# ESTIMATING THE IMPACT OF TUNISIAN TRAINING PROGRAMS ON WAGE, USING A SIMULTANEOUS EQUATIONS MODEL WITH SELF-SELECTIVITY

RIHAB BELLAKHAL*

MOHAMED-BADRANE MAHJOUB**

AVRIL 2010

ABSTRACT. The aim of this paper is to assess the impact of vocational programs offered in Tunisia on the wages of individuals. The data we use are from a study carried out in Tunisia by the Ministry of Vocational Training and Employment on the graduates of the national vocational training in 1998. The estimated model corresponds to three simultaneous equations determining the participation in training, the insertion into the labor market and the wages observed, where the residuals of the two selection equations (training and insertion) are assumed to be correlated with that of the wage equation. The results show that the individuals who have received professional training in Tunisia have on average higher wages than those who had not benefited.

JEL Classification: C31, J18.

Keywords: Public Policy, impact evaluation, self selectivity, simultaneous equations, training programs, Tunisia.

*Faculté des Sciences Economiques et de Gestion de Tunis, B.P 248 El Manar II 2092 Tunis. Email: bellakhal_rihab@yahoo.fr.
**Université Paris1 Panthéon-Sorbonne, CES, 106, boulevard de l'Hôpital, 75647 Paris cedex 13, France. Email: gbadrane@univ-paris1.fr.

## 1. INTRODUCTION

In this paper we evaluate the impact of vocational training programs offered in Tunisia on the wages of individuals. Several studies on the evaluation of public policies, especially those of training and employment have been conducted in these last years. The common objective of these papers is to ask about the efficiency of policies studied whose implementation requires important financial and human resources and the results of which are always in question.

In practice, evaluating a given policy is not easy to achieve, because in addition to the questions we must ask about the efficiency of the policy studied, other questions are raised about the choice of the method to be used for evaluating this policy. This method must enable us to identify the effects which are only due to the studied policy. The majority of the evaluation studies has been performed on non-experimental data such as those we use in this article (see, for example, Angrist and Krueger (1991), Bonnal, Fougère and Sérandon (1997), Heckman, Ichimura and Todd (1997 , 1998), Heckman and Smith (1998), Dehejia and Wahba (1999, 2002), Fougère, Goux and Maurin (2001)). In comparison to experimental data, estimating the impact of a given policy on the basis of non experimental data is not easy to achieve because of the problem of selection bias present in such data. Any evaluation process should take into account this problem of selection bias.

A problem of selection bias exists when people's participation in the training program is the result of a decision taken by those most eligible. This decision depends on both observable characteristics (such as the place of residence, the education level, the age,...) and unobservable ones (such as the willingness to work, the individual ability,...). Then, The assignment of individuals to the program is by self-selection and not by random assignment.

From the econometric point of view, this corresponds to a problem of endogeneity of the variable of interest (training) in the outcome equation that we want to study (wage here) (see Heckman (1978) and Heckman and Hotz (1989) ).

In literature, several methods have been proposed to deal with this problem of endogeneity or self selection. Heckman (1976) suggests using instrumental variables to correct this problem. This method was subsequently used in several studies (see,for example, Angrist and Krueger (1991), Card (1993), Imbens and Angrist (1994), Heckman and Smith (1998), Heckman and Vytlacil (2000)). However, the difficulty in using this method lies in choosing the appropriate instrumental variable. In the context of employment policies, this variable should affect the selection into the program without directly affecting the outcome variable.

Another way to deal with the problem of self-selectivity is the use of parametric selection models. In such models, we simultaneously estimate the equation of treatment and observed outcome with supposing that we know the true joint distribution of error terms of these equations. The parametric selection model most commonly used in literature is the selection model with normal disturbances. It is this approach that we adopt in our paper. The advantage of this model is that it takes into account the phenomena of selection on observables and unobservables.

It allows dependency between the various disturbances of the different equations conditionally on observable characteristics.

For our empirical framework, we use a non-experimental micro-data from a study conducted in 2001 by the Tunisian Ministry of Vocational Training and Employment. This study has focused on the graduates of initial vocational training in 1998. Tunisia established several years ago many vocational training programs aiming to provide young people leaving school very early, qualifications and training to enable them to integrate into the labor market. However, despite efforts, until now there is not a clear idea about the effectiveness of these programs. All the studies made on this subject, including the one we use the data, provide both summary and general results, far from the scientific evaluation studies carried out in other countries (the United States and France for example). Hence the contribution of our paper.

In total the survey covered a sample of 1,002 individuals and has provided a number of relevant information concerning the characteristics of individuals, their situation on the labor market and the characteristics of the job that they occupy at the time of the study. The information has been collected from two surveys, a survey carried out among the main beneficiaries of vocational training (treatment group), and further investigation with non-beneficiaries (control group).

