

Estimating the Economic Return to Schooling on the Basis of Panel Data

Authors	Kalwij,A.S.
Published in	CentER Discussion Paper
Publication Date	1996
Link	https://research.tilburguniversity.edu/en/publications/e9ee1abf-081f-48fc-8525-970de1970498
Citation	Kalwij, A S 1996 'Estimating the Economic Return to Schooling on the Basis of Panel Data' CentER Discussion Paper, vol. 1996-55, Econometrics, Tilburg.
Download Date	2026-05-11 02:09:20
Rights	<p>General rights</p> <p>Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.</p> <ul style="list-style-type: none"> - Users may download and print one copy of any publication from the public portal for the purpose of private study or research. - You may not further distribute the material or use it for any profit-making activity or commercial gain - You may freely distribute the URL identifying the publication in the public portal" <p>Take down policy</p> <p>If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.</p>

Estimating the Economic Return to Schooling on the basis of Panel Data.

Adriaan Kalwij¹

June 1996, Tilburg University²

Abstract

This paper is concerned with estimating the economic return to schooling of men in the Netherlands. We adopt an IV approach to estimate a panel data model with random individual effects. We exploit the fact that older individuals have relatively less schooling compared to younger individuals to construct instruments and include GNP per worker at the time an individual turned 16 to control for birth-cohort effects. The estimated return to schooling is about 15%. Ignoring the endogeneity of schooling results in a lower return to schooling. Ignoring birth-cohort effects results in a lower return to work experience.

¹ The author likes to thank Rob Alessie, David Card, Guido Imbens, Arie Kapteyn and the anonymous referee for their valuable comments and discussions. Financial support has been provided by the NWO.

² Mailing address: Tilburg University, CentER for Economic research, P.O.BOX 90153, 5000 LE Tilburg, The Netherlands. E-mail: A.KALWIJ@KUB.NL

1. Introduction

In this paper we investigate the effect of the outcome of the schooling decision on earnings of an individual in the Netherlands. In the literature this causal effect of years of schooling on earnings is often referred to as the return to schooling. We use the human capital theory as theoretical framework, Becker (1964). To determine the return to schooling we choose a conventional earnings function approach, see for instance Mincer (1974).

The number of years of schooling will typically be one of the explanatory variables of the earnings of an individual. The number of years of schooling (referred to as schooling) is the result of the schooling decision made earlier in life. Ability related characteristics, for instance intelligence and motivation, are likely to play a crucial role in the outcome of the schooling decision. Ability is also likely to influence the earnings of an individual, once controlled for schooling. This, together with the fact that ability is not observed, causes schooling to be a potential endogenous explanatory variable in the earnings function. We control for the endogeneity of schooling by making use of a Generalized Instrumental Variable estimator. For the choice of instruments we exploit the fact that older individuals have relatively less schooling compared to younger individuals. The fact we have data on individuals of different birth-cohorts causes the observed work experience-earnings profile to be downward sloping at higher ages. One possible explanation for this so called birth-cohort effect is the increase in the marginal productivity of labor over time. This caused an increase in starting wages. The fact we have panel data makes it possible to identify both the birth-cohort effect and the effect of work experience on earnings. Conventionally, some function of age is used to control for birth-cohort effects. We, on the other hand, include the Gross National Product (GNP) per worker at the time an individual turned 16 to control for the differences in starting wages³. This macro-economic variable controls for the increase in the marginal productivity of labor. Hence, for the differences in starting wages between individuals of different birth-cohorts. Moreover, this makes it possible to use some flexible function of age as an instrument for schooling.

Most of the recent literature on the return to schooling is surveyed in Card (1994) and mainly covers empirical studies using U.S. data⁴. Instead of using direct measurements of ability, for instance the results of an IQ-test, e.g. Griliches (1977), recent studies make use of Instrumental Variable (IV) techniques to take ability into account. For instance, Angrist and Krueger (1991) use

³ This method was suggested by Kapteyn, Alessie and Lusardi (1995).

⁴ Most of the earlier literature is surveyed by Griliches (1977) and Willis (1986). Returns to education across countries is surveyed by Psacharopoulos (1985).

as an instrument for schooling the quarter of birth of a respondent. Empirical evidence shows that individuals born early in the year have relatively low levels of schooling compared to individuals born later in the year. This is attributed to compulsory schooling laws⁵. Some recent studies on the return to schooling make use of information on twins or siblings to take ability and family background into account. This is not the approach we take in this paper. Therefore we only refer to a discussion on these studies by Miller *et al.* (1995). In studies adopting an IV approach, estimated returns to schooling in the U.S. vary from 8% in Angrist and Krueger up to 15% in Hausman and Taylor (1981). The central finding in almost all empirical studies trying to take ability into account is that ignoring ability results in an underestimation of the return to schooling. Most studies that adopt an IV approach use as an instrument for schooling some variable that is unique to their dataset, e.g. 'quarter of birth' in Angrist and Krueger or 'nearby college in county of residence' in Card (1993). Such a variable we refer to as a unique instrument.

The main contribution of this paper is that we take the endogeneity of schooling into account without making use of a unique instrument. As discussed above, we exploit the fact that older individuals have relatively less schooling compared to younger individuals to construct instruments and include GNP per worker at the time an individual turned 16 to control for birth-cohort effects. In practise this means we need information on the wage rate, the years of schooling, the years of work experience and the age for all individuals. GNP is a macro-economic variable and is reported in the National Accounts. Panel data is necessary to disentangle the effects of GNP per worker at the time an individual turned 16, schooling and work experience on earnings. We estimate a return to schooling of about 15%. Controlling for birth-cohort effects appears to be at least as important as controlling for the endogeneity of schooling. Ignoring the endogeneity of schooling causes an underestimation of the return to schooling. Ignoring birth-cohort effects causes an underestimation of the return to work experience.

