once again—Germany, whose governments of the past two decades invested heavily in building a community able to evaluate labour market policies.

A.2 Target population, geographical scope and time coverage

This CIE pilot study will focus on a limited segment of the complex array of training programmes co-financed by ESF in Italy. In particular, the study will focus on "**post-diploma (PD)" training courses targeted mainly to 20-29 year old unemployed individuals with a high school degree** (known as "diploma"). This segment stands out for being relatively more homogeneous than any other segment of the training system. While publicly-financed training programmes vary enormously in terms of duration, employment status at program entry, eligibility and characteristics of the students they attract, the post-diploma courses we target are more homogeneous: they are all intensive in terms of classroom hours as well as total hours; they are mostly attended by young workers who became unemployed by job loss or by protracted spells of joblessness alternating with periods of precarious employment. In both cases, these individuals are likely to see job training as their "last chance" significant investment in human capital, with the intent to increase substantially their chances of finding a stable and qualified job. A lot of importance seems to be at stake for the individuals choosing these courses than for those picking other "lighter" form of investment in human capital, particularly **internships** activated after the high school diploma. The fact that a lot is at stake for the trainees of these courses, in addition to their higher costs, makes them a worthwhile candidate for an in-depth CIE-based cost-effectiveness analysis.

Do these courses effectively improve the outcome of interest for those who decide to enrol? Are they a cost-effective way of investing public money? Is it possible to conclude that enrolling in these courses for a holder of high school diploma is more effective than continuing job search combined with marginal employment? How do *internships* enter the picture, considering that most post-diploma courses include internship at the end of training?

The impact of the program will be estimated for two main **outcomes**: (i) the probability of obtaining a job with an open-ended contract at any year after the training; (ii) the number of

weeks with a job in the following 12-month period after the program below explains the content and the accessibility of the data sources for these variables. A third outcome (annual earnings) might become available from the Italian Social Security Agency (INPS).

Geographical scope

The effectiveness of these training courses will be evaluated using administrative data coming from five Italian regions, each representing a separate ESF Managing Authority, namely Piemonte, Veneto, Lazio, Puglia and Provincia di Trento. Although not intended to be a nationally representative sample, conducting the same analysis on the same typology of interventions in different regions could have some beneficial effects: (i) it actively involves more regional managing authorities, contributing to develop capacity; (ii) requiring that these regions share the same data sources for analytic purpose, might favour the emergence of nation-wide standards; (iii) pooling the data for several regions enables estimating the possible heterogeneity of the impact of comparable courses across different socio-economic contexts.

Time coverage

The analysis will focus on the PD courses started in 2007, 2008 and 2009. Thus for the 2007 cohort, we will have at least six years of post-enrolment data, four for the 2009 cohort, assuming on having outcome data until the end of 2013. Since these courses last at most two years, with an average under one year, the lock-in effect that is almost always observed for training programmes targeted to those out of work should be absorbed after at most two years, leaving enough time for positive effects to emerge, if indeed this type of training has any positive effect. We will produce results aligned in two different ways: relative to the onset of training (for example, impact of training of the number of weeks worked in the first 12

months, in the second 12 months) and in a given calendar year (number of weeks worked in 2011, in 2012). The relative measures are the most common approach, but the calendar year measures allow placing the results in the context of the evolving macroeconomic situation. Table 1 shows some approximate numbers of enrolees in post-secondary courses. An issue that will need to be solved is whether to include in the analysis courses that started in 2007 (and 2008) but funded under the 2000-2006 programming period. Excluding them would greatly reduce the sample size for Lazio and Puglia and to reach meaningful conclusions, would force us to pool data across years and across regions. To exacerbate the problem, when the "n+2 rule" is no longer a concern in 2009, the supply of training is sharply reduced, as resources are siphoned away to finance income replacement measures.

Table 1: Number of enrolees in post-secondary training courses, by year and region

Fonte; Monit 2000-2006 and Monit 2007-2013

A.3 Suitability of the intervention for a CIE

The suitability of an intervention for CIE hinges critically on the possibility of finding a control group that plausibly represents the average outcome we would have observed if the trainees been exposed to the training. The crucial step is to identify the eligible units at the time of enrolment, so that they can be partitioned into eligible participants and eligible nonparticipants. The control group must be found among the latter. The eligibility requirements for the courses we intend to evaluate are only two: holding a high school diploma, and being out of work at the time of enrolment. The first is easy to observe, albeit with some non trivial amount of missing data; the latter is always more difficult to establish, even more so several years after enrolment.

