Regional disparities in the effect of training on employment

Regional disparities in the effect of training on employment

Andrea Filippetti
National Research Council, Institute for the Study of Regionalism, Federalism and Self-Government (ISRFA), Via dei Taurini, 19, 00185 Rome, Italy, and London School of Economics and Political Science, Department of Geography and Environment, Houghton Street, London, WC2A 2AE, UK
Email: a.filippetti@cnr.it

Frederick Guy
Birkbeck, University of London, Department of Management, Malet Street, Bloomsbury, London WC1E 7HX, UK
Email: **F.Guy@bbk.ac.uk**

Simona Iammarino*
London School of Economics and Political Science, Department of Geography and Environment, Houghton Street, London, WC2A 2AE, UK
Email: s.iammarino@lse.ac.uk

Abstract

In this paper we investigate one particular aspect of human capital formation: the relative effectiveness of training, as reflected in its effect on the probability of securing continued employment during the recent financial crisis. We use a panel on 3,983 individuals for the period 2008-2011 and focus on how the effects of training differ between the South and the North of Italy and across workers with different levels of education. Our most striking result is that the effect of training on continued employment is notably stronger in the South than in the North of the country.

Keywords: Training in employment, South and North of Italy, Education, Economic Crisis

JEL Codes: R1, R23, J2

*Corresponding author

Regional Disparities in the Effect of Training on Employment

Introduction

In this paper we investigate the relative effectiveness of employee training in securing continued employment in the context of the recent economic and financial downturn. By many measures, territorial differences have grown since the global financial crisis of 2008. Such uneven impacts of the crisis are not unique to any particular country, and have led some scholars and policy makers to call for more economic geography-based perspectives on the consequences of economic crises (e.g. Martin, 2011; EC, 2013; Bailey and De Propis, 2014; McCann, 2016; Turok et al., 2017).

By training, we mean relatively short-term instruction in some well-defined set of employment-related skills. The provision of training is a matter of growing public and policy concern for two main reasons. One is that the factors that shift demand – changes in consumer taste (the speed of the product life cycle), in production technology, and in the spatial organization of the division of labour – have accelerated. A serious economic downturn – as in the period 2008-2011 studied here – adds a further immediate demand shock, and can accelerate the ongoing structural changes. The second factor is that private actors may lack sufficient incentive to pay for training: for the individual, it is an undiversified and risky investment; employers, especially large ones, are somewhat better able to self-insure against this risk, but they generally face free-rider problems when inter-firm employee mobility is high. In recent decades, as contingent employment has grown, employers' motivation to invest in human capital has weakened and, at the same time, public resources for education and training have diminished. Together, these factors both accentuate the need for training, and threaten to undermine the supply.

Italy is characterized by profound geographical differences in the types of industrial specialization, quality of education and human capital, and labour market performance. These differences – part of the historical socio-economic and institutional differentiation of the country as a nation state (e.g., among the most recent, Trigilia and Burroni, 2009; De Blasio and Nuzzo, 2010; Mauro and Pigliaru, 2011; Dias and Tebaldi, 2012; Crescenzi et al., 2013; Cellini and Torrisi, 2014; D'Agostino and Scarlato, 2015) – have in recent decades been exacerbated by internal migration of highly-educated labour force from the Southern to the Central and Northern regions (e.g. Piras, 2005; D'Antonio and Scarlato, 2007; Etzo, 2011; Marinelli, 2013; Nifo and Vecchione, 2014).

In the frame offered by this ancient and deeply embedded dualism, we set out to examine differences in the effect of a specific measure of skill upgrading – training given to people who are already employed – on a specific outcome – remaining employed. The data we use to study this issue spans 2008 to 2011, and thus happens to coincide with the recent financial crisis.

This paper's contribution to the existing literature is twofold: it focusses on the effectiveness of training in ensuring employability at the sub-national level, a dimension that has been so far neglected in the limited research on training and labour market outcomes; it considers a period of severe recession, 2008-2010, followed in Italy by the short-term recovery of 2011 (Faggian et al., 2017). As the economic crisis has shown unusual regional dynamics – trespassing the typical partition of the country in North versus South (Centro Studi Confindustria, 2017) – changes in demand and supply for training may also be territorially uneven, in either their incidence or their impact (Martin, 2012). Hence, territorial differences in the provision and effectiveness of training could play a role either in the erosion of Italy's historic geographical dualism, or in its reproduction. This is all the more important when

considering a major economic shock, as training may facilitate workers' transition to new labour demand conditions.

We use a panel from the PLUS survey (Survey on Labour Participation and Unemployment Survey) carried out by the Institute for the Development of Vocational Training for Workers (ISFOL), which includes 3,983 individuals covering the years 2008, 2010 and 2011, in order to estimate the effects of training on the likelihood of staying employed. Our focus is in particular on how the effects of training differ between the South and the North of Italy, and across workers with different levels of education. Doing so requires to address the identification problem posed by the possible endogeneity of training, which we tackle by exploiting the unique richness of information contained in the dataset, that allows controlling for both job-level features and a wide variety of individual level characteristics, and by adopting an instrumental variable approach. Our most striking finding is that training has a clear positive effect on the likelihood of a worker's continued employment in the Italian South, but only a small and statistically insignificant effect in the North. This applies at all levels of formal educational attainment.

The paper is divided into 6 sections. The following section reviews the literature on the relationship between training, education and geography, providing the background for the empirical analysis. We then present the data and define the variables used in the analysis. After describing the estimation strategy and reporting the main results, the robustness of the estimates are discussed. The conclusion summarises the implications of the study and directions for future research.

Background of the study

A considerable body of research has addressed the relationship between training and earnings and/or productivity (among many, Blundell et al., 1996; Dearden et al., 2006; for a review see Leuven, 2005; for Italy see Conti, 2005; Colombo and Stanca, 2008); a somewhat smaller research stream has addressed training and employment outcome (e.g. Ok and Tergeist, 2003; Sanders and De Grip, 2004; Budría and Telhado-Pereira, 2009), but none of this latter has addressed sub-national differences and territorial specific patterns.

In the rest of this section we review first the evidence on the incidence of training across employees by levels of education, and the effects of training on continuity of employment; second, the sub-national geography of training incidence and effects; and third, sub-national variations within Italy which may affect the incidence or effectiveness of training.

Training incidence and employment outcomes

Training provides human capital after schooling, and is often associated with some set of skills useful for a particular occupation or industry, or complementary to a particular set of technologies. We might think of training as something which compensates for deficient human capital, yet a consistent finding in the literature is that there is a positive correlation between prior education and training while employed. According to the International Adult Literacy Survey, 13% of employees in low-skilled occupations participate in employer-sponsored training courses, compared to 38% in high-skilled occupations (Bassanini et al., 2004). Other studies have found that the amount of training is positively correlated with the level of formal education (Heckman, 1999; Kuckulenz and Zwick, 2003; Hughes et al., 2004); that firms with a more highly qualified workforce and advanced work organization train more (Lynch and Black, 1998); and that training incidence is higher in countries with more educated labour forces (Brunello, 2001). Jobs for which previous school-provided

vocational training is important in selecting new hires tend to involve much more training on-the-job than jobs for which previous school-provided training is not important (Bishop, 1996).

Training while employed is also correlated with job characteristics. Workers with permanent or long-term contracts tend to receive more training than those on temporary contracts. Arulampalam et al. (2004) find, across Europe, a negative association between fixed-term contracts and training, and that public sector workers get more training. Workers in large firms receive almost twice as many hours of training as workers in small firms (Bassanini et al., 2004).

Training generally improves the likelihood of continued employment. In a cross-country study Ok and Tergeist (2003) find that, after controlling for individual characteristics, training reduces the job-loss rate by 3.5%; this effect is particularly large in Italy, and is driven by the higher rate of re-employment for trained workers. Sanders and De Grip (2004) show that training of low-skilled workers contributes only to retention of existing jobs, while it does not contribute to their external employability. Budría and Telhado-Pereira (2009) suggest that schooling and training are complementary in their contributions to remaining employed.

The effects of both education and training on employment outcomes may vary with the unemployment rate. When high unemployment puts workers with different skill levels in competition for the same jobs, the more highly educated or better trained are favoured both in labour hoarding and in hiring, as well as in retaining the current job (van Ours and Ridder, 1995).

Sub-national geographies of training and skills

From the evidence we have on territorial variance in labour market structures – as for example between the North and South of the UK (e.g. Turok and Edge, 1999; Sunley et al., 2001; Gardiner et al., 2013; Green and Livanos, 2015), or the South and North of Italy (e.g. Brunello et al., 2001; Contini and Trivellato, 2005; ISFOL, 2014; Meliciani and Radicchia, 2016) – we should expect a territorial differentiation on both provision and effect of training with respect to employment and labour market outcomes (Fingleton et al., 2015).

The complementarity of training with prior education suggests that training might have a very limited role in improving the relative performance of regions in which education is comparably poor, or of improving the lot of poorly educated workers. This raises an issue for scholars of regional economic development, the literature of which is couched in terms of learning dynamics and creation and diffusion of knowledge (e.g. Paci and Usai, 1999, 2000). In this literature, while skills are typically understood as an important variable, they are usually not theorized or problematized. As Rutherford (2001, p. 1874) observes, learning-region approaches view geographical labour markets "either as secondary or as analytically indistinct from learning based on direct interfirm and institutional networks": in other words, the focus on inter-organisational and institutional knowledge exchange and innovation-intensive networks – a crucial contribution of the innovation systems perspective – has led to neglect the role played in regional economic development by the education and training of the broad labour force. Increasingly, however, scholars have seen a need for a more geographically-specific perspective on both such labour market dynamics and the consequences of economic crises and shocks (e.g. Shuttleworth et al., 2005; Quatraro, 2009; Martin, 2011). Processes of skill change are in fact extremely 'sensitive to the contingencies of place' (Peck and Haughton, 1991, 829; see also Turok and Edge, 1999; Sunley et al., 2001; Wixe, 2015; Kasabov and Sundaram, 2016).