The outline of the paper is as follows. Section 2 discusses the theoretical framework and a very simple model of the schooling decision. This provides some insight in the relation between unobservables (including ability) in the wage equation, schooling and earnings. Section 3 describes the econometric framework and the data used for the empirical analyses. Estimation results are in section 4. This section also investigates the quality of the instruments. Section 5 summarizes the main results.

⁵ In principle all persons born in the same calendar year start attending school at the same time. As a result, persons born earlier in the year reach minimum school-leaving age at a lower grade compared to persons born later in the year.

2. Theoretical Framework: The Human Capital Theory ⁶

2.1 Investment in Human Capital and Earnings

The three principal elements of human capital we consider in this paper are initial human capital, schooling investment and post-school investment.

Already at the time of birth individuals are different with respect to intelligence, motivation, the social and economic environment, and other ability related characteristics. In the literature these endowments are often lumped together and referred to as ability. A particular combination of these endowments can be best described by initial human capital. Initial human capital is given at birth. Individuals accumulate human capital during the schooling period and continue to learn and improve on their skills during working life. The human capital theory suggests there is a relation between accumulated human capital and earnings. The more human capital an individual accumulates the higher his or her productivity. The individual receives a wage according to his or her productivity. We loosely formalize this relation between the investments in human capital and earnings as follows:

$$\ln(Y_t) = f(A,S,E_t) + \epsilon_t \quad , \quad t \geq S. \quad (2.1)$$

We denote t as the age of an individual minus the duration of the pre-schooling period. Ability is denoted by A . Y_t denotes some measure of earnings of an individual at time t , for instance hourly wages, after E_t years of work experience and S years of schooling. During the schooling period ($t < S$) the earnings are assumed to be zero. We assume that the first order derivatives of $f(A,S,E_t)$ with respect to the arguments A , S and E_t are all positive. A random income shock with expectation equal to zero is denoted by ϵ_t .

2.2. The Schooling Decision: a Simple Model

In this section we discuss a simple economic model of the schooling decision of an individual to give a rough impression on the relation between ability, earnings and the optimal choice of schooling. This model provides some insights in the direction of the bias that may arise if we use a Least Squares instead of an Instrumental Variable estimator to estimate the economic return to

⁶ Important and useful references are Becker (1964) and Mincer (1974).

schooling. This bias is often referred to as the ability bias. As discussed in the introduction, recent studies find downward biased estimates of the return to schooling. This suggests a negative correlation between the error term (including ability) and schooling, see for instance Griliches (1957). Using a very simple model of schooling decision, we show that it is possible to have a negative relation between ability and schooling. Griliches (1977) already demonstrated this by assuming that some of the costs of schooling are subsidized. We do not intend to produce a realistic schooling decision equation. Therefore the resulting equations should be interpreted with caution.

We assume that an individual maximizes expected life-time utility. An individual's life-time utility is assumed to be the sum of life-time earnings and non-monetary benefits from schooling. At the beginning of his life, the individual makes a schooling decision based on beliefs (expectations) about the future. We assume perfect capital markets. The benefits of one additional year of schooling are the higher earnings during working life and the joy of attending school. This last benefit we refer to as a non-monetary benefit and may depend on ability. The costs of one additional year of schooling are the foregone earnings during that year. We assume an infinite horizon and time constant discount factor denoted by ρ . The earnings at time t are denoted by Y_t and the non-monetary benefits from attending school at time t are denoted by $\phi(A,t)$. At $t=0$ the individual determines the optimal number of years of schooling and starts attending school. In this setting, each individual is assumed to solve the following maximization problem:

$$\max_S E_0 \left\{ \int_0^S \phi(A,t) e^{-\rho t} dt + \int_S^\infty Y_t e^{-\rho t} dt \right\} ;$$

subject to

$$\ln(Y_t) = f(A,S,E_t) + e_t \quad \text{if } t \geq S .$$

There are no earnings during the schooling period ($t < S$) and no non-monetary benefits during working life ($t \geq S$). To guarantee the individual does not decide to stay in school forever, we assume ρ is sufficiently large⁷. The first order condition of this maximization problem is given by

$$E_0 \left\{ \phi(A,S) e^{-\rho S} + \int_S^\infty \frac{dY_t}{dS} e^{-\rho t} dt \right\} = E_0 \left\{ Y_S e^{-\rho S} \right\} .$$

On the left hand side we have the expected marginal benefits of schooling and on the right hand side the expected marginal costs of schooling. The individual chooses to stay in school until the

⁷ Formally we impose the transversality condition $\lim_{S \rightarrow \infty} Y_S e^{-\rho S} = 0$.

expected marginal benefits equalizes the expected marginal costs of one additional year of schooling. Differences in ability across individuals causes the schooling choices to differ across individuals. We consider a special case of equation (2.1). We assume a flat experience-earnings profile and that ability only influences the intercept of log-earnings⁸. In this case, the wage equation is given by

$$\ln(Y_t) = \alpha(A) + \beta S + \epsilon_t \quad .$$

We assume that the non-monetary benefits are the same for all individuals and can be written as $\phi e^{-\gamma t}$. Furthermore we assume that there are decreasing marginal non-monetary benefits over time ($\gamma \geq 0$). Solving the first order condition for S yields the optimal years of schooling, denoted by S_{opt} :

$$S_{opt} = \frac{1}{\beta + \gamma} [\ln(\phi) + \ln\left(\frac{\rho}{\rho - \beta}\right) - \alpha(A) - \ln(E_0\{e^{\epsilon_t}\})] \quad . \quad (2.2)$$

This equation shows the relation between ability and the optimal years of schooling. In the case that ability is not observed the econometrician has to include $\alpha(A)$ in the error term of the wage equation which is used to estimate the effect of schooling on earnings. In this case the individual specific part of the error term in the wage equation equals $\alpha(A)$. The individual specific part of the error term is negatively correlated with the optimal years of schooling and positively correlated with earnings. Standard econometric theory tells us that in this case one can expect a downward biased Least Squares estimate of the effect of schooling on earnings (see e.g. Griliches (1957)). The last term at the right hand side of equation (2.2) gets in because of uncertainty of future wages. Hereby we assumed that the ϵ_t 's are independently distributed over time. For instance, if ϵ_t is $N(0, \sigma^2)$ distributed then $E_0\{\exp(\epsilon_t)\} = \exp(1/2\sigma^2)$. The larger is σ^2 the smaller is the optimal years of schooling.

