The evolution of wage structures in Portugal 1982-1992
Cabral Vieira, J.A.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 4

Returns to Education: a comparison of alternative estimators

4.1 Introduction

There is a long-running debate in empirical labour economics regarding the interpretation of the positive association between education and wages. This association has been used to estimate the rate of return to education through the human capital earnings function proposed by Mincer (1974). Typically, empirical applications of this function have relied on the ordinary least squares (OLS) estimator. However, it has been argued that OLS estimates of the rates of return to education may be biased because they are unable to isolate the contribution of schooling from the contribution of an individual’s (unobserved) ability to earn. A related problem is the recognition that schooling is a potentially endogenous explanatory variable in the wage function. Moreover, it has been argued that measuring education with error might be a serious problem. Each of these problems violates the OLS assumption of a zero-correlation between the error term and regressors.

Several studies have tried to correct for the potential bias in the estimated returns to education by using either instrumental variables, fixed-effects and/or two-step methods. Most of the studies have revealed that the return to education is as high or higher than the OLS estimate. Despite this similarity, a lively international debate on the issue persists. The objective of this chapter is twofold. We first intend to shed further light on the aforementioned debate. And second, we will examine the situation in Portugal. Existing studies for Portugal rely on OLS estimates and although most do mention the possibility of a bias, insufficient data has precluded the use of alternative methods.

The chapter compares OLS, IV and a two-step method estimates of the return to education in Portugal. Identification in the wage equation is achieved by considering exogenous variations in individual’s attained education brought about by changes in the compulsory level of education.

The structure of the chapter is as follows. Section 4.2 gives an overview of the literature. Section 4.3 provides a short description of changes in compulsory schooling and educational attainment in Portugal. Section 4.4 outlines the estimation methods. Section 4.5 presents the data set and the estimation results. Section 4.6 offers an interpretation. Finally, section 4.7 concludes.

4.2 Overview of the literature

As suggested above, OLS estimates of the return to education may be biased. First, attained education depends on educational decisions. In a model of optimal investment in schooling, we
would expect a positive correlation between education and its return. Second, a bias may arise from unobserved - and therefore not controlled for in the conventional OLS regression - ability which is correlated with education and wages. Third, the bias may originate from measurement errors in schooling. Finally, OLS estimates may be subject to discount rate bias because individuals with a high discount rate choose less education (Card, 1994). It is normally reported that the first source would bias the OLS estimate upward. The same would occur with respect to the omission of ability (Griliches, 1977, Blackburn and Neumark, 1995). However, Griliches (1977) also argues that in certain situations the OLS bias due to omitted ability can be downward. With regard to the third and fourth sources, OLS would yield a downward biased estimate (Griliches, 1977, Card, 1994).

The accuracy of OLS returns to schooling has recently been the object of great interest. Several authors have attempted to correct for the potential bias in the OLS estimate of that parameter. Their studies have used one (or more) of four econometric approaches. We will give a short overview of this literature below.

Several studies have added some type of ability measure to the wage equation regressors' list to eliminate ability bias (Griliches, 1977, Blackburn and Neumark, 1993, 1995, Uusitalo, 1996). Typically, the inclusion of measures of ability reduces the return to education. This upholds the idea that ability also determines wages and correlates positively with an individual's education. In such a case, omitting ability overstates the rate of return estimate.

A second approach uses instrumental variables (IV) that exploit variations in education caused by exogenous influences. Several studies have used very ingenious mechanisms to instrument education. Angrist and Krueger (1991) use the individual’s season of birth. This variable is related to educational attainment because of school starting age policy and compulsory attendance laws. The IV estimate of the return to education is 28% larger than the OLS counterpart. Angrist and Krueger (1992) utilise a lottery number assigned during the Vietnam draft era. The idea behind the choice of this instrument is that the draft lottery led to higher college enrollment rates because enrolled students could obtain draft exemptions. The IV estimate is 10% larger than the corresponding OLS. Butcher and Case (1993) use an indicator variable on the presence of sisters. The choice of this instrument is based on a previous analysis of educational attainment that led them to conclude that women with one or more sisters have less education than women from the same sized families who have only brothers. The IV estimate of the returns to education is approximately 100% above the OLS.