Regional differences in market and industry structure can affect returns to employers from investments in training, and this may either reinforce or offset geographical agglomeration effects. Firms are expected to invest more in training when product markets are imperfectly competitive (Acemoglu and Pischke, 1999), which in general would be more likely in regions with lower employment density and/or larger firms. The same firms are more inclined to hoard skilled labour in periods of slack demand (Sepulveda, 2002; Brunello, 2009); labour hoarding during such times should, in turn, promote further training, since the opportunity costs of providing training is lower when demand is slack. Even in the absence of hoarding, market structure can affect the nature of skills produced in a locality: large firms with elaborate internal divisions of labour have different skill needs than smaller firms (Kamnungwut and Guy, 2012). On the other hand, regions with industry structures characterised by specialised agglomerations, technological complementarities and localised knowledge spillovers may encourage the provision of adaptable skills, despite labour market pooling tends to generate higher wages. In a study on training and local economic conditions across the Italian provinces, a negative relationship is found between the provision of training and employment density, which however is reversed in the case of highly specialised agglomerations (Brunello and De Paola, 2004).

On the sub-national impact of training on employment outcomes the literature so far is silent: one should expect that the differences in market and industry structures that explain the provision of training may similarly affect its impact on employability.

Regional differences in labour markets and training in Italy

Italy has always inspired scholars from various disciplines interested in investigating territorial differences while holding formal national institutions constant. In this paper, we divide Italy into two areas, or macro-regions: South, or Mezzogiorno, and North. Table A1 in

Appendix A, the variable list (for this and all other tables and figures with the prefixes "A" and "B", see the online appendix), defines the regions we classify as South and as North.

Such a territorial division of Italy is customary adopted in the literature and can be justified on the basis of the historical background (most of the South was a colony of one foreign power or another throughout recorded history, until the creation of Italy in the late 19th century – see Iammarino, 2005, for a brief review); of historical and recent social and economic statistics (e.g. Iuzzolino et al., 2013); and of any of numerous measures of institutional quality (e.g. D'Agostino and Scarlato, 2015).

There is an alternative territorial division often reported in the literature in which the North is itself divided into two, with some regions in the Centre seen as an institutionally and economically distinctive 'Third Italy', renowned for rich social capital (e.g. Putnam, 1993) and clusters of SMEs (e.g. Becattini, 1979; Pyke et al., 1990), in contrast to the supposed large firm prevalence elsewhere in the North. Yet, the usefulness of this North-Centre distinction has been questioned on the grounds that the SME networks of the Centre (including also, for these purposes, the Northeast) are in fact closely integrated with the larger multinational firms and business services of the North (or, more precisely, the Northwest: see Dunford and Greco, 2005). The South exists largely outside of these networks, except to the extent that it hosts branch plants of Centre-North firms (a decreasing presence over the past two decades, as many such branch plants have been relocated abroad). Local social networks are not as extensive in the South, which lacks 'bridging' social capital (Crescenzi et al., 2013); consequently, southern firms tend to internalize more functions and more stages of production, and thus to be less networked than firms in the Centre-North (e.g. Passaro, 1994; Lazerson and Lorenzoni, 1999). The relatively stand-alone nature of southern firms – located in areas of lower inter-firm collaborative linkages and employment density – should increase their monopsony power in relation to workers and, for the reasons discussed above, this

should be a factor encouraging them both to provide more training and to retain skilled workers when demand is slack.

With respect to human capital, the South is comparable to the North in terms of levels of formal educational qualifications, but it scores lower in terms of internationally standardized measures of literacy and numeracy (e.g. Bratti et al., 2007, on Pisa scores). Following the 2001 constitutional reform, the Italian vocational education and training system was modified in the direction of attributing more responsibility at the regional level, with the idea that training policies should reflect the specific territorial specialization. This process of decentralisation has been further strengthened by the shift of the European cohesion policy towards place-based and smart-specialization approaches – the bulk of resources for training deriving from the Structural Funds. As a result, the regions have major responsibility in the area of vocational training and its integration with the educational system: regions are in charge of identifying the curricula and experimenting new tools for vocational training and apprenticeship. Italy has thus experienced an increase in the variety of local training instruments going from incentives to employability to professional re-training (Teselli, 2016).

Figure A1 shows average GDP per capita and unemployment rates in the Italian South and North from 2000 to 2011, and the percentage of temporary on total employment (excluding self-employed) between 2004 and 2011. The GDP gap between the two areas is considerable, and fairly steady over this period. The unemployment rate in the South is always much higher, but substantial convergence (and overall decline in the rate) occurs between 2000 and 2007, while a steady difference in the rates (in the context of rising unemployment overall) happens between 2008 and 2011. The share of temporary workers is also considerably higher in the Mezzogiorno regions, though the difference has slightly reduced since the outset of the crisis (ISTAT, 2014).

The steadiness in the difference in unemployment rates has been achieved by a reduction in labour force participation in the South (ISFOL, 2014). In addition, higher unemployment in the Mezzogiorno regions has been associated with a higher degree of worker turnover (see, for all, Mussida and Pastore, 2015): this regards in particular young, female, less educated and temporary contract-holder individuals. Figure A2 disaggregates the unemployment rate for 2011 by the twenty Italian administrative regions; in keeping with the averages from Figure A1, unemployment rates are distinctly higher in the South.

Figure A3 compares regional rates for remaining employed in 2011 if employed in 2008 (our dependent variable in the empirical analysis below): the rate tends to be higher in the North. Figure A4 shows rates of training while employed by region over the period 2008-2010: this is the share of those employed who reported participation in training in the PLUS surveys of 2008, 2010, or both. Here there is no clear South-North pattern: both in the South and in the North we see a mix of regions with low, middling and high levels of training participation.

The effectiveness of training at sub-national level is not addressed by the literature; the literature does suggest a number of possible relationships between regional development and training, but leaves us without clear expectations. Lower levels of economic development in the South may be associated with lower quality of human capital, and thus a reduced effectiveness of further training; differences in industry structure and intensity of localised externalities could produce more firm-specific kinds of training in the South, but whether this would be associated with greater or lesser effectiveness is not clear.

Data sources and sample

We employ data from the PLUS Survey (Participation Labour Unemployment Survey), a sample survey on the Italian labour market supply developed and administered by ISFOL, a

national research institute reporting to the Italian Ministry of Labour and Social Policy (see, for a detailed description, Gianmatteo, 2009; Meliciani and Radicchia, 2011, 2016; Mandrone and Radicchia, 2012). The PLUS Survey annually samples about 40,000 individuals, contacted through a dynamic CATI system without proxy interviews. The survey sample design is stratified over the Italian population aged 18-64: strata are defined by region (20 administrative regions), type of city (metropolitan/not metropolitan), age (5 classes), sex, and employment status of the individual (employed, unemployed, student, retired, other inactive/housewife). The reference population is derived from the annual averages of the ISTAT Labour Force Survey.

We use the 2008, 2010 and 2011 survey waves, including 12,593 individuals in each wave. The panel we have is balanced. We do not have information on any participants who did not respond to all three surveys. ISFOL provides weights to account for the probability of attrition based on surveyed characteristics (Mandrone et al., 2014). All estimates reported in the paper use those weights.

The survey is extremely rich in information on individual job features, employer characteristics, types of training activities, income, and educational history, at the same time providing detailed information on other crucial aspects of the respondents such as, for instance, family background, residential mobility, geographical location, self-confidence, and health.

Our sample consists of those who were employed in 2008. We exclude self-employed workers, because we expect that there is considerable unobserved variation in how fully or gainfully employed these workers really are. This leaves us with observations on individuals included in the panel who were employed in 2008, and had not retired, gone in parental leave or full-time education in 2011. After excluding observations for which any of our regressors is missing, this gives us a sample of 3,983 individuals.

Our dichotomous dependent variable, *remain employed*, takes a value of 1 when the respondent is (i) recorded as employed in all three surveys, 2008, 2010, and 2011, and (ii) does not shift from full-time employment in 2008 to part-time employment in 2011. It takes a value of 0 – an adverse outcome – for those who were employed in 2008, and in 2011 were either (i) unemployed, or (ii) out of the labour market but not retired, on parental leave or in full-time education, or (iii) working part-time, if they had been working full-time in 2008.[1] We exclude workers reported as retired even though some of these are, of course, actually adverse labour market outcomes (unplanned, involuntary retirement): we have no way of distinguishing such cases from planned retirement, and in the latter case we would expect strong endogeneity of training, since both employee and employer anticipate the employee's departure from the labour market. In all, 344 (9%) of the 3,983 observations were coded as adverse outcomes. For consistency in the construction of our dependent variable we exclude those individuals who were employed in 2008 and 2011 but unemployed in 2010 (1.3% of the sample).

The variable *training* takes the value 1 if the individual respondent reported both being employed and receiving training in the 2008-2010 period, 0 otherwise. Here we define training by using a conservative measure which includes training in the classroom, distance learning and on-the-job training, and excludes short-term forms of training, i.e. seminars, conferences, workshops, fairs, and other.

[1] In Italy, full-time workers forced into involuntary part-time following the crisis in 2008 have steadily increased over time (see, for instance, Associazione Bruno Trentin, 2015). This is the reason behind our choice to treat the transition full-time to part-time as an adverse outcome. We have however checked the robustness of our estimates by dropping the transitions to part-time altogether and the results are qualitatively unchanged.

The choices made in the definition of the sample and key variables may, of course, affect our results. Alternative formulations of both are discussed below, with estimates reported in Appendix B.

