Abstract

In the last two decades, France has experienced an increase in mismatches between education and work. This article studies twenty two years of French productivity to highlight the causes and effects of over-education on the employee wages and the national income. From the INSEE and Cereq data, this analysis shows a positive effect in the short term on wages of the least qualified and overeducated worker. Furthermore, over-education phenomenon does not penalize the higher graduates. Paradoxically, if it is always profitable for individuals to increase their education investment; in term of growth, over-education of the higher graduates produce an unfavourable short term effect on GDP.
1. Introduction

In the last four decades, the French growth of student number has increased at an unprecedented rate: in one half-century, the studies duration doubled (Estrade and Minni, 1996). This evolution is explained by the lengthening of compulsory schooling, the training diversification of educational system and the selectivity decrease. This large extent of student number - and the employment crisis in these last decades – has changed the functioning of the labour market: for example, Forgeot and Gautié (1997) and Guironnet (2006a) showed a large overeducation extent for younger graduates.¹

Freeman (1976) was one of the pioneers in the research upon overeducation topic. In the USA, he finds that the education returns have significantly decreased in the seventies. Freeman assigned this fall to an excess of graduates offer. The interest of his results is that they call into question the belief that a higher degree represents a profitable investment and a virtual guarantee of economic success. Overeducation literature then develops the idea of a mismatch between supply and demand of graduates (Hartog, 2000; McGuiness, 2006).

The overeducation evolution reflects some rigidity upon the labour market as the problem of information asymmetry described by several theoretical models (e.g. Thurow, 1972; Arrow, 1973 and Spence, 1973). It also highlights the existence of labour underutilisation. This situation can be connected to the opposite phenomenon preached by Akerlof (1984): the theory of efficient wages. This idea of this approach is that the productive effectiveness of a worker mainly depends of the feeling which it has "to be well treated" by his employer. Therefore, higher wages represent a kind of gratification to encourage the individuals to increase their effort level. In contrast, an overeducated worker can tend to become under productive by a fall in this motivation (Tsang, 1987). Generalized at the productive system, it is the whole of the economy which is concerned by this phenomenon: Overeducation can be a transitory phenomenon in the professional careers of the individuals but permanent for the economy (Rubb, 2003). The majority of framework however measures the unfavourable effect of overeducation upon the individual wages. These past researches do not identify the causes, or consequences, of overeducation from an aggregate level.

Therefore, this framework proposes a new research perspective in analyzing the links between economic growth and overeducation phenomenon. In particular, we seek to know how these two entities evolve, one compared to the other. We thus try to answer at a key question: is the rise of overeducation phenomenon due to the unfavourable economic situations or does the general crisis of the economic system result from the overeducation development?

From this viewpoint, we choose to work in term of causality (Granger, 1969; Engle and Granger, 1987) which it is particularly adapted to these problems (Jaoul, 2004a). This framework is decomposed as follows: section #2 presents the database and the used methodology; section #3 studied the links between education and wages and section #4 identified the links between overeducation and growth.

2. Data, Indicators and Methodology

In contrast to the previous studies, the interest of our approach - in term of causality - is to take the needs of labour market and the intensities of the use of worker skills into account. For this purpose, we choose to re-use the time series elaborate in a previous framework (Guironnet, 2009).

¹ i.e. for a given job, the required skills are lower than those accredited by educational system.
2.1. Time Series of Overeducation and Undereducation Phenomenons

To get data over the period 1980-2002\(^2\), our series are built with a wage downgrading measure (see framed #1 in appendices), initially investigated by Nauze-Fichet and Tomasini (2002). This proxy of overeducation rate can be seen as an indicator of stress between supply and demand of graduates.

Our time series are based on the INSEE surveys, entitled (“Emploi” database). From this sample, part-time workers are excluded. Furthermore, individuals are collected following their age: lesser of 25 years, 25-39 years, 40-49 years, 50-59 years and more of 60 years. In order to assess worker mismatches, a ranking subdivided in six educational attainments is used: the graduates in higher education (three, four, five and six years in higher education); the graduates of short courses in higher education (the two first years), the bachelors, the graduates of secondary school, i.e. vocational degrees (CAP or BEP obtained at least sixteen), those of school, i.e. vocational degree (BEPC obtained at least fourteen) and the workers without degrees. It would obviously be better to work upon lesser aggregated ranking, in particular for the higher education. Nevertheless, the number of students from higher education is not sufficient (in particular for the higher level of degrees), in the eighties, to consider a more detailed list.