Card (1993) utilises an indicator variable of a nearby college. This instrument is justified by the fact that men who grow up near a four-year college attend significantly more education. He concludes that the IV estimate is 80% higher than the OLS counterpart. Harmon and Walker (1995) explore changes in the minimum schooling leaving age in the U.K. The IV return to education is twice as large as the corresponding OLS estimate. Uusitalo (1996) utilises family background variables for Finland. The IV is 60% larger than the OLS estimate. In a study for

---

23 Unless otherwise stated, the results summarised in this section use U.S. data.
France, Boumahdi and Plassard (1992) use the education of the parents (in dummy variables) and the number of brothers and sisters in the family at the time the individual left school. The exogeneity of education was rejected. The IV estimate for the return to education is 30% larger than the OLS counterpart.

A third approach uses an extension of the two-step method developed by Heekman (1979). Harmon and Walker (1995) used this type of method for the U.K., and Uusitalo (1996) used it for Finland. The method consists of estimating a first stage equation for schooling attainment (by an ordered probit to account for the ordinal nature of schooling). This allows us to calculate the truncated means, which are added in a second stage to the wage equation list of covariates. This type of estimation normally makes use of at least one identifying variable. That is a variable that affects education, but can legitimately be removed from the wage equation.24 Harmon and Walker (1995) and Uusitalo (1996) utilise as identifiers the same variables they use as instruments. In both of these studies, two-step method estimates are higher than the corresponding OLS. In the case of Harmon and Walker (1995), it is twice as large.

A fourth approach uses a fixed effect estimator (FE). The studies have used panel data for either twins, siblings, or observations of the same individuals at two points in time. The idea behind this method is that it is possible to “difference-out” a person’s ability and/or family background by first differentiating the wage equation. Angrist and Newey (1991) use repeated observations of the same individuals over time. Ashenfelter and Zimmerman (1993) use data on pairs of brothers. Ashenfelter and Krueger (1994) use a data set for pairs of American twins. Miller et al (1995) use a data set for pairs of Australian twins. In all the studies except one, the FE estimates of the return to education are larger than the OLS. The exception is Ashenfelter and Zimmerman (1993), where the FE is below the OLS estimate; furthermore, the study by Miller et al (1995) produces mixed evidence. In this case, the FE estimate is lower than the OLS counterpart once measurement error in the schooling variable is not taken into account. However, after correcting for measurement error in self-reported education levels, the FE estimate of the return to education increases considerably and often exceeds the OLS estimate.

Two groups of results can be singled out when they are compared with the OLS baseline. First, the simple inclusion of some measure of ability to the list of regressors in the wage equation seems to reduce the (OLS) estimate of the rate of return to education. This is in accordance with conventional wisdom. Second, the use of either IV, fixed effects, or two-step methods tends to produce estimates that exceed the OLS estimates. Although widely reported in empirical studies, this result is at variance with earlier conventional wisdom. Some of the studies referred to above use different estimation methods. Furthermore, within a given

---

24 As Heekman (1987) points out, identification in this type of estimation is possible without excluding variables. In this particular case, the nonlinearity of the ordered probit will generate the identification of the parameters in the wage equation. However, the inclusion of variables in the schooling decision equation, which can be excluded from the wage equation, may aid in the identifying of the parameters.
method, they use different identification strategies. The similarity of the results is nevertheless striking.

The cross-study regularity of the IV and FE estimates as compared to the OLS baseline motivated Card (1994) to understand the findings. He takes as reference the optimal schooling model developed by Becker (1967) and derives explicit formulae for the conventional estimate of the return to schooling and for the alternative IV and FE estimators. He shows that OLS provides an estimate of the rate of return on average, while the IV provides an estimate of the rate of return for marginal individuals affected by a specific intervention. The IV estimate of the return to education from a particular intervention can then exceed the conventional OLS if the intervention affects a sub-population with a sufficiently high marginal rate of return to schooling (e.g. those with a high discount rate). With respect to the empirical regularity of the FE being higher than the OLS, the most likely explanation is that variation in the facility of access to funds or in tastes for schooling is reduced within families. If this is the case, schooling choices are more correlated with ability within families than across the overall population, thus leading to an upward bias in the within family fixed effects estimate relative to OLS.

