Instrument Choice and the Returns to Education: New Evidence from Vietnam*

Jean-Louis ARCAND† Béatrice D'HOMBRES††

Paul GYSELINCK††

†CERDI-CNRS, Université d'Auvergne & European Development Network

††CERDI-CNRS, Université d'Auvergne

July 2004

Abstract

This paper focuses on instrument choice while consistently estimating the returns to education in Vietnam. Using data culled from the 2 rounds of the Vietnam Living Standards Survey (VLSS), we explore different sets of exogenous instruments that rely on demand and supply side sources of variation in schooling as well as the matrix of instruments proposed by Hausman and Taylor (1981). Instrument validity tests suggest that many variables do not satisfy the necessary conditions allowing them to be used as instruments. As in several studies, we find that IV estimates of the returns to education are substantially higher than the corresponding OLS estimate. We show how the Hausman-Taylor matrix of instruments, when combined with other instruments, may be a useful way of consistently estimating an average return to education rather than a local average treatment effect (Imbens and Angrist, 1994).

Keywords: rate of return, instrumental variables procedures, instrument choice, Hausman-Taylor estimator, Hahn-Hausman test, Vietnam

JEL classification: J31, I21, C30

*We thank participants at the IZA European Summer School in Labor Economics, the 53^{th} Annual Meeting of the French Economic Association, the 14^{th} Meeting in Mathematics and Statistics Applied to Economic Modeling and the development seminar at the University of Auvergne. We are also grateful to Christopher Grigoriou for useful comments and suggestions.

Corresponding author: B.D'Hombres, CERDI, 65 bld Mitterrand, 63000 Clermont-Ferrand, France.
E-mail: B.D_Hombres@u-clermont1.fr
Tel + 33 4 73 17 74 18/Fax + 33 4 73 17 74 28

1 Introduction

Measuring the returns to schooling has been the object of considerable interest in the empirical labour economics literature.[1] In recent years, several new methods have emerged to deal with different biases associated with estimates of the causal impact of education on earnings. This paper contributes to the debate by (i) using IV estimators in order to obtain consistent estimates of the returns to education in Vietnam and (ii) exploring the validity of different sets of instruments using various criteria recently proposed in the econometric literature.

Finding instruments that are orthogonal to the disturbance term in Mincerian wage equations has been the topic of a great deal of debate: it is widely recognized that it is difficult to identify demand-side variation in schooling uncorrelated with individual earnings. The focus has therefore shifted to supply-side sources of variation in schooling (such as changes in the minimum school-leaving age, schooling reform or the geographic proximity of schools) which should allow one to identify exogenous variation in schooling decisions.[2] The problem is that the condition that the instruments be strongly correlated with the endogenous variables, a condition that is necessary if one is to avoid the "weak instruments problem" (leading to finite sample bias), as emphasized by Staiger and Stock (1997), is often not satisfied in studies that rely on supply side variation. When one has panel data, the generalized instrumental variables procedure proposed by Hausman and Taylor (1981, henceforth HT) constitutes another tool that should allow one to consistently identify the effect of time-invariant variables correlated with unobserved individual effects in the absence of external instruments.

In this paper, we estimate the returns to education in Vietnam using these different sets of instruments, and a battery of tests is performed in order to test their validity. To the best of our knowledge, this is the first study that takes unobservable heterogeneity into account in an earnings equation estimated on Vietnamese data. Surprisingly, the panel nature of the VLSS as well as the wealth of the data available have never been used in an effort to obtain consistent estimates of the returns to schooling.[3]

[1] See Card (2001) for a review of the literature on this subject.

[2] Another solution to the endogeneity problem, in the absence of panel data, is to possess a proxy for ability (as in Griliches (1977) or Boissiere *et al* (1985)) or twins data (as in Ashenfelter and Zimmerman (1997), Behrman *et al* (1994) or Bound and Solon (1999)). The use of twins data causes some problems in the presence of measurement error on schooling variables and imposes strong assumptions concerning the measure of ability.

[3] Note that this paper is the first study that we are aware of that simulatenously uses these three types of instruments. Some studies have used two of the sets of instruments we propose: one set based on demand-side variation and another based on supply-side variation. See for example Callan and Harmon (1999), Dearden (1999) and Ichino and Winter-Ebmer (1999). See Guillotin and Sevestre (1994) for an application of the HT instrument set.

A priori, the transition from a planned to a market economy should lead to an increase in the returns to schooling insofar as it is expected that wages will be more closely connected to productivity.[4] A number of studies (Moock, Patrinos and Venkataraman, 2003, Nguyen, 2002) have shown, as in the majority of planned economies, that the returns to education in Vietnam are still low despite the movement towards greater liberalization. One of the goals of this paper is to ascertain whether this is still the case once one controls for unobservable individual heterogeneity.[5]

The rest of the article is organized as follows. Section 2 discusses different matrices of instruments with a particular focus on that proposed by HT. Section 3 describes the data used in the paper and the estimation results. We compare generalized IV results (GIV) with OLS results and those traditionally obtained on Vietnamese data. Section 4 concludes.

2 "External" versus "internal" instruments

Consider the following earnings equation:[6]

$$\begin{cases} y_{it} = \beta_0 + X_{it}\beta_1 + Z_{1i}\gamma_1 + Z_{2i}\gamma_2 + \eta_{it}, .. \\ \eta_{it} = f_i + \varepsilon_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T. \end{cases}, \tag{1}$$

where X_{it} is a $[NT \times K]$ vector of time-varying explanatory variables and $Z_i = [Z_{1i}; Z_{2i}]$ is a $[NT \times (G_1 + G_2)]$ vector of time-invariant explanatory variables. N is the number of individuals and T is the number of periods over which they are observed. We assume that X_{it} and Z_{1i} are "doubly exogenous" in that they are uncorrelated with the disturbance term ε_{it} and the unobserved individual effects f_i. On the other hand, Z_{2i} is "singly exogenous" in that it is assumed to be correlated only with the individual effect. Typically, Z_{2i} will correspond to years of schooling. In the present context, we ignore all correlation between the explanatory variables and the error term ε_{it} which could stem, for example, from measurement error in years of completed schooling.[7] The preceding assumptions may then be expressed in formal terms as follows:

[4] See, among others, Halpern and Krosi (1998), Orazem and Vodopivec (1997) or Rutkowski (1997) for evidence on the returns to schooling in transition economies.