Following our approach, we choose in our analysis to favour the comments upon the higher educational system. Figure #1 thus presents the evolution of relative numbers (compared to the working population in employ) of overeducated and undereducated workers from higher education.

Figure. 1. Evolution of the share of mismatched workers of higher education and unemployment rate.

Source: Employment surveys, INSEE.
Field: working population in employment gathered by age groups.

From figure #1, the mismatch evolution of the higher graduates is increasing. Paradoxically, these leavers of educational system – better trained - have greater difficulties to get a job

\(^2\) We are conscious that the number of observations is weak for a study in time series. Since this framework is a new track of research, we do not have yet a more packed database.
corresponding to their adequate educational level. In addition, the increasing gap between the two curves shows a growing release of the links between educational levels and wages. Nowadays, the overeducated share of higher education - compared to the working population - reached approximately 10% in 2002 against 2.6% in 1980. Furthermore, the growth of undereducated population of the higher education moves from 0.9% in 1980 to 3.3% in 2002. These two results lead us to think that the higher degrees avoid some overeducation risks. This phenomenon however is today not negligible for the higher levels of education.

In comparison to unemployment rate, the overeducation phenomenon displays comparable evolutions with a weak time delay. Nevertheless, the changes in the economic situation are not the only explanatory factors: Some evolutions in opposite direction of the unemployment rate and the overeducated population of higher education can be seen over these last years. The better economic situations do not stop the upward trend of overeducation phenomenon (Fondeur, 1999). This evidence leads us to think that the overeducation evolution would come from a more structural phenomenon than the economic situation (Guironnet, 2006b): It would be the result of an increasing mismatch between supply and demand of graduates.

Following this theoretical background, our objective is to study the causes and consequences of the structural growth of overeducation phenomenon upon the individual wages and the GDP growth.

2.2. Methodology

Causality analysis requires the use of VAR modelling (Vector Auto Regressive). The advantages of over classic modelling non-structural VAR modelling are first that it allows better dynamic analysis of the systems, taking into account the intrinsic structure of the series and the dynamic effects between the variables and secondly that it makes it possible to envisage all causal relations between two variables without any a priori assumptions with regard to the exogeneity of any of them.

The VAR models form a continuation of the work of Granger (1969) on the causal relation between two variables. Using this viewpoint, Sims proposes modelling that extends analysis of causality to a system of several variables. For this, he proposes to treat all the variables in an identical manner without a condition of exclusion or exogeneity and selecting the same delay for each of them in all the equations.

VAR models nevertheless have their limits. The first is the problem of the number of variables to be included in the model and the resulting estimation problem. Indeed, VAR models differ from theory-based structural models in the greater scope left for empiricism, but how many variables should be chosen in this case? The number of variables to be included in the model thus brings the problem of “vanishing degrees of freedom”, (Johnston, 1999). Indeed, considering 20 variables and 4 delays leads to estimating 80 coefficients per equation and the number of unknown coefficients often approaches the size of the sample analysed.

Another criticism often aimed at VAR models is the small amount of theory to which they refer, describing them as a-theoretical models. This 'theory versus measurement' debate had already opposed economists in the 1920s following the work of Mitchell (1913) and reappeared in the 1980s with that of Sims. This debate however is far from settled, and if VAR models are criticised for their lack of theory, the theoretical models of supporters of the

---

3 The intrinsic structure of the series is related to its identification in the ARIMA classification (Box and Jenkins, 1976).

4 The 'theory versus measurement' debate started in the analysis of Mitchell cycles (1913) that laid the empirical foundations of modern macroeconomic theory.
Cowles Commission are also criticised for their lack of flexibility (Lucas, 1976). In the face of these viewpoint differences, our approach proposes the reconciliation of theory and measurement in proportions providing both the theoretical and empirical debate required in economics.

In these models, each equation of the model describes the evolution of a variable in function of its last values and of last values of the other variables of the system. Once the optimal lag \( p \) has been determined, the analysis can then be taken in two directions: on the one hand, study of the dynamics of the model and, on the other, analysis of causality which has two forms: analysis of the long-run relationship using cointegration and the study of causality relationships.