This simple theoretical exercise shows that if individuals obtain utility from attending school then it is possible to find a downward bias in the return to schooling when using a Least Squares estimator. However, changing one of the assumptions of the model, for instance making ϕ dependent on A, may lead to different results. Therefore, it is a priori not clear in which direction the Least Squares estimate of the return to schooling is biased and empirical evidence is necessary.

⁸ Assuming earnings increase with years of work experience does not alter the main conclusion of this section. Also one can argue that the discount factor (ρ) and slope parameter of the effect of schooling on log-earnings (β) differ across individuals. An extension in this direction is given by Card (1994). Almost all empirical studies, however, do not allow for this because ability is not observed. Usually one estimates a log-earnings equation linear in schooling, work experience and work experience squared.

3. Empirical Specification

To determine the causal effect of schooling on earnings we need to parameterize the wage function (equation 2.1). To be able to compare our results with the findings in the studies mentioned in the introduction, we assume a conventional Mincerian wage equation.⁹ Section 3.1 discusses the econometric framework and the estimation procedure. Section 3.2 discusses the data used for estimation.

3.1 A Panel Data Model with Random Individual Effects

For the empirical analysis we use a panel data model with random individual effects, see for instance Hsiao (1986). We have N individuals and T time periods, indexed by respectively i and t . Schooling is assumed to be an endogenous time constant regressor and work experience is by construction, an endogenous time varying regressor¹⁰. The error term, denoted by ϵ_{it} , consists of two parts, a random individual specific effect, denoted by α_i , and a random effect, denoted by η_{it} . We allow the intercept to vary across time periods (α_t). We include the variable C_i to control for differences in starting wages between individuals of different birth-cohorts. S_i denotes the number of years of schooling and E_{it} the number of years of work experience in period t . We refer to this model as the cohort-schooling model:

$$(3.1) \quad \ln(Y_{it}) = \alpha + \alpha_t + \theta C_i + \beta S_i + \gamma_1 E_{it} + \gamma_2 E_{it}^2 + \epsilon_{it} \quad ,$$

$$\epsilon_{it} = \alpha_i + \eta_{it} \quad ,$$

with the stochastic specification:

$$E\{\epsilon_{it}|C_i\} = 0 \quad ; \quad E\{\epsilon_{it}|S_i, E_{it}\} \neq 0 \quad ;$$

$$E\{\epsilon\epsilon^T\} = \sigma_\alpha^2(\mathbf{1}_T \mathbf{1}'_T \otimes I_N) + \sigma_\eta^2(I_T \otimes I_N) \quad .$$

$\mathbf{1}_T$ is a $(T \times 1)$ -vector containing ones and I_N (I_T) is the identity matrix of rank N (T) and ϵ is a $(NT \times 1)$ -vector. We make no distributional assumptions concerning the error term. As discussed in

⁹ This means we use a wage equation where the logarithm of hourly wage is linear in the years of schooling. This linearity assumption has been tested by Harmon and Walker (1996). Their preliminary results indicate that linearity is not rejected by the data.

¹⁰ When comparing two individuals of the same age, the one with the most years of schooling has the least work experience.

the previous sections, schooling is correlated with ability (included in α_i). For this reason schooling is a potentially endogenous explanatory variable. A consistent estimate is obtained by using an Instrumental Variable (IV) estimator. We take the variance-covariance structure into account to get more efficient estimators. We refer to this estimator as the Estimated Generalized Instrumental Variables (EGIV) estimator. We follow Magnus (1978) and use a two-step procedure in order to get consistent estimates of σ_α^2 and σ_η^2 .¹¹

We assume that birth-cohort effects arise only due to differences in starting wages between individuals of different birth-cohorts. We control for these birth-cohort effects by including C_i in the wage equation. If we include a dummy variable for each year of birth to control for birth-cohort effects then we have an identification problem. We get around this identification problem by including GNP per worker at the time an individual turned 16 to control for birth-cohort effects. This macro-economic variable controls for the increase in the marginal productivity of labor. Hence, for the differences in starting wages between individuals of different birth-cohorts. In the next section we discuss this in somewhat more detail. Panel data is necessary to disentangle the effects of GNP per worker at the time an individual turned 16, schooling and work experience on earnings. Notice that both GNP per worker at the time an individual turned 16 and schooling are time constant regressors and work experience is a time-varying regressor.

3.2 Data: the Socio-Economic Panel

The micro data we use for the empirical analyses are data from the Socio-Economic Panel (SEP). About 5000 households participate in this survey. The survey is conducted twice a year by the Central Bureau of Statistics in the Netherlands, a wave in April and a wave in October. Only the October waves have detailed information on earnings. There can be more than one respondent¹² per household. To each respondent questions about their socio-economic and demographic situations are asked. We use the October waves of the years 1986, 1987, 1988 and 1989. We select men between 16 and 60 years of age who are working full-time in 1986 and kept on working full-time during the following years. We use a balanced panel and panel attrition is about 16% per year. This leaves us with 1398 observations in each of the four waves. Table 1 reports the sample statistics of the relevant variables.

¹¹ The EGIV estimator and the two-step estimation procedure are described in appendix C.

¹² A respondent is a person at least 16 years old. In principle each person in the household over 15 should complete the questionnaire.

