ELSEVIER
European Economic Review 43 (1999) 889–901
EUROPEAN ECONOMIC REVIEW

Lower and upper bounds of returns to schooling: An exercise in IV estimation with different instruments

Andrea Ichino[a,c,d,*], Rudolf Winter-Ebmer[b,d,e]

[a] *European University Institute, I-50016 San Domenico di Fiesole, Firenze, Italy*
[b] *University of Linz, A-4040 Linz, Austria*
[c] *IGIER, Milan, Italy*
[d] *CEPR, London, UK*
[e] *WIFO, Vienna, Austria*

Abstract

Several recent studies based on 'exogenous' sources of variation in educational outcomes show IV estimates of returns to schooling that are substantially higher than the corresponding OLS estimates. Card (1995a, Earnings, schooling, and ability revisited. Research in Labor Economics 14, 23–48) suggests that these results are explained by the existence of heterogenity in individual returns and by the fact that these studies are based on instruments that influence only the educational decision of individuals with high marginal returns due to either liquidity constraints or to high ability. This conclusion is consistent with the local average treatment effect (LATE) interpretation of IV (Imbens and Angrist, 1994, Identification and estimation of local average treatment effects. Econometrica 62, 467–475) according to which IV identifies only the average returns of those who comply with the assignment-to-treatment mechanism implied by the instrument. We show evidence for Germany suggesting that returns to schooling are heterogeneous, instruments matter and the LATE interpretation of IV makes sense. With an appropriate choice of instruments we also show how IV can be used to approximate the range of variations of returns to schooling in Germany.

* Corresponding author. E-mail: ichino@datacomm.iue.it.

0014-2921/99/$ – see front matter
PII: S0014-2921(98)00102-0

JEL classification: J31; I21; J24; C25

Keywords: Returns to schooling; Wage determination; Instrumental variables

1. Introduction

At the origin of the empirical literature on the causal effect of education on earnings, instrumental variables (IV) techniques were adopted with the goal of finding a consistent estimate of the return to schooling which was thought to be a unique parameter in the population.[1] The need for IV was motivated by at least two sources of distortion potentially affecting OLS estimates. First, the 'ability' bias causing OLS to overestimate the true return to schooling if individuals with higher income generating capacities are also individuals who choose higher education. Second, the 'measurement error' bias causing OLS to underestimate the true return to schooling if the educational attainment is not observed precisely.

More recently, however, the joint consideration of three different sets of contributions to the literature suggests a different interpretation of the results obtained with IV techniques. First, building on Becker (1967), Card (1995a, 1998) has shown how much mileage is offered by a theoretical model in which 'the return to education is not a single parameter in the population, but rather a random variable that may vary with other characteristics of individuals, such as family background, ability, or level of schooling'. In his model, educational choices are optimally made by equating marginal returns to marginal costs of schooling, where both are assumed to be heterogenous in the population. As a result, at each individually optimal level of education, marginal returns can be expected to vary widely. Given such heterogeneity, one might wonder whether the estimation of the *average return to schooling* in the population is feasible in the first place. Card (1995a, 1998) shows that in data generated by his model, not only OLS but also IV techniques would provide biased estimates of the *average return* in the population.

A second set of contributions comes from the literature on the evaluation of 'treatment' effects (see e.g. Heckman, 1997), that applies to returns to schooling when the treatment is defined as the acquisition of additional education and the outcome is defined as labor earnings. A crucial insight of this literature is that when treatment effects are idiosyncratic and participation into treatment is not random, the estimation of the effect of treatment on a random person in the population is not only impossible in the absence of randomized controlled experiments, but, perhaps more importantly, it is *not even* an interesting research strategy. It is indeed hard to think of meaningful positive or normative questions

[1] See for example Willis (1986).

