

Economic Returns to Education in France: OLS and Instrumental Variable Estimations

Sajjad Haider Bhatti*, Jean Bourdon**, and Muhammad Aslam***

Abstract

This article estimates the economic returns to schooling as well as analyzing other explanatory factors for the French labor market. It addresses the issue of endogeneity bias and proposes two new instruments for use in the instrumental variable two-stage least squares technique. Our results show that the proposed instruments are relevant and adequate, based on evidence from the available literature. After using the proposed instruments, we find that the OLS coefficients for schooling are biased downwards. Finally, we choose between the two proposed instruments.

Keywords: Endogeneity bias, instrumental variable, Mincerian model, two-stage least squares, wage regression, France.

JEL classification: C1, I2, J3, P5.

1. Introduction

In the labor market, individuals are rewarded depending on their skills, competencies, or knowledge. These skills and competencies are more typically referred to as human capital. Human capital theory (Mincer, 1958, 1974; Becker, 1964) states that education and training increase the productivity of individuals by augmenting their skills and knowledge. Education and training are, therefore, key factors in determining the economic performance of an individual. The rates of return to different human capital factors can help individuals make decisions regarding their investment in education, based on their possible future earnings. It is also important for policymakers to allocate resources such that they discourage or eliminate discrimination in economic rewards on the basis of geography, ethnicity, sex, or age.

* Institut de Recherche sur l'Éducation: Sociologie et Économie de l'Éducation (IREDU), Université de Bourgogne, Dijon, France.

** Directeur de recherche émérite au CNRS, Institut de Recherche sur l'Éducation: Sociologie et Économie de l'Éducation (IREDU), Université de Bourgogne, Dijon, France.

*** Assistant Professor, Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan.

Mincer's (1974) econometric wage regression model has been commonly used to estimate the returns to schooling. Some notable early work on this model includes studies by Chiswick (1983), Kuch and Haessel (1979), and Tomes (1983). Despite its wide-ranging applications, the model's simple estimation can be biased due to econometric problems of endogeneity, measurement error, and sample selection bias (Ashenfelter & Krueger, 1994; Card, 1999; Griliches, 1977, 1979).

In this article, we focus on the bias arising from the endogeneity of schooling, which can restrict the estimation of the true causal effect of schooling on wages (Griliches, 1977). The endogeneity problem arises due to the violation of an assumption under the Mincerian model that treats all individuals as identical (with respect to ability, opportunities, and environment) other than differences in education and training. This violation is inevitable given that different people cannot be identical with respect to certain unobservable characteristics such as social environment, location, and family background. For example, ability can be seen as a determinant of wages in the labor market on the one hand, and may also be correlated with schooling on the other, i.e., more able people tend to acquire more schooling, will be more productive at work, and hence paid better.

If unobserved ability affects both schooling and wages, then an ordinary least squares (OLS) estimation will yield biased results (Griliches, 1977; Card, 2001). In such a situation, the schooling variable will be correlated with the error term in the wage equation and, as a result, the coefficient associated with schooling will be biased. This type of bias is known as an endogeneity bias and can be tackled by using the instrumental variable two-stage least squares (IV2SLS) estimation approach. For this, we need instrumental variables that affect schooling but are otherwise uncorrelated with wages.

Different types of instruments have been used to counter the endogeneity bias, including parental or spouse education (see Blackburn & Neumark, 1993; Ozdural, 1993; Trostel, Walker, & Woolley, 2002; Zhang, 2011, among others) and the availability of (or distance to) a nearby college or school (see Card, 1993; Maluccio, 1998, among others). Many other instruments can also be found in the literature (see Angrist & Krueger, 1991; Brunello & Miniaci, 1999; Harmon & Walker, 1995; Ismail, 2007). The consensus in the literature is that the returns to schooling obtained from IV2SLS estimation are typically higher than those from OLS estimation.

Inspired by numerous studies on the relationship between earnings/wages and their determinants in different countries, a number of studies have been carried out using French data to investigate different aspects of the French labor market (see, for example, Abdelkarim & Skalli, 2005; Boumahdi & Plassard, 1992; Fougère, Goux, & Maurin, 2001; Selz & Thélot, 2004). However, few studies address the issue of the endogeneity of schooling in this context (e.g., Boumahdi and Plassard, 1992; Steunou, 2003; Viger, 2007).