Differences in gross wages between individuals are assumed to be a better measurement of productivity differences than differences in net wages. Therefore we use gross hourly wages uncorrected for inflation. During the period 1986-1989 inflation has been very low¹³. In the SEP the net wages and the working hours per week are asked. Therefore gross hourly wages is a constructed variable¹⁴. The date of schooling completion and the starting date of the first paid job are explicitly asked to the respondent. The years of schooling are the years spent at school until the respondent obtained his highest level of education¹⁵. Work experience is the number of years of employment since the respondent started with his first paid job¹⁶. We regard the years of schooling during working life as post-school investment, i.e. work experience.

Figure 1 shows the Gross National Product (GNP) per worker per year over the period 1941-1986. GNP per worker increased rapidly during the second half of this century. As discussed in the introduction, a possible explanation for the birth-cohort effect in the work experience-earnings profile is the increase in the starting wages due to an increase in the marginal productivity of labor over time. As suggested by Kapteyn *et al.* (1995), we use changes in the logarithm of the GNP per worker over time as an approximation for the changes in the marginal productivity of labor over time. To be more specific, we include the logarithm of GNP per worker at the time an individual turned 16 as an explanatory variable in the wage equation¹⁷. We refer to this variable as the GNP per worker.

Figure 2 shows a non-parametric relation between schooling and age. This non-parametric relation is based on a Kernel regression and we use uniform confidence bands. The sharp increase in schooling at the younger ages is caused by the fact that higher educated individuals complete schooling at a later age. Therefore, in this range, the less educated are over represented. We have only a few individuals over 55 years of age. More than 90% of the individuals in the sample are between 25 and 55 years of age. In this range there is a decrease in schooling with age. One possible explanation for this is the reduction in the cost of schooling over time. Therefore, we conclude that older individuals possibly have relatively less schooling compared to younger

¹³ The Consumption Price Index in the Netherlands increased by 1.4% over the period 1986-1989.

¹⁴ Gross wages are a non-linear function of net wages, taken into account all the relevant characteristics of the households and individuals, such as number of children and mortgage interest. However, the most important non-linearity is caused by a tax exemption for earnings. For instance, the exemption is f11123,- for a single person household in 1989.

¹⁵ We assume that all individuals start attending school in the year they turned 6.

¹⁶ A job does not include a vacation job, military service (compulsory conscription) alternative national service and 'occasional' work.

¹⁷ The difference with our method is that they used GNP per inhabitant at the time an individual turned 22.

individuals. An important implication of this relation is that some function of age may be a suitable instrument for schooling.

4. Estimation Results

We estimate the cohort-schooling model as discussed in section 3.1. The dependent variable is the logarithm of gross hourly wages and the explanatory variables are a set of year dummy variables, the logarithm of GNP per worker, schooling, work experience and work experience squared. The latter three variables are assumed to be endogenous. The instruments are the exogenous explanatory variables and a fifth order polynomial in age. The estimation results of the cohort-schooling model are reported in table 2. Before discussing these results we discuss some specification tests of the cohort-schooling model.

The Hausman test-statistic for the endogeneity of schooling and work experience, Hausman (1978), is equal to 11.9¹⁸. This means we reject the null-hypothesis of exogeneity of schooling and work experience. The Sargan test-statistic for the over-identifying restrictions, Sargan (1958) is equal to 0.021¹⁹. This means we do not reject the null-hypothesis that the orthogonality conditions of the instruments hold. IV estimates are often based on a low correlation between the excluding instruments and the endogenous explanatory variables. Our EGIV estimates are no exception on this (see appendix B). For this reason IV approaches to estimate the return to schooling have been criticized recently. Bound *et al.* (1995) show that if there is even a weak correlation between the instruments and the error term then the finite sample bias of the IV estimates approaches that of the OLS estimates as the R^2 between potential endogenous explanatory variable and the instruments approaches 0. They suggest that when IV estimates are reported both the partial R^2 and the F-statistic on the excluded instruments in the first stage regression should be reported as well. These are rough guides to the quality of the IV estimates. The results of the first stage regressions are reported in appendix B. The F-statistic on the excluded instruments for schooling is equal to 11²⁰. The corresponding partial R^2 equals 0.010. The F-statistic on the excluded instruments for work experience is equal to 129. The corresponding partial R^2 equals 0.10. Although both of these compare favorably with those reported in other studies adopting an IV approach (see for instance

¹⁸ The null-hypothesis is exogeneity of schooling and work experience. The critical value is $\chi_{0.95}(3) = 7.8$.

¹⁹ The null-hypothesis is that the orthogonality conditions of the instruments hold. Critical value is $\chi_{0.95}(7) = 14.1$.

²⁰ The null-hypothesis is that the estimated coefficients of the excluding instruments are equal to 0. The critical value is $F_{0.95}(5,5587) = 4.36$.

Bound *et al.*, Harmon and Walker (1995) or Lusardi (1996)) intuitively the partial R^2 in the schooling equation seems rather small. Another issue is the fact that we have more than one endogenous variables. Therefore it is less clear what the partial R^2 's of first stage regressions are telling us about the overall performance of the IV estimator. A simultaneous test on the quality of instruments is provided in Bekker (1994). He gives a quality measure for the instruments and a lower bound on the quality of instruments²¹. We report the quality measure in Appendix B. The quality measure equals 0.0051 and is higher than the lower bound on this quality measure. Based on these results we conclude that the quality of the instruments are sufficient.

Next, we turn to the estimation results of the cohort-schooling model (table 2). Controlling for the endogeneity of schooling results in a return to schooling of approximately 15%. Assuming schooling is an exogenous explanatory variable (the EGLS estimates) results in an estimate equal to 6.9%. This provides some empirical evidence of a downward ability-bias of about 8% points in case we do not control for endogeneity. The direction of the bias is in line with the findings in recent studies using U.S. or U.K. data. The maximum of the work experience-earnings profile is calculated at 49 years of work experience. The return to work experience for an individual who just completed schooling is equal to 6.8%. GNP per worker has a significant effect on wages. A 1% increase in the GNP per worker results in a 0.24% increase in starting wages. The dummy variables for each year of the panel are not significantly different from zero.