4.3 Educational attainment and changes in compulsory schooling

This section will focus on legal changes in the compulsory level of schooling in Portugal. Completion of the first three years of education corresponded to the compulsory level from the 1920s until 1956, when it was established at four years, but for boys only (entailing education until the age of 12 or less if the person completed the level before turning 12 years old). Compulsory education increased to six years in 1964 for children beginning school that year (entailing education until the end of age 14 or less if the person completed the level prior to turning 15 years old). After the mid-eighties a new law which outlined the structure of the educational system increased the compulsory level of education to nine years. The data set

---

25 However, a recent paper by Harmon and Walker (1997) casts doubt upon this reason. They use a set of instruments appropriate for capturing different marginal groups. They find that the OLS is much lower than their alternative estimates. Furthermore, their estimates are stable with respect to the choice of instruments. In particular, instruments which might be expected to affect children likely to be constrained by the minimum school leaving age seem to imply the same estimated returns, as do instruments likely to be the most important in determining participation in higher levels of schooling.

26 Compulsory schooling of four years was introduced for girls in 1960. The 1956 law recognised exceptions with respect to enrolment in the fourth year of education due to distance from the nearest school or economic status of the family. For instance, those living more than 3 km from the nearest school could be exempt if there was no free transportation. The distance could be 4 Km if, among other things, there was a canteen in the school. Those from (rural) families who relied on the work of the child for the necessity of the family could also be exempt depending on the family’s economic situation, but the law prescribed that such exemptions should be limited. Meanwhile, several restrictions on employment were imposed. For instance, employers in the industry and services could not hire workers under 21 who did not have four years of completed education after the start of 1959. Furthermore, those with less than four years of formal education could not receive a driving licence, participate in official sports activities, or be hired as public servants.
used in this study does not capture the effects of this alteration. Because of this, we focus only on the changes that occurred in the late 1950s and early 1960s.

Table 4.1: Changes in the compulsory level of schooling (men)

<table>
<thead>
<tr>
<th>time period</th>
<th>compulsory level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920s - 1955</td>
<td>3 years</td>
</tr>
<tr>
<td>1956 - 1963</td>
<td>4 years</td>
</tr>
<tr>
<td>1964 - mid-1980s</td>
<td>6 years</td>
</tr>
<tr>
<td>after the mid-1980s</td>
<td>9 years</td>
</tr>
</tbody>
</table>

If we look ahead to Table 4.2, most who entered school during the 4-year compulsory period completed at least this level. Although the law may have influenced the enrolment rate, it has been argued that the 4-year extension most likely occurred when a vast majority of youngsters were already attending at least that level of school (see Grácio, 1986). The situation is clearly different with respect to the change to six years in 1964. A majority of those who entered the first year of education after 1963 at most have obtained four years of education. It is likely that some individuals reached the mandatory school-leaving age without completing the compulsory level of six years due to failing some years of education. Indeed, the failure rate was quite high in the first years of education. However, the proportions with fewer than six years of education are much higher for those who enrolled in school before 1964 (73.3% and 72.6% in the samples of 1986 and 1992, respectively).

Table 4.2: Education attainment by period of entering school (men)

<table>
<thead>
<tr>
<th>cohort</th>
<th>% below their compulsory level</th>
<th>average years of education</th>
</tr>
</thead>
<tbody>
<tr>
<td>before 1956</td>
<td>3.5</td>
<td>3.4</td>
</tr>
<tr>
<td>1956-1963</td>
<td>53.5</td>
<td>45.4</td>
</tr>
<tr>
<td>after 1964</td>
<td>53.5</td>
<td>45.4</td>
</tr>
</tbody>
</table>


Increases in the compulsory level coincide with increases of the average years of completed education. Obviously, such an aggregate analysis alone does not imply a cause-effect relation. Indeed, the literature documents an increasing time trend of the educational attainment in many countries largely because on average, younger cohorts attend school longer than the older ones.

---

27 One must be cautious here. The data correspond to two samples of employees in 1986 and 1992 and include no information on the individuals' educational career. In the following analysis, it is assumed that it was made by the individual without interruptions after age 6. However, it may be the case that some individuals who attended school during the 3-year compulsory period concluded the 4th-year later, from within the labour market, by participating in adult education programs.
(see Simon and Boggs, 1997). This issue will be addressed in section 6.6, where we control for the individual’s age in a regression analysis.