For purposes of analysis, we divide Italy into two macro-regions: South and North. We have also run our basic models on the mentioned alternative territorial division of three areas (South, Centre and North), but we found little difference between North and Centre. We obtain data on regional characteristics from various National Institute of Statistics (ISTAT) sources. All variables used in the analysis, with definitions and sources, are listed in Table A1.

Table A2 gives means and standard deviations for the sample, and breaks this down both by trained/not trained and South/North, reporting the t-test. In the panel at the end of Table A2 (Mean participation in training), we see that training rates rise with the higher level of education – about 40% for university-educated workers, 30% for those who completed secondary school but not university, and 15% for those who had not completed secondary school; for each education level, the training rates are approximately the same in North and South.

Empirical strategy and estimates

The estimation strategy for the effect of training while employed on the probability of remaining employed starts with this simple baseline equation:

$$remain\ employed_i = c + training_i + south_i + edu_i + reg_r + ind_i + job_i + v_i \quad (1)$$

South is 1 for the Mezzogiorno, 0 for North and Centre. *Edu* consists of indicator variables for two of the three levels of educational attainment. *Reg* is a set of macroeconomic controls which vary across Italy's twenty regions; *Ind* is a vector of controls for individual

characteristics; *Job* is a set of controls for characteristics of the employer or of the individual's job. Both *Ind* and *Job* are based on the 2008 survey, except for change of place of residence (moved for job) as noted below; *c* is a constant, and *v* is the error term.

As in any multiple regression, the inclusion of controls is an effort to approximate experimental conditions when using non-experimental data. Our treatment – *training* – is not randomly assigned, meaning that its incidence may vary with characteristics of the employee, of the employer, or with the economic and institutional environment in which the employment relationship is situated. As discussed above, such differences are quite evident in Table A2. The question then is whether the controls adequately account for unobserved individual characteristics which may be correlated with the outcome and, if not, whether another strategy is available to redress this problem.

Table 1 presents the results from our baseline probit estimates. We can see how the controls affect the estimated impact of training on remaining employed, as we move across the columns of the table. Model 1A gives the probit for training with no covariates; model 1B introduces a dummy for South, and also the macroeconomic variables measured at the level of the twenty administrative regions. *A priori,* we would expect that variables such as the regional unemployment rate, or its change, would be important, but they are not statistically significant and do not produce any substantial change in the coefficient of training. Model 1C adds a number of additional variables reflecting characteristics of the individual worker. Higher levels of education are strong predictors of training, as we expect; so are being male, years of work experience and being a foreign citizen. Most of the other controls are not statistically significant – including, contrary to our priors, any of the age categories, in comparison with the excluded group of 18-24 years, and the other two proxies for worker's ability beyond education and work experience, i.e. top-grade and mother's education. The inclusion of variables for these personal characteristics reduces the estimated effect of

training on continued employment by about 25%. Model 1D then adds characteristics of the employment contract, change of job and move for job-related reasons, a public sector dummy, and the size and industry (15 industry categories) of the firm. Permanent contracts, part-time work, public sector employment, and larger firms are all predictive of continued employment – and, with these variables included, the direct effect of training on continued employment is a bit more than half of its original value.

[Table 1 about here]

In order to study differences in the effect of training between South and North, and between education levels, we introduce interaction terms as follows:

$$remain\ employed_i = c + training_i + south_i + edu_i + reg_r + ind_i + job_i + \boldsymbol{training_i} * \boldsymbol{south_i} + v_i \quad (2)$$

$$remain\ employed_i = c + training_i + south_i + edu_i + reg_r + ind_i + job_i + \boldsymbol{training_i} * \boldsymbol{south_i} + \boldsymbol{training_i} * \boldsymbol{edu_i} + \boldsymbol{south_i} * \boldsymbol{edu_i} + \boldsymbol{training_i} * \boldsymbol{south_i} * \boldsymbol{edu_i} + v_i \quad (3)$$

Table 2 presents the results of the probit estimates for Models 2 and 3. Both equations are estimated with the full set of controls (the same as in Table 1, Model 1D), but these are not reported, as they do not change materially from the previous estimates.

[Table 2 about here]

Model 2 tells us that the effect of training on the probability of remaining employed is not statistically significant in the North (the base case); the interaction of training with South is positive and statistically significant. The right hand side of Table 2 reports differences, between South and North, in the marginal effect of training on continued employment: here Model 2 shows that training gives workers in the South an increase of 0.103 (10.3 percentage

points) in the probability of remaining employed; this is 0.081 greater than the point estimate for the effect of training in the North. Both the effect of training in the South, and the difference between the effects in South and North, are statistically significant at the 0.001 level.

Model 3 allows us to examine the effect of training, and South-North differences in this effect, across three levels of educational attainment. With the exception of medium-educated workers in the North, the estimated effects of training are all a bit smaller than in Model 2. The relative values between South and North remain unchanged, and overall training is more effective at higher educational levels.

Robustness of the estimates

Our estimates could be fragile or unreliable in a number of different ways: if our controls are inadequate, we could have a problem of endogeneity in the selection of workers into training; our dependent variable is constructed via choices we made defining certain labour market outcomes as good, others as bad, and excluding still others altogether; we pool disparate groups of workers; the South-North difference we find in the effectiveness of training is not explained by the variables in our model, but are there other variables suggested in other studies which may explain it?

For reasons of space, here below we focus in particular on the first and the last of these issues.

IV strategy

We do not have information on who decided that a worker was to be trained – it could be the employer, the worker, or the outcome of a joint decision. If the choice is the employer's, it is

likely that she will opt to train workers who are expected to be kept, whether because they are regarded as more capable or because it is more costly to dismiss them.

Our strategy has been to address this problem with a very rich set of controls for characteristics of the individual worker, the employer, the contract, and regional macroeconomic conditions. In addition, we apply here an instrumental variable approach. In keeping with previous works which have studied the effects of training, we use a region-level aggregate measure: average expenditure for active labour marker policies per unemployed, source ISTAT, census 2011 (*ALMP*).[2]

Conventional instrumental variable regression where both the dependent variable and the endogenous regressor are binary, is not straightforward: standard one-step maximum likelihood estimators will not accommodate it; use of probit fitted values in a two stage estimator is not valid (Angrist and Pischke, 2009, 190-193). We use the control function approach (Wooldridge, 2010, 126-137) which consists of first estimating a first stage linear regression with the potentially endogenous regressor as the dependent variable, and the instrument along with all other regressors on the right hand side:

$$training_i = c + ALMP_r + reg_r + edu_i + ind_i + job_i + e_i \qquad (4)$$

[2] ALMP expenditure includes: apprenticeship, work-entry contract, incentives for taking on long-term unemployed and for hiring those registered in mobility list, temporary contracts, and conversion of temporary contracts into permanent ones.

The residual from Model 4, e_i, is then included as an additional regressor in Model 2. The statistical significance of the coefficient on e_i serves as a test for the endogeneity of training; if e is not significant, it may be dropped from the final estimates: these are reported in Table 3.

[Table 3 about here]

ALMP is significant at the 1 per cent level in Model 4, and the Kleibergen-Paap statistic of 13.80 indicates that the instrument is strong (rule of thumb > 10). We estimate Model 2, with the residual from Model 4 now included. The residual is not statistically significant, and the coefficients on training and its interaction with South are little changed from what reported in Table 2. To the extent that we can have confidence in our instrument, the test provides evidence that endogeneity is not a problem for our estimates.

Explaining the South-North difference

The South-North difference in the effectiveness of training is surprisingly strong, given the large number of controls included in the models. This suggests that an answer should be sought in differences, at the regional or sub-regional level, in institutions or in industry structure. As discussed in the review of the literature above, previous research offers a broad menu of explanations for the size and persistence of South-North differences in Italy. Some of these can be addressed by incorporating additional region-level variables into the regression models above (all variables listed in Table A1 in the online Appendix).

First, many authors have suggested that differences in the quality of social capital produce differences between South and North in the functioning of both formal institutions and commercial relationships. Tabellini (2010) provides region-level survey-derived measures of trust ("most people can be trusted"), tolerance (scaling for importance of "tolerance for other people") and respect (importance of "tolerance for others").

Second, although levels of formal educational attainment are similar in South and North, measures of the quality of basic education are not, as reflected in regional averages on the Pisa tests and of PIAAC literacy scores; we also consider the number of technical high schools per capita, as an indicator of investment in industry-specific human capital (Filippetti and Guy, 2015). The regional pool of human capital is also affected by the prevalence of training in the region, as measured by the regional average of our training variable.

Third, differences in both regional education systems and labour market performance may be reflected in the regional share of youth not in education, in employment, or training – NEET.

Finally, there are broad area differences in the industry structure, as reflected in employment shares in the primary, secondary and tertiary sectors.

[Table 4 about here]

Incorporating each of these variables in turn as an additional regressor in Model 2 above as reported in Table 4, we did find that two of them – Manufacturing and Service shares – had statistically significant effects on employment; however, in no case did the inclusion of these controls lead to any substantial change in the size or statistical significance of the interaction of training and South.

Variable definitions and pooling

To test the robustness of our results, we have estimated our models with different definitions of employment and of training, and on sub-samples. The main findings reported in the online Appendix B are summarized here below:

(1) the South-North difference in the effectiveness of training on continued employment is robust to alternative specifications of the dependent variable (i.e. including or omitting the self-employed, and coding full-time to part-time transitions and workers whose continued

employment is publicly subsidized (Cassa Integrazione) alternately as good or bad outcomes);

(2) separating public and private sectors: the estimated South-North difference in training effectiveness is larger in the private sector than in the public; in the public sector, the difference is not statistically significant. From this we conclude that the difference in training effectiveness is not an artefact of differences in public sector employment practice between South and North.

