AFC ASSOCIATION FRANÇAISE DE CLIOMÉTRIE

WORKING PAPERS

Nr. 4, 2008

The Impact of Surplus Schooling on Earnings: Some Additional Findings II

Jean-Pascal GUIRONNET

The Impact of Surplus Schooling on Earnings: Some Additional Findings II

Jean-Pascal Guironnet∗
∗ *LAMETA, University of Montpellier I, Department of Economics, Espace Richter, Avenue de la Mer, France*
E-mail: jp.guironnet@lameta.univ-montp1.fr.

Version: 25 January 2008

Abstract

This paper proposes a new statistical measure of the overeducation phenomenon. From Stochastic Production Frontiers (SPF), we use efficiency score between schooling and earning to assess worker mismatches. Based on the French case, our measure reveals an overestimation of the overeducation extent, in comparison to the traditional measures. Following our econometric model, SPF approaches provide better estimation results on probit and earning function. Moreover, our approach shows that traditional measures overestimate also the wage penalties of overeducation. Thus, we point out that previous studies overestimate the dimishing returns of education and they have, conversely, understated the role of human capital.

Key Words: Overeducation; Stochastic Production Frontiers; Measurements.
Subject Classification: [*JEL*] I2; J24; J3.

1. INTRODUCTION

At the beginning of the seventies, the overinvestment idea in education was firstly investigated by Freeman (1971 and 1976) and Dore (1976). Freeman found that education returns have significantly decreased in the United States. He assigned these diminishing returns to an excess of graduate offer. The results called into question the belief that college degrees mirror a profitable investment and a virtual guarantee of economic success. Freeman's work has catalysed a large literature on overeducation. Since this study, a large number of empirical researches investigate the impact of surplus schooling on earnings. These papers showed significant

overeducation rates in several developed countries (Sloane, 2003). Today, the overeducation concept can be defined as follow: a worker is considered as overeducated if his educational level exceeds that required for a particular job.[1]

Duncan and Hoffman (1981) were the first economists to carry out[2] an empirical study which estimated the education return with a correction for education surplus. Height years later, Verdugo and Verdugo (1989) published a new measure of the surplus schooling. The Verdugo and Verdugo (VV) method instigated a large number of critical and comments (see next section for more details and the dialogue initiated by Cohn, 1992; Gill and Solberg, 1992). Numerous studies fave explored the topic of overeducation measurement (see next section) and discussed about the various methods to appraise the worker mismatches, like Chevalier (2003) published in *Economica*.

Our purpose is to extent these seminal works and to suggest a new measure of overeducation, following a previous work (Guironnet and Peypoch, 2007) which used "*Data Envelopment Analysis*" (DEA) to assess education surplus. However, this approach is not without limits: our analysis was restricted to small samples. To avoid this problem, we use the SPF approach. This econometric model allows to generalize our results and to obtain better political recommendations in the "benchmarking" of human resources.

In this paper, we apply our overeducation measure on French case. Thus, we use two surveys, conducted by the Centre d'Etudes et de Recherches sur les Qualifications (Céreq), in 1987 and 1999. The first section explains the various available methods to evaluate overeducation extent. The second section presents our measure and a comparison of overeducation extent with traditional measures. Finally, our last section studies the impact of surplus schooling on earnings.

2. OVEREDUCATION MEASURES

Determining what constitute a "normal" relation between initial training and employment remains a difficult task. Moreover, technical progress and "diploma

[1] Conversely, an individual can be considered as undereducated if his educational level falls below that required by the job.

[2] Published in the first issue of *The Economics of Education Review*.

inflation" increase the assessment difficulties. However, the relation between educational mismatch and earnings seems not to be sensitive to the applied measurement method (Hartog, 2000).

There are two main way to evaluate the overeducation extent: the "subjective" approach and the "objective" approach (Rumberger, 1987). This latter can be decomposed into the normative and statistical methods.

2.1. Subjective Approach

The subjective approach is based on how individuals feel about their work, using what are called "self-assessment" techniques. The various surveys use very different questions to evaluate this feeling. More precisely, this measurement method follows two distinct purposes and it can be subdivided in two subcategories: on one hand, we can ask which educational level is required to get the job (e.g. Duncan and Hoffman, 1981); on the other hand, we can ask which is the required education to do the job (e.g. Hersch, 1991).

However, the weakness of this approach remains its dependence on the objectivity of the respondent and the query statements. For example, two types of bias can work against each other. First, overeducated workers may have a negative opinion of their job and overestimate their overeducation extent. Secondly, "the diploma inflation" leads workers to overestimate the required education. Thus, they internalize the feeling of overeducation (Giret and Lemistre, 2004), without there being any real change in the tasks to be conducted by the worker.

In this context, we choose not to look further the subjective approach (more fitted to the sociological analysis) and we choose to privilege, in this paper, the objective approach. Thus, we present in the following section the statistic and normative methods, respectively.

2.2. Objective Approach

This approach involves comparing an analysis of the skills that are *a priori* required for a given job and the occupation-type for which initial training prepares the holder. This fact implies to elaborate "matches grid" between degrees and

occupations. Thus, overeducation measurement is based on a comparison between the individual's level of education and that which is "normally" required for the job he/she does. Two types of "matches grids" can be distinguished: one normative and the other statistical.

The first is constructed from an analysis expert of the labour market, who establishes what the requirements should be for a range of jobs. This method, much used in the United States, is essentially based on the "*Dictionary of Occupational Titles*". For the French case, this type of matches grid has been developed by Affichard (1981). The normative approach seems to provide more reliable results and it is "*a very attractive source for defining job requirements, because of its explicit goal of objectivity, clear definitions and detailed measurement instructions*" (Hartog, 2000). However, this method assumes that all workers doing a given job carry out tasks of the same complexity. In this case, the heterogeneous nature of the roles which are required within an aggregated occupation can bias the overeducation measure (Halaby, 1994). Moreover, this latter demands a frequent redefinition of the required skills of the occupations with the changes of productive system. Consequently, a lot of researchers use the statistical method as a substitute of the normative measure.

