Evaluation of training providers for a new governance of the training system for the unemployed: a proposal for Italian Regions

Elena Cappellini*
Silvia Duranti*
Valentina Patacchini*
Carla Rampichini**
Nicola Sciclone*

* *Irpet* * * *University of Florence*

Introduction

In the last years, the trend towards activation has been one of the major issues in welfare and labour market reforms in Europe. Italy has lacked for a long time a strong net of activation policies for the unemployed, but the latest reforms (Fornero law in 2012, Jobs Act in 2014) have placed great emphasis on the need to invest in Public Employment Services to make labour market more inclusive. At the same time the European Union, through the European Social Fund, has made available the necessary financial resources to promote active policies, especially those related to training, crucial for the development of human resources in a knowledge-based economy.

The increasing focus on activation policies and in particular on training also raises the issue of better evaluation. Over recent years there has been an increasing amount of research on the effectiveness of labour market training programmes in many countries (Oecd, 2005; Kluve et al. 2007), but this field of research is still developing in Italy (Gori et al., 1993, Regione Toscana, Giunta Regionale, Osservatorio Regionale del mercato del lavoro (1997), Paccagnella, 2006, Maitino et al. 2011), where the lack of the relevant data often prevents evaluation studies. Still less attention has been given to the analysis of the effectiveness of different institutional frameworks for the provision of training (Gori, 2000).

However, the governance structure for activation and training services is a major determinant of success (Oecd, 2005). Being that of training a natural quasi-market, evaluation[1] is the key factor for shaping an efficient system. For this reason, this paper proposes for the Italian context a governance framework inspired to the principles of regulation of providers and efficiency of resources allocation.

In Italy training is competence of the Regions, which, within certain standards defined at the national level[2], have defined models characterised by different levels of planning centralisation and training providers

[1] It is important to stress that we speak here of relative evaluation. The objective is not to evaluate the absolute effectiveness of training programmes or training providers, but the relative performance of each provider compared to the others.

[2] For example, the basic principles of accreditation, inspired to the European framework of reference for quality assurance of education and training (EQARF), are defined at the national level in order to provide some indicators that guarantee on the quality of providers of education and training. In particular, the Italian system is based on five criteria that provide specific instructions to regional authorities, competent in training issues, to define the specific use of accreditation:

A - Resources and logistics infrastructure;
B - Reliability economic and financial;
C - Capacity management and professional resources;
D - Effectiveness and efficiency;
E - Relations with the territory.

regulation. The proposed model could therefore be applied to any Regional system, although it is here simulated for the context of an Italian Region, Tuscany.

The current Tuscan model is based on a bottom up planning of training supply, since the content of tenders is mainly based on the needs expressed by social parties and gathered at the Regional or Provincial level. As regards regulation of training providers, the current model is based on accreditation rules where the performance of agencies plays no role. In particular, necessary and sufficient conditions to obtain accreditation is the possession of a minimum requirement, based on four criteria: organizational and administrative structure, logistics, relations, performance. The latter however, does not consider any indicator of job placement, being made up of 3 indicators concerning: rate of drop out, rate of trained, efficiency of spending. Once obtained accreditation, penalties affect training providers only in the event of formal errors and deficiencies and the revocation or suspension of accreditation only occurs when minimum requirements relating to formal criteria are not respected; any real role is attributed to indicators related to effectiveness.

A different model is the one implemented in Lombardy, especially in what concerns regulation. Indeed, the role of demand in shaping supply is maximum, because a system of vouchers (Dote), guarantees to individuals freedom of choice between the alternatives autonomously proposed by training providers, which therefore have incentives to adapt to demand. In this context regulation is extremely relevant. For this reason, the regulation of providers is based on a system of Rating, in which employment performance of providers are essential for the allocation of resources. In particular, the rating model is configured as a further development of the accreditation system and aims to measure the relative positioning of each training provider with respect to employment performance of trainees. So, the model is not aimed at identifying agencies to be accredited or accredited providers that do not meet the required conditions: under investigation in the rating are the only agencies that have achieved accreditation to provide training services.