whose answers require the estimation of the average return to education in the population, because every policy measure is likely to influence only a certain population subgroup. Certainly more interesting is the estimation of the 'effect of treatment on the treated' which in the case of education would be the return to schooling for those who decide to acquire more schooling. But Heckman (1997) shows that not only OLS but also IV techniques require very restrictive assumptions in order to provide estimates of the average 'effect of treatment on the treated'. Angrist et al. (1996) suggest convincingly that the only treatment effect that IV can consistently estimate is the average treatment effect for those who change treatment status because they comply with the assignment-to-treatment mechanism implied by the instrument: they call this parameter local average treatment effect (LATE). For example, according to this interpretation, IV estimates of returns to schooling based on college proximity as an instrument (as in Card, 1995b) should be interpreted as the average return to schooling for a person that acquires an additional year of education only because home is close to college but would drop out of school if no college could be found nearby. In the absence of heterogeneity of returns, the LATE would obviously be equal to the true (and unique) return to schooling in the population at large. But in the presence of heterogeneity, the LATE would in general be very different from both the 'effect of treatment on a random person' and the 'effect of treatment on the treated'. One crucial consequence of this interpretation of IV is that using different instruments one should expect to estimate different returns to schooling, inasmuch as the different instruments (i.e. assignment-to-treatment mechanism) change, at the margin, the educational decisions of individuals in different subgroups of the population.

The third set of relevant contributions collects some recent empirical studies in which IV estimates of returns to schooling are based on 'exogenous' sources of variation in educational outcomes. The common theme of these studies, surveyed in Card (1995a, 1998), is that they all obtain IV estimates of returns to schooling that are substantially higher than the corresponding OLS estimates. Since errors in the measurement of schooling would cause a downward bias for OLS but not for IV, this difference could in principle be attributed to the imprecise observation of educational attainments. But the size of this difference is too large to be explicable by measurement error alone. An alternative possibility, again proposed by Card (1995a, 1998), is that this difference might be caused by heterogeneity of returns to schooling. Indeed, all these studies are based on instruments that are likely to influence the educational decision of individuals who have high marginal returns at the optimal decision. This may be so because liquidity constraints increase their marginal cost of schooling or because higher ability increases their marginal return to schooling. Although Card does not stress this point explicitly, his interpretation of the IV–OLS difference found in these studies is consistent with the LATE interpretation of IV proposed by Angrist et al. (1996).

The joint consideration of these three sets of contributions provides *indirect* evidence supporting the validity of a framework for the analysis of the causal effect of education on earnings, based on the following ingredients: (1) returns to schooling are heterogenous in the population, and (2) different instruments should generate different estimates of average returns for different subgroups in the population. However, the existing '*direct*' evidence on the heterogeneity of returns and on the variability of IV estimates obtained with different instruments does not seem to support this framework: in particular, Harmon and Walker (1996) find no evidence for the UK that different instruments affect different decision margins and conclude that heterogeneity of returns is not a convincing explanation of the difference between IV and OLS estimates in the studies surveyed by Card (1995, 1998). In contrast to Harmon and Walker (1996), in this paper we present some direct and indirect evidence based on Germany that supports the existence of heterogeneous returns to schooling and the validity of the LATE interpretation of IV. Furthermore, with an appropriate choice of instruments we also show how IV can be used to approximate upper and lower bounds of returns to schooling in Germany.

2. A model for the choice of instruments when returns are heterogeneous

In this section we draw heavily on Card (1998) to guide our search for instruments that should be expected to identify different average returns to schooling in the presence of heterogeneity.

Individuals are assumed to choose the optimal number of years of schooling S to maximize the utility function

$$U(S,Y) = \log(Y) - h(S), \tag{1}$$

where $Y = Y(S)$ is the income generating function for an individual with S years of schooling and h is an increasing convex function of S. Following Becker (1967) we interpret the assumption of strict convexity of h as implying that the marginal cost of each additional year of schooling rises by more than the foregone earnings of that year because of liquidity constraints.[2]

The optimal number of years of schooling is obtained from the solution of the first order condition

$$\frac{Y'(S)}{Y(S)} = h'(S). \tag{2}$$

[2] The utility function (1) can be shown to generalize a standard discounted present value objective function.