The term 'cointegration' first appeared in 1964 in Sargan's work but only received true theoretical coverage in 1987 by Engle and Granger. Cointegration encompasses the idea that two or more series evolve together in time and generate statistical equilibrium in the long term, whereas the variables may move in different directions in the short term. If they however continue to move far from each other in the long term, economic forces such as a market mechanism or government intervention makes it possible to bring them towards each other. The existence of cointegration between variables implies that the causality is analysed by using a VECM (Vector Error Correcting Model). This corrected model takes into account this long run relationship in order to avoid spurious regressions (Granger & Newbold, 1974). Indeed, in the case of cointegrated series, the “goodness of the fit” of the model is only due to the fact that series are non stationary; in this case, forecasts from the model are not robust.

In addition to the identification of the generating process of each variable of the model with unit root tests, the finality of a model VAR is the identification of the relations of causality between the variables. Demonstrating causal links between the economic variables enables better understanding of economic phenomena and brings additional information in regard to the anteriority of the events. This also engenders the establishment of an optimised economic policy.

The definition of causality is given by Granger (1969): variable \( y_{2t} \) causes variable \( y_{1t} \) if the forecasting of the latter is improved by incorporating in the analysis information concerning \( y_{2t} \) and its past. There are two approaches of causality: Granger (1969) and Sims (1980). Causality within the meaning of Granger (1969) relates to the propagation of interpretable deterministic impulses like modifications related to structural changes. In contrast, the analysis of Sims (1980) is based on the propagation of stochastic impulses representative of "surprises". On the level of the results the two approaches are generally equivalent (Bruneau, 1996) but one chose here a test of Granger, because we consider that it is legitimate to associate the labour market and the higher education to a non stochastic context.

With two variables, the VAR(p) model is as follows:

\[
\begin{bmatrix}
    y_{1,t} \\
    y_{2,t}
\end{bmatrix} = \begin{bmatrix} A_0^1 & A_1 & A_2 & \cdots & A_p \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\
    y_{2,t-1} \end{bmatrix} + \begin{bmatrix} B_0 & B_1 & B_2 & \cdots & B_p \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\
    y_{2,t-2} \end{bmatrix} + \cdots + \begin{bmatrix} C_0 & C_1 & C_2 & \cdots & C_p \end{bmatrix} \begin{bmatrix} y_{1,t-p} \\
    y_{2,t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\
    \varepsilon_{2,t}
\end{bmatrix}
\]

The test proceeds in two stages:

---

5 Lucas (1976) used strong theoretical bases to argue that these models are fundamentally imperfect for assessing the consequences of the results of political alternatives. He puts forward the reason that, for example, their functioning plans little advice for political managers with regard to predicting changes of effect in economic policy because it is improbable that the parameters of the models remain stable under alternative economic policies.
- \( Ho \) is tested: \( y_{2t} \) does not cause \( y_{1t} \), that is to say that the coefficients of the matrix blocks \( B \) are nil.

- \( H'0 \) is tested: \( y_{1t} \) does not cause \( y_{2t} \), that is to say that the coefficients of the matrix blocks \( C \) are nil.

It is then possible of calculates the sign of the relation of causality. The sign of the relation of causality from \( Y_1 \) to \( Y_2 \) is calculated by:

\[
\eta_{Y_1 \rightarrow Y_2} = \frac{\sum_{i=1}^{p} C_i}{1 - \sum_{i=1}^{p} D_i}
\]

If \( \eta > 0 \), the relation of causality is positive and a complementarity between the two entities can be confirmed; in the other case, it is a relation of substitutability which is found.

### 3. Links between Education and Wages

The present section leads a micro-econometric approach to analyze the links - in terms of causality - between the structural growth of overeducation phenomenon and worker wages.

Several frameworks have showed a positive correlation between educational attainment and wages (for a survey upon the education returns, see for example, Psacharapoulos, 1994). These studies however neglect the demand side of the labour market. Indeed, what do these results become when we take mismatches between job and training into account?

