IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

Proxying ability by family background in returns to schooling estimations is generally a bad idea

Erik Mellander
Sofia Sandgren-Massih

WORKING PAPER 2008:22

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market policies, studies of the functioning of the labour market, the labour market effects of educational policies and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P O Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

Proxying ability by family background in returns to schooling estimations is generally a bad idea*

Erik Mellander† Sofia Sandgren-Massih‡

October 20, 2008

Abstract

A regression model is considered where earnings are explained by schooling and ability. It is assumed that schooling is measured with error and that there are no data on ability. Regressing earnings on observed schooling then yields an estimate of the return to schooling that is subject to positive omitted variable bias (OVB) and negative measurement error bias (MEB). The effects on the OVB and the MEB from using family background variables as proxies for ability are investigated theoretically and empirically. The theoretical analysis demonstrates that the impact on the OVB is uncertain, while the MEB invariably increases in magnitude. The empirical analysis shows that the MEB generally dominates the OVB. As the measurement error increases and/or more family background variables are added, the total bias rapidly becomes negative, driving the estimated return further and further away from the true value.

JEL codes: C13, C20, C52, J31
Keywords: missing data, proxy variables, measurement error, consistent estimates of omitted variable bias and measurement error bias.

---

*Financial support from the Swedish Agency of Innovation Systems and the Swedish Research Council is gratefully acknowledged. Thanks to Colm Harmon for comments on an early version of this paper. Seminar participants at the Institute for labour market policy evaluation, in particular Björn Öckert, provided constructive criticism on a later version. Finally, the paper was considerably improved by useful comments from two anonymous referees. We also thank Anna Sjögren for help with a graphical illustration.

†Institute of Labour Market Policy Evaluation (IFAU), P.O. Box 513, SE – 751 20 Uppsala, Sweden. Email: erik.mellander@ifau.uu.se

‡Dept. of Economics, Uppsala University, Uppsala, Sweden, and Centre Banking and Finance, Dept. of Infrastructure, Royal Institute of Technology, Stockholm, Sweden. Email: sofia.sandgren@nek.uu.se

1 Introduction

Many practitioners estimating the return to schooling have noted the tendency for the return estimates to fall when, for want of ability measures, family background variables are included in the earnings equation. Could this be a general property, i.e. is it possible to demonstrate analytically that it holds under a large variety of circumstances?

Lam and Schoeni (1993) claim that this is indeed the case. Referring to Welch (1975) and Griliches (1977), they note that if there is measurement error in the schooling variable, the inclusion of a variable that is correlated with a worker's schooling may increase the measurement error bias as well as reduce the omitted variable bias. To take this into account, they provide equations for the asymptotic bias in the estimated return to schooling where the total bias is additively decomposed into omitted variable bias and measurement error bias, before and after the inclusion of a family background variable. Without proof, they claim (op. cit., p. 719) that "... under plausible assumptions ...": i) the omitted variable bias is positive in both cases but smaller after the inclusion of the family background variable and ii) the measurement error bias is negative in both cases but larger in magnitude when the family background variable is included. The addition of the family background variable is thus claimed to affect the omitted variable bias and the measurement error bias in the same direction; both changes lower the estimated return to schooling.

In a subsequent empirical analysis, Lam and Schoeni (op cit) implicitly extend these theoretical conclusions, drawn in the context of a single family background variable, to the case with many family background variables.

The purpose of this paper is threefold. The first purpose is to correct an error in Lam and Schoeni's theoretical analysis of the effects of including a single family background variable in the earnings regression. Section 2 demonstrates that while their conclusion ii) is right, conclusion i) is in general incorrect. It is shown, however, that there are restrictive conditions under which i) does hold true. We also consider related results that have been established earlier in another literature, focussing specifically on omitted variable bias.

The second purpose is to extend the theoretical analysis to an arbitrary number *(K)* of family background variables. Section 3 shows, i.a., that the property that there are conditions under which the omitted variable bias is reduced towards zero does not carry over to the case with two or more family background variables. The measurement error bias results obtained in the context of one family background variable do extend to the general case, however.

The final purpose is to provide an empirical assessment of the effects of proxying ability by means of family background variables. How is the omitted variable bias and the measurement error bias affected? What does this imply for the total bias?

To answer these questions we need data on ability, i.e., the information whose absence is the very reason for the problem considered. It might seem odd to address the problem of proxying ability when data on ability are actually available, but only in this way can conclusions be drawn about the consequences of proxying when ability data are *not* at hand. Second, as schooling is treated as predetermined in the theoretical analysis we either need information substantiating this assumption or, otherwise, an instrument for observed schooling.

The unique data set that we employ satisfies these requirements. It is a panel data set, covering 555 males, born 1928 in the city of Malmö in the south of Sweden. Details on the data are provided in Section 5. In Section 4 we derive estimates of the omitted variable bias and the measurement error bias conditional on a given ratio of the measurement error variance to the total variance in schooling, thus establishing that explicit information on the amount of measurement error in the schooling variable is not a prerequisite for the empirical analysis in Section 6. Concluding comments are provided in Section 7.

2 The case with one family background variable

To facilitate comparison with the results in Lam and Schoeni (op cit) we derive our basic results in the context of the stylized model that they employ. Next, in Section 2.2, we show how the analysis can be extended to accommodate an arbitrary number of control varibles.

2.1 Correction of the results in Lam and Schoeni (1993)

The starting point in Lam and Schoeni (henceforth LS) is the following equation, giving the "true" relation between (log) earnings, Y, schooling, S, and (unknown) ability, A, for the ith individual

$$Y_i = \beta_0 + \beta_s S_i + \beta_a A_i + u_i, \qquad \text{where} \qquad \beta_s, \beta_a > 0, \tag{1}$$

and u_i is a random disturbance with zero mean and constant variance.[1] For simplicity, the individual observations will be treated as random draws from the same underlying population. The u_i are thus viewed as realizations of the random variable u, characterized by $E(u) = 0$ and $Var(u) = \sigma_u^2$. In addition, it is assumed that the schooling variable is measured with error, such that observed schooling, S^*, can be expressed according to

$$S_i^* = S_i + w_i, \tag{2}$$

where w represents pure measurement error uncorrelated with S, i.e. $E(w) = 0$, $Var(w) = \sigma_w^2$, and $Cov(S, w) = 0$. Finally, LS implicitly take w to be uncorrelated with A and u, as well, and both u and w to be uncorrelated with the family background variable, F. Thus:

$$Cov(w, S) = Cov(w, A) = Cov(w, u) = Cov(w, F) = Cov(u, F) = 0. \tag{3}$$

[1] Lam and Schoeni do not explicitly state the positivity constraints on β_s and β_a. They consistently use these restrictions in their discussion about omitted variable bias and measurement error bias, however.

Another assumption implicitly made by LS is that schooling can be treated as a predetermined variable. In general, this is a strong assumption; see, e.g., Card (1999). However, in the present context it is merely a simplifying assumption which allows us to focus on the problems of omitted variable bias and measurement error bias.[2]

LS first consider the case when Y is simply regressed on S^*, i.e. when the unobserved ability variable is disregarded and the measurement error in schooling ignored. The probability limit of the estimated return to education is then given by

$$plim\,\widehat{\beta}_S = \beta_s - \beta_s\lambda + \beta_a\hat{\beta}_{AS}(1-\lambda), \tag{4}$$

where λ is the noise-to-signal ratio, i.e.

$$\lambda \equiv \frac{Var(w)}{Var(S^*)}, \qquad 0 \leq \lambda < 1, \tag{5}$$

and $\hat{\beta}_{AS}$ is the coefficient from a hypothetical regression of ability on true schooling

$$\hat{\beta}_{AS} \equiv \frac{Cov(A,S)}{Var(S)}, \qquad \hat{\beta}_{AS} > 0 \tag{6}$$

The second term on the RHS of (4) is the measurement error bias and the third term is the omitted variable bias. It can be seen that the measurement error bias is negative, and increasing in magnitude with the variance of the measurement error. Since it is assumed that schooling and ability are positively correlated, cf (6), the omitted variable bias is positive. It should be noted that, in general, one cannot rule out the possibility that the measurement error bias dominates the omitted variable bias, in which case the total bias is negative.

Given (4), LS consider how the probability limit of the estimate $\hat{\beta}_S$ is affected if a measure of family background, F, is added to the regression. Their result for this case contains an error, however. The correct expression is provided in Proposition 1. LS's equation is considered immediately after the proposition. Three corollaries to Proposition 1 are then given. The last of these provides a bridge between the general result in Proposition 1 and the equation suggested by LS.

Proposition 1 *Given (1), (2), and (3), OLS regression of Y on S^* and a family background measure, F, yields an estimate of β_s whose probability limit is given by*

$$plim\hat{\beta}_{S\cdot F} = \beta_s - \beta_s\frac{\lambda}{1-R^2_{S^*F}} + \beta_a\hat{\beta}_{AS}\,(1-\lambda)\left(1-\theta\cdot\rho^2_{AF\cdot S^*}\right) \tag{7}$$

where λ and $\hat{\beta}_{AS}$ are defined by (5) and (6), respectively. Further, $R^2_{S^*F}(<1)$ is the squared correlation of S^* and F, and $\rho^2_{AF\cdot S^*}$ is the squared partial correlation of ability

[2] It is always possible to think of S as an instrument for schooling – rather than schooling itself – and w as an associated random error. Thus, if schooling is endogenous our results can be applied once an instrument has been substituted for the schooling variable.

and the family background measure when one controls for schooling, i.e.

$$\rho^2_{AF\cdot S^*} = \left(\frac{\rho_{AF} - \rho_{AS^*} \cdot \rho_{S^*F}}{\sqrt{1-\rho^2_{AS^*}}\sqrt{1-\rho^2_{S^*F}}} \right)^2,$$

while θ is defined according to

$$\theta = \frac{\rho_{S^*F}/\rho_{AS^*} - \rho_{AS^*} \cdot \rho_{S^*F}}{\rho_{AF} - \rho_{AS^*} \cdot \rho_{S^*F}}; \qquad \rho_{AF} - \rho_{AS^*} \cdot \rho_{S^*F} \neq 0,$$

where ρ_{S^*F}, ρ_{AS^*}, and ρ_{AF} denote bivariate correlations.

Proof. See Appendix.

