We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,400
Open access books available

117,000
International authors and editors

130M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Returns to Education and Experience Within the EU: An Instrumental Variable Approach for Panel Data

Inmaculada García-Mainar and Víctor M. Montuenga-Gómez
University of Zaragoza
Spain

1. Introduction

Estimating the returns to education and experience has been a topic for labour economists for decades, with a significant volume of research being devoted to appraising the causal effect of schooling on earnings. One central interest when estimating these returns has been to study whether differences exist across several demographic sectors; essentially, distinguishing between males and females, whites and non-whites, or natives and immigrants, to assess the possibility of wage discrimination, and to compute the extent of the wage gap between the groups (Harmon et al., 2003).

A rapidly growing literature examines differences in the return to education, distinguishing between the self-employed and wage earners (Evans & Leighton, 1989; Hamilton, 2000; Lazear & Moore, 1984; Rees & Shah, 1986). Fundamentally, these studies have set out not only to investigate earning differentials between the two employment groups *per se*, but also to test competing views about the relationship between earnings and education, on the basis that these groups are subject to different economic incentives. In this context, the self-employed can be used as a control group to distinguish between human capital and sorting models of wage determination, assuming that signaling or screening functions are much less relevant for the self-employed (Riley, 1979; Wolpin, 1977). Returns to experience for the self-employed have also been estimated against those for wage earners, in order to test different theories of the labor market, such as those of agency and risk hypotheses, against the learning and matching models, and against the compensating differentials premises, for example. Thus, so long as the self-employed have fewer incentives to shirk on the job, or to quit it, they should exhibit flatter earnings-experience profiles, since wage earners obtain higher earnings than the self-employed as they grow older (Salop and Salop, 1976).

In this chapter, we set out to estimate the returns to education and to experience for both the self-employed and wage earners, with our aim being to address some of the issues raised above. In doing so, we provide evidence of such returns for three EU countries, namely Germany, Italy and the UK, using panel data information. Examining cross-country data is helpful in identifying common features that are not considered in a single-country analysis. Moreover, these countries cover a wide range of variation across Europe. Germany represents those countries with self-employment rates well below the EU average; Italy is an example of those Southern and peripheral countries with self-employment rates over 20%,

and the UK stands for those countries exhibiting the average. Table 1 shows the 20-year evolution of the self-employment rate within the EU-15.

	1987	1995	2000	2007
Austria	-	10.8	10.5	14.4
Belgium	15.3	15.4	13.6	12.8
Denmark	9.2	8.4	8.0	8.9
Finland	-	14.3	12.6	12.6
France	12.7	11.6	10.0	9.0
Germany	9.1	9.4	9.7	12.0
Greece	35.4	33.8	31.3	35.9
Ireland	21.8	20.8	16.5	16.8
Italy	24.4	24.5	23.6	26.4
Luxembourg	9.2	10.0	8.7	6.1
Netherlands	10.1	11.5	10.0	12.4
Portugal	27.2	25.8	20.2	24.2
Spain	23.5	21.8	18.0	17.7
Sweden	-	11.3	9.8	10.6
United Kingdom	12.5	13.0	10.9	13.8
EU15	15.9	15.0	13.6	15.6

Table 1. Self-employment rates (Eurostat Labour Force Survey). Note: Percentage of self-employed persons over total employed.

Using a homogeneous database, we investigate whether differences exist in the profitability of schooling and experience, both between the two employment statuses, and across the three sample countries. The database used in this work has been obtained from the European Community Household Panel (ECHP), from 1994 to 2000, which provides abundant information about the personal and labour characteristics of individuals, and has the advantage that this information is homogeneous across the sample countries, given that the questionnaire is the same and the collection process is coordinated by the Statistical Office of the European Community (EUROSTAT). Additionally, the application of panel data techniques offers some clear advantages over the traditional cross-sectional approaches. First, individual unobserved heterogeneity is controlled for. Second, biases arising from aggregation, selectivity, measurement error, and endogeneity can be dealt with in an appropriate form. Third, dynamic behavior, such as the movements into and out of self-employment, may be explicitly accounted for.

The two usual estimators used in panel data, that is to say fixed and random effects, are, however, inadequate in this setting if the objective is to obtain consistent and efficient measures of the profitability of education and experience. Thus, the presence of time-invariant and possibly endogenous regressors (e.g. education), would make it impossible to estimate their associated parameters when a time or mean-differencing approach, i.e. the fixed effects, is applied. Additionally, the probable correlation between these time-invariant regressors and the unobserved effects causes the random effects estimator to be inconsistent. Altogether, this points to the advisability of using a hybrid technique that lies between both extremes. Moreover, the potential existence of measurement errors and/or endogeneity requires instrumental variables (IV) to obtain consistent estimates of the coefficients. As a consequence, the Hausman & Taylor (1981) estimator is the most adequate, since it has been

shown to be an Efficient Generalised Instrumental Variable (EGIV). This procedure allows simultaneous control for the correlation between regressors and unobserved individual effects (as fixed effects) and to identify the estimates for the time-invariant covariates, such as education, as a random effects estimator. Furthermore, it eliminates the uncertainty associated with the choice of instruments, since exogenous included variables, and their means over time, are used as efficient instruments.

Our results show that returns to education are greater for workers in paid work, with non-linearities in the relationship between wages and educational levels (the so-called sheepskin effect). Both findings point to the relevance of signalling in the earnings of workers. Earnings experience profiles are, however, not very different across groups, especially when experience is not very great, indicating absence of delays in remuneration for workers.

The rest of the chapter is structured as follows. In the next section, we consider the theoretical aspects of the returns to education and experience. Section 3 is devoted to the empirical model and to a discussion of the estimation procedure, the EGIV technique proposed in Hausman and Taylor (1981). Section 4 describes the data in the sample countries. Section 5 offers the estimates of the rates of return and examines the differences across countries to cast some light on labour-status differences. Finally, Section 6 provides a summary of the main results.

2. Theoretical aspects of returns to schooling and experience in relation to self-employment

A new-born child enjoys an initial endowment of human capital (a conglomerate of intelligence, ability, motivation, characteristics of the social and economic environment, etc.) that can be improved upon by knowledge accumulation, both during the schooling period and through on-the-job experience. According to the human capital theory (Becker, 1962, 1964), there exists a positive relationship between investment in human capital and earnings, in such a way that a greater accumulation of human capital is rewarded in the labor market with higher earnings.