In order to introduce the possibility of heterogeneous schooling decisions, Card (1998) assumes differences in the individual marginal returns to schooling (the slopes of the income generating function Y) and in the individual marginal costs of schooling (the slopes of the cost function h). This can be shown in a linearized form:

$$\frac{Y'(S)}{Y(S)} = b_i - k_b S, \qquad h'(S) = r_i + k_r S. \tag{3}$$

The parameter b_i captures differences in individual ability defined as a factor that increases the marginal return to schooling.[3] The parameter r_i captures the possibility of differences in the cost of additional schooling that each individual faces: r_i is larger for those who are more liquidity constrained. Furthermore, for each individual, marginal returns decrease with schooling while marginal costs increase.

Substituting Eq. (3) into Eq. (2), the optimal amount of schooling differs across individuals and is equal to

$$S_i^* = \frac{(b_i - r_i)}{k_b + k_r}. \tag{4}$$

For each individual it is now possible to characterize the marginal return to schooling β_i *at her optimal level of schooling* defined by Eq. (4):

$$\beta_i = b_i - k_b S_i^* = (1-\phi) b_i + \phi r_i, \tag{5}$$

where $\phi = k_b/(k_b + k_r)$. Therefore, in this model the marginal return to schooling differs across individuals and is equal to a weighted average of the two parameters that generate the heterogeneity of ability and costs. In order for the marginal return at the optimum to be the same across individuals ($\beta_i = \bar{\beta}$), either the marginal costs or the marginal returns have to be identical across individuals and constant w.r.t. years of schooling ($h'(S) = \bar{r}$ or $Y'(S)/Y(S) = \bar{b}$).[4]

To make our point more clear, we assume that each heterogeneity parameter can take only two values: $b_H > b_L$ and $r_H > r_L$. The four possible combinations of these two values for each parameter characterize four groups of individuals in the population, denoted by $g = \{LL, LH, HL, HH\}$. Hence, there are four possible optimal returns to schooling β_g. Finally assume that the distribution of types in the population is given by the four probabilities $\{P_{LL}, P_{LH}, P_{HL}, P_{HH}\}$.

[3] Note that an alternative definition would characterize ability as a factor that increases incomes at all schooling levels; but with this second definition, more able individuals could end up choosing less schooling, which is perhaps counterfactual. This outcome is instead excluded by the assumption that ability increases only the marginal returns to schooling.

[4] The exact conditions are, either (a) $b_i = \bar{b}$ and $k_r = 0$, or (b) $r_i = \bar{r}$ and $k_b = 0$.

Do all four types exist in the population? Card (1995a, 1998) suggests that the correlation between b_i and r_i is likely to be negative if ability tends to persist across generations and if more able dynasties tend to be richer and therefore less liquidity constrained. But even a negative correlation, which we find convincing, would nevertheless be compatible with positive probabilities for all the four combinations of abilities and costs, provided that the two extreme cases corresponding to $g = LL$ (e.g. the 'stupid rich') and to $g = HH$ (e.g. the 'smart poor') are less frequent. The existence of these two extreme cases is certainly plausible and will play a crucial role in the following:

Consider a dichotomous exogenous source of variation in schooling Z_i such that

$$\mathrm{E}\,(S_i|Z_i = 1) \neq \mathrm{E}\,(S_i|Z_i = 0).$$

The IV estimator based on the instrument Z_i of the schooling coefficient in the regression $\log(Y) = \alpha + \beta S + \varepsilon$ would have probability limit

$$\mathrm{Plim}\ \beta_Z^{\mathrm{IV}} = \frac{\mathrm{E}\,(\log Y_i|Z_i = 1) - \mathrm{E}\,(\log Y_i|Z_i = 0)}{\mathrm{E}\,(S_i|Z_i = 1) - \mathrm{E}\,(S_i|Z_i = 0)} = \frac{\mathrm{E}_g(\beta_g \Delta S_{g|Z})}{\mathrm{E}_g(\Delta S_{g|Z})}, \tag{6}$$

where E_g is the expectation taken on the distribution of the four groups g, and $\Delta S_{g|Z}$ is the exogenous change in schooling induced by the instrument Z among group g individuals who all share, at the optimum, the same marginal return to schooling β_g.