We will make the following crucial simplifying assumption: we will identify as eligible only those who, in addition to holding a high school diploma, were also **registered as unemployed at the PES office** at the time they made the decision to enrol or not to enrol. However, while such time is well defined for those who decided to enrol, it is not for those who did not enrol. A further complication is due to the fact that people change their decisions overtime, so that a non-enrolled at time t might become enrolled at t+1. Following Sianesi (2004) the solution to both these problems is to focus on the set of people eligible for participating in the wave of courses starting at time t – i.e. holding a diploma and being registered as unemployed at the PES at *that* time. Then we compare (according to the design described in Section C of this proposal) participants and non-participants within *this* pool of people ending up with the estimated causal effect of taking part into wave t of the courses vs. taking part in a *subsequent* wave of the courses or not taking part in a training course altogether.

A.4 The theory of change behind the intervention—and its pitfalls

The theory of human capital in the Becker-Mincer tradition provides a straightforward argument pro-training. Workers with more general education and more specific job training are more productive; as a result the labour market should them pay more than it pays those with less education and less training, and they should be hired first, thus reducing the time they spend unemployed. Moreover, if the supply of training is well planned on the basis on accurate manpower forecasts, there should be only minimal discrepancies between the number trained for a specific occupation and the vacancies filled in that specific occupation.

The reality of training programmes suggests a more nuanced view of the mechanisms at work. There are many reasons why training, particularly vocational training, might fail to deliver positive impacts. First, not everyone can complete every training: most programmes have implicit prerequisites; i.e., some combination of soft skills (self-discipline, self-control, positive attitude), raw intelligence, physical skills (strength, stamina, dexterity). The process driving the unemployed to enrol in training might ignore these requirements, driven by

equitable access, rather than efficiency motives. When prerequisites are in limited supply, supply cannot expand much, so high returns to training for some groups do not imply that returns will be high for other groups. Second, vocational schools provide openings based on administrative decisions concerning available teachers, budgets, and potential enrolment. Although administrators take some account of market demands, the schools seem largely insulated from the job market.

Based on these arguments, the expected results of the intervention are uncertain, and the return on the investment possibly very low. The existing literature is full of examples of such uncertain results. It should be noted that no specific estimates are available for this specific segment of the training system, so that comparison with prior work will not be possible. On the other hand, the 5 regions comprise almost 30 provinces, which are the territorial units that have the final decision on the allocation of ESF money for training. Some exploratory analyses will be conducted to test whether different rules for funds allocation translate into different outcomes.

B. Describe the action and where it will be implemented

The action will be implemented at two—connected but distinct—levels. At the **National level**, a *Management and Methods Coordinating Committee* (MMCC) will have the specific objectives of: (i) maintaining and further developing a common approach to be followed by the five Local Working Groups (LWG); (ii) monitoring the work being conducted by the LWG; (iii) synthesize the findings; (iv) disseminate the results and lessons learned in using CIE. The MMCC must ensure that difficulties encountered are reported and openly discussed with the MMCC and representatives of the other LWGs, and do not end up representing "surprises" at the close of the project.

The second level is represented, in each of the five regions, by the *Local Working Group*, whose core staff is made by researchers based in the three institutions that support the effort by the Ministry of Labour and Social Policies in mainstreaming CIE: ASVAPP will work closely with the ESF-MA of Piemonte; IRVAPP will cover both the Veneto Region and the Province of Trento; while MLPS ISFOL will work with Lazio and Puglia. It is crucial that a collaborative effort and attitude prevails within the *Local Working Groups*. A positive climate between researchers and MA staff will be crucial in determining the ultimate success of this action.

C. Methodology to be followed

Of the four quasi-experimental methods indicated in the Call—namely, difference-indifferences, regression discontinuity design, instrumental variables, and propensity score matching—the last one seems the only applicable approach to PD training courses. While there is no obvious instrument available—some event/characteristic that affects the probability of enrolment but does not directly affect the outcome—the typical case is that there is also no rationing that would possibly lead to a discontinuity design. Eligibility itself does not imply any discontinuity. The choice of limiting the analysis to the 20-29 year-old is a design choice, not a constraint on who can access the programme. Occasionally some rationing for some types of courses might induce management to create waiting lists, but an initial exclusion based on ranking applicants is often followed by some form of "late admission", for example to replace no-shows or early drop-outs, this way weakening the design based on discontinuities. Finally, difference-in-differences applied to training programmes has a long history of dubious applications, due a phenomenon known as "Ashenfelter's dip"—the empirical regularity that the mean earnings of participants in employment and training programmes generally decline during the period just prior to participation, bouncing back naturally afterwards. Such pattern had originally been identified in Ashenfelter (1978) and has since been observed for participants in many other employment and training programmes.