The model of Lombardy is a good benchmark for the improvement of the effectiveness of the Tuscan training system. However, the latter would be less market-oriented and based on a top-down planning of training supply, developed through calls for tenders for training courses. The idea is to establish a central governance of the content of supply (which courses to finance and how), leaving to the market the provision. In order to provide incentives for an effective training system, the current model of selection of providers is a good starting point, but it requires upgrading to make an assessment that goes beyond mere respect of formal criteria. Indeed, it is necessary to associate to formal selection a more substantial rating mechanism, which, as in the experience of Lombardy, gives more importance to the performance of providers. Being one of the primary goals of training (in particular training for the unemployed) the increase in employability, the evaluation of performance of a provider should be based on its capacity to train people and get them back to work. As in the Lombardy experience, the rating model is configured as a further development of the accreditation system (which would continue to exist selecting agencies on the base of formal or financial criteria), aimed at promoting an allocation of resources which reflect performance and stimulates an overall improvement of the system[3].

The rest of the paper is organised as follows. Section 1 will describe the necessary data to implement a governance framework for the training system which is based on the reward of employment outcomes of providers. In section 2 a new governance framework for the Italian regional training system is designed while the subsequent sections describe in detail the different steps to implement it. In particular, Section 3 describe how to create a system of profiling, Section 4 how to rank providers according to their adjusted performance and Section 5 simulates the allocation of resources under the new system.

[3] Lam (2007) reports for Australia an improvement of 15 percentage points in the average outcome rate of Employment Services after only 18 months of implementation of the “Star Rating System”.

1. Sources of data

The primary source of data for evaluating the training system is the database of the European Social Fund, which represents the main source of funding for the regional training system. From this we have extracted the training courses financed by Tuscany in the programming period 2007-2013, selecting only those for the unemployed[4]. We thus consider a population of 337 providers, 2,466 courses and 21,908 unemployed trainees having terminated the course by the end of 2013. The dataset contains information on the trainees, specifically concerning sex, age, nationality, education level, previous work experience and duration of unemployment. Information on courses are also very rich and concern duration, thematic content, class size, cost to public finance, data of beginning and end. Unfortunately, the dataset contains only the name and the VAT identification number of training providers. The original data set has thus been enriched with the information contained in Irpet's database on firms: employees, income, type of provider (firm, no profit).

Moreover, we use the Compulsory Communications System of administrative data on employment dynamics, which record all the activations, transformations, fixed-term extensions and anticipated terminations of employment relationships between any worker and employer since the beginning of 2008.

Through a merge of labour market administrative data with the dataset on trainees, courses and providers it is possible to check the employment outcomes of all unemployed people having attended a training activity. A limitation of the joined data set is the lack of information on self-employment: placement rates are therefore net of activation as self-employed. The choice to use only administrative data comes from the need of developing standard and replicable procedures in a new framework of governance. Another advantage of administrative data instead of ad hoc surveys for the analyses of employment outcomes is the accuracy of information: indeed, SIL contains information on all jobs following the end of the course, their timing and length and many other characteristics (sector of activity, contract and qualification).

2. A proposal for a new governance framework of the training system

The management of the training system first requires the establishment of an adequate model of selection of training providers, able to select the most successful ones to effectively distribute resources.

The accreditation process is a good starting point in this process, but for the sake of the system's effectiveness and efficiency, this first selection needs to be complemented by a more substantial system of rating, based on the results of funded activities. The rating model is a development of the accreditation system and aims to measure the relative positioning of each training provider in a performance ranking net of external factors. Therefore, the model is not aimed at identifying sites to be accredited or accredited providers that do not meet the required conditions: only the providers that have already achieved accreditation are under investigation.

The rating can be used to select the best performers, to whom a certain amount of resources can be reserved. Since new entrants in the training system cannot be evaluated on the basis of past performances, it is necessary to leave part of total resources to be contended by all providers.

The following is an example of how to allocate the total available resources: 40% to all operators (including outsiders) irrespective of previous re-employment performance, 40% to insiders having had a high or medium performance in the last year and 20% to best performers.

There are several options to rank operators on the basis of their performance. Among the exiting options at the international level[5], the Australian methodology is the nearest to our proposal (Lam,2007, Australian Governament-Disability Employment Services, 2013), since it uses regression-based estimates of performance to rank providers. The aim is to enable fair comparisons between providers, by taking into

[4] Moreover, the database has been cleaned from courses specifically addressed to disabled, identified by those having a percentage of disabled higher than 50% of total trainees. For this courses employability is only one of the goals, having as primary aims social inclusion and acquisition of basic skills.