The second type of objective approach defines "normal" matches on the basis of what appear from the statistical analysis as being the most frequent. In the United States, the most common statistical method was developed by VV. This consists of calculating the average level of education on the basis of an occupation range. Then, workers are classified as overeducated if his/her educational level exceeds a standard deviation of the average level for the occupation. However, it seems somewhat arbitrary to fix a threshold for overeducation on the basis of a standard deviation. If the proportion of overeducated workers is high (or low) then the average levels of education will be high (or low). In this case, overeducation will be underestimated (or overestimated).[3] Furthermore, the statistical method is not immunized about the heterogeneous tasks inside occupations.

[3] Other criterion can be used in this statistical approach like the modal values instead of a standard deviation (Kiker, *et al.* 1996).

3. STOCHASTIC PRODUCTION FRONTIER *VS.* NORMATIVE APPROACH

Using the Céreq data on higher education leavers, our comparison of overeducation measures is based on two surveys carried out in 1987 and 1999.

3.1. Data and SPF approach

The database of 1987 represents the oldest data available and the one of 1999 is the most recent built with the same methodology. These surveys are constructed using an identical procedure: investigating a graduate cohort three years after students left (cohort graduations in 1984 and 1996 respectively). These longitudinal data provides extensive information on graduate careers. Our study is based on leavers of bachelor and master degrees, technological universities, engineer graduates, and business schools accredited by the government. Medical, paramedical, social or artistic training, graduates of foreign nationalities and postgraduate students are not interviewed. Our analysis aims to measure the drop in graduate incomes: for this purpose, we excluded unemployed graduates at the survey's year, part-time workers, people of more 35 year and those who pursued their studies.

We overall retained, respectively for 1987 and 1999, a sample of 8 410 observations which represents a weighted population of 36 309 individuals and a sample of 3 348 observations which represents a weighted population of 137 843 individuals. This change of the leaver number of educational system reflects the real evolution of the graduate offer in France: over twelve years, the possibilities to follow higher studies have largely improved. For example, graduates of higher education represented 36% of the leavers of educational system in 2000 whereas twenty years ago the proportion was around 15%.

In 1984, the French economy was characterized by a weak growth recovery, after oil crisis. Since 1987, GNP rates had risen and fallen in a recession due to the Gulf war. Unemployment rate rose from 1982 until 1990. In 1999, business activities was also prosperous but in contrast with 1987, the integration difficulties of young graduates into the job market increased: their unemployment rate, three years after they left college, are estimated at 9.71% in 1999 whereas twelve years beforehand it was only about 5.06% (Guironnet, 2006). This high number of unemployed

graduates is especially significant since it makes pressure on starting wages.

Moreover, the level of qualifications is sensitive to the educational offer: same degree obtained by a young worker is not strictly equivalent in terms of skills to that obtained by older workers. In order to correct the bias of generalized qualification depreciation, subsequent from an education supply side effect, we used a method that was investigated by Jarousse and Mingat (1986). The individual position P_i in diploma hierarchy, inside its cohort, corresponds to a centred and reduced value of the number of educational years (NEY_{it}) for each cohort t. Then, the standardized number of educational years ($SNEY$) corresponds to the degree position inside the cohort adjusted by the distribution parameters of NEY for the cohort of reference ($t = 0$):

$$SNEY_{it} = P_{it}\sigma_0 + \overline{NEY}_0 \tag{1}$$

where $P_{it} = (NEY_{it} - \overline{NEY}_t)/\sigma_t$, $\overline{NEY}_t$ is the mean number of educational years of the French leavers of higher education for a time period t and σ_t its standard deviation. Thus, we obtain a proxy of relative qualification levels that removes cohort effects.

From table 1, we observe an increase of the NEY average for all occupations except executive workers. For this last, it seems that the needed skills decreased. Consequently, we can expect that the demand effect should play in favour of a decrease of overeducation rate in this profession. If we choose 1987 as a base year[4] to calculate $SNEY$, we can evaluate the offer effect.[5]

Thus, over twelve years, we perceive significant cohort effects: the large increase of the graduate offer has a negative impact on all qualification levels. In each profession, the offer effect reduces the information capacity of the degrees to identify the productive potential of the holders, particularly for the least qualified occupations. Obviously, the weakest degrees, and indirectly the least qualified occupations, are the most affected by cohort effects. Thus, we observe a positive correlation between offer effect and professional hierarchy. However, the intermediate professions in business or administrative career (IPTA) display a different

[4] Thus, $SNEY$ replaces NEY for the cohort of 1999.

[5] Offer effect can be obtained by the growth rate: $(SNEY - NEY)/NEY$ (Lemistre, 2003).

TABLE 1: Higher Educational Years by Occupations

Occupations/ *Variables*	NEY_{1987}	NEY_{1999}	$SNEY_{1999}$	*Offer Effect*
Executive	4.29	4.16	3.89	-6.51%
Engineer	4.63	4.82	4.58	-5.10%
Manager	2.25	3.38	3.08	-8,87%
IPTA	2.73	3.27	2.96	-9.29%
Technician	2.18	3.39	3.09	-8.82%
Employee	2.29	2.74	2.42	-11.75%
Worker	2.09	2.20	1.85	-15.61%

trend: these professions set a skill structure badly specified (Guironnet, 2006). In this sector, employers recruit workers holding some educational levels, or training specialities, with strong disparities. Consequently, the workers in commercial careers hold a large range of qualifications which are probably the most devalued.

As we discussed in the introduction, our objective is to measure the education surplus with SPF approach. Thus, we measure overeducation phenomenon in terms of the difference between potential income, determined on the basis of the production frontier boundary, and real income (Jensen, 2003). Our brief survey of traditional overeducation measures (*cf.* section 2) situates the benefits of a measure that is based on the matches between skills and earnings. For example, how can a graduate who works as a technician and earns an executive wage be considered as overeducated? The main benefit of SPF measure is that is not biased by the homogeneous tasks assumption within an occupation. Furthermore, a salary-based approach avoids the problem of a socio-occupational groups ranking while wage is clearly a continuous ordinal variable (Nauze-Fichet and Tomasini, 2002).