The equation provided by LS [eq. (8) in their paper] is

$$plim\hat{\beta}_{S\cdot F} = \beta_s - \beta_s \frac{\lambda}{1 - R^2_{S^*\cdot F}} + \beta_a \hat{\beta}_{AS} (1-\lambda) \left(1 - \rho^2_{AF\cdot S^*}\right). \tag{8}$$

This equation differs from equation (7) with respect to the final term, i.e. the expression for the omitted variable bias. More specifically, the last parenthesis reads $\left(1 - \rho^2_{AF\cdot S^*}\right)$ instead of $\left(1 - \theta \cdot \rho^2_{AF\cdot S^*}\right)$ in (7). Since $\rho^2_{AF\cdot S^*} \in [0,1]$ by construction, and $\rho^2_{AF\cdot S^*} \in]0,1[$ by assumption, (8) implies that the omitted variable bias invariably decreases towards zero upon the inclusion of a family background variable. Corollary 2 shows, however, that the omitted variable bias may well increase, thus driving the estimate of β_s upward, rather than downward. This possibility was, for obvious reasons, overlooked by LS.

Corollary 2 *If schooling and the family background variable are correlated, i.e. if $R^2_{S^*F} > 0$, then the inclusion of the family background variable unambiguously increases the measurement error bias, compared to when no family background variable is included. If $R^2_{S^*F} = 0$ the measurement error bias will be unchanged. The effect on the omitted variable bias cannot be determined a priori; the omitted variable bias may decrease or it may increase. This is true also in the absence of measurement error.*

Proof The first part of the corollary follows trivially from the facts that, by construction, $R^2_{S^*F} \in [0,1]$ and, by assumption, $R^2_{S^*F} < 1$. The second part follows from the fact that θ may be both negative and positive. Moreover, θ is not bounded, either from below or from above. If θ is negative the omitted variable bias increases with certainty. If θ is positive the omitted variable bias may decrease – if θ is small enough to ascertain that $\theta \cdot \rho^2_{AF\cdot S^*} < 1$. But it might also increase – if θ is large enough to yield $\theta \cdot \rho^2_{AF\cdot S^*} < 1$.[3] Finally, comparison of (4) and Proposition 1 shows that the effect on

[3]In principle, $\theta = 0$ is also a possibility. However, that requires $\rho_{AS} = 1$, which is a pathological case in the sense that it implies that the parameters in (1) cannot be identified.

the omitted variable bias is manifested in the factor $(1 - \theta \cdot \rho^2_{AF \cdot S^*})$; it is independent of the factor $(1 - \lambda)$, i.e. the extent of measurement error. Q.E.D.

To provide some intuition for Corollary 2, note that controlling for F means purging the schooling variable of "A – factors" *and* of other factors as well. At some point, the latter, negative, effect will outweigh the former, positive, effect.

In Corollary 3 we discuss a special case of the general situation considered in Corollary 2. A constraint on θ is provided which ascertains that the omitted variable bias stays positive *and* is reduced towards zero. A condition is also given which is necessary, but not sufficient, for this constraint to be satisfied.

Corollary 3 *If, and only if,* $0 < \theta \leq 1/\rho^2_{AF \cdot S^*}$ *then the positive omitted variable bias will remain positive and be reduced towards zero when the family background variable is added to the earnings equation. A necessary, but not sufficient, condition for these inequalities to hold is that* $sign(\rho_{AF}) = sign(\rho_{S^*F})$.

Proof. That the constraint implies that the omitted variable is reduced while staying positive follows directly from the fact that the change in the bias is determined by $(1 - \theta \cdot \rho^2_{AF \cdot S^*})$ where $\rho^2_{AF \cdot S^*} \in]0, 1]$. For values of θ above the value of $\rho^2_{AF \cdot S^*}$ the omitted variable bias changes sign. For $\theta = 0$ the omitted variable bias is unaffected and thus not reduced. For $\theta < 0$ the omitted variable bias increases.

To prove the necessary condition, first consider the case when $\rho_{S^*F} > 0$. In this case the numerator of θ is unambiguously positive; cf. the definition of θ in Proposition 1 and remember that $\rho_{AS^*} = \rho_{AS} \in]0, 1[$, by assumption. A necessary requirement for θ to be positive, which in turn is necessary for θ to belong to $]0, 1/\rho^2_{AF \cdot S^*}]$, is thus that the denominator of θ is positive, too. This requires $\rho_{AF} > 0$. But it may be that $0 < \rho_{AF} < \rho_{AS^*} \cdot \rho_{S^*F}$ in which case $\theta < 0$. Hence, for $\rho_{S^*F} > 0$ the condition $\rho_{AF} > 0$ is necessary but not sufficient for the omitted variable bias to remain positive and be reduced towards zero. In a perfectly analogous way it can be shown that if $\rho_{S^*F} < 0$ then $\rho_{AF} < 0$ is a necessary but not sufficient condition for maintaining the omitted variable bias positive and decreasing it towards zero. The case $\rho_{S^*F} = 0$ can be disregarded because it implies $\theta = 0$. Putting the results for the cases $\rho_{S^*F} > 0$ and $\rho_{S^*F} < 0$ together one obtains the necessary but not sufficient condition stated in the corollary. Q.E.D.

Corollary 4 considers a special case of the special case characterized in Corollary 3, namely when $\theta = 1$, the constraint implicitly imposed by LS. Corollary 4 provides an interpretation of this constraint, in terms of the correlation between schooling and family background, conditional on ability.

Corollary 4 *If the correlation between schooling and family background is equal to zero when ability is controlled for, i.e. if* $\rho_{S^*F \cdot A} = 0$*, then* $\theta = 1$*. This is a sufficient, but not necessary, condition for the positive omitted variable bias to decrease towards zero when one family background variable is included in the earnings regression.*

Proof. Note that, by definition,

$$\rho_{S^*F\cdot A} = \frac{\rho_{S^*F} - \rho_{AS^*} \cdot \rho_{AF}}{\sqrt{1-\rho_{AS^*}^2}\sqrt{1-\rho_{AF}^2}} \tag{9}$$

Thus, if $\rho_{S^*F\cdot A} = 0$ then $(\rho_{S^*F}/\rho_{AS^*}) = \rho_{AF}$ and, consequently, $\theta = 1$. Given $\theta = 1$, the second part follows directly from Corollary 3. Q.E.D.

At this point, a relevant question is whether the assumption $\theta = 1$ is not implicit in the model which provides the starting point of LS's analysis, i.e. (1) – (3)? The answer is no. The best way to see this is to note that just as the specification (1) – (3) allows F to be an instrument for A, it also allows F to be an instrument for S^*. That F is an admissible instrument for A follows from the properties $Cov(u, F) = 0$, given in (3), and $\rho_{AF\cdot S^*} \neq 0$, which follows from Proposition 1. Likewise, $Cov(u, F) = 0$ and $\rho_{S^*F\cdot A} \neq 0$ together imply that F is a valid instrument for S^*, too. The only way to exclude the latter possibility is to explicitly specify that $\rho_{S^*F\cdot A} = 0$ and this is certainly not done by LS.[4]

An alternative to the conditions provided in Corollary 3 and Corollary 4 is to impose constraints on F that imply $\rho_{S^*F\cdot A} = 0$. One (trivial) example is provided by McCallum (1972) and Wickens (1972) in the context of omitted variable bias only. Applied to the present context, their analyses assume that $F = A + \eta$ where η is a random error exhibiting zero correlation with all of the model's explanatory variables and all of its random terms.[5] Under this assumption $\rho_{S^*F} = \rho_{AS^*}$ and $\rho_{AF} = 1$, implying that $\rho_{S^*F\cdot A} = 0$.

Yet another alternative is to consider if the assumption $\rho_{S^*F\cdot A} = 0$ can be justified from an empirical point of view. Again, the answer is no. On the contrary, it is quite natural to think of family background as reflecting both nature (ability) and nurture (schooling). To the extent that this view is justified, $\rho_{S^*F\cdot A}$ is most likely to be non-zero. To illustrate this, we provide three examples of $\rho_{S^*F\cdot A}$ at the end of Section 5.

2.2 Allowing for control variables

Frost (1979) provides (without proof) results that are qualitatively equivalent to Proposition 1, in the context of omitted variable bias only.[6] In so doing, he allows for an arbitrary number of control variables, X_1, $X_2, ...,$ X_M. Frost's approach amounts to preceding the above analysis by purging the right hand side variables from the influences of the control variables. To this end, consider estimating the following regressions

[4] It should be noted that $\rho_{S^*F\cdot A}$ is not affected by the fact that F is not included in equation (1). Leaving out F amounts to assuming $\rho_{YF\cdot S,A} = 0$ which is equivalent to $\rho_{uF} = 0$. The latter implies $Cov\,(u, F) = 0$ which we have explicitly accounted for in (3).

[5] I.e. $Cov\,(\eta, S^*) = Cov\,(\eta, A) = Cov\,(\eta, u) = Cov\,(\eta, w) = 0$.

[6] We thank one of the referees for pointing out Frost's work to us.

by OLS:

$$\begin{array}{rcl} S_i & = & \gamma_0 + \gamma_1 X_{1i} + \; .... + \gamma_M X_{Mi} + e_{S_i} \\ A_i & = & \delta_0 + \delta_1 X_{1i} + \; .... + \; \delta_M X_{Mi} + e_{A_i} \\ F_i & = & \kappa_0 + \kappa_1 X_{1i} + \; .... + \; \kappa_M X_{Mi} + e_{F_i} \end{array} \tag{10}$$

What Frost's analysis shows is that for $\lambda = 0$ the estimated residuals $\widehat{e}_{S_i}$, $\widehat{e}_{A_i}$, and $\widehat{e}_{F_i}$ can be substituted for S_i, A_i, and F_i, respectively, in Proposition 1 and Corollaries 2 – 4, and the results will still go through – although they must, of course, be reinterpreted as "control-variables-corrected".

Fortunately, allowing for $\lambda \neq 0$ is unproblematic. In this case, the first regression is replaced by

$$S_i^* = \gamma_0 + \gamma_1 X_{1i} + \; .... + \gamma_M X_{Mi} + (e_{S_i} + w_i) \tag{11}$$

and $\widehat{e}_{S_i^*} = \widehat{e_{S_i} + w_i}$ is substituted for S_i^*. Since $Cov\,(e_{S_i}, w_i) = 0$, by (3), we have $Var\,(e_{S_i} + w_i) = Var\,(e_{S_i}) + Var\,(w_i)$. This means that given an assumption about λ (cf. Section 6), which yields an estimate of $Var\,(w_i)$, an estimate of $Var\,(e_{S_i})$ can also be determined, which would be needed in the "control-variables-corrected" version of (16) in Section 4.