### 4.4 Estimation methods

Consider the following two-equation model describing wages and education attainment:

\[ \ln w_i = \beta' X_i + \alpha E_i + u_i \]  \hspace{1cm} (4.1a)

\[ E_i = \gamma' Z_i + v_i \]  \hspace{1cm} (4.1b)

where \( E_i \) denotes the years of education of individual \( i \), \( w_i \) is a measure of wages, \( X_i \) is a vector of control variables (e.g. age, region, etc.) and \( u_i \) and \( v_i \) represent a pair of residuals. \( Z_i \) is a vector of exogenous variables that influence individuals’ education decisions. It is well-known that unless \( u_i \) and \( v_i \) are uncorrelated, OLS gives a biased estimate for the rate of return \( \alpha \). As mentioned in section 4.2, there are several reasons to suppose that the pair of errors may indeed be correlated.

One way of dealing with the correlation between \( u_i \) and \( v_i \) is to use the IV method. It consists of estimating (4.1b) by OLS in the first stage and predicting \( E_i \). Equation (4.1a) is estimated in a second stage, where \( E_i \) is replaced by its predicted value. Variables in \( X_i \) and \( Z_i \) may be overlapping, but some variables in \( Z_i \) must be excluded from \( X_i \) for identification.

Education is supposed to be a continuous variable in (4.1b). However, this may be restrictive since what we observe in the data is an ordered set of different levels of schooling. This ordinal nature of education can be acknowledged by an ordered probit estimation method. Furthermore, the estimates from the wage equation can be corrected by using a two-step method. This approach was used by Harmon and Walker (1995) and corresponds to an extension of the two-step method developed by Heckman (1979). In this case the model for education and wages can be written as:

\[ \ln w_i = \beta' X_i + \alpha E_i^* + u_i \]  \hspace{1cm} (4.2a)

\[ E_i^* = \gamma' Z_i + v_i \]  \hspace{1cm} (4.2b)

\[ E_i = j \text{ if } \mu_{j-1} < E_i^* \leq \mu_j \]  \hspace{1cm} (4.2c)

where \( E_i^* \) is a latent variable assumed to depend linearly on the vector \( Z_i \). The model structures individuals’ education attainment (\( E_i \)) via the latent variable \( E_i^* \) and interval thresholds \( \mu \). These thresholds are such that \( \mu_i > \mu_m \) for \( k>m \). The lowest \( \mu \) equals \(-\infty \) and
the highest equals $+\infty$. The pair of error terms $u_i$ and $v_i$ is assumed to be $\text{BVN}(0, 0, \sigma_u^2, \sigma_v^2, \rho_{uv})$ distributed.

The two-step method involves evaluating the expected value of wages conditional on the attained level of education:

$$E(\ln w_i \mid E_i = j) = \beta' X_i + \alpha E_i + E(u_i \mid E_i = j)$$

$$= \beta' X_i + \alpha E_i + \rho_{uv} \sigma_v E\left(v_i \left| \frac{\mu_{j-1} - \gamma' Z_i}{\sigma_v} < \frac{u_i}{\sigma_v} < \frac{\mu_j - \gamma' Z_i}{\sigma_v}\right.\right)$$

The last expectation is the mean of a double truncated univariate normal. This expectation can be written as:

$$\lambda_i = \frac{\phi\left(\frac{\mu_{j-1} - \gamma' Z_i}{\sigma_v}\right) - \phi\left(\frac{\mu_j - \gamma' Z_i}{\sigma_v}\right)}{\Phi\left(\frac{\mu_j - \gamma' Z_i}{\sigma_v}\right) - \Phi\left(\frac{\mu_{j-1} - \gamma' Z_i}{\sigma_v}\right)}$$

(4.4)

where $\phi$ and $\Phi$ denote, respectively, the density function and the cumulative distribution function of the standard normal variate.\(^{28}\) Therefore, we can write

$$E(\ln w_i \mid E_i = j) = \beta' X_i + \alpha E_i + \theta \lambda_i,$$

(4.5)

where $\theta = \rho_{uv} \sigma_v$. If $\rho_{uv} \neq 0$ the estimation of equation (4.2a) by OLS will yield a biased estimate of $\alpha$. The selectivity-corrected equation consists of adding $\lambda_i$ to the list of covariates in equation (4.2a). We estimate by OLS the equation:

$$\ln w_i = \beta' X_i + \alpha E_i + \hat{\theta} \lambda_i + \varepsilon_i, \quad i = 1, \ldots, N$$

(4.6)

The variable $\lambda_i$ is constructed according to expression (4.4), where $\gamma$ and the $\mu$'s are replaced by their estimates. These estimates are obtained by an ordered probit model for schooling whose log-likelihood function is given by:

\[ \text{Log } L = \sum_{i=1}^{N} \sum_{j=0}^{6} D_{ij} \log \left\{ \frac{\Phi(H_{ij} - \gamma'Z_{ij}) - \Phi(H_{ij} - \gamma'Z_{ij})}{\sigma_v} \right\} \]

(4.7)

where \( D_{ij} = 1 \) if individual \( i \) has completed the level of education \( j \), 0 otherwise.