We aim to eliminate the possible endogeneity bias in returns to schooling when applied to French data with the help of certain new instruments. Generally, researchers working with French data have tackled the issue of endogeneity bias by using parental education and family background variables as instruments. However, such instruments have been justifiably criticized due to their direct influence on wage determination in the labor market (Flabbi, 1999). This study is different from previous studies in that it tries to address the endogeneity problem using new instrumental variables that are expected not to have a direct impact on wages in the work market.

2. New Instrumental Variables and French Data

As with the correlation between an individual's schooling and his/her parents' education, it is reasonable to suppose that the former could also depend on the general trend in and motivation for schooling among that individual's family members. Given the likely similarities in environment (social and otherwise) and era, one's education may have a higher correlation with that of one's siblings relative to one's parents; this qualifies the variable as a more reasonable instrument than parents' education. The natural question to ask is which sibling's education should be used as an instrument, which leads us to use average household schooling as an instrument for endogenous schooling.

Based on this notion, we introduce the first instrument (Z_1) as "the household's average number of years of education." This variable reflects the fact that different persons in a family may have different educational inclinations and levels of ability. Thus, averaging the schooling of all family members eliminates the bias of ability. Since it is a data-generated instrument, its key benefit is that it can be used in any study that provides some information on household education. The instrument combines the effects of parents' and siblings' education on the schooling of an individual.

The educational decisions of an individual may also be affected by different factors at the time that he/she decides whether to work or to continue schooling. Thus, our second proposed instrument (Z2) is “the average years of schooling for a specific gender of a specific age group in the particular year that they join the labor force.” In this case, the ability bias is eliminated as different people come from different areas and different families (some with low ability and some with high ability). Any effect of school proximity is also balanced out by using this overall average. Technically, Z2 combines several instruments: ability, school proximity, family background, school quality, etc. The average schooling for the period 1950–2010¹ has been calculated by Barro and Lee (2010).

To estimate the Mincerian wage regression for the French labor market, we use data from the Labor Force Survey (Enquête d’Emploi Continu) carried out by the Institut National de la Statistique et des Etudes Économiques in 2007. Although the Mincerian model’s initial formulation was used to regress only the log of wages on a linear term for schooling and linear and quadratic terms for experience (or potential experience), most studies now use an extended version of the model, which contains other explanatory variables in addition to schooling and experience.

Our set of explanatory variables therefore includes different variables that can also affect wage determination, especially in the context of the French labor market. Table 1 defines all the variables used in our estimation. However, to compute the second instrument (Z2), Barro and Lee’s (2010) measures interact with the gender, year, and age group variables relevant to an individual joining the labor force.

¹ These measures are available at <http://www.barrolee.com/> for many countries of the world, including France.

Table 1: Variables in estimation of Mincerian model for French data

Variable	Description
Response variable	
LNWAGE	Natural logarithm of the monthly wage of an individual from main job
Explanatory variables	
SCH	Variable for education, measured in number of years of completed schooling
BEFEX	Experience (in labor market before current job), measured in number of years
BEFEX2	Experience squared
EXP	Current job seniority (labor market experience within present job), measured in number of years
EXP2	Current job seniority squared
HOURS	Number of hours associated with monthly salary (i.e., hours worked per month)
DGENDER	Gender of individual (male = 1, female = 0)
RDRURAL0	Dummy variable indicating whether individual is resident of rural area or not (reference category)
DNPARI	Dummy variable indicating whether individual is resident of urban area other than Paris region or not (non-Paris urban = 1, otherwise = 0)
DPARI	Dummy variable indicating whether individual is resident of Paris region or not (Paris region = 1, otherwise = 0)
DTYPDIP	Dummy variable indicating whether individual has degree/diploma in professional education or general education (professional = 1, general = 0)
DPUBLIC	Dummy variable indicating whether individual works in public sector or private sector (public = 1, private = 0)
RDTMPCT	Dummy variable indicating that individual is engaged in temporary work (reference category)
DFIXCT	Dummy variable indicating that individual works under fixed-term contract (fixed-term contract = 1, otherwise = 0)
DPERCT	Dummy variable indicating that individual works under permanent contract (permanent contract = 1, otherwise = 0)

3. Results and Discussion

Table 2 presents the results of human capital wage regressions based on the French data, using OLS and IV2SLS. In the first-stage regressions, the lower p-values associated with the instruments suggest the relevance of both instruments. The results show that the endogeneity problem causes a downward bias in returns to schooling for the OLS estimates. The schooling coefficient in the IV2SLS estimation (with Z1) is about 14 percent higher than that in the OLS estimation. This finding is in accordance with the literature, where endogeneity-corrected (IV2SLS) estimates are 10–30 percent (or approximately two percentage points) higher than OLS estimates (see Card, 1994; Ashenfelter, Harmon, & Oosterbeek, 1999). The schooling coefficient in the IV2SLS estimation (with Z2) is nearly double the corresponding OLS coefficient. The higher returns to schooling yielded by the IV2SLS approach are confirmed by several other researchers using French data (see Boumahdi & Plassard, 1992; Viger, 2007).