3 Several family background variables

In this section the number of family background variables will be taken to be equal to $K \geq 1$. For analytical simplicity we abstract from control variables; from Section 2.2 it should be clear that this simplification can be made without loss of generality. The K-variable counterpart to Proposition 1 is given by the following proposition.

Proposition 5 *Given (1), (2), and (3), OLS regression of Y on S^* and a $K \times 1$ vector $\boldsymbol{F}$ of family background variables yields an estimate of β_s whose probability limit is given by*

$$plim\hat{\beta}_{S\cdot F} = \beta_s - \beta_s \cdot \frac{\lambda}{1 - R^2_{S^*F}} + \beta_a \hat{\beta}_{AS} \frac{(1-\lambda)}{1 - R^2_{S^*F}} \left[1 - \sum_{j=1}^{K} \frac{\rho_{AF_j}}{\rho_{AS^*}} \frac{\sqrt{Var\,(F_j)}}{\sqrt{Var\,(S^*)}} plim\,(\hat{\alpha}_j) \right]$$

where λ and $\hat{\beta}_{AS}$ are defined by (5) and (6), respectively, and $\hat{\alpha}_j$ is the OLS estimate of the coefficient for F_j in the linear regression of S^* on $\boldsymbol{F}$.

Proof. See Appendix.

There are several features of Proposition 5 that are worth noting. The first is that the result for the measurement error bias is a straightforward extension of the result in the case with one family background variable. Inclusion of family background variables will always increase the negative measurement bias, thus driving the estimate of β_s downward.

The second interesting property is that, like the measurement error bias, the omitted variable bias is inversely related to $1 - R^2_{S^*\mathbf{F}}$. Thus, the larger the part of the variance in S^* explained by the family background variables the higher is the probability that the omitted variable bias increases, compared to when family background variables are disregarded. This tendency may be balanced by the sum within brackets but, in general, it is impossible to say anything about the relative weights of these opposing forces.

To illustrate how difficult it is to say anything *a priori* about how the omitted variable bias is affected in the general case, it is instructive to consider the case $K = 2$. This is done in Example 1, below. The example also enables a simple, albeit non-stringent, demonstration of the equivalence between Proposition 1 and Proposition 5 when $K = 1$.

Example 1 (The omitted variable bias for K=2) *By means of standard results, the probability limits of the coefficients in the regression of S^* on F_1 and F_2 can be expressed as*

$$plim\left(\hat{\alpha}_1\right) = \frac{\left(\rho_{S^*F_1} - \rho_{F_1F_2} \cdot \rho_{S^*F_2}\right)}{\left(1 - \rho^2_{F_1F_2}\right)} \frac{\sqrt{Var\left(S^*\right)}}{\sqrt{Var\left(F_1\right)}}$$

and

$$plim\left(\hat{\alpha_2}\right) = \frac{\left(\rho_{S^*F_2} - \rho_{F_1F_2} \cdot \rho_{S^*F_1}\right)}{\left(1 - \rho^2_{F_1F_2}\right)} \frac{\sqrt{Var\left(S^*\right)}}{\sqrt{Var\left(F_2\right)}}.$$

By Proposition 5, the omitted variable bias thus equals

$$\beta_a \hat{\beta}_{AS} \frac{(1-\lambda)}{1 - R^2_{S^*F}} \left\{ 1 - \left[\frac{\rho_{AF_1}\left(\rho_{S^*F_1} - \rho_{F_1F_2} \cdot \rho_{S^*F_2}\right) + \rho_{AF_2}\left(\rho_{S^*F_2} - \rho_{F_1F_2} \cdot \rho_{S^*F_1}\right)}{\rho_{AS^*} \cdot \left(1 - \rho^2_{F_1F_2}\right)} \right] \right\}.$$

Concentrating on the ratio within brackets, we see that the denominator is unambiguously positive, as $\rho_{AS^*} \in]0,1[$, by assumption. A necessary, but not sufficient, requirement to create a downward pressure on the omitted variable bias is thus that the numerator is positive, too. However, the sign of the numerator depends to a large extent on the signs and relative magnitudes of ρ_{AF_1} and ρ_{AF_2}, both of which are unknown and can lie anywhere in the closed interval $[-1, 1]$.

Example 1 enables us to check that Proposition 1 and Proposition 5 yield the same results for $K = 1$. To this end, set $\rho_{F_1F_2} = \rho_{S^*F_2} = \rho_{AF_2} = 0$ and remember that, for $K = 1$, $R^2_{S^*F}$ is equal to $\rho^2_{S^*F_1}$. This reduces the expression for the omitted variable bias in Example 1 to

$$\beta_a \hat{\beta}_{AS}(1-\lambda) \times \frac{1}{1 - \rho^2_{S^*F_1}} \left(1 - \frac{\rho_{AF_1} \cdot \rho_{S^*F_1}}{\rho_{AS^*}}\right).$$

By means of Table 3 it can be seen that numerical evaluation of the factor after "×" yields the same result as evaluation of $\left(1 - \theta \cdot \rho^2_{AF \cdot S^*}\right)$ which demonstrates that the two propositions are equivalent.

4 Estimation of the omitted variable bias and the measurement error bias

The analysis in this section corresponds to the following thought experiment: Assume that, initially, the econometrician does not have access to information about ability and, therefore, tries to proxy ability by means of family background variables. At a later stage, (s)he gets access to a measure of ability and wants to use this information to estimate the omitted variable and the measurement error biases associated with the initial estimates. We take the initial stage as given here, i.e. we assume that we have results from regressions of Y on S^* and of Y on S^* plus family background variables.

¿From (4) and Proposition 1 we know that two unknown parameters involved in the omitted variable bias and the measurement error bias are β_s and β_a. Given data on true schooling, S, and ability, A, these parameters could be estimated by application of OLS to (1). However, while we know A we lack information about S; what we have is S^*. Accordingly, we can run a regression which is very close to (1), namely:

$$Y_i = \beta_{0^*} + \beta_{s^*} S_i^* + \beta_{a^*} A_i + u_i^*, \qquad \text{where} \qquad \beta_{s^*}, \beta_{a^*} > 0. \tag{12}$$

The only difference between (12) and (1) is that in (12) observed schooling, S_i^*, replaces true schooling, S_i. This substitution will change the model's parameters and its stochastic disturbance, compared to (1); to reflect this we have attached a * to the parameters and the disturbance.

Intuitively, it seems likely that OLS estimates of the parameters in (12) can provide information about the parameters in (1) and, thus, be useful in the construction of estimates of the omitted variable bias and the measurement error bias. Proposition 6 and Corollary 7 support this intuition.

Proposition 6 *The OLS estimates of the parameters β_{s^*} and β_{a^*} in (12) have the following probability limits*

$$\begin{aligned} plim\ \widehat{\beta}_{s^*} &= \beta_s - \beta_s \frac{\lambda}{1-\rho^2_{AS*}} \\ plim \widehat{\beta}_{a^*} &= \beta_a + \beta_s \frac{\lambda \cdot \widehat{\beta}_{S^*A}}{1-\rho^2_{AS*}} \end{aligned} \tag{13}$$

where

$$\widehat{\beta}_{S^*A} = \frac{Cov\,(A, S^*)}{Var\,(A)} \tag{14}$$

is the coefficient from a regression of observed schooling on ability.

Proof. See Appendix.

Corollary 7 *Conditional on a consistent estimate of λ, denoted $\lambda^{\dagger}$, consistent estimates of β_s and β_a in (1) can be constructed as follows*

$$\begin{aligned}
\widetilde{\beta}_{s|\lambda^{\dagger}} &= \widehat{\beta}_{s^*} \cdot \left(1 - \frac{\lambda^{\dagger}}{1-\rho^2_{AS*}}\right)^{-1} \\
\widetilde{\beta}_{a|\lambda^{\dagger}} &= \widehat{\beta}_{a^*} - \widetilde{\beta}_{s|\lambda^{\dagger}} \cdot \frac{\lambda^{\dagger} \cdot \widehat{\beta}_{S^*A}}{1-\rho^2_{AS*}}
\end{aligned}$$

Proof. Implied by Proposition 6 and the properties of the plim operator.

By means of (4), Proposition 5 and Corollary 7 we can construct consistent estimates of the omitted variable bias conditional on $\lambda^{\dagger}$, without and with family background variables, according to:

$$\begin{aligned}
\widehat{OVB}_{S^*|\lambda^{\dagger}} &= \widetilde{\beta}_{a|\lambda^{\dagger}} \widehat{\beta}_{AS|\lambda^{\dagger}} \left(1-\lambda^{\dagger}\right) \\
\widehat{OVB}_{S^*,\mathbf{F}|\lambda^{\dagger}} &= \widetilde{\beta}_{a|\lambda^{\dagger}} \widehat{\beta}_{AS|\lambda^{\dagger}} \left(1-\lambda^{\dagger}\right) \\
&\quad \times \tfrac{1}{1-R^2_{S^*\mathbf{F}}} \left[1 - \sum_{j=1}^{K} \tfrac{\rho_{AF_j}}{\rho_{AS^*}} \tfrac{\sqrt{Var(F_j)}}{\sqrt{Var(S^*)}} \left(\widehat{\alpha}_j\right)\right].
\end{aligned} \tag{15}$$

Note that in the last row of (15) the consistent estimate $\widehat{a}_j$ has been substituted for the corresponding probability limit in Proposition 5. This is OK because consistency of $\widehat{a}_j$ is sufficient to ascertain consistency of $\widehat{OVB}_{S^*,\mathbf{F}|\lambda^{\dagger}}$, *ceteris paribus.* It still remains, however, to determine the parameter $\widehat{\beta}_{AS|\lambda^{\dagger}}$. While the original parameter $\widehat{\beta}_{AS}$ in (6) cannot be computed unless true schooling, S, is known, this is, fortunately, not the case when we condition on $\lambda^{\dagger}$. We have:

$$\begin{aligned}
\widehat{\beta}_{AS|\lambda^{\dagger}} &\equiv \left.\tfrac{Cov(A,S)}{Var(S)}\right|_{\lambda^{\dagger}} &&= \left.\tfrac{Cov(A,S^*)}{Var(S^*)-\lambda^{\dagger}\cdot Var(S^*)}\right|_{\lambda^{\dagger}} \\
&= \left.\tfrac{Cov(A,S^*)}{Var(S^*)\left(1-\lambda^{\dagger}\right)}\right|_{\lambda^{\dagger}} &&= \left.\rho_{AS^*} \tfrac{\sqrt{Var(A)}}{\sqrt{Var(S^*)}} \tfrac{1}{\left(1-\lambda^{\dagger}\right)}\right|_{\lambda^{\dagger}}
\end{aligned} \tag{16}$$

where the first equality follows from (2) and (3) and the third equality follow directly from the definition of the coefficient of correlation.