Card (1998) shows that, if $\Delta S_{g|Z} = \Delta S_Z$, i.e. the instrument induces the same marginal change in schooling for all the four groups, clearly

$$\mathrm{Plim}\ \beta_Z^{\mathrm{IV}} = E_g(\beta_g) \tag{7}$$

and IV would estimate consistently the average return to schooling in the population. Furthermore, if $\beta_g = \beta$, i.e. the marginal return to schooling is identical in the four groups, then IV estimates consistently the unique return to education in the population. But let aside these two special cases, it does not make sense to talk about a unique marginal return to schooling in the population and there is no reason to expect that IV should consistently estimate the *average* return in the population, even if this were an interesting parameter.[5]

Given this result, it is natural to wonder what can be consistently estimated using IV. In general, Eq. (6) suggests that the probability limit of the IV estimator based on Z is a weighted average of the marginal returns to schooling in the four groups where the weights depend on the impact of Z on S, $\Delta S_{g|Z}$, which can differ across groups – and on the size of the four groups. It is,

[5] As we argued in the introduction we agree with Heckman (1997) in considering this parameter relatively uninteresting.

therefore, evident that, if a different instrument W affected the schooling decision of the four groups differently than Z, the probability limit of the IV estimator based on W would in general be different. In other words, IV estimates based on different instruments could very well differ substantially one from the other and would not have anything to do with the unweighted average return to schooling in the population at large.

This intuition suggests that if one could find an instrument affecting only the schooling decision of one of the four groups, IV would estimate consistently the marginal return to schooling for that particular group. And, more importantly, if one could find instruments capable of identifying the highest and the lowest returns, the corresponding estimates would allow us to bracket the range of variation of marginal returns to schooling in the population.

The instruments used by Ichino and Winter-Ebmer (1998) to estimate the loss of earnings suffered by the individuals who received less education because of World War II seem to serve well the purpose of estimating the upper limit of the range of returns to schooling in Germany. Let $Z_i = 1$ indicate the fact that the father of individual i was involved in WWII, while $Z_i = 0$ for those individuals whose father was not involved in the war. We refer to Ichino and Winter-Ebmer (1998) for a discussion of the validity of this instrument and for the evidence showing its effects on schooling decisions. Here, we first argue that this instrument is likely to cause an exogenous variation in schooling for the individuals in the group $g = HH$. These are liquidity constrained individuals who, thanks to their ability, choose more schooling in the absence of the war constraint but drop out of school if constrained by the war.[6] We further conjecture that none of the other groups is likely to be affected by this instrument. First of all, individuals in $g = LL$ and $g = HL$ (i.e. the 'rich dynasties') suffer limited liquidity constraints and therefore the war is likely to be irrelevant for their schooling decision.[7] Furthermore, also group $g = LH$ is likely to be unaffected by the war instrument because these individuals have limited ability and they are liquidity constrained. Because of this combination of factors they are likely to acquire only the minimum amount of schooling regardless of the war.[8]

If these conjectures are correct we should expect $\Delta S_{LL|Z} = \Delta S_{HL|Z} = \Delta S_{LH|Z} \approx 0$ and therefore

$$\text{Plim}\, \beta_Z^{\text{IV}} \approx \beta_{HH}\,, \tag{8}$$

which is the highest return in the population.

[6] In the LATE framework of Angrist et al. (1996) these are the *compliers*, i.e. those who comply with the assignment mechanism implicit in the war instrument. Note that given the instrument, in this case the treatment is defined as a reduction of schooling.

[7] Referring again to the framework of Angrist et al. (1996) these are the *never-takers*, i.e. those who refuse the treatment independently of the assignment.

[8] These are the *always-takers* because they choose less schooling independently of the war.

At the other extreme, to estimate the lowest return in the population, β_{LL}, we conjecture that family background could offer instruments that, with some caveats, would serve our goal.[9] Card (1998) argues extensively and convincingly that family background is likely to provide upwardly biased estimates of the *average* return to schooling in the population not only because family background is very likely to have an independent causal effect on earnings, but also because ability is likely to persist across generations. But let us abstract for one moment from this problem, on which we will come back later, and let us concentrate instead on the population group whose schooling decision is likely to be affected by an instrument based on family background.