The individual chooses to stay in school until the expected marginal benefit equals the expected marginal costs of one additional year of schooling, and differences in ability among individuals cause schooling choices to also differ. Empirically, a linear relationship is usually assumed between the log of the earnings and the set of regressors. This implies that ability influences only the intercept of log-earnings. Following this reasoning, we apply the widely-used wage equation (Mincer, 1974) that can be expressed as:

$$\ln w_t = \alpha\, f(A_t) + \beta\, g(Edu_t) + \gamma\, h(Exp_t) + \eta'Char_t + \varepsilon_t \qquad (1)$$

where w denotes earnings, A the initial human capital, or ability, Edu is the education variable, Exp is the experience and $Char$ a set of personal and labor characteristics (such as gender, age, occupation, type of contract, etc.) which can be time-constant or time-varying. Since ability is usually unobservable to the researcher, this must be included in the error term. However, this ability may be correlated with schooling, in such a way that standard least squares yield biased estimates (Griliches, 1977). This issue will be discussed further in Section 3.2.

Although specification (1) has been derived on the grounds of human capital theory, competing perspectives may generate similar conclusions. In particular, the sorting or signaling model also predicts that higher earnings are positively related to higher

educational attainments. However, in this case, greater human capital does not lead to higher productivity (and thus, higher earnings), but greater human capital is acquired in order to signal higher productivity (Spence, 1973; Stiglitz, 1975). In other words, firms do not reward productivity in a direct way because this is not observed *a priori*; rather, they infer productivity from education, and students choose an education level to signal their productivity to potential employers. Similarly, firms offer higher wages for the highly educated, since education acts as a screening device, so long as education is positively correlated with the unobserved productivity.

As a consequence, estimating equation (1) does not help to discriminate between human capital and the sorting models; while it may be viewed as a good approach to assessing the effect of schooling on earnings, it is not completely satisfactory in elucidating which view prevails in the process of wage determination (Weiss, 1995). However, considering the self-employed as a control group may serve as a device to investigate the question, since signaling and screening purposes seem to be unimportant for this group of workers (Riley, 1979; Wolpin, 1977). The null hypothesis adopted by these authors is that returns to education will be higher in those occupations that exhibit signaling, on the basis that it is difficult to reconcile the idea that education for the self-employed could act as a sorting mechanism. As a consequence, returns to schooling for those in paid employment should be higher, since those individuals benefit from the dual effect of education: the productive and the informative functions. By contrast, the self-employed are only remunerated for the productive nature of education and, thus, returns are lower.

However, although the theoretical implications seem quite clear-cut, the empirical evidence is not conclusive. Focusing on the US, some authors report that self-employed earnings are less responsive to human capital variables than wage-employed earnings (Hamilton, 2000), thereby favouring the sorting hypothesis., whereas others (Evans and Jovanovic, 1989; Evans and Leighton, 1989; Kawaguchi, 2003) find that self-employed earnings equations have larger schooling coefficients than those corresponding to the wage-employed, rejecting the sorting hypothesis.

Distinguishing between self-employed and wage-earner returns may also be helpful in providing insights into the features of theoretical labour market models. Thus, studying the experience profile in earnings may serve to ascertain whether agency issues, learning and matching models, or compensating differentials theory, for example, better fit the labour market. A number of studies predict that earnings-experience profiles are flatter for the self-employed. Under the agency or risk theories (Lazear, 1981; Lazear & Moore, 1984), employers should pay less than the marginal productivity to workers when they are young, and more when they grow older, to avoid shirking on the job, contrary to the case of the self-employed, given that these individuals have no incentive to shirk. Similarly, asymmetric information models (Salop & Salop, 1976) argue that, because employers are interested in minimizing the quits of more productive workers, they offer tilted-up wage profiles as a screening device, in such a way that only workers with low probabilities of quitting apply for jobs. By contrast, since the self-employed are not willing to quit, they have flatter earnings profiles than those of wage earners. In the same vein, learning models claim that, due to sector-specific abilities that are unknown for the individual, workers may not match themselves to the appropriate sector. Those who realize they have a poor match quit their jobs, and only those with relatively good matches stay. This situation causes experience profiles to increase over time (Jovanovic, 1979, 1982). Furthermore, since the self-employed habitually invest strongly at the start-up of their businesses, they are not able to move out of

their poor match, and therefore their experience profiles are flatter (Dunn & Holt-Eakin, 2000). All of these studies suggest steeper experience profiles for wage earners, especially at the very end of the years of experience distribution, to ensure long seniority in the firm.

Contrary results, however, are found in the investment model, for example, which validates the claim that the self-employed obtain steeper earnings profiles because physical and human capital investments are not shared with an employer (Hashimoto, 1981). Similarly, steeper earning profiles of the self-employed can be observed in the cases where average returns to experience are distorted by the existence of a few, but very successful, entrepreneurs (the so-called "superstars"), whereas the bulk of the self-employed remain with low returns or leave for paid employment (Rosen, 1981).

In summary, undisputed conclusions about the magnitude of the returns to education and experience for the self-employed and for salaried employees have not been achieved. The majority of prior analyses have focused on investigating only one country, without offering any kind of comparative study. Furthermore, only a limited number of articles have used data for a period of more than one year. Even when they have done so, they have estimated returns by pooling the data, an approach which does not allow for control for the unobserved characteristics of the individuals, nor for movements into or out of self-employment. Additionally, only the recent availability of panel data and the development of new statistical packages have permitted the application of EGIV techniques in order to gain efficiency in estimations. The aim of this chapter is precisely to address some of these gaps in the literature by computing returns to education and experience for a set of EU countries, using information provided in panel data form, which are statistically efficient.

3. The empirical model and estimation procedure

This section focuses on the empirical specification of the earnings equation and the methodology used for its estimation. The first sub-section is devoted to determining the most appropriate empirical model for our study, while the second describes the logic supporting the use of the Hausman-Taylor procedure in the estimation.

3.1 Empirical specification

As discussed in Section 2, estimates of the returns to education and experience for the self-employed and wage earners are usually obtained from Mincer-type wage regressions. Dating from the mid 20th century, a body of empirical work has investigated these returns across countries on the basis of such a specification (Psacharapoulos, 1973, 1981, 1985, 1994; Trostel et al., 2002). Cross-sectional information has normally been used, with the IV approach, progressively substituting for the traditional OLS estimation to account for endogeneity and ability biases, as well as measurement errors. This has resulted in estimates of the rates of return well above those obtained from OLS (Card, 1999, 2001).

Our estimated model is an extended version of the Mincerian-baseline equation (1), in which earnings rewarding more education can be seen as the combined effect of human capital accumulation and the effect of being identified as a graduate rather than as a dropout. It takes the following form:

$$\ln w_{it} = \beta Edu_i + \mu_1 Exp_{it} + \mu_2 Exp^2_{it}/100 + X'_{it} \delta + Z'_i \gamma + u_{it}, \qquad (2)$$

where i and t stand for the N individuals and the T time periods, respectively. As indicated before, w denotes earnings; Edu is the education variable (that is considered time-invariant);

Exp is the experience; *X* is a set of time-varying regressors, and *Z* is a set of time-invariant regressors. The *β* coefficient expresses the rate of return to education; μ_1 and μ_2 represent the earnings-experience profile, and *δ* and *γ* are the set of parameters accompanying the remaining regressors.