Appendix A reports the estimates without controlling for differences in starting wages. This results in an underestimation of the return to work experience²². The return to work experience for an individual who just completed schooling is equal to 4.1%. Also in this case we find that we underestimate the return to schooling when not controlling for the endogeneity of schooling.

Estimates of the return to schooling in previous studies using data of the Netherlands are comparable with EGLS estimates reported in appendix A. For instance, Hartog *et al.* (1993) provide a survey of estimated returns to schooling in the Netherlands over the years 1962-1989. The return to schooling drops from 11% in 1962 to 5% in 1986 and is slightly increasing thereafter. These results are based on a Mincerian wage equation using cross-section data and a Least Squares estimator. Using this specification we estimate a return to schooling of 5.9% (see appendix A). In a discussion on whether or not there is a situation of over-investment in schooling in the Netherlands, Theeuwes (1993) argues that the estimated returns to schooling in the Netherlands are presumably upward biased. The arguments are based on a positive relation between

²¹ Because we are not aware of any application using this quality measure we extent on this in appendix D.

²² In this case, the maximum of the work experience-earnings profile is calculated at 45 years of work experience.

ability and both schooling and earnings. In contrast with these theoretical predictions, there is an increasing number of papers that provide empirical evidence on downward biased Least Squares estimates. This has been discussed in the introduction and in section 2.2 we provided a possible theoretical explanation for this. Also our empirical results indicate that the Least Squares estimate of the return to schooling is downward biased in the Netherlands. Comparing the results of the EGIV estimates of the cohort-schooling model with the EGLS estimates the schooling model shows the net effect of controlling for both the endogeneity of schooling and birth-cohort effects. This net effect on the return to schooling is equal to 8.9% points (14.8% versus 5.9%).

5. Conclusions

We used the conventional Mincerian approach to estimate the economic return to schooling for men in the Netherlands. We adopted an IV approach to estimate a panel data model with random individual effects. We did not make use of a unique instrument for schooling but exploited the negative relation between age and schooling to construct instruments and included GNP per worker at the time an individual turned 16 to control for birth-cohort effects. We paid special attention to the quality of the instruments. We needed panel data to disentangle the effects of GNP per worker at the time an individual turned 16, schooling and work experience on earnings.

The estimated economic return to schooling is equal to 15%. This estimated return is considerably larger than the returns estimated in previous studies using data from the Netherlands. Ignoring the endogeneity of schooling causes an underestimation in the return to schooling. This result is in line with the findings in recent studies using U.S. data. The quality of the instruments was shown to be sufficient. Furthermore, we found that controlling for birth-cohort effects is at least of as much importance as controlling for the endogeneity of schooling. Ignoring birth-cohort effects causes an underestimation of the return to work experience. This shows the importance of using a panel data model and therefore be able to control for both the endogeneity of schooling and birth-cohort effects.

References

- Angrist, J.D., and A.B.Krueger (1991): "Does compulsory school attendance affect schooling and earnings?", *The Quarterly Journal of Economics*, November, 979-1014.
- Becker, G. (1964): "Human Capital: a theoretical and empirical analysis, with special references to education", *Columbia University Press*, New York.
- Bekker, P.A. (1994): "Alternative Approximations to the Distributions of Instrumental Variables Estimators", *Econometrica*, 62, 657-681.
- Bound, J., D.A. Jaeger and R.M. Baker (1995): "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak", *Journal of the American Statistical Association*, 90, 443-450.
- Card, D. (1993): "Using Geographic Variation in College Proximity to Estimate the Return to Schooling", *NBER Working Paper*, no. 4483, Cambridge Ma.
- Card, D. (1994): "Earnings, schooling and ability revisited", *NBER working paper*, no. 4832, Cambridge Ma.
- Griliches, Zvi (1957): "Specification bias in estimates of production functions", *Journal of farm economics*, 39, 8-20.
- Griliches, Zvi (1977): "Estimating the returns to schooling: some econometric problems", *Econometrica*, 45, 1-22.
- Härdle, W. (1990): "Applied Nonparametric Regression", *Cambridge University Press*, New York.
- Harmon, C. and I. Walker (1995): "Estimates of the Economic Return to Schooling for the UK", *American Economic Review*, 85, 1278-1286.
- Harmon, C. and I. Walker (1996): "The Marginal and Average Returns to Schooling", *working paper*, Maynooth College, Ireland.
- Hartog, J., Oosterbeek, H and C. Teulings (1993): "Age, wages and education in the Netherlands". In: P.Johnson and K. Zimmermann (eds), *Labor markets in an aging Europe*, Cambridge University Press, Cambridge, 182-211.
- Hausman, J.A. (1978): "Specification tests in econometrics", *Econometrica*, 46, 1251-1271.
- Hausman, J.A. and W.E. Taylor (1981): "Panel data and unobservable individual effects", *Econometrica*, 49, 1377-1398.
- Hsiao, C. (1986): "Analysis of Panel Data", *Cambridge University Press*, Cambridge.
- Kapteyn, A., Alessie, R.J.M. and A. Lusardi (1995): "The Wealth Accumulation of Cohorts: A Panel Analysis", *CentER discussion paper*, Tilburg University.