(3) estimating for training in the 2010 survey only (excluding 2008): more recent training has a stronger effect on continued employment. The South-North difference in the effect is also larger for 2010 training than for 2008 training, and is not statistically significant for 2008 training;

(4) distinguishing training paid for by the worker versus somebody else (i.e. the employer or the state): training paid for by the worker has a somewhat stronger effect on continued employment; controlling for self-payment in a pooled regression increases the estimated South-North difference in the training effect but also increases the standard error, and so makes the South-North difference statistically insignificant;

(5) using a broader definition of the training variable: the inclusion of short forms of training increases its overall effect on employment, but reduces the South-North difference in that effect, which however remains statistically significant.

Conclusion

This paper has shown that the effect of training on continued employment differs notably across sub-national geography. In the South of Italy, training has a substantially larger effect on the probability of remaining employed through the financial crisis and its aftermath, than

it does in the Northern regions. The effect is not explained by differences, across Italy's twenty regions, in initial unemployment, the change in unemployment, or GDP per capita, all of which we control for. It is most pronounced in the private sector, but a South-North difference is evident in the public sector as well. It remains strong when we instrument training to correct for endogeneity; and it remains so when we add further region-level controls for industry structure, quality of education, and social capital.

To understand any of these differences we will need to know more about the local demand side of the labour market. Literature on industry structure and inter-firm relations tells us that Southern firms internalise more functions, have more limited external networks, and are less likely to be located in specialised clusters or industrial districts. The broad literature on industrial dynamics and agglomeration economies suggests that this is not a good recipe for innovation or for job creation; but, since it also describes an environment in which inter-firm job mobility is likely to be lower as it is the threat of poaching, this may be more conducive to training by employers, or to a stronger association between such training and remaining with that same employer. It may be speculated that, under such circumstances, provision of training might be more firm-specific in southern firms, making the costs of dismissal relatively higher.

This does suggest, much as Shuttleworth et al. (2005) concluded in the case of the UK, the need for a geographical approach to labour market policy. Where demand is low and inter-firm mobility is slight, it is difficult to maintain a pool of skilled labour: skills obtained outside the firm, notably in universities, tend to leave the area, while those obtained inside the firm tend to stay put in that same firm. Under such conditions, an effective skills policy must either go together with the easing of inter-firm barriers and the emergence of specialised agglomerations, or work with employers to support their internal needs (Giunta et al., 2012) and technological upgrading (Goldstein et al., 2012). There is some hope that the latter may

aid the former: Klepper (2011) and Feldman (2014), among others, show how many large specialized agglomerations have grown from single successful firms that provided good training and inadvertently fostered the creation of competitors, staffed by their own ex-employees. Yet, hoping to see such an outcome flowing from one or more large employers in the South of Italy may be vain: the culture and other informal institutions of the Mezzogiorno are known to be unfriendly to it, and some of the competing regions benefit not simply from Marshallian pools of skilled labour, but from inter-firm cooperation in training which actively feeds these pools.

This, however, is speculative. Before we can draw definite policy conclusions from our findings, we must understand what drives them, and specifically how they are connected to sub-national variation in industry structure, informal institutions and the effectiveness of formal institutions. We are in the process of assembling a dataset which has finer geographical resolution and a better representation of the demand side of the labour market, and with the aid of that we hope to be able to shed more light on the phenomenon documented in this paper.

Acknowledgements

The authors would like to thank for their helpful comments and suggestions three anonymous referees, and the participants to: the Workshop 'University-Industry Linkages: From Theory to Policy', University of Rome Tre, 26-28 May 2014; the 2014 Eu-SPRI Conference, Manchester Institute of Innovation Research, 18-20 June 2014; the 'Scientific Labour Markets and Innovation Systems' Workshop, Birkbeck Centre for Innovation Management Research (CIMR), Birkbeck University of London, July 4, 2014; the ERSA 54th Congress, Saint Petersburg, Russia, 26-29 August 2014; the Work-in-Progress Seminar, Department of Geography & Environment, LSE, 20 May 2015; and the Italian Association of Regional Science (AISRe), Cagliari (CA), 20-22 September 2017. We are also grateful to Ron Smith for econometric advice, though any errors of course remain ours.

Andrea Filippetti gratefully acknowledges the Marie Curie Intra-European Fellowship project FP7-PEOPLE-2011-IEF-298167-REGIO_SPIN, under the EC Grant Agreement No: PIEF-GA-2011-298167.

The authors are solely responsible for any errors contained in the paper.

REFERENCES

Acemoglu, D. and Pischke, J. (1999) The structure of wages and investment in general training, *Journal of Political Economy*, 107, 539–572.

Angrist, J. D. & J.-S. Pischke. 2009. *Mostly Harmless Econometrics*. Princeton: Princeton University Press.

Arulampalam, W., Booth, A. L. and Bryan, M. L. (2004) Training in Europe, *Journal of the European Economic Association*, 2, 346–360.

Associazione Bruno Trentin - CGIL (2015) Gli effetti della crisi sul lavoro in Italia, https://www.fondazionedivittorio.it/sites/default/files/content-attachment/TerzoTrimestre2014_rev.pdf

Bailey, D. and De Propis, L. (2014) Editorial: Recession, Recovery and Resilience?, *Regional Studies*, 48:11, 1757-1760.

Bassanini, A. and Ok, W. (2004) How Do Firms' and Individuals' Incentives to Invest in Human Capital Vary across Groups?, Paper presented at the joint EC-OECD conference on 'Human Capital and labour market performance'.

Becattini, G. (1979) Dal settore industriale al distretto industriale: alcune considerazioni sull'unità di indagine dell'economia industriale, *Rivista di Economia e Politica Industriale*, 1, 7-21.

Bishop, J. H. (1996) What We Know About Employer-Provided Training: A Review of Literature, CAHRS Working Paper #96-09, Ithaca, NY: Cornell University, School of Industrial and Labor Relations.

Blundell, R., Dearden, L., Meghir, C. (1996) The Determinants and Effects of Work-Related Training in Britain, The Institute of Fiscal Studies, April.

Bratti, M., Checchi, D. and Filippin, A. (2007) Geographical differences in Italian students' mathematical competencies: Evidence from PISA 2003, *Giornale degli Economisti e Annali di Economia,* 66, 3, 299-333.

Brunello, G. (2001) On the Complementarity between Education and Training in Europe, IZA Discussion Paper 309.

Brunello, G. (2009) The Effect of Economic Downturns on Apprenticeships and Initial Workplace Training: A Review of the Evidence, *Empirical Research in Vocational Education and Training*, 1, 145–171.

Brunello, G. and De Paola, M. (2004) Training and the Density of Economic Activity: Evidence from Italy, IZA Discussion Paper No. 1173, July.

Brunello, G., Lupi, C. and Ordine, P. (2001) Widening differences in Italian regional unemployment, *Labour Economics*, 8, 103–129.

Budría, S. and Telhado-Pereira, P. (2009) The Contribution of Vocational Training to Employment, Job-Related Skills and Productivity: Evidence from Madeira, *International Journal of Training and Development,* 13, 53–72.

Cellini, R. and Torrisi, G. (2014) Regional Resilience in Italy: A Very Long-Run Analysis. *Regional Studies,* 48**,** 1779-1796.

Centro Studi Confindustria (2017) *Innovazione: effetti sull'industria and performance*, No. 8, November, SIPI Editor, Rome.

Colombo, E. and Stanca, L. (2008) The Impact of Training on Productivity: Evidence from a Large Panel of Firms, Department of Economics, University of Milan-Bicocca, Working Paper 134, January.

Conti, G. (2005) Training, Productivity and Wages in Italy, *Labour Economics*, 12, 557-576.

Contini, B. and Trivellato, U. (Eds) (2005) *Eppur si muove. Dinamiche e persistenze nel mercato del lavoro italiano*, Il Mulino, Bologna.

Crescenzi, R., Gagliardi, L. and Percoco, M. (2013) Social capital and the innovative performance of Italian provinces, *Environment and Planning A*, 45 (4): 908-929.

D'Agostino, G. and Scarlato, M. (2015) Innovation, Socio-institutional Conditions and Economic Growth in the Italian Regions, *Regional Studies*, 49:9, 1514-1534.

D'Antonio, M. and Scarlato, M. (2007) I Laureati del Mezzogiorno: una Risorsa Sottoutilizzata o Dispersa, Quaderni SVIMEZ, SVIMEZ, Roma.

De Blasio, G. and Nuzzo, G. (2010) Historical traditions of civicness and local economic development, *Journal of Regional Science*, 50:4, 833–857.

Dearden, L., Reed, H. and Van Reenen, J. (2006) The Impact of Training on Productivity and Wages: Evidence from British Panel Data, *Oxford Bulletin of Economics and Statistics*, 68, 397–421.

Dias, J. and Tebaldi, E. (2012) Institutions, human capital, and growth: the institutional mechanism, *Structural Change and Economic Dynamics*, 23, 300–312.

Dunford, M. and Greco, L. (2005) *After the Three Italies: Wealth, Inequality and Industrial Change*, Blackwell Publishing, Oxford.

Etzo, I. (2011) The determinants of the recent interregional migration flows in Italy: a panel data analysis, *Journal of Regional Science*, 51, 948–966.

European Commission (2013) *The urban and regional dimension of the crisis*, Eighth progress report on economic, social and territorial cohesion, Luxembourg, European Union.

Faggian, A., Gemmiti, R., Jaquet, T. and Santini, I. (2017) Regional economic resilience: the experience of the Italian local labour systems, *Annals of Regional Science*, DOI 10.1007/s00168-017-0822-9

Feldman, M. (2014) The character of innovative places: entrepreneurial strategy, economic development and prosperity, *Small Business Economics*, 43:1, 9-20.

Filippetti, A. and Guy, F. (2015) Skills and social insurance: evidence from the relative persistence of innovation during the financial crisis in Europe, *Science and Public Policy*, 43**,** 505-517.