Rather than using the education measure normally employed in the literature, years of schooling, we consider the educational attainment of each worker, providing two clear advantages. First, it does not impose the annual marginal effect of schooling as being the same in each year of education, and second, the level of education is a more appropriate measure, since multiple education streams characterize European countries, and salary profiles used to be largely linked to the education category attained. In other words, they capture the possibility that credentials matter more than years of schooling *per se* in the wage determination. This hypothesis is commonly known as the "sheepskin effect" and attempts to explain the discontinuous changes in earnings associated with completion of elementary school, high school or college (Belman & Heywood, 1991; Hungeford & Solon, 1987; Jaeger & Page, 1996).

Educational attainment, which is considered time-invariant in our sample, represents the last completed level of schooling, classified as primary, secondary, and high. Primary includes elementary and below elementary school, secondary includes vocational and middle school, and tertiary or high includes university studies (in either short or long cycles). Consequently, the fragment "*β Edu$_i$*" in equation (2) would be represented by "*$β_1$ EduS$_i$* + *$β_2$ EduH$_i$*". The category of reference is *EduP$_i$*, omitted in the estimation.

$$\ln w_{it} = \beta_1 \, EduS_i + \beta_2 \, EduH_i + \mu_1 \, Exp_{it} + \mu_2 \, Exp^2_{it}/100 + X'_{it} \, \delta + Z'_i \, \gamma + u_{it}, \qquad (3)$$

The dependent variable is the natural log of net earnings, where these are defined as gross earnings less tax, expressed in per hour real terms. The earnings-experience profiles are analyzed by considering the number of years that an individual has been working, and its squared value divided by 100 to take care of the decreasing returns. Specifically, experience is measured as the difference between the current age and the age of initiation at work, thereby expressing potential experience. The remaining independent variables, represented in equation (3) by *X* and *Z*, contain dummies for gender, marital status, training, occupation, sector, seniority, and a set of time fixed effects, as described in Section 4.

3.2 Estimating the earnings equation: the Hausman-Taylor procedure

The general model in (3) assumes that the error term u_{it} consists of the sum of two components, i.e. $u_{it} = a_i + v_{it}$, where a_i represents the random individual-specific effect that characterizes each worker and is constant through time, and v_{it} is a random disturbance varying over time and individuals. This latter stochastic term is assumed to be uncorrelated with all included variables. Similarly, it is also assumed that the random disturbance is a sequence of i.i.d. random variables with mean zero and variance σ^2_v; that v_{it} and α_i are mutually independent, and that α_i is i.i.d. over the panels with mean zero and variance σ^2_α. Thus, the variance-covariance matrix of the system has the random effects structure that can be represented as $E(UU') = \sigma^2_\alpha \, (i_T i_T' \otimes I_N) + \sigma^2_v \, (I_T \otimes I_N)$, where i_T is a *T*x1 vector containing 1s, I_N (I_T) is the identity matrix of rank *N* (*T*), and *U* is an *NT*x1 vector of disturbances. Thus, random-effects or Generalised Leasts Squares (GLS) produce consistent estimators.

However, the presence of measurement errors and unobserved variables, such as ability, motivation, etc., that may be correlated with schooling, bias GLS estimates. Specifically, it

has been shown that measurement errors bias downwards the GLS estimates (Angrist and Krueger, 1999) although recent evidence (Card, 2001) only attributes a 10% gap, at most, to this source of bias. By contrast, since schooling and any unobserved ability may be positively correlated, omitting measures of ability results in the schooling coefficient being biased upwards (Griliches, 1977). Consequently, some effort must be made to alleviate such an ability bias as much as possible. When a direct indication of ability, such as IQ score tests, or information from twins or siblings, is not available (see Ashenfelter & Krueger, 1994, and Miller et al., 1995), the most appropriate exercise is to select an instrumental variables estimator, through which schooling is instrumented with variables that are correlated with it, but not with errors. A broad range of instruments have been proposed in the literature. Typical examples are those known as natural experiments (see Rosenzweig and Wolpin, 2000, for a summary) which include: i) school reforms and features of the school system (Harmon and Walker, 1995); ii) the proximity to College of the place of residence (Card, 1995); iii) other supply-side instruments capturing features of the education system (see Card, 2001 for a survey of the literature); and iv) the season of birth of the individual (Angrist and Krueger, 1991).

When using IV in cross-sections, a common finding is that estimates are 20% higher, or even more, than OLS estimates. This is a rather unexpected result, since OLS is already believed to provide upward-biased estimates arising from the ability bias. Some reasons have been proposed to explain such a result. Apart from the positive publication bias (Ashenfelter et al., 1999), IV estimates may be biased upwards further than OLS due to the existence of unobserved differences between the characteristics of the treatment and comparison groups implicit in the IV scheme (Bound et al., 1995). Specifically, when treatment effects are used, since returns to education are heterogeneous across individuals, the IV estimates tend to recover the returns to education of the population group most affected by the intervention, so that IV estimates are then a better approximation for the returns to education of the affected group, rather than for the whole population (Card, 1999, 2001). Similarly, IV estimates will tend to be biased towards the returns to schooling attainments that are most common in the sample data (see Belzil and Hansen, 2002).

Both the available data structure and the existence of problems associated with the choice of instruments have influenced the procedure applied in this study. On the one hand, the ECHP is in panel data form, but does not provide information on IQ tests, and the presence of twins is not especially accounted for. On the other hand, although the number of alternative instruments routinely considered in the literature is sufficiently wide, their application to our data is quite complex. This has led us to consider an alternative procedure for estimation, in which the availability of panel data is taken into account, namely the IV-type model proposed by Hausman and Taylor (1981). Our selection of this procedure is motivated by several considerations. As is well known, the availability of panel data allows us to control for individual unobserved heterogeneity (possibly correlated with other included variables), since this factor may be eliminated by mean or time-differencing, i.e. by applying a fixed effects-type estimator (Polachek & Kim, 1994). Although this within estimator is probably not fully-efficient, it produces consistent estimates. However, when operating in this way, coefficients of the time-constant variables (e.g. the level of education) cannot be estimated, since they disappear when mean or time-differences are constructed. For its part, a pure random effects estimator, the GLS, produces biased and inconsistent estimates, assuming as it does that there is no correlation between any of the regressors and

the individual effects. In our case, the GLS estimator is not valid because education and experience may be correlated with individual effects.

