THE IMPACT OF TUNISIAN VOCATIONAL TRAINING PROGRAMS ON EMPLOYMENT AND WAGE

RIHAB BELLAKHAL*

MOHAMED-BADRANE MAHJOUB**

FEBRUARY 2013

ABSTRACT. The aim of this paper is to estimate the impact of Vocational Training Programs offered in Tunisia on employment and wages of individuals. The data we use come from a study carried out in Tunisia in 2001 by the Ministry of Vocational Training and Employment on the graduates of the national vocational training. The estimated model corresponds to three simultaneous equations determining the participation in training, the insertion in the labor market and the wages observed. The two residuals of the treatment and selection equations (training and employment) are assumed to be correlated with that of the outcome equation (wage). The results show that professional training in Tunisia improves employability and increases potential wages.

JEL CLASSIFICATION: C31, J18.

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

*Tunis Graduate School of Business, Campus Universitaire de La Manouba 2010, La Manouba, Tunisie. Phone : 0021671600615 - 0021698546019, Fax: 0021671601311. Email: bellakhal_rihab@yahoo.fr.
**University of Paris1 Pantheon-Sorbonne, CES, 106, boulevard de Hopital, 75647 Paris cedex 13, France. Email: mb.mahjoub@gmail.com.

1. INTRODUCTION

In this paper we evaluate the impact of Vocational Training Programs offered in Tunisia on employment and wages of individuals. During the last decades, Tunisia established many employment and vocational training programs. While the establishment of these programs is not recent, the objective and the potential results assigned to these policies are today different. These programs have to play an important role against the increase of unemployment rate by providing qualification allowing young people to find a job and to have a salary. Insertion to the labor market and wage are then, two important objectives of these policies and the main criteria of their efficiency.

Tunisian employment and vocational training programs were developed gradually over several years and have undergone several reforms since the 1990s, following especially international commitments signed by Tunisia. Two types of training programs are offered in Tunisia; in-service training which is generally offered for graduates of higher education who are already employed, and initial vocational training which is received by young people after dropping out of the general education system. Our paper focuses on the initial vocational training.

Despite the development and the diversity of the vocational training policies in Tunisia, evaluation studies of these policies are unavailable. To the best of our knowledge, there is no serious attempt to evaluate 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 carried out in other countries such as in US and France. Hence the contribution of our paper.

Several studies on the evaluation of public policies, especially those of training and employment have been conducted in these last years. 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 caused by 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 carefully take into account this problem.

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 place of residence, education level, age,...) and unobservable ones (such as willingness to work, individual ability,...). Then, the assignment of individuals to the program is by self-selection and not by random assignment. From an econometric point of view, this corresponds to a problem of endogeneity of the variable of interest (training) in the outcome equations that we want to study (employment and wage here) (see Heckman (1978) and Heckman and Hotz (1989)).

In the literature, several methods have been proposed to deal with this problem of selection bias. Rubin (1977) and Rosenbaum and Rubin (1983) propose in this context the matching method. Dehejia and Wahba (1999, 2002) and Heckman, Ichimura and Todd (1997, 1998) use this method to evaluate American training programs. Note, however, that this non parametric method takes into account only the phenomena of selection on observable. 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 participation in 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 by making parametric assumptions on the 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 the approach 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 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. 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 their job at the time of the study. The information has been collected from two surveys. A first survey carried out among the main beneficiaries of vocational training (treatment group), and a second survey carried out on a sample of non-beneficiaries (control group).

Using the information contained in these surveys, we estimate the impact of programs on the employment and the wages of individuals. As we mentioned earlier, the approach we use is parametric. It is based on modeling simultaneously the participation decision and the outcome variables, by specifying a 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 a job we 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 (called also treatment equation) is the first selection equation and the insertion into the labor market is the second. It is comparable with 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 Maximum Likelihood Estimator. The likelihood function of our model depends on the conditional densities of the disturbances of the two selection equations with respect to the error term of the wage equation. The estimation results show that the participation of individuals in training programs significantly increases their probability to find a job and their monthly wage. The results also show that unobserved factors of participation in training were correlated with unobserved characteristics affecting employment and wage.

In Section 2 we present the data and some descriptive statistics. In section 3, we present the model and derive its likelihood function. In section 4 we discuss the estimation results, and in section 5 we conclude.

2. The Data

The data used in this work are from the survey of Vocational Training Program graduates conducted by the Ministry of Vocational Training and Employment in Tunisia in 2001. The survey covered a sample of 499 individuals graduating from the different vocational training programs in 1998[1] and a sample of 503 individuals serving as a control group.