Two types of analysis are then produced: firstly, one studied the overeducated workers and, secondly, a similar study is undertaken for the undereducated workers. The time series analysis is lead with the sub group because we want to analyse the impact of each category of individuals (with and without diploma) upon wages. The database and the methodology described in the section #1 are now used. The following variables are then considered:

- Overeducated workers of the Higher Degree (OHD);
- Overeducated workers Without Higher Degree (OWHD);
- Annual average Wages of the Working Population (AWPE);

Unit root tests (Elliott, Rothenberg & Stock, 1996) show that the variables are stationary. A necessary condition of cointegration between two variables is that the variables are integrated in the same order. The series are stationary: they are then not integrated and there is no relation of cointegration between the variables. We can thus show an absence of long term relation between the studied variables. These results are in accordance with Rubb’s idea (2003) concerning the transitory character of the overeducation influence upon the professional careers. The test of causality highlights only one significant relation, with a threshold of 10%:

\[ + \overset{\rightarrow}{\text{OWHD}} \to \text{AWPE} \]

This link is confirmed by the variance decomposition: wages are more sensitive to a variation of the overeducated workers without any degree of the higher education (95 %) that the latter are not sensitive to a wage change (2 %). The wage downgrading measure is thus not sensitive to the wage deviation: this result confirms the validity of our applied measurement. Furthermore, this result confirms the stylized fact from ORU function: an overeducated worker earns less that their counterparts adequately matched with the same educational

---

6 We reason in growth rate for each variable.
attainment (i.e. our measurement definition) but they earn more than those adequately matched with his job (i.e. our causality result). Therefore, an increase in overeducated workers - with a lower level of education - produces a positive effect on the corresponding wages. This result can be explained by the exogenous evolution of the French minimum wage. For example, the guaranteed minimum wage (SMIC) increases with monetary inflation controlled, of 13% over the period 1990-1998. Despite overeducation risks, the individuals must pursue their school course to the baccalaureate level (or a training of a similar level) since the profitability of education is guaranteed by the minimum wage. In contrast, we do not find any significant link between the overeducated workers of higher education and the wages. The mismatches, for the highest levels of degree, does not modify the profit prospects, in particular their relative advantage in term of wages (high schools and 3rd cycles).

In our second analysis, the variables taken into account are identical but concern the undereducated workers. The econometric analysis does not reveal any relation between these variables. Therefore, mismatches between training and job do not produce any unfavourable wage effect. In contrast, our analysis in term of causality reveals a wage benefit for the overeducated workers holding a qualification level lower - or equal - to the academic baccalaureate. In long period, the workers must adopt a strategy to protect themselves from their non-adaptability to anticipate the future technological improvements and the new needs of skills.

4. Links between Education and Growth

The object of this section is to study the impact of the overeducation, either on the individual wages, but upon the national income. In a previous study (Jaoul, 2004b), a link between higher education and economic growth have been showed over the period 1820-2000. None relation however has been found since 1950. This result shows a lack of coordination between the economic and educational policies: France has some difficulties to adapt its higher educational system to the economic system. A similar problem also occurs in Switzerland: Wirz and Atukeren (2005) showed that a better utilization of some workers investment in education (especially for women) should contribute positively to swiss economic growth.

A more specific analysis in terms of degree (Jaoul, 2002) shows the impact of the degrees of higher education upon the French economic growth. Among them, PhD - and more particularly those of the health sector (Medicine and Pharmacy) - has a positive effect upon the national income (Jaoul, 2004b). This framework however takes not the feasible effects of the mismatches between training and job into account.

Therefore, the main objective of this part is to assess in which way the overeducation evolution has an impact on the economy. For this purpose, the growth rates of the following variables are considered:

- GDP;
- The Share of Overeducated workers of the Higher Education (SOHE);
- Overeducated workers of the Higher Education (OHE);
- The Share of Overeducated workers Without any degree of Higher Education (SOWHE);
- Overeducated workers Without any degree of Higher Education (OWHE);

7 Source: “Séries longues sur les salaires (2000), Insee, Résultats Emploi-Revenus, n° 72”. 
Unit root tests reveal a stationnarity of all variables. They do not present any relation of long term. This result tends to contradict the idea of a persistent overeducation phenomenon upon economy (Rubb, 2003). Concerning the short term, only a causality relationship is highlighted (significant threshold of 10 %):

\[ \text{SOHE} \rightarrow \text{GDP} \]

Following this result, an increase of the share of overeducated workers of the higher education produces an unfavourable effect upon the economic growth. Over these two last decades, the recruitment of overqualified workers would then produce a growth deceleration rather than unfavourable economic conditions produce an increase of overvaluation skills during job recruitments. In addition, we find a similar result than a previous study with the same database (Guironnet, 2009): overeducated workers of higher education decrease significantly the national productivity. From an individual level, students have some benefits to get the higher educational level but for the government this behaviour seems to have negative effect: on one hand, the training lengthening represents a surplus of cost for the educational system; on the other hand, mismatches in skill allocations produce a suboptimal growth of national income.