[5] However we underline that this article focuses more on the methodology used than on the specifics of developing countries.

[6] For the sake of simplicity, we adopt HT's notation throughout the remainder of this paper.

[7] See Ashenfelter and Krueger (1994), Card (2001) or Ashenfelter *et al* (1999) for summaries of the sources and consequences of measurement errors in the schooling variable.

$$\begin{cases} E[\mathbf{X}'_{it} f_i] = E[Z'_{1i} f_i] = 0, E[Z'_{2i} f_i] \neq 0, \\ E[\mathbf{X}'_{it} \varepsilon_{it}] = E[Z'_i \varepsilon_{it}] = 0, E[f'_i \varepsilon_{it}] = 0. \end{cases} \tag{2}$$

Because individual effects (attributed to innate ability and motivation in our basic Mincerian wage equation) are unobservable, estimating equation (1) using the pooling estimator will yield biased estimates of the coefficients. One therefore has to estimate (1) via IV methods. Using traditional fixed effect methods (such as the within or first difference transformations) is not a viable solution in that, while they allow one to control for f_i, they also sweep out all time-invariant variables, thereby rendering it impossible to identify γ_1 and γ_2. In order to implement an IV procedure, we have to rely on (i) the availability of external instruments and/or (ii) the matrix of instruments proposed by HT.

2.1 "Traditional" instruments

The usual instrumental variables procedure relies on a matrix of external instruments W_{it} of dimension $[NT \times W]$ with $W \geqslant G_2$ which satisfies the following conditions:

$$E[\mathbf{W'}_{it} \eta_{it}] = 0, \; E[\mathbf{W'}_{it} \mathbf{Z}_{2i}] \neq 0. \tag{3}$$

We possess two broad sets of excluded instruments used in the literature.

First, we employ instruments that rely on DEMAND-SIDE VARIATION IN SCHOOLING such as parental education and smoking habits. Formally speaking, the first IV estimator (IV_1) will be based on two variables corresponding to the number of years of education of individual i's mother and father. More educated parents can assist their children in reaching higher levels of education. Because of higher wealth, they may also be faced with reduced liquidity constraints that may otherwise limit their children's educationnal attainment. Glewwe and Patrinos (1999) highlighted the role played by parental education in terms of its impact on educational attainment, while Glewwe and Jacoby (2004) underscored that human capital investment in Vietnam is constrained by household resources. Even if family characteristics are often considered to be potentially correlated with earnings (thus failing to satisfy one of the necessary conditions for instrument validity), they are widely used. In the case of Vietnam, we believe that these instruments should meet the orthogonality conditions because, under communism, the intergenerational transmission of wealth and social background only obtained through educational attainment.[8]

Our second IV (IV_2) estimator, also based on demand-side variation in schooling, stems from heterogeneity in individual discount rates. More specifically, we use information

[8] Note also that such instruments are not immune from measurement error.

concerning smoking habits in the past (a dummy variable indicating whether an individual ever smoked for at least six months) as a predictor of educational attainment.[9] As stated by some authors such as Fersterer and Winter-Ebmer (2003), there are many reasons that could render this instrument invalid. Note for example that if smoking is a normal good, its consumption will increase with earnings. Smoking may also be negatively correlated with education because of heightened awareness concerning the risks involved. This instrument will therefore be carefully tested so as to check its validity.

The second set of instruments considered relies on SUPPLY-SIDE SCHOOLING VARIATION. As in Card (1995) or Malluci (1998) we use the proximity of primary schools and colleges as instruments, denoted by IV_3.[10] While these instruments should induce exogenous variations in education, we remain suspicious concerning their relevance.[11] As noted by Bound *et al* (1995), even if the orthogonality condition is satisfied, weak correlation between the endogenous variable and the set of instruments W_{it} leads to finite sample bias in the same direction as the OLS estimate, the magnitude of which depends upon the correlation between the endogenous variable and the excluded instruments.[12]

2.2 Hausman Taylor instruments

An alternative to excluded instrumental variables is provided by the HT estimator (rarely employed in the literature perhaps because of the paucity of panel data) which provides consistent and efficient estimates of the coefficients γ_2 associated with "singly exogenous" time-invariant variables Z_{2i} despite the absence of external instruments.[13] Their approach involves using individual-specific means, as well as deviations with respect to the individual-specific means as instrumental variables. More precisely, the set of instruments proposed by Hausman-Taylor (1981) is:

[9] See Fersterer and Winter-Ebmer (2003), Chevalier and Walker (1999) or Evans and Montgomery (1994)).

[10] We use a dummy variable indicating if there was a primary school in the village when the concerned individuals were of school age. Similarly, we use a dummy variable indicating whether there was a secondary school.

[11] Other exogenous sources of variation in education have been considered in the literature. The best-known example is constituted by Angrist and Krueger (1991) in which an individual's quarter of birth (and interactions with state of birth) are used as instruments. Harmon and Walker (1995, 1999), Ichino and Winter-Ebmer (1998), Meghir and Palme (1999) and Ferterer and Winter-Ebmer (2003) also use exogenous sources of variation in schooling outcomes.

[12] For a discussion concerning the properties of IV estimators and finite sample bias, see for example Nelson and Startz (1990a,b), Buse (1992) or Staiger and Stock (1997). The bias in finite samples is due to the fact that the coefficients from the reduced form equation are estimated. For an excellent survey concerning the weak instruments problem, its consequences and potential solutions, see Stock *et al* (2002) or Hahn and Hausman (2003).

[13] See Hausman and Taylor (1981) for an application of this estimator to the returns to schooling.

$$A_{HT} = [Q_v X_{it}; P_v X_{it}; Z_{1i}], \tag{4}$$

where P_v and Q_v are the idempotent matrices that perform the between and within transformations respectively. The basic intuition behind the HT estimator is that only the f_i component of the error term is correlated with Z_{2i}, which allows one to decompose X_{it} and use $Q_v X_{it}$ and $P_v X_{it}$ as instruments since $E[(Q_v X_{it})' \varepsilon_{it}] = E[(P_v X_{it})' \varepsilon_{it}] = 0$, under the assumptions given in (2). The HT estimator therefore allows one to control for unobservable correlated individual effects, while identifying the parameter of interest (γ_2) in our Mincerian equation. A necessary condition for identification is that the number of elements of X_{it} be greater than the number of elements of Z_{2i}.[14] This IV approach is an ingenious manner of artificially multiplying the number of available instruments and thereby side-stepping the identification issue. This matrix of instruments will be denoted by IV_4.