- Lusardi, A. (1996): "Permanent Income, Current Income and Consumption: Evidence from Panel Data", *Journal of Business and Economic Statistics*, 14, no.1, 81-90.
- Magnus, J.R. (1978): "Maximum likelihood estimation of the gls model with unknown parameters in the disturbance covariance matrix", *Journal of Econometrics*, 7, 281-312.
- Mincer, J. (1974): "schooling, experience, and earnings", *Columbia University Press*, New York.
- Miller, P., Mulvey, C., and N. Martin (1995): "What do Twin Studies Reveal About the Economic Returns to Education? A Comparison of Australian and U.S. Findings", *The American Economic Review*, Vol.85, no.3, 586-599.
- Pesaran, M.H and R.J. Smith (1994); "A generalized R^2 criterion for regression models estimated by the instrumental variable method", *Econometrica*, 62, 705-710.
- Psacharopoulos, G. (1985): "Returns to Education: A Further International Update and Implications", *Journal of Human Resources*, 20, 583-604.
- Sargan, G.D. (1958): "The estimation of economic relationships using instrumental variables", *Econometrica*, 26, 393-415.
- Statistics Netherlands (1989): "1899-1989, negentig jaren statistiek in tijdreeksen", Voorburg/Heerlen.
- Theeuwes, J.J.M. (1993): "Het rendement van scholing en training", *Economisch Statistische Berichten*, 845-859.
- Willis, R.J. (1986): "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings functions", In Orley Ashenfelter and Richard Layard, editors, *Handbook of Labour Economics*, North Holland, New York NY, 525-602.

Appendix A: Estimation results of the schooling model. No control for differences in starting wages. Dependent variable is the logarithm of gross hourly wages. The number of observations is equal to 5592 (4x1398). Standard errors are in parenthesis.

Estimator	EGLS	EGIV
Variable	Coefficient	Coefficient
constant	1.939 (0.034)	0.962 (0.324)
1987	0.031 (0.010)	0.028 (0.016)
1988	0.050 (0.010)	0.050 (0.016)
1989	0.071 (0.010)	0.071 (0.017)
schooling	0.059 (0.002)	0.144 (0.028)
experience	0.041 (0.002)	0.036 (0.004)
experience squared	-0.0006 (5×10^{-5})	-0.0004 (1×10^{-4})
s_{η}^2	0.07	0.18
s_{α}^2	0.03	0.03
R^2	0.31	0.14
Hausman test $\chi_{0.95}(3)=7.8$		10.2
Sargan test $\chi_{0.95}(7)=14.1$		0.06

Appendix B: Estimation results of the first stage regressions of the IV estimator in table 2. Regressions of schooling, work experience and work experience squared on a fifth order polynomial in age, time-dummy variables and ln(GNP). The number of observations is equal to 5592 (4x1398). Standard errors are in parenthesis.

Dependent Variable: ----- Explanatory Variables	Schooling coefficient	Work Experience Coefficient	(Work Experience) ² Coefficient
Constant	-32.1 (34.4)	29.4 (35.4)	263 (160)
Age	4.67 (5.06)	-4.14 (5.21)	-42.6 (23.4)
Age ² /10	-2.40 (2.77)	2.72 (2.85)	24.1 (12.9)
Age ³ /100	0.66 (0.74)	-0.76 (0.76)	-6.65 (3.43)
Age ⁴ /1000	-0.09 (0.10)	0.11 (0.10)	0.94 (0.45)
Age ⁵ /10000	0.01 (0.01)	-0.01 (0.01)	-0.05 (0.02)
1987	-0.07 (0.19)	0.07 (0.20)	-0.58 (0.88)
1988	-0.13 (0.29)	0.14 (0.30)	-1.17 (1.35)
1989	-0.19 (0.41)	0.19 (0.42)	-1.84 (1.89)
ln(GNP)	0.92 (1.35)	-1.08 (1.39)	3.90 (6.28)
R ²	0.020	0.86	0.85
Partial* R ²	0.010	0.10	0.39
Partial F-test, F(5,5578)=4.31	11.0	129	720
Bekker's quality measure	0.0051		

* The partial R² and the partial F-test are based on the excluding instruments age, age², age³, age⁴ and age⁵.

** A lower bound for the quality measure is given by K/NT = 0.0009 (see appendix D).

Appendix C: Estimation Procedure.

In this appendix we briefly discuss the EGIV estimator and the estimation procedure. We have N individuals and T time periods. Let Y denote the dependent variable, a (NTx1)-vector, and X the explanatory variables, a (NTxG)-matrix). We can write our model of section 3.1 as follows:

$$Y = X\beta + \epsilon \quad ;$$

$$E\{X^T\epsilon\} \neq 0 \quad , \quad E\{\epsilon\epsilon^T\} = \sigma_\alpha^2(\mathbf{1}_T\mathbf{1}_T^T \otimes I_N) + \sigma_\eta^2(I_T \otimes I_N) \quad .$$

I_N (I_T) is the identity matrix of rank N (T), and $\mathbf{1}_T$ is a Tx1 vector containing ones. The set of instruments is denoted by Z, a (NTxK)-matrix, with $E\{\epsilon^T Z\} = 0$. A necessary condition is that $K \geq G$. The derivation is based on Sargan (1958). The estimate of β is denoted by b_{GIV} and is chosen in such a way that it minimizes the weighted sum of the residuals:

$$b_{GIV} = \min_\beta E\{(Y - X^T\beta)^T Z W Z^T (Y - X\beta)\} \quad ,$$

with

$$W^{-1} = E\{Z^T \epsilon \epsilon^T Z\} \quad .$$

Using the first order condition, we obtain:

$$b_{GIV} = (X^T Z W Z^T X)^{-1} (X^T Z W Z^T Y)$$

Using the variance-covariance structure we can rewrite W^{-1} as follows:

$$W^{-1} = T\sigma_\alpha^2 Z^T L Z + \sigma_\eta^2 Z^T Z \quad .$$

The L operator is defined as $L = (\mathbf{1}_T \mathbf{1}_T^T \otimes I_N) / T$. The variance-covariance matrix of the estimate b_{GIV} is given by:

$$\text{var}(b_{GIV}) = (X^T Z W Z^T X)^{-1} \quad .$$

In our case, σ_η^2 and σ_α^2 are unknown and we have to be replaced them by consistent estimates, denoted by s_η^2 and s_α^2 . For this purpose we use a two-step estimation procedure and follow Magnus (1978):

First step : The IV estimator gives a consistent estimate of β , denoted by b_{IV} . This means we assume $E\{\epsilon\epsilon^T\} = \sigma^2 I_{NT}$.