Following our approach, SPF measures the maximum potential ouput for any particular input vector and given technology.[6] Consequently, this method is in conformity with human capital theory that assumes that income maximization is a major determinant of education investment. For example, it determines for every value of experience input (EXP) the minimum amount of NEY input to obtain

[6] So we estimate one SPF for each survey to allow a production technology to change over the twelve years.

the desired earning level. In measuring education efficiency, this approach can appraise the education fully utilized and that may be underutilized (i.e. the surplus schooling). Technical Efficiency (TE) is measured as the ratio[7] of realised output to the maximum potential output of the worker with unchanged input (output level on the production frontier).

There are two conceivable approaches to estimate TE: DEA which uses non parametric approach and SPF which uses parametric function.[8] The production frontier in DEA is deterministic, so any deviations from it are related to inefficiency. In contrast, SPF approach is sensitive to random shocks by including a random error term to the production frontier. Consequently, only deviations caused by controllable decisions can be assigned to inefficiency. Obviously, this latter approach is more adapted to modelize worker choices in the labour market. In this paper, the used model is based on the Battese and Coelli (1995) specification:

$$\ln(y_i) = \beta X_i + v_i - u_i \tag{2}$$

where y_i is the wages of the *i-th* worker, X_i is a vector of production input like NEY, EXP, TEN (tenure), β is a column vector of coefficient to estimate and v_i refers to independent identically distributed random variables $N(0, \sigma_v^2)$. In order to identify some reasons for differences in predicted efficiency between workers, u_i are assumed to account for TE in production. In effect, like our measure is affected by some factors which produce wage heterogeneities of the population, we estimate a variable vector which may influence efficiency and indirectly the worker earning. This random error is assumed to be normally truncated such as $u_i = |U_i|$ where $U_i \sim N(\delta Z_i, \sigma_u^2)$. Thus, inefficiency effect, u_i, is expressed as an explicit function of a vector of worker-specific variables and a random error, w_i. The inefficiency determinants function can be written as:

$$u_i = \delta_0 + \delta Z_i + w_i \tag{3}$$

where Z_i is a vector of characteristics affecting the efficiency level, δ is a vector of

[7] This ratio is inferior to the unity.

[8] For a survey on SPF, the reader can consult Coelli *et al.* (1998) and Kumbhakar, Lovell (2000).

parameters to be estimated. Following Battese and Corra (1977), variance terms are parameterized by replacing σ_v^2 and σ_u^2 with $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\gamma = \sigma_u^2/\sigma^2$. The maximum output is achieved when $U_i = 0$. TE of the *i-th* worker is defined as:

$$\widehat{TE}_i = E(Y_i \mid U_i, X_i)/E(Y_i \mid U_i = 0, X_i) = \exp(-u_i) \tag{4}$$

$\widehat{TE}$ measures the part of efficiency differences (across workers) that cannot be explained by Z variables. Consequently, our measure is not affected by individual unobserved heterogeneity. Equations (2) and (3) can be estimated by maximum likehood.[9] For this purpose, we used *Frontier 4.1* software to obtained final estimates of maximum likehood (MLE). A three-stepwise procedure[10] in estimating the MLE estimates of the SPF parameters will then be used.

3.2. Rate of overeducation

Following Mincerian Model (1974), wages (in logs) are determined by five factors: NEY (or $SNEY$ for 1999) and two quadratic forms of TEN and EXP variables (coded in months). A lot of specifications were tested and we have retained this significant model for two theoretical reasons : first, these human capital components can be considered as substitutes (Alba-Ramirez, 1993), thus our comments are only restricted in terms of surplus of human capital; second, these characteristics are modifiable, thus we can made some feasible recommendations for the "benchmarking" of human resources. However, other individual heterogeneities influence the earnings. In this way, we introduce inalterable characteristics in inefficiency model: sex, to have an executive father and regional rate of unemployment (RRU). As suggest overeducation literature, inefficiency model include also the individual choices on French labour market; i.e., firm size (base-case small), public sector, worker's unemployment months over the three survey years ($UNEMPLOY$) and worker's socio-professional categories (base-case employee). Lastly, we introduce the other educational choices that influence indirectly our main human capital determinants: vocational degrees and discipline of initial training (base-case is social science and law).

[9] The log-likehood function of this model is presented in Battese and Coelli (1993).

[10] Using OLS starting values, this software combines a grid search and a Quasi-Newton method.

TABLE 2: Stochastic Production Frontier from 1987 and 1999

Variable/Occupation	SPF 1987	Std. error	SPF 1999	Std. error
Constant	8.93	0.05	10.29	9.95
NEY	0.07	0.003	0.06	0.004
TEN	0.008	0.001	0.005	0.001
TEN^2	-0.00001	0.00003	$3e^{-8}$	0.00004
EXP	0.004	0.001	0.001	0.002
EXP^2	0.0001	0.00004	0.0001	0.00007
Inefficiency model				
Constant	0.46	0.05	1.61	9.96
Vocational degree	-0.02	0.01	-0.02	0.01
Sex (male=1)	-0.11	0.01	-0.08	0.01
Medium firm	-0.04	0.01	-0.06	0.01
Large Firm	-0.08	0.01	-0.11	0.01
Public sector	0.08	0.01	0.09	0.01
Natural	0.08	0.01	0.06	0.02
Humanities	0.07	0.01	0.05	0.01
Exact science	0.01	0.01	0.02	0.01
Executive Father	-0.07	0.01	-0.04	0.01
UNEMPLOY	0.03	0.01	0.03	0.01
RRU	0.02	0.002	0.01	0.002
Executive	-0.32	0.01	-0.34	0.01
Engineer	-0.40	0.02	-0.38	0.02
Manager	-0.16	0.02	-0.13	0.02
IPTA	-0.16	0.01	-0.11	0.02
Technician	-0.14	0.01	-0.11	0.01
Worker	-0.01	0.02	0.05	0.03
σ^2	0.06	0.0009	0.04	0.001
γ	0.04	0.04	0.53	1.20
Log-Likehood	253		471	
LR test, $\chi^2_{mix}(19)$	2976		1596	

Following equation 2 estimates, we find that NEY explains the greater part of the wage variance. Naturally, schooling is the main determinant to start one's career. In contrary, we have not enough longer professional careers (can exceed three years in our database) to reach diminishing returns on the quadratic terms of TEN and EXP variables. The estimated coefficients suggest that tenure and experience improve efficiency and reduce, consequently, overeducation: this phenomenon is temporary. It can be registered as a logic of starting professional career of the workers (Guironnet, 2006).[11]