To take into account the composite structure of our stochastic error term, since $\Omega = \sigma_f^2 I_{T_N} + \sigma_\varepsilon^2 J_{T_N}$ (where $I_{T_N} = I_N \otimes I_T$ and $J_{T_N} = I_N \otimes J_T$ with J_T a $T \times T$ matrix of ones), and to obtain a more efficient estimator, a generalized IV estimator (GIV) is implemented.[15]

[14] Note that we could decompose X_{it} into $X_{it} = [X_{1it}; X_{2it}]$ where X_{it} is a $[NT \times (K_1 + K_2)]$ vector of time-varying explanatory variables and X_{1it} is assumed to be "doubly exogenous" while X_{2it} is "singly exogenous". In this case, the matrix of instruments proposed by HT is $A_{HT} = [Q_v X_{it}; P_v X_{1it}; Z_{1i}]$. A necessary condition for identification is then that $K_1 > G_2$. These results have been extended by Amemiya and McCurdy (1986) and Breusch, Mizon and Schmidt (1989) who suggest the wider set of instruments given respectively by $A_{AM} = [Q_v X_{it}; X_{1it}^*; Z_{1i}]$ and $A_{BMS} = [Q_v X_{it}; (Q_v X_{it})^*, P_v X_{1it}; Z_{1i}]$ where X_{1it}^* is a $(1 \times TK_1)$ vector where X_{1it} stands for X_1 in each period. The Amemiya-McCurdy instrument set assumes that the doubly exogenous variables are uncorrelated with the individual effects in each period. The Breusch-Mizon-Schmidt instrument set assumes that these correlations are the same in each period. Their approach requires one to have a panel data set where $T > 2$. If one only has two periods, as we do, their matrix of instruments is the same as that proposed by Hausman and Taylor.

[15] Equation (1) is premultiplied by $\Omega^{-1/2}$ which, as indicated by Hausman (1978), is equal to :

$$\Omega^{-1/2} = I_N - (1-\theta) P_v \text{ with } \theta = \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_f^2} \right)^{\frac{1}{2}}.$$

Formally speaking, we use a two-step procedure so as to obtain consistent estimates of θ and then of Ω. The matrix of instruments is applied to equation (1) yielding a consistent estimate of the vector of coefficients. The residuals are then used in order to carry out the transformations allowing one to obtain consistent estimates of σ_ε^2 and σ_f^2. Finally equation (1) is transformed by $\Omega^{-1/2}$and one applies 2SLS on the transformed equation using the matrix of instruments under consideration. As noted in Wooldridge (2002), if Ω is not random, other variance-covariance matrices should be considered since they would potentially produce more efficient estimates. Our program is available upon request.

3 Empirical Results

3.1 Data

The data is derived from two waves of the Vietnam Living Standards Surveys (VLSS) household survey collected by the General Statistical Office of Vietnam with the assistance of the World Bank. The first survey was undertaken in $1992-93$ and the second in $1997-98$. The panel structure of the VLSS allows us to track 324 wage-earning males over the period $1992-1998$. Only males were retained in order to avoid issues of labor market participation. Table 1 reports the sample statistics associated with the relevant variables.

Table 1: Summary statistics

	$1993-1997$
Number of observations	648
Age: mean	30.22 (11.15)
Hourly earnings (in thousand Dong)	2.51 (2.64)
Years of completed schooling	8.43 (4.12)
Region of residence: North(%)	40.12
Urban Area (%)	39.35
Ethnic group (%): Kinh + Chinese	93.51
Public sector (%)	40.74
Instruments	
Mother's Education	2.97 (3.14)
Father's Education	5.29 (4.08)
Smoking Habits (% having ever smoked)	67.59
Proximity of a primary school	0.663
Proximity of a secondary school	0.519

Standard errors in parentheses.

Our dependent variable is the logarithm of the hourly wage of an individual's main job in the week prior to the survey.[16] It includes all reported wages, bonuses, work subsidies, and income in kind. It is equal on average to 1.46 and 3.56 thousand Dongs (VSD) in

[16] We are conscious of the fact that most Vietnamese work in agriculture or are self-employed. This explains why a sizable portion of the work force is dropped from our sample. However, the transition from a planned to a market economy involves an expansion of wage employment and it is thus of great interest to identify those forces driving wage growth.

1992 – 93 and 1997 – 98, respectively.[17] Accounting for an inflation rate of 8% per year between the two surveys, this represents an increase in real terms of 10.7% per year. This performance of the wage sector can be compared to an average growth rate of real GDP per capita of 6.5% during the same period. These numbers illustrate the success of the liberalization movement (Doi Moi) implemented in the late 1980s.

Our measure of education is the highest year of education completed by the individual.[18] Note that this measure of education is not the actual number of years spent at school allowing us to account for grade repetition. The average level is 8.43 years over our sample. This is very high for a low income country such as Vietnam. Moreover, only 5% of our sample has no education while 60% and 30% finished lower and upper secondary school, respectively. These figures highlight the importance ascribed to in education in Vietnam, as compared with other countries having a similar GDP per capita (411$ in 2001). Roughly 40% of our sample lives in northern Vietnam. The relative importance of the public sector remained stable between the two surveys, representing about 40% of wage employment.

All estimates of the returns to schooling presented below are taken from a basic Mincerian wage equation that links the log of wages to years of schooling. Control variables include age, age squared, a regional dummy (north versus south), an urban area dummy, an occupational dummy (public versus private sector), an ethnic dummy and a time dummy. The ethnic minority dummy has been introduced in order to control for inequality between the two main economic groups (Kinh and Chinese) and other minority groups (see Van de Walle and Gunewardena 2001 and Baulch and Minot 2002).[19] We have also included a public sector dummy so as to control for the remaining effect of the old public sector wage structure.