To be specific, assume that $W_i = 1$ when the father of individual i has a degree higher than high-school, and $W_i = 0$ otherwise. We conjecture that in this case only $\Delta S_{LL|W}$ would be different from zero, and in particular positive. This is because the group $g = LL$ (the *compliers*) includes rich individuals with limited ability who may be helped/pushed to reach a higher education if their parents are highly educated, but would not do so otherwise. By way of contrast, the groups $g = HL$ and $g = HH$ (the *always-takers*), given the higher ability of their members, are likely to continue into higher education independently of the education of the father. Finally, individuals in group $g = LH$ may not be sufficiently able and may be too heavily liquidity constrained to continue into higher education independently of the father's education (the *never-takers*).

If these conjectures are correct, we should expect the probability limit of the instrument based on W to be

$$\text{Plim}\ \beta_W^{\text{IV}} \approx \beta_{LL} + N, \tag{9}$$

where $N > 0$ is the potential bias caused by the existence of a direct causal effect of family background on earnings.[10] Inasmuch as the bias N exists, Eq. (9) could be considered as an upper bound of the lowest return to schooling in the population.

In Section 3 we search for evidence in favor of these conjectures. More precisely, we want to know if $\beta_Z^{\text{IV}} > \beta_W^{\text{IV}}$ in which case we could safely conclude that $\beta_{HH} > \beta_{LL}$.

Note that the existence of a positive bias N affecting the IV estimate based on family background W would just reinforce our conclusion. This evidence would therefore support the existence of heterogenous returns to schooling which would be ranging within the limits approximated by the IV estimates based on the instruments Z and W defined above.

---

[9] Intergenerational correlation in education has been documented consistently for different countries, see Ichino et al. (1996).

[10] Angrist et al. (1996) show how to compute the IV bias, with respect to the LATE parameter, caused by violations of the exclusion restrictions. In our case, the LATE parameter is β_{LL} i.e. the average return to schooling of those who choose more schooling only because their father has higher education.

3. IV estimates with different instruments in Germany

Using data for male German workers from the 1986 wave of the German Socio-Economic Panel, we estimate returns to schooling based on the two dichotomous instruments Z and W described in Section 2: $Z_i = 1$ indicates the fact that the father of individual i served actively in WWII, and $Z_i = 0$ otherwise; $W_i = 1$ indicates instead that the father of individual i has reached a degree higher than high-school, and $W_i = 0$ otherwise. We use a human capital specification of the earnings function in which the logarithmic hourly wage is regressed on years of schooling and on a polynomial in age.[11]

Results are presented in Table 1. Columns 1–4 present IV estimates, whereas the corresponding OLS estimate is shown in column 5 for comparison. In the first column, using the father-in-war instrument, the return to one further year of schooling is estimated equal to 14.0%. In the second column, instead, using the father's-education instrument the estimate is substantially lower, being equal to 4.8%. As we argued in Section 2, these two estimates can be considered as an approximation of the upper and lower bounds of the returns to schooling in Germany. The true range of variation is actually likely to be larger on the bottom side because the estimate in column 2 – where father's education is used as an instrument – is likely to be biased upward if (i) parental background has an independent impact on earnings or (ii) if individual-specific returns to education or individual-specific earnings components are themselves correlated with family background. This wide range of variation represents a puzzle unless the existence of individual heterogeneity of returns to schooling is accepted[12] and IV estimates are interpreted as LATE estimates of the returns for different subgroups in the population.

The IV estimate based on the father-in-war instrument is consistent with the conjecture that this instrument changed the educational decision of students who had high marginal returns and high marginal costs of education, i.e. the group $g = HH$. Some individuals in this group would have proceeded to higher education in peaceful times, but could not do so, because of the constraints imposed by the involvement of their fathers in the war.[13] On the other hand, the estimates based on father's education as an instrument are consistent with the

[11] Note that our estimates of the 'return to schooling' are therefore lower than the ones we would have obtained using the standard specification based on potential experience instead of age.

[12] Of course, an alternative hypothesis would be, that our instruments are bad. As we argued before, there are good reasons to assume that the instrument 'father's education' gives upward biased results, whereas we argue elsewhere (Ichino and Winter-Ebmer, 1998) that the 'father-in-war' instrument should be considered a valid instrument.