More generally, γ is not significant suggesting an absence of inefficiency effects in our model. However, we find a strong worker heterogeneity: all estimated coefficients of inefficiency model are highly significant. Moreover, we use the generalised likelihood ratio test statistic, i.e. $LR = -2[\log L(H_0) - \log L(H_1)]$ where $L(H_0)$ denotes the value of likelihood function under the null hypothesis (H_0) whereas $L(H_1)$ denotes alternative hypotheses. Under the null hypothesis $H_0 : \gamma = \delta_0 = \delta_1 = ... = \delta_{18} = 0$ (inefficient effects are absent from the model), LR is distributed according to a "mixed" chi-square distribution with a degree of freedom equal to the difference between the parameters involved in the null and alternative hypothesis[12]. In the present case[13], LR tests suggest that we reject the H_0 hypothesis. Consequently, TE appears in the earning function and our model is well specified.

From inefficiency model, we find the expected estimated coefficients. In particular, we can see a positive correlation with inefficiencies for the following variables: academic courses; women; public sector; small firms; humanities; if we have not an executive father and RRU (see next section for a more precisely comment of this results).

As suggested by Thurow's approach, job characteristics were major determinants of the worker productivity. The highest occupations in professional hierarchy allow to use more fully the worker skills. Consequently, well-qualified professions

[11] Groeneveld and Hartog (2004) found a positive effect of overeducation on job promotion.

[12] In this case, critical values for the generalised likelihood-ratio test are obtained from Table 1 of Kodde and Palm (1986).

[13] Due to the length of this paper, we don't present the results for the constrained model. However, these estimates are available upon request.

largely decrease inefficient component. Contrary to the least qualified jobs, workers in the well-qualified job make their educational levels more profitable. Another explanation concerns the technology evolution: non-neutral technologies benefit to the more educated individuals (Daly *et al.* 2000). We find a similar result than previous studies: with downward wage rigidities in France (Guironnet and Peypoch, 2007), the estimated coefficients of the lowest qualified professions of inefficiencies model indicate that the education requirement was largely overestimated for these professions (Dolton and Silles, 2003). This last interpretation is consistent with a "skill bumping" effect: under the pressure of unemployment rate in 1999, firms preferred to employ overeducated workers rather than to adjust wages for the lower-level occupations (Sloane, 2003).

In order to appraise mismatches, we measure the overeducation extent from efficiency score. In base 100, efficiency mean is around 74% in 1987 whereas twelve years later this value falls of thirteen percentage points (i.e. 61%). This large efficiency decrease suggests an increase of overeducation phenomenon over this period. The main explanation of this result is that the new graduate cohorts are always better trained but the downward wage rigidities in France yield dismishing education returns. From efficiency score, we choose[14] to consider a worker as overeducated if this efficiency score is below the efficiency mean (EM) of the samples minus standard deviation (σ_e), i.e. $1 - \exp(-u_i) + \sigma_e$.[15] Thus, an individual is adequately matched if his efficiency score is between $EM \pm \sigma_e$.

From this overeducation definition, we find that 30% of the weighted number of the graduate leavers of higher education was overeducated whereas twelve years ago this rate was around 18%. In comparison to the normative and statistical measure[16] (quite similar to the VV method), overeducation extent is assessed to

[14] We are conscious that this definition is arbitrary and that this score can be used to evaluate overeducation extent. However, our definition allows to take account the overeducation endogeneity (see section 4.1).

[15] Conversely, we can define an individual as undereducated if his efficiency score is higher of $EM + \sigma_e$.

[16] For French case, the statistical and normative measures leads to the same "matches grid" (consultable in appendix B) for the graduates of higher education. Consequently, our wage downgrading approach is compared to these two objective measures.

35% in 1999 and 27% twelve years ago. Why such differences?

Two main biases can explain our results. Firstly, our measure is not affected by the different technology used because each frontier is estimated separately: we allow a given technology for each year. Secondly, as we have evoked in section 2.2, the other objective measures can consider a worker as overeducated whereas that he/she earns a significant wage. In contrary, our approach appraises overeducation from the "inputs-outputs" ratio, i.e. matches between earning and schooling. These two biases tend to overestimate the overeducation extent measured by traditional approaches and they understate the role of human capital.

The next step estimates the overeducation effect on wages. For this purpose, we always compare the various overeducation measurements.

4. THE IMPACT OF SURPLUS SCHOOLING ON EARNING

For the French case, numerous studies have showed that degrees[17] are most significant that NEY. Consequently, in your earning functions, we use a binary code for degree and overeducation variables.

4.1. Econometric model

Our earning function is specified as follow:

$$\ln(y_i) = \beta_0 + \beta_{1k} X_{ki} + \beta_2 Over_i + \varepsilon_i \tag{5}$$

where X is a row vector of exogenous characteristics (degrees, sex, etc.) and β_{1k} a column vector of estimated coefficients. ε_i represents the random error, $Over_i$ is the observed result of overeducation and β_2 its return. As we used dummy variable for overeducation, the expected coefficient β_2 should be negative: overeducated worker is compared to worker adequately matched with the same educational level.

However, overeducated workers are unlikely aleatory selectioned in the population (Dolton and Silles, 2003). Unobservable characteristics are partially included in the residuals and they are correlated to the overeducation phenomenon. Thus, a OLS estimation gives $\widehat{\beta}_2 = \beta_2 + \frac{Cov(Over_i, \varepsilon_i)}{Var(Over_i)}$. In other words, some characteristics

[17] The reader can consult appendix A for some explanations on French educational system.

explaining the overeducation are linked to those affecting earnings. The existence of a strong unobservable individual heterogeneity (see for example, section 3.2) implies to esteem an econometric method for a treatment effect[18]. In our case, we choose an endogenous model based on Barnow *et al.* (1981) specification, estimated in two steps.

