bluecol	-1.1591**	-1.0117*	-1.1492*
	(-2.08)	(-1.71)	(-1.90)
age × father-whitecol	0.0518**	0.0485**	0.0481**
	(3.36)	(2.99)	(3.03)
age × father-bluecol	0.0235*	0.0199	0.0266*
	(1.78)	(1.42)	(1.88)
f & child-same-educ	-2.7045**	-2.6832**	-2.6043**
	(-16.12)	(-15.24)	(-15.16)
city-childhood	0.6757**	0.7073**	0.4921**
	(3.42)	(3.41)	(2.41)
Control Variables			
constant	10.5693**	10.9447**	12.0371**
	(12.02)	(11.84)	(12.55)
age	-0.0124	-0.0210	-0.0342
	(-0.60)	(-0.96)	(-1.53)
female	-0.4702**	-0.4718**	-0.3924**
	(-3.58)	(-3.44)	(-2.94)
school-reform	-1.4306**	-1.5357**	-1.4229**
	(-6.03)	(-6.13)	(-5.82)
R^2	0.380	0.371	0.363
$\bar{R}^2$	0.371	0.362	0.354
# observations	1694	1573	1606
Heckman's λ	0.7037**	0.6008*	0.8554**
	(2.21)	(1.86)	(2.51)
R^2	0.373	0.362	0.367
# observations	1523	1406	1573

Table 5: Education equations

Equation	(3)	(3)	(3)
Sample	1984 A	1984 B	1986
Human Capital Variables			
peducy	0.0183	0.0085	0.1125**
	(0.51)	(0.23)	(2.73)
noexp	-0.9101**	-0.8811**	-0.6264
	(-3.46)	(-3.26)	(-1.61)
exp	0.1458**	0.1518**	0.1532**
	(8.45)	(8.49)	(7.54)
exp^2	-0.0023**	-0.0024**	-0.0024**
	(-6.62)	(-6.71)	(-5.69)
Social Capital Variables			
mother-educ0	-0.2644	-0.2622	-0.2986
	(-1.58)	(-1.48)	(-1.49)
mother-educ1	-0.4638*	-0.4377*	-0.3744
	(-1.85)	(-1.69)	(-1.34)
mother-educ2	0.0617	0.0643	-0.6014**
	(0.26)	(0.26)	(-2.41)
mother-educ3	-0.1886	-0.2187	-0.3617
	(-0.60)	(-0.67)	(-1.06)
mother-educ4	0.0190	0.00048	-0.6551*
	(0.05)	(0.001)	(-1.91)
father-whitecol	-0.3010	-0.2945	0.0785
	(-0.71)	(-0.67)	(0.18)
father-bluecol	0.8452**	0.9163**	0.4710
	(2.37)	(2.47)	(1.17)
age × father-whitecol	0.01133	0.0126	-0.0026
	(1.02)	(1.05)	(-0.22)
age × father-bluecol	-0.0175**	-0.0196**	-0.0106
	(-2.03)	(-2.18)	(-1.06)
swedish-parents	0.3447**	0.2511	-0.0084
	(2.31)	(1.56)	(-0.04)
married	0.4999**	0.5026**	0.1369
	(2.39)	(2.33)	(0.60)
children-le15	-0.1095	-0.0734	-0.0170+
	(-1.39)	(-0.89)	(-0.16)
external contacts	-	0.00013	-
	-	(0.36)	-
introvert activities	-	-0.000049	-
	-	(-0.18)	-
extrovert activities	-	-0.00145	-
	-	(-1.54)	-

Equation, cont.	(3)	(3)	(3)
Control Variables			
constant	0.2779	0.3595	-0.4279
	(0.48)	(0.58)	(-0.63)
age	-0.0278**	-0.0266**	-0.0106
	(-2.28)	(-2.10)	(-0.75)
female	0.3421	0.3773	0.2044
	(0.94)	(1.01)	(0.52)
age × female	-0.00030	-0.0016	-0.0078
	(-0.03)	(-0.16)	(-0.71)
female × married	-0.4170	-0.4247	-1373
	(-1.49)	(-1.48)	(-0.46)
female × spouse-dead	-0.7559*	-0.8048**	-0.5422
	(-1.93)	(-1.99)	(-1.16)
female × children-le15	-0.1147	-0.1222	-0.2265*+
	(-1.14)	(-1.17)	(-1.87)
city-counties	0.2612**	0.2894**	-0.0395
	(2.23)	(2.38)	(-0.31)
forest-counties	-0.1552	-0.1618	-0.5393**
	(-0.89)	(-0.91)	(-2.76)
female × forest-counties	0.1683	0.1406	0.4159*
	(0.78)	(0.64)	(1.77)
correct predictions			
"wage earners"	99.2 %	99.1 %	99.4 %
"others"	20.9 %	22.0 %	9.6 %
total	91.3 %	90.9 %	91.9 %
# observations	1695	1574	1608