Table 2 displays OLS and IV estimates of the returns to schooling.[20] The OLS results give a rate of return to education of 2.6%. In comparison, using the 1992 – 93 VLSS, Moock *et al* (2003) found, for male workers, an average return of 3.4%. Similarly, Nguyen (2002) obtained returns to education of 3.1% and 3.9%, using respectively the 1992 – 93 VLSS and the 1997 – 98 VLSS data. Our lower OLS estimate can be explained either by the use of a more complete specification or by the exclusion of women from our sample in order to avoid sample selection issues. In all cases, these numbers are *extremely* low compared to those typically obtained in developing countries. It has been shown that returns to

[17] 1 USD was exchanged for about 10 000 VSD in 1993 and 13 000 VSD in 1997.

[18] The Vietnamese educational system is composed of general, vocational and higher education (universities and colleges). General education includes primary school (5 years), lower secondary school (4 years), and upper secondary school (3 years). There are vocational schools after each of these levels.

[19] However, as stated by these authors, the main source of ethnic inequality stems more from an unequal distribution of endowments - particularly in terms of geographical concentration in poorer areas - than from a pure discrimination effect.

[20] The complete results are presented in Table $A.4$ of the Appendix.

Table 2: Return to education

	OLS	**IV$_1$**	**IV$_4$**	**IV$_1$ + IV$_4$**
Returns to Education	0.026 (0.004)	0.070 (0.001)	0.051 (0.169)	0.065 (0.001)
Statistical tests :				
Hansen test of overid. restrict χ^2	$n.a.$	0.001 (0.99)	0.99 (0.91)	1.19 (0.98)
Partial R^2	$n.a.$	0.13	0.058	0.176
Number of excluded instruments	$n.a.$	2	5	7
Hahn-Hausman-m_1	$n.a.$	-0.027	-0.072	-0.082
Hahn-Hausman-m_2	$n.a.$	-0.000	0.678	1.085
Donald-Newey $MSE\ test$	$n.a.$	2.354	16.458	1.382
σ	0.64	0.65	0.61	0.62
$\overline{R}^2$	0.43	0.41	0.44	0.43

Sets of instruments used : IV_1 : Parent's education, IV_2 : Smoking habits, IV_3 : School proximity, IV_4 : HT set of instruments.

P-values in parentheses.

schooling are usually low in transition economies but tend to increase as economic reforms deepen. For example, Maurer-Fazio (2002) found an average return of 2.9% and 3.7% in China for male workers in 1989 and 1992.[21] This phenomenon can be explained by the persistence of the egalitarian wage structure of the pre-reform period. More specifically, in the case of Vietnam, restrictions on worker mobility allow enterprises to underpay their workers compared to their marginal productivity.[22] Thus, if the gap between the wage and marginal productivity is larger for educated workers, returns to education will be low by construction[23] Note that we also tested a specification in which education was interacted with time (column 2, Table A.2). No evidence emerged indicating the the returns to education have significantly increased over time. This may be because wages increased uniformly across the schooling distribution. In the remainder of the paper, we therefore focus on the more restrictive specification in which the returns to education are constant over time.

3.2 IV estimates

[21] See also Orazem and Vodopivec (1995) and Varga (1995) for studies on the returns to education in China, Slovenia and Hungary.

[22] See Guest (1998).

[23] Using data on rural enterprises, Fleisher and Wang (2004) provided evidence that this gap was relatively larger for skilled workers than for production workers in China.

In our attempt to correct for the potential endogeneity of schooling, we implemented the IV estimators proposed above. However, as a preliminary step, we carried out three sets of tests concerning the validity and relevance of our proposed instruments.

3.2.1 Instrument orthogonality

We first rely on the traditional HANSEN TEST OF OVERIDENTIFYING RESTRICTIONS (which tests the orthogonality of the instruments) despite the fact that this test is now well-known to be potentially inconclusive. This is because the Hansen test is based on the hypothesis that at least one instrument in each instrument set is exogenous (Wooldridge 2002). Moreover, its power is particularly low in the presence of weak instruments (Baum, Schaffer, and Stillman 2003). Adding weak instruments may lead one not to reject the null hypothesis of orthogonality just by increasing degrees of freedom (Sevestre 2002).

In order to take this weakness of the HANSEN TEST into account, our strategy consisted in (i) repeating the test using various combinations of instruments and (ii) implementing the "DIFFERENCE-HANSEN TEST" so as to check the validity of subsets of instruments (see Hayashi 2000). This last statistic is simply the difference between two Hansen statistics, the first statistic being that computed from the restricted model which uses only the "non-suspect" instruments while the second is that associated with the unrestricted specification which includes the instruments "under suspicion".

The first four rows of table $A.1$ in the Appendix present results of overidentification tests on the four instrument sets proposed above. In *all* cases, the overidentifying restrictions are not rejected. However, these results may be misleading as our various sets of instruments are by definition homogeneous. Thus, for example, if the father's education were to fail the Hansen test, one should expect the same to obtain for the mother's education. Moreover, because it is just identified, we cannot test the orthogonality of our second set of instruments (smoking habits), as noted in row 2 of table $A.1$. Rows 5 to 14 present Hansen tests on a large number of possible combinations of instruments. From row 5 to row 8, we use different combinations of IV_1, IV_2 and IV_3, while rows 9 to 14 systematically include the matrix of instruments proposed by HT (IV_4). As with the Hansen tests on each set of instruments individually, these results would lead one not to reject. It is interesting to note, however, that while the p-values associated with the various Hansen tests never fall below 20%, *inclusion of* $IV2$ *and* $IV3$ *always reduces the p-value associated with the test.*

We next implemented the Difference-Hansen test. Our suspicions concerning IV_2 and IV_3 were then confirmed. As shown in table $A.1$, when the subsets of instruments being tested are those related to school proximity and individual discount rates (IV_2 and IV_3), the null hypothesis that these instruments are valid is marginally rejected (see rows $5, 6, 7$ and 10 to 14). This leads us to be cautious concerning the validity of these instrument sets, and to lean towards dropping them in our attempt to consistently estimate the returns to schooling in Vietnam. By contrast, as shown in rows 8 and 9, parental education appears once again to be strongly exogenous. This confirms our expectations concerning returns to education and the intergenerational transmission of human capital in Vietnam. Private returns to schooling under communism are usually held to have been non-existent. Note that $IV4$ also passes the Hansen test.