[5] See Ocse, 2005 for a more general review on Public Employment Service evaluation.

account of factors outside their control. Among these factors there is certainly the type of user. For this reason, such a governance model requires a profiling of trainees in order to allow an evaluation of training providers' performance net of characteristics of users, thus preventing creaming phenomena that might arise in similar contexts.

In summary, a model of governance based on the selectivity and transparency in allocating resources between providers should therefore be based on three steps.

First, the creation of system of profiling that summarizes the level of disadvantage in the labour market of each trainee. Such a tool is necessary to prevent creaming or cherry peaking phenomena; a proposal for this is described in section 3, where a statistical approach is used.

Second, the creation of a system of rating of providers based on their net performance in terms of employment outcomes; an example of how this can be made is provided in section 4 where an econometric approach has been adopted to obtain a ranking of training providers.

Third, the allocation of resources based on the reward of best performers in terms of employment outcomes; a simulation of this procedure for the programming period 2007-2013 is proposed in section 5, where the new hypothetical distribution of resources is compared to the actual one.

3. Profiling system

To summarize the characteristics of training users we have created a tool for profiling users based on their distance from the labour market. Indeed, the methodology of profiling is based on the idea that employability is predictable given unemployed characteristics and, in particular, the length of unemployment. The goal is to identify the characteristics that best define the profiles of people furthest away from the labour market and therefore harder to be reallocated and summarize them in a single indicator (O'Connell et al. 2009; Rudolph and Konle-Seidl, 2005; Lam 2007; Hasluck, 2004).

The methodology used is rather similar to the one adopted in Lombardy and described in Arifil (2013).The profiling score has been attributed on the basis of a linear probability model for the probability to find a job within 12 months since arrival at the Public Employment Service[6].

Table 1 reports the estimated coefficients. The luckiest subject is a young Italian male, with an education level higher than compulsory education, a work experience and a short duration of unemployment (less than 6 months), with a predicted probability of re-employment equal to 69.4%.

The coefficients of covariates represent the marginal effect of each characteristic the probability of re-employment ranges from 10.9% (unluckiest subject, the one having the worse characteristics for re-employment) to 69.4% (luckiest subject, or baseline subject). The estimated difference is thus 59%. Standardising this difference into a range 0-100, where 0 represents minimum risk and 100 maximum risk of remaining unemployed[7], we obtain the profiling scores for each characteristics as shown in Table 2. Summing up the scores, we calculate a total profiling score for every unemployed according to his distance from the labour market.

[6] The population on which the logit regression has been estimated is represented by all people signed up to unemployment list by Public Employment Services in the period 2008-2012.

[7] The profiling score has been standardized as: 100*(score-MIN)/MAX.

Table 1.
Results of the linear probability model

	Coef.
Woman	-0,023***
30-40 years old	-0,047***
40-50 years old	-0,068***
More than 50 years old	-0,221***
Only compulsory education	-0,059***
Foreigner	-0,038***
No work experience	-0,173***
Duration of unemployment: 6-12 months	-0,078***
Duration of unemployment: 12-24 months	-0,133***
Duration of unemployment: more than 24 months	-0,243***
Constant	0,694***

Note The baseline subject (i.e. a subject with all the covariates equal to zero) represents the luckiest unemployed, having characteristics favourable to a fast re-employment: he is male, young, Italian, with an education level higher than compulsory education, with work experience and a short duration of unemployment.

Table 2.
Profiling scores

SEX	Man	0
	Woman	4
AGE	Less than 30 years old	0
	30-40 years old	8
	40-50 years old	12
	More than 50 years old	38
EDUCATION	Only compulsory education	10
	Secondary or tertiary education	0
NATIONALITY	Italian	0
	Foreigner	7
PREVIOUS WORK EXPERIENCE	With work experience	0
	No work experience	30
DURATION OF UNEMPLOY-MENT	Duration of unemployment: less than 6 months	0
	Duration of unemployment: 6-12 months	13
	Duration of unemployment: 12-24 months	23
	Duration of unemployment: more than 24 months	42

The total score is then calculated for each trainee recorded in the database for the period 2007-2013.