The full set of parameters in Log-\( L \) is not identified. Because of this, we use a common normalisation procedure that is to set \( \mu_0 = 0 \) and \( \sigma_v = 1 \). Since the error term in equation (4.6) is heteroscedastic, the covariance matrix will be computed by using the method of White (1980). This type of correction procedure is proposed in Lee (1982) and Davidson and MacKinnon (1993, p. 544).

4.5 Estimation results

This section includes estimates for the models outlined in section 4.4. The data were drawn from Quadros de Pessoal for 1986 and 1992, and are as described in the previous chapter. The nature of the data prevents any correction for potential biases arising from decisions to participate in the labour market. Because this problem might be more relevant for women, only male workers are considered. \(^{29}\) In order to create identifiers, external information on the number of years of compulsory education when the individual began schooling was added to the data set. These are defined as binary variables:

<table>
<thead>
<tr>
<th>variable</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956 - 1963</td>
<td>binary variable taking value 1 if the individual entered school after 1955</td>
</tr>
<tr>
<td></td>
<td>and before 1964, 0 otherwise.</td>
</tr>
<tr>
<td>after 1963</td>
<td>binary variable taking value 1 if the individual entered school after 1963,</td>
</tr>
<tr>
<td></td>
<td>0 otherwise.</td>
</tr>
</tbody>
</table>

The dependent variable in the wage equation is the log of hourly wages. For the sake of comparison, the specification is as close as possible to that of Harmon and Walker (1995). The list of regressors includes years of education, age, age squared, and four region dummies. With respect to the education equations, the list of covariates includes age, age squared, four region dummies, and the identifying variables. \(^{30}\) The variables associated with age may capture the time trend effect in education attainment. Estimated results are in Tables A-4.1 and A-4.2 in

\(^{29}\) Indeed, we applied the same technique to the sample of female workers but the results were quite abnormal. In most of the cases the coefficient of education appears negative, although not significant.

\(^{30}\) Since we do not have the individual's experience, age instead of the conventional measure of potential experience is used. Harmon and Walker (1995) did the same. The reason is that if education is endogenous, potential experience is also endogenous. Moreover, if education is measured with error, potential experience will be too.
the appendix. The following analysis begins with the IV estimates and ends with those obtained with the two-step method. OLS estimates are used as the baseline.

When applying the IV technique some crucial statistical tests are also reported in Tables A-4.1 and A-4.2 in the appendix. In a recent paper, Bound et al (1995) warns of problems with instrumental variables estimation when the correlation between instruments and the endogenous explanatory variable is weak. As they show, a weak correlation, even in apparently large samples, can result in a large bias in the IV estimates. They propose two tests to check the quality of the IV estimates. First, the F-statistic on excluded instruments in the reduced form schooling equation needs to indicate statistical significance. Second, the researcher must examine the adjusted partial $R^2$ obtained from regressing education against the potential instruments once the common exogenous variables have been partialled out. The validity of the instruments is also analysed through an overidentification test requiring that all instruments are orthogonal to the error term of the wage equation. The idea behind this test is that the excluded instruments should have no explanatory power in the wage equation once we have controlled for the other regressors. The statistic of this test is calculated as outlined in Davidson and MacKinnon (1993 p. 236) and has a chi-squared distribution with degrees of freedom equal to the number of overidentification restrictions. Finally, a Hausman t-test which allows one to check whether IV and OLS estimates are truly different is also reported.

Of particular interest in the education equations is the effect of changes in the compulsory schooling law. After controlling for the individual's age, all coefficients have the expected positive sign. Furthermore, the more recent the change for longer compulsory schooling, the larger the respective coefficient. Altogether this indicates that changes in compulsory schooling laws probably have lengthened the average schooling attainment.