Finally, we get the following consistent estimates of the measurement error bias before and after the inclusion of family background variables, again combining Corollary 7 with (4) and Proposition 5:

$$\begin{aligned}
\widehat{MEB}_{S^*|\lambda^{\dagger}} &= -\widetilde{\beta}_{s|\lambda^{\dagger}} \lambda^{\dagger} \\
\widehat{MEB}_{S^*,\mathbf{F}|\lambda^{\dagger}} &= -\widetilde{\beta}_{s|\lambda^{\dagger}} \frac{\lambda^{\dagger}}{1-R^2_{S^*\mathbf{F}}}.
\end{aligned} \tag{17}$$

5 Data and variable specifications

Our data set stems from a unique longitudinal survey, initiated in 1938 in the city of Malmö, in the south of Sweden. For the purpose of disentangling the effects of cognitive ability and social background on student achievement, a Swedish sociologist conducted a survey involving all of the city's third grade pupils. As the common school starting age at the time was seven years of age, the children were generally in their tenth year when they were interviewed in the spring of 1938. Altogether, 1 542 children were surveyed. These individuals have been followed and recurrently interviewed until 1993, when they were 65 years old, the Swedish retirement age. Presumably, this makes the Malmö survey one of the longest panels in the world.[7]

The variables that we use in this paper are defined in Table 1.

Table 1: Variable definitions

Y	=	ln of gross income in 1000s of SEK, 1968
S^*	=	observed years of schooling, 1964
A	=	IQ, measured at age 10, i.e. 1938
F_1	=	Father's education, in years, 1938
F_2	=	Family income in SEK, 1937
F_3	=	Family size in 1938

With respect to schooling, the original data contain information about the type and level of education, and, in many cases, about whether the respondent completed the education or not; altogether 42 alternatives are specified. Because of the large number of detailed alternatives, the likelihood that individuals have been wrongly categorized is small. A measurement error arises, however, when we transform the categorical data to number of years of schooling. This transformation is based on information about the stipulated number of years for each level and type – combination (compared to the corresponding next lower level).

Specifically, individuals with completed educations have been assigned the stipulated number of years. To individuals reporting incomplete educations we have assigned the stipulated years minus 1. As the difference in stipulated years between two subsequent educational levels is at least two years, this means that the number of years we assign to reported incomplete educations correspond to at least half of the stipulated completion times. By doing so, we (strongly) increase the probability that the continuous measurement error that we generate is symmetric around zero. To see this, consider the following argument.

First, note that educations for which there is reported information on "completed" / "non-completed" the combinations of true and reported status give rise to four possibilities: reported completed & truly completed (A), reported completed & truly non-completed (B), reported non-completed & truly completed (C), and reported non-completed & truly non-completed (D). For A our procedure yields a measurement error equal to zero, corresponding to a degenerated symmetric distribution centered

[7]For information about the Malmö survey, see Fägerlind (1975) and Furu (2000).

on zero. For D we know that all individuals have gone some way towards completing the education. In the absence of information about the distribution of the true years of schooling this means that the best estimate of the expected years must not be less than half of the stipulated number of years. Assigning incomplete educations an expected duration equal to the number of stipulated years - 1 fulfills this requirement since the number of stipulated years is mostly > 2 (and never < 2). With this choice of expected years we can hope to obtain measurement errors that are both positive and negative and hopefully symmetrical around zero. And for the B (C) individuals the choice will mean that the invariably positive (negative) measurement error will be smaller (in absolute terms) than it would have been if the expected duration had been set to the number of stipulated years $\div$ 2.

Regarding the ability measure we follow the preferred practice in the literature and use IQ test results.[8] The test employed contained four themes: "Word opposites", "Sentence completion", "Perception of identical figures", and "Disarranged sentences".[9]

Regarding the age at which the test has been conducted, it has been argued by several authors that it is important to measure IQ at a young age to avoid the ability measure from being contaminated by schooling.[10] On the other hand, constructing reliable tests for very young children is difficult. Age 10 strikes a balance between these opposing considerations. Moreover, to the extent that education does influence test results we want the influence to be similar across individuals. This, also, is an argument for conducting the test at age 10 because, at the time, education was comprehensive only up to the fourth grade.

The data include extensive information about family background. We have chosen information about the father's education, family income, and family size for two reasons. The first reason is that these variables are standard choices in the literature. Second, they are measured very closely in time to the IQ test. This should be an advantage in view of the fact that we want to use the family background variables as proxies for ability/IQ.

Information about (actual) work experience is missing in our data. However, since we are considering a cohort of male individuals, differences in work experince will, for given schooling, be determined by unemployment spells only. As the period between WWII and 1968 was characterized by almost full employment in Sweden, neglecting differences in unemployment across individuals is likely to be a matter of minor importance.[11]

[8]In principle, it could conceivably be argued that there might a measurement error in the IQ test results, just as in the schooling variable. However, since the concept of ability defies a strict definition, ability has to be made operational on the basis of what can be measured. Given that, and the fact that there is agreement upon IQ test results being the best available measure, it does not make sense to talk about IQ (test results) measuring ability subject to a measurement error.

[9]A pilot test had been carried out on third-graders in municipalities in the south of Sweden in the year preceeding the survey. Moreover, when the real test was conducted great effort was taken to ascertain that the test conditions were the same for all students.

[10]See, e.g., Hansen et al. (2004), and the references contained therein.

[11]Regarding military service, it was compulsory in Sweden at the time. The lenght of duty varied

Descriptive statistics are provided in Table 2.

Table 2: Descriptive statistics for the males in the Malmö study

	Sub-sample used (N=555)		Original Sample (N=834)	
	Mean	Std.dev.	Mean	Std.dev.
Y	10.29	0.664	10.28	0.652
S^*	8.867	2.571	8.853	2.554
A	98.26	15.75	97.73	16.02
F_1	7.690	1.492	7.728	1.575
F_2	3653	5780	3956	6861
F_3	4.500	1.557	4.500	1.579

It can be seen in the table that the average length of schooling was close to nine years for the 1928 cohort that we study. Since the school-starting age was 7 this means that, on average, the individuals finished their studies in 1944. The average time span between school end and the point in time when income is measured, 1968, is thus 24 years. With this extensive time lag it is natural to assume Schooling to be a predetermined variable. This assumption is supported by Sandgren (2005) who, utilizing a very similar data set from the Malmö survey, cannot reject the hypothesis that schooling is exogenous in the earnings regression.

Due to attrition and incomplete data for some individuals, the sample we use comprises 2/3 of the original sample. According Table 2, the loss of observations has barely affected means and standard deviations, however.

Coefficients of correlations are provided in Table 3.

Table 3: Coefficients of correlations

	S^*	A	F_1	F_2	F_3
A	0.4894				
F_1	0.4443	0.2177			
F_2	0.2767	0.1663	0.3605		
F_3	−0.1784	−0.1710	−0.0743	−0.0111	
Y	0.4489	0.3386	0.1785	0.1722	−0.0609

Table 3 can be used to compute examples of a partial correlation which plays a crucial role in Section 2.1, the partial correlation between schooling and family background, controlling for ability; cf. Corollary 4. For simplicity, we here only consider one background variable at a time, cf Table 4.

Table 4: Partial correlations between schooling and the family backgrund variables (F_i), controlling for ability, $\rho_{S^*F_i \cdot A}$

F_i	$(\rho_{S^*F_i},\ \rho_{AF_i},\ \rho_{AS^*})$	$\rho_{S^*F_i \cdot A}$
F_1	(0.4443, 0.2177, 0.4894)	0.3968
F_2	(0.2767, 0.1663, 0.4894)	0.2271
F_3	(−0.1784, −0.1710, 0.4894)	−0.1102

Note: For the computation of $\rho_{S^*F_i \cdot A}$ see (9)

somewhat across individuals, but not much.

These examples clearly show that our data do not satisfy $\rho_{S^*F \cdot A} = 0$, which would ascertain a reduction in omitted variable bias when a family background variable is used as proxy for ability in the earnings equation.

6 Empirical analysis

In this section we apply the results derived in Section 4. As all our empirical results are conditional on the unknown noise-to-signal ratio, we begin by considering the relevant range for this parameter. We then compute the estimates of the return to schooling and ability. Finally, we estimate the omitted variable bias, the measurement error bias and the total bias when earnings are regressed on i) observed schooling only or ii) observed schooling and various constellations of our three family background variables.

6.1 The noise-to-signal ratio

An estimate of the noise-to-signal ratio, i.e. λ, in Swedish data is provided by Isacsson (1999). He estimates λ to be 0.12 for imputed years of schooling, subject to a continuous (classical) mesurement error of the type (2).

For our purposes, it is not important to have a precise estimate of λ. Since we can estimate the omitted variable bias (OVB) and the measurement error (MEB) bias associated with any λ we only need some idea about its range. Thus, it doesn't matter that the method we have used to impute years of schooling differs somewhat from the method applied by Isacsson (1999).

We report estimates of OVB and MEB for three values on λ: 0.08, 0.13 and 0.18.[12] For the total bias, which is of primary interest, we report estimates when λ varies continuously between 0 and 0.20.

6.2 Estimates of the return to schooling and ability

The result of regressing ln(earnings) on schooling, subject to measurement error, and on ablity is reported in Table 5.

