Table 1 – Target population by certification type and active participation to labour market policies. Absolute and % values. The paper focus on foreign BAS and SPE student (grey area).

	OI	BAS	SPE	TOT by LMP	% by LMP
No	2711	1952	2482	7145	74.4
Yes	1078	617	765	2460	25.6
TOT by certification	3789	2569	3247	9605	
% by certification	39.5	26.7	33.8		

Being the data extracted from monitoring and administrative archives, a careful pre-processing is needed for a correct quantification of the target population[7], which finally counts 9,605 individuals. This number includes the experimental VET activity in compulsory education, which is out of the purpose of this paper.

Notwithstanding the local peculiarities in VT policy programming, preliminary work advised against a sampling stratification by territory and action (Benati *et al.*, 2014). Hence, the sample is stratified by type of certification (compulsory education, qualification, specialization) and active participation to labour market policies (LMP)[8].

3.2 Sampling design and quality assessment

The optimal sampling strategy is not unique, rather it depends on the evaluation objectives.

In the present case, several tasks have to be satisfied:

1. Reliable estimate of VT students' follow-up (accountability purposes);
2. Focus on the main aspects of local VT policies (evaluation and programming purposes);
3. Focus on individual characteristics and outcomes (target evaluation and programming);
4. Net impact estimate (improve policy effectiveness);
5. Investigate labour market transitions.

Clearly, some tasks are partially in contrast, e.g. point 1 claims for a huge treated sample and point 4 for a large counterfactual sample, but no more than 2000 interviews can be globally collected by terms of contract.

In the end, a 2-dimensional sampling strategy is implemented, accounting for the type of certification and the active participation to any LMP after VT enrolment (6 strata overall). This allows the researcher to focus on the peculiarities of each training action, accounting for the effect of other labour market policies. At the stratum-level, a proportional allocation is performed, controlling for individual characteristics that influence his employment outcome (gender, citizenship, age). Practically, the ratios observed in each stratum of the target

---

[7] Duplicates have been reduced to single records prioritising successful and longer treatments, while incomplete records have been matched to administrative SILP data.

[8] This is due to a special interest in the transition from training to the labour market. Obviously, administrative data can solely pinpoint labour market services offered by institutional subjects (employment agencies, town and Province services), neglecting all informal activity (training and temp agencies, private employment agencies, labour union, religious and voluntary associations).