One way to obtain consistent estimates of the returns to education and experience would be to find instruments for these variables which are potentially correlated with the individual effects. The choice of the appropriate instruments is, however, not an easy task, since the use of instruments that are weakly correlated with endogenous variables may produce downward-biased estimates, even with large samples (see Bound et al, 1995; Chamberlain & Imbens, 2004; Staiger & Stock, 1997), generating uncertainty in the selection of instruments. Consequently, what we require is a procedure that controls for the endogeneity of education (and possibly other variables), but which is still able to recover the coefficient of time-invariant regressors. Hausman & Taylor (1981) propose a model where some of the regressors may be correlated with individual effects, as opposed to the random effects model, where no regressor can be correlated with the individual effect, and to the fixed effects model, where all the regressors may be correlated with individual effects. If, in addition, this procedure does not require instruments excluded in the regression but the instruments used are precisely those included in the wage regression, the Hausman-Taylor estimator is, potentially, the best choice.

This Hausman-Taylor estimator is an instrumental variables estimator that uses both the between and within variations of the strictly exogenous variables as instruments. More specifically, the individual means of the strictly exogenous regressors are used as instruments for the time invariant regressors correlated with individual effects. This procedure is implemented in the following steps. First, equation (3) is estimated by pooled Two Stages Least Squares (2SLS), where the set of variables mentioned above act as instruments. Second, the pooled 2SLS residuals are used to obtain estimates of σ^2_α and σ^2_v, which can then be used to construct the weights for a Feasible Generalized Least Squares estimator. Third, these weights are used to transform (by quasi-time demeaning) all the dependent, explanatory, and instrumental variables. Finally, the transformed regression is again estimated by pooled 2SLS, where the individual means, over time, of the time-varying regressors, and the exogenous time-invariant regressors, are the instruments. Under the full set of assumptions mentioned in the previous sub-section, this Hausman and Taylor estimator becomes an Efficient Generalized Instrumental Variables (EGIV) and coincides with the efficient GMM estimator.

Formally, the Hausman-Taylor model can be represented in its most general form as follows:

$$\ln w_{it} = \alpha_i + X'_{it}\delta + Z'_i\gamma + v_{it}, \tag{4}$$

where $i = 1, \ldots, N$ and $t = 1,\ldots, T$. The Z_i are individual time-invariant regressors, whereas the X_{it} are time-varying. α_i is assumed to be i.i.d.$(0, \sigma^2_\alpha)$ and v_{it} i.i.d.$(0, \sigma^2_v)$, both independent of each other. The matrices X and Z can be split into two sets of variables $X=[X_1, X_2]$ and $Z=[Z_1, Z_2]$, such that X_1 is $NT \times k_1$, X_2 is $NT \times k_2$, Z_1 is $NT \times g_1$, and Z_2 is $NT \times g_2$. The X_1 and Z_1 are assumed exogenous and not correlated with α_i and v_{it}, while X_2 and Z_2 are endogenous due to their correlation with α_i but not with v_{it}. Hausman & Taylor (1981) suggest an instrumental variables estimator which pre-multiplies expression (4) by $\Omega^{1/2}$, where Ω is the variance- covariance term of the error component $\alpha_i + v_{it}$, and then performing 2SLS using $[Q, X_1, Z_1]$ as instruments. Q is the within transformation matrix with $X^* = QX$ having a typical element $X^*_{it} = X_{it} - \bar{X}_i$ and $\bar{X}_i$ is the individual mean. This is

equivalent to running 2SLS with $[X^*, X_1, Z_1]$ as the set of instruments. If the model is identified, in the sense that there are at least as many time-varying exogenous regressors X_1 as there are individual time-invariant endogenous regressors Z_2, i.e. $k_1 \geq g_2$, this Hausman-Taylor estimator is more efficient than fixed effects. If the model is under-identified, i.e. $k_1 < g_2$, then one cannot estimate γ and the Hausman-Taylor estimator of δ is identical to fixed effects (Hausman & Taylor, 1981; Wooldridge, 2002).

In the case under consideration, education is a potentially endogenous, time-invariant regressor, whereas the experience variables may also be endogenous, but time-varying. Since we are interested in the coefficients of these variables, all the exogenous variables (either time-invariant or time-varying), plus the individual means over time of all the time-varying regressors can be used as instruments to obtain consistent estimates of the returns to education and experience. These instruments are chosen based on Hausman specification tests (Hausman, 1978) in a sequential procedure according to (Baltagi et al., 2003). Specifically, a first Hausman test is the standard to distinguish between the random and fixed effects estimators. A second Hausman test contrasts the Hausman-Taylor against the fixed effects model. Although the fixed effects estimator is not an option in our study, since it does not allow for the estimation of the coefficients of the time invariant regressors, it is useful in order to test the strict exogeneity of the regressors used as instruments in the Hausman-Taylor estimation. Thus, when strict exogeneity for a set of regressors is rejected, others must be considered in the estimation to act as instruments. Once the second Hausman test has identified the regressors that are strictly exogenous, they are used as instruments in the Hausman-Taylor estimation. Additionally, the variance-covariance structure can be taken into account to obtain more efficient estimators (Im et al., 1999), so that the Hausman-Taylor procedure is a good alternative to pure IV estimation when panel data is available.

4. The data and descriptive statistics

The data used in this study come from the ECHP for the period 1994-2000. As stated earlier, this is the only database that provides individual information that is comparable for all EU countries, since the design and organization of the survey is coordinated by EUROSTAT. Individual or micro data is preferred to more aggregate data, both because they provide more flexibility in creating sample restrictions, and because they allow us to directly control for individual-level characteristics in our regression.

At the time of the interview, individuals are requested to indicate whether they are working in a job for at least 15 hours a week. If so, workers identify themselves as either self-employed or employed when asked about their main labor market activity (paid apprenticeships and unpaid work in a family enterprise are excluded from the analysis). As a consequence, the job status of a particular worker may vary from year to year. In the sample, we have selected those workers, either self-employed or wage earners, who have provided information for all variables under consideration. These variables include personal and labor characteristics such as gender, marital status, schooling, experience, earnings, seniority, occupation, whether the individual works in the private or in the public sector, the number of hours worked per week, and if the worker has taken some training course during the last year.

Table 2 illustrates the main characteristics of our samples for the three countries. The number corresponding to wage earners in the sample ranges from about 6,500 in the UK to

more than 8,000 in Germany. For the self-employed, the figures are considerably lower, varying between less than 1,000 in Germany to almost 2,500 in Italy. Sample proportions are not very different from population rates (see Table 1). Bearing in mind that self-employed earnings are commonly believed to be under-reported (Hamilton, 2000), wage earners appear to earn a little more than the self-employed in Italy, whereas the opposite occurs in Germany and in the UK. Note that dispersion in earnings is higher for the self-employed, reflecting a greater heterogeneity in these types of activities, from low-ability jobs (retailers and basic services) to those of professionals, such as doctors or lawyers. People living in Germany and in the UK obtain higher earnings than those residing in Italy.