Fingleton, B., Garretsen, H. and Martin, R. (2015) Shocking aspects of monetary union: the vulnerability of regions in Euroland, *Journal of Economic Geography*, 15, 907–934.

Gardiner, B., Martin, R., Sunley, P. and Tyler, P. (2013) Spatially unbalanced growth in the British economy, *Journal of Economic Geography*, 13, 889–928.

Gianmatteo, M. (2009) L'indagine campionaria Isfol-Plus: contenuti metodologici e implementazione, Studi Isfol 2009/3, Collana Studi ISFOL: Rome.

Giunta, A., A. Nifo & D. Scalera (2012) Subcontracting in Italian Industry: Labour Division, Firm Growth and the North–South Divide, *Regional Studies*, 46**,** 1067-1083.

Goldstein, H.A., Lowe, N. and Donegan, M. (2012) Transitioning to the New Economy: Individual, Regional and Intermediation Influences on Workforce Retraining Outcomes, *Regional Studies*, 46:1, 105-118.

Green, A.E. and Livanos, I. (2015) Involuntary Non-Standard Employment and the Economic Crisis: Regional Insights from the UK, *Regional Studies,* 49:7, 1223-1235.

Heckman, J. J. (1999) Doing It Right: Job Training and Education, *Public Interest*, 135, 86–107.

Hughes, G., O'Connell, P. J. and Williams, J. (2004) Company Training and Low-Skill Consumer-Service Jobs in Ireland, *International Journal of Manpower*, 25, 17–35.

Iammarino, S. (2005) An evolutionary integrated view of regional systems of innovation. Concepts, measures and historical perspectives, *European Planning Studies*, 13(4), 495–517.

ISFOL (2014) *Rapporto di Monitoraggio del Mercato del Lavoro*, ISFOL, Rome, http://www.isfol.it/primo-piano/diffuso-il-rapporto-di-monitoraggio-isfol-2014

ISTAT (2014) *Rapporto noiitalia*, ISTAT, Rome.

Iuzzolino, G., Pellegrini, G. and Viesti, G. (2013) 'Regional convergence', G. Toniolo (ed.), *The Oxford Handbook of the Italian Economy Since Unification*, 571-598, Oxford University Press, Oxford.

Kamnungwut, W., and Guy, F. (2012) Knowledge in the air and cooperation among firms: traditions of secrecy and the reluctant emergence of specialization in the ceramic manufacturing district of Lampang, Thailand, *Environment and Planning A*, 44 (7):1679 – 1695.

Kasabov, E. and Sundaram, U. (2016) Conceptualizing Clusters as Dynamic and Path-Dependent Pools of Skills, *Regional Studies*, 50**,** 1520-1536.

Klepper, S. (2011) Nano-economics, spinoffs, and the wealth of regions, *Small Business Economics*, 37 (2):141-154.

Kuckulenz, A. and Zwick, T. (2003) The Impact of Training on Earnings: Differences between Participant Groups and Training Forms, ZEW Discussion Paper 03-57.

Lazerson, M.H. and Lorenzoni, G. (1999) The firms that feed industrial districts: A return to the Italian source, *Industrial and Corporate Change*, 8 (2), 235-266.

Leuven, E. (2005) The Economics of Private Sector Training: A Survey of the Literature, *Journal of Economic Surveys*, 19, 91–111.

Lynch, L. M. and Black, S. E. (1998) Beyond the Incidence of Employer-Provided Training, *Indus. & Lab. Rel. Rev.*, 52, 64-81.

Mandrone, E, Corsetti, G and Spizzichino, A. (2014) L'INDAGINE ISFOL PLUS, *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXVIII no. 3/4 July-December.

Mandrone, E. and Radicchia, D. (2012) Indagine Plus : il mondo del lavoro tra forma e sostanza : terza annualità / ISFOL - 475 p. I libri del Fondo sociale europeo, Rubbettino.

Marinelli, E. (2013) Sub-national graduate mobility and knowledge flows: An exploratory analysis of onward-and return-migrants in Italy, *Regional Studies*, 47 (10), 1618-1633.

Martin, R. (2011) The Local Geographies of the Financial Crisis: From the Housing Bubble to Economic Recession and beyond, *Journal of Economic Geography*, 11, 587–618.

Martin, R. (2012) Regional economic resilience, hysteresis and recessionary shocks, *Journal of Economic Geography*, 12, 1–32.

Mauro, L. and Pigliaru, F. (2011) Social Capital, Institutions and Growth: Further Lessons from the Italian Regional Divide, Centro Ricerche Economiche Nord Sud-CRENoS Working Paper 3, University of Cagliari.

McCann, P. (2016) *The UK Regional-National Economic Problem*, Routledge, Abington UK, and New York.

Meliciani, V. and Radicchia, D. (2011) The informal recruitment channel and the quality of job-worker matches: an analysis on Italian survey data, *Industrial and Corporate Change*, 20 (2): 511-554.

Meliciani, V. and Radicchia, D. (2016) Informal networks, spatial mobility and overeducation in the Italian labour market, *Annals of Regional Science*, 56:2, 513-535.

Mussida, C. and Pastore, F. (2015) Worker turnover across Italian regions, in C. Mussida and F. Pastore (eds.) *Geographical labour market imbalances*, Springer, Berlin.

Nifo, A. and Vecchione, G. (2014) Do Institutions Play a Role in Skilled Migration? The Case of Italy. *Regional Studies*, 48**,** 1628-1649.

Ok, W. and Tergeist, P. (2003) Improving Workers' Skills: Analytical Evidence and the Role of the Social Partners, OECD Publishing, Paris.

Paci, R. and Usai, S. (1999) Externalities, knowledge spillovers and the spatial distribution of innovation, *GeoJournal*, 49: 381.

Paci, R. and Usai, S. (2000) The Role of Specialisation and Diversity Externalities in the Agglomeration of Innovative Activities, *Rivista Italiana degli Economisti*, 2, 237-268.

Passaro, R. (1994) Le strategie competitive delle piccole impress di una area interna del Mezzogirono, *Piccola Impresa / Small Business*, 3, 85-112.

Peck, J.A. and Haughton, G.F. (1991) Youth training and the local reconstruction of skill: evidence from the engineering industry of North West England, 1981-88, *Environment and Planning A*, 23, 813-832.

Piras, R. (2005) Il contenuto di capitale umano dei flussi migratori interregionali: 1980-2002, *Politica economica*, 21, 461-491

Putnam, R. D. (1993) *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton, Princeton University Press.

Pyke, F., Becattini, G. and Sengenberger, W. (eds) (1990) *Industrial Districts and Inter-Firm Co-operation in Italy*, ILO, Geneva.

Quatraro, F. (2009) Diffusion of Regional Innovation Capabilities: Evidence from Italian Patent Data, *Regional Studies*, 43**,** 1333-1348.

Rutherford, T.D. (2001) The state of training: learning, institutional innovation, and local boards for training and adjustment in Ontario, Canada, *Environment and Planning A*, 33, 1871-1891.

Sanders, J. and de Grip, A. (2004) Training, Task Flexibility and the Employability of Low-Skilled Workers, *International Journal of Manpower*, 25, 73–89.

Sepulveda, F. (2002) *Training and Business Cycles*, Michigan State University.

Shuttleworth, I., Tyler, P. and McKinstry, D. (2005) Redundancy, readjustment, and employability: what can we learn from the 2000 Harland &Wolff redundancy?, *Environment and Planning A*, 37, 1651-1668.

Sunley, P., Martin, R. and Nativel, C. (2001) Mapping the New Deal: local disparities in the performance of welfare-to-work, *Transactions of the Institute of British Geographers*, 26, 484-512.

Tabellini, G. (2010) Culture and Institutions: Economic Development in the Regions of Europe, *Journal of the European Economic Association*, 8: 4, 360-432.

Teselli, A. (Ed.) (2016) *Formazione professionale e politiche attive del lavoro*, Carocci editore, Rome.

Trigilia, C. and Burroni, L. (2009) Italy: rise, decline and restructuring of a regionalized capitalism, *Economy and Society*, 38:4, 630-653.

Turok, I., Bailey, D., Clark, J., Du, J., Fratesi, U., Fritsch, M., Harrison, M., Kemeny, T., Kogler, D., Lagendijk, A., Mickiewicz, T., Miguelez, E., Usai, S. and Wishlade, F. (2017) Global reversal, regional revival?, *Regional Studies*, 51:1, 1-8.

Turok, I. and Edge, N. (1999) *The Jobs Gap in Britain's Cities: Employment Loss and Labour Market Consequences*, Policy Press, Bristol.

Van Ours, J. C. and Ridder, G. (1995) Job Matching and Job Competition: Are Lower Educated Workers at the Back of Job Queues?. *European Economic Review*, 39, 1717–1731.

Wixe, S. (2015) The Impact of Spatial Externalities: Skills, Education and Plant Productivity. *Regional Studies*, 49**,** 2053-2069.