Firstly, we determined the individual probabilities to be overeducated by a Probit model:

$$Over_i = \begin{cases} 1 \text{ si } Over_i^* > 0 \\ 0 \text{ si } Over_i^* \leq 0 \end{cases} \text{ with } Over_i^* = \gamma w + u_i \tag{6}$$

where $Over_i^*$ is a latent variable explaining overeducation and w is a set of exogenous individual characteristics influencing overeducation. Selection bias implies that ε_i and u_i are correlated. We assume that these two random errors follow a bivariate distribution of zero mean; variance and covariance are σ_ε^2, σ_u^2, $\sigma_{\varepsilon u}$. Following this method, estimated coefficients $\widehat{\gamma}$ are used to set up estimated conditional expectation of u_i.

Secondly, equation (5) is estimated with a corrector term of selectivity effect:

$$\ln(y_i) = \beta_0 + \beta_{1k} X_{ki} + \alpha Over_i + \delta \widehat{\lambda} + \mu_i \tag{7}$$

β, α and δ coefficients are estimated with OLS procedure. $\widehat{\lambda}$ yields estimated conditional expectation of u_i. Following (6), $S = 1$ is equivalent to $-\gamma w < u_i$. We assume $\sigma_u^2 = 1$ and it follows a normal (centred and reduced) distribution, independent of X. Thus, we obtain:

$$\begin{cases} E[u_i \mid X, Over_i^* > 0] = E[u_i \mid X, u > -\gamma w] = \frac{\varphi(\gamma w)}{\phi(\gamma w)} \\ E[u_i \mid X, Over_i^* \leq 0] = E[u_i \mid X, u \leq -\gamma w] = \frac{-\varphi(\gamma w)}{1-\phi(\gamma w)} \end{cases} \tag{8}$$

where φ and ϕ are respectively the density function and cumulative distribution function of the standard distribution.

[18] For an examination of the various biases of the link between training and earning, the reader can consult Card (2001).

If we note $\varphi(\gamma w)$ and $\phi(\gamma w)$ respectively φ and ϕ; we obtain:

$$\begin{aligned} E[u_i \mid X, Over_i] &= Over_i \times E[u_i \mid X, Over_i = 1] + (1 - Over_i) \times E[u_i \mid X, Over_i = 0] \\ &= Over_i \times \tfrac{\varphi}{\phi} - (1 - Over_i) \times \tfrac{\varphi}{1-\phi} = \tfrac{\varphi(1-\phi)}{\phi(1-\phi)} = \lambda \end{aligned} \tag{9}$$

Finally, we estimate:

$$E[Y \mid X, Over_i] = \beta_0 + \beta_{1k} X_{ki} + \alpha Over_i + \delta\lambda + \mu_i \text{ with } \delta = corr(\varepsilon_i, u_i) \tag{10}$$

For identification model, we exclude from X the father profession variables. This latter is only used in Probit specification. Thus, we consider that the father can find a starting job for his children but he can not influence the children earning.[19]

4.2. Results

In appendixes, table 5 and 6 show the estimated coefficients from Probit specifications. Globally, these results are in conformity with our previous results from SPF estimates (*cf.* section 3.2).

From the estimated degree coefficients, Business School confers the higher benefits with a lesser probabilities to be overeducated in 1987. However, this advantage is largely reduced twelve years later: today, overeducation influences each educational levels. Firm sizes also influence mismatches: overeducation is negatively correlated with firm sizes. Workers who have an executive father hold a significant advantage in their professional career. This latter can probably beneficiate from father's relational network. Another explanation is that this individual has better academic results with a better home environment.

As suggested by theory of differential overqualification (Frank, 1978), women have higher probabilities to be overeducated (Sloane *et al.* 1999; Büchel and Battu, 2003). Lastly, *RRU* can explain the sensivity of overeducation phenomenon with the economic situation. In effect, overeducation may be influenced by local context. Following the competitions model (Thurow, 1975), we can consider *RRU* as a proxy

[19] We have tried to include this variable in the estimation of earning function. However, the correspondent coefficient is insignificant, testifying of the validity of our assumption.

of "*labour queue*" in the local market. In this context, overeducation phenomenon is positively correlated to *RRU* (Groot and Maassen van den Brink, 2000).

Globally, SPF measure provides better results for each year than the other traditional measures: the concordant predicted probabilities in 1987 (resp. 1999) are evaluated to 82.1% (resp. 72.6%) with traditional measures whereas with SPF measures, this percent is 94.6% (resp. 82.5%). In particular, we find with our measure the expected estimated coefficients for public sector and humanity variables: overeducation measured by SPF produce more reliable results.

Concerning public sector, one explanation is that civil servants prefer job safety than higher wages. Thus, they have a higher risk to be mismatched. About the training faculty, a student in humanity courses has lower probabilities to be matched with the required education of their job. On one hand, student in humanities substitutes higher amounts of "non-relevant" education for a lack of job relevant education. Consequently, they have higher probability to be overeducated (McGuinness, 2003). On the other hand, biased technical progress can explain the advantage for student in exact science over the eighties. In particular, computer innovation (Haskel and Heden, 1999; Falk and Koebel, 2004) increases probably the demand of skill in this training.

From tables 3 and 4, we can see results from earning function. Globally, *adjusted R-squared* shows not significant differences in the overeducation measurement in 1987. However, this criterion is largely higher for our measure in 1999, suggesting better result of SPF measure. One explanation is that there are probably more task heterogeneities in 1999, explaining the fall of adjusted *R-squared* for traditional measures: a SPF approach seems to be better for the most recent years.

From a general viewpoint, omitted unobserved heterogeneities seems to underestimate ($\delta > 0$) overeducation impact with the others objective measures.[20] In contrary, with SPF measures, this bias seems to overestimate the overeducation impact on wages ($\delta < 0$). This latter result is mostly in accordance with overeducation literature issues (e.g. Chevalier, 2003 and McGuinness, 2006), supporting the fact that the wage downgrading is a better measure.

[20] In the same way of the footnote #13, the OLS estimations are available upon request.