The group of graduates was interviewed 36 months after leaving the training (i.e., in 2001). The questionnaire for this survey was designed to collect the detailed information on the individual characteristics and their professional situation at the time of the survey, especially in terms of employability and income. The information on family characteristics is also provided in this survey, describing the family and social context of the individual at the moment of his enrollment in the training program.

The sample of control group was selected from a list of job seekers registered in the files of the Tunisian Agency for Employment in 1996 or 1997. It includes students who leaved the regular education system before 1998. They did not enter a training program in spite of their eligibility. Investigators checked also that none of this group has participated in another training program or benefited from employment assistance in order to avoid contamination bias. These individuals were interviewed at the same time as the treatment group and have answered the same questionnaire. Thus, a number of information mainly concerning their individual

[1]Most common Vocational Training Cycles in Tunisia are Vocational Middle Education, Vocational High Education and Vocational College. These three levels are respectively attested by CAP, BTP and BTS degrees.

characteristics, their employment status and their occupational integration in the labor market have been collected.

In order to ensure comparability between the two groups, we exclude from the data base individuals who left school before 1980, whose age is upper than 40 and whose mensual wage is less than 40. These observations are considered to be outliers. Observations with missing data were also deleted. Our final sample includes 880 individuals whose 462 of them benefited from the vocational training programs.

Table 1 shows the main characteristics of the sample of graduates compared to those of non-participants. We see in this table that 58% of individuals who receive training have found a job versus 42% for those who don't receive it. Participants have also a higher average wage than non-participants. Men participate more in the training programs than women. Participants are younger than control group individuals. Regarding the educational level[2], there is no big differences between the two groups: Individuals with high school level of education are the majority in the two groups (respectively 49% and 43% for treatment and control groups), followed by those with primary school level (21% versus 22%). As for the year of leaving school, 56% of participants left school between 1990 and 1995 and only 4% before 1990, which is different for non-participants (respectively 38% and 35%). We also see that participants come from larger family than non-participants. Regarding the fathers's occupation, the table does not show any obvious difference between the characteristics of control and treatment groups (Individuals whose father is inactive or dead are the majority in the two groups). As for the residence area, it seems that the proportion of individuals living in big cities is slightly higher in the control group. The table shows also that most of the individuals are employed in the Services Sector in the two groups, compared with the agriculture and manufacturing sectors. Generally speaking, we can not exclude the assumption of similarity between the two groups, and then our identification strategy based fundamentally on the comparison between them can be considered as valid.

[2]Before college, the General or Regular Education System in Tunisia is organised as follows: 9 years of Basic Education (Including 6 years in Primary School and 3 years in Middle-High School) and 4 years of High School Education.

TABLE 1: DESCRIPTIVE STATISTICS
(Treatment Group vs Control Group)

	Treatment Group		Control Group	
	Mean	Std.D	Mean	Std.D
Employed	0.584	0.493	0.423	0.494
Wage	281.663	151.811	250.090	108.854
Man	0.623	0.485	0.557	0.497
Age	25.119	3.282	28.200	4.080
Educational Level				
None	0.010	0.103	0.009	0.097
Primary School	0.216	0.412	0.224	0.418
Middle-High School	0.036	0.188	0.071	0.258
Two years of High School	0.138	0.345	0.153	0.360
Four years of High School	0.495	0.500	0.430	0.495
College or More	0.101	0.302	0.110	0.313
Year of Leaving School				
Before 1990	0.045	0.208	0.358	0.480
Between 1990 and 1995	0.560	0.496	0.380	0.486
After 1995	0.393	0.489	0.260	0.439
Family Size				
Under than 6	0.374	0.484	0.511	0.500
Between 6 and 8	0.512	0.500	0.421	0.494
More than 8	0.112	0.316	0.066	0.250
Number of active members in the family	2.196	1.365	1.782	1.190
Head's Occupation				
Inactive, Other (dead)	0.538	0.499	0.502	0.500
Unemployed	0.051	0.222	0.081	0.273
Blue Collar	0.173	0.378	0.172	0.378
White Collar	0.110	0.313	0.136	0.343
Middle Manager, Technician	0.056	0.230	0.038	0.192
Executive, Lawyer, Doctor, Engineer	0.069	0.254	0.069	0.254
Residence				
Big City	0.452	0.498	0.645	0.478
Small or Medium City	0.452	0.498	0.315	0.465
Rural Area	0.095	0.293	0.038	0.192
Industry				
Agriculture	0.044	0.206	0.045	0.208
Manufacturing	0.318	0.466	0.293	0.456
Services	0.637	0.481	0.661	0.474