Wooldridge, J.M. (2010) *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

Table 1 - Baseline probit estimates - dependent variable: remain employed

	1A	1B	1C	1D
training	0.458***	0.448***	0.363***	0.260**
	(0.122)	(0.124)	(0.108)	(0.0998)
South		0.0746	0.111	0.0802
		(0.313)	(0.295)	(0.305)
unemployment 2008		0.00571	-0.00308	-0.0110
		(0.0438)	(0.0413)	(0.0383)
change unemp. 2008-11		0.0817	0.0873	0.0759
		(0.0609)	(0.0552)	(0.0477)
regional GDPpc 2008		0.710	0.673	0.451
		(0.551)	(0.456)	(0.503)
medium-educated			0.431***	0.264**
			(0.117)	(0.0960)
high-educated			1.034***	0.847***
			(0.141)	(0.150)
age 25-29			-0.0868	-0.242
			(0.129)	(0.150)
age 30-39			0.00305	-0.171
			(0.207)	(0.193)
age 40-49			-0.0491	-0.291
			(0.280)	(0.283)
age 50-64			-0.0151	-0.301
			(0.327)	(0.341)
male			0.333**	0.410**
			(0.122)	(0.130)
top-grade			-0.236	-0.246
			(0.182)	(0.192)
work experience			0.0260*	0.0244*
			(0.0108)	(0.0101)
health			-0.0527	-0.0395
			(0.0986)	(0.106)
life satisfaction			0.0752	0.0718*
			(0.0409)	(0.0350)
children			-0.0488	-0.166
			(0.144)	(0.138)
foreign			0.948***	1.085***
			(0.284)	(0.290)
father_job: mixed cognitive			-0.0974	-0.0593
			(0.120)	(0.106)
father_job: intellectual			-0.177	-0.113
			(0.0948)	(0.0927)
mother_edu.: medium-educated			0.0652	0.0416
			(0.116)	(0.114)
mother_edu.: high-educated			-0.0106	-0.0106
			(0.191)	(0.169)
urban			0.0747	0.0259
			(0.135)	(0.120)
permanent contract				0.542***
				(0.101)
part-time in 2008				0.372**
				(0.115)
changed job				-0.210
				(0.166)
moved for job				0.104
				(0.174)
employed in public 2008				0.814***
				(0.116)
employed in medium firm				0.359**
				(0.121)
employed in large firm				0.504***
				(0.153)
industry dummies	no	no	no	yes
Constant	1.034***	-6.546	-7.389	-5.331
	(0.0711)	(5.930)	(4.846)	(5.308)
N	3983	3983	3983	3983
Pseudo R^2	0.018	0.030	0.090	0.150

Note: regional clustered standard errors, in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Table 2 - Interaction probit estimates: dependent variable: remain employed

	Model 2	Model 3
training	0.15	0.0071
	(0.12)	(0.26)
South	-0.0029	-0.27
	(0.31)	(0.33)
training x South	0.43**	0.24
	(0.16)	(0.40)
medium-educated	0.26**	0.074
	(0.095)	(0.088)
high-educated	0.84***	0.80***
	(0.15)	(0.17)
training x medium-educated		0.23
		(0.23)
training x high-educated		0.25
		(0.37)
South x medium-educated		0.54***
		(0.11)
South x high-educated		-0.0073
		(0.19)
training x South x medium-educated		0.015
		(0.45)
training x South x high-educated		0.35
		(0.48)
N	3983	3983
Pseudo R-squared	0.152	0.159

Marginal effects of training on probability of remaining employed, at means of covariates

	Model 2	Model 3
all workers, North	0.0221	
	(0.0169)	
all workers, South	0.103***	
	(0.0144)	
South-North difference, all workers	0.081***	
	(0.022)	
low-educated, North		0.0014
		(0.053)
low-educated, South		0.071
		(0.076)
South-North difference, low-educated		0.069
		(0.095)
medium-educated, North		0.039**
		(0.013)
medium-educated, South		0.079***
		(0.017)
South-North difference, medium-educated		0.040
		(0.023)
high-educated, North		0.015
		(0.012)
high-educated, South		0.091**
		(0.031)
South-North difference, high-educated		0.076*
		(0.033)

Note: industry dummies included; regional clustered standard errors
* p<0.05, ** p<0.01, *** p<0.001

Table 3 - Control function estimates

Dependent variable	**Training** First stage linear probability model	**Remain employed** Control function estimates (Model 2 + 1st stage residuals)	**Remain employed** Model 2 (drop 1st stage residuals)
ALMP	0.31**		
	(0.083)		
Stage 1 residual		1.47	
		(1.05)	
training		-1.30	0.15
		(1.04)	(0.12)
South	0.0046	-0.0035	-0.0040
	(0.050)	(0.32)	(0.31)
training x South		0.42*	0.43*
		(0.17)	(0.17)
	(All controls included)		
N	3983	3983	3983
R-squared	0.128		
Pseudo R sq		0.153	0.152
South-North difference in effect of training		0.076	0.083***
		(0.044)	(0.025)

Standard errors (clustering by region) in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Kleibergen-Papp statistic on the first stage:

(1) ALMP = 0; F(1, 18) = 13.80; Prob > F = 0.0016

Table 4 – Regressions with additional regional controls – dependent variable: remain employed

Panel A	(1)	(2)	(3)	(4)	(5)
VARIABLES	TShare	NEET	TechSch	PISA	PIAAC
training	0.148	0.153	0.152	0.151	0.148
	(0.125)	(0.120)	(0.119)	(0.124)	(0.125)
South	-0.040	-0.004	0.090	0.017	0.019
	(0.291)	(0.284)	(0.286)	(0.312)	(0.302)
training	0.433***	0.427***	0.429***	0.431***	0.434***
x South	(0.167)	(0.164)	(0.163)	(0.166)	(0.168)
additional regional	0.047	-1.572	-0.000	-0.001	0.003
Var.†	(0.070)	(1.474)	(0.000)	(0.003)	(0.006)

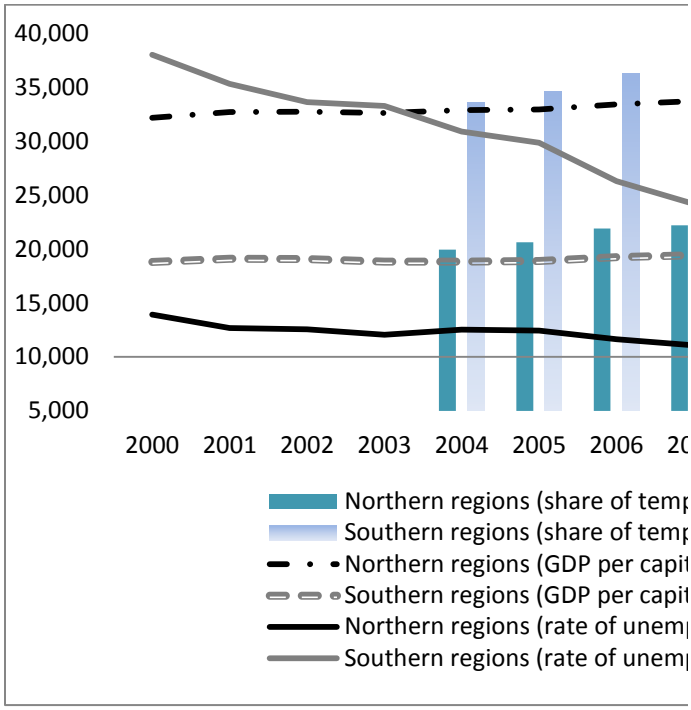
Panel B	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	Trust	Respect	Tolerance	Agri	Mfg	Service
training	0.151	0.151	0.148	0.150	0.159	0.159
	(0.124)	(0.126)	(0.122)	(0.125)	(0.118)	(0.119)
South	-0.008	0.065	-0.065	0.061	0.279	0.464
	(0.311)	(0.456)	(0.308)	(0.333)	(0.329)	(0.353)
training	0.430***	0.432**	0.433***	0.433***	0.421***	0.422***
x South	(0.165)	(0.169)	(0.164)	(0.166)	(0.162)	(0.162)
additional	0.001	0.002	0.005	-1.629	-1.978**	2.567***
regional var†	(0.008)	(0.010)	(0.013)	(2.565)	(0.891)	(0.953)

N=3,983. †: reports the coefficient for the additional regional variables specific to the column (e.g. TShare; NEET, etc.) Regional clustered errors in parentheses; all controls included
* p<0.05, ** p<0.01, *** p<0.001

Appendix A

FIGURES

Figure A1 - Gross domestic product (GDP) at market prices per inhabitant, left axis (2000-2011); Unemployment rate, right axis (2000-2011); share of temporary workers, right axis (2004-2011)