TABLE 3: Earning Function in 1987

Variable	Earning Function	Std. error	Earning Function with SPF Measure	Std. error
Constant	8.48	0.25	8.38	9.95
Sex	0.12	0.01	0.14	0.01
TEN	0.01	0.0004	0.01	0.0004
EXP	0.01	0.0004	0.01	0.004
DEA	0.37	0.18	0.40	0.02
DESS	0.38	0.15	0.40	0.01
Business School	0.40	0.01	0.44	0.01
Master	0.25	0.01	0.27	0.01
Licence	0.17	0.01	0.17	0.01
Deug	0.11	0.01	0.12	0.01
Dut	0.01	0.01	0.01	0.01
Natural	-0.05	0.01	-0.03	0.01
Humanities	-0.03	0.01	-0.05	0.01
Exact science	0.03	0.01	0.06	0.01
Medium Firm	0.03	0.01	0.03	0.01
Large Firm	0.06	0.01	0.07	0.01
Public sector	-0.07	0.01	-0.07	0.002
Paris	0.13	0.01	0.13	0.01
Over	-0.19	0.03	-0.04	0.02
λ	0.007	0.02	-0.09	0.01
δ	0.03		-0.36	
Adjusted R-squared	0.54		0.53	

TABLE 4: Earning Function in 1999

Variable	Earning Function	Std. error	Earning Function with SPF Measure	Std. error
Constant	8.94	0.06	8.66	0.04
Sexe	0.06	0.01	0.09	0.01
TEN	0.006	0.0007	0.007	0.0005
EXP	0.004	0.0007	0.006	0.0006
DEA	0.28	0.03	0.34	0.03
DESS	0.26	0.02	0.34	0.02
Business School	0.31	0.02	0.38	0.02
Master	0.21	0.02	0.24	0.02
Licence	0.17	0.02	0.18	0.02
Deug	0.08	0.03	0.10	0.03
Dut	-0.01	0.02	0.02	0.01
Natural	-0.06	0.02	-0.01	0.02
Humanities	-0.07	0.02	-0.01	0.01
Exact science	-0.03	0.02	0.05	0.01
Medium Firm	0.04	0.01	0.05	0.01
Large Firm	0.08	0.01	0.11	0.01
Public sector	-0.08	0.01	-0.04	0.01
Paris	0.12	0.01	0.10	0.01
Over	-0.39	0.04	-0.14	0.04
λ	0.10	0.02	-0.14	0.02
δ	0.45		-0.66	
Adjusted R-squared	0.55		0.63	

For each year, all estimated coefficients from X increased with SPF measure. In this case, overeducation variable allows a better observation of earning discrimination of the exogenous worker characteristics. Consequently, this dummy variable better captures the overeducation effects than the other objective measures. From each specification, we must add estimated coefficients from *Over* and λ variables to reach the real overeducation incidence on earning. Thus, overeducation reduces wages of around 18% (resp. 29%) whereas with SPF measure this percentage is 13% (resp. 28%) in 1987 (resp. 1999). Consequently, we find that traditional measures overestimate slightly the wage penalties and that our overeducation measurement is quite similar, in terms of wage penality, to the issues of other applied measures (Hartog, 2000). Like several studies on overeducation topic (e.g. Alba-Ramirez, 1993), we find that the job match between worker and job characteristics is highly relevant.

Despite lesser wage penalties from overeducation, all estimated degree coefficients in 1999 are lesser than in 1987 (excepted for the *Dut* degree). However, the differences between 1987 and 1999 are lesser than those with traditional measure. This result suggests that previous studies have overestimated the dismishing return of education.

Lastly, we briefly comment the other results. In this way, the estimated coefficient shows that civil servants have a wage penalty for each year. Following human capital theory, more competitive private sector allows more flexible earnings than public sector. As expected, women earn less than men. However, this penality is significantly lesser in 1999. In the same way, firm size is positively correlated to earnings. One explanation is that worker selectivity was probably better in large firms (Dupray, 2001). As expected, working in Paris confers an earning benefit.

5. CONCLUSION

Several studies on France and other developed country have shown significant overeducation extent and devoted wage penalties. However, this paper expresses some doubts about these results: overeducation rates are probably lower than the predicted values of the previous studies and wage penalties from mismatches are

probably also overestimated.

In 1987, overeducation extent is assessed to 18% whereas traditional measures assess this rate to 27%. Despite better economic situation, overeducation extent increases to 30% with SPF measure and 35% with the other objective measures in 1999. Thus, our measure suggests a lower overeducation rate but a higher rate of overeducation growth. However, we find a slightly overestimation of wage penalties from overeducation with traditional measures. These two main results always suggest (like other studies) that educational policy makers must have better respect on overeducation phenomenon. Like United States, French government policy has decided to pursue a 50% higher education participation. However, these paper results suggest that this policy should surely be highly questionable.

In particular, we find the evidence of a large "skill bumping" effect on French labour market. The main consequence of this phenomenon is the employer overestimation of the educational requirement for the least qualified occupation. From a "benchmarking" of human resources viewpoint, more vocational courses should decrease this effect: a better allocation of the least qualified workers would probably produce some benefits like better worker satisfaction (Tsang *et al.* 1991), higher productivity, higher national welfare, etc.

Of course, additional analysis is warranted to confirm our results, especially on the overeducation measurement topics. For example, SPF measure can be applied on other countries to confirm the better results as well on Probit model that on ORU function. Moreover, we could estimate this econometric model for more disaggregated occupation and for each sex, if database allows this analysis-type.

APPENDIX A: FRENCH EDUCATIONAL SYSTEM

DEA: academic degree in master second year.

DESS: vocational degree in master second year.

Busines School: private school in commercial career.

This training is equivalent to the master level.

Master: first year of master.

Licence: third year of licence.

DEUG: academic degree in licence second year.

DUT, BTS: vocational degree in licence second year.