Table 6: Probit equations. *Note*: + The 1986 sample refers to children less than 14 years

Equation	(11)	(11)	(11)
Sample	1984 A	1984 B	1986
Dep. var.: log hourly wage			
Human Capital Variables			
educyear	0.0369**	0.0367**	0.0411**
	(16.97)	(16.17)	(18.48)
noexp	0.1767**	0.1965**	0.0792
	(2.00)	(2.14)	(0.66)
exp	0.00025	-0.0018	0.0071
	(0.06)	(-0.41)	(1.46)
exp^2	-0.000046	-0.000027	-0.000057
	(-0.63)	(-0.35)	(-0.68)
tenure	0.0045**	0.0046**	0.0042**
	(4.79)	(4.59)	(4.60)
Control Variables			
constant	3.0179**	2.9479**	3.3179**
	(51.65)	(44.69)	(50.69)
age	0.0116**	0.0128**	0.0028*
	(6.64)	(6.94)	(1.68)
female	-0.0410	-0.0280	-0.1305**
	(-0.72)	(0.47)	(-2.32)
female × age	-0.0046**	-0.0052**	-0.0031**
	(-3.57)	(-3.82)	(-2.44)
female × married	-0.0089	-0.0089	0.0325
	(-0.34)	(-0.32)	(1.29)
few-children	0.0270*	0.0253	0.0262*
	(1.78)	(1.60)	(1.81)
many-children	-0.0236	-0.0137	-0.0301
	(-0.38)	(-0.21)	(-0.51)
city-counties	0.0076	0.0071	0.0329**
	(0.49)	(0.43)	(2.25)
forest-counties	-0.0314	-0.0293	0.0039
	(-1.35)	(-1.21)	(0.16)
female × forest-counties	0.0437	0.0450	-0.0182
	(1.41)	(1.39)	(-0.59)
part-time1	0.0714**	0.0787**	0.3209**
	(2.00)	(2.05)	(7.83)
part-time2	0.0667**	0.0716**	0.0523**
	(3.78)	(3.86)	(3.02)
not-paid-overtime	0.1083**	0.1040**	0.0577*
	(3.43)	(3.20)	(1.79)
paid-overtime	-0.0737**	-0.0764**	-0.1027**
	(-3.94)	(-3.92)	(-5.54)
public-sector	-0.0682**	-0.0774**	-0.0932**
	(-3.23)	(-3.53)	(-4.51)
female × public-sector	0.1018**	0.1068**	0.1320**
	(3.70)	(3.72)	(4.89)
work	-	0.0085**	-
	-	(2.34)	-
Heckman's λ	-0.3291**	- 0.3536**	-0.1747*
	(-4.13)	(-4.37)	(-1.82)
R^2	0.394	0.401	0.388
$\bar{R}^2$	0.386	0.391	0.379
# observations	1522	1405	1472