C.1 Describe and motivate the choice of CIE method(s) and data sources

The matching approach is not without problems of its own, as it is based on a rather strong *Conditional Independence Assumption*. This assumption implies that after accounting for a suitable set of observable characteristics of the eligible individuals the difference between participants and non participants in the programme with respect to the outcome variable is due *only* to programme participation. This condition places a heavy burden on the analyst requiring her ability to identify and properly measure *all* the individual characteristics relevant both for the participation decision and for the outcome. As discussed above, the decision to undertake training or not among those eligible *is almost entirely determined by the participants themselves.* Some external constraints might exist on the availability of certain types of courses, but not enough to be exploited for identification of the causal effects.
After restricting the analysis to the pool of individuals sharing the eligibility conditions for the wave t of training courses as described in section A.3, to improve the comparability of the two groups of individuals – those taking part into one of the wave t courses and those not taking part in any wave t course – we rely on the information available in the administrative archives. In particular, we draw from the information on the labour market history of eligible individuals *prior* to their choice to take or not to take part into a wave t course. This is known to be a key information to balance the two groups with respect to individual characteristics relevant for the post-program labour market outcomes (see Paggiaro, Rettore and Trivellato, 2009, for evidence from Italy). The intuition is straightforward: past labour market history is affected by the same observable and unobservable individual characteristics relevant for the future labour market history. In a sense, it is a summary of all the relevant individual

characteristics affecting one's labour market outcomes. As a consequence, conditioning on that summary is an effective way of conditioning on all the individual characteristics entering the summary.
Then, for the implementation of the matching estimator we rely on standard econometric practice of estimating a propensity score, checking its balance property and finally experimenting with a variety of matching procedure to assess the robustness of the results. Besides, new strategies to implement the matching estimator not relying on the propensity score (Iacus, King, Porro, 2011) will be implemented for the sake of exploring the robustness of results to alternative choice of matching procedure. The report will include the details of this robustness analysis.
A possible limit to the validity of this design arises because a certain fraction of **trainees** does not enrol at the PES, but applies directly to the vocational schools. In principle, this would create a potential bias in the impact estimate, undermining internal validity. Note however that our main claim regarding the balancing property of the past labour market history should safeguard with respect to this threat. There might also be a problem of external validity due to this problem of partial coverage in that results for those enrolled in a PES might not carry over to those not enrolled. Preliminary information for Piemonte and Veneto indicates that over 70% of trainees in 2008-09 were also enrolled in a PES office. Were that finding applicable to the other Regions as well, one obstacle to the external validity of the results will be removed.
An additional problem with validity might arise due to partial lack of the so-called common support. This is a problem arising when some individuals taking part into the program feature peculiar characteristics such that there is no individual in the comparison group comparable to them. The analysis is then restricted to the sub-set of participants within the common support, namely those to whom it is possible to match a comparison individual. Strictly speaking, CIE results apply only to this sub-set of individuals, hence a problem of external validity arises. There is no prior result on Italian training courses on which we can rely to predict whether this is likely or not to be a problem. We shall carefully check for its existence and provide full details in the report.

C.2 Data sources, accessibility and content

Figure 1 concisely illustrates the three archives from which the data needed for the analysis will be extracted and then merged to produce an analysis file:

1. the archive built to deliver and monitor publicly funded training programmes;
2. the Public Employment Service archives;
3. the COB ("Comunicazioni OBbligatorie") archive, containing the information that every private employer or its agent is legally bound to communicate on-line to the COB

archive in order to initiate, modify or terminate any work-related contract.

Preliminary agreements have been reached with almost all holders of all three archives—which in all cases are different branches of the regional government ("*assessorati*"). The letters of intent obtained so far are from Regione Veneto, Regione Puglia and Regione Lazio.