[13] Note that because of the very low high-school completion as well as university enrollment rates in Germany at that time, only the most talented youngsters from poor families could go to higher education in the first place.

Table 1
IV estimates of returns to schooling with different instruments in Germany

	IV				OLS
	Instrument: father-in-war	Instrument: father highly educated	Instrument: father-in-war	Instrument: father highly educated	
Years of education	0.140	0.048	0.117	0.048	0.055
	(0.078)	(0.013)	(0.053)	(0. 014)	(0.005)
Age (yrs)	0.106	0.215	0.141	0.215	0.208
	(0.101)	(0.039)	(0.070)	(0.039)	(0.033)
$Age^2/100$	− 0.183	− 0.434	− 0.263	− 0.434	− 0.418
	(0.235)	(0.093)	(0.164)	(0.094)	(0.084)
$Age^3/10\,000$	0.106	0.291	0.165	0.290	0.279
	(0.175)	(0.007)	(0.123)	(0.008)	(0.007)
Father is a blue-collar work (0,1)	–	–	0.058	− 0.001	0.004
			(0.051)	(0.031)	(0.026)
Father is self-employed (0,1)	–	–	− 0.032	− 0.041	− 0.041
			(0.043)	(0.042)	(0.037)
Father has more than high-school education (0,1)	–	–	− 0.209	–	− 0.019
			(0.172)		(0.052)
Constant	− 0.684	− 1.080	− 0.909	− 1.075	− 1.060
	(0.619)	(0.483)	(0.517)	(0.484)	(0.411)
R^2	0.071	0.207	0.148	0.207	0.205
No observations	1822	1822	1822	1822	1822
Partial R^2 for instrument in 1st stage	0.003	0.114	0.006	0.085	–
F-test on instrument in 1st stage	5.53	211.2	14.2	189.2	–

Standard errors in parentheses. The sample is taken from the 1986 wave of the German Socio-Economic Panel. The dependent variable is the log of hourly wages. The 'father-in-war' instrument is an indicator that takes value 1 if the father has been involved in WWII. The 'father highly educated' instrument takes value 1 if the father has obtained a degree higher than high-school.

conjecture that students with lower marginal returns and lower marginal costs reacted to the assignment mechanism implied by this instrument (i.e. the group $g = LL$).

It is of course possible that each of our two instruments affects more than one group in the population. For example, having the father-in-war could reduce the education not only of the $g = HH$ group, but also of the other groups. In this case, the probability limit of the IV estimate based on the father-in-war instrument would be a weighted average of β_{LL}, β_{HL}, β_{LH} and β_{HH}. Given the above assumptions, this weighted average would certainly be lower than β_{HH}.[14] Similarly the probability limit of the IV estimate based on father's education could be a weighted average of the four returns that would certainly be larger than β_{LL}.[15] For this reason, the interval delimited by our IV estimates (4.8%–14.0%) is likely to be smaller than the interval between the highest and the lowest returns (β_{LL}–β_{HH}). This argument, therefore, reinforces the conclusion that returns to schooling are heterogeneous in Germany.

Bound and Jaeger (1996) argue that IV estimates could be biased upward by unobserved differences between the characteristics of the treatment and the control groups implicit in the IV scheme. This would for example happen if treatment and control groups came from different social backgrounds. Following a suggestion by Card (1998) we therefore include also information on parental background as control variables in our columns 3 and 4. As information on parental background we use father's social status, which is captured by dummies for self-employment and white-collar versus blue-collar jobs. These last variables were measured at the time the student was 10 years old. For the estimate based on the father-in-war instrument we also add father's education to better control for parental background. Note that the elimination of upward bias suggested by Bound and Jaeger (1996) and Card (1998) is particularly important for the estimate based on the father-in-war instrument because this is

[14] But there is some evidence suggesting that only children from less-educated parents reduced their educational attainment because of the father's involvement in the war. In fact, the father-in-war instrument reduces schooling by 1.59 (0.39) years for those students whose father had only compulsory education, but only by 0.49 (0.82) years for other students (standard errors in parentheses).