Table 5: OLS estimates of the parameters in eq. (12)

	Estimate	Std. error	t value
β_{0^*}	8.7910	0.1580	55.65
β_{s^*}	0.0961	0.0111	8.64
β_{a^*}	0.0066	0.0018	3.63
		$N = 555$	$R^2 = 0.22$

The estimated return to schooling is 0.096. For comparison, Björklund and Kjellström (2002) obtained a return estimate of 0.087 in an earnings equation for Swedish males,

[12] That we have chosen $\lambda = 0.13$ as our middle estimate, rather than Isacsson's 0.12 estimate, is due expository reasons.

based on cross-section data for 1968, and using as regressors years of schooling and years of work experience.[13]

To construct consistent estimates of β_s and β_a from the estimates in Table 5, we first use Tables 2 and 3 to obtain $\rho^2_{AS^*} = 0.2395$ and $\widehat{\beta}_{S^*A} = \rho_{AS^*}\left(\sqrt{Var\,(S^*)}/\sqrt{Var\,(A)}\right) = 0.0799$. Next, Corollary 7 yields:

Table 6: Estimates of β_s and β_a, conditional on $\lambda^\dagger$

	$\lambda^\dagger = 0.08$	$\lambda^\dagger = 0.13$	$\lambda^\dagger = 0.18$
$\widetilde{\beta}_{s\|\lambda^\dagger}$	0.10740	0.11592	0.12590
$\widetilde{\beta}_{a\|\lambda^\dagger}$	0.00568	0.00501	0.00420

Comparing Tables 5 and 6, we see that, as expected, the OLS estimate of the return to schooling is biased downwards, due to the measurement error in schooling. Given that the return estimate is biased downwards, the OLS coefficient for ability must be biased upwards. just as we find it to be.

6.3 Conditional estimates of omitted variable bias, measurement error bias, and total bias

Table 7 provides estimates of the omitted variable bias (OVB) and the measurement error bias (MEB), corresponding to (15) and (17), respectively. With respect to the OVB, the most important thing to note is that for a given noise-to-signal ratio inclusion of family background variables has very small effect on the omitted variable bias. Moreover, the bias is not monotonically decreasing in the number of family background variables included. Of course, this is not surprising, given our theoretical results. With increasing measurement error in schooling the bias decreases somewhat, but not very much. For example, more than doubling $\lambda^\dagger$ (going from 0.08 to 0.18) we obtain a decrease of less than 0.5 percentage points, or slightly more than a quarter of the initial bias (at $\lambda^\dagger = 0.08$).

Table 7: Estimates of omitted variable bias (OVB) and measurement error bias (MEB), conditional on $\lambda^\dagger$, for different sets of family background variables proxying for ability

	$\lambda^\dagger = 0.08$		$\lambda^\dagger = 0.13$		$\lambda^\dagger = 0.18$	
	OVB	MEB	OVB	MEB	OVB	MEB
No F_i's	.0170	-.0086	.0150	-.0151	.0126	-.0227
F_1 = father's educ.	.0170	-.0107	.0150	-.0188	.0126	-.0282
F_2 = family income	.0167	-.0093	.0147	-.0163	.0124	-.0245
F_3 = family size	.0165	-.0089	.0145	-.0156	.0122	-.0234
F_1 & F_2	.0168	-.0109	.0148	-.0191	.0125	-.0288
F_1 & F_3	.0165	-.0110	.0145	-.0193	.0122	-.0290
F_2 & F_3	.0161	-.0096	.0142	-.0169	.0119	-.0254
F_1 & F_2 & F_3	.0163	-.0112	.0143	-.0197	.0120	-.0296

Note: The estimates have computed using (15) and (17)

[13] The difference between the two estimates is not statistically significant.

Turning to the estimates of the measurement error bias (MEB), we see that these are larger (in absolute value) than the omitted variable bias estimates, except when the measurement error is small ($\lambda^{\dagger} = 0.08$). The inclusion of family background variables induces sizeable increases in the MEB. For instance, when $\lambda^{\dagger} = 0.13$ and all three of the family background variables are included the (absolute value of the) MEB increases by almost 0.5 percentage points compared to when no background variables are included. The OVB is virtually unaffected by the same operation. And, in contrast to the OVB, the MEB is monotonically related to the number of included family background variables. The MEB is also much more sensitive to $\lambda^{\dagger}$ than is the OVB. Going from $\lambda^{\dagger} = 0.08$ to $\lambda^{\dagger} = 0.18$ increases the MEB by more than 250 percent, i.e. the relative change in the bias is larger than the relative change in the noise-to-signal ratio.

The total bias, TB, is equal to OVB + MEB. Since the MEB dominates the OVB, the TB has the same general properties as the MEB. In particular, inclusion of additional family background variables will always make the TB fall. Referring back to very beginning of this paper, this means that Lam and Schoeni (1993) were right in claiming that the estimated return to schooling can always be made to decrease, by inclusion of family background variables.

A decrease in the TB can be a good thing if the TB is positive to begin with. It's a bad thing if the decrease occurs when the TB is negative already; in this case the change implies a move further away from the true return to schooling. This case is illustrated in Table 7 for $\lambda^{\dagger} = 0.13$. When no family background variables are included the TB is negative, but very small – virtually zero – because the OVB and the MEB cancel out. Adding family background variables merely has the effect of making a very good return estimate increasingly more biased.

To see when the use of family background variables turns from a useful to a wasteful practice we need to consider the total bias as a continuous function of $\lambda^{\dagger}$. This is done in Figure 1, which shows the estimated TB, in percent of $\widetilde{\beta}_{s|\lambda^{\dagger}}$, in two cases. In the first case, no family background variables are included in the regression and in the second case F_1, F_2, and F_3 are included. The figure also shows the difference between the absolute values of the relative TBs in the two cases. The vertical dotted line delimits the areas where the inclusion of the family background variables decreases the TB – to the left of the line – and where the bias increases – to the right.

Figure 1: Total bias, in % of $\widetilde{\beta}_{s|\lambda}$, as a function of λ, in the absence of family background variables, $TB(\lambda)|_{S^*}$, and when F_1, F_2, and F_3 are included, $TB(\lambda)|_{S^*,\mathbf{F}}$

about here

First, note how little the inclusion of the background variables decreases the relative TB when $\lambda = 0$. The reduction equals $20.5 - 19.6 = 0.9$ percentage points of the return estimate. With λ strictly positive, the family background variables at first yield an increasing advantage in terms of relative TB. But this advantage is present only when $\lambda < 0.117$. Moreover, it is quite small; only when λ varies between 0.04 and 0.11 is the decrease at least 2 percent of the corresponding return estimate. And the gain is largest when the bias is small to begin with. The maximum reduction occurs when the bias without family background variables is less than 4 percent of the estimated true return (at $\lambda \approx 0.105$). If λ grows beyond 0.12 the relative TB rapidly becomes much larger with family background variables than without. For instance, when $\lambda = 0.18$ the relative TB is 6 percentage points larger with family background variables included, than when they are left out.

Qualitatively, the results in Figure 1 are the same if, instead, only one or two family background variables are included. The quantitative difference is that the gains/losses in the total bias are smaller (in absolute terms) if the number of background variables is reduced.

7 Concluding comments

This study was inspired by a remarkable claim, implied by an analysis in Lam and Schoeni (1993): using family background variables as proxies for unobserved ability in earnings regressions you can drive the estimated return to schooling down to arbitrarily low levels. A compact way to summarize our findings is that Lam and Schoeni were right – but for the wrong reasons.

We have shown that Lam and Schoeni's assertion that inclusion of family background variables will reduce the positive omitted variable bias (OVB) and increase the negative measurement error bias (MEB) is partly incorrect. The OVB may increase, as well as decrease. We also demonstrate, though, that in the context of a single family background variable the OVB will indeed decrease for sure if the correlation between schooling and family background is identically zero when one controls for ability. But this a restrictive assumption; in our empirics we find that when ability is controlled for, the correlation between schooling and family background is far from zero.

Unlike Lam and Schoeni, we also conduct a theoretical analysis of the case with several family background variables. We show that the MEB result in the one variable case can be extended to the K - variable case. The indeterminacy of the OVB, on the other hand, becomes even larger; the conditions which ascertain that the OVB is reduced in the one variable case can not be extended to the case when $K \geq 2$.

Theoretically, the reduction in the estimated return induced by increased MEB can thus be counteracted by increases in OVB. The effect on the total bias, i.e. MEB + OVB, hence becomes an empirical matter.

For the empirical analysis, we derive OVB and MEB estimates that are consistent, conditional on the ratio of measurement error variance to total variance in observed schooling (λ). The estimators are applied to a unique Swedish data set that is extremely well suited to our analysis.

Our empirical results yield three conclusions. First, to the extent that inclusion of family background variables leads to increases in OVB, these increases are very small. Secondly, *all* changes in the OVB are very small, irrespective of whether they are positive or negative. Thirdly, except for small values on λ – below 0.13 in our application – the OVB is dominated by the MEB, implying a negative total bias.

Furthermore, the MEB is much more sensitive than the OVB to changes in the number of family background variables and in λ. And additional family background variables and increases in λ both monotonically increase the magnitude of the MEB.

Proxying ability by family background is thus generally a bad idea because the bias that one wants to influence – the OVB – is barely affected, while the side-effect – an increase in the MEB – is substantial and generally makes the total bias larger than in the absence of the background variables. For example, if $\lambda = 0.13$ the total bias when family background variables are excluded is in our data estimated to be virtually zero; -0.08 percent of the estimated true return. Adding three family background variables increases the relative bias to -4.6 percent. Moreover, when the inclusion of family background variables indeed does reduce the total bias, the reduction is largest when the bias is small to begin with, i.e. when a reduction is not very important.

It might be argued that our analysis is too stylized to provide useful insights about how omitted variables and measurement error affect estimates of the rate of return to education. We don't think so. Essentially, there are two potential problems: endogeneity of schooling and the lack of variables beside schooling and ability.

Regarding the lack of variables beside schooling and ability, we have shown how an arbitrary number of control variables can be accounted for; cf. Section II.

Concerning the endogeneity problem, assume that schooling, S, is endogenous. This means that S is a stochastic variable that is correlated with the stochastic disturbance in the wage equation. The standard remedy to this problem is to find an instrument for S. Such an instrument is another stochastic variable that is correlated with S but not with the disturbance term. Our S^* variable can be interpreted in precisely this way. Accordingly, our analysis can either be interpreted as concerning the case when the measure of schooling is predetermined *or* as taking place after an instrument has been found for the endogenous schooling variable.

8 References

Björklund, A. and C. Kjellström (2002), "Estimating the return to investments in education: how useful is the standard Mincer equation?", *Economics of Education Review*, Vol. 21, pp. 195-210.

Card, D. (1999), "The causal effect of education on earnings", in *Handbook of Labor Economics*, Vol. 3, O. Ashenfelter & D. Card (eds), Elsevier, Amsterdam.

Carroll, R., J. Ruppert, and L. Stefanski (1995), *Measurement Error in Nonlinear Models*, Chapman and Hall, London.