The years of experience are clearly higher in the self-employed sector than in that of wage earners. The majority of individuals in Italy present low educational levels, the percentage being somewhat higher among the self-employed. By contrast, in Germany, workers in general are highly educated, with more than 40% of the self-employed having attained a tertiary degree. The case of the UK is especially appealing since workers attaining a secondary level are clearly fewer than those of primary or higher education, indicating some kind of a bi-modal distribution. Roughly speaking, where self-employment rates are higher, the self-employed themselves are less educated, as against countries with low self-employment rates, which exhibit a higher proportion of workers, either wage earners or the self-employed, who have obtained at least a secondary level of education.

		Earnings per hour	Exp.	Primary educ.	Secondary educ.	Higher educ.	Obs. per year
Germany	Self-employed	9.06 (8.26)	22.94 (10.98)	12.31	46.19	41.50	840
	Wage earner	8.48 (4.30)	20.48 (11.16)	19.48	57.88	22.64	8066
Italy	Self-employed	6.21 (5.31)	23.69 (13.34)	56.82	31.95	11.23	2494
	Wage earner	6.73 (3.55)	18.03 (11.07)	44.13	44.63	11.24	7865
United Kingdom	Self-employed	8.71 (9.73)	25.33 (13.03)	46.48	13.74	39.78	1053
	Wage earner	7.88 (5.59)	20.02 (12.86)	45.99	13.70	40.31	6433

Table 2. Sample averages (ECHP 1994-2000). Note: Standard errors between parentheses. Earnings are expressed in terms of the PPP. Observations per year is the average, since figures vary from year to year.

5. Estimation results

This section presents the empirical evidence which is then assessed in the light of the aspects mentioned in Section 2, with the aim of providing some insights into the functioning of the European labour markets. The results from EGIV Hausman-Taylor estimations are shown in Table 3, along with the tests for choosing the appropriate instruments (Baltagi et al., 2003). In column H1, a standard Hausman test rejects the random effects hypotheses in favour of

the fixed effect estimator. A second Hausman test contrasting the Hausman-Taylor against the fixed effects model (column H2), is useful in order to test the strict exogeneity of the regressors used as instruments in the Hausman-Taylor estimation. Once this second Hausman test has identified which regressors are strictly exogenous, they are then used as instruments in the Hausman-Taylor estimation.

Comparing the coefficients of Table 3 with those from the GLS estimation (not shown, but available from the authors upon request), we can note that the Hausman-Taylor estimation provides coefficients of education and experience that, in general, are consistently much higher. This is in accordance with the typical finding reported in the literature when using instrumental variables of upward bias in IV-type, compared to OLS-type estimations.

Regarding the returns to education, wages increase with educational attainment, with returns higher among wage earners in Germany and Italy and quite similar in the UK, and the coefficients for the self-employed with secondary education level in Germany and the UK are non-significant. The percentage changes across educational categories, computed as the difference in the percentage change in wage for group i-1 relative to the group i-1, $e^{\beta i}- e^{\beta i-1}$, where β_i is the coefficient for the dummy variable for group I, show that returns increase as we move up the qualification ladder, especially from secondary to higher education, which supports a convex configuration of earnings on the returns to education, thus confirming the importance of the sheepskin effect. Both results, higher returns to education for paid workers, and increasing non-linearities in the relationship between wages and educational attainment, indicate some degree of a sorting or signalling role played by education.

Country	Labour status	Exp.	Exp.^{2}/100	Second. education	Higher education	H1	H2
Germany	Self-employed	0.053**	-0.068**	0.654	1.243*	234.96	8.44
		(3.72)	(-2.83)	(0.89)	(2.45)	(0.0000)	(0.9986)
	Wage earners	0.065**	-0.084**	0.848**	1.291**	1117.89	10.51
		(22.24)	(-22.21)	(8.83)	(17.31)	(0.0000)	(0.9921)
Italy	Self-employed	0.037**	-0.053**	0.339**	0.631**	138.47	34.66
		(6.30)	(-4.66)	(2.78)	(6.35)	(0.0000)	(0.0946)
	Wage earners	0.047**	-0.089**	0.551**	0.847**	2103.26	49.63
		(27.52)	(-23.60)	(15.09)	(19.07)	(0.0000)	(0.0752)
United Kingdom	Self-employed	0.053**	-0.055**	0.629	0.714**	275.90	5.92
		(5.16)	(-3.41)	(0.74)	(3.54)	(0.0000)	(0.9999)
	Wage earners	0.063**	-0.106**	0.524**	0.706**	2064.11	4.58
		(27.27)	(-27.28)	(3.41)	(15.14)	(0.0000)	(1.0000)

Table 3 Estimated Coefficients of Mincerian Earnings Function by Hausman-Taylor. Notes: t-ratios between parentheses (p-values in H1 and H2). Both panels are unbalanced, since the employment status may vary across individuals over time. Controls used. Gender: 1 for male and 0 for female. Marital status: married, single, divorced, widow or separated. Training: if the worker has realized some course of occupational training. Eight dummies that indicate occupation. A dummy indicating whether the individual works in the private or the public sector. Dummies that indicate seniority: less than two years, between 2 and 10 years, and more than 10 years. Dummies that indicate the year. * Significant at the 5% level. ** Significant at the 1% level. H1 tests the random effects estimator against the fixed effects. H2 tests the Hausman-Taylor estimator against the fixed effects.

Across countries, experience increases human capital accumulation during the life cycle. At first sight, returns to experience are greater for wage earners, even though they depreciate at a faster rate than in the case of the self-employed. In order to extract more robust conclusions, a series of indicators are used. First, the maximum return, i.e. the point where experience ceases to add positively to earnings, which is defined by $\partial \ln w / \partial Exp$ from earnings equation (3), that is to say, the number of years that equals $\mu_1 + \mu_2 \, Exp/50$ to 0 provided $\mu_2 < 0$. The third column in Table 4 is viewed as always being greater for the self-employed, but with marked differences across countries. Thus, the maximum number of years is almost equal in Germany, close to 39 years, with differences around ten years in Italy and almost 20 in the UK. In this latter case, experience is continually adding to earnings during the whole working life of the self-employed. The effects of experience are less long-lasting in Italy, especially among wage earners.