5. Conclusion

With an increasing phenomenon of overeducation over these two last decades, the overeducation literature shows an unfavourable effect of this phenomenon on wages. The overeducation phenomenon however is not without benefits: for a given job, the overeducated workers have higher wages compared to the workers adequately matched (Sloane and al. 1999). Furthermore, overschooling can lead to a better situation with the valorization of job training.8

From twenty two years of French productivity, our analysis seems to confirm that the overeducation phenomenon does not produce any unfavourable effect from an individual level. On the one hand, this phenomenon can produce some advantages for the least qualified workers. On the other hand, it would not affect, or slightly, the most qualified workers: they preserve their wage advantage (in particular for high schools and 3rd cycles). Thus, the credentialism scenario9 - or the qualification inflation - does not seem irrational: the individual prospect, which consists to get the higher educational level, always involves a rentable investment. Our analysis however exceeds this microeconomic viewpoint in analyzing the links between overeducation and GDP growth.

Following this approach, this phenomenon seems to reduce the GDP growth. This result is pregnant in term of policy recommendations: over these last decades, the French policy-makers tend to reduce the unemployment of young people, in particular for the least qualified (see “young employment”10, and “apprenticeship contract”11). Nevertheless, they must now improve the professional insertion of the most qualified to find an adequate job in

8 Some empirical works show that overeducation would produce some benefits upon the development of professional careers, as well in term of promotions than in wages (Groeneveld and Hartog, 2004; Dekker et al., 2002).
9 See Spence (op. cit.).
10 Employment contract in definite duration (maximum 5 years), for young people (from 18 to 26 years old) in unemployment, to finance job of public utility.
11 Employment contract in definite duration (between 1 and 3 years) for young people (from 16 to 26 years old) which enables to young people to acquire a professional degree.
respect to their educational attainment. Such prospect would allow a decrease of the underutilization of the most qualified and it would improve the economic growth.\footnote{For an approximate evaluation of the profits of expected growth, the concerned reader can consult Guironnet’s paper (2006b).}

This framework however can be showed as a new research topic and our results can be considered as intuitive. Our objective is to underline a feasible unfavourable effect of the overeducation phenomenon upon growth and to have a particular attention for the professional insertion of the most graduates. Other studies could supplement this analysis: for example, the series developments over a longer period would undoubtedly make it possible to specify our results and new macroeconomic models would also supplement our approach.

6. References


**APPENDIX**

**UNDER- AND OVER-EDUCATION MEASUREMENT**

Two methods are available to measure objectively the overeducation extent:

- the normative method with experts of labour market which define the requirements for a range of job (see dictionary of occupational titles for the USA case).
- statistical method which defines “normal” matches on the basis of what emerge from the statistical analysis as being the most frequent. Several studies have used different statistical measures like the modal values (Kiker et al. 1996), the standard deviation from the mean (Verdugo and Verdugo, 1989), etc.

As the normative method demands a frequent redefinition of the required skills of the occupations with the changes of productive system, the statistical method is often used as substitute. In the present paper, we not choose to investigate further the normative method which will probably be largely biased by technical evolution. We thus favour the statistical method despite the arbitrary feature of this measurement. In our case, we choose a wage
statistical approach which has the main advantage to not consider – for example – a technician who earns an executive wage as overeducated.

Following this approach, a worker will be considered as overeducated (resp. undereducated) if more than 50% of holder individuals of the immediately lower degree (resp. higher) earn more (resp. less). For example, if a worker with five years in higher education earns 1500€ but that 50% of the worker with four years in higher education earn more, this worker can be considered as overeducated.

Following this definition, we determine a relational indicator between “degree and job” which is homogeneous over a long period. This proxy is not affected by the "biased technical progress" and the "diploma inflation" process. This measurement however remains affected by other factors which contribute to the wage heterogeneity of population. Thus and as evoked in the section #2.1, the part-time workers are excluded from our sample. Our analysis however has not been carried out following the gender (the number of women student is not sufficient for the distinction). Therefore, our overeducation rates can be overestimated by other wage discriminations.