ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Pfeiffer, Friedhelm; Pohlmeier, Winfried

Working Paper

Causal returns to schooling and individual heterogeneity

IZA Discussion Papers, No. 6588

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Pfeiffer, Friedhelm; Pohlmeier, Winfried (2012) : Causal returns to schooling and individual heterogeneity, IZA Discussion Papers, No. 6588, Institute for the Study of Labor (IZA), Bonn,
https://nbn-resolving.de/urn:nbn:de:101:1-201303221775

This Version is available at:
https://hdl.handle.net/10419/58467

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 6588

Causal Returns to Schooling and Individual Heterogeneity

Friedhelm Pfeiffer
Winfried Pohlmeier

May 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Causal Returns to Schooling and Individual Heterogeneity

Friedhelm Pfeiffer
ZEW and IZA

Winfried Pohlmeier
University of Konstanz

Discussion Paper No. 6588
May 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 6588
May 2012

ABSTRACT

Causal Returns to Schooling and Individual Heterogeneity*

In this paper, human capital investments are evaluated by assuming heterogeneous returns to schooling. We use the potential outcome approach to measure the causal effect of human capital investments on earnings as a continuous treatment effect. Empirical evidence is based on a sample of West German full-time employed males (BIBB/IAB survey on educational and vocational attainment and career 1998/99). Our estimate of the average partial effect (APE) of an additional year of schooling amounts to 8.7%, which is higher than OLS estimates and quite similar to conventional instrumental variable estimates.

JEL Classification: J21, J24, J31

Keywords: returns to schooling, human capital, heterogeneity

Corresponding author:

Friedhelm Pfeiffer
Centre for European Economic Research (ZEW)
P.O. Box 103443
68034 Mannheim
Germany
E-mail: pfeiffer@zew.de

* Financial support by the DFG through research group "Heterogeneous Labor" at the University of Konstanz and the ZEW is gratefully acknowledged. Friedhelm Pfeiffer furthermore thanks the German Federal Ministry of Education and Research (BMBF) for financial support under the grant 01PW11001. We would like to thank Michael Maier for competent research assistance and Derya Uysal for valuable comments. The usual disclaimer applies.

1 Introduction

Two major issues concerning the impact of schooling on earnings have been raised in the empirical literature on human capital investments.[1] First, education as the individual's choice parameter is an endogenous choice variable in the standard earnings function. Therefore, coefficients on schooling from conventional Becker-Mincer type of earnings functions can only be interpreted as causal returns to schooling if the schooling level is uncorrelated with unobserved individual factors, i.e. if schooling is more or less assigned to individuals randomly. Therefore, ordinary least squares estimates are only of explorative nature and their usefulness with respect to policy recommendations is limited. Second, as a choice parameter the individual's schooling level is determined by the individual's observed and unobserved marginal benefits and costs of schooling. Thus, the return of an additional year of schooling varies across individuals and is driven by observable factors (e.g. family background, school quality etc.) as well as unobservable factors (e.g. cognitive and noncognitive skills, peer group and network effects).

This paper presents empirical results for the causal returns to schooling in Germany accounting for endogeneity and heterogeneity. In order to assess the causal effect of education on earnings, we adopt the concept of the average partial effect (APE) developed in the econometric evaluation literature for causal effects on continuous treatments. This can be used to quantify the expected earnings difference between two otherwise identical individuals if they had been assigned randomly to S and S + 1 years of schooling, respectively. Contrary to previous studies on the APE in earnings functions with heterogeneous returns which rest on a

[1] See Card (1999), Flossmann and Pohlmeier (2006), Gebel and Pfeiffer (2010), or Wooldridge (2002), among others.

a control function approach (Garen (1984) and Heckman and Vytlacil (1998) for the US), we apply the conditional moment independence (CMI) approach proposed by Wooldridge (2004) to identify the APE on schooling. The approach applied in this study is easy to implement and reveals several advantages over other popular econometric approaches to estimate causal returns to schooling in the presence of heterogeneity. Aside from the problem that the classical instrumental variable approach to estimate the Local Average Treatment Effect (LATE) only provides estimates of causal returns to schooling for the potentially very small subgroup of compliers, it suffers in real world applications from the drawback of the availability of an appropriate and valid instrument (see Heckman and Urzua (2010)). Moreover, this binary instrument should have a relevant policy interpretation. Alternative valid instruments lead to estimates for different groups of compliers and therefore have a different interpretation. Similar to the standard instrumental variable approach, control function approaches to random coefficient models (Garen (1984), Heckman and Vytlacil (1998), Wooldridge (2008)) require strong assumptions on exclusion restrictions, which are often difficult to justify.