Before interpreting the coefficients in relation to the other explanatory variables, we need to choose between the two instruments Z1 and Z2. From the IV2SLS estimates based on Z1 and Z2, it is clear that the OLS estimates are downward-biased. The magnitude of difference between the OLS and IV2SLS estimates for the effect of education depends on the instrument used for schooling in the first-stage schooling equation. The explanatory power of Z1 is greater than that of Z2 in terms of the first-stage R^2 (see Table 2).

Table 2: OLS and IV2SLS estimates

Variable	OLS	IV2SLS-Z1		IV2SLS-Z2	
		First-stage	Second-stage	First-stage	Second-stage
INTERCEPT	4.7342 (<0.0001)	0.9943 (<0.0001)	4.6386 (<0.0001)	6.9604 (<0.0001)	4.0385 (<0.0001)
SCH	0.0670 (<0.0001)	–	0.0766 (<0.0001)	–	0.1371 (<0.0001)
BEFEX	0.0133 (<0.0001)	-0.0494 (<0.0001)	0.0145 (<0.0001)	-0.1482 (<0.0001)	0.0227 (<0.0001)
BEFEX2	-0.0002 (<0.0001)	0.0005 (<0.0001)	-0.0002 (<0.0001)	0.0014 (<0.0001)	-0.0003 (<0.0001)
EXP	0.0184 (<0.0001)	-0.0121 (0.0001)	0.0189 (<0.0001)	-0.0089 (0.1277)	0.0221 (<0.0001)
EXP2	-0.0001 (<0.0001)	-0.0003 (0.0003)	-0.0001 (<0.0001)	-0.0008 (<0.0001)	-0.0001 (0.0086)
HOURS	0.0064 (<0.0001)	0.0012 (<0.0001)	0.0063 (<0.0001)	0.0047 (<0.0001)	0.0060 (<0.0001)
DGENDER	0.1687 (<0.0001)	-0.2163 (<0.0001)	0.1734 (<0.0001)	-0.4695 (<0.0001)	0.2027 (<0.0001)
DNPARI	0.0117 (0.0159)	0.0621 (0.0045)	0.0082 (0.0915)	0.3552 (<0.0001)	-0.0137 (0.0312)
DPARI	0.2212 (<0.0001)	0.2631 (<0.0001)	0.2056 (<0.0001)	1.5910 (<0.0001)	0.1077 (<0.0001)
DTYPDIP	-0.0498 (<0.0001)	0.6557 (<0.0001)	-0.0619 (<0.0001)	1.2536 (<0.0001)	-0.1378 (<0.0001)
DPUBLIC	0.0049 (0.4736)	0.0594 (0.0555)	0.0014 (0.8423)	0.3518 (<0.0001)	-0.0210 (0.0129)
DFIXCT	0.3415 (<0.0001)	0.6011 (<0.0001)	0.3231 (<0.0001)	1.4442 (<0.0001)	0.2079 (<0.0001)
DPERCT	0.4963 (<0.0001)	0.4792 (<0.0001)	0.4760 (<0.0001)	1.5706 (<0.0001)	0.3487 (<0.0001)
Instrument	–	0.8916 (<0.0001)	–	0.2922 (<0.0001)	–
Hausman exogeneity test		262.58 (<0.0001)		64.66 (<0.0001)	
R ²	0.5528	0.7459	0.5435	0.2856	0.4355

Note: p-values given in parentheses.

Source: Authors' calculations.

In order to choose the more appropriate instrument, we look at the nature and behavior of both Z1 and Z2. Since Z1 represents the average number of years of education for the household, there may be some issues

when using this instrument. For example, an individual's ability may be related to his/her family situation or to a genetically determined inclination to pursue education. Furthermore, some household members might face similar schooling costs or opportunities (in terms of distance or the availability of educational institutions, etc.). Common demographics and geography can also affect schooling attainment because members of the same household are more likely to share a social and community environment. On the other hand, the impact of similar demographics, geography, schooling costs, opportunities, and genetic ability when using Z1 is eliminated in the case of the second instrument (Z2) because it is an overall average (as defined in Section 2).