APPENDIX B: MATCHES GRID FOR FRENCH CASE (1987,1999)

Occupation/*Degree*	*Business School* *Master 2*	*Master 1* *Licence*	*DEUG* *DUT, BTS*
Executive Engineer	Adequatly	Adequatly	Under
Intermediate Profession	Over	Adequatly	Adequatly
Technician	Over	Over	Adequatly
Skilled Employee	Over	Over	Over
Unskilled Employee	Over	Over	Over
Skilled Worker	Over	Over	Over
Unskilled Worker	Over	Over	Over

APPENDIX C: PROBIT ESTIMATION (1987, 1999)

Variable-1987	Overeducated	Std. error	Overeducated with SPF Measure	Std. error
Constant	0.73	0.13	-0.06	0.18
Sex	-0.06	0.04	-1.66	0.07
TEN	-0.02	0.002	-0.03	0.004
EXP	-0.02	0.003	-0.04	0.004
DEA	-0.63	0.12	-1.68	0.24
DESS	-0.52	0.09	-1.33	0.15
Business School	-1.38	0.07	-1.76	0.14
Master	-0.57	0.06	-1.11	0.08
Licence	-0.17	0.06	-0.83	0.08
Deug	-0.20	0.08	-0.08	0.11
Dut	-0.14	0.05	-0.20	0.07
Natural	-0.44	0.07	-0.40	0.09
Humanities	0.11	0.05	0.40	0.06
Exact science	-0.87	0.05	-1.08	0.08
Medium Firm	-0.07	0.04	-0.30	0.06
Large Firm	-0.17	0.04	-0.81	0.06
Public sector	-0.04	0.04	0.62	0.05
Indépendant	-0.06	0.05	-0.14	0.07
Executive	-0.30	0.05	-0.74	0.07
Intermediate	-0.16	0.06	-0.24	0.08
Employee	-0.16	0.06	-0.23	0.09
RRU	0.02	0.01	0.15	0.01
Log Likehood	2028		3369	
Predicted Probabilities *Percent Concordant*	82.1		94.6	

Variable-1999	Overeducated	Std. error	Overeducated with SPF Measure	Std. error
Constant	1.13	0.21	0.78	0.24
Sex	-0.44	0.05	-0.76	0.06
TEN	-0.02	0.003	-0.02	0.003
EXP	-0.02	0.004	-0.02	0.004
DEA	-0.52	0.14	-1.14	0.18
DESS	-0.82	0.10	-1.24	0.12
Business School	-0.72	0.10	-1.23	0.12
Master	-0.39	0.08	-0.91	0.09
Licence	-0.22	0.11	-0.75	0.12
Deug	-0.19	0.16	-0.32	0.17
Dut	-0.23	0.09	-0.17	0.10
Natural	-0.37	0.11	-0.05	0.12
Humanities	0.40	0.08	-0.11	0.09
Exact science	-0.66	0.06	-0.47	0.07
Medium Firm	-0.16	0.05	-0.46	0.06
Large Firm	-0.35	0.06	-0.88	0.08
Public sector	-0.31	0.06	0.05	0.07
Indépendant	-0.04	0.08	-0.23	0.08
Executive	-0.41	0.07	-0.66	0.08
Intermediate	-0.20	0.08	-0.37	0.08
Employee	-0.09	0.09	-0.19	0.10
RRU	0.04	0.01	0.08	0.02
Log Likehood	532		896	
Predicted Probabilities Percent Concordant	72.6		82.5	

ACKNOWLEDGMENTS

We would like to thank Thierry Blayac and Claude Diebolt for this comments and suggestions. We are also grateful to Patrice Bougette for useful remarks.

REFERENCES

[1] Affichard, J., 1981. Quels emplois après l'école : la valeur des titres scolaires depuis 1973. Économie et Statistique 173, 7-26.

[2] Alba-Ramirez, A., 1993. Mismatch in the Spanish Labor Market. Journal of Human Resources 28(2), 259-78.

[3] Barnow, B., Cain, G., Goldberger, A., 1981. Issues in the Analysis of Selctivity Bias. In Evaluation Studies Review Annual. ed. Stromsdorfer E. and Farkas G. Calif: Beverly Hills. 5, 43-59.

[4] Battese, G.E., Coelli, TJ., 1995. A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data. Empirical Economics 20, 325-32.

[5] Battese, G.E., Coelli, T.J., 1993. A Stochastic Frontier Production Function Incorporating a Model for Technicial Inefficiency Effects. Working Papers in Econometrics and Applied Statistics n°69, University of New England, Armidale.

[6] Battese, G.E., Corra, G.S., 1977. Estimation of a Production Frontier Model: With Application to the Pastoral Zone of Eastern Australia. Australian Journal of Agricultural Economics 21, 169-79.

[7] Büchel, F., Battu, H., 2003. The Theory of Differential Overqualification: Does it Work. Scottish Journal of Political Economy 50(1), 1-16.

[8] Card, D., 2001. Estimating the return to Schooling: Progress on some Persitent Econometric Problems. Econometrica 69(5), 1127-60.

[9] Chevalier, A., 2003. Measuring Overeducation. Economica 70, 509-31.

[10] Coelli, T.J., Rao, P.D.S., Battese, G.E., 1998. An Introduction to Efficiency and Productivity Analysis. Boston: Kluwer Academic Publishers.

[11] Cohn, E., 1992. The Impact of Surplus Schooling on Earnings: Comments. The Journal of Human Resources 27(4), 679-82.

[12] Cohn, E., Khan, S.P., 1994. The Wage Effects of Overschooling Revisited. Labour Economics, 12(2), 67-76.

[13] Daly, M., Büchel, F., Duncan, G., 2000. Premiums and Penalties for Surplus and Deficit Education: Evidence from the United States and Germany. Economics of Education Review 19, 169-78.

[14] Dolton, P., Silles, M., 2003. The Determinants and Consequences of Graduate Overeducation. In Overeducation in Europe: Current Issues in Theory and Policy. ed. Büchel F., de Grip A. and Mertens A. Cheltenham: Edward Elgar 189-216.

[15] Dore, R., 1976. Diploma Disease: Education, Qualification and Development. University of California Press, Berkeley.

[16] Duncan, G., Hoffman, S., 1981. The Incidence and Wage Effects of Overeducation. Economics of Education Review 1, 75–86.

[17] Dupray, A., 2001. The Signalling Power of Education by Size of Firm and the Long Term Effects on Workers Career. International Journal of Manpower 22, 13-38.

[18] Falk, M., Koebel, B., 2004. The Impact of Office Machinery, and Computer Capital on the Demand for Heterogeneous Labour. Labour Economics,11(1), 99-117.

[19] Frank, R.H., 1978. Why Women Earn Less: The Theory and Estimation of Differential Overqualification. American Economic Review 68, 360-73.

[20] Freeman, R., 1971. The Market for College-Trained Manpower. A Study in the Economics of Career Choice. Harvard University Press, Cambridge.

[21] Freeman, R., 1976. The Overeducated American. Academic Press, New-York.