Using cross-sectional data on 7,722 male full-time workers in Germany we present evidence that the causal returns to schooling estimated by the conditional moment approach are well above the figures one obtains by ordinary least squares estimates, but lower than the two-stage least squares estimates for models assuming inessential heterogeneity in the returns to schooling.

Our paper is organized as follows: In Section 2, we develop the idea of random returns to education based on Card's schooling model (Card, 1999). Following Wooldridge (2004), we identify the average treatment effect via CMI assumptions and demonstrate that the APE for the continuous treatment variable schooling can be estimated by means of auxiliary

regressions. Section 3 describes the data and provides some information on the institutional settings in Germany. Our empirical findings are presented in Sections 4 and 5, while Section 6 concludes with an outlook on future research.

2 The Conditional Moments Approach

For many years the literature on returns to schooling has emphasized that schooling is a choice variable depending on observable and unobservable factors which determine the individual's marginal costs and benefits of schooling. For the econometrician, this implies that the individual return rate itself is a random variable correlated with these observable and unobservable factors and therefore returns to schooling vary across individuals. These basic features can easily be captured by Card's (1999) simple model of schooling and earnings, which we will use in the following as a specification device. The individual is assumed to choose the optimal amount of schooling, S, and earnings, Y, which maximize his lifetime utility depending on earnings and the disutility of schooling, $\varphi(S)$:

$$\max_{S,Y} U(S,Y) = \ln Y - \varphi(S) \text{ with } \varphi'(S) > 0 \text{ and } \varphi''(S) > 0. \tag{2.1}$$

Let the benefits of schooling (schooling-earnings relationship) be $Y = Y(S)$ with $Y'(S) > 0$. This yields the first order conditions

$$\frac{Y'(S)}{Y(S)} = \varphi'(S). \tag{2.2}$$

The conventional log earnings function linear in schooling arises if marginal benefits are constant:[2]

$$MB \equiv \frac{Y'(S)}{Y(S)} = \beta \,. \tag{2.3}$$

Note, that in this framework β is a fixed parameter for the optimizing individual. However, from the viewpoint of the econometrician it is assumed to be a correlated random variable depending on observable and unobservable factors. If marginal costs (MC) are linear in schooling, then:

$$MC \equiv \varphi'(S) = \gamma + \kappa S, \quad \kappa > 0 \,. \tag{2.4}$$

Given these functional form assumptions optimal schooling is given by the schooling function of the form

$$S = \frac{\beta - \gamma}{\kappa} \,. \tag{2.5}$$

Integration of the marginal benefit function (2.3) yields a log linear earnings function with random coefficients, an individual specific intercept and an individual specific slope coefficient:

$$\ln Y = \alpha + \beta S \,. \tag{2.6}$$

[2] Assuming a linear marginal benefit function results in a log earnings function that contains an additional quadratic schooling term.

The intercept α captures the absolute productivity (ability) advantages of the agent. Similar to β, it is also treated as a correlated random coefficient. Observable factors and unobserved heterogeneity in the absolute and marginal benefits of schooling as well as factors driving the marginal costs of schooling enter the earnings function through the coefficients α, β and γ, respectively. Let α be presented by the linear predictor function

$$\alpha = \alpha_0 + X_1'\alpha_1 + \eta_\alpha , \tag{2.7}$$

where X_1 is a vector of observables and the random variable η_α captures unobserved heterogeneity in the absolute productivity term (e.g. individual work experience, regional factors etc.) which serve as additional control variables in the classical Becker-Mincer framework. Likewise, marginal productivity may depend on the same set of factors:

$$\beta = \beta_0 + X_1'\beta_1 + \eta_\beta , \tag{2.8}$$

while marginal cost depends on the X_1 variables as well as on additional cost driving factors X_2:

$$\gamma = \gamma_0 + X_1'\gamma_1 + X_2'\gamma_2 + \eta_\gamma . \tag{2.9}$$

Inserting (2.7) - (2.9) in the schooling equation (2.5) yields a reduced form of the schooling equation:

$$\begin{aligned} S &= \frac{1}{\kappa}\left[(\beta_0 - \gamma_0) + X_1'(\beta_1 - \gamma_1) - X_2'\gamma_2 + \eta_\beta - \eta_\gamma\right] \\ &= \pi_0 + X_1'\pi_1 + X_2'\pi_2 + \xi \end{aligned} \tag{2.10}$$