3. THE MODEL

For each individual i in the sample, we observe simultaneously three variables. Denote by D_i a dummy variable equal 1 if the individual i has participated in the training program and 0 otherwise; E_i a dummy variable equal 1 if the individual i has found a job and 0 otherwise; 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 set of exogenous variables that may explain the participation in training; X_{2i} is the set of exogenous factors that may explain the employment and X_{3i} are exogenous variables that determine the wage. Note that X_{2i} includes the treatment variable D. Otherwise, some exogenous variables such as age, sex and educational level can belong to X_1, X_2, and X_3. These variables are important determinants of participation, employment and wage.

To estimate this model, we assume that the vector of disturbances $(\varepsilon_{1i}, \varepsilon_{2i}, \nu_i)$ follows a trivariate normal disturbances with mean zero and covariance matrix Ω 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 ρ_{12} is the correlation coefficient between ε_1 and ε_2; ρ_1 is the correlation coefficient between ε_1 and ν; ρ_2 is the correlation coefficient between ε_2 and ν; and σ^2 is the variance of ν.

The estimation method we use is that of maximum likelihood. The likelihood function is based on the joint density of the perturbations $(\varepsilon_{1i}, \varepsilon_{2i})$ conditional on the error term ν_i. 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[3]:

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

[3] See the proof in Appendix.

Let $\mu_1^* = \dfrac{\rho_1}{\sigma}\nu_i$ be the the conditional expectation of $\varepsilon_{1i} \mid \nu_i$; $\mu_2^* = \dfrac{\rho_2}{\sigma}\nu_i$ the conditional expectation of $\varepsilon_{2i} \mid \nu_i$; $\sigma_1^* = \sqrt{1-\rho_1^2}$ the conditional standard deviation of $\varepsilon_{1i} \mid \nu_i$; $\sigma_2^* = \sqrt{1-\rho_2^2}$ the conditional standard deviation of $\varepsilon_{2i} \mid \nu_i$; and $\rho_{12}^* = \dfrac{\rho_{12}-\rho_1\rho_2}{\sqrt{(1-\rho_1^2)(1-\rho_2^2)}}$ the correlation coefficient of ε_{1i} and ε_{2i} conditionally on ν_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. Four situations can occur depending on the values taken by the three endogenous variables (D_i, E_i and Y_i).

- *Did not participate in training, Is not employed:* $D_i = 0$; $E_i = 0$; Y_i *not observed*
  $\mathcal{L}_i = \mathrm{Prob}(D_i = 0,\ E_i = 0)$
  $= \mathrm{Prob}\big(\varepsilon_{1i} \leqslant -X_{1i}\beta_1,\ \varepsilon_{2i} \leqslant -X_{2i}\beta_2)$
  $= \Phi_2(-X_{1i}\beta_1,\ -X_{2i}\beta_2,\ \rho_{12})$
  We set: $A_i = -X_{1i}\beta_1$; $B_i = -X_{2i}\beta_2$; So that $\Rightarrow \mathcal{L}_i = \Phi_2(A_i,\ B_i,\ \rho_{12})$; where Φ_2 is the distribution function of bivariate normal distribution.
- *Participated in training, Is not employed:* $D_i = 1$; $E_i = 0$; Y_i *not observed*
  $\mathcal{L}_i = \mathrm{Prob}(D_i = 1,\ E_i = 0)$
  $= \mathrm{Prob}(\varepsilon_{1i} > -X_{1i}\beta_1,\ \varepsilon_{2i} \leqslant -X_{2i}\beta_2)$
  $= \Phi_2(-A_i,\ B_i,\ -\rho_{12})$
- *Did not participate in training, Is employed:* $D_i = 0$; $E_i = 1$; Y_i *observed*
  $\mathcal{L}_i = \mathrm{Prob}(D_i = 0,\ E_i = 1,\ Y_i = y_i)$
  $= \mathrm{Prob}(D_i = 0,\ E_i = 1/Y_i = y_i) \times \mathrm{Prob}(Y_i = y_i)$
  $= \mathrm{Prob}(\varepsilon_{1i} \leqslant -X_{1i}\beta_1,\ \varepsilon_{2i} > -X_{2i}\beta_2/Y_i = y_i) \times \mathrm{Prob}(Y_i = y_i)$
  $= \dfrac{1}{\sigma}\phi(\dfrac{C_i}{\sigma}) \times \Phi_2(F_i,\ -G_i,\ -\rho_{12}^*)$;