[15] However, collateral evidence suggests that parental education mostly affects the schooling decision of richer students. We do not have income information on parents, but use indicators of social status instead. If the father has a degree higher than high-school, the education of the child rises by 3.84 (0.66) years in households with self-employed heads, by 2.98 (0.31) years in households with white-collar heads and only by 0.49 (0.96) years in households with blue-collar heads (standard errors in parentheses). If this interpretation were correct, we expect the weights attached to the lower returns in the IV estimand to be large. Accordingly, for the father-in-war instrument, we can suspect the weight of the highest returns to be large. In conclusion, we believe that even if the probability limits of our IV estimates are weighted averages of the four returns, these averages are not too different from (respectively) β_{LL} and β_{HH}.

our estimate of the upper bound of the returns to education. Adding these controls reduces the IV estimate based on the father-in-war instrument from 14.0% to 11.7%, but leaves the other IV estimate unchanged. Hence, according to these results the interval of variations of returns to schooling is somewhat smaller then the one implied by columns 1 and 2, but still substantial since it ranges at least between 4.8% and 11.7%.

4. Conclusion

In this paper we presented evidence supporting the validity of a framework for the analysis of the causal effect of education on earnings in which (1) returns to schooling are heterogenous in the population, and (2) different instruments generate different estimates of average returns for different subgroups in the population. With an appropriate choice of instruments we show how a local average treatment effect interpretation of IV[16] can be used to approximate the range of variations of returns to schooling in Germany. A conservative estimate of this range suggests that the lower bound of returns to schooling is not higher than 4.8% while the upper bound is not lower than 11.7%.

Acknowledgements

We would like to thank Joshua Angrist, David Card, Luca Flabbi and Wilhelm Kohler for helpful comments. This research was supported by a EU-TSER grant (SOEZ-CT98-2044).

References

Angrist, J.D., Imbens, G.W., Rubin, D.B., 1996. Identification of causal effects using instrumental variables. Journal of the American Statistical Association 91, 444–455.

Becker, G.S., 1967. Human Capital and the Personal Distribution of Income. University of Michigan Press, Ann Arbor, MI.

Bound, J., Jaeger, D.A., 1996, On the validity of season of birth as an instrument in wage equations: A comment on Angrist and Krueger's 'Does compulsory school attendance affect schooling and earnings?' Working paper no. 5835. NBER, Cambridge, MA.

Card, D., 1995a. Earnings, schooling, and ability revisited. Research in Labor Economics 14, 23–48.

[16] Note that the LATE interpretation of IV estimates in the presence of two instruments for the same relationship is fundamentally different from a conventional one: the latter interpretation would assess one of the two instruments as invalid, if the two results are not the same.

Card, D., 1995b. Using geographic variation in college proximity to estimate the returns to schooling, In: Christofides, L.N., Grant, E.K., Swidinsky, R., (Eds.), Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp, University of Toronto Press, Toronto, pp. 201–222.

Card, D., 1998. The causal effect of education on earnings. In: Ashenfelter, O., Card, D., (Eds.), Handbook of Labor Economics, vol. 3, North-Holland, Amsterdam, forthcoming.

Harmon, C., Walker, I., 1996. The marginal and average returns to schooling, Mimeo. Institute for Fiscal Studies, London.

Heckman, J.J., 1997. Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. Journal of Human Resources 32, 441–462.

Ichino, A., Rustichini, A., Checchi, D., 1996. More equal but less mobile? Education financing and intergenerational mobility in Italy and in the US. Discussion paper no. 1466. CEPR, London.

Ichino, A., Winter-Ebmer, R., 1998. The long-run educational cost of world war II: An example of local average treatment effect estimation. Discussion paper no. 1895. CEPR, London.

Imbens, G., Angrist, J.D., 1994. Identification and estimation of local average treatment effects. Econometrica 62, 467–475.

Willis, R.J., 1986. Wage determinants: A survey and reinterpretation of human capital earnings functions. In: Ashenfelter, O., Layard, R. (Eds.), Handbook of Labor Economics. North-Holland, New York.