Using the information contained in these surveys, we estimate the impact of programs on the wages of individuals which is a valorization criterion of accepted employment . As we reported earlier, the approach we use is parametric. It is based on modeling simultaneously the participation decision and the outcome variable, and the specification of the joint distribution of disturbances. As part of our study, three variables are observed simultaneously for each individual in the sample. First, eligible individuals decide whether or not to participate in a vocational program. Following this involvement, people can find a job or not. Finally, for those who have found employment we can observe the wage. This model corresponds to a system of simultaneous equations determining the participation in training, the insertion into the labor market and the wage. The disturbances of the first two equations are assumed to be correlated with that of the wage equation, which takes into account the presence of unobserved heterogeneity in the data. The estimated model is a model with a double selection (see Lee (1978), Maddala (1983)) where the equation of participation is the first selection equation and the insertion into the labor market is the second. It is comparable whith that used by Fougère, Goux and Maurin (2001) to evaluate the impact of training sponsored by employers on employees mobility and wage in France, whose results show that training within firm has no significant impact on the wages of workers.

The estimation method we use is that of maximum likelihood. Note that the likelihood function of our model is a bit complex because it dépends on the conditional densities of the disturbances of the two selection equations with respect to the error term of the wage equation. We also have introduced into the equation of participation an exclusion variable necessary to estimate the impact on wages. The estimation results show that the passage of individuals through a training program significantly increase their monthly wage. The results also show that unobserved

characteristics facilitating the participation in training were not correlated with unobserved characteristics affecting wages.

In Section 2 we present the data and some descriptive statistics. Section 3 presents the model and the likelihood function we estimate. Section 4 presents the estimation results, and Section 5 concludes the text.

## 2. The Data

The data used in this work are from the survey of graduates of the initial vocational training conducted by the Ministry of Vocational Training an Employment in Tunisia in 2001. The survey covered a sample of 499 individuals graduated from the vocational training of 1998 in its different types[1] and a sample of 503 individuals who had not participated in training serving as a control group, is therefore a total sample of 1,002 individuals.

The group of graduates was interviewed 36 months after leaving the training (i.e., in 2001). The questionnaire for this survey was designed so that the detailed information on the individual characteristics and their professional situation at the time of the survey can be obtained, especially in terms of employability and income. The information on the family characteristics were also provided by this survey, which allows to relate the family and social context of individual at the time of his enrollment.

The sample of non-beneficiaries was selected from a list of job seekers registered for the first time at the Offices of the Tunisian Agency for Employment in 1996 or 1997. These are students who dropped out of the general education when the graduate starts training, and who possess the qualifications required to enter a training program. This group is composed of individuals who were eligible for the programs studied but who, for one reason or another, did not participate. Investigators have also ensured that none of this group has participated in another training program or employment assistance in order to avoid contamination bias. These individuals, the number of which is 503, were interviewed at the same time as the group of beneficiaries and have answered the same questionnaire as this group. Thus, a number of information mainly concerning their individual characteristics, their employment status and their possible insertion in the labor market have been collected.

Table 1 shows the main characteristics of the sample of graduates that we compare to those of the group of non-participants. This allows us to see if these two groups are similar, which is a necessary condition for the validity of the control group and its use in estimating the impact of the programs studied.

The comparison between the main characteristics of the treatment group and the control group shows that 37.6% of the individuals of treatment group are women against 45.1% for the control group. The average age in this group is 25 years against 29 years for the control group. Regarding the educational level, the majority

[1] Four types of programs corresponding to four levels of qualifications (CAP, BTP, BTS and Learning) are offered in Tunisia. However, these programs are aggregated into one because of small sample sizes representing each.

(49.6%) of participants has an educational level equivalent to secondary second cycle, which is similar to the non-participants whose 40.7% of them have the level of graduate school. The proportion of individuals whose educational level does not exceed the sixth year of basic school is respectively equal to 21.6% and 26.4% for the treatment group and control group.The participants come from families whose average size is 6 people, for the non-participants, this value is 5 people. Furthermore, 16.3% of the fathers of the members of the treatment group are active workers, 12.1% are inactive and 43.2% died. These proportions are very similar to those observed for members of the control group (17.8% for workers, 7.3% for inactive and 42.7% for deceased).