Table 3 reports the trainees classified into 4 groups, on the basis of the estimated profile score, reflecting employability, where 1 represents the lowest risk to be unemployed after 12 months and 4 the highest.

Table 3.
Training users, by unemployment risk group

Risk group	Score range	Percentage distribution of users	Average score
Low risk group	0-16	25,2	9
Medium to low risk group	18-33	25,3	25
Medium to high risk group	31-46	25,3	39
High risk group	47+	24,3	61
TOTAL		100%	33

4. A system of rating of training providers

4.1 Theoretical framework and methodology

To rank training providers, the approach of the school-effectiveness research (SER) was used. In particular, the goal is to rank training providers according to their Type B effectiveness (Raudenbush and Willms (1995), that is the actual performance of providers, net of factors that are exogenous and then not controllable by schools themselves. These studies usually make use of multilevel methodologies to take into account the hierarchical nature of data and to exploit second level residuals to measure school effectiveness (Raudenbush and Willms, 1995; Goldstein, 1997; Grilli and Rampichini, 2009). Indeed, the second level variance that cannot be explained by control variables (the residual) is considered as the "net effect", of training providers. This approach is not new for the evaluation of effectiveness of courses and training providers in the Italian context. The same has been adopted by Gori (2000) for the Lombardy Region, by Gori et al. (1993) and Regione Toscana, Giunta Regionale, Osservatorio Regionale del mercato del lavoro (1997) for Tuscany, and, more recently, by Paccagnella (2006) for the Autonomous Province of Bolzano.

The measurement of the training providers' performance first requires the identification of one or more measurable dimensions for which the ranking has to be produced (placement rates, time to find the first job, days worked in the year following the course). International and regional experiences offer several examples of different outcome measures of effectiveness upon which to evaluate providers (Lam, 2007; Finn, 2009). In this version of the paper a model has been proposed where the outcome variable is the placement rate in the 12 months following the end of the course.

The methodology used consists-to fit a two-level random intercept logit model (Goldstein, 2003) to properly take into account the hierarchical structure of data, i.e. trainees nested into training providers.

Let Y_{ij} be the binary response, i.e. $Y_{ij} = 1$ if the i-th subject of the j-th agency find a job and zero otherwise, where $i = 1, \cdots, n_j$ denotes the number of trainees (level 1 units) nested within the second level unit (or cluster) j, i.e. the training provider, $j = 1, \cdots, J$, J total number of considered agencies. Given the success probability $\pi_{ij} = P(Y_{ij} = 1 | \mathbf{x}_{ij}, u_j)$, the model is specified as follows:

$$logit(\pi_{ij}) = log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \boldsymbol{\beta}\mathbf{x}_{ij} + u_j$$

where $\mathbf{x}_{ij}$ is the vector of covariates for the i-th subject of the j-th cluster (including a constant term for the intercept) and $\boldsymbol{\beta}$ is the vector of fixed parameters (including the intercept).The residuals u_j are assumed to be

independent and identically distributed across clusters with a normal distribution and common variance σ_u^2.

Model fitting is performed using the melogit procedure of Stata (StataCorp, 2015).

The model allows to consider covariates at both levels. At the trainee level we considered individual characteristics and job status, as reported in Table 4, while as second level explanatory variables only those factors which are believed to affect the outcome but are outside of the control of the organization. In particular, we considered aggregated characteristics of trainees and labour market conditions, proxied by the year and the area where the course has taken place. The characteristics of trainees are included both at the individual and at the agency level.

The characteristics of the course in terms of content (vocational versus non vocational) and length have been included in the model, consistently with a top-down governance framework, where the Region plans the allocation of funding between types of training (for young people or mature ones; vocational or general, structured, light, or intensive), leaving to providers only the detailed planning of single courses. For this reason, providers cannot be penalised for having organized light courses, which notoriously guarantee lower employment outcomes, but which have been planned and financed by the regional government. On the contrary, a more market-oriented system, where the planning and supply of training is fully in charge of providers (which only have the goal of employability, whatever this goal is achieved), would require a rating system where performance is considered gross of course characteristics[8].

Several other characteristics of the providers could be considered in the analysis: type of provider (school, firm, local governments, non profit organization), size in terms of income or employees, degree of courses specialisation, employees, frequency of funding with public resources. However, for the purpose of this analysis they do not need to be controlled for when estimating a net performance: each training provider has as primary goal the increase in the employability of trainees, irrespective of how this goal is pursued.

Our estimation procedure follows four steps: in the first step, we estimate an “empty” model, to decompose the total variance into the trainee-level (within) variance and training provider-level (between) variance and test the relevance of the latter, while in the second, third and fourth steps we add explanatory variables respectively at individual, course and provider level.

The results of model fitting are reported in Table 4. The reported likelihood-ratio test compares the random intercept model to ordinary logistic regression (Berkhof and Snijders, 2001) and is highly significant for these data, confirming an “agency effect”, which can change the employment probability of a trainee from 48% to 75% according to the gross training provider effectiveness (model without covariates).

Once controlled for exogenous variables, providers still appear to perform significantly differently in regards to rate of re-employment. Indeed, in the full model (Column D of Table 4) the likelihood-ratio test is still highly significant, indicating unexplained second level variability. Consistently, the 95% coverage interval of the predicted employment probability for the reference trainee ranges from 75% to 89%, according to the training provider net effectiveness (net of users’, courses’ and labour market characteristics). To give an idea of the effect of unobserved factors at the training provider level Table 5 reports the predicted probability of employment considering three types of providers (an average effective provider, a low effective provider and a highly effective one) and four typical profiles of trainee.

[8] It is the case of Australia, but also of Lombardy.
A similar model has been tested with our data and results differ significantly from the basic model. In particular, more than 20% of providers resulting best performers in the basic model, leave the top of the ranking in the alternative one.

Table 4.
Results of the multilevel regression

	Empty	Individual	Individual+Course+Contest	Individual+Course+Contest+Agency
	A	*B*	*C*	*D*
Less than 25 years		0,014	-0,0223549	-0,0132042
35-45 years		-0,1927859***	-0,1796916***	-0,1823279***
45-55 years		-0,4404111***	-0,4189482***	-0,4275741***
More than 55 years		-1,012609***	-0,9728384***	-0,9862505***
Compulsory education		-0,2815189***	-0,2671207***	-0,2645606***
Tertiary education		-0,0716924*	-0,0885232**	-0,0896434**
Duration of unemployment: 6-12 months		-0,388839***	-0,3939155***	-0,3973648***
Duration of unemployment: 12-24 months		-0,5394318***	-0,5409294***	-0,5469298***
Duration of unemployment: more than 24 months		-1,048683***	-1,051738***	-1,059172***
Foreigner		-0,304549***	-0,2436235***	-0,2589435***
Woman		0,051	0,0458642	0,0419189
No work experience		-0,8753572***	-0,8670168***	-0,8616595***
Disabled		-0,3079433**	-0,2668043*	-0,3038765**
Profiling score		0,0064824***	0,0068313***	0,0073261***
Vocational course			0,0244834	0,025684
Dilute			-0,0462696	-0,0567303
Intensive			0,0027701	0,0059383
Light			-0,2051863***	-0,2144***
Strctured			0,3328132***	0,3443698***
Employment rate in Local Labour System			-0,00000713	-0,0020702
Ended in 2010			0,0986962	0,0972724
Ended in 2011			-0,2006249***	-0,1962192***
Ended in 2012			-0,1936943**	-0,1934334***
Ended in 2013			-0,0798713	-0,0746722
Average profiling score in the agency				-0,0626949
Percentage of disabled in the agency				0,0214126**
Percentage of unemployed without work experience in the agency				-0,0057285**
Percentage of Italians in the agency				-0,0037241**
Percentage of qualified people (secondary or tertiary education) in the agency				0,0009565
Percentage of over55 people in the agency				-0,002842
Percentage of long-term unemployed in the agency				-0,0066193
Probability for the reference person in an average agency (u=0)	***48%***	***63%***	***63%***	***75%***
Probability for the reference person in a high performing agency (u=2*sd)	***23%***	***37%***	***39%***	***55%***
Probability for the reference person in a low performing agency (u=-2*sd)	***75%***	***83%***	***82%***	***89%***
LR test vs. logistic regression: chibar2(01)	1075,97	668,6	475,7400	334,4
Prob>=chibar2	0,0000	0,0000	0,0000	0,0000
ICC	***9,1%***	***7,9%***	***7,0%***	***6,1%***

Note. The reference subject has the following characteristics: male, aged between 25 and 35 years, Italian, with secondary education, with work experience and a short duration of unemployment, thus having a profiling score equal to zero. The reference course has the following characteristics: non vocational, average duration in terms of length and hours, ended in 2009, with no disabled trainees, no long-term unemployed, only foreigners, only trainees with more than compulsory education, only under55 and no trainees looking for their first job.