Frost, P. (1979), "Proxy Variables and Specification Bias", *Review of Economics and Statistics*, Vol. 61, pp. 323-325.

Furu, M. (2000), "The Malmö Study", in *Seven Swedish Longitudinal Studies*, Forskningsrådsnämnden (the Swedish Research Council), Stockholm

Fägerlind, I. (1975), *Formal Education and Adult Earnings: A Longitudinal Study on the Economic Benefits of Education*, Almqvist & Wicksell International, Stockholm.

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

Hansen, K.T, J.J. Heckman, and K.J. Mullen (2004), "The Effect of Schooling and Ability on Achievement Test Scores", *Journal of Econometrics*, Vol. 121, pp. 39-98.

Isacsson, G. (1999), "Estimates of the return to schooling in Sweden from a large sample of twins", *Labour Economics*, Vol. 6, pp. 471-489.

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.

Maddala, G.S., 1977, *Econometrics*, McGraw-Hill, New York.

McCallum, B.T. (1972), "Relative Asymptotic Bias from Errors of Omission and Measurement", *Econometrica*, Vol. 40, pp. 757-758.

Sandgren, S. (2005), "Earnings of Swedish young men", in Sandgren (2005): "Learning and earning: Studies on a cohort of Swedish men", Doctoral thesis, Royal Institute of Technology, Dept. of Infrastructure, Stockholm, Sweden.

Welch, F. (1975), “Human Capital Theory: Education, Discrimination, and Life Cycles”, *American Economic Review Papers and Proceedings*, Vol. 65, pp. 63-73.

Wickens, M.R. (1972), ”A Note on the Use of Proxy Variables, *Econometrica*, Vol. 40, pp. 759-761.

A Appendix: Proof of Propositions

Proof of Proposition 1. Denote mean sums of squares and cross-products according to

$$q_{s*s*} = \frac{1}{N}\sum_{i=1}^{N}\left(S_i^* - \bar{S}^*\right)^2 \qquad \text{and} \qquad q_{s*f} = \frac{1}{N}\sum_{i=1}^{N}\left(S_i^* - \bar{S}^*\right)\left(F_i - \bar{F}\right).$$

Then, by standard results for regressions involving two regressors and in accordance with well-known properties of probability limits,

$$plim\hat{\beta}_{S.F} = \frac{plim\left(q_{ff}\right)plim\left(q_{s*y}\right) - plim\left(q_{s*f}\right)plim\left(q_{fy}\right)}{plim\left(q_{s*s*}\right)plim\left(q_{ff}\right) - \left[plim\left(q_{s*f}\right)\right]^2}$$

To evaluate this expression, first note that,

$$plim\left(q_{ff}\right) = Var\left(F\right). \tag{18}$$

Next, by (2) and (3)

$$plim\left(q_{s*s*}\right) = Var\left(S^*\right) = Var\left(S\right) + var\left(w\right). \tag{19}$$

$$plim\left(q_{s*f}\right) = Cov\left(S^*, F\right) = Cov\left(S, F\right) \tag{20}$$

Further, by (1) and (3),

$$plim\left(q_{fy}\right) = \beta_s Cov\left(S, F\right) + \beta_a Cov\left(A, F\right) = \beta_s Cov\left(S^*, F\right) + \beta_a\left(A, F\right), \tag{21}$$

where the last equality follows from (20). Finally, by (1) – (3),

$$plim\left(q_{s*y}\right) = \beta_s Var\left(S\right) + \beta_a Cov\left(A, S^*\right) = \beta_s Var\left(S\right) + \beta_a Cov\left(A, S\right). \tag{22}$$

Collecting results and rearranging we get

$$\begin{aligned} plim\hat{\beta}_{S.F} \quad = \quad & \frac{\beta_s\left\{Var(S)Var(F) - [Cov(S^*,F)]^2\right\}}{Var(S^*)Var(F) - [Cov(S^*,F)]^2} \\ & + \frac{\beta_a[Var(F)Cov(A,S) - Cov(S^*,F)Cov(A,F)]}{Var(S^*)Var(F) - [Cov(S^*,F)]^2} \end{aligned} \tag{23}$$

To simplify the first term in (23), first substitute $[Var\left(S^*\right) - Var\left(w\right)]$ for $Var\left(S\right)$, then divide the numerator and the denominator by $[Var\left(S^*\right)Var\left(F\right)]$, and, finally, use (5). This yields

$$\frac{\beta_s\{Var\left(S\right)Var\left(F\right) - \left[Cov\left(S^*, F\right)\right]^2\}}{Var\left(S^*\right)Var\left(F\right) - \left[Cov\left(S^*, F\right)\right]^2} = \beta_s - \beta_s\frac{\lambda}{1 - \rho^2_{S^*F}}. \tag{24}$$

To rewrite the second term in (23) first divide the numerator and the denominator by $[Var\,(S^*)\,Var\,(F)]$ and note that, by (2)

$$\frac{Cov\,(A,S)}{Var\,(S^*)} = \hat{\beta}_{AS}\,(1-\lambda)\,, \tag{25}$$

where $\hat{\beta}_{AS}$ and λ are defined by (6) and (5), respectively. This yields

$$\frac{\beta_a[Var\,(F)\,Cov\,(A,S) - Cov\,(S^*,F)\,Cov\,(A,F)]}{Var\,(S^*)\,Var\,(F) - [Cov\,(S^*,F)]^2} = \beta_a\hat{\beta}_{AS}\,(1-\lambda)\cdot\zeta \tag{26}$$

where

$$\zeta = \frac{1-\rho^2_{S^*F}\frac{Cov(A,F)Var(S^*)}{Cov(S^*,F)Cov(A,S)}}{1-\rho^2_{S^*F}} \tag{27}$$

Using the equality $Cov\,(A,S) = Cov\,(A,S^*)$ and rearranging one can rewrite ζ according to

$$\zeta = 1 - \frac{\rho^2_{S^*F}\frac{\rho_{AF}-\rho_{S^*F}\cdot\rho_{AS^*}}{\rho_{S^*F}\cdot\rho_{AS^*}}}{1-\rho^2_{S^*F}}. \tag{28}$$

It remains to prove that the second term on the RHS of (28) is equal to the product $\theta\cdot\rho^2_{AF\cdot S^*}$. Multiplication of the numerator and the denominator by $\left(1-\rho^2_{AS^*}\right)$ yields

$$\begin{aligned}\frac{\rho^2_{S^*F}\frac{\rho_{AF}-\rho_{S^*F}\cdot\rho_{AS^*}}{\rho_{S^*F}\cdot\rho_{AS^*}}}{1-\rho^2_{S^*F}} &= \frac{[(\rho_{S^*F}/\rho_{AS^*}) - \rho_{S^*F}\cdot\rho_{AS^*}]\,(\rho_{AF}-\rho_{S^*F}\cdot\rho_{AS^*})}{\left(1-\rho^2_{S^*F}\right)\left(1-\rho^2_{AS^*}\right)} \\ &= \theta\cdot\rho^2_{AF\cdot S^*},\end{aligned} \tag{29}$$

where θ and $\rho_{AF\cdot S^*}$ are defined in Proposition 1. Now, substitute this equality in (28) and, subsequently, (28) in (26). The resulting expression and (24) can then be used in (23). Finally, to get (7), note that in the case with only one family background variable $\rho^2_{S^*F} = R^2_{S^*F}$. Q.E.D.

Proof of Proposition 5. Let the $(K+1)$ square matrix $\mathbf{Q_{xx}}$ be defined as

$$\mathbf{Q_{xx}} = \begin{pmatrix} q_{s^*s^*} & q_{s^*\mathbf{f}} \\ (1\times 1) & (1\times K) \\ \mathbf{q}_{\mathbf{f}s^*} & \mathbf{Q_{ff}} \\ (K\times 1) & (K\times K)\end{pmatrix}$$

where

$$q_{s^*s^*} = \frac{1}{N}\sum_{i=1}^{N}(S_i^* - \bar{S}^*)^2,$$

and the typical elements of the vector $\mathbf{q}_{\mathbf{f}s^*}(=\mathbf{q}'_{s^*\mathbf{f}})$ and the matrix $\mathbf{Q_{ff}}$ are

$$\mathbf{q}_{\mathbf{f}s^*} = (q_{f_js^*}) = \left[\frac{1}{N}\sum_{i=1}^{N}(F_{ij}-\bar{F}_j)(S_i^*-\bar{S}^*)\right],$$

and

$$\mathbf{Q_{ff}} = (q_{f_k f_j}) = \left[\frac{1}{N}\sum_{i=1}^{N}(F_{ik} - \bar{F}_k)(F_{ij} - \bar{F}_j)\right],$$

respectively. Similarly, denote by $\mathbf{q}_{\mathbf{x}_y}$ the $(K+1)\times 1$ vector whose first element is

$$q_{s^*y} = \frac{1}{N}\sum_{i=1}^{N}(S_i^* - \bar{S}^*)(Y_i - \bar{Y})$$

and whose following elements are

$$q_{f_j y} = \frac{1}{N}\sum_{i=1}^{N}(F_{ij} - \bar{F}_j)(Y_i - \bar{Y}), \quad j = 1, \dots K.$$

The OLS estimate of β_s is given by the first element of $(K+1)$ vector

$$\mathbf{Q}_{\mathbf{xx}}^{-1}\mathbf{q}_{\mathbf{x}_y} = \frac{1}{\det(\mathbf{Q_{xx}})}\mathrm{adj}(\mathbf{Q_{xx}})\mathbf{q}_{\mathbf{x}_y} \tag{30}$$

where

$$\mathrm{adj}(\mathbf{Q_{xx}}) = \begin{pmatrix} C_{s^*s^*} & C_{f_1 s^*} & \dots & C_{f_K s^*} \\ C_{f_1 s*} & C_{f_1 f_1} & \dots & C_{f_1 f_K} \\ \vdots & \vdots & \ddots & \vdots \\ C_{f_K s^*} & C_{f_1 f_K} & \dots & C_{f_K f_K} \end{pmatrix}$$

is the transpose of the matrix of cofactors of $\mathbf{Q_{xx}}$. Thus, $C_{s^*s^*}$ =det($\mathbf{Q_{ff}}$) and, e.g., $C_{f_3 s^*}$ is (-1) times the determinant of the matrix obtained by deleting the first column and the fourth row of $\mathbf{Q_{xx}}$. Accordingly,

$$plim\widehat{\beta}_{S\cdot\mathbf{F}} = \frac{plim(q_{s^*y})\cdot plim(C_{s^*s^*}) + \sum_{j=1}^{K} plim(q_{f_j y})\, plim(C_{f_j s^*})}{plim[det(\mathbf{Q_{xx}})]}$$