  where

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

  $$F_i = \frac{(-X_{1i}\beta_1 - \mu_1^*)}{\sigma_1^*} = \frac{(-X_{1i}\beta_1 - \rho_1 C_i/\sigma)}{\sqrt{1-\rho_1^2}};$$

  $$G_i = \frac{(-X_{2i}\beta_2 - \mu_2^*)}{\sigma_2^*} = \frac{(-X_{2i}\beta_2 - \rho_2 C_i/\sigma)}{\sqrt{1-\rho_2^2}};$$

  and ϕ is the density function of the standard normal distribution.
- *Participated in training, Is employed:* $D_i = 1$; $E_i = 1$; Y_i *observed*
  $\mathcal{L}_i = \mathrm{Prob}(D_i = 1,\ E_i = 1,\ Y_i = y_i)$
  $= \mathrm{Prob}(D_i = 1,\ E_i = 1/Y_i = y_i) \times \mathrm{Prob}(Y_i = y_i)$
  $= \dfrac{1}{\sigma}\phi(\dfrac{C_i}{\sigma}) \times \Phi_2(-F_i,\ -G_i,\ \rho_{12}^*)$.

4. RESULTS

We estimate the model presented above using our sample of 880 individuals. The exogenous variables introduced in the equations are age, sex, level of education, place of residence, head of the household's occupation and sector of activity. Our data do not allow us to introduce variables on the characteristics of firm. Concerning the outcome equation, the wage used corresponds to the logarithm of monthly net salary recorded for each individual in the job he holds at the time of the survey.

In addition to theses exogenous variables, we introduce in the equation of participation two instrumental variables to identify the impact of training programs on employment and wage. These variables are the number of active members in the family and the year of leaving school. These variables are supposed to determine the participation of the individual in training without having any direct effect on his employment or wage, every thing being equal. In fact, a member of the family already in the job market gives the individual a better chance to participate in a training without monetary constraints. This working member not only can be in charge of living expenses of the family, but also he/she can finance the training of our participant individual. That's why we consider that more we have active members in the family, less the individual will support pressure to enter job market right after dropping out from school. The second instrument is the year of leaving school. This variable is considered as an important determinant of participation in training programs. In fact those who left school before 1990 were less likely to participate in vocational program because of a limited supply. As we mentioned in the introduction, Tunisian vocational training programs have undergone several reforms since the 1990s, so we consider that before 1990 training programs and training centers were not well developed especially in medium cities. Now, for these two variables to be considered as valid instruments, we should check that they do not have any direct impact on employment and wage. From a pure statistical point of view, the data show negligible or non significant correlation *intra* in each one of the couples (Employment, Number of Active members in the Family), (Wage, Number of Active members in the Family), (Employment, Year of Leaving School), (Wage, Year of Leaving School). From an economically intuitive point of view, we do not believe that the year of leaving school or number of active members in the family can have any obvious relation with productivity. Employment and wage are the consequences of observed traits such as education and training and unobserved ones such as ability and motivation, rather than by demographic aspects such as age or family composition.

In addition to these two instruments in the treatment equation, the employment equation contains an exclusion variable facilitating the identification of the selection mechanism. We choose here the variable Family Size as an exclusion restriction. In fact, when the family is large, individuals have more chance to find a job owing to connections that may have the family members. Wage obtained in the labor market remains always a consequence of the individual productivity and training. We 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 an exclusion restriction. The need for such restriction is mitigated in the case of normality by the nonlinearity of the functional form adopted. However, in practice the introduction of such relationship is often preferable. It ensures that identification of the policy parameter does not depend only on the distributional assumption made, making hence the estimator more robust.

Before presenting the results, we precise that for estimating the model we have conducted a decomposition of "Cholesky" on the covariance matrix of disturbances (Ω) . This decomposition is necessary in practice to guarantee that the variances are positive and the correlation coefficients ρ_i are in the range $[-1, 1]$[4].

Table 2 gives the results of the simultaneous estimation of the three equations and the correlation matrix. The first part of the table 2 gives the estimates of the treatment equation parameters. The coefficients of our instrumental variables are significant with the expected signs. The number of active members in the family increases significantly the probability of enrolment in a vocational training program. And those who left school before 1990 have less chance to be enrolled in a training program. Younger people have more chance to participate, without any significant gender gap. Other results show that blue and white collar parents do not encourage their children to go training and that, comparatively to rural area, living in a big city decreases the probability of participating in training programs.

Regarding the employment equation (second part of table 2) we can see that the coefficient of participation variable is highly significant and positif which means that individuals who participate in training programs are more likely to find a job than those who didn't. This result on the treatment effect of vocational training on employment is unique in the Tunisian context. It needs to be confirmed and reinforced by other studies on other data sets to provide a strong recommendation in terms of vocational training public policy. Concerning the exclusion variable parameter, we see that people belonging to small or medium family (size under 8) have more chance to integrate the labor market. Large family is generally a family living in rural area with low level of parent's education and with difficulty to find a job. Other results show that the probability of employment increases with age in big cities and for individuals with middle-high school level of education.

The third part of the table 2 gives the parameter estimates of the wage equation. As we can see, the salary has the classical concave function of age, and the usual gender gap in favor of men. Individuals who have college degree or more, who live in small or medium cities and whose father is middle manager or technician have on average higher salaries than other individuals. The most important result of our study concerns the impact of training on wage. This sign of α_y is positive and significant. Thus, the individuals who participated in training have on average higher wages than those who did not participate, controlling for socio-demographic characteristics and taking into account the selection bias.

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

TABLE 2: ESTIMATION RESULTS

Dependent variable: Training		
Number of active members in the family	0.215***	(0.030)
Year of Leaving School (Ref : After 1995)		
Before 1990	-0.827***	(0.151)
Between 1990 and 1995	-0.042	(0.086)
Age	-0.111***	(0.014)
Man	0.088	(0.093)
Residence (Ref : Rural Area)		
Big City	-0.797***	(0.210)
Small or Medium City	-0.372*	(0.212)
Head's occupation (Ref : Inactive, Other (dead))		
Unemployed	-0.276	(0.193)
Blue Collar	-0.360***	(0.131)
White Collar	-0.476***	(0.145)
Middle Manager, Technician	0.041	(0.210)
Executive, Lawyer, Doctor, Engineer	0.070	(0.189)
Constant	3.407***	(0.435)
Dependent variable: Employment		
Training (D)	1.694***	(0.092)
Family Size (Ref : More than 8)		
Under than 6	0.212*	(0.124)
Between 6 and 8	0.255**	(0.121)
Age	0.081***	(0.012)
Man	-0.011	(0.082)
Residence (Ref : Rural Area)		
Big City	0.351**	(0.170)
Small or Medium City	0.044	(0.170)
Educational level (Ref : None)		
Primary School	0.294	(0.381)
Middle-High School	0.770*	(0.406)
Two years of High School	0.202	(0.382)
Four years of High School	0.247	(0.375)
College or More	0.266	(0.388)
Head's occupation (Ref : Inactive, Other (dead))		
Unemployed	-0.063	(0.169)
Blue Collar	0.052	(0.112)
White Collar	-0.013	(0.130)
Middle Manager, Technician	-0.176	(0.192)
Executive, Lawyer, Doctor, Engineer	-0.041	(0.164)
Constant	-3.742***	(0.539)

TABLE 2 (CONTINUED): ESTIMATION RESULTS

Dependent variable : Wage		
Age	0.344***	(0.065)
Age Squared	-0.005***	(0.001)
Man	0.229***	(0.043)
Educational level (Ref : None)		
Primary School	0.150	(0.249)
Middle-High School	0.210	(0.258)
Two years of High School	0.055	(0.250)
Four years of High School	0.142	(0.245)
College or More	0.450*	(0.252)
Residence (Ref : Rural Area)		
Big City	0.161	(0.098)
Small or Medium City	0.155*	(0.089)
Head's occupation (Ref : Inactive, Other (dead))		
Unemployed	-0.043	(0.097)
Blue Collar	-0.030	(0.060)
White Collar	0.093	(0.070)
Middle Manager, Technician	0.257***	(0.099)
Executive, Lawyer, Doctor, Engineer	0.124	(0.081)
Industry (Ref : Agriculture and Manufacturing)		
Services	-0.028	(0.042)
Constant	-0.675	(1.075)

α_y	0.552***	(0.171)
ρ_{12}	-0.898***	(0.071)
ρ_1	-0.480***	(0.168)
ρ_2	0.551***	(0.179)
σ	0.457***	(0.047)
Number of observations :	880	
Log-likelihood :	-1269.037	
Wald chi2 (12) :	208.470	
Prob > chi2 :	0.000	

Notes : (***) 1% significance level, (**) 5% significance level , (*)10% significance level. Standard errors are in brackets.

The last part of the results table gives the estimated parameters of the components of Ω. The correlation coefficient ρ_{12} between the residuals of participation and employment equations is negative and significant and very high in absolute value. It says that the unobserved determinants of training are correlated with those of employment. The negative sign of this parameter is not in contradiction with the positive sign of the coefficient of D in the employment equation. Individuals go to a vocational training program when they think that their chance to find a job without any training is small. It can be the case of students without any family networks facilitating employment or students who failed in job interviews because of their limited skills, or because anything that is not observable to us. Vocational training is then a remedial treatment initiated by the individual himself. The correlation coefficients between the wage equation residual and the training residual has also a negative sign, perhaps for the same reasons discussed above, and especially reasons concerning individual ability and skills. As for the correlation coefficient between employment and wage, it reveals the importance of the selection process in our model and confirms the need to take it into account, otherwise our policy parameter α_y would be biased. The positive and significant sign of ρ_2 says that unobserved factors increasing the probability of employment are positively correlated with those increasing productivity. Here we think that it concerns motivation and the desire to succeed. Although all the individual and family obstacles, the individual perseveres to get a job and works hard to improve his wage. The estimation of the components of Ω is then very informative in terms of unobserved features of our individuals. A linear estimation, without correlations and selection control would not give us such information. Hence the main motivation of a model with double selectivity.

5. Conclusion

This work estimated the impact of vocational training programs offered in Tunisia on employment and wage of individuals. For this purpose we use a simultaneous equations model for training, employment and wage. The variable corresponding to participation in the treatment is considered as endogenous and the variable corresponding to employment considered as the dependent variable of the selection equation. The basic result obtained using a sample of 880 individuals is that vocational training in Tunisia has a positive treatment effect on the probability of employment and on wage. Our estimation results show also that assignment to the programs depends on the observable and the unobservable characteristics of the individual.

The results obtained in this study must, however, be confirmed and deepened. Because of the sample size, we aggregate the different branches of vocational programs into one, which does not allow us to evaluate the relative effectiveness of each one. 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 important issues than the marginal impact that we have considered in this work, such as the proportion of individuals who have benefited from the participation or the categories of individuals who have more benefited than others.

In Tunisia, before the Arab spring, there were no willingness to evaluate the public policies. The surveys were conducted by official administrations with the help and complicity of international organisations like IMF and World Bank. The data were not available to researchers, but were aimed to praise the regime. We believe that the situation is changing. During the last two years, more and more data are leaving the *safes* of the ministries providing researchers a vast field of applied econometric studies. We hope that our work could improve in the future by taking advantage of richer data sets (providing better instruments for example), in order to propose a framework for the evaluation of vocational training policies in Tunisia.

APPENDICES

A 1: Conditional distribution of disturbances
We define $\varepsilon_1,\ \varepsilon_2$ and ν 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 ε_1 and ε_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 ε knowing ν 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,

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

A 2: Decomposition of "Cholesky"

We seek to estimate the covariance matrix of disturbances Ω which in our model, takes the following form:

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

The decomposition of "Cholesky" is to determine the different values of Ω from the values of a triangular matrix A as $\Omega = AA'$ in order to ensure that the matrix Ω 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:

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

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

By identifying (B) to (C), we obtain:

$a_1^2 = 1 \Rightarrow a_1 = 1$

$a_1a_2 = \rho_{12} \Rightarrow a_2 = \rho_{12}$

$a_1a_4 = \rho_1\sigma \Rightarrow a_4 = \rho_1\sigma$

$$a_4^2 + a_5^2 + a_6^2 = \sigma^2 \Rightarrow \sigma = \sqrt{a_4^2 + a_5^2 + a_6^2}$$

$$a_4 = \rho_1 \sigma \Rightarrow \rho_1 = \frac{a_4}{\sigma} \Rightarrow \rho_1 = \frac{a_4}{\sqrt{a_4^2 + a_5^2 + a_6^2}}$$

$$\rho_2 \sigma = a_2 a_4 + a_3 a_5 \Rightarrow \rho_2 = \frac{a_2 a_4 + a_3 a_5}{\sqrt{a_4^2 + a_5^2 + a_6^2}}.$$

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.

Rosenbaum, P. and Rubin, D. (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70(1), pp.41-55.

Rubin, D. (1977), "Assignment to a Treatment Group on the Basis of a Covariate", *Journal of Educational Statistics*, 2, pp. 1-26.