Note that the returns of an additional year of schooling now is a random variable depending on the level of schooling and the marginal costs of schooling, i.e. the returns to schooling vary across the population. The APE of an additional year of schooling is the mean across all individual returns for an additional year of schooling:

$$APE = E\left[E\left[\ln Y \middle| S = s+1, \alpha, \beta\right] - E\left[\ln Y \middle| S = s, \alpha, \beta\right]\right] = E\left[\beta\right]. \tag{2.11}$$

The literature on causal effects in non-experimental settings is largely focused on binary treatments and to some extend to discrete multivalued treatments, while estimation issues of causal effects for continuous treatments has gained comparatively little attention. One can distinguish between three approaches to the treatment evaluation problem for continuous treatments. Garen (1984), Heckman and Vytlacil (1998) and Wooldridge (2008) propose an IV or control function approach that makes use of control functions such as (2.7) and (2.8) to estimate the APE from the reduced forms for earnings and schooling. The major drawback of this approach is the limited availability of a reasonable exclusion restriction (instrument) which identifies the causal treatment effect. Based on the conditional independence assumption the propensity score approach has been generalized to the case of continuous treatments by Imai and van Dyk (2004) and Hirano and Imbens (2005). Contrary to the binary case their generalizations require strong distributional assumptions. Here, we follow a suggestion by Wooldridge (2004, 2008). The APE is estimated in a random coefficient framework by assuming conditional mean independence. Under this assumption, treatments can be ignored conditional on a set of confounding variables. The APE can be identified if the following assumptions hold (ignorability conditions):

Identifying Assumptions for the APE (Wooldridge (2004):

A.1 Equation (2.6) holds.

A.2 For a set of covariates X, the following ignorability assumption holds:

$$E\left[\ln Y \middle| S, \alpha, \beta, X\right] = E\left[\ln Y \middle| S, \alpha, \beta\right]$$

A.3 Conditional on X, α and β are ignorable in the first two conditional moments of S:

$$E\left[S \middle| X, \alpha, \beta\right] = E\left[S \middle| X\right] \text{ and } \mathrm{V}\left[\mathrm{S} \middle| \mathrm{X}, \alpha, \beta\right] = V\left[S \middle| X\right] > 0$$

Identification condition *A.2* obviously holds since the control variable X enters the earnings function through α , β and S only. The linear predictor specification used for illustrative purposes in (2.7) and (2.8) is not required to identify the APE. In fact, the conditional mean independence approach uses identification conditions different from the control function approaches in correlated random coefficient models. Identification condition *A.3* denotes that conditional on the controls, expected schooling is mean independent of α and β . Thus no new information is gained in projecting schooling if there are sufficient controls. This is the crucial identification condition (ignorability condition) needed to identify the APE.

Proposition 2.1 (APE) (Wooldridge (2004):

Under the identifying assumptions *A.1 – A.3*, the average treatment effect for all X in the relevant population is given by

$$E[\beta] = E[E[\beta \mid X]] = E\left[\frac{Cov[S, \ln Y | X]}{V[S|X]}\right].$$

In the following analysis, we estimate V[S|X] and Cov[S, ln Y |X] by means of linear regression. Replacing the population parameters with the regression estimates yields a consistent estimate of the average treatment effect under the assumption of independent, identically distributed observations:

$$\hat{E}[\beta] = \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{Cov}[S_i, \ln Y_i | X_i]}{\hat{V}[S_i | X_i]}.$$

Contrary to the instrumental variable or control function approach, the CMI approach does not require exclusion restrictions for instrumental variables in such a way that the instruments drive the selection process (choice of the optimal years of schooling) and are uncorrelated with the error term of the earnings function. Since the APE is nothing but the mean of the ratio of second moments and cross-moments of schooling and earnings conditional on X, more insights into the causal effects of schooling can be obtained by analyzing other distributional properties of this ratio in addition to the mean.