Table 5.

Predicted re-employment rates for typical profiles of trainees in differently effective training providers. Full model (column D of Table 4)

	Average agency (sd=0)	Low performing agency (u=-2*sd)	High performing agency (u=-2*sd)	*Difference*
Over 55, unemployed from 6 to 12 months, with only compulsory education	34%	17%	56%	39%
Foreign woman, 40 years old, unemployed since less than 6 months, with only compulsory education	67%	44%	84%	39%
28years old man, qith a degree and no work experience, unemployed since 6-12 months	40%	21%	63%	42%
Foreign man, 40 years old, unemployed since 12-24 months, with only compulsory education	34%	17%	57%	40%

Note. Predicted probabilities with observable course and agency characteristics equal to zero, i.e. the base categories of level 2 covariates: non vocational course, with average duration in terms of length and hours, ended in 2009, and a training providers with no disabled, no long-term unemployed, no Italian, no qualified people, no over55, and no people looking for their first job.

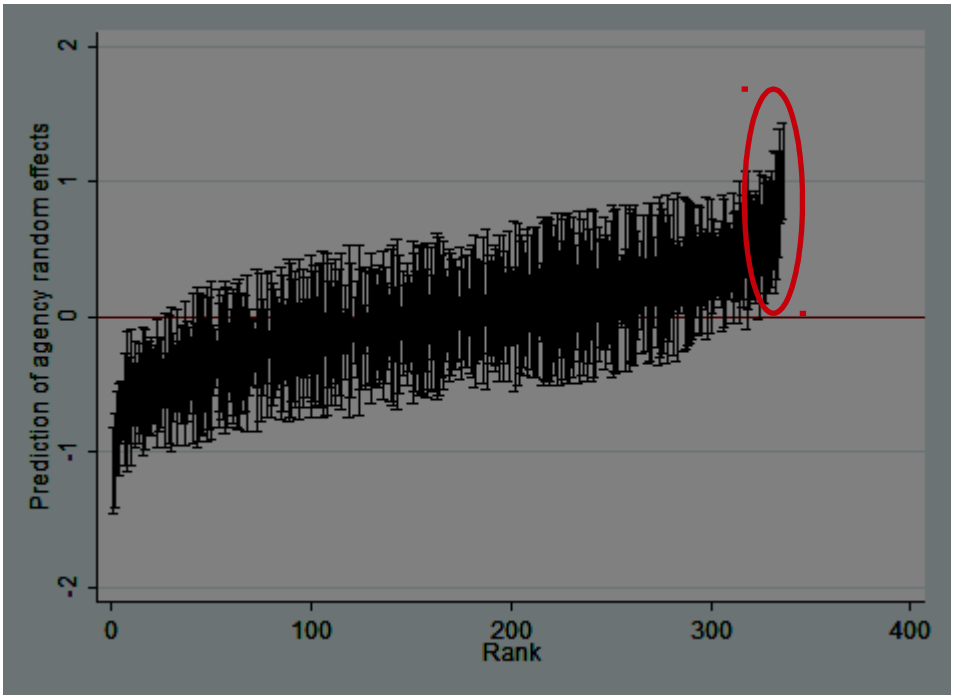
4.2 Ranking of training providers

We ranked Tuscan publicly funded training providers according to their effectiveness, measured by Empirical Bayes (level 2) residuals of the full model (colum D of Table 4). Figure 1 shows the predictions of level 2 residual from the full model alongside with their comparative confidence intervals. The training providers, in the lower left and upper right part of the graph, have a predicted residual significantly different from zero. The latter, are those to be considered the best performers of the Tuscan training system, because they show placement rates significantly higher than those predicted by the model on the basis of observable characteristics of users and courses. These 45 agencies are those to be rewarded by deserving them a quota of public resources.