**Table 1: Descriptive Statistics**
**(Treatment Group - Control Group)**

| | **Treatment Group** | | **Control Group** | |
|---|---|---|---|---|
| | Mean | Std.D | Mean | Std.D |
| Women (1 if women, 0 otherwise) | 0.376 | 0.485 | 0.451 | 0.498 |
| Age | 25.180 | 3.473 | 29.232 | 5.142 |
| Educational Level | | | | |
| none | 0.012 | 0.109 | 0.007 | 0.088 |
| (1-6) th year of basic school | 0.216 | 0.412 | 0.264 | 0.441 |
| (7-9) th year of basic school | 0.036 | 0.186 | 0.063 | 0.244 |
| Secondary 1st cycle | 0.140 | 0.347 | 0.149 | 0.356 |
| Secondary 2nd cycle | 0.496 | 0.500 | 0.407 | 0.491 |
| Higher | 0.098 | 0.297 | 0.107 | 0.309 |
| Year of leaving school | 1994 | 2.363 | 1990 | 6.476 |
| Family Size | 6.158 | 2.046 | 5.380 | 2.011 |
| Head's occupation | | | | |
| Executive | 0.066 | 0.249 | 0.068 | 0.253 |
| Senior technician | 0.056 | 0.231 | 0.034 | 0.182 |
| Qualified worker | 0.111 | 0.314 | 0.127 | 0.334 |
| Worker | 0.163 | 0.370 | 0.178 | 0.383 |
| Unemployed | 0.048 | 0.215 | 0.089 | 0.285 |
| Inactive | 0.121 | 0.326 | 0.073 | 0.260 |
| Other (died) | 0.432 | 0.495 | 0.427 | 0.495 |
| Father's education | | | | |
| Illiterate | 0.331 | 0.471 | 0.334 | 0.472 |
| Primary | 0.360 | 0.480 | 0.354 | 0.478 |
| Secondary | 0.230 | 0.421 | 0.246 | 0.431 |
| Higher | 0.076 | 0.265 | 0.064 | 0.245 |
| Residence | | | | |
| Big city | 0.458 | 0.498 | 0.649 | 0.477 |
| Small/ medium town | 0.450 | 0.498 | 0.318 | 0.466 |
| Outside communes | 0.090 | 0.287 | 0.032 | 0.176 |

Regarding the educational level of the father, the proportions for both groups are very similar. 36% of the fathers of the participants have an education equal to primary school against 35.4% for the non-participants. Regarding the place of residence, 45.8% of the participants leave in big cities. This proportion is slightly higher for the non-participants, it is of the order of 64.9%.

In conclusion, we can say that the members of the comparison group have observable characteristics very close to those of the sample of the beneficiaries of the training, which confirms the validity of the comparison group constructed.

## 3. The Model

For each individual $i$ in the sample, we observe simultaneously three variables. Let $D_i$ be a dummy variable equal 1 if the individual $i$ has participated in the training program; $E_i$ a dummy variable equal 1 if the individual $i$ has found a job and $Y_i$ the variable representing the wage offered to the individual $i$ for the found job.

The econometric model is then a double selection model (Lee ((1978), Maddala (1983)) which corresponds to a system of three equations specified as follows:

$$D_i = \begin{cases} 1, & \text{if } D_i^* = X_{1i}\beta_1 + \varepsilon_{1i} > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

$$E_i = \begin{cases} 1, & \text{if } E_i^* = X_{2i}\beta_2 + \varepsilon_{2i} > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

$$\ln Y_i = X_{3i}\beta_3 + \alpha_Y D_i + \nu_i; \tag{3}$$

where $X_{1i}$ represents the exogenous variables that may explain the participation in training; $X_{2i}$ is the set of exogenous factors that may explain the insertion in the labor market and $X_{3i}$ are exogenous variables that determine the wages.

Note that some exogenous variables such as the age, the sex and the educational level can both belong to $X_1$, $X_2$, and $X_3$. These variables are important determinants of both the participation as well as the insertion and the wages.

Furthermore, it should be noted that the aim of this model is to estimate the impact of the participation on the wage (the variable of interest here). The equation of insertion is in this model, a selection equation since we can not observe the wages for the individuals who don't work.

Let us note, however, the absence of the treatment variable ($D$) in the employment equation that makes it clear that the passage through the training programs has not affected the employability of the participant. This specification is however justified by the difficulties inherent in adding this variable in the employment equation. Indeed, the impact of training on the insertion could be identified only if there is exogenous variables qualified as exclusion variables, explaining the participation in training without having an effect on insertion. However, finding such variables is not easy in practice.

To estimate this model, we assume that the vector of disturbances $(\varepsilon_{1i}, \varepsilon_{2i}, \nu_i)$ follows a trivariate normal with mean zero and covariance matrix $\Omega$ such as:

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \nu_i \end{pmatrix} \rightsquigarrow N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} & \rho_1\sigma \\ \rho_{12} & 1 & \rho_2\sigma \\ \rho_1\sigma & \rho_2\sigma & \sigma^2 \end{pmatrix}\right] \tag{4}$$

where:
$\rho_{12}$: correlation coefficient between $\varepsilon_1$ and $\varepsilon_2$;
$\rho_1$: correlation coefficient between $\varepsilon_1$ and $\nu$;
$\rho_2$: correlation coefficient between $\varepsilon_2$ and $\nu$;
$\sigma^2$: variance of $\nu$.