To simplify this expression, first use (22), (19), and (21) to get

$$\begin{aligned} plim\widehat{\beta}_{S\cdot\mathbf{F}} \quad &= \quad \beta_s \frac{[Var(S^*) - Var(w)]\, plim(C_{s^*s^*}) + \sum_{j=1}^{K} Cov(S^* F_j)\, plim(C_{f_j s^*})}{plim[\det(\mathbf{Q_{xx}})]} \\ &\quad + \beta_a \frac{Cov(A,S)\cdot plim(C_{s^*s^*}) + \sum_{j=1}^{K} Cov(A, F_j)\, plim(C_{f_j s^*})}{plim[\det(\mathbf{Q_{xx}})]} \end{aligned} \tag{31}$$

We now further simplify the two terms in (31) in turn. Concerning the first term, note that in accordance with the rules for Laplace expansions of determinants

$$Var(S^*) plim(C_{s^*s^*}) + \sum_{j=1}^{K} Cov(S^*, F_j) plim(C_{f_j s^*}) = plim[\det(\mathbf{Q_{xx}})] \tag{32}$$

Thus, by (32), (5), and (30)

$$\beta_s \frac{[Var(S^*)-Var(w)]plim(C_{s^*s^*})+\sum_{j=1}^{K} Cov(S^*,F_j)plim(C_{f_js^*})}{plim[\det(\mathbf{Q_{xx}})]} = \beta_s - \beta_s \lambda Var\left(S^*\right) plim\left(Q^{-1}_{s^*s^*}\right) \tag{33}$$

where $Q^{-1}_{s^*s^*}$ denotes the first element in the first row of $\mathbf{Q}^{-1}_{\mathbf{xx}}$, i.e.

$$Q^{-1}_{s^*s^*} = C_{s^*s^*}/\det(\mathbf{Q_{xx}}) = \det(\mathbf{Q_{ff}})/\det(\mathbf{Q_{xx}}) \tag{34}$$

It remains to show that $Var(S^*)plim(Q^{-1}_{s^*s^*}) = (1-R^2_{S^*\cdot\mathbf{F}})^{-1}$. Using (32) and the rules for the plim operator we get

$$Var(S^*)plim(Q^{-1}_{s^*s^*}) = \left[1 - \frac{\sum_{j=1}^{K} Cov(S^*,F_j)plim\left[\frac{-C_{f_js^*}}{\det(\mathbf{Q_{ff}})}\right]}{Var(S^*)}\right]^{-1} \tag{35}$$

As the (asymptotic) $R^2_{S^*\cdot\mathbf{F}}$ can be written

$$R^2_{S^*\cdot\mathbf{F}} = \frac{\sum_{k=1}^{K} Cov(S^*,F_j)plim\hat{\alpha_j}}{Var(S^*)} \tag{36}$$

where $\hat{\alpha_j}$ denotes the OLS estimate of the jth slope coefficient in the regression of S^* on F [cf. Maddala(1977, p. 107)], the final step amounts to demonstrating that $[-C_{f_js^*}/\det(\mathbf{Q_{ff}})] = \hat{\alpha_j}$. To this end, write the minor of the element $q_{s^*f_j}$ in $\mathbf{Q_{xx}}$ as $\det(\mathbf{M}_{s^*f_j})$ and denote by $\mathbf{\Psi}_j$ the matrix obtained by replacing the jth column of $\mathbf{Q_{ff}}$ by the column vector $\mathbf{q_{f}}_{s^*}$. Then

$$\begin{aligned} \frac{-C_{f_js^*}}{\det(\mathbf{Q_{ff}})} &= \frac{-[(-1)^{(j+1)+1}\det(\mathbf{M}_{s^*f_j})]}{\det(\mathbf{Q_{ff}})} \\ &= \frac{-[(-1)^{(j+1)+1}(-1)^{j-1}\det(\mathbf{\Psi}_j)]}{\det(\mathbf{Q_{ff}})} = \frac{\det(\mathbf{\Psi}_j)}{\det(\mathbf{Q_{ff}})} = \widehat{\alpha}_j \end{aligned} \tag{37}$$

The first equality follows directly from the definition of the cofactor $C_{f_js^*}$. The second equality is due to the fact that $\mathbf{\Psi}_j$ can be obtained by (*j-1*) interchanges of the columns in $\mathbf{M}_{s^*f_j}$, each of which results in the associated determinant being multiplied by (-1). The third equality follows because $(-1)^{2(j+1)} = 1\forall j$. The final equality is just an application of Cramer's rule to the system $\mathbf{Q_{ff}}\hat{\boldsymbol{\alpha}} = \mathbf{q_f}_{s^*}$. Substituting (37) in (35), using (36), and inserting the result in (33) we get the two first terms on the RHS in Proposition 5.

To rewrite the second term in (31) first use (25), (34), and the equality

$$Var(S^*)plim(Q^{-1}_{s^*s^*}) = (1-R^2_{S^*\mathbf{F}})^{-1}$$

implied by (35) – (37) to get

$$\begin{aligned}\beta_a \frac{Cov(A,S)\cdot plim(C_{s^*s^*})+\sum_{j=1}^{K} Cov(A,F_j) plim\left(C_{f_js^*}\right)}{plim[\det(\mathbf{Q_{xx}})]} &= \\ \beta_a \hat{\beta}_{AS} \frac{(1-\lambda)}{1-R^2_{S^*\mathbf{F}}}[1+\Phi]\end{aligned} \tag{38}$$

The variable Φ is given by

$$\begin{aligned}\Phi &= \sum_{j=1}^{K} \frac{Cov(A,F_j)}{Cov(A,S^*)} \frac{plim[C_{f_js^*}/\det(\mathbf{Q_{xx}})]}{plim\left(Q^{-1}_{s^*s^*}\right)} \\ &= -\sum_{j=1}^{K} \frac{Cov(A,F_j)}{Cov(A,S^*)} plim\left[\frac{-C_{f_js^*}}{\det(\mathbf{Q_{ff}})}\right]\end{aligned} \tag{39}$$

where $Cov(A,S) = Cov(A,S^*)$ has been used to obtain the first equality. To get the second equality, (34) has been employed and the sign of $C_{f_js^*}$ has been changed, whereupon the whole expression has been multiplied by -1. By (37), the term within brackets is equal to $\hat{\alpha_j}$. Finally, some straightforward manipulations yield

$$\frac{Cov(A,F_j)}{Cov(A,S^*)} = \frac{\rho_{AF_j}}{\rho_{AS^*}} \frac{\sqrt{Var(F_j)}}{\sqrt{Var(S^*)}} \tag{40}$$

Substituting (40) in (39) and inserting the result in (38) we get the last term on the RHS in Proposition 5. Q.E.D.

Proof of Proposition 6. The proof in the following can alternatively be replaced by a proof based on application og the results in Carroll et al. (1995, Ch. 2.2.3).

Using the same notation and the same arguments as in the proof of Proposition 1, we can write the OLS estimate of β_{S^*} as

$$plim\ \widehat{\beta}_{S^*} = \frac{plim\,(q_{aa})\ plim\,(q_{s^*y}) - plim\,(q_{s^*a})\ plim\,(q_{ay})}{plim\,(q_{s^*s^*})\ plim\,(q_{aa}) - [plim\,(q_{s^*a})]^2} \tag{41}$$

To evaluate (41), first note that by analogy with (18):

$$plim\,(q_{aa}) = Var\,(A)\ , \tag{42}$$

while $plim(q_{s^*y})$ is given by (22). Further, by analogy with (20) and (21), respectively,

$$plim\,(q_{s^*a}) = Cov\,(S^*, A) = Cov\,(S, A) \tag{43}$$

and

$$plim\,(q_{ay}) = \beta_s Cov\,(A, S) + \beta_a Var\,(A) = \beta_s Cov\,(A, S^*) + \beta_a Var\,(A)\,, \tag{44}$$

while $plim(q_{s^*s^*})$ is given (19). Using (42), (22), (43), (44), and (19) in (41), we obtain, after some simplifications

$$plim\ \widehat{\beta}_{S^*} = \beta_s \frac{Var\,(S)\,Var\,(A) - [Cov\,(S,A)]^2}{\left\{Var\,(S)\,Var\,(A) - [Cov\,(S,A)]^2\right\} + Var\,(w)\,Var\,(A)}\,. \tag{45}$$

Noting that $Var(S) = Var(S^*) - Var\,(w)$, and using (43) and (5), we can rewrite (45) according to

$$\begin{aligned} plim\ \widehat{\beta}_{S^*} &= \beta_s - \beta_s \tfrac{Var(w)Var(A)}{Var(S^*)Var(A)-[Cov(S^*,A)]^2} \\ &= \beta_s - \beta_s \frac{\lambda}{1 - \frac{[Cov(S^*,A)]^2}{Var(S^*)Var(A)}} \\ &= \beta_s - \beta_s \frac{\lambda}{1-\rho^2_{AS^*}} \end{aligned} \tag{46}$$

where the last equality follows from the definition of the coefficient of correlation between A and S^*. This proves the first part of Proposition 3.