3 Data

Our empirical study is based on a sample of full-time employed male workers from the so-called BIBB/IAB survey on educational and vocational attainment and career (BIBB/IAB (1999)), conducted in 1999. The BIBB/IAB[3] survey is a 0.1 % representative survey of German workers which has been conducted every five to six years since 1979. The objective of the survey is to produce "differentiated, actual data on workers in Germany, their qualifications and working conditions" (Dostal and Jansen (2002)). The data are processed and documented by the Central Archive for Empirical Social Research ("Zentralarchiv für empirische Sozialforschung", ZA), Cologne. Neither BIBB, IAB nor ZA take any responsibility for the analysis or the interpretation of the data presented here.

The 1999 survey contains comprehensive information on the number of years spent in the educational and vocational education system in Germany. In particular, our data contain extensive information on the successful completion of schooling levels (basic schooling, vocational and university education) and the actual years spent in the educational system to obtain the degree. Hence, our schooling variable is more closely related to the definition of an input variable compared to the standard measurement using either the minimum years required by the individual to receive his/her highest educational attainment or the average years of schooling necessary to attain a degree.

Table 1 contains selected summary statistics on the number of observations, on the number of years of schooling and on earnings for four different educational groups: workers without any

[3] BIBB: Federal Institute for Vocational Training ("Bundesinstitut für Berufsbildung"), IAB: Institute for Labour Market and Occupational Research of the Federal Labour Office (Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit).

formal occupational degree, workers with an apprenticeship degree ("Geselle"), workers with senior craftsmen qualifications ("Meister") or a degree from a university of applied sciences ("Fachhochschule") and workers with at least a university degree (for more details on the German educational system compare Blechinger and Pfeiffer (2000), among others). The overall years of standard schooling for these groups are 10, 13, 15-16 and 18. The actual number of years of schooling spent to capture a university degree is 1.7 years higher than standard years. For the quantitatively important group of workers with an apprenticeship degree, actual and standard years are not that different.

Table 1: Summary Statistics of Schooling and Earnings by Skill Group

Sample	Observations		Earnings [DM]		Schooling [years]	
	Freq.	Percent	mean	std.-dev.	mean	std.-dev.
Overall sample	7,722	100	4,697	1,986.4	14.3	3.5
Unskilled	762	9.9	3,689	1,667.7	10.6	2.4
Vocational training	4,988	64.6	4,302	1,572.0	13.7	2.5
Foreman, senior craftsman	1,330	17.2	5,627	2,017.5	16.5	3.6
University graduate	642	8.3	7,028	2,628.3	19.7	2.7

Data source: BIBB/IAB 1999; own calculation; for definitions and sample selection see text.

For our empirical analysis, we select a sample of German male full-time workers from the 1989/99 survey. We concentrate on full-time male workers because men by and large have an

inelastic labor supply and we can disregard selection into the labor force. Our earnings variable refers to the natural logarithm of gross monthly earnings before taxes. We end up with 7,722 observations. Summary statistics of the covariates are presented in Table A1 in the Appendix.

4 Empirical Findings

As a benchmark for our estimates of the APE, we first present the results of more traditional two-stage least squares estimates of the earnings function assuming homogenous returns to schooling (Table 2 and 3). The instruments used are the unemployment rate at graduation and its interaction terms with age and the squared age variable. This gives us three overidentifying restrictions. The reasoning behind the use of these instruments lies in some specific institutional features of the German vocational system. By opting for the elementary vocational year ("Berufsgrundbildungsjahr"), youths, especially those without an apprenticeship training position, have the opportunity to prepare for vocational training by attending a full-time school year (optional as part-time school). The preparation year for vocational training ("Berufsvorbereitungsjahr") basically serves the same purpose as the elementary vocational year, but in a somewhat broader sense. It prepares youths without an apprenticeship position for vocational training.[4] If unemployment reflects opportunity costs, an individual is more likely to stay in the educational system if employment prospects are low. This argument seems particularly relevant for the case of Germany where tuition and fees for general schooling and vocational training are rare exceptions or negligible.

[4] Franz et al. (2000) study the impact of vocational training on youth unemployment duration.

Table 2 presents the reduced form estimates for the schooling equation. Given the large value of the F-Test (195.84), we can reject the null of weak instruments in terms of the relative 2SLS bias (> 10%) and the actual size of the 2SLS t-test (> 15%) according to the critical values presented in Stock et al. (2002). The unemployment rate at graduation has a significant impact on the schooling level, and its impact varies across cohorts. Our specification explains 36% of the variation in schooling in the sample. Using the Hausman test (auxiliary regression specification) we can reject the hypothesis that schooling can be treated as an exogenous explanatory variable.