Country	Labour Status	Maximum rate	At sample average
Germany	Self-employed	39	2,18
	Wage earner	38,7	2,36
Italy	Self-employed	35	1,61
	Wage earner	26	1,49
United Kingdom	Self-employed	48	2,51
	Wage earner	30	2,06

Table 4. Returns to experience. Note: Own calculations from the estimated coefficients obtained in Tables 3.

We have computed the series of rates of return as $\mu_1 + \mu_2 \, Exp/50$, with "Exp" playing the role of a variable. Column 4 in Table 4 reports the rate of return evaluated at the sample average in each country. Values are quite similar in Germany and the UK, and clearly lower in Italy. The greatest difference between the self-employed and wage earners is found in the UK, close to 0.5.

Earnings-experience profiles are constructed from the series of rates of return to experience. Figure 1 displays these profiles for both types of workers in the sample countries. It is clear from the evidence that the profiles are very similar during the first years, (14 in Italy, around 20 in the UK and almost 38 in Germany), with the profiles being slightly steeper for wage earners; then the profiles switch position, revealing the long-lasting effects of experience for the self-employed.

Overall, the body of evidence seems to indicate that investment considerations may be at work, al least in the long run. This can be due to the fact that returns to other capital accumulation, physical or technological, are more long-lasting than those from human capital alone. Alternatively, it can be argued that, if mobility costs are reduced, only well-matched self-employed workers remain as such, with less successful entrepreneurs leaving self-employment and undertaking paid employment. Taken together, it may be reasoned that competitive functioning of the labour market may be at work in these countries. While the different theories cannot be compared one with another in the absence of a more detailed analysis, it nevertheless appears that imperfections in the labour market play a less important role than expected.

In summary, as regards the functioning of the labour market in the set of EU countries considered in this Chapter, two basic ideas emerge. First, returns to education are, in
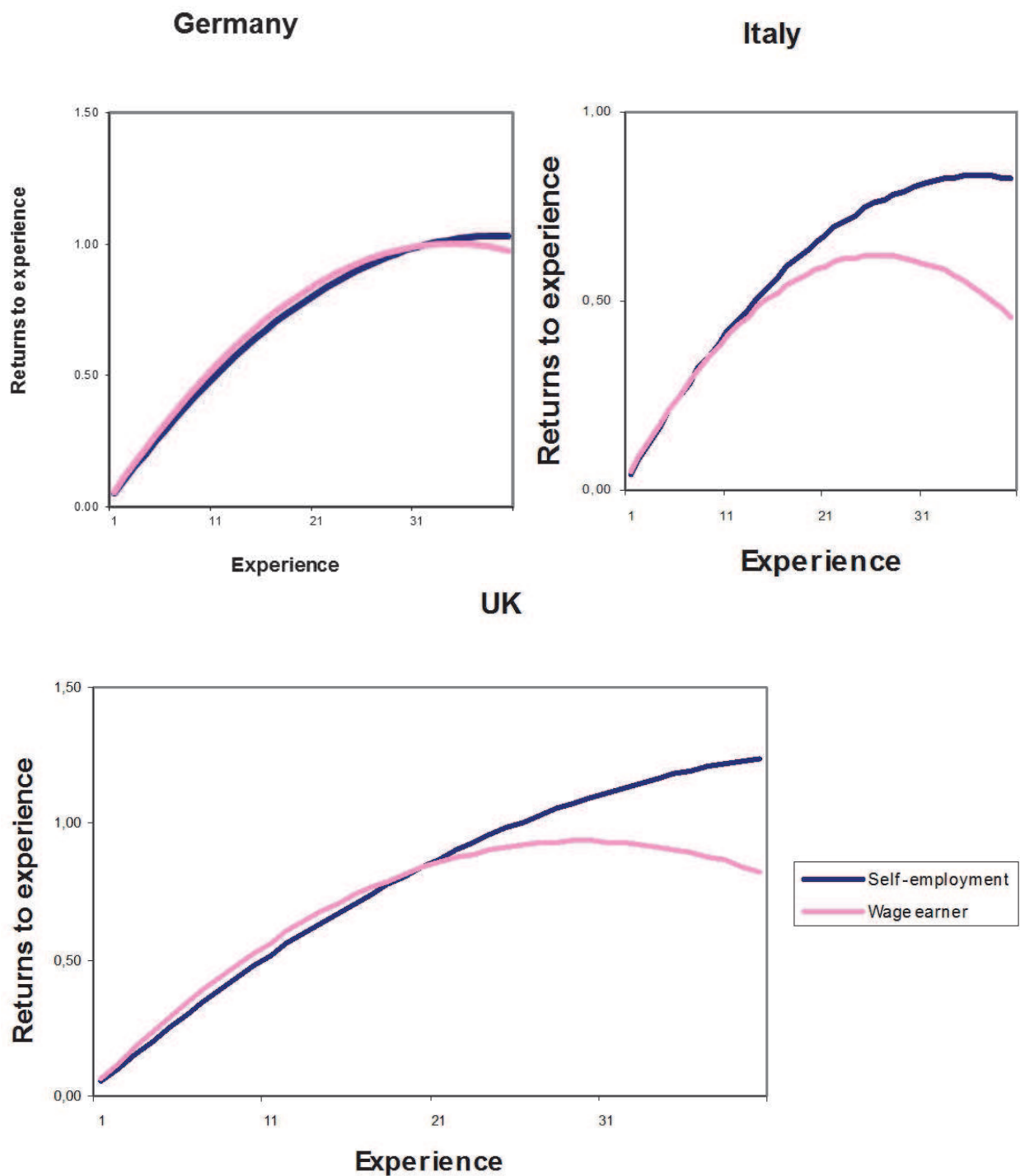
Germany

Italy

UK

Fig. 1. Earnings experience profiles for the three sample countries

general, found to be higher for wage earners, which can be interpreted as an indication of the relevance of the signalling role of education in determining earnings. This latter result was expected, bearing in mind the prevalent wage rates in the EU countries, where earnings are usually linked to the education level attained by the worker. Second, according to the evidence shown by the earning-experience profiles, which tend to be steeper in the case of the self-employed, traits of competitiveness can be discerned, with little or no evidence of imperfections.

6. Conclusions

The aim of this Chapter has been to extend the existing research on the returns to human capital accumulation that differentiates between the self-employed and wage earners. This has been carried out by providing evidence in a cross-country framework using a homogenous database, which mitigates the problems associated with the existence of different data sources across countries, by using a panel data approach that is useful in dealing with endogeneity and selectivity biases, as well as unobserved heterogeneity, and by applying an efficient estimation method that allows for the correlation between individual effects and time-invariant regressors, and that avoids the insecurity associated with the choice of the appropriate instruments.