FIGURE 1: THE DATA ARCHIVES NEEDED FOR CIE OF TRAINING PROGRAMMES

The three archives play quite a different role in the process of estimating impacts. The training archives have a minimal role, despite being traditionally at central stage, because they contain all the *monitoring* results. These archives are still important for CIE, because they allow to identify the individual trainees, as well as the title of the training course. The PES archive become crucial once our *ad hoc* definition of eligibility is deemed acceptable. If eligibility consists of being enrolled in PES, it is easy enough to separate the eligible participants from the eligible non-leading its first essential contribution, since the possibility of matching hinges on the possibility of including substantial information on the personal characteristics of treated and non-treated units, particularly their pre-enrolment work history. Prior studies, both in Italy and elsewhere, have repeatedly shown that prior work history, rather then demographic or family–level information, is the key information to balance treated and controls. Finally, once the control group has been identified, the COB are essential in providing employment-related events collected by the same mechanism for treated and controls.

D. Expected results and their use

The value-added of this Call—as well as the motivation for the submission of a bid by the Italian Ministry of Labour and Social Policies—is given by the emphasis on capacity building and increased visibility of CIE among managing authorities. To be sure, the whole evaluation exercise would lose meaning if the analytical work was not of the highest standard and the estimates obtained were not based on scientific methods. Therefore the first intended result of this action will certainly be the production of "*robust evidence on the impacts of the ESF interventions*". However, we believe the project would lose entirely its meaning if it would just produce "robust evidence" but it failed to deliver "*increased awareness and knowledge of counterfactual analysis*" and even more if it were unable to generate "*enhanced capacity to effectively implement counterfactual analysis*": capacity which implies "being able to write competent terms of reference in order to formulate questions that are relevant for policy decisions", even without knowing all the technicalities that are the unavoidable by-product of any advancement of knowledge.

Although it would be impossible (or a futile exercise when possible) to actually measure "awareness", "knowledge" and "capacity", we believe that the members of any professional community know how to recognize them and act accordingly to maximize them.

We intend to give meaning to this project in four ways: (i) by focusing most of its resources on basic applications of propensity score matching, rather than pursuing applications of the most sophisticated and not fully tested techniques; (ii) by producing accessible, non-technical writing on the methods and onhe results, conveying the message of the potential as well as limited applicability of CIE; (iii) by disseminating this material to the staff of the managing authority as well as to the top management of the regional governments, including members

of the regional assemblies: (iv) by offering training opportunities for the staff of managing authorities.

E. Project management

The action will involve several organizations, playing different roles: the Italian Ministry of Labour and Social Policies, who is the official applicant and provider of the co-financing of the project, together with its in-house research organization, MLPS ISFOL; five regional ESF managing authorities (Piemonte, Veneto, Lazio, Puglia, Trento) from whom full cooperation is expected in making the data available; two organizations (ASVAPP and IRVAPP) who have been at the forefront of the attempt to introduce rigorous CIE in Italy, both with experimental and non experimental methods.

The first step for the management of the project will be the official appointment in Rome by the Ministry of the *Management and Methods Coordinating Committee* (MMCC) who will be in charge of coordinating the whole research effort, guaranteeing a common analytic approach, monitoring the work being conducted by the local teams, synthesizing the findings and disseminating the results. The MMCC will meet approximately for a total of eight one-day

meetings, of which one will be held, respectively, in Torino, Venice, Trento and Bari, so that the MMCC members can meet the local ESF managers; and the remaining three, including the first and the last one, in Rome, so that they can be attended easily by the representatives of the Ministry. After each meeting, the coordinator of the project will issue a Progress Report, to be delivered to the Project Manager (Marianna D'Angelo) of the Ministry of Labour. The first meeting will be held in Rome within the first month after signing the contract with the EC DG Employment and serve as kick-off meeting.

The Ministry intends to appoint to the MMCC five nationally recognized experts, who will report directly to the Ministry. The members of the MMCC would be: Bruno Anastasia (Agenzia Veneto Lavoro), for his expertise in the use of administrative data in support of the analysis of labour markets; Alberto Martini (Università del Piemonte Orientale), for his international experience in CIE and his expertise in simplifying complex problems for nontechnical audiences; Enrico Rettore (Università di Padova) for his internationally recognized statistical expertise applied to evaluation issues; Ugo Trivellato (Università di Padova) for his in-depth knowledge of labour markets, evaluation methods as well as data confidentiality issues; Paolo Severati (MLPS ISFOL), who will be the scientific coordinator of the project, for his extensive expertise in the planning and management of evaluation in the ESF.