Table 2: Reduced Form Estimates of the Schooling Equation

Variable	β	t value
Experience	-.330	-16.65
Experience squared	-.003	-4.96
Age	.267	2.43
Age squared	.003	2.12
Unemployment ratio at graduation	-.331	-1.10
Unemployment ratio at graduation * Age	.028	1.79
Unemployment ratio at graduation * Age squared	.000	-1.41
Constant	6.685	2.94

N =7,722; F(22, 7699) = 195.84; R^2 = 0.36

Additional controls: sectoral dummies, regional dummies, firm size and a dummy variable for handicapped workers.
Dependent variable: Years of schooling; Data source: BIBB/IAB 1999; own estimates.

The 2SLS estimates of the fixed coefficient earnings function are given in Table 3. In addition to the typical covariates schooling, experience and experience squared, sector dummies, regional dummies, firm size and a dummy variable for handicapped workers are used as additional controls. The return to an additional school year is 8.3%, which is in line with previous estimates for Germany (see Flossmann and Pohlmeier, 2006). Ignoring the endogeneity of schooling by estimating the equation using ordinary least squares results in a lower estimate of 4.2 % (Maier et al., 2003). These differences confirm the international evidence that the return rates obtained from instrumental variable estimators often are above the ones from ordinary least squares estimates (see Card, 1999).

Table 3: 2SLS Estimates of the Earnings Equation

Variable	β	t value
Schooling	.083	33.57
Experience	.030	18.09
Experience squared	.000	-9.48
Handicapped	-.039	-1.88
Constant	6.920	159.9

N =7,722; F(18, 7703) = 132.19; R^2 = 0.236

Hausman test (N(0,1)) = 20.76

Additional controls: sector dummies, regional dummies, firm size dummies. Dependent variable: Logarithm of wage,
Data source: BIBB/IAB 1999, own estimates

The estimates of the expected rate of return to an additional year of schooling based on the CMI approach are reported in Table 4. In the first row of Table 4, we report the APE using all observations. Outliers turn out to have a significant effect on the reported APE values. Therefore, the second line presents estimates based on a trimmed sample, where observations below the 1% and above the 99% APE quantiles were dropped. Trimming obviously leads to more plausible estimation results.

Table 4: Estimates of the APE

$\hat{E}[\beta]$	t-value	Quantiles					
		10 %	25%	50%	75%	90%	
1.060	1.22	-.103	.005	.076	.158	.294	without trimming
0.087	29.94	-.091	.007	.076	.156	.283	trimmed

Data source: BIBB/IAB 1999; own estimates.

The CMI approach reveals an average treatment effect of an additional year of schooling (APE) at 8.7%, which is significantly different from zero. This estimate does not differ much from the 2SLS results reported above. Angrist and Imbens (1995) show that for models with variable treatment intensity, the 2SLS estimator identifies a weighted average of the treatment effect in the population whose educational attainment was changed by the instrument. Hence, there is no reason to expect ex ante quantitatively similar estimates. Using a control function approach, Gebel and Pfeiffer (2010) obtained a value of 7.2% for APE based on SOEP data 1999, which lies slightly above their OLS estimates (5.8%).

The quantiles of the individual return rates reported in Table 4 reveal that the impact of educational attainment on earnings is far from being homogeneous. For a quarter of the individuals, the causal return rate is more than 15.6%, and for the 90% quantile, it is 28.3%. On the other hand, for a quarter of the individuals, there are very low or even negative causal return rates. For example, negative return rates may result from a restricted entry into the labour market, in which case education serves as means of bridging over waiting queues in times of unemployment. They can be the result of a suboptimal matching between heterogeneous students and teaching institutions as well.

More descriptive evidence on the distribution of the heterogeneous returns is given by the kernel density estimates of the conditional average partial effect $E[\beta \mid X]$ depicted in Figure 1.