Table 7: Wage equations, accounting for sample selection

Equation	(12)	(12)	(12)
Sample	1984 A	1984 B	1986
Dep. var.:log hourly wage			
Human Capital Variables			
educyear	0.0369**	0.0364**	0.0388**
	(15.08)	(14.21)	(15.60)
noexp	0.1483	0.1759*	0.1074
	(1.63)	(1.84)	(0.88)
exp	0.0023	-0.000046	0.0085*
	(0.51)	(-0.01)	(1.65)
exp^2	-0.000077	-0.000051	-0.000084
	(-1.00)	(-0.60)	(-0.95)
tenure	0.0044**	0.0044**	0.0042**
	(4.70)	(4.35)	(4.62)
Social Capital Variables			
mother-home	-0.0097*	-0.0082	0.00056
	(-1.88)	(-1.52)	(0.10)
age × mother-home	0.00026**	0.00021	-0.000047
	(2.00)	(1.58)	(-0.35)
father-educ2	-0.0127	-0.0174	-0.0029
	(-0.41)	(-0.54)	(-0.10)
father-educ3	-0.0526	-0.0583	-0.0537
	(-1.46)	(-1.55)	(-1.47)
father-educ4	-0.0524	-0.0496	0.0165
	(-1.41)	(-1.28)	(0.47)
mother-educ2	-0.0370	-0.0448	0.0131
	(-1.37)	(-1.58)	(0.49)
mother-educ3	-0.0047	-0.0093	0.0447
	(-0.11)	(-0.21)	(1.10)
mother-educ4	-0.0070	-0.0274	-0.0329
	(-0.14)	(-0.52)	(-0.71)
f & m-educ2-4	0.0438	0.0495	0.0388
	(1.12)	(1.20)	(1.00)
f & child-same-educ	0.0059	0.0044	-0.0228
	(0.34)	(0.24)	(-1.28)
father-whitecol	0.0114	-0.0039	0.0712
	(0.19)	(-0.06)	(1.21)
age × father-whitecol	0.0010	0.0014	-0.0013
	(0.68)	(0.92)	(-0.94)
city-childhood	0.0242	0.0247	0.00081
	(1.16)	(1.13)	(0.04)
external contacts	-	0.000045	-
	-	(0.79)	-
introvert activities	-	-0.000078**	-
	-	(-2.05)	-
extrovert activities	-	-0.000022	-
	-	(-0.15)	-

Equation, cont.	(12)	(12)	(12)
Control Variables			
constant	3.1299**	3.0613**	3.2858**
	(37.90)	(33.67)	(35.70)
age	0.0077**	0.0094**	0.0043*
	(2.97)	(3.44)	(1.67)
female	-0.0439	-0.0409	-0.1190**
	(-0.77)	(-0.68)	(-2.09)
female × age	-0.0046**	-0.0052**	-0.0034**
	(-3.58)	(-3.75)	(-2.65)
female × married	-0.0061	0.00025	0.0315
	(-0.23)	(0.001)	(1.24)
few-children	0.0276*	0.0281*	0.0248*
	(1.80)	(1.75)	(1.69)
many-children	-0.0338	-0.0218	-0.0363
	(-0.54)	(-0.34)	(-0.61)
city-counties	0.0021	0.0020	0.0294*
	(0.13)	(0.11)	(1.88)
forest-counties	-0.0313	-0.0301	0.00098
	(-1.34)	(-1.23)	(0.040)
female × forest-counties	0.0479	0.0500	-0.0147
	(1.54)	(1.54)	(-0.48)
part-time1	0.0666*	0.0756**	0.3101**
	(1.86)	(1.96)	(7.53)
part-time2	0.0637**	0.0687**	0.0531**
	(3.60)	(3.68)	(3.05)
not-paid-overtime	0.1057**	0.0997**	0.0563*
	(-3.34)	(3.06)	(1.74)
paid-overtime	-0.0743**	-0.0784**	-0.1014**
	(-3.96)	(-4.01)	(-5.47)
public-sector	-0.0718**	-0.0809**	-0.0932**
	(-3.40)	(-3.68)	(-4.49)
female × public-sector	0.1085**	0.1154**	0.1296**
	(3.93)	(4.00)	(4.78)
work	-	0.0087**	-
	-	(2.19)	-
Heckman's λ	-0.2949**	-0.3183**	-0.1593
	(-3.35)	(-3.48)	(-1.53)
R^2	0.401	0.409	0.394
$\bar{R}^2$	0.388	0.393	0.380
# observations	1522	1405	1472