The estimation method we use is that of maximum likelihood which is a bit complicated here because it involves the conditional density of the perturbations $(\varepsilon_{1i}, \varepsilon_{2i})$ compared to the error term $\nu_i$, which we must determine to explain the likelihood function.

To determine the conditional densities $(\varepsilon_{1i}, \varepsilon_{2i}) \mid \nu_i$, we use the theorem of marginal and conditional normal distributions (Greene (2005)). From this theorem, we can prove that[2]:

$$(\varepsilon_{1i}, \varepsilon_{2i}) \mid \nu_i \rightsquigarrow N\left[\begin{pmatrix} \frac{\nu_i\rho_1}{\sigma} \\ \frac{\nu_i\rho_2}{\sigma} \end{pmatrix}, \begin{pmatrix} 1-\rho_1^2 & \rho_{12}-\rho_1\rho_2 \\ \rho_{12}-\rho_1\rho_2 & 1-\rho_2^2 \end{pmatrix}\right] \tag{5}$$

We define:

$\mu_1^* = \frac{\rho_1}{\sigma}\nu_i$ : the conditional expectation of $\varepsilon_{1i} \mid \nu_i$;

$\mu_2^* = \frac{\rho_2}{\sigma}\nu_i$: the conditional expectation of $\varepsilon_{2i} \mid \nu_i$;

$\sigma_1^* = \sqrt{1-\rho_1^2}$: conditional standard deviation of $\varepsilon_{1i} \mid \nu_i$;

$\sigma_2^* = \sqrt{1-\rho_2^2}$: conditional standard deviation of $\varepsilon_{2i} \mid \nu_i$;

$\rho_{12}^* = \frac{\rho_{12}-\rho_1\rho_2}{\sqrt{(1-\rho_1^2)(1-\rho_2^2)}}$: the correlation coefficient between $\varepsilon_{1i}$ and $\varepsilon_{2i}$ conditionally on $\nu_i$.

With these new parameters, we can explicit the likelihood of the model. We give in the following, the individual contributions to the likelihood, conditional on observable and unobservable specifications. Four situations can occur depending on the values taken by the three endogenous variables ($D_i$, $E_i$ and $Y_i$).

[2]See demonstration in Appendix.

Thus, the contribution to the likelihood of an individual not involved in training, not finding a job and who therefore did not pay $\big(D_i = 0;\ E_i = 0;\ Y_i \text{ unobservable}\big)$ is:

$\mathcal{L}_i = \mathrm{Prob}(D_i = 0,\ E_i = 0);$
$= \mathrm{Prob}\big(\varepsilon_{1i} \leqslant -X_{1i}\beta_1,\ \varepsilon_{2i} \leqslant -X_{2i}\beta_2\big);$
$= \Phi_2(-X_{1i}\beta_1,\ -X_{2i}\beta_2,\ \rho_{12}).$

We set:
$A_i = -X_{1i}\beta_1;$
$B_i = -X_{2i}\beta_2;$
$\Rightarrow \mathcal{L}_i = \Phi_2(A_i,\ B_i,\ \rho_{12});$
where $\Phi_2$ is the distribution function of bivariate normal distribution.

Similarly, for an individual who has participated in training program but did not get a job$\big(D_i = 1;\ E_i = 0;\ Y_i \text{ unobservable}\big)$, we have:

$\mathcal{L}_i = \mathrm{Prob}(D_i = 1,\ E_i = 0);$
$= \mathrm{Prob}(\varepsilon_{1i} > -X_{1i}\beta_1,\ \varepsilon_{2i} \leqslant -X_{2i}\beta_2);$
$= \Phi_2(-A_i,\ B_i,\ -\rho_{12}).$

The contribution to the likelihood of an individual not involved in training, having found a job and receiving a salary $\big(D_i = 0;\ E_i = 1;\ Y_i \text{ observable}\big)$ is written as:

$\mathcal{L}_i = \mathrm{Prob}(D_i = 0,\ E_i = 1,\ Y_i = y_i);$
$= \mathrm{Prob}(D_i = 0,\ E_i = 1/Y_i = y_i) \times \mathrm{Prob}(Y_i = y_i);$
$= \mathrm{Prob}(\varepsilon_{1i} \leqslant -X_{1i}\beta_1,\ \varepsilon_{2i} > -X_{2i}\beta_2/Y_i = y_i) \times \mathrm{Prob}(Y_i = y_i);$
$= \frac{1}{\sigma}\phi(\frac{C_i}{\sigma}) \times \Phi_2(F_i,\ -G_i,\ -\rho^*_{12});$

where $C_i = \ln Y_i - X_{3i}\beta_3 - \alpha_y D_i = \nu_i;$

$$F_i = \frac{(-X_{1i}\beta_1 - \mu^*_1)}{\sigma^*_1} = \frac{(-X_{1i}\beta_1 - \rho_1 C_i/\sigma)}{\sqrt{1-\rho_1^2}};$$

$$G_i = \frac{(-X_{2i}\beta_2 - \mu^*_2)}{\sigma^*_2} = \frac{(-X_{2i}\beta_2 - \rho_2 C_i/\sigma)}{\sqrt{1-\rho_2^2}};$$

$\phi$ is the density function of the standard normal distribution.