[22] Gill, A.M., Soldberg, E.J., 1992. Surplus Schooling and Earning: A Critique. Journal of Human Resources 27(4), 683-89.

[23] Giret, J-F., Lemistre, P., 2004. Déclassement à l'embauche des jeunes : vers un changement de la valeur des diplômes ? Brussels Economic Review 47, 483-503.

[24] Groeneveld, S., Hartog, J., 2004. Overeducation, wages and promotions within the firm. Labour Economics, 11(6), 701-714.

[25] Groot, W., Maassen van den Brink, H., 2000. Overeducation in the labor market: A Meta-Analysis. Economics of Education Review 19(2), 149-58.

[26] Guironnet, J-P., 2006. La suréducation en France : vers une dévalorisation des diplômes du supérieur ? Économie Appliquée 59(1), 93-120.

[27] Guironnet, J-P., Peypoch, N., 2007. Human Capital Allocation and Overeducation: A Measure of French Productivity. Economics Modelling 24, 398–410.

[28] Halaby, C., 1994. Overeducation and Skill Mismatch. Sociology of Education 67, 47-59.

[29] Hartog, J., 2000. Over-Education and Earnings: Where are we, Where should we go? Economics of Education Review 19, 131–48.

[30] Haskel, J., Heden, Y., 1999. Computers and the Demand or Skilled Labour: Industry-and Establishement-Level Panel Evidence for the UK. The Economic Journal 454(109), C68-C79.

[31] Hersch, J., 1991. Education Match and Job Match. Review of Economics and Statistics 73, 140-44.

[32] Jarousse, J-P., Mingat, A., 1986. Un réexamen du modèle de gain de Mincer. Revue Économique 37(6), 999-1031.

[33] Jensen, U., 2003. Measuring Overeducation with Earnings Frontiers and Panel Data. In Overeducation in Europe: Current Issues in Theory and Policy. ed. Büchel, F., de Grip, A., Mertens, A. Edward Elgar 155-69.

[34] Kiker, B.F., Santos, M.C., De Oliveira, M., 1996. Overeducation and Undereducation: Evidence for Portugal. Economics of Education Review 16, 111-25.

[35] Kodde, D.A., Palm, F.C., 1986. Wald Criteria for Jointly Testing Equality and Inequality Restrictions. Econometrica 54(5), 1243-48.

[36] Kumbhakar, S.C., Lovell, C.A.K., 2000. Stochastic Frontier Analysis. New York: Cambridge University Press.

[37] Lemistre, P., 2003. Dévalorisation des diplômes et accès au premier emploi. Revue d'Économie Politique 113(1), 37-58.

[38] McGuinness, S., 2003. Graduate Overeducation as a Sheepskin Effect: Evidence from Nothern Ireland. Applied Economics 35, 597-608.

[39] McGuinness, S., 2006. Overeducation in the Labour Market. Journal of Economic Surveys 20(3), 387-418.

[40] Mincer, J., 1974. Schooling, experience and earnings. National Bureau of Economic Research, New York.

[41] Nauze-Fichet, E., Tomasini, M., 2002. Diplôme et insertion sur le marché du travail: approches socioprofessionnelles et salariale du déclassement. Économie et Statistique 354, 21-48.

[42] Rumberger, R.W., 1987. The Impact of Surplus Scholing on Productivity and Earnings. The Journal of Human Resources 22, 24–50.

[43] Sloane, P.J., 2003. Much Ado About Nothing? What Does the Overeducation Literature Really Tell Us? In Overeducation in Europe: Current Issues in Theory and Policy. ed. Büchel, F., de Grip, A., Mertens, A. Edward Elgar Publishers 11-45.

[44] Sloane, P.J., Battu, H., Seaman, P., 1999. Overeducation, Undereducation and the British Labour Market. Applied Economics 31, 1437-53.

[45] Tsang, M., Rumberger, R., Levin, H., 1991. The Impact of Surplus Schooling on Worker Productivity. Industrial Relations 30(2), 209-28.

[46] Thurow, L.C., 1975. Generating inequality. Basic Books, New York.

[47] Verdugo, R.R., Verdugo, N.T., 1989. The Impact of Surplus Schooling on Earnings: Some Additional Findings. The Journal of Human Resources 22, 629–43.

WORKING PAPERS DE L'AFC

Année 2008

WP2008-1	Nadir ALTINOK, Saloua BENNAGHMOUCH "School Resources and the Quality of Education: Is there a Link?"
WP2008-2	Cédric DOLIGER "La fécondité et ses déterminants. Revue de la littérature et application économétrique au cas français"
WP2008-3	Claude SCHWOB "La politique commerciale des États-Unis et la question du leadership entre 1900 et 1944"
WP2008-4	Jean-Pascal GUIRONNET "The Impact of Surplus Schooling on Earnings: Some Additional Findings II"
WP2008-5	Joan ROSÉS, Kevin O'ROURKE, Jeffrey G. WILLIAMSON "Globalization, Growth and Distribution in Spain 1500-1913"
WP2008-6	Branko MILANOVIC, Peter H. LINDERT, Jeffrey G. WILLIAMSON "Measuring Ancient Inequality"
WP2008-7	Jeffrey G. WILLIAMSON "Globalization and the Great Divergence: Terms of Trade Booms and Volatility in the Poor Periphery 1782-1913"
WP2008-8	Claude DIEBOLT, Magali JAOUL-GRAMMARE "Econométrie historique des salaires en France : une relecture des années charnières"
WP2008-9	Claude DIEBOLT, Karine PELLIER "Analyse spectrale de l'évolution de longue période des brevets en France, en Allemagne, en Grande-Bretagne, aux Etats-Unis et au Japon (18ème-20ème siècles)"
WP2008-10	Jean-Luc DEMEULEMEESTER, Olivier DEBANDE "Quality and Variety Competition in Higher Education"
WP2008-11	Claude DIEBOLT, Jamel TRABELSI "Human Capital and French Macroeconomic Growth in the Long Run"
WP2008-12	Tapas MISHRA, Mamata PARHI, Claude DIEBOLT "Human Capital Accumulation and Spatial TFP Interdependence"

La liste complète des Working Papers est disponible sur le site www.cliometrie.org