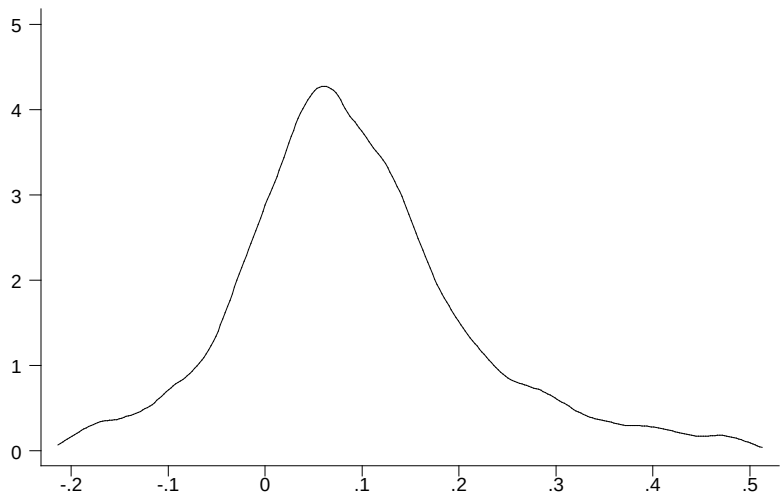

Figure 1: Kernel density estimates of the conditional average partial effect $E[\beta \mid X]$

The kernel density plot of the conditional average partial effect produces some insights on the heterogeneity of the returns to schooling, although it should not be mixed up with variance of

the partial effect or variance of the causal effect of the returns to schooling, $V[\beta]$, which cannot be identified given the assumptions on conditional moment independence. The distribution of the conditional average partial effect reveals fat tails indicating that for certain socio-economic background factors the returns to schooling can even be negative or very large. Moreover, the conditional APE turns out to be skewed to the right.

For economic policy reasons it is important to know who benefits from education. To get some empirical insights into this question, the sample has been divided into workers with a APE-value above the median, and workers below the median APE-value. Table 5 presents descriptive statistics on some relevant socio-economic characteristics in the two sub-samples.

Table 5: Socio-economic characteristics of two groups of workers

Variables	Estimated conditional ATE		t-value of difference
	below median	above median	
Schooling	13.92	14.76	-10.65
Earnings	4576.17	4817.15	-5.34
Age	37.53	37.40	.68
Qualification:			
Unskilled	.11	.09	3.97
Vocational Training	.68	.61	7.16
Foremen etc.	.16	.18	-2.53
University graduates	.04	.12	-13.34

Data source: BIBB/IAB 1999; own calculations.

The differences in the two samples suggests that on average the group with higher returns to education workers have acquired more years of schooling, are better qualified (the share of university graduates for example is 12% in the high returns group compared to only 4% in the low returns group) and do have higher earnings. These statistics seem to suggest that the law of diminishing returns to investments in human capital does not necessarily hold for all persons and all educational institutions at the chosen levels of schooling. If this interpretation is valid, then a variety of hitherto not fully exploited investment opportunities in schooling might exist for significant groups of individuals in Germany.

5 Concluding Remarks

In this paper, the returns on educational investments are assessed based on the potential outcome for continuous treatments. The estimate of the causal effect of schooling on earnings presented in this study is based on the CMI approach, taking into account heterogeneity of costs and benefits among individuals. Our estimate of the average causal effect of an additional year of schooling is 8.7%, which is close to the two stages least square estimate of the rate of return in a traditional fixed coefficient earnings function. Heterogeneity in the returns does matter and the monetary benefits of an additional year of schooling vary largely across the population. For 20 to 30% of the male workers in our sample, an additional year of schooling yields negative returns. For more than 25%, the returns are above 15%. Negative return rates may result from restricted entry into the labor market in which case education is a mean of bridging over waiting queues in times of unemployment, for example. The large positive returns may result from individual differences in learning abilities, educational costs and educational quality, among other reasons.

Since practical experience with the CMI approach to correlated random coefficient models is limited, our results, although plausible, should be treated with caution. More evidence is needed to assess the estimator. In particular, the estimation procedure applied is very sensitive to outliers. Trimming, for instance, can be used to reduce the impact of outliers. Another alternative would be the use of a shrinkage estimator, especially when a large number of confounding variables is available. For policy analysis, other treatment effects such as the effect of treatment on the treated, the treatment on the non-treated and the local average treatment effect should also be evaluated. Clearly, more research is needed to disentangle individual heterogeneity and institutional diversification for the economic returns of education.

References

Angrist, J. D., and G. W. Imbens (1995): Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity, *Journal of the American Statistical Association*, 90, 431 – 442.

BIBB/IAB (1999): Erwerb und Verwertung beruflicher Qualifikationen von Erwerbstätigen, *BIBB/IAB Strukturerhebung 1998/99. Erhebungsmaterialien zur Hauptstudie*. BIBB, Berlin.