Finally, for an individual who participated in the training program and obtained employment $\big(D_i = 1;\ E_i = 1;\ Y_i \text{ observable}\big)$, we have:

$\mathcal{L}_i = \mathrm{Prob}(D_i = 1,\ E_i = 1,\ Y_i = y_i);$
$= \mathrm{Prob}(D_i = 1,\ E_i = 1/Y_i = y_i) \times \mathrm{Prob}(Y_i = y_i);$
$= \frac{1}{\sigma}\phi(\frac{C_i}{\sigma}) \times \Phi_2(-F_i,\ -G_i,\ \rho^*_{12}).$

## 4. RESULTS

The estimation of the model presented above to our sample composed of 944 individuals[3] has been performed on "STATA". The exogenous variables introduced in the equations are the sex, the age, the level of education, the place of residence, the level of education of the father and the situation of the household head. Unfortunately, the available data do not allow us to introduce variables on the characteristics of the firm. These variables, however, enrich the specification of the equations of insertion and wage.

Furthermore, we have introduced in the equation of participation an exclusion variable which is the number of active member in the family. This variable determines, indeed, the participation of the individual in training but has no effect on the wages received after his participation. When the number of people who work in a family increases, the individual is not obliged to find a job in the immediate and can gain access to a training program that the active members of the family can finance. Let us note, however, that theoretically, in the case of a selection model with normal disturbances, it is possible to identify the model parameters without strictly having to use a relationship of exclusion. This need is somewhat masked in the case of normality by the nonlinearity of the functional form adopted. However, in practice the introduction of such relationship is often necessary. It assures that the impact of the program is not only identified by the assumption that we have made on the joint distribution of disturbances and therefore makes the estimators more robust.

Concerning the wage equation, the variable of interest used corresponds to the logarithm of monthly net salary recorded for each individual to the job he occupies at the time of the survey.

Before presenting the results, it is advisable to note that for estimating the model we have conducted a decomposition of "Cholesky" on the covariance matrix of disturbances ($\Omega$) . This decomposition is necessary in practice, it guarantees that the variances are positive and correlation coefficients $\rho_i \in [-1, 1]$[4].

Table 2 gives the results of the simultaneous estimation of all three equations of the model. The estimated correlation coefficients of disturbances of the three equations are also presented in this table. These values can confirm the absence or existence of the phenomena of selection on unobservables in the model.

We present in what follows briefly the results of estimating the equation of participation and of insertion before stating those relating to the wage equation, especially the impact of training on wages, object of our empirical study.

Regarding the participation equation, we note that several variables have a significant impact on the individual decision to participate on training programs in Tunisia. This decision is favored when the number of active member in the family is greater than or equal to three. Instead, the relatively low education levels, not exceeding the secondary, acts negatively on the decision. The oldest individuals, who reside in urban areas and whose fathers are less educated are less involved

[3]Observations with missing data were deleted.

[4]The details of this decomposition are reported in Appendix.

in the programs. Some situations of the head act positively on the participation decision. This is the case when the head is inactive or deceased.

Regarding the insertion equation we can see that only two variables have a significant impact on insertion. It is the education level and the occupation of the household head. Individuals with little education and whose head is unemployed are less likely to find employment. Other variables such as sex, age or place of residence have no impact on insertion.

Several variables have a significant effect on salary. These include age, sex and place of residence. The older people and those living in a big city or a small town have on average higher salaries than other individuals. The coefficients associated with these variables are indeed positive and very significant. Regarding gender, we find that being a female acts negatively on the wages received. Women have on average lower wages than their male counterparts.

**Table 2: Simultaneous estimation (Training - Insertion - Wage)**

| **Explained variable: Participation in training program** | | |
|---|---|---|
| Number of active member in the family (Réf : None) | | |
| One | 0.274 | (0.199) |
| Two | 0.392* | (0.202) |
| Three or more | 0.559*** | (0.209) |
| Age | -0.430*** | (0.083) |
| Women | -0.143 | (0.094) |
| Educational level (Réf : Higher) | | |
| None | -0.262 | (0.561) |
| (1-6)th year of basic school | -0.589*** | (0.131) |
| (7-9)th year of basic school | -1.010*** | (0.227) |
| Secondary | -0.559*** | (0.142) |
| Father's education (Réf : Higher) | | |
| Illitrate | -0.573** | (0.224) |
| Primary | -0.603*** | (0.223) |
| Secondary | -0.499** | (0.208) |
| Résidence (Réf : Outside communes) | | |
| Big city | -0.809*** | (0.209) |
| Small/ medium town | -0.402* | (0.209) |
| Head's occupation (Réf : Unemployed) | | |
| Executive | 0.188 | (0.281) |
| Senior technicien | 0.372 | (0.294) |
| Qualified worker | 0.060 | (0.232) |
| Worker | 0.106 | (0.218) |
| Inactive | 0.609*** | (0.237) |
| Other (died) | 0.431** | (0.201) |
| Constant | 8.793*** | (1.328) |