With respect to the wage equation, the results obtained from the OLS estimation are broadly in line with results from OLS for other studies. Rates of return around 7.5-8.2% are in line with, for example, the studies referred to in Card (1994) and the studies of Harmon and Walker (1995) and Uusitalo (1996). While they may be marginally at the upper end of the scale, they are certainly not at variance with existing evidence in the area.

The IV estimates for the return to education have a substantially higher standard error than the OLS baseline. This type of result is also very common in other works (see e.g. Angrist and Krueger, 1991, 1992, Butcher and Case, 1993, Card, 1993). Large standard errors may result from a weak correlation between instruments and the variable of concern. To examine this correlation we rely on the tests suggested by Bound et al (1995). The F-statistic requirement on the identifying instruments in the first stage equation is clearly satisfied. With respect to the partial $R^2$, the values are somewhat lower than that of Harmon and Walker (1995), who found a value of .0046, but are still higher than those reported by Bound et al (1995) in a replication of the work of Angrist and Krueger (1991). In any case we must recognise that they are not high.

The IV point estimate is lower than the OLS counterparts, but the standard errors are very large. Indeed, a Hausman test does not reject the equality of OLS and IV coefficients at the
5% level in some cases. However, the results for 1992 are not very accurate: the hypothesis that the instruments are orthogonal to the error term in the wage equation is rejected in column (A) of Table A-4.2. Apart from this case, the results seem to indicate that IV estimates of the return to education are as high or lower than the OLS counterparts. This is clearly at odds with most empirical evidence reported in other studies for other countries.

Now consider the two-step method. We must notice that contrary to the IV estimator, identification can be achieved in the two-step method without exclusion of variables. Identification in the wage equation is generated through the nonlinearity of the ordered probit (see footnote 24). Because of this property, it may be a preferred method when the identifiers are weak. However, the outcome is mixed when compared with the OLS baseline. The coefficient of the selection term ($\hat{\alpha}$) shows statistical significance in 1986 and 1992. This indicates that OLS and two-step method estimates are different. However, the sign is negative in 1986 and positive in 1992. The results for 1986 suggest that OLS return to education exceeds the two-step method estimate. The selectivity-corrected estimates are much more homogeneous than those obtained by IV identifiers (this may suggest identification through nonlinearity and weakness of the identifiers based on changes in the compulsory schooling level). They are nearly 1.3 percentage points lower than the OLS value. This is not corroborated in the 1992 data, however. In this case, the estimates for the return to education are slightly higher (about 0.9 percentage points) than in the OLS.

But this method relies on quite strong assumptions. In particular, the assumption of joint normality of the errors may be strong. A semiparametric analysis, where the selection process can be nonparametrically modelled and the selectivity correction made with weaker assumptions, is an appealing way to proceed in the future.

Finally, it is worth noting that all methods point to an increase of the returns to education over time.

4.6 Towards an interpretation

OLS can differ from IV estimates because of endogeneity of education as a result of individuals' optimal choices, measurement error in this variable, and omitted variables. Since the effect of measurement error is to introduce a downward bias in OLS estimates, and in the previous section OLS exceeds or equals the IV estimates, this source of bias is ruled out. The other sources of bias may, however, be operative. In particular, if omitted ability has a positive effect on wages and correlates positively with education, OLS estimates may be biased upwards.
Card’s analysis (1994) provides another interesting line of thought. He argues that an IV procedure based on an intervention affecting only a sub-sample of individuals (the treatment group), with no effect on an otherwise identical sub-sample (the control group), estimates the return to education of the specific group affected by the (binary) instrument. On the other hand, OLS yields the average return for the entire sample. This argument reads as follows. Consider a pooled regression for those two groups of log-wage on schooling, using treatment group status as an instrument for education. The resulting IV estimate of the return to education has the following probability limit:

$$\rho_v = \frac{\ln w_t - \ln w_c}{S_t - S_c}$$

(4.8)

where the subscripts $t$ and $c$ denote the treatment and the control groups, respectively. If the rates of return vary across individuals, the instrumental variables estimate that results from the intervention equals the marginal rate of return to schooling of the sub-group affected.