Blechinger, D., and F. Pfeiffer (2000): Technological Change and Skill Obsolescence: the Case of German Apprenticeship Training, in: *Education and Training in a Knowledge-Based Economy*, ed. by H. Heijke, and J. Muysken, St. Martin's Press, New York, 243–276.

Card, D. (1999): The causal effect of education and earnings, *Handbook of Labour Economics*, Vol. 3 Part 1, ed. by O. C. Ashenfelter and D. Card, Elsevier, 1801-1863

Card, D. (2001): Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems, *Econometrica*, 69 (5), 1127–1160.

Dostal, W., and R. Jansen (2002): Qualifikation und Erwerbssituation in Deutschland - 20 Jahre BIBB/IAB Erhebungen, *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 35 (2), 232–253.

Flossmann, A. and W. Pohlmeier (2006), Causal Returns to Education: A Survey on Empirical Evidence for Germany, *Jahrbücher für Nationalökonomie und Statistik*, 226, 6-23.

Franz, W., J. Inkmann, W. Pohlmeier, and V. Zimmermann (2000): Young and Out in Germany: On Youths' Chances of Labor Market Entrance in Germany, in: *Youth Employment and Joblessness in Advanced Countries*, ed. by D. G. Blanchflower, and R. B. Freeman, University of Chicago Press, 381–425.

Garen, J. (1984): The Returns to Schooling: A Selectivity Bias Approach with a Continous Choice Variable, *Econometrica*, 52, 1199 – 1218.

Gebel, M. and F. Pfeiffer (2010), Educational Expansion and its Heterogeneous Returns for Wage Workers, *Schmollers Jahrbuch - Journal of Applied Social Science Studies*, 130(1), 19-42.

Hirano, K. and G.W. Imbens, G. W. (2005), The propensity score with continuous treatments, in A. Gelman & X.-L. Meng, eds, *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family*, Wiley, 73–84.

Heckman, J. J., and E. Vytlacil (1998): Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Rate of Return to Schooling When the Return is Correlated with Schooling, *Journal of Human Resources*, 23, 974–987.

Heckman, J. J. and S. Urzua (2010), 'Comparing IV with structural models: What simple IV can and cannot identify', *Journal of Econometrics*, 156, 27 – 37.

Imai, K. and D. A. van Dyk (2004), 'Causal inference with general treatment regimes', *Journal of the American Statistical Association*, 99, 854–866.

Maier, M., F. Pfeiffer, and W. Pohlmeier (2003): Overeducation and Individual Heterogeneity, in: *Overeducation in Europe: Current issues in theory and policy*, ed. by F. Büchel, A. de Grip, and A. Mertens, Edward Elgar, 133-152.

Stock, J. H., J. H. Wright, and M. Yogo (2002): A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments, *Journal of Business and Economic Statistics*, 20, 518–529.

Wooldridge, J. M. (2004): Estimating Average Partial Effects under Conditional Moment Independence Assumption, Centre for Microdata Methods and Practice, Institute for Fiscal Studies, CEMMAP Working Paper CWP03/04

Wooldridge, J. M. (2008): Unobserved Heterogeneity and Estimation of Average Partial Effects, in Advances in Econometrics, Vol 21 (Modeling and Evaluating Treatment Effects in Econometrics), D. Millimet, J. Smith and E. Vytlacil (eds.), Amsterdam: Elsevier, 93-117.

Appendix

Table A1: Summary Statistics of the Covariates

Variable	mean	std. dev.
Age	37.47	8.352
Age squared	1474	629.5
Experience	18.68	9.46
Experience squared	438.4	374.0
Handicapped	.038	.191
Sector		
Manufacturing	.318	.466
Craft	.213	.409
Trade	.103	.303
Public Service	.232	.422
Agriculture	.009	.094
Others	.120	.325
Firm size		
small	.414	.493
medium	.341	.474
big	.229	.420
City size		
small	.360	.480
medium	.276	.447

big	.365		.481
Federal state			
Schleswig Holstein and Lower Saxony	.160		.367
Hamburg and Bremen	.033		.180
North-Rhine Westphalia	.286		.452
Rhineland-Palatinate, Hessen and Saarland	.166		.372
Baden Wuerttemberg and Bavaria	.320		.467
West Berlin	.034		.182
Unemployment ratio	4.16		2.83
Number of observations		7,722	