**Table 2 (following): Simultaneous estimation (Training - Insertion - Wage)**

| **Explained variable: Insertion** | | |
|---|---|---|
| Women | -0.138 | (0.101) |
| Age | 0.128 | (0.109) |
| Educational level (Réf : Higher) | | |
| None | -0.265 | (0.487) |
| (1-6)th year of basic school | -0.295** | (0.131) |
| (7-9)th year of basic school | 0.598** | (0.243) |
| Secondary | -0.251* | (0.149) |
| Résidence (Réf : Outside communes) | | |
| Big city | -0.094 | (0.206) |
| Small/ medium town | -0.163 | (0.207) |
| Head's occupation (Réf : Other (died)) | | |
| Executive | -0.021 | (0.184) |
| Senior technicien | -0.475** | (0.234) |
| Qualified worker | -0.431*** | (0.145) |
| Worker | -0.135 | (0.123) |
| Unemployed | -0.419** | (0.180) |
| Inactive | 0.051 | (0.149) |
| Constant | -1.507 | (1.607) |
| **Explained variable : log (Wage)** | | |
| Age | 0.334*** | (0.060) |
| $Age^2$ | -0.005*** | (0.001) |
| Women | -0.239*** | (0.049) |
| Educational level (Réf : Higher) | | |
| None | -0.131 | (0.252) |
| (1-6)th year of basic school | -0.074 | (0.073) |
| (7-9)th year of basic school | 0.146 | (0.105) |
| Secondary | -0.096 | (0.078) |
| Résidence (Réf : Outside communes) | | |
| Big city | 0.261*** | (0.100) |
| Small/ medium towns | 0.237** | (0.101) |
| Constant | -0.628 | (0.943) |
| $\alpha_y$ 0.606*** (0.113) | | |
| $\rho_{12}$ -0.059 (0.156) | | |
| $\rho_1$ -0.230 (0.162) | | |
| $\rho_2$ 0.978*** (0.070) | | |
| $\sigma$ 0.580*** (0.025) | | |
| Number of observations : 944 | | |
| Log-likelihood : -1371.716 | | |

Notes : (***) significant at level of 1% , (**) significant at level of 5% , (*) significant at level of 10%.
Standard deviations are given in brackets.

The most important result of our study concerns the impact of training on wages. This impact $(\alpha_y)$ is positive and highly significant. Thus, the individuals who have received training receive on average higher wages than those who have not benefited. Going through a training program significantly increases their monthly salary.

Furthermore, the results show the partial absence of selection bias in the data. The estimated value of $\rho_1$ is negative but not significant. This means that unobserved characteristics facilitating the participation in training are not correlated with unobserved characteristics affecting wages. Similarly, the assumption of correlation between the error term of the participation equation and the the error term of the equation of insertion is rejected ($\rho_{12}$ is negative but not significant).

In conclusion, we can say that for our data, the selection mechanisms depend on the observable characteristics of the individual and that the unobserved heterogeneity does not play an important role.

## 5. Conclusion

The aim of our work is to estimate the impact of vocational programs offered in Tunisia on the wages of individuals. For this, we use a simultaneous equation model, considering both the training participation, the insertion into the labor market and the wage, where the variable corresponding to the participation is regarded as endogenous in the wage equation. The results obtained using a sample of 944 individuals, show:

(a) The participation of individuals in training programs in Tunisia significantly increases their wages. These programs are beneficial to participants.

(b) The selection in the programs depends on the observable characteristics of the individual rather than on the unobservable heterogeneity.

The results obtained in this study must, however, be confirmed and deepened. We couldn't, in fact, study the impact of training on insertion. Our data are not rich enough to afford it. Besides, because of the sample size, we aggregate the different types of programs into one, which does not allow us to judge the relative effectiveness of each. This can be studied through an analysis in terms of multiple treatments.

Moreover, other questions may be raised, particularly regarding the impact of the programs on welfare, whose estimation requires the study of the distribution of wages among the beneficiaries. The study of the distribution can answer other as important issues as the average impact that we have considered in this work, such as the proportion of individuals who have benefited from the participation or the categories who have benefited most.