Table 8: Wage Equations; accounting for sample selection and omitted variables

Equation	(13)	(13)	(13)
Sample	1984 A	1984 B	1986
Dep. var.:log hourly wage			
Human Capital Variables			
peduc	0.0438**	0.0414**	0.0380**
	(4.58)	(3.99)	(3.59)
noexp	0.1643*	0.2117**	0.1650
	(1.68)	(2.07)	(1.25)
exp	-0.0054	-0.0075	-0.0011
	(-1.08)	(-1.38)	(-0.20)
exp^2	-0.0000087	0.0000074	0.0000045
	(-0.10)	(0.08)	(0.05)
tenure	0.0045**	0.0045**	0.0040**
	(4.50)	(4.24)	(4.14)
Social Capital Variables			
mother-home	-0.0113**	-0.0095	0.0018
	(-1.97)	(-1.59)	(0.30)
age × mother-home	0.0003**	0.00025	-0.000078
	(2.06)	(1.63)	(-0.51)
father-educ2	-0.0481	-0.0470	-0.0160
	(-1.21)	(-1.13)	(-0.42)
father-educ3	-0.0825*	-0.0810*	-0.0672
	(-1.86)	(1.75)	(-1.49)
father-educ4	-0.0839*	-0.0745	0.0011
	(-1.81)	(-1.54)	(0.024)
mother-educ2	-0.0534*	-0.0607*	0.0193
	(-1.74)	(-1.87)	(0.62)
mother-educ3	-0.0282	-0.0305	0.0368
	(-0.61)	(-0.62)	(0.82)
mother-educ4	-0.0506	-0.0675	-0.0388
	(-0.89)	(-1.08)	(-0.71)
f & m-educ2-4	0.0504	0.0561	0.0370
	(1.21)	(1.28)	(0.88)
f & child-same-educ	0.0294	0.0212	-0.0114
	(1.03)	(0.69)	(-0.37)
father-whitecol	0.0494	0.0377	0.0801
	(0.75)	(0.54)	(1.26)
age × father-whitecol	-0.00018	0.00017	-0.0016
	(-0.11)	(0.10)	(-1.03)
city-childhood	0.0087	0.0098	-0.0095
	(0.37)	(0.40)	(-0.41)
external contacts	-	0.000067	-
	-	(1.11)	-
introvert activities	-	-0.00011**	-
	-	(-2.64)	-
extrovert activities	-	-0.000033	-
	-	(-0.20)	-

Equation, cont.	(13)	(13)	(13)
Control Variables			
constant	2.9788**	2.9165**	3.1846**
	(24.20)	(21.54)	(21.92)
age	0.0121**	0.0142**	0.0103**
	(4.46)	(4.90)	(3.79)
female	0.0241	0.0193	-0.0303
	(0.40)	(0.30)	(-0.50)
female × age	-0.0061**	-0.0066**	-0.0057**
	(-4.49)	(-4.53)	(-4.22)
female × married	-0.0154	-0.0071	0.0323
	(-0.55)	(-0.24)	(1.18)
few-children	0.0363**	0.0368**	0.0307*
	(2.22)	(2.15)	(1.93)
many-children	-0.0599	-0.0544	-0.0416
	(-0.90)	(-0.79)	(-0.65)
city-counties	0.0106	0.0092	0.0442**
	(0.61)	(0.50)	(2.63)
forest-counties	-0.0355	-0.0358	0.0103
	(-1.42)	(-1.38)	(0.39)
female × forest-counties	0.0517	0.0585*	-0.0289
	(1.56)	(1.69)	(-0.87)
part-time1	0.0447	0.0511	0.2806**
	(1.17)	(1.24)	(6.34)
part-time2	0.0495**	0.0560**	0.0529**
	(2.63)	(2.82)	(2.82)
not-paid-overtime	0.1743**	0.1652**	0.1340**
	(5.23)	(4.81)	(3.89)
paid-overtime	-0.0688**	-0.0751**	-0.0901**
	(-3.44)	(-3.61)	(-4.52)
public-sector	-0.0146	-0.0231	-0.0441**
	(-0.66)	(-1.00)	(-2.00)
female × public-sector	0.0826**	0.0908**	0.1107**
	(2.81)	(2.96)	(3.79)
work	-	0.0102**	-
	-	(2.64)	-
Heckman's λ	-0.3847**	-0.4112**	-0.2696**
	(-3.91)	(-4.02)	(-2.38)
R^2	0.319	0.330	0.298
$\bar{R}^2$	0.304	0.311	0.281
# observations	1522	1405	1472

Table 9: Wage equations; accounting for sample selection, omitted variables and measurement error

Appendix B: Definitions of Variables

Endogenous Variables

LHWAGE	Logged hourly wage (before taxes). The hourly wage is defined as the total monthly wage divided by working hours per month. The total monthly wage includes wage income derived from extra work (secondary occupation) and from overtime work. In some cases direct measures of the wage income are not available. In these cases the wage income has been imputed by means of the previous year's wage income and the present year's employment hours.
EDUCYEAR	Number of years of schooling and vocational training. Part-time schooling is converted to full-time equivalent.