Second step : We calculate the residuals $U_{IV} = Y - X b_{IV}$ and carry out the transformations LU_{IV} and LU_{IV} , where $M = I_{NT} - L$.
A consistent estimate of σ_η^2 is given by $s_\eta^2 = MU_{IV}' MU_{IV} / N$ and a consistent estimate of σ_α^2 is given by $s_\alpha^2 = LU_{IV}' LU_{IV} / N - s_\eta^2 / T$.

Substituting s_η^2 and s_α^2 for σ_η^2 and σ_α^2 results in a consistent estimate of W, say W^* . Substituting W^* in (A1) and (A3) gives $b_{EGIV} = (X' Z W^* Z' X)^{-1} X' Z W^* Z' Y$ and $\text{var}(b_{EGIV}) = (X' Z W^* Z' X)^{-1}$. We refer to this estimator as the EGIV estimator. In the special case that Z equals X, this estimator reduces to the well-known EGLS estimator.

Appendix D: the quality of the instruments.

In this appendix we give a quality measure for the instruments and a lower bound for this quality measure, as proposed by Bekker (1994). We refer to Bekker for a more thorough discussion on this and for an excellent discussion on the finite sample asymptotics of the LS and IV estimator in relation to the quality of the instruments.

We consider an equation with G endogenous explanatory variable, denoted by X , K_1 exogenous explanatory variables, denoted by Z_1 , and an additional set of instruments of rank K_2 , denoted by Z_2 . Y denotes the dependent variable. The instrument set is $Z=(Z_1, Z_2)$ and $K=K_1+K_2$. We write this equation as follows:

$$(D1) \quad Y = X\beta + Z_1\gamma + \epsilon \quad .$$

where $E(X'\epsilon) \neq 0$, $E(Z'\epsilon)=0$ and $\text{rank}(Z) = K$. We have NT observations. So Y is a $(NT \times 1)$ -vector, X a $(NT \times G)$ -vector and Z a $(NT \times K)$ -matrix. The relation between the instruments and the endogenous variables X is given by:

$$(D2) \quad X = Z\delta + R \quad .$$

$D2$ is a system of G simultaneous equations. It is assumed that $E(Z'R)=0$. We define $P_Z = Z(Z'Z)^{-1}Z'$ and $P_{Z_1} = Z_1(Z_1'Z_1)^{-1}Z_1'$. The starting point in Bekker's paper is an inequality that says, quote, "the instruments do a better job than do arbitrary variables in explaining $E(X)$ ". This is formalized by the inequality:

$$(D3) \quad \frac{E'(X)(P_Z - P_{Z_1})E(X)}{K} > \frac{E'(X)E(X)}{NT} \quad .$$

Departing from this inequality one can derive the quality measure:

$$(D4) \quad q = \min_l \left\{ \left(1 + \frac{l'R'l}{l'\delta'Z_2'(P_Z - P_{Z_1})Z_2\delta l} \right)^{-1} \right\} \quad .$$

A lower and upper bound of this quality measure is given by:

$$(D5) \quad \alpha \leq q \leq 1 \quad \text{with} \quad 0 \leq \alpha \leq K/NT \quad .$$

Bekker shows how the asymptotic bias of the LS and IV estimators can be related to the quality of instruments in a finite sample. Note when studying large-sample asymptotics it is assumed that $\alpha = 0$. In this case the IV estimate is unbiased. However, in case $\alpha > 0$ the bias of the IV estimator is shown to increase when the quality of the instruments decreases. The quality measure depends on the explanatory power of the excluding instruments, as one may expect. The lower bound on the quality of the instruments increases in the number of instruments. In appendix B we report the quality measure as proposed by Bekker, q , and the lower bound which is given by K/NT .

Table 1: Sample Statistics (years 1986 up to and including 1989, 1398 observations per year).

Variable	Mean	Std.Dev.	Minimum	Maximum
Gross hourly wage ¹⁾	28	17	5	250
Schooling ²⁾	11.6	3.8	5	27
Work experience ²⁾	20.0	9.9	0	41
Age ²⁾	37.8	8.7	17	60