In the current case, IV estimates are as high or lower than OLS estimates. It is likely that the IV estimate is to some extent capturing the return of a specific group, i.e. low achievers. Indeed, changes in the compulsory level of schooling were designed to constrain this education group. If, in this particular case, low investment in education is more related to low marginal benefits (i.e. low ability) rather than to high marginal costs, IV estimates may be reflecting (lower) returns to education among the less able. In fact, the (OLS) rate of return in Portugal increases as the level of education increases; it is lower within the less-educated groups (see the information on Table A-4.3). The inclusion of a squared term for education in the wage-equation yields a positive coefficient for this variable. Hence, the interpretation has some support.

### 4.7 Conclusions

In this chapter we have tried to elaborate on the current debate in empirical labour economics regarding the potential bias in OLS estimates of the returns to education. In particular, we have examined the Portuguese case. To do so, we compared IV and two-step method estimated returns to education with the OLS baseline.

Two key issues emerge from the analysis. Under a first heading, we must say that the results obtained from OLS estimation are broadly in line with results from OLS estimation for other studies in other countries. While they may be marginally at the upper end of the scale, they are certainly not at variance with existing evidence in the area. Under a second heading, most of the results except one seem to indicate that OLS return to schooling is as high or higher than the IV and the two-step method estimates. This result is at odds with recent findings for other countries. It is consistent with the omitted ability bias but not with measurement errors in the
education variable. This could suggest that OLS estimates provide a reasonable upper-bound for the return to education for men in Portugal. Given the particular nature of the instruments used, we must be cautious about generalising this result onto the entire population.

Finding identifying variables is a difficult task. In part because of this, the study may suffer a few drawbacks. First, we rely on changes in the compulsory schooling level since this was the *only* candidate available at the time. This identifying strategy is very close to that of Harmon and Walker (1995) but our instruments for Portugal may not be as good as those for the U.K. The first change to four years of education in 1956 may have caused a low exogenous variation in education. Most likely the extension occurred when a vast majority of youngsters were already attending four years of schooling. The extension to six years in 1964 also deserves comment. Harmon and Walker (1995) report that compulsory schooling laws were strictly adhered to in the U.K. This was not really the case in Portugal. Nevertheless, empirical analysis points to some influence of the legal changes on school attainment. Second, given the nature of the instruments, the IV estimate may reflect the rate of return of a specific sub-population. The interventions were aimed at influencing the level of education of low achievers, which apparently have a low rate of return in Portugal. The development of a set of instruments able to capture different marginal groups is a necessary goal for the near future.
## Appendix

### Table A-4.1: Estimations Results - 1986

#### 1. Wage equations

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>IV</th>
<th>Two-step method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)</td>
<td>(B)</td>
<td>(A')</td>
</tr>
<tr>
<td>education</td>
<td>.0748* (118)</td>
<td>.0303 (1.50)</td>
<td>.0608* (21.1)</td>
</tr>
<tr>
<td></td>
<td>.0303 (1.30)</td>
<td>.0511</td>
<td></td>
</tr>
<tr>
<td>(\hat{\lambda})</td>
<td>.0492* (5.08)</td>
<td>.0477* (4.90)</td>
<td></td>
</tr>
<tr>
<td>adj. R(^2)</td>
<td>.445</td>
<td>.376</td>
<td>.426</td>
</tr>
<tr>
<td>Hausman test (t-statistic)</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
</tr>
<tr>
<td>Partial R(^2) (excluded instruments)</td>
<td>.0013</td>
<td>.0012</td>
<td>.0013</td>
</tr>
<tr>
<td>F-test (excluded instruments)</td>
<td>21.61</td>
<td>10.56</td>
<td>21.61</td>
</tr>
<tr>
<td>Chi-square (overidentification)</td>
<td>1.17</td>
<td>1.17</td>
<td>1.17</td>
</tr>
</tbody>
</table>


#### 2. Education equations

<table>
<thead>
<tr>
<th></th>
<th>Reduced Form OLS</th>
<th>Ordered Probit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)</td>
<td>(A')</td>
</tr>
<tr>
<td>1956-1963</td>
<td>.3430* (5.72)</td>
<td>1956-1963</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>after 1963</td>
<td>.6781* (6.41)</td>
<td>after 1963</td>
</tr>
<tr>
<td></td>
<td>.2287* (3.23)</td>
<td></td>
</tr>
<tr>
<td>adj. R(^2)</td>
<td>.0717</td>
<td>Log-likelihood</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Log-lik. (slopes=0)</td>
</tr>
</tbody>
</table>