Exogenous Variables

Human Capital Variables

EDUCYEAR	See endogenous variables.
EXP	Number of years that the respondent has worked three months or more.
EXP2	EXP**2
NOEXP	If the respondent has EXP=0 then NOEXP:=1 else NOEXP:=0.
TENURE	Number of years since the respondent was hired by the current employer

Social Capital Variables

MOTHER_HOME	Number of years that the respondent's mother has been working at home until the respondent turned 16. Can only take on the values: 16 - never worked outside home, 13 - worked outside home for 1-5 years, 8 - worked outside home for 6-10 years, 3 - worked outside home for more than 10 years.
AGE*MOTHER_HOME	Interaction variable
FATHER_EDUC0	Father's highest level of education is... ...elementary school, at least six years. (binary)

FATHER_EDUC1	...vocational school. (binary)
FATHER_EDUC2	...intermediate school-leaving examination / comprehensive school / adult education. (binary)
FATHER_EDUC3	...gymnasium (upper secondary school). (binary)
FATHER_EDUC4	...university, college. (binary)
MOTHER_EDUC0	See above
MOTHER_EDUC1	See above
MOTHER_EDUC2	See above
MOTHER_EDUC3	See above
MOTHER_EDUC4	See above
F&M_EDUC2-4	Both parents' education correspond to groups 2, 3 or 4. (binary)
F&CHILD_SAME_EDUC	The respondent and his/her father have the same level of education. (binary)
FATHER_WHITECOL	The respondent's father is a white-collar worker. (binary)
FATHER_BLUECOL	The respondent's father is a blue-collar worker. (binary)
AGE*FATHER_WHITECOL	Interaction variable.
AGE*FATHER_BLUECOL	Interaction variable.
CITY_CHILDHOOD	The respondent grew up in a large metropolitan area in Sweden. (binary)
SWEDISH_AT_HOME	While the respondent was growing up Swedish was usually spoken at home. (binary)
SWEDISH_PARENTS	If both parents were Swedish citizens when the respondent was born then SWEDISH_PARENTS:=1, if both parents were foreign citizens then SWEDISH_PARENTS:=0 and if one and only one parent was a Swedish citizen then SWEDISH_PARENTS:=0.5.
MARRIED	Respondent is married or living with a partner. (binary)
CHILDREN_LE15	Number of children in the household aged 0-15.
(CHILDREN_LE14	Number of children in the household aged 0-14.)
EXTERNAL_CONTACTS	Number of minutes per day spent on entertaining guests at home, visiting friends, meals at restaurants, travel in connection with taking care of someone, telephone calls, and conversation with persons other than household members. **Note:** This variable is only available for 1984.
EXTROVERT_ACTIVITIES	Number of minutes per day spent on organizational activities, religious activities, spectator activities outside the home, and other recreational activities. **Note:** This variable is only available for 1984.
INTROVERT_ACTIVITIES	Number of minutes per day spent on at-home activities related to hobbies, TV or radio, reading, doing nothing, writing or reading letters, and conversation among household members. **Note:** This variable is only available for 1984.

Control Variables

CONSTANT	
AGE	Respondent's age.
FEMALE	Respondent's sex. 1 if respondent is female else 0.
FEMALE*AGE	Interaction variable.
FEMALE*MARRIED	Interaction variable.
FEMALE*SPOUSE_DEAD	Interaction variable. SPOUSE_DEAD:=1 if the respondent is widowed else 0.
FEMALE*CHILDREN_LE15	Interaction variable.
FEW_CHILDREN	Number of children in the household is 1-3. (binary)
MANY_CHILDREN	There are 4 children or more in the household. (binary)
CITY_COUNTIES	Respondent is living in a large metropolitan area, i.e., Stockholm, Göteborg or Malmö. (binary)
FOREST_COUNTIES	Respondent is living in a one of the counties, i.e., Värmland, Kopparberg, Gävleborg, Västernorrland, Jämtland, Västerbotten or Norrbotten. (binary)
FEMALE*FOREST_COUNTIES	Interaction variable.
PART-TIME1	Respondent is working 1-19 hours/week. (binary)
PART-TIME2	Respondent is working 20-34 hours/week. (binary)
SCHOOL_REFORM	Respondent is born 1957 or later (binary), implying that he/she spent at least one year in mandatory school after the mandatory school reform 1971.
NOT_PAID_OVERTIME	Respondent is working overtime, but is not compensated for the extra hours. (binary)
PAID_OVERTIME	Respondent is working overtime, and is compensated for the extra hours. (binary)
PUBLIC	Respondent is employed in the public sector. (binary)
FEMALE*PUBLIC	Interaction variable.
WORK	Respondent's assessment of his/her gainful employment on a scale from 0 to 10, where 10 is the highest rating. **Note:** This variable is only available for 1984.