But who are best performers? Descriptive statistics show that they are mainly firms (instead of schools, nonprofit organizations, local governments), whose core activity concerns training. This information is not trivial, since around half of training providers have not as core business training itself. Looking at public resources in the observation period, best performers do not seem to have received on average more than other providers, and this is consistent with the lack of any kind of performance evaluation in the current model. An interesting result of our analysis concerns the composition of the users; indeed, for the best training providers the rate of drop-out is slightly lower (14% versus 18% on average): rewarded agencies are not only able to re-employ people, but also to limit drop out. A final remark on the composition of users in best performing agencies: here problematic users are slightly under-represented (this agencies have a lower percentage of trainees in the high risk group and a lower average profiling score), indicating that even controlling for individual and composition variables, differences in users' are not fully accounted for[9].

[9] This is not surprising and reflects the ongoing debates about how well the profiling instrument captures a client's real level of disadvantage (Lam, 2007).

Fig. 1.
Ranking of training providers according to second level residuals from multilevel regression



5. Allocation of resources

To simulate the impact of such a governance framework of training providers on the allocation of resources between insiders, the ranking has been used to select best performers as those showing significantly higher placement rates than those predicted by the model.

The simulation of the new allocation of resources has been done relying on the assumption of complete elasticity of supply: the more resources are available, the more agencies use them.

Table 6 shows the allocation of public resources applying the rating system. Only 40% of total resources (about 27million euros) are contestable between all training providers (including new entrants); 40% of resources (about 27million euros) are contestable only between those agencies having proved more performing than the average (agency residual higher than zero), while the remaining 20% (about 14million euros) is reserved to best performers (agency residual significantly higher than zero). This means that each provider theoretically disposes of 81,071 euros, which become 175,280 euros (94,210 euros more) for those providers not selected as worst performers and jump to 478,844 (303,564 euros more) for best performers.

Simulating this new allocation on effective agency data for the period 2009-2013, we found that the major changes in the distribution of resources are concentrated on best and worst performers, while the majority of training providers do not change significantly the amount of public resources received. In the new model, financial resources available to best performers totally increase by more than 12millions euros (+89%) compared to the effective distribution in the period 2009-2013. On the contrary, financial resources available to worst performers totally decrease by more than 4millions euros (-25%). Thus, the distribution of resources under the new model is far less homogeneous.

Table 6
Allocation of public resources applying the rating system

	All	Not worst	Best
Total reserved amount	€ 27.320.763	€ 27.320.763	€ 13.660.381
Number of obs.	337	290	45
Individual average amount	€ 81.071	€ 94.210	€ 303.564
Individual cumulated amount	**€ 81.071**	**€ 175.280**	**€ 478.844**

Conclusions

The governance structure for activation and training services is a major determinant of success. For this reason, this paper proposes for the Italian context a governance framework inspired to the principles of regulation of providers and efficiency of resources allocation.

The idea is to establish a central governance of the content of supply (which courses to finance and how), leaving to the market the provision. In order to provide incentives for an effective training system, it is necessary to associate to formal selection a more substantial rating mechanism, which gives more importance to the performance of providers. Since the increase in employability is one of the primary goals of training for the unemployed, the evaluation of performance of a provider should be based on its capacity to train people and get them back to work. Therefore, the proposed methodology is based on a regulation of the market which rewards agencies with better employment performances, in order to stimulate an overall improvement of the system.

The proposed methodology, although based on clear and transparent criteria, leaves room for political choices.

First of all, a different outcome variable can be chosen. In this paper we chose a simple measure (at least a job entry in the 12 months following the course), but it is also possible to consider more specific, or qualitative measures. Possible future developments of this work include ratings based on alternative outcomes, which consider time (e.g. time to find a job or duration of the job found) and qualitative aspects; in particular the type of contract would be an interesting aspect to be considered, possibility through the use of a multinomial multilevel regression.

Secondly, also the inclusion of covariates in the model leaves room for political choices, because it depends on the type of training “market” one wants to develop. In this paper, we assumed a centrally planned training supply, where, therefore, some characteristics of the course are exogenous to the agency, and thus need to be controlled for. On the contrary, a more market-oriented system, where the planning and supply of training is fully in charge of providers (which only have the goal of employability, whatever this goal is achieved), would require a rating system where performance is considered gross of course characteristics.