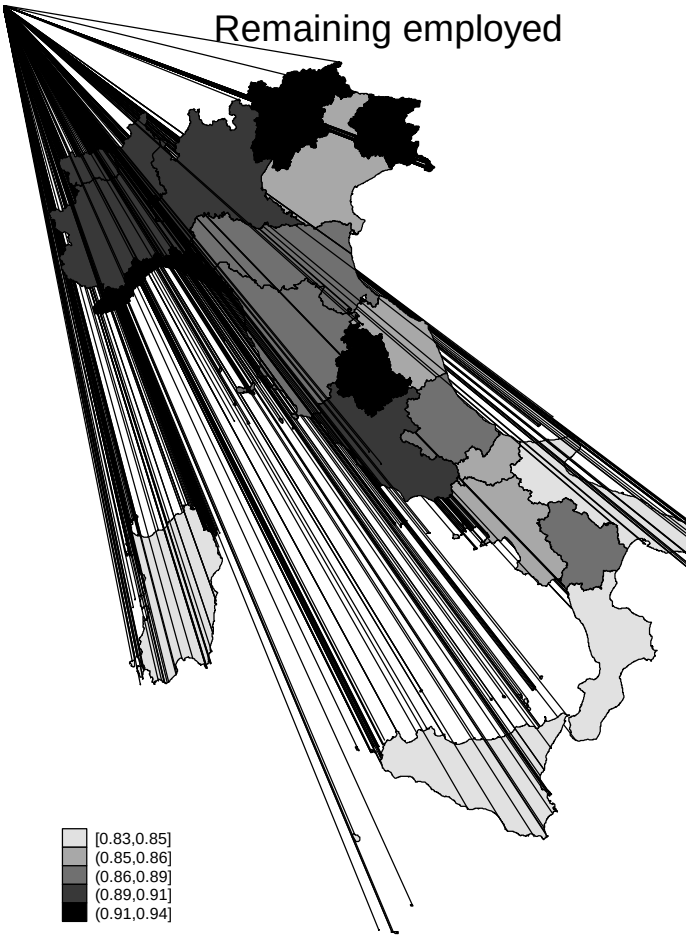


Source: ISTAT data warehouse, and Labour Force Survey

Figure A2 – Regional unemployment rates in 2011

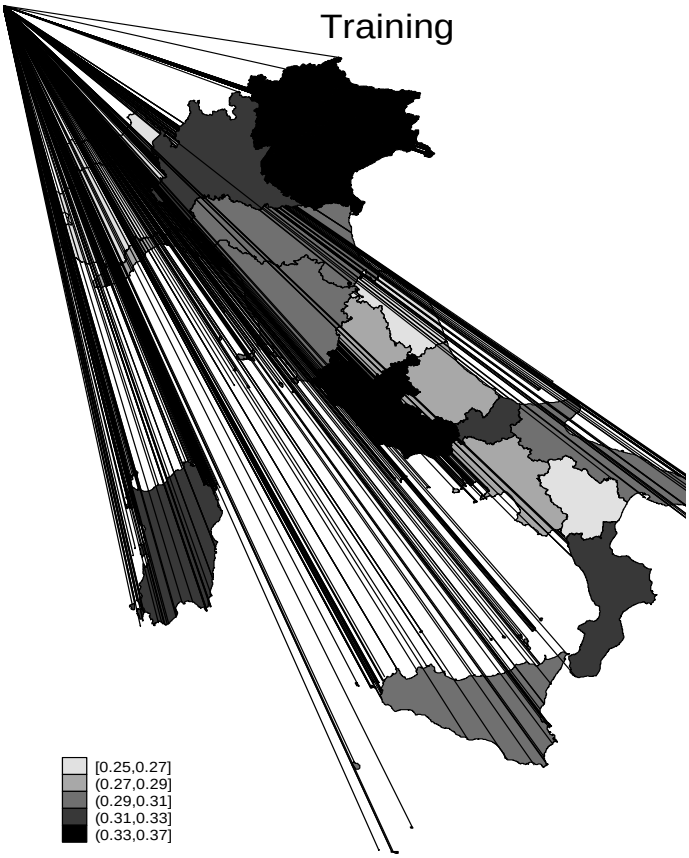
Source: Istat data warehouse

Figure A3 – Remaining employed over the period 2008-2011, regional averages



Source: PLUS Survey

Figure A4 – Participation in training while employed over the period 2008-2010, regional averages



Source: PLUS Survey

TABLES

Table A1 - Variable List

Name/acronym	Description	Type	Source
	Dependent variable		
Remain employed (y)	1=employed over the period 2008-2011; 0=employed over the period 2008-2010 and unemployed (or shifted from full-time to part-time) in 2011	dummy	PLUS Q #D6bis
	Training while employed		
Training	1= trained while employed in the 2008-2010 period; 0=otherwise. Includes training in the classroom, distant learning, on-the-job; excludes short-term forms of training such as seminars, conferences and workshops	dummy	PLUS Q #D175
	Regional characteristics		
Unemployment 2008	regional rate of unemployment, 2008	continuous	ISTAT
Change in unemployment 2008-11	regional change in unemployment over 2008-2011	continuous	ISTAT
Regional GDPpc 2008	regional GDP per capita, 2008	continuous	ISTAT
	Individual characteristics		
South	1=living in a Southern region; 0=otherwise North includes: Piemonte, Val D'Aosta, Lombardia, Trentino Alto Adige, Veneto, Friuli Venezia Giulia, Liguria, Emilia Romagna; Toscana, Umbria, Marche, Lazio; South includes: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna.	dummy	PLUS
Education: low-educated, medium-educated, high-educated	low-educated (primary and middle school, omitted category); medium-educated (high school); high-educated (university or above)	dummies	PLUS Q #D.88
Age	Years; classes in estimates: 18-24 (omitted category); 25-29; 30-39; 40-39; 50-64	dummies	PLUS Q #D.cleta5_1
Sex	1=male; 0=female	dummy	PLUS Q #D. 8
Top-grade in the last education level	1=top-grade achiever; 0=otherwise	dummy	PLUS Q #D.89
Work experience	number of years since started working	continuous	PLUS Q #D.90
Health	1=no smoking and doing sport; 0=otherwise	dummy	PLUS Q #D.109_2
Life satisfaction	composite indicator built on 5 dimensions of life satisfaction each ranked from 1 (low) to 7 (high)	continuous	PLUS Q #D.183ter
Children	1=has children; 0=otherwise	dummy	PLUS Q #D.97bis
Foreign	1=foreign (non-Italian) citizenship; 0=otherwise		PLUS Q #D.110
Father_job	father's occupation: manual (omitted category); mixed cognitive; intellectual and scientific*	dummies	PLUS Q #D.105bis
Mother_edu	mother's education (see education above)	dummies	PLUS Q #D.105b
Urban	1=living in an urban area; 0=otherwise	dummy	PLUS Q #D.v9142

	Job characteristics		
Permanent contract	1=permanent contract; 0=otherwise	dummy	PLUS Q #D.11bis_2
Part-time in 2008	1=part-time contract in 2008; 0=otherwise	dummy	PLUS Q #D.14
Changed job	1=changed job during 2008-2011; 0=otherwise	dummy	PLUS Q #D.46
Moved for job	1=moved in a different province for job during 2008-2011; 0=otherwise	dummy	PLUS Q #D.37
Employed in public sector 2008	1=employed in public sector in 2008; 0=otherwise	dummy	PLUS Q #D.39
Size of firm	employed in: large firm (>499); medium firm (10-499); small firm (<10)	dummies	PLUS Q #D.40
Employer industry (ATECO ISTAT)	15 industries	dummy	PLUS Q #D.sett15_n_11
Additional regional controls			
Training and skills	regional share of firms doing on-the-job training	continuous	ISTAT
	region's share of young NEET (not in education, employment or training) residents 2008-11	continuous	ISTAT
	share of regional technical hi-schools	continuous	ISTAT
	results of PISA test	continuous	ISTAT
	results of PIAAC - Survey of Adult Skills - test	continuous	OECD
Culture and institutions	trust (percentage of respondents in each region answering "most people can be trusted")	continuous	Tabellini
	tolerance (percentage of respondents in each region indicating "tolerance for other people" as important)	continuous	Tabellini
	respect (percentage of respondents in each region indicating "respect for other people" as important)	continuous	Tabellini
Regional specialization	regional share of agriculture # of workers	continuous	ISTAT
	regional share of manufacture # of workers	continuous	ISTAT
	regional share of services # of workers	continuous	ISTAT

Notes: * This classification was obtained by grouping ISTAT occupational categories reported in PLUS

Table A2 – Summary statistics for the main variables used in the analysis

		participation in training		T-test	geographical area		T-test
	Total	**no**	**yes**		**North**	**South**	
REGIONAL CHARACTERISTICS (based on mean values for nineteen regions)							
unemployment 2008	6.7	6.73	6.64	N	4.4	11.2	Y***
	3.59	*3.59*	*3.6*		*1.21*	*2.09*	
change in unemployment 2008-11	1.58	1.58	1.57	N	1.65	1.45	Y***
	0.61	*0.62*	*0.6*		*0.46*	*0.82*	
regional GDPpc 2008 (log)	10.4	10.4	10.4	N	10.6	10.1	Y***
	0.25	*0.25*	*0.25*		*0.091*	*0.1*	
ALMP (IV)	0.63	0.64	0.64	N	0.71	0.48	Y***
	0.18	*0.18*	*0.18*		*0.16*	*0.12*	
INDIVIDUAL CHARACTERISTICS							
remain employed(y)	0.89	0.86	0.94	Y***	0.9	0.86	Y***
	0.32	*0.35*	*0.23*		*0.3*	*0.35*	
training	0.31	0	1		0.32	0.31	N
	0.46	*0*	*0*		*0.47*	*0.46*	
age	41.3	40.6	42.9	Y***	40.3	43.3	Y***
	12.4	*12.5*	*12*		*12.1*	*12.7*	
sex	0.49	0.51	0.46	Y***	0.47	0.54	Y***
	0.5	*0.5*	*0.5*		*0.5*	*0.5*	
low-educated	0.18	0.22	0.095	Y***	0.19	0.16	Y***
	0.39	*0.42*	*0.29*		*0.4*	*0.37*	
medium-educated	0.52	0.52	0.52	N	0.52	0.52	N
	0.5	*0.5*	*0.5*		*0.5*	*0.5*	
high-educated	0.29	0.25	0.38	Y***	0.28	0.32	Y***
	0.46	*0.44*	*0.49*		*0.45*	*0.46*	
top-grade	0.18	0.16	0.24	Y***	0.17	0.2	Y***
	0.39	*0.36*	*0.43*		*0.38*	*0.4*	
work experience	19.5	19.1	20.2	Y***	19.1	20.2	Y***
	12.5	*12.7*	*12*		*12.4*	*12.6*	
health	0.15	0.14	0.17	Y***	0.17	0.11	Y***
	0.36	*0.35*	*0.38*		*0.38*	*0.31*	
life satisfaction	5.22	5.22	5.22	N	5.21	5.24	Y***
	1.11	*1.13*	*1.06*		*1.09*	*1.15*	
children	0.58	0.57	0.62	Y***	0.57	0.62	Y***
	0.49	*0.5*	*0.49*		*0.5*	*0.49*	
foreign	0.009	0.01	0.006	Y***	0.013	0.0007	Y***
	0.095	*0.1*	*0.08*		*0.11*	*0.027*	
father job: manual	0.64	0.66	0.6	Y***	0.64	0.64	N
	0.48	*0.47*	*0.49*		*0.48*	*0.48*	
father job: mixed cognitive	0.23	0.22	0.25	Y***	0.23	0.23	N
	0.42	*0.41*	*0.43*		*0.42*	*0.42*	
father job: intellectual & scientific	0.13	0.13	0.14	N	0.13	0.13	N
	0.34	*0.33*	*0.35*		*0.34*	*0.34*	