By analogy with (41), the OLS estimate of β_{a^*} can be written

$$plim\widehat{\beta}_{a^*} = \frac{plim\,(q_{s^*s^*})\ plim\,(q_{ay}) - plim\,(q_{s^*a})\ plim\,(q_{s*y})}{plim\,(q_{s^*s^*})\ plim\,(q_{aa}) - [plim\,(q_{s^*a})]^2}\,. \tag{47}$$

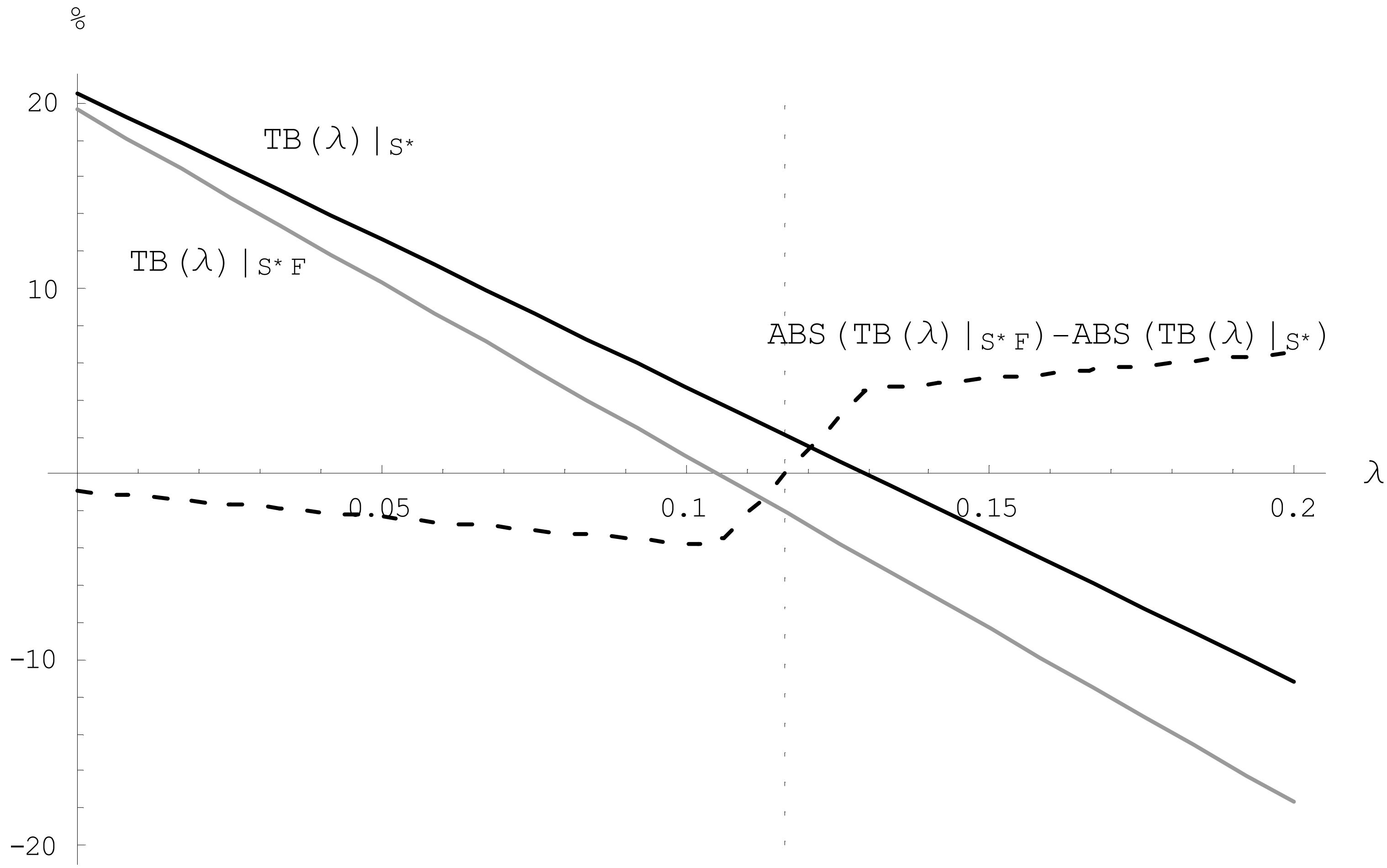
Using (19), (44), (43), (22), and (42), and simplifying, we get

$$plim\widehat{\beta}_{a^*} = \beta_a + \beta_s \frac{[Var\,(S^*)\,Cov\,(A,S^*) - Var\,(S)\,Cov\,(A,S^*)]}{Var\,(S^*)\,Var\,(A) - [Cov\,(S^*,A)]^2}\,. \tag{48}$$

Simplifying further by means of the equality $Var\,(S^*) = Var\,(S) + Var\,(w)$ and the definitions of ρ_{AS^*}, λ, and $\widehat{\beta}_{S^*A}$ we obtain

$$\begin{aligned} plim\widehat{\beta}_{a^*} &= \beta_a + \beta_s \tfrac{Var(w)Cov(A,S^*)}{Var(S^*)Var(A)-[Cov(S^*,A)]^2} \\ &= \beta_a + \beta_s \frac{\frac{Var(w)}{Var(S^*)} \cdot \frac{Cov(A,S^*)}{Var(A)}}{1-\rho^2_{AS^*}} \\ &= \beta_a + \beta_s \tfrac{\lambda \cdot \widehat{\beta}_{S^*A}}{1-\rho^2_{AS^*}}\,, \end{aligned} \tag{49}$$

which yields the last part of Proposition 6. Q.E:D.


%
20
10
-10
-20
0.05
0.1
0.15
0.2
λ
TB(λ)|S*
TB(λ)|S* F
ABS(TB(λ)|S* F)-ABS(TB(λ)|S*)

Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapporter/Reports

2008:1 de Luna Xavier, Anders Forslund and Linus Liljeberg ”Effekter av yrkesinriktad arbetsmarknadsutbildning för deltagare under perioden 2002–04”

2008:2 Johansson Per and Sophie Langenskiöld ”Ett alternativt program för äldre långtidsarbetslösa – utvärdering av Arbetstorget för erfarna”

2008:3 Hallberg Daniel ”Hur påverkar konjunktursvängningar förtida tjänstepensionering?”

2008:4 Dahlberg Matz and Eva Mörk ”Valår och den kommunala politiken”

2008:5 Engström Per, Patrik Hesselius, Bertil Holmlund and Patric Tirmén ”Hur fungerar arbetsförmedlingens anvisningar av lediga platser?”

2008:6 Nilsson J Peter “De långsiktiga konsekvenserna av alkoholkonsumtion under graviditeten”

2008:7 Alexius Annika and Bertil Holmlund ”Penningpolitiken och den svenska arbetslösheten”

2008:8 Anderzén Ingrid, Ingrid Demmelmaier, Ann-Sophie Hansson, Per Johansson, Erica Lindahl and Ulrika Winblad ”Samverkan i Resursteam: effekter på organisation, hälsa och sjukskrivning”

2008:9 Lundin Daniela och Linus Liljeberg ”Arbetsförmedlingens arbete med nystartsjobben”

2008:10 Hytti Helka och Laura Hartman ”Integration vs kompensation – välfärdsstrategier kring arbetsoförmåga i Sverige och Finland”

2008:11 Hesselius Patrik, Per Johansson och Johan Vikström ”Påverkas individen av omgivningens sjukfrånvaro?”

2008:12 Fredriksson Peter and Martin Söderström ”Vilken effekt har arbetslöshetsersättningen på regional arbetslöshet?”

2008:13 Lundin Martin ”Kommunerna och arbetsmarknadspolitiken”

2008:14 Dahlberg Matz, Heléne Lundqvist and Eva Mörk ”Hur fördelas ökade generella statsbidrag mellan personal i olika kommunala sektorer?”

2008:15 Hall Caroline ”Påverkades arbetslöshetstiden av sänkningen av de arbetslösas sjukpenning?”

2008:16 Bennmarker Helge, Erik Mellander and Björn Öckert ”Är sänkta arbetsgivaravgifter ett effektivt sätt att öka sysselsättningen?”

2008:17 Forslund Anders ”Den svenska jämviktsarbetslösheten – en översikt”

2008:18 Westregård Annamaria J. ”Arbetsgivarens ökade ansvar för sjuklön och rehabilitering kontra arbetstagarnas integritet – Går det att förena?”

2008:19 Svensson Lars ”Hemmens modernisering och svenska hushålls tidsanvändning 1920–90”

2008:20 Johansson Elly-Ann and Erica Lindahl ”Åldersintegrerade klasser – bra eller dåligt för elevernas studieresultat?”

Working Papers

2008:1 Albrecht James, Gerard van den Berg and Susan Vroman “The aggregate labor market effects of the Swedish knowledge lift programme”

2008:2 Hallberg Daniel “Economic fluctuations and retirement of older employees”

2008:3 Dahlberg Matz and Eva Mörk “Is there an election cycle in public employment? Separating time effects from election year effects”

2008:4 Nilsson J Peter ”Does a pint a day affect your child’s pay? The effect of prenatal alcohol exposure on adult outcomes”

2008:5 Alexius Annika and Bertil Holmlund “Monetary policy and Swedish unemployment fluctuations”

2008:6 Costa Dias Monica, Hidehiko Ichimura and Gerard van den Berg ”The matching method for treatment evaluation with selective participation and ineligibles”

2008:7 Richardson Katarina and Gerard J. van den Berg “Duration dependence versus unobserved heterogeneity in treatment effects: Swedish labor market training and the transition rate to employment”

2008:8 Hesselius Patrik, Per Johansson and Johan Vikström “Monitoring and norms in sickness insurance: empirical evidence from a natural experiment”

2008:9 Verho Jouko, “Scars of recession: the long-term costs of the Finnish economic crisis”

2008:10 Andersen Torben M. and Lars Haagen Pedersen “Distribution and labour market incentives in the welfare state – Danish experiences”

2008:11 Waldfogel Jane “Welfare reforms and child well-being in the US and UK”

2008:12 Brewer Mike “Welfare reform in the UK: 1997–2007”

2008:13 Moffitt Robert “Welfare reform: the US experience”

2008:14 Meyer Bruce D. “The US earned income tax credit, its effects, and possible reforms”

2008:15 Fredriksson Peter and Martin Söderström “Do unemployment benefits increase unemployment? New evidence on an old question”

2008:16 van den Berg Gerard J., Gabriele Doblhammer-Reiter and Kaare Christensen ”Being born under adverse economic conditions leads to a higher cardiovascular mortality rate later in life – evidence based on individuals born at different stages of the business cycle”

2008:17 Dahlberg Matz, Heléne Lundqvist and Eva Mörk “Intergovernmental grants and bureaucratic power”

2008:18 Hall Caroline “Do interactions between unemployment insurance and sickness insurance affect transitions to employment?”

2008:19 Bennmarker Helge, Erik Mellander and Björn Öckert “Do regional payroll tax reductions boost employment?”

2008:20 Svensson Lars “Technology, institutions and allocation of time in Swedish households 1920–1990”

2008:21 Johansson Elly-Ann and Erica Lindahl “The effects of mixed-age classes in Sweden”

2008:22 Mellander Erik and Sofia Sandgren-Massih “Proxying ability by family background in returns to schooling estimations is generally a bad idea”

Dissertation Series

2007:1 Lundin Martin “The conditions for multi-level governance: implementation, politics and cooperation in Swedish active labor market policy”

2007:2 Edmark Karin “Interactions among Swedish local governments”

2008:1 Andersson Christian “Teachers and student outcomes: evidence using Swedish data”