Finally, given the predicted agency's econometrically-adjusted performance, the allocation of resources between training providers depends on how strict the identification of best performers is (only those with a performance significantly different from the average, or all those with a positive residual, or the first quartile in the ranking ecc.) and on the quota reserved to them.

The possibility to test different versions of the proposed model (along the aforementioned aspects) makes it flexible enough for the effective implementation by the policy maker.

References

Arifil (2013), *La Dote Unica Lavoro Il sistema delle fasce ad intensità di aiuto Descrizione della metodologia*, Newsletter Arifl N.2 - Ottobre 2013**.**

Australian Governament-Disability Employment Services (2013), *DES Star Ratings Methodology Advice.* https://docs.employment.gov.au/system/files/doc/other/disability_employment_services_star_ratings_methodology_advice.pdf

Berkhof, J. and Snijders, T. A. B. (2001), Variance Component Testing in Multilevel Models, *Journal of Educational and Behavioral Statistics*, Vol. 26, No. 2 (Summer, 2001), pp. 133-152.

Finn, D. (2009), *Differential pricing in contracted out employment programmes: review of international evidence.* In: Research report . Department for Work and Pensions.

Goldstein, H. (1997) "Methods in School Effectiveness Research", *School Effectiveness & School Improvement*, Vol. 8, Issue 4, pp.369-395.

Goldstein, H. (2003), *Multilevel statistical models*, 3rd edn. Arnold, London.

Grilli, L. and Rampichini, C. (2009), "Multilevel models for the evaluation of educational institutions: a review", in Bini M., Monari P., Piccolo D., Salmaso L., *Statistical Methods for the Evaluation of Educational Services and Quality of Products*, pp. 61-80, Heidelberg: Physica-Verlag HD.

Hasluck, C. (2004), *Targeting Services in the Individual Customer Strategy: the role of profiling. A review of research evidence*, for Department for Work and Pensions.

Kluve, Jochen & Schneider, Hilmar & Uhlendorff, Arne & Zhao, Zhong (2007). "Evaluating continuous training programs using the generalized propensity score," Technical Reports 2007,39, Technische Universität Dortmund, Sonderforschungsbereich 475: Komplexitätsreduktion in multivariaten Datenstrukturen.

Lam, M. (2007), *Performance measures for welfare-to-work programmes: The relevance of Australian star ratings to the UK*, Work Directions UK, London.

Maitino, M.L., Mariani, M., Mealli, F., Sciclone, N. (2011) What kind of training works for the unemployed and first-time jobseekers? Differential effects of a regional program. Irpet.

Nuremberg, January 12-14, 2005 Supported by EU Commission (DG Empl) VP/2004/007.

O'Connell, P. J., MCGuinness, S., Kely, E., Walsh, J. (2009), National profiling of the unemployed in Ireland, *Research series number 10, July 2009,* The Economic snd Social Research Institute, Dublin.

OECD (2005), "Labour Market Programmes and Activation strategies: Evaluating the Impacts", in OECD Employment Outlook, OECD Publishing.

Paccagnella, O. (2006), Comparing vocational training courses through a discrete-time multilevel hazard model, *Statistical Modelling*; 6: 119–139.

Raudenbush S. and Willms, J. (1995), "The estimation of school effects", *Journal of Educational and Behavioral Statistics,* Vol. 20, pp. 307-335.

Rudolph, H. and Konle-Seidl, R. (2005), Profiling for Better Services . Report on the European Profiling Seminar.

StataCorp (2013), *Multilevel Mixed-Effects Reference Manual*, Stata: Release 13. Statistical Software. College Station, TX: StataCorp LP.

Struyven, L. and Steurs, G. (2005), 'Design and redesign of a quasi-market for the reintegration of jobseekers: empirical evidence from Australia and the Netherlands', *Journal of European Social Policy*, 15, 3, 211–29.

Wooldridge (2013), "A Binary Dependent Variable: The Linear Probability Model". *Introductory Econometrics: A Modern Approach*, 5th international ed.. Mason, OH: South-Western. pp. 238–243)