Source: Socio-Economic Panel

1) measured in Dutch Guilders, no correction for inflation

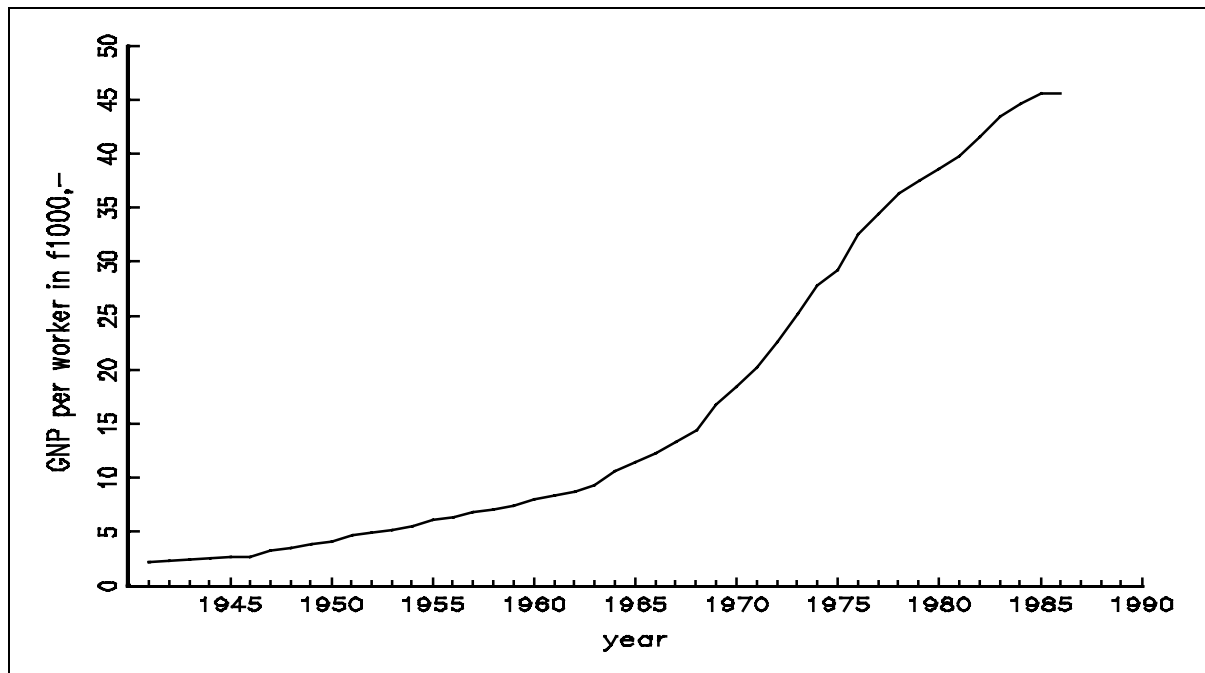
2) measured in years

Table 2: Estimation results of the cohort-schooling model. Dependent variable is the logarithm of gross hourly wages. The number of observations is equal to 5592 (4x1398). Standard errors are in parenthesis.

Estimator	EGLS	EGIV
Variable	Coefficient	Coefficient
1987	0.021 (0.012)	0.008 (0.017)
1988	0.038 (0.015)	0.010 (0.024)
1989	0.053 (0.019)	0.012 (0.031)
schooling	0.069 (0.006)	0.148 (0.025)
experience	0.052 (0.007)	0.068 (0.014)
(experience) ²	-0.0006 (6x10 ⁻⁵)	-0.0006 (1x10 ⁻⁴)
ln(GNP per worker)	0.101 (0.058)	0.239 (0.100)
constant	1.122 (0.472)	-0.805 (0.814)
s_{η}^2	0.07	0.14
$s_{\alpha 2}$	0.03	0.03
R^2	0.31	0.14
Hausman-test $\chi_{0.95}(3)=7.8$		11.9
Sargan-test $\chi_{0.95}(3)=14.1$		0.021

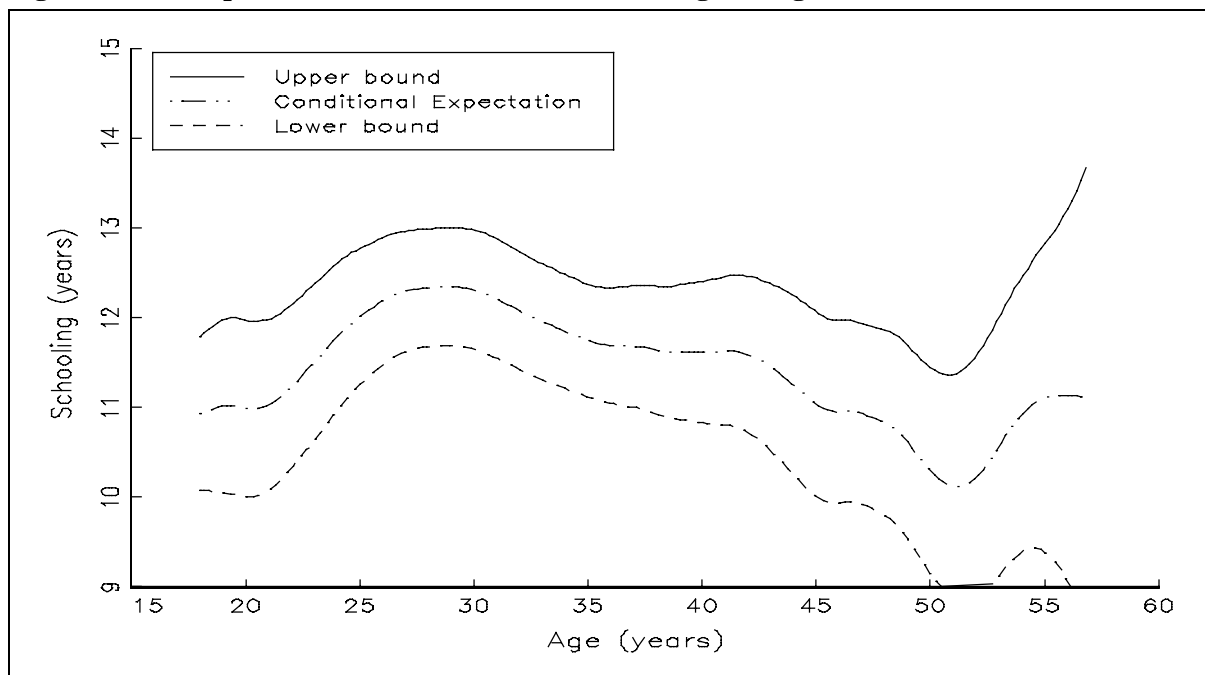
* The R^2 is based on the second step of the IV estimation procedure, e.g. Pesaran and Smith (1994).** The excluding instruments are age, age², age³, age⁴ and age⁵. The results of the first stage regression are reported in appendix B.

Figure 1: Gross National Product (GNP) per worker per year during the period 1941-1986. GNP is deflated with consumption price index (base year 1985).



Source: Statistics Netherlands (1989).

Figure 2: A non-parametric relation between schooling and age *.



* We use a Quartic Kernel and uniform confidence bands (Härdle, 1990). The bandwidth is set to 4 years and we use only the year 1986. This leaves us with 1398 observations.