## APPENDICES

A 1: Conditional distribution of disturbances

We define $\varepsilon_1,\ \varepsilon_2$ and $\nu$ three random variables such that:

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \nu \end{pmatrix} \rightsquigarrow N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 1 & \rho_{12} & \rho_1\sigma \\ \rho_{12} & 1 & \rho_2\sigma \\ \rho_1\sigma & \rho_2\sigma & \sigma^2 \end{pmatrix}\right] \qquad (A)$$

We define $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ the vector of disturbances $\varepsilon_1$ and $\varepsilon_2$.

$$\varepsilon \rightsquigarrow N(\mu_\varepsilon, \Omega_{\varepsilon\varepsilon});$$

where $\mu_\varepsilon = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\Omega_{\varepsilon\varepsilon} = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}$.

from another side:

$$\nu \rightsquigarrow N(\mu_\nu, \Omega_{\nu\nu})$$

where $\mu_\nu = 0$ and $\Omega_{\nu\nu} = \sigma^2$.

Furthermore, we define the following matrices:

$$\Omega_{\varepsilon\nu} = \begin{pmatrix} \sigma\rho_1 \\ \sigma\rho_2 \end{pmatrix} \quad \text{et} \quad \Omega_{\nu\varepsilon} = \Omega'_{\varepsilon\nu} = \begin{pmatrix} \sigma\rho_1 & \sigma\rho_2 \end{pmatrix}.$$

The conditional distribution of the vector $\varepsilon$ knowing $\nu$ is a normal (Greene (2005) p.845) as:

$$\varepsilon \mid \nu \rightsquigarrow N(\mu_{\varepsilon.\nu}, \Omega_{\varepsilon\varepsilon.\nu});$$

where $\mu_{\varepsilon.\nu} = \mu_\varepsilon + \Omega_{\varepsilon\nu}\Omega_{\nu\nu}^{-1}(\nu - \mu_\nu)$;
$\Omega_{\varepsilon\varepsilon.\nu} = \Omega_{\varepsilon\varepsilon} - \Omega_{\varepsilon\nu}\Omega_{\nu\nu}^{-1}\Omega_{\nu\varepsilon}$

- Determination of the conditional expectation $\mu_{\varepsilon.\nu}$:

$$\mu_{\varepsilon.\nu} = \mu_\varepsilon + \Omega_{\varepsilon\nu}\Omega_{\nu\nu}^{-1}(\nu - \mu_\nu).$$

under (A),

$$\begin{aligned} \mu_{\varepsilon.\nu} &= \Omega_{\varepsilon\nu}\Omega_{\nu\nu}^{-1}\nu; \\ &= \frac{1}{\sigma^2}\begin{pmatrix} \sigma\rho_1 \\ \sigma\rho_2 \end{pmatrix}\nu \\ \Rightarrow \mu_{\varepsilon.\nu} &= \begin{pmatrix} \frac{\nu\rho_1}{\sigma} \\ \frac{\nu\rho_2}{\sigma} \end{pmatrix} \end{aligned}$$

- Determination of conditional covariance matrix $\Omega_{\varepsilon\varepsilon.\nu}$:

$$\begin{aligned}
\Omega_{\varepsilon\varepsilon.\nu} &= \Omega_{\varepsilon\varepsilon} - \Omega_{\varepsilon\nu}\Omega_{\nu\nu}^{-1}\Omega_{\nu\varepsilon} \\
&= \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix} - \begin{pmatrix} \sigma\rho_1 \\ \sigma\rho_2 \end{pmatrix} \frac{1}{\sigma^2} \begin{pmatrix} \sigma\rho_1 & \sigma\rho_2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix} - \frac{1}{\sigma^2} \begin{pmatrix} \sigma^2\rho_1^2 & \sigma^2\rho_1\rho_2 \\ \sigma^2\rho_1\rho_2 & \sigma^2\rho_2^2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix} - \begin{pmatrix} \rho_1^2 & \rho_1\rho_2 \\ \rho_1\rho_2 & \rho_2^2 \end{pmatrix} \\
\Rightarrow \Omega_{\varepsilon\varepsilon.\nu} &= \begin{pmatrix} 1-\rho_1^2 & \rho_{12}-\rho_1\rho_2 \\ \rho_{12}-\rho_1\rho_2 & 1-\rho_2^2 \end{pmatrix}
\end{aligned}$$

Finally,

$$(\varepsilon_1, \varepsilon_2) \,|\, \nu \quad \rightsquigarrow \quad N\left[\begin{pmatrix} \frac{\nu\rho_1}{\sigma} \\ \frac{\nu\rho_2}{\sigma} \end{pmatrix}, \begin{pmatrix} 1-\rho_1^2 & \rho_{12}-\rho_1\rho_2 \\ \rho_{12}-\rho_1\rho_2 & 1-\rho_2^2 \end{pmatrix}\right].$$

A 2: Decomposition of "Cholesky"
We seek to estimate the covariance matrix of disturbances $\Omega$ which in our model, takes the following form:

$$\Omega = \begin{pmatrix} 1 & \rho_{12} & \rho_1\sigma \\ \rho_{12} & 1 & \rho_2\sigma \\ \rho_1\sigma & \rho_2\sigma & \sigma^2 \end{pmatrix} \qquad (B)$$