3.2.2 Instruments relevance

In order to address concerns about the weakness of our instruments, we then carried out the partial F-test of the joint significance of the instruments and calculated the partial R^2, for the first stage regressions.[24] Results are reported in the last two columns of Table $A.1$ and suggest that IV_2 (smoking habits), IV_3 (school proximity) and IV_4 (HT) are not sufficiently correlated with schooling (rows 2, 3 and 4). As shown by Staiger and Stock (1997), and even in the presence of a very large data set, an F-statistic below 10 when there is a single endogenous regressor means that one is potentially facing a weak instruments problem.[25] While we did not reject their orthogonality on the basis of the Hansen or Diff-Hansen tests, the F-test size leads one to be very cautious concerning IV_4. On the other hand, there is a strong correlation between schooling and parental education, as illustrated by a partial R^2 of 13% and an F-statistic of 47.9 (row 1). When we combine IV_1 and IV_4, the partial R^2 rises to 17% (row 9). This shows that adding the HT matrix of instruments can potentially improve the efficiency of more traditional IV methods based on excluded demand side instruments. Note that when we combine IV_1 with IV_2 or IV_3 the hypothesis of weak instruments is also rejected. However, as shown earlier, IV_2 and IV_3 may not satisfy the orthogonality condition, leading us to consider their combination with IV_1 as inadmissible. As a consequence of these findings, we are led to focus our attention, in what follows, on IV_1, IV_4 and IV_{14}.

Another way of evaluating the relevance of our instruments involved implementing the Hahn-Hausman specification test (2002).[26] In contrast to the partial R^2 and F-statistics which test the null hypothesis of "weak" instruments, the Hahn Hausman test (henceforth HH) is based upon the null of "strong" instruments.[27] This test is constructed

[24] This means that we compute the F-test and the R^2 of the reduced form once the other covariates have been partialled out.

[25] More precisely, from table 1 presented in Stock and Yogo (2002, p. 522), given the partial F-test size and the number of excluded instruments, we can infer that the 2SLS bias will probably exceed 10% with IV_2, IV_3 and IV_4.

[26] To the best of our knowledge, this is one of the first empirical uses of this test. Our program is available upon request.

[27] To be precise, the Hahn Hausman test is a joint test of the orthogonality and relevance of the instruments. However, as in the empirical example in Hahn and Hausman (2002), we apply this test to instruments which pass a first screening in terms of the test of the overidentifying restrictions.

by running the 2SLS regression in its usual "forward" form, and comparing the result to that obtained by running the "reverse" regression, in which the jointly endogenous right-hand-side (RHS) variable is moved to the LHS, and the dependent variable is entered on the RHS.[28] The basis for their first test (m_1) is that, if the specification is correct and the instruments are "strong", *standard first-order asymptotics imply that there will be very little difference between the results one obtains using the forward or reverse regressions.*[29] Things are complicated somewhat because of the need to adjust for second-order bias in the estimators, and the test is standardized by using a second-order expression for the variance of the difference between the forward and reverse estimators.

The test, which can be read as simple t-statistic, is presented in Table $A.2$. The null hypothesis of strong instruments is clearly not rejected for IV_1, IV_4 and the combination of both (IV_{14}). We also implemented another version of the test which involves use of the forward and reverse bias-adjusted 2SLS (Nagar) estimator proposed by Donald and Newey (2001). As stated by Hahn and Hausman, this test is somewhat simpler to implement because of the absence of the (second-order) bias term. Here, the m_2 statistic based on the Nagar estimator does not reject IV_1, IV_4.or IV_{14}. Note also that the point estimate of the returns to education is remarkably stable whatever estimator is used (forward 2SLS, reverse 2SLS, forward bias-adjusted 2SLS, reverse bias-adjusted 2SLS).

3.2.3 Instrument choice

As a final set of diagnostics, and in order to single out one instrument set as our preferred choice (among IV_1, IV_4 or IV_{14}), we carried out the DONALD AND NEWEY (2001) "CHOICE OF INSTRUMENTS" TEST. The Donald and Newey test is based upon choosing, from within *a number of valid instrument sets*, the one which *minimizes the mean-squared error (MSE)* of each estimator.[30] As can be seen in the first column of the table $A.3$, IV_{14} minimizes the MSE of all three estimators. The conclusion, on the basis of the Donald and Newey test, is that IV_{14} is the preferred instrument set, whether one uses 2SLS, Nagar or LIML.

[28] For example, in the wage regressions we consider here, the reverse regression involves putting educational attainment on the left and log wage on the right, with the position of the included predetermined variables remaining unchanged.

[29] See Hahn and Hausman (2002).

[30] In the first stage of the test procedure, one identifies the instrument set that minimizes the Mallows or Cross-Validation reduced form goodness of fit criterion. This instrument set is then used to compute initial estimates of the variance of the reduced form and structural equation residuals, as well as their covariance, which enter into the expressions for the Mallows (and Cross-Validation criterion) criterion and the MSEs. In the second stage of the procedure, one recomputes the Mallows criterion (or the Cross-Validation), for each instrument set. This is then plugged into the appropriate expression for the MSE of each estimator (see Donald and Newey, 2001, pp. 1164-5), which itself depends upon the choice of instrument set. The instrument set which minimizes the MSE is then the one that should be used with the corresponding estimator.

The upshot of these procedures is that, despite having various matrices of instruments at our disposal, few are able to satisfy the two conditions that are necessary for them to be admissible. A sequence of test procedures led us to settle on IV_1, IV_4 and IV_{14} as being the best potential candidates, with a marginal preference, based on the Donald and Newey choice of instruments test, for IV_{14}.

3.2.4 IV estimates

As in most studies of the returns to education, we obtain IV estimates that are substantially higher than the corresponding OLS estimate. Using IV_1, the estimated return to education increases to 7% (from 2.6% using OLS), while using the matrix of instruments proposed by HT (column 2) yields an estimated return of 5.1%, though it is estimated less precisely. The specification using a combination of IV_1 and IV_4 (reported in column 4) gives a point estimate of 0.065 which lies between the figures estimated with IV_1 and IV_4 alone.

Given that the correlation between unobservable heterogeneity (such as innate ability) and educational attainment is likely to be positive, the OLS estimate should be biased upwards. Card (1999), Bound and Jaeger (1996) and Ichino and Winter-Ebmer (1999) have proposed an explanation for this phenomenon that is based on the hypothesis that the returns to schooling are heterogeneously distributed across the population.[31] This heterogeneity may appear at two levels.

First, it is likely that the marginal returns to education are decreasing in the level of schooling. Thus, if parental education mainly influences the educational choice of individuals at the lower end of the distribution of schooling, IV procedures will yield a return that is higher than it should be in the population as a whole.[32] Second, for a given level of education, parental education will potentially affect educational choice (i) for more able individuals (educated parents will more easily spot abler children and encourage them to pursue their studies) and/or (ii) individuals with high discount rates (due for example to low taste for education).[33]

In short, heterogeneity in the returns to education implies that IV estimates will produce different results that are functions of the set of instruments used, the difference

[31] Note that several other explanations have been suggested. In particular, Griliches (1977) and Angrist and Krueger (1991) pointed out the potential downward bias caused by measurement error in OLS. However, it is now well-known that the magnitude of the error required to explain the observed differences between OLS and IV estimates is much larger than what had previously been established in the litterature. Ashenfelter *et al* (1999) also furnished an explanation in terms of publication bias.

[32] See Card (2001).

[33] In a constrained situation, parents will choose to finance more skilled children.

stemming from which subpopulation is most affected by the instruments in question. The weighted marginal return we estimate is akin to what Imbens and Angrist (1994) called a *local average treatment effect* (LATE).[34] We believe that this phenomenon explains why our IV estimates exceed the OLS estimate.

Combining IV_1 with IV_4 (HT) should produce an estimated return closer to the true average return insofar as the instruments used in the HT procedure should influence educational decisions more uniformly through the distribution of schooling. This would appear to be confirmed by the results reported in column 4 of Table 2 where the point estimate is slightly lower than what is obtained using IV_1 alone. While the HT procedure may not allow one to precisely identify the average return to education, we believe that it does offset the effect of instruments that are only correlated with part of the distribution of schooling. Combining demand side (IV_1) and HT (IV_4) instruments may therefore constitute a good compromise solution, especially here, where the instruments in question are not rejected by a succession of tests designed to assess their exogeneity, relevance, and performance in terms of the mean-squared error or information criteria of the resulting estimators.[35]

4 Conclusion

In this paper, we have studied the economic returns to education in Vietnam using a panel data set. We used the instrumental variables procedure proposed by Hausman and Taylor (1981), as well as various matrices of instruments commonly used in the literature. A series of tests led us to conclude that few instrumental variables met the two conditions necessary for them to be admissible. In the case of Vietnam, only parental education and the matrix of instruments proposed by HT satisfy the two usual requirements. When the endogeneity of schooling is taken into account, the return to an additional year of schooling increases substantially: in line with international evidence, we find that OLS under-estimates the returns to schooling. As has been suggested by a number of authors, we believe that we are potentially facing a local average treatment effect (LATE) problem, which the HT matrix of instruments can partly help one to solve.

References

AMEMYA, T., AND T. MACURDY (1986): "Instrumental-Variable Estimation of an Error-Components Model," *Econometrica*, 54(4), 869–881.

[34] On this topic, see Heckman and Vytlacil (1998) and Imbens and Angrist (1994).

[35] However, note that the IV estimates of the return to education is higher than the corresponding OLS estimate. Consequently the solution is only partial.

Angrist, J., and A. Krueger (1991): "Does Compulsory Schooling Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics*, 106(4), 979–1014.

Ashenfelter, O., C. Harmon, and H. Oosterbeek (1999): "A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias," *Labour Economics*, 6(4), 453–470.

Ashenfelter, O., and A. B. Krueger (1994): "Estimates of the Economic Returns to Schooling from a New Sample of Twins," *American Economic Review*, 84(5), 1157–1173.

Ashenfelter, O., and C. Rouse (1998): "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins," *Quarterly Journal of Economics*, 113(1), 253–284.

Ashenfelter, O., and D. Zimmerman (1997): "Estimates of the Return to Schooling from Sibling Data: Fathers, Sons and Brothers," *Review of Economics and Statistics*, 79(1), 1–9.

Baulch, B., and N. Minot (2002): "The Spatial Distribution of Poverty in Vietnam and the Potential for Targeting," World Bank Policy Research Working Paper 2829.

Baum, C. F., M. E. Schaffer, and S. Stillman (2003): "Instrumental Variables and GMM: Estimation and Testing," *Stata Journal*, 3(1), 1–31.

Behrman, J. R., M. Rosenzweig, and P. Taubman (1994): "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment," *Journal of Political Economy*, 102(6), 1131–1174.

Boissiere, M., B. J. Knight, and R. H. Sabot (1985): "Earnings, Schooling, Ability, and Cognitive Skills," *American Economic Review*, 75(5), 1016–1030.

Bound, J., D. Jaeger, and R. 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(430), 443–50.

Bound, J., and G. Solon (1999): "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling," *Economics of Education Review*, 18(2), 169–182.

Breusch, T., G. Mizon, and P. Schmidt (1989): "Efficient Estimation Using Panel Data," *Econometrica*, 57(3), 695–700.

Buse, A. (1992): "The Bias of Instrumental Variables Estimators," *Econometrica*, 60(1), 173–180.

Callan, T., and C. Harmon (1999): "The Economic Return to Schooling in Ireland," *Labour Economics*, 6(4), 543–550.

Card, D. (1995): "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," pp. 201–222, Aspects of labour market behaviour: Essays in honour of John Vanderkamp.

——— (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69(5), 1127–1160.

Chevalier, A., and I. Walker (1999): "Further Results on the Returns to Education in the UK," mimeo, University of Warwick.

Dearden, L. (1999): "The Effects of Families and Ability on Men's Education and Earnings in Britain," *Labour Economics*, 6(4), 551–567.

Donald, S., and W. Newey (2001): "Choosing the Number of Instruments," *Econometrica*, 69(5), 1161–1191.

Evans, W., and E. Montgomery (1994): "The Effects of Cigarette Smoking on Wages," *Industrial and Labor Relations Review*, 50(3), 493–509.

Falaris, E. M. (2003): "The Effect of Survey Attrition in Longitudinal Surveys: Evidence from Peru, Côte D'Ivoire and Vietnam," *Journal of Development Economics*, 70(1), 133–157.

Fersterer, J., and R. Winter-Ebmer (2003): "Smoking, Discount Rates, and Returns to Education," *Economics of Education Review*, 22(6), 561–566.

Fleisher, B., and W. X. (2004): "Skill Differentials, Return to Schooling, and Market Segmentation in a Transition Economy: The Case of Mainland China," *Journal of development economics*, 73(1), 315–328.

Glewwe, P., M. Gragnolati, and H. Zaman (2002): "Who Gained from Vietnam's Boom in the 1990s?," *Economic Development and Cultural Change*, 50(4), 773–792.

Glewwe, P., and H. Jacoby (1998): "School Enrollment and Completion in Vietnam: An Investigation of Recent Trends," in *Household Welfare and Vietnam's Transition*, ed. by G. P. Dollar, D., and J. Litvack, Washington, D.C. World Bank Regional and Sectoral Studies.

Glewwe, P., and H. Jacoby (2004): "Economic Growth and the Demand for Education: Is There a Wealth Effect?," *Journal of Development Economics*, 74(1), 33–51.

Glewwe, P., and H. A. Patrinos (1999): "The Role of the Private Sector in Education in Vietnam: Evidence From the Vietnam Living Standards Survey," *World Development*, 27(5), 887–902.

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

Guest, P. (1998): "The Dynamics of Internal Migration in Vietnam," *United Nations Development Program*, Hanoi, United Nations Development Program, Discussion Paper 1.

Guillotin, Y., and P. Sevestre (1994): "Estimations de Fonctions de Gains sur Données de Panel : Endogénéité Du Capital Humain et Effets de la Sélection," *Économie et Prévision*, 116(4), 119–135.

Hahn, J., and J. Hausman (2002): "A New Specification Test for the Validity of Instrumental Variables," *Econometrica*, 70, 163–189.

——— (2003): "Weak Instruments: Diagnosis and Cures in Empirical Econometrics," *American Economic Review*, 93(2), 118–125.

Halpern, L., and G. Korosi (1998): "Labour Market Characteristics and Profitability: An Econometric Analysis of Hungarian Exporting Firms, 1986-95," *Economics of Transition*, 6(1), 145–162.

Harmon, C., and I. Walker (1995): "Estimates of the Economic Return to Schooling for the United Kingdom," *American Economic Review*, 85(5), 1278–1286.

——— (1999): "The Marginal and Average Returns to Schooling in the UK," *European Economic Review*, 43(4), 879–887.

Hausman, J., and W. Taylor (1981): "Panel Data and Unobservable Individual Effects," *Econometrica*, 49(6), 1377–1398.

Hayashi, F. (2000): *Econometrics*. Princeton University Press, Princeton, NJ, 1st edn.

Heckman, J., and E. Vytlacil (1998): "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return Is Correlated with Schooling," *Journal of Human Resources*, 33(4), 974–987.

Ichino, A., and R. Winter-Ebmer (1999): "Lower and Upper Bounds of Returns to Schooling: An Exercice in IV Estimation with Different Instruments," *European Economic Review*, 43(4-6), 889–902.

Imbens, G., and J. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), 467–476.

Malluci, J. (1998): "Endogeneity of Schooling in the Wage Function: Evidence from the Rural Philippines," FCNCD discussion paper 54, International Food Policy Research Institute, Washington, D.C.

Maurer-Fazio, M., and N. Dinh (2002): "Differential Rewards to, and Contributions of, Education in Urban China's Segmented Labor Markets," WDI Working Paper 508.

Meghir, C., and M. Palme (1999): "Estimating the Effect of Schooling on Earnings Using a Social Experiment," *IFS Working Paper*, 99/12, London, UK.

Moock, P., H. Patrinos, and M. Venkataraman (2003): "Education and Earnings in a Transition Economy: The Case of Vietnam," *Economics of Education Review*, 22(5), 503–510.

Nelson, C. R., and R. Startz (1990a): "Some Further Results on the Exact Small Properties of the Instrumental Variable Estimator," *Econometrica*, 58(4), 967–976.

——— (1990b): "The Distribution of the Instrumental Variables Estimator and Its T-Ratio When the Instrument is a Poor One," *Journal of Business*, 63(2), S125–S140.

Nguyen, N. (2002): "Trends in the Education Sector from 1993-8," World Bank Policy Research Working Paper 2891.

Orazem, P., and M. Vodopivec (1997): "Unemployment in Eastern Europe, Value of Human Capital in Transition to Market: Evidence from Slovenia," *European Economic Review*, 41(3-5), 893–903.

Rutkowski, J. (1997): "Low Wage Employment in Transitional Economic of Central and Eastern Europe," *MOST, Economic Policy and Transitional Economics*, 7(1), 105–130.

Sevestre, P. (2002): *Econométrie Des Données de Panel*. Dunod, Paris.

Staiger, D., and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65(3), 557–586.

Stock, J.H., W. J., and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in GMM," *Journal of Business and Economic Statistics*, 20(4), 518–529.

Van de Walle, D. (2003): "Are Returns to Investment Lower for the Poor? Human and Physical Capital Interactions in Rural Vietnam," *Review of Development Economics*, 7(4), 636–653.

Van de Walle, D., and D. Gunewardena (2001): "Sources of Ethnic Inequality in Vietnam," *Journal of Development Economics*, 65(1), 177–207.

Varga, J. (1995): "Returns to Education in Hungary," *Review Acta Oeconomica*, 47(1-2), 203–215.

Wooldridge, J. (2002): *Econometric Analysis of Cross-Section and Panel Data*. MIT Press, Cambridge, MA.

A Appendix

A.1 Validity and relevance of our instruments

	Excluded instruments	Hansen test	Diff-Hansen test	Subset of instruments tested	Partial R^2	F-test
1	IV_1	0.000 (0.99)	*n.a.*	*n.a.*	0.131	47.9 (0.000)
2	IV_2	*n.a.*	*n.a.*	*n.a.*	0.005	3.41 (0.07)
3	IV_3	0.92 (0.34)	*n.a.*	*n.a.*	0.001	0.15 (0.86)
4	IV_4	0.99 (0.91)	*n.a.*	*n.a.*	0.058	7.30 (0.000)
5	IV_1+IV_2	1.31 (0.52)	1.31 (0.25)	IV_2	0.136	33.41 (0.000)
6	IV_1+IV_3	4.14 (0.25)	4.14 (0.13)	IV_3	0.131	24.1 (0.000)
7	IV_1+IV_2+IV_3	4.96 (0.29)	4.96 (0.17)	IV_2+IV_3	0.136	20.02 (0.000)
8	IV_1+IV_2+IV_3	4.96 (0.29)	0.46 (0.8)	IV_1	0.136	20.02 (0.000)
9	IV_4+IV_1	1.19 (0.98)	0.22 (0.9)	IV_1	0.176	19.34 (0.000)
10	IV_4+IV_2	2.69 (0.75)	1.71 (0.19)	IV_2	0.062	7.00 (0.000)
11	IV_4+IV_3	5.42 (0.49)	4.43 (0.11)	IV_3	0.058	5.60 (0.000)
12	IV_4+IV_1+IV_3	5.4 (0.71)	4.21 (0.12)	IV_3	0.177	15.00 (0.000)
13	IV_4+IV_1+IV_2	2.73 (0.91)	1.55 (0.21)	IV_2	0.180	17.31 (0.000)
14	*All*	6.42 (0.41)	5.23 (0.11)	IV_2+IV_3	0.180	13.79 (0.000)

Sets of instruments used : IV_1, Parent's education, IV_2 : Smoking habits,
IV_3 : School proximity, IV_4 : HT set of instruments.
P-values in parentheses.

A.2 The Hahn Hausman (2002) test

Instrument set	IV_1	IV_4	$IV_1 + IV_4$
2SLS	0.069	0.051	0.065
	(0.023)	(0.036)	(0.020)
2SLS reverse	0.069	0.076	0.072
	(0.023)	(0.046)	(0.021)
Bias-Adjusted 2SLS	0.069	0.054	0.067
	(0.023)	(0.037)	(0.020)
Reverse Bias-Adjusted 2SLS	0.069	−0.002	0.042
	(0.023)	(0.001)	(0.012)
Hahn-Hausman m_1 test statistic	−0.027	−0.072	−0.082
Hahn-Hausman m_2 test statistic	−0.000	0.678	1.085

Standard errors in parentheses. When the number of instruments is equal to 2, the B2SLS estimator (reverse B2SLS) boils down to the 2SLS estimator (2SLS reverse).

A.3 The Donald and Newey (2001) instrument selection criteria

Instrument set	Mallows criterion on reduced form $\widehat{R}^m$	MSE of estimator based on $\widehat{R}^m$		
		$\widehat{S}_{2SLS}$	$\widehat{S}_{B2SLS}$	$\widehat{S}_{LIML}$
IV_1	5.732	2.354	2.415	2.416
IV_4	39.338	16.458	16.595	16.562
IV_{14}	3.288	1.271	1.382	1.389
	Cross-validation criterion on reduced form $\widehat{R}^{cv}$	MSE of estimator based on $\widehat{R}^{cv}$		
		$\widehat{S}_{2SLS}$	$\widehat{S}_{B2SLS}$	$\widehat{S}_{LIML}$
IV_1	5.800	2.383	2.444	2.445
IV_4	40.788	17.067	17.207	17.172
IV_{14}	3.284	1.270	1.380	1.388

A.4 Returns to education in Vietnam

	OLS	**OLS**	**IV$_1$**	**IV$_4$**	**IV$_1$+IV$_4$**
Coefficient					
Years of education	0.026 (0.004)	0.015 (1.49)	0.070 (0.001)	0.051 (0.169)	0.064 (0.001)
Years of education interacted with time		0.020 (1.61)			
Age	0.040 (0.006)	0.043 (3.30)	0.025 (0.137)	0.031 (0.111)	0.026 (0.109)
Age, squared	−0.001 (0.008)	−0.000 (−3.32)	−0.000 (0.099)	−0.000 (0.077)	−0.000 (0.080)
Urban dummy	0.302 (0.000)	0.300 (5.47)	0.235 (0.000)	0.264 (0.000)	0.243 (0.000)
North dummy	−0.420 (0.013)	−0.420 (−7.47)	−0.501 (0.000)	−0.467 (0.000)	−0.492 (0.000)
Public sector dummy	−0.013 (0.028)	−0.017 (−0.27)	−0.145 (0.125)	−0.090 (0.484)	−0.130 (0.147)
Ethnic dummy	−0.028 (0.779)	−0.025 (−0.22)	−0.106 (0.333)	−0.072 (0.540)	−0.096 (0.373)
Time dummy	0.951 (0.000)	0.780 (6.62)	0.973 (0.000)	0.964 (0.000)	0.970 (0.000)
Constant	−0.688 (0.008)	−0.658 (−2.86)	−0.589 (0.028)	−0.584 (0.022)	−0.556 (0.027)
Statistical tests :					
Hansen test of overid. restrict χ^2 $(p - value)$	*n.a.*	*n.a.*	0.0001 (0.99)	0.99 (0.91)	1.19 (0.98)
Partial R^2 (excluded instruments)	*n.a.*	*n.a.*	0.131	0.058	0.176
F on excluded instruments	*n.a.*	*n.a.*	47.90	7.30	17.60
Number of excluded instruments	*n.a.*	*n.a.*	2	5	7
Hahn-Hausman-m_1	*n.a.*	*n.a.*	−0.027	−0.072	−0.082
Hahn-Hausman-m_2	*n.a.*	*n.a.*	−0.000	0.678	1.085
σ	0.64	0.64	0.65	0.61	0.62
$\overline{R}^2$	0.43	0.42	0.41	0.44	0.43
Number of observations	648	648	648	648	648

P-values are in parentheses.