The MMCC will draw advice from a larger Scientific Committee (SC), that will never physically meet, but provide comments on request and meet if needed through conference calls. The SC will be chaired by Antonio Schizzerotto (Università di Trento) and include Daniele Checchi (Università di Milano), Claudio Lucifora (Università Cattolica), Erich Battistin (Università di Padova), Daniele Bondonio (Università del Piemonte Orientale), Andrea Ichino (Università di Bologna) and Kevin Hollenbeck (Upjohn Institute, Kalamazoo, Michigan) who has conducted the largest non-experimental evaluation of the Workforce Investment Act, the training and employment services in the US that come closer to the ESF (Hollenbeck, 2009).

The main task of the MMCC will be to coordinate and monitor the work of the local research units, hosted by ASVAPP for Piemonte (under the responsibility of Luca Mo Costabella), IRVAPP for Veneto and Trentino (under the responsibility of Slavica Zec), MLPS ISFOL for Lazio and Puglia (under the responsibility of Paolo Severati).

The local research units will be established within the second month from the start of the project. By the fourth month they must be able obtain the data described in section C.2 and perform a first round of impact estimates, according to the methodology outlined in section C.1 and subsequently refined by the MMCC.

A second round of refinements will be necessary before the results reach their final form. The goal is to have all empirical work completed by the ninth month of the contract, so as to allow

a sufficiently long period for the dissemination phase of the action.
A detailed work programme is presented in Annex D2.
11

F. Arrangement for monitoring/supervision of the operation

Monitoring will be the first order of business of the MMCC, who will need to focus on the soft aspects of climate and conflict resolution, in addition to guaranteeing that impact estimates are produced on schedule using state-of-the-art econometric techniques. The risk to the project are foreseen coming not from the technical side, given the proven experience and expertise of the individuals doing the work, but from its ability to produce awareness, knowledge and capacity. Special attention should be devoted to progress being made on these three fronts. Furthermore, direct and indirect control systems, analogous to those of the Ministry of Labour, will be implemented in order to monitor and supervise effectively the status of the activities.

Lastly, within 60 days after the closing date of the action, we will submit to the Commission the final report on implementation of the action, which will be presented based on the template available in the call, together with the final financial statement of all actual expenditure and actual revenue for the project. Both the above mentioned reports will be submitted on-line via SWIM and in paper version.

G. Sustainability of the project's achievements

The long-run goal is to remedy a situation in which claim of effectiveness of labour market policies are based on wishful thinking rather than on empirical evidence. The counterfactual view brought into policy analysis and evaluation a *constructively sceptical* attitude toward "effectiveness": from one that presumes any policy is effective as long as the money is spent without irregularities; to one in which any policy is subject to scrutiny for their effectiveness using data and rigorous methods. The intent of the Ministry of Labour and Social Policies is to trigger a more wide-reaching effort to induce (and enable) regional ESF Managing Authorities to rigorously evaluate some of their labour market programmes. More resources will need to be devoted to capacity building, but this project will be a litmus test of the feasibility and chances of success of such wider effort.

REFERENCES

Ashenfelter, Orley. (1978). "Estimating the Effect of Training Programmes on Earnings. " *Review of Economics and Statistics,* 60, 47-50.

Hollenbeck, Kevin (2009) "Workforce Investment Act (WIA) Net Impact Estimates and Rates of Return." Paper presented at EC-Sponsored Conference on "What the European Social Fund Can Learn from the WIA Experience" with the Workforce Investment Act. Washington, DC: November 2009.
12
Iacus, Stefano M, King, Gary Porro, Giuseppe (2011) "Multivariate Matching Methods That Are Monotonic Imbalance Bounding," *Journal of the American Statistical Association*, American Statistical Association, vol. 106(493), pages 345-361.

Kahn, Lisa B., (2010). "The long-term labour market consequences of graduating from college in a bad economy," *Labour Economics, E*lsevier, vol. 17(2), pages 303-316, April. ,"

Paggiaro, Adriano Enrico Rettore, Ugo Trivellato (2009), "The Effects of Temporary Job Experiences On Short-term Labour Market Outcomes in Italy" , Rapporto di Ricerca IRVAPP

Sianesi, Barbara (2004), "An evaluation of the Swedish system of active labour market programmes in the 1990s", *Review of Economics and Statistics*, Vol. 86, No. 1, pp. 133-155