The decomposition of "Cholesky" is to determine the different values of $\Omega$ from the values of a triangular matrix $A$ as $\Omega = AA'$ in order to ensure that the matrix $\Omega$ is positive definite. Matrix $A$ can be written as follows:

$$A = \begin{pmatrix} a_1 & 0 & 0 \\ a_2 & a_3 & 0 \\ a_4 & a_5 & a_6 \end{pmatrix}$$

$\Omega = AA'$ then written as follows:

$$\begin{aligned}
\Omega &= \begin{pmatrix} a_1 & 0 & 0 \\ a_2 & a_3 & 0 \\ a_4 & a_5 & a_6 \end{pmatrix} \times \begin{pmatrix} a_1 & a_2 & a_4 \\ 0 & a_3 & a_5 \\ 0 & 0 & a_6 \end{pmatrix} \\
\Omega &= \begin{pmatrix} a_1^2 & a_1a_2 & a_1a_4 \\ a_1a_2 & a_2^2+a_3^2 & a_2a_4+a_3a_5 \\ a_1a_4 & a_2a_4+a_3a_5 & a_4^2+a_5^2+a_6^2 \end{pmatrix} \qquad (C)
\end{aligned}$$

By identifying (B) to (C), we obtain:
$a_1^2 = 1 \Rightarrow a_1 = 1$
$a_1a_2 = \rho_{12} \Rightarrow a_2 = \rho_{12}$
$a_1a_4 = \rho_1\sigma \Rightarrow a_4 = \rho_1\sigma$

$$\alpha_4^2+\alpha_5^2+\alpha_6^2=\sigma^2 \Rightarrow \sigma=\sqrt{\alpha_4^2+\alpha_5^2+\alpha_6^2}$$

$$\alpha_4=\rho_1\sigma \Rightarrow \rho_1=\frac{\alpha_4}{\sigma} \Rightarrow \rho_1=\frac{\alpha_4}{\sqrt{\alpha_4^2+\alpha_5^2+\alpha_6^2}}$$

$$\rho_2\sigma=\alpha_2\alpha_4+\alpha_3\alpha_5 \Rightarrow \rho_2=\frac{\alpha_2\alpha_4+\alpha_3\alpha_5}{\sqrt{\alpha_4^2+\alpha_5^2+\alpha_6^2}}.$$

## REFERENCES


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

Bonnal, L., Fougère, D. and Sérandon, A. (1997), "Evaluating the impact of French employment policies on individual labour market histories", *Review of Economic Studies*, 64, pp.683-713.

Card, D. (1993): "Using Geographic Variation in College Proximity to Estimate the Returns to Schooling", *Working Paper*, n°4483, NBER.

Dehejia, R. and Wahba, S. (1999), "Causal Effects in Non-Experimental Studies: Re-evaluating the Evaluation of Training Programs", *Journal of the American Statistical Association*, 94(448), pp.1053-1062.

Dehejia, R. and Wahba, S. (2002), "Propensity Score Matching Methods for Non Experimental Causal Studies", *Review of Economics and Statistics*, 84, pp.151-161.

Fougère, D., Goux, D. and Maurin, E. (2001), "Formation continue et carrières salariales: une évaluation sur données individuelles", *Annales d'Economie et de Statistique*, 62, pp. 49-70.

Greene, W. (2005), "Econométrie", édition Française, 5ème édition, Pearson Education, France.

Heckman, J. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", *Annals of Economic and Social Measurement*, 5(4), pp.475-492.

Heckman, J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System", *Econometrica*, 46(4), pp.931-959.

Heckman, J. and Hotz, J. (1989), "Alternative Methods for Evaluating the Impact of Training Programs", *Journal of the American Statistical Association*, 84(408), pp.862-880.

Heckman, J., Ichimura, H. and Todd, P. (1997), "Matching As an Econometric

Evaluation Estimator: Evidence on its Performance Applied to the JTPA Program, Part I. Theory and Methods", *Review of Economic Studies*, 64(4), pp.605-654.

Heckman, J., Ichimura, H. and Todd, P. (1998), "Matching As an Econometric Evaluation Estimator: Evidence on its Performance Applied to the JTPA Program, Part II. Empirical Evidence", *Review of Economic Studies*, 65(2), pp.261-294.

Heckman, J. and Smith, J. (1998), "Evaluating the Welfare State", *Frisch Centenary Econometric Monograph Series*, Cambridge University Press.

Heckman, J. and Vytlacil, E. (2000), "Local Instrumental Variables", *Working Paper*, n° T0252, NBER.

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

Lee, L. (1978), "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables", *International Economic Review*, 19, pp. 415-433.

Maddala, G.S. (1983), "Limited-Dependent and Qualitative Variables in Econometrics", *Econometric Society Monographs (3)*, Cambridge University Press.