Information from the ECHP for the period 1994-2000 has been used, allowing us to apply an Efficient Generalised Instrumental Variable estimator that provides consistent estimates of the rates of return to education and experience. Education has been represented by dummies of qualification levels (primary, secondary and higher), and experience has been measured as the difference between the current age and the age of initiation at work. The results have been presented in a reduced form, with the aim being to provide both comparisons across countries about the earnings differentials between the two employment statuses analyzed, and evidence as to whether such differences are consistent with the predictions offered by a variety of theoretical models.

The self-employed have been used as a control group to help in assessing the true impact of credentials achieved in the process of wage determination, as well as in determining which type of theoretical structure underlies labour market behavior. We have operated under the premise that, on the basis that signalling is of much less importance for the self-employed, comparing across both types of employment statuses should show that, for the sorting hypothesis to be accepted, returns to education for wage earners are significantly higher than those for the self-employed, as well as possibly increasing in a non-linear way. Similarly, most labour market models based on imperfect information predict steeper experience-earnings profiles for wage earners, whereas competitive traits in the labor market would imply similar or flatter profiles for this category of worker.

The evidence that emerges for the sample countries tends to support the view that signalling theory is indeed relevant in determining individual earnings, in that, first, returns to education are lower where signalling is expected to play a less important role, i.e. in the case of the self-employed, and, second, certain non-linearities appear. Furthermore, earnings-experience profiles are found to be steeper for the self-employed in the long-run, indicating a certain significance of competitiveness in the labour markets.

Some aspects of the investigation have been omitted or require further attention. We are conscious that selectivity issues should be carefully dealt with, when the development of a reliable instrument makes this possible (Semykina & Wooldridge, 2010). Furthermore, obtaining structural estimates for the returns to education and experience would probably require dynamic programming models of occupational choice (Belzil & Hansen, 2002). Finally, the availability of richer panel data sets is of particular importance to control for the movements into and out of self-employment. These topics are all matters for future research, and will undoubtedly be helpful in carrying out a more in-depth investigation into the behaviour of the labour market and wage determination in the EU countries, in such a way that we can more fully assess their degrees of competitiveness.

7. Acknowledgements

8. References

Angrist, J. & Krueger, A (1991). Does Compulsory Schooling Attendance Affect Schooling and Earnings? *Quarterly Journal of Economics*, Vol. 106 (November 1991), pp. 979-1014, ISSN 0033-5533.

Angrist, J. & Krueger, A (1999). Empirical Strategies in Labor Economics. In *Handbook of Labor Economics*, 3A, O. Ashenfelter & D. Card (Eds.), pp. 1277-1366, Elsevier Science, ISBN 978-044-4534-68-2, Amsterdam, The Netherlands.

Ashenfelter, O. & Krueger, A. (1994). Estimates of the Economic Return to Schooling from a New Sample of Twins. *American Economic Review*, Vol.84, (December 1994), pp. 1157-1173, ISSN 0002-8282.

Ashenfelter, O.; Harmon, C. & Oosterbeek, H. (1999). A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias. *Labour Economics*, Vol.6 (November 1999), pp. 453-470, ISSN 0927-5371.

Baltagi, B.; Bresson, G. & Pirotte, A. (2003). Fixed Effects, Random Effects or Hausman-Taylor? A Pretest Estimator. *Economics Letters* Vol.79 (June 2003), pp. 361-369, ISSN 0165-1765.

Becker, G. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, Vol.70, pp. 9-49, ISSN 0022-3808.

Becker, G. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* National Bureau of Economics Research, ISBN, 978-0226041209, New York, USA.

Belman, D. & Heywood, J. (1991). Sheepskin Effects in the Return to Education. *Review of Economics and Statistics*, Vol.73 (November 1991), pp. 720-724, ISSN 0034-6535.

Belzil, C. & Hansen, J. (2002). Unobserved Ability and the Return to Schooling. *Econometrica*, Vol.70 (September 2002), pp, 2075-2091, ISSN 0012-9682.

Bound, J.; Jaeger, D. & Baker, R. (1995). Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and Endogenous Explanatory Variable Is Weak. *Journal of the American Statistical Association*, Vol.90 (June 1995), pp. 443-450, ISSN 0162-1459.

Card, D. (1995). Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp* L. Christofides, K. Grant & R. Swidinsky (Eds.), pp. 201-22, University of Toronto Press, ISBN 0802028640, Toronto, Canada.

Card, D. (1999). The Causal Effect of Education on Earnings." In *Handbook of Labour Economics*, vol. 3A, O. Ashenfelter & D. Card (Eds.), pp. 1801-1863, Elsevier Science, ISBN 978-044-4534-68-2, Amsterdam, The Netherlands.

Card, D. (2001). Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, Vol.69 (September 2001), pp. 1127-1160, ISSN 0012-9682.

Chamberlain, G. & Imbens, G. (2004). Random Effects Estimators With Many Instrumental Variables. *Econometrica*, Vol.72 (January 2004), pp. 295-306, ISSN 0012-9682.

Dunn, T. & Holtz-Eakin, D. (2000). Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links. *Journal of Labor Economics*, Vol.18 (April 2000), pp. 282-305, ISSN 0734-306X.

Evans, D. & Jovanovic, B. (1989). An Estimated Model of Entrepreneurial Choice under Liquidity Constraints. *Journal of Political Economy*, Vol.97 (August 1989), pp. 808-27, ISSN 0022-3808.

Evans, D. & Leighton, L. (1989) Some Empirical Aspects of Enterpreneurship. *American Economic Review*, Vol.79 (June 1989), pp. 519-35, ISSN 0002-8282.

Griliches, Z. (1977). Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica*, Vol.45 (January 1977), pp. 1-22, ISSN 0012-9682.

Hamilton, B. (2000). Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment. *Journal of Political Economy*, Vol. 108, (June 2000), pp. 604-631, ISSN 0022-3808.

Harmon, C. & Walker, I. (1955). Estimates of the Economic Return to Schooling for the United Kingdom. *American Economic Review*, Vol.85 (December 1995), pp. 1279-1286, , ISSN 0002-8282.

Harmon, C; Oosterbeck, H. & Walker, I. (2003). The Return to Education: Micreoconomics. *Journal of Economic Surveys*, Vol.17 (April 2003): 115-155, ISSN 0950-0804.

Hashimoto, M. (1981). Firm-Specific Human Capital as a Shared Investment. *American Economic Review*, Vol.71 (June 1981), pp. 475-482, ISSN 0002-8282.

Hausman, J. (1978). Specification Test in Econometrics. *Econometrica*, Vol.46 (November 1978), pp. 1251-1271, ISSN 0012-9682.

Hausman, J. & Taylor, W. (1981). Panel Data and Unobservable Individual Effects. *Econometrica*, Vol.49 (November 1981), pp. 1377-1398, ISSN 0012-9682.