A further comparison of Z1 and Z2 can be made with the help of a correlation matrix, given in Table 3. The correlation between Z1 and SCH is higher than that between Z2 and SCH. However, the correlation between the response variable (LNWAGE) and Z1 is also greater than that between LNWAGE and Z2. Although Z2 has a relatively low correlation with schooling, it also has a very low correlation with the response variable. Finally, keeping in view the correlation matrix and the more compact definition of Z2, we find this instrument more suitable for IV2SLS estimation. We refer to IV2SLS-Z1 and IV2SLS-Z2 in the IV2SLS estimations using the instruments Z1 and Z2, respectively.

Table 3: Correlation matrix for response, endogenous variable, and instruments

Variable	LNWAGE	SCH	Z1	Z2
LNWAGE	1.0000	0.3399	0.3360*	-0.1360
SCH		1.0000	0.8481*	0.2728*
Z1			1.0000	0.2633
Z2				1.0000

Note: * = significant at 5 percent level of significance.

Source: Authors' calculations.

Having selected the second instrument, Z2, as the more appropriate of the two, we now discuss the results of IV2SLS-Z2. The returns to schooling are about 13.7 percent for each additional year of schooling, which is twice that obtained from the OLS estimation. The returns to labor market experience are also higher than the corresponding OLS estimate. Each additional year of experience yields a 2.27 percent increment. Similarly, each extra year of current job seniority increases wages by approximately 2.21 percent.

The impact of work duration is almost identical in all specifications. The results reveal that women earn 20 percent less than their male counterparts and this gap is three percentage points higher in the IV2SLS-Z2 estimation. The wage differential due to regional differences changes considerably in the IV2SLS-Z2 estimation when compared to the OLS results. The differential between non-Paris urban and rural workers favors rural workers by a magnitude of 1.4 percent; in the OLS estimation, however, this differential favors non-Paris urban workers. The wage gain for workers in the Paris region over rural workers is 10.77 percent, which is 11–12 percentage points lower than the OLS estimate.

The coefficient of the professional diploma dummy variable indicates that professional diploma holders earn 13.8 percent less than those with a general education degree. This finding contradicts the previous evidence from other studies that have used French data (see Tansel, 1994; Simonnet, 1996). Individuals working in the private sector earn roughly 2 percent more than those working in the public sector; this difference is significant only in the IV2SLS-Z2 specification. We also find that individuals on fixed-term and permanent contracts earn 21 percent and 35 percent more, respectively, than their counterparts who work as temporary workers. These wage premiums increase by about 15 percent when endogeneity is taken into account. The direction of these findings is consistent with Arai, Ballot, and Skalli (1996) but the magnitude of the penalty for temporary workers is much larger in our investigation.

4. Conclusion

When estimating the economic returns to education using the Mincerian wage regression model, the problem of endogeneity bias is liable to arise. The 2SLS method with some potential instruments is commonly used in this situation. For the French data, we have used two new instruments: (i) the household's average number of years of education and (ii) the average schooling for a specific gender and specific age group in the particular year that an individual joins the labor force. The second instrument is found to be more appropriate.

Using this instrument, we find that the returns to schooling are about 13.7 percent for each additional year of schooling. Each year of experience yields an increase of about 2.27 percent in wages. French men are likely to earn 20 percent more than women. Workers in the public sector earn more than in the private sector, while temporary workers earn considerably less.

References

- Abdelkarim, O. B., & Skalli, A. (2005). Wage inequality in France (Working Paper ERMES). Paris, France: Université Panthéon-Assas Paris II.
- Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, *106*, 979–1014.
- Araï, M., Ballot, G., & Skalli, A. (1996). Différentiels intersectoriels de salaire et caractéristiques des employeurs en France. *Économie et Statistique*, *299*, 37–58.
- Ashenfelter, O., & Krueger, A. (1994). Estimates of the economic return to schooling from a new sample of twins. *American Economic Review*, *84*, 1157–1173.
- Ashenfelter, O., Harmon, C., & Oosterbeek, H. (1999). A review of estimates of the schooling/earnings relationship, with tests for publication bias. *Labor Economics*, *6*, 453–470.
- Barro, R. J., & Lee, J.-W. (2010). A new dataset of educational attainment in the world, 1950–2010 (Working Paper No. 15902). Cambridge, MA: National Bureau of Economic Research.
- Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. Chicago, IL: University of Chicago Press.
- Blackburn, M. L., & Neumark, D. (1993). Omitted-ability bias and the increase in the return to schooling. *Journal of Labor Economics*, *11*, 521–544.
- Boumahdi, R., & Plassard, J.-M. (1992). Note à propos du caractère endogène de la variable éducation dans la fonction de gains. *Revue Économique*, *43*, 145–156.
- Brunello, G., & Miniaci, R. (1999). The economic returns to schooling for Italian men: An evaluation based on instrumental variables. *Labor Economics*, *6*, 509–519.

- Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling (Working Paper No. 4833). Cambridge, MA: National Bureau of Economic Research.
- Card, D. (1994). Earnings, schooling, and ability revisited (Working Paper No. 4832). Cambridge, MA: National Bureau of Economic Research.
- Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (vol. 3-A, chap. 30, pp. 1801–1863). Amsterdam: North-Holland.
- Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69, 1127–1160.
- Chiswick, B. R. (1983). The earnings and human capital of American Jews. *Journal of Human Resources*, 18, 313–336.
- Flabbi, L., Paternostro, S., & Tiongson, E. R. (2008). Returns to education in the economic transition: A systematic assessment using comparable data. *Economics of Education Review*, 27(6), 724–740.
- Fougère, D., Goux, D., & Maurin, É. (2001). Formation continue et carrières salariales: Une évaluation sur données individuelles. *Annales d'Économie et de Statistique*, 62, 49–69.
- Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica*, 45, 1–22.
- Griliches, Z. (1979). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy*, 87, S37–S64.
- Harmon, C., & Walker, I. (1995). Estimates of the return to schooling for the United Kingdom. *American Economic Review*, 85, 1278–1286.
- Ismail, R. (2007). *The impact of schooling reform on returns to education in Malaysia* (Munich Personal RePEc Archive Paper No. 15021). Munich, Germany: University Library of Munich.
- Kuch, P., & Haessel, W. (1979). *An analysis of earnings in Canada*. Ottawa: Statistics Canada.

- Maluccio, J. (1998). *Endogeneity of schooling in the wage function: Evidence from the rural Philippines* (Working Paper No. 54). Washington, DC: International Food Policy Research Institute.
- Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66, 281–302.
- Mincer, J. (1974). *Schooling, experience and earnings*. New York, NY: Columbia University Press.
- Ozdural, S. (1993). Intergenerational mobility: A comparative study between Turkey and the United States. *Economics Letters*, 43, 221–230.
- Selz, M., & Thélot, C. (2004). The returns to education and experience: Trends in France over the last thirty-five years. *Population*, 59, 9–47.
- Simonnet, V. (1996). Mobilité professionnelle et salaire: Des différences entre hommes et femmes. *Économie et Statistique*, 299, 59–71.
- Steunou, P.-Y. (2003). *Capital humain et hétérogénéité: Théorie et application*. Doctoral thesis, UFR de Droit et Sciences Économiques, Université du Maine.
- Tansel, A. (1994). Wage employment, earnings and returns to schooling for men and women in Turkey. *Economics of Education Review*, 13, 305–320.
- Tomes, N. (1983). Religion and the rate of return on human capital: Evidence from Canada. *Canadian Journal of Economics*, 16, 122–138.
- Trostel, P., Walker, I., & Woolley, P. (2002). Estimates of the economic return to schooling for 28 countries. *Labor Economics*, 9, 1–16.
- Viger, E. (2007). *Les effets de la démocratisation de l'enseignement en France: Une étude empirique*. Doctoral thesis, UFR de Sciences Économiques, Université Paris 1 Panthéon–Sorbonne.
- Zhang, X. (2011). The rate of returns to schooling: A case study of urban China. *International Journal of Humanities and Social Science*, 1, 173–180.

*Appendix***Table A1: Summary statistics**

Variable	Mean	Median	SD	SE	Kurtosis	Skewness
LNWAGE	7.26	7.24	0.49	0.003	1.50	-0.33
SCH1	11.66	11.00	2.93	0.02	0.58	-0.05
BEFEX2	9.51	6.00	9.76	0.06	0.31	1.08
BEFEX22	185.62	36.00	305.92	1.86	5.77	2.31
EXP2	10.23	6.60	10.19	0.06	0.15	1.09
EXP22	208.33	43.56	336.95	2.05	3.67	2.05
HOURS3	144.65	151.00	31.21	0.19	3.52	-1.76
Z1	11.51	11.25	2.54	0.02	0.52	0.16
Z2	10.21	10.77	1.95	0.01	0.13	-0.95

N = 27,136