Absolute t-values in parentheses. All equations also include covariates for age, age squared and four region dummies. *significant at the 1% level. ** significant at the 5% level. N=39329 observations.
Table A-4.2: Estimations Results - 1992

1. Wage equations

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>IV</th>
<th>Two-step method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)</td>
<td>(B)</td>
<td>(A')</td>
</tr>
<tr>
<td>education</td>
<td>.0815*</td>
<td>.0348**</td>
<td>-.0286*</td>
</tr>
<tr>
<td></td>
<td>(108)</td>
<td>(2.33)</td>
<td>(2.48)</td>
</tr>
<tr>
<td></td>
<td>.0565*</td>
<td>(3.85)</td>
<td>-.0290*</td>
</tr>
<tr>
<td></td>
<td>(26.2)</td>
<td>(26.2)</td>
<td></td>
</tr>
<tr>
<td>λ</td>
<td>.0897*</td>
<td>.0898*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.48)</td>
<td>(2.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.0897*</td>
<td>.0898*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.48)</td>
<td>(2.53)</td>
<td></td>
</tr>
<tr>
<td>adj. R²</td>
<td>.401</td>
<td>.341</td>
<td>.407</td>
</tr>
<tr>
<td>Hausman test (t-statistic)</td>
<td>3.29</td>
<td>1.74</td>
<td></td>
</tr>
<tr>
<td>Partial R² (excluded instruments)</td>
<td>.0029</td>
<td>.0027</td>
<td></td>
</tr>
<tr>
<td>F-test (excluded instruments)</td>
<td>49.0</td>
<td>94.5</td>
<td></td>
</tr>
<tr>
<td>Chi-square (overidentification)</td>
<td>50.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


2. Education equations

<table>
<thead>
<tr>
<th></th>
<th>Reduced Form OLS</th>
<th>Ordered Probit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)</td>
<td>(A')</td>
</tr>
<tr>
<td></td>
<td>(B)</td>
<td>(B')</td>
</tr>
<tr>
<td>1956-1963</td>
<td>.1261***</td>
<td>.1894*</td>
</tr>
<tr>
<td></td>
<td>(1.90)</td>
<td>(6.01)</td>
</tr>
<tr>
<td>after 1963</td>
<td>.7638*</td>
<td>.1852*</td>
</tr>
<tr>
<td></td>
<td>(8.20)</td>
<td>(7.91)</td>
</tr>
<tr>
<td>adj. R²</td>
<td>.0872</td>
<td>.0871</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Log-likelihood</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-47505</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Log-lik.(slopes=0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-49438</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-49438</td>
</tr>
</tbody>
</table>

Absolute t-values in parentheses. All equations also include covariates for age, age squared and four region dummies. *significant at the 1% level  ** significant at the 5% level  ***significant at the 10% level. N=34870 observations.

Table A-4.3: Wage equations: OLS estimates for the entire sample and for selected sub-samples

<table>
<thead>
<tr>
<th></th>
<th>year: 1986</th>
<th></th>
<th>year: 1992</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>by groups of schooling</td>
<td>whole sample</td>
<td>by groups of schooling</td>
<td>whole sample</td>
</tr>
<tr>
<td></td>
<td>≤4 years</td>
<td>≤6 years</td>
<td></td>
<td>≤4 years</td>
</tr>
<tr>
<td>education</td>
<td>.0451*</td>
<td>.0664*</td>
<td>.0748*</td>
<td>.0347*</td>
</tr>
<tr>
<td></td>
<td>(25.9)</td>
<td>(49.5)</td>
<td>(118)</td>
<td>(13.5)</td>
</tr>
<tr>
<td>adjusted R²</td>
<td>.2728</td>
<td>.3032</td>
<td>.4458</td>
<td>.2162</td>
</tr>
<tr>
<td></td>
<td>.3531</td>
<td>.3661</td>
<td>.3871</td>
<td>.4013</td>
</tr>
<tr>
<td>F-value</td>
<td>1408</td>
<td>1958</td>
<td>4520</td>
<td>777</td>
</tr>
<tr>
<td></td>
<td>26274</td>
<td>31492</td>
<td>39329</td>
<td>19711</td>
</tr>
</tbody>
</table>

Absolute t-values in parentheses. All equations also include covariates for age, age squared and four region dummies. *significant at the 1% level.