Hungerford, T. & Solon, G. (1987). Sheepskin Effects in the Return to Education. *Review of Economics and Statistics*, Vol.69 (1987), pp. 176-178, ISSN 0034-6535.

Im, K.; Ahn, S.; Schmidt, P. & Wooldridge, J. (1999). Efficient Estimation of Panel Data Models with Strictly Exogenous Explanatory Variables. *Journal of Econometrics*, Vol.93 (November 1999), pp. 177-201, ISSN 0304-4076.

Jaeger, D. & Page, M. (1996). Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education. *Review of Economics and Statistics*, Vol.78 (November 1996), pp. 733-40, ISSN 0034-6535.

Jovanovic, B. (1979). Job Matching and the Theory of Turnover. *Journal of Political Economy*, Vol.87 (October 1979), pp- 972-90, ISSN 0022-3808.

Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica*, Vol.50 (May 1982), pp. 649-670, ISSN 0012-9682.

Kawaguchi, D. (2003). Human Capital Accumulation of Salaried and Self-Employed Workers. *Labour Economics*, Vol.10 (February 2003), pp. 55-71, ISSN 0927-5371.

Lazear, E. (1981). Agency, Earnings Profiles, Productivity, and Hours Restrictions. *American Economic Review*, Vol.71 (September 1981), pp. 606-620, ISSN 0002-8282.

Lazear, E. & Moore, R. (1984). Incentives, Productivity, and Labor Contracts. *Quarterly Journal of Economics*, Vol.99 (May 1984), pp. 275-296, ISSN 0033-5533.

Miller, P.; Mulvey, C. & Martin, N. (1995). What do Twin Studies Reveal About the Economic Returns to Education? A Comparison of Australian and U.S. Findings. *American Economic Review*, Vol.85 (June 1995), pp. 586-599, ISSN 0002-8282.

Mincer, J. (1974). *Schooling, Experience and Earnings.* Columbia University Press, ISBN 0-870-14265-8, New York, USA.

Polachek, S. & Kim, M. (1994). Panel Estimates of the Gender Earnings Gap: Individual-Specific Intercept and Individual-Specific Slope Models. *Journal of Econometrics*, Vol.61 (March 1994), pp. 23-42 ISSN 0304-4076.

Psacharopoulos, G. (1973). *Returns to Education: An International Comparison.* Elsevier, Jossey-Bass, San Francisco, USA.

Psacharopoulos, G. (1981). Returns to education: an updated international comparison. *Comparative Education*, Vol.17 (October 1981), pp. 321-341 ISSN 0305-0068.

Psacharopoulos, G. (1985). Returns to Education: A Further International Update and Comparison. *Journal of Human Resources*, Vol.20 (Fall 1985), pp. 583-597 ISSN 0022-166X.

Psacharopoulos, G. (1994). Returns to Investment in Education: A Global Update. *World Development*, Vol.22 (September 1994), pp. 1325-1343 ISSN 0305-750X.

Rees, H. & Shah, A. (1986). An Empirical Analysis of Self-Employment in the UK. *Journal of Applied Econometrics*, Vol.1 (January 1986),pp. 95-108 ISSN 1099-1255.

Riley, J. (1979). Testing the Educational Screening Hypothesis. *Journal of Political Economy*, Vol.87, (October 1979), pp. S227-S251, ISSN 0022-3808.

Rosen, S. (1981). The Economics of Superstars. *American Economic Review*, Vol.71, (December 1981), pp. 845-858, ISSN 0002-8282.

Rosenzweig, M. & Wolpin, K. (2000). Natural "Natural Experiments" in Economics. *Journal of Economics Literature*, Vol.38, (2000), pp. 827-874, ISSN 0022-0515.

Salop, J. & Salop, S. (1976). Self-Selection and Turnover in the Labor Market. *Quarterly Journal of Economics*, Vol.90, (November 1976), pp. 619-627, ISSN 0033-5533.

Semykina, A. & Wooldridge, J. (2010). Estimating Panel Data Models in the Presence of Endogeneity and Selection. *Journal of Econometrics*, Vol.157, (June 2010), pp. 375-380, ISSN 0304-4076.

Spence, M. (1973). Job Market Signaling. *Quarterly Journal of Economics*, Vol.87, (August 1973), 355-374, ISSN 0033-5533.

Staiger, D. & Stock, J. (1997). Instrumental Variables Regressions With Weak Instruments. *Econometrica*, Vol.65 (May 1997), pp. 557-586, ISSN 0012-9682.

Stiglitz, J. (1975). The Theory of Screening, Education and the Distribution of Income. *American Economic Review*, Vol.65, (June 1975), pp. 283-300, ISSN 0002-8282.

Trostel, P.; Walker, I. & Woolley, P. Estimates of the Economic Return to Schooling for 28 Countries. *Labour Economics*, Vol.9, (February 2002), pp. 1-16, ISSN 0927-5371.

Weiss, A. (1995). Human Capital vs. Signalling Explanation of Wages. *Journal of Economic Perspectives*, Vol.9, (Fall 1995), pp. 133-154, ISSN 0895-3309.

Wolpin, K. (1977). Education and Screening. *American Economic Review*, Vol.67, (December 1977), pp. 949-958, ISSN 0002-8282.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, ISBN 0-262-23219-7, Cambridge, MA, USA.
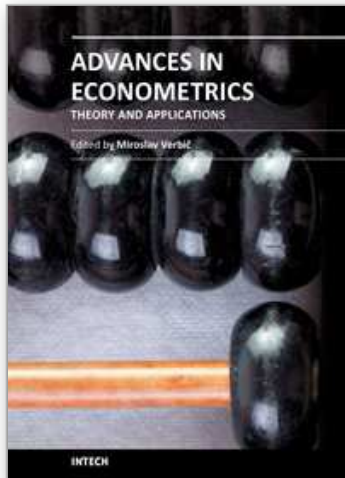
Advances in Econometrics - Theory and Applications

Edited by Dr. Miroslav Verbic

Econometrics is becoming a highly developed and highly mathematicized array of its own sub disciplines, as it should be, as economies are becoming increasingly complex, and scientific economic analyses require progressively thorough knowledge of solid quantitative methods. This book thus provides recent insight on some key issues in econometric theory and applications. The volume first focuses on three recent advances in econometric theory: non-parametric estimation, instrument generating functions, and seasonal volatility models. Additionally, three recent econometric applications are presented: continuous time duration analysis, panel data analysis dealing with endogeneity and selectivity biases, and seemingly unrelated regression analysis. Intended as an electronic edition, providing immediate â€œopen accessâ€ to its content, the book is easy to follow and will be of interest to professionals involved in econometrics.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds