

A CORPUS LINGUISTIC ANALYSIS ON VOCATIONAL TRAINING

Lecturer PhD. Madalina Ecaterina POPESCU¹

Abstract

The topic of impact evaluation of vocational training on the labour market has recently become extremely popular, especially as the European Commission has increased its concern upon conducting counterfactual impact evaluation to all EU-funded programs. Although most studies aiming to evaluate the impact of vocational training on employability focus predominantly on quantitative measures and totally neglect other research perspectives, the novelty of this paper consists in conducting a corpus linguistic analysis on the topic of vocational training, in order to highlight some particularities of the literature review in the field. The corpus linguistic analysis will thus focus on three relevant aspects, regarding general concepts on one hand, the methodology applied to quantify the impact of vocational training, as well as the main findings of the empirical research in the field.

Keywords: *vocational training, corpus, linguistic analysis, impact evaluation*

JEL Classification: J24, I21

1. INTRODUCTION

Recently, the topic of impact evaluation of Active Labour Policy Measures upon employability has become extremely popular, especially as the European Commission has increased its concern on counterfactual impact evaluation of all EU-funded programs. Among the Active Labour Policy Measures, vocational training plays a major role, as there is still uncertainty upon its real effects on employability. During the last decade the empirical research in the field has yielded quite controversial results. Since training programs increase human capital, positive effects on both income and employability are to be expected.

Although most studies aiming to evaluate the impact of vocational training focused predominantly on quantitative measures and totally neglect other research perspectives, the novelty of this paper consists in conducting a corpus linguistic analysis on the topic of vocational training, in order to highlight some particularities of the literature review in the field.

2. THE CORPUS LINGUISTIC ANALYSIS

The corpus consisted of 10 articles on the topic of vocational training, summing up 4975 types and 74875 tokens. The ANTCOnc software was used in order to conduct the corpus linguistic analysis. When considering the rank of the words most frequently listed in the corpus, after excluding the connectors and considering just the most content-relevant words in the corpus, the following most relevant terms were identified in the corpus: *training* (1287 times), *employment* (573 times), *impact* (552 times), *program* (469 times), *group* (444 times), *treatment* (299 times), *data* (283 times), *control* (233 times), *participants* (230 times), *sample* (228 times), *job* (226 times), *market* (225 times), *results* (203 times), *earnings* (201 times), *education* (167 times), *unemployment* (163 times), *programs* (161 times), *comparison*

¹Dep. Informatics and Economic Cybernetics, Academy of Economic Studies and Researcher, The National Scientific Research Institute for Labour and Social Protection, madalina.andreica@gmail.com

(156 times), *income* (146 times), *positive (effects)* (139 times), *evaluation* (137 times), *skills* (137 times) and *matching* (135 times).

The corpus linguistic analysis then focused on the following three main aspects:

- General concepts used in the analysis
- The methodology applied to quantify the impact of vocational training
- The main findings of the empirical research

Rank	Freq	Word
1	5860	the
2	2966	of
3	2307	in
4	2149	and
5	1655	to
6	1287	training
7	1094	a
8	1007	for
9	921	on
10	827	is
11	785	that
12	596	are
13	573	employment
14	552	impact
15	536	as
16	512	x
17	469	program
18	464	by
19	459	with
20	444	n

Fig. 2.1. The rank of the words most frequently listed in the corpus

Regarding the way the general concepts on vocational training are perceived and used in the corpus, the following print-screens suggest a clear and common meaning of the concept of training. First, the concordance list on the word ‘*training*’ was checked.

Hit	KWIC	File
1	receiving complete expression and communication training. However, the pros and cons of	Dynamic Eff
2	to 90% in Month 18 after receiving the training for a period of time. With	Dynamic Eff
3	also reduced the cost of personnel training as well as enhanced corporate industrial	Dynamic Eff
4	, which increased the cost of personnel training as well as further obstructed corporate	Dynamic Eff
5	between schools and enterprises or vocational training institutions enables students to learn b	Dynamic Eff
6	The Role of Training and Skills Development in Active Labour	Role of Train
7	\x92s creation of employer-led Training and Enterprise Councils in 1989 (the mai	Role of Train
8	overlap, namely in the provision of training and skills development to participants i	Role of Train
9	measures; active measures include labour market training, job creation measures, support for activ	Role of Train
10	on the supply side, measures include: 1. training schemes: these are the classic elements	Role of Train
11	as Jobcentre Plus. The Role of Training and Skills Development in Active Labour	Role of Train
12	-Saxon nations. 3\x85 and how do training measures fit in? We now turn	Role of Train
13	now turn to the role of training as part of the portfolio of (Role of Train
14	relatively small share of GDP on training and skills-related active measures (Figur	Role of Train
15	creasing extent sanctions for non- participation. Training measures for the unemployed and inactive,	Role of Train
16	would seem that the importance of training measures has been decreasing in the	Role of Train
17	7 shows that in 2005 the expenditure on training-related active measures fell below 0.1 p	Role of Train
18	ntury. Similarly, Figure 8 indicates that whereas training measures accounted for around 45 per cen	Role of Train
19	cent by 2005. This declining emphasis on training within active measures is not confined	Role of Train
20	in the share accounted for by training measures (from 31 to 24 per cent). De	Role of Train

Fig. 2.2. The concordance list on the word ‘*training*’

The main content-relevant collocates that were identified when searching for the collocates of the family-word of *training* by considering a 4 word window-frame were: “*impact of training*”, “*vocational training*”, “*training programs*”, “*employment and training*”, “*job training*”, as well as “*occupational training*”.

One interesting concordance found in just 2 of the 10 articles, but with a rather high frequency of 18 times consisted in the structure “*training vouchers*”. According to Lauringson et al. (2011) a personalised training voucher could become a quite flexible tool for matching the training as precisely as possible to the individual requirements of the job-seeker.

Rank	Total No. of Collocates	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	464	445	19	3.17513	of	
2	378	210	168	1.90820	the	
3	259	101	158	2.85917	in	
4	179	22	157	3.50272	on	
5	169	60	109	2.16457	and	
6	167	152	15	4.14739	impact	
7	110	63	47	1.95748	to	
8	89	87	2	5.34795	vocational	
9	70	63	7	4.45460	after	
10	69	1	68	4.62319	programs	
11	69	14	55	2.77129	employment	
12	68	53	15	2.34313	that	
13	68	24	44	1.94845	for	
14	66	4	62	2.22138	is	
15	58	22	36	1.60478	a	
16	54	51	3	3.80692	job	
17	53	51	2	5.71749	occupational	
18	51	1	50	2.66812	program	
19	49	42	7	3.38364	who	
20	43	1	42	3.16405	was	

Fig. 2.3. Collocates of the word “training”

The idea consists in the fact that a training voucher normally allows the unemployed to choose a course organized by a training provider approved by the Unemployment Insurance Fund in order to get specific training in areas where the demand for training is low.

Concordance Hits 18

Search Term: training vouchers

Search Window Size: 50

Level 1: 1R, Level 2: 2R, Level 3: 3R

File: 8_Impact_Ev

an integrated sequence; the use of training vouchers; and significant changes in gove
 Services Department or by using personalised training vouchers. The training manager uses the
 formed from the job-seekers. Personalised training vouchers were introduced as a parallel
 training is not very high. The training vouchers can be used to choose
 was available on the basis of training vouchers, but from 2011 they can be
 up to EUR 2 500 worth of personalised training vouchers (until 31 July 2011 the ceiling
 had just started operating trainings, the training vouchers had not been introduced and
 10. These constitute trainings arranged after the training vouchers were launched and the new
 applicable (however, at that time the training vouchers could not yet be used
 April and October 2010 constituted trainings with training vouchers8. In both periods studied the
 impact of occupational training with personalised training vouchers and procured training (in 2010)
 organised on the basis of personalised training vouchers. Thus it is possible to
 impact of training based on personalised training vouchers and procured training (the same
 income of training based on personalised training vouchers and procured training finished i
 employment in the group that used training vouchers is almost 12% higher than in
 trainings except those based on personalised training vouchers, i.e. also professional examin
 organised on the basis of personalised training vouchers in addition to procured trainin
 employment in the group that used training vouchers was almost 12% higher than in

Fig. 2.4. The concordance list for “training vouchers”

Regarding the main methods applied in order to evaluate the vocational training impact on employability, the following results were obtained, as presented in figure 2.5.

Concordance Hits 44

Search Term: method

Search Window Size: 44

Level 1: 1R, Level 2: 2R, Level 3: 3R

File: 8_Impact_Ev

scales, which are standard and proven method-ologies in psychology to measure personali
 raised about the choice of the method to be used for evaluating this
 used for evaluating this policy. This method must enable us to identify the
 (1983) propose in this context the matching method. Dehejia and Wahba (1999, 2002) and Heckma
 Heckman, Ichimura and Todd (1997, 1998) use this method to evaluate American training programs. No
 . Note, however, that this non parametric method takes into account only the phenomena
 variables to correct this problem. This method was subsequently used in several studies (
 (2000)). However, the difficulty in using this method lies in choosing the appropriate instru-
 the wages of workers. The estimation method we use is Maximum Likelihood Estimator.
 ?2 is the variance of ?. The estimation method we use is that of maximum
 by using propensity score matching, a method widely used in similar evaluations. Persons
 measures. The choice of the evaluation method depends primarily on the objective of
 approach is to find a suitable method for using non-experimental data. In
 conclusion that there is no universal method that would suit each and every
 before-after estimator, difference-in-differences method, the method of instrumental variables and
 estimator, difference-in-differences method, the method of instrumental variables and the selectio
 instrumental variables and the selection model method belong to the second group. The
 policies. This analysis uses the matching method for evaluating the impact of training.
 (Caliendo and Kopeinig 2008): Although the PSM method is widely used in policy evaluation
 policy. In this analysis the PSM method is implemented by using Stata modules

Fig. 2.5. The concordance list for the word “method”

A brief selection of the main aspects debated on the methods required to conduct the analysis is presented next:

”This analysis uses the matching *method* for evaluating the impact of training. As mentioned above, matching has several advantages over regression analysis and can be used when selection over observables is assumed and a rich dataset is available. For matching on a higher number of observable characteristics, matching on some balancing scores tends to be more used than cell matching. Propensity score matching as a balancing score is particularly popular in evaluating labour market measures.”

”Based on the treatment of selection bias the evaluation methods fall into two broad groups: 1) methods assuming that selection is based on observable characteristics; 2) methods assuming that selection is based on both observable and unobservable characteristics” (Caliendo, 2006).

“The more popular methods in the first group are matching and linear regression analysis. Methods like the before-after estimator, difference-in-differences *method*, the *method* of instrumental variables and the selection model *method* belong to the second group. The assumption that selection is based on observables means that selection to treatment is assumed to be determined by observable characteristics, but the selection to treatment does not depend on outcomes in the absence of treatment. Selection based on un-observables means that unobservable characteristics are also used to determine selection to treatment” (Smith, 2004).

Regarding the results of the empirical studies, most findings suggested that trainings have a rather positive impact on employability, although delays in the process of job finding are also to be expected. On the short run, however, the effects on participants’ outcomes on the labour market might turn out to be negative (especially for a one year delay).

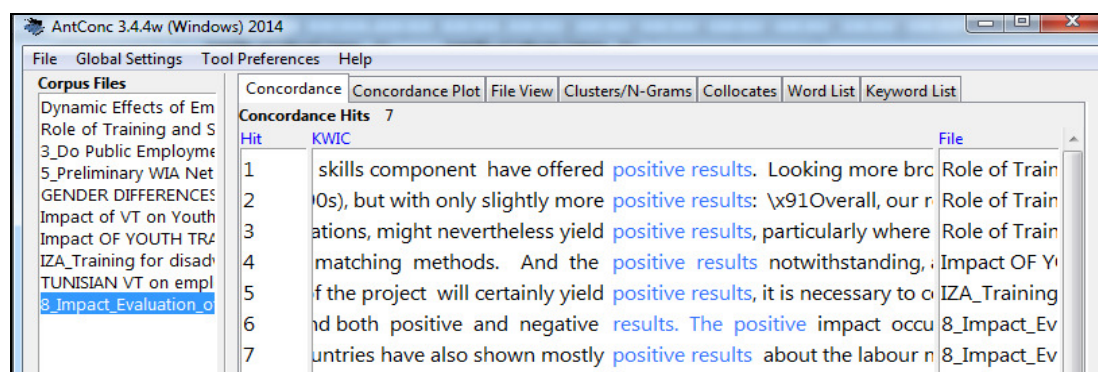


Fig. 2.6. The concordance list for “positive results”

However, over the past decade a number of long-term follow-up studies have been conducted and indicated that impacts followed over a sufficiently long time period after the training can become quite strongly positive. These findings are also easily confirmed by the frequency of the word “positive” of about 139 times, as compared to the term “negative” that only appears one time in the corpus.

4. CONCLUSIONS

Based on the corpus used in the linguistic analysis, we can state that the topic of impact evaluation of vocational training upon the labour market has recently become extremely popular. Although most studies aiming to evaluate the impact of vocational training on employability focused predominantly on quantitative measures and totally neglect other research perspectives, the novelty of this paper consisted in conducting a corpus linguistic analysis on the topic of vocational training, in order to highlight some particularities of the literature review in the field.

The corpus linguistic analysis focused on three relevant aspects, regarding general concepts on one hand, the methodology applied to quantify the impact of vocational training, as well as the main findings of the empirical research in the field.

To conclude, the main findings of the corpus linguistic analysis on the topic of vocational training suggest a common understanding of the general concepts used in the field, a use of mostly quantitative analysis based on matching techniques, along with a general tendency towards concluding that vocational trainings have a rather positive effect on employability, especially on the long run.

5. REFERENCES

- Heinrich, C.J., Mueser, P.R., Troske, K.R., Jeon K.S., Kahvecioglu, D.C., (2013) *Do Public Employment and Training Programs Work?*, Heinrich et al. IZA Journal of Labor Economics, 2:6, <http://www.izajole.com/content/2/1/6>
- Hollenbeck, K, King, C.T., Schroeder, D., (2003) *Preliminary WIA Net Impact Estimates: Administrative Records Opportunities and Limitations*, New Tools for a New Era! Symposium, Bureau of Labor Statistics & the Workforce Information Council, Embassy Row Hilton Hotel Washington, D.C.
- Chen, P.C (2014) *Research on Dynamic Effects of Employability of Vocational College Students in Taiwan*, International Journal of Computer Science & Information Technology (IJCSIT) Vol 6, No 1, DOI : 10.5121/ijcsit.2014.6113 161
- Meager N, (2009) *The Role of Training and Skills Development in Active Labour Market Policies*, International Journal of Training and Development 13:1, Wiley-Blackwell
- Lauringson, A., Villsaar, K., Tammik, L., Luhavee, T., (2011) *IMPACT EVALUATION OF LABOUR MARKET TRAINING*, Estonian Unemployment Insurance Fund
- Cho,Y., Kalomba, D., Mobarak, A.M., Orozco, V., Wolfson, D., (2015) *GENDER DIFFERENCES IN THE EFFECTS OF VOCATIONAL TRAINING: CONSTRAINTS ON WOMEN AND DROPOUT BEHAVIOR*
- Batchuluun, A., Batbekh, S., Dalkhjav, B., Sanjmyatav,A., Baldandorj, T.E., (2013) *The Impacts of Vocational Training Programs on Youth Employment in Mongolia*, RESEARCH PROPOSAL Presented to Partnership for Economic Policy (PEP)
- Card, D., Ibarraran, P., Regalia, F., Rosas, D., Soares, Y. (2007), *The labor market impacts of youth training in the Dominican Republic: Evidence from a randomized evaluation*, NBER Working Paper Series, Working Paper 12883, <http://www.nber.org/papers/w12883>
- Ibarraran, P., Ripani, L., Taboada, B., Villa, J.M., Garcia, B. (2014), *Life skills, employability and training for disadvantaged youth: Evidence from a randomized evaluation design*, IZA Journal of Labor& Development, 3:10, <http://www.izajold.com/content/3/1/10>
- Bellakhal, R., Mahjoub, M.B. (2013) *The impact of Tunisian vocational training programs on employment and wage*, *Proceedings of the Labor Markets and Labor Policy in MENA Countries*, Tunisia, 2013, http://www.iza.org/conference_files/mena_2013