mother education: low-educated	0.54	0.55	0.51	Y**	0.5	0.6	Y***
	0.5	*0.5*	*0.5*		*0.5*	*0.49*	
mother education: medium-educated	0.25	0.25	0.26	N	0.27	0.23	Y***
	0.44	*0.44*	*0.44*		*0.44*	*0.42*	
mother education: high-educated	0.21	0.2	0.23	Y***	0.23	0.18	Y***
	0.41	*0.4*	*0.42*		*0.42*	*0.38*	
urban	0.29	0.29	0.29	N	0.31	0.25	Y***
	0.46	*0.46*	*0.46*		*0.46*	*0.44*	
JOB CHARACTERISTICS							
permanent contract	0.81	0.79	0.86	Y***	0.81	0.82	N
	0.39	*0.41*	*0.34*		*0.39*	*0.38*	
part-time in 2008	0.16	0.18	0.11	Y***	0.17	0.13	Y***
	0.37	*0.39*	*0.31*		*0.38*	*0.34*	
changed job	0.12	0.14	0.08	Y***	0.13	0.1	Y***
	0.32	*0.34*	*0.27*		*0.33*	*0.3*	
moved for job	0.14	0.12	0.16	Y***	0.12	0.17	Y***
	0.34	*0.33*	*0.37*		*0.32*	*0.38*	
employed in public sector 2008	0.42	0.36	0.54	Y***	0.37	0.51	Y***
	0.49	*0.48*	*0.5*		*0.48*	*0.5*	
employed in small firm	0.64	0.62	0.68	Y***	0.59	0.72	Y***
	0.48	*0.49*	*0.47*		*0.49*	*0.45*	
employed in medium firm	0.31	0.33	0.26	Y***	0.35	0.23	Y***
	0.46	*0.47*	*0.44*		*0.48*	*0.42*	
employed in large firm	0.05	0.05	0.06	Y***	0.053	0.048	Y***
	0.22	*0.21*	*0.24*		*0.22*	*0.21*	

Note: standard deviations are here in italics. T-test are added to test for the difference between participation in training (Yes vs No) and between geographical area (North vs South): Y means reports of difference significant at *5%, **1%, or ***0.1% of statistical significance; N means absence of statistical difference.

Mean participation in training, by educational level and region

Regions	**Education level**		
	low-educated	**medium-educated**	**high-educated**
North	0.16	0.32	0.41
South	0.15	0.30	0.40

Appendix B

Choices in variable and population specification: general considerations. The construction of the dependent variable *remain employed* and the *training* variable both involve numerous choices by us: to include or exclude the self-employed; to treat full-time-to-part-time transitions as favourable outcomes or adverse ones; to include all forms of training while employed, or to exclude (or include) from our measure various combinations of training,. Taken together, these choices present a substantial number of researcher's degrees of freedom, or what Gelman and Loken call a 'garden of forking paths' (Simons et al., 2011; Gelman and Loken, 2013)[3]; we must be alert to the possibility that our results could be artefacts, whether inadvertent or opportunistic, of choices made when defining of these variables. Similar issues pertain to the choice to pool disparate groups of workers. We report in what follows on steps we have taken to verify that our results are not artefacts of such choices – or, to put it another way, that our results are robust to a range of plausible re-specifications of key variables and important pooling assumptions.

Construction of the dependent variable. We have treated various groups of workers differently: self-employed have been included in the sample and treated as employed; all who were inactive in 2011 have been excluded from the sample on the grounds that they are out of the labour market; and full-time to part-time transitions have been grouped with all others who have continued in employment. However, it is worth noting that despite a considerable change in the dependent variable definition (and the loss of about a quarter of our sample – the net difference from the exclusion of self-employed and inclusion of transitions to inactive), we obtain comparable results with regard to the North-South difference in the effect of training.

A similar issue arises with workers who remain employed in 2011, but are classified in that year as *Cassa Integrazione*: these are workers for whom wage payments are continued, with public subsidy, while they

[3] Simmons, J. P., Nelson L. D. and Simonsohn, U. (2011) False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science,* 22**,** 1359-1366.

Gelman, A. & E. Loken (2013) The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. Technical report, New York, NY: Department of Statistics, Columbia University.

are on temporary layoff (which may be a full layoff, or a partial layoff while continuing part-time work for full-time pay). There were no workers in *Cassa Integrazione* in 2008, and 23 in our sample in 2011. In the main paper we treated these as adverse outcomes, together with the unemployed and the transitions from full-time to part-time. Coding them instead with the employed makes a small difference in the estimates but does not affect our conclusions (see Table B1 below).

Public vs. private sector employment. Italian public sector employment is considerably more secure than private sector employment; the effect of training on continued employment may not be the same in the two sectors. The estimates reported in the main paper include a dummy for public sector employment; such a specification, however, imposes the restriction that coefficients on all other variables – including the training effect and South-North differences in that effect – are the same for public and private sector employees.

To ascertain the effect of this restriction, we run Model 2 separately for public and private sector employees; the pooled estimates are also presented for comparison. Results are shown in Table B2: probit coefficients are in the top half of the table, and marginal effects for South and North in the bottom half; the latter are easier to interpret. In none of the estimates – pooled, private, or public - is the effect of training in the North statistically distinguishable from zero. Similarly, the effect of training in the public sector is never significantly different from zero, although the South-North difference in point estimates for the public sector shrinks only slightly.

That leaves us, finally, with private sector training in the South. This has a positive effect on continued employment which is 50% larger than that in the pooled estimate, and which remains significant at the 0.01 level. In short, private sector employment is driving our results.

Does it matter when the training occurred? Our training variables are constructed from survey responses in both 2008 and 2010 (in all cases we omit training reported in 2011, the survey in which employment status is determined): anybody reporting training in the period 2008-10 was treated as trained. Does the effect of training decay with time? It seems that it does – estimates are reported in Table

B3 below. Coefficients for 2008 training are smaller in all cases; and, for 2008 training, the South-North difference in the marginal effect is no longer significant, though the simple training effect in the South remains significant at the 0.01 level (bottom part of Table B3). This result suggests that the decay, or depreciation, of training may be an important issue that would be worth future study. As with the public and private sector question, the result raises the issue of whether training in the two surveys should be treated differently and yet also confirms the qualitative differences found in the estimates.

Does it matter who pays for training? Survey respondents indicate whether they paid for their own training, or it was paid for by somebody else, such as their employer or a public training program. Self-payment may indicate differences in the worker's motivation, and also in the employer's judgement about the worker. We test this by augmenting Model 2 from the main paper with a dummy variable Worker Pays, which is equal to 1 if the worker has been trained *and* paid for the training. The variable is interacted with the South dummy to allow for differential effects between South and North. The coefficient on Worker Pays is positive, small and statistically insignificant. Other coefficients change somewhat, but the most notable effect is a rise in the standard errors. The estimated marginal effects for South and North, and the South-North difference in those effects, are very close to the Model 2 estimates, though the standard errors are larger and statistical significance is lost. These results are shown in Table B4.

What kind of training? Survey respondents report a number of different kinds of training (Q #D176 in the ISFOL-PLUS Survey). In the results reported in the paper, we have defined training as training courses in the classroom or distance learning plus on-the-job training; this definition excludes short-term forms of training, such as attendance at seminars, conferences and fairs, and workshops. In Table B5, we report estimates of Model 2 with these short forms included coded as training_all. The result is a somewhat larger (and now statistically significant) training effect in the North; a marked, and statistically significant, South-North difference remains.

Table B1 - Cassa Integrazione status

	Model 2	**Model CI**
	Cassa Integrazione treated as unemployed	***Cassa Integrazione treated as employed***
training	0.15	0.10
	(0.12)	(0.13)
South	-0.0039	-0.024
	(0.31)	(0.27)
training x South	0.43**	0.43**
	(0.17)	(0.16)
Observations	3983	3983
Pseudo R-squared	0.152	0.162
Marginal effects of training on probability of remaining employed		
North	0.022	0.014
	(0.017)	(0.017)
South	0.10***	0.088***
	(0.014)	(0.013)
South-North difference	0.080***	0.074***
	(0.022)	(0.021)

Note: all controls included; regional clustered standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table B2 - Private vs. Public Sector Employers

	pooled	private sector	public sector
training	0.15	0.25	-0.3
	(0.12)	(0.15)	(0.2)
South	-0.0	-0.11	0.4
	(0.31)	(0.31)	(0.51)
training x South	0.43**	0.46	0.62*
	(0.17)	(0.24)	(0.28)
Observations	3983	2315	1660
Pseudo R-squared	0.15	0.14	0.27
Effects of training on probability of remaining employed, at means of covariates			
North	0.024	0.043	-0.027
	(0.019)	(0.024)	(0.021)
South	0.11***	0.16***	0.027
	(0.015)	(0.035)	(0.016)
South - North difference	0.083***	0.12**	0.054
	(0.024)	(0.039)	(0.028)

Note: all controls included; regional clustered standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table B3 – Training Effect Fades Over Time

	pooled	2008	2010
training	0.15		
	(0.12)		
training 2008		0.066	
		(0.17)	
training 2010			0.18
			(0.14)
training x South	0.43**		
	(0.17)		
training 2008 x South		0.22	
		(0.23)	
training 2010 x South			0.89***
			(0.23)
South	-0.0039	0.040	0.025
	(0.31)	(0.30)	(0.32)
Observations	3983	3983	3983
Pseudo R-squared	0.15	0.15	0.152
Effects of training on probability of remaining employed, at means of covariates			
North	0.022	0.0097	0.025
	(0.017)	(0.023)	(0.018)
South	0.10***	0.054**	0.13***
	(0.014)	(0.021)	(0.015)
South - North difference	0.081***	0.04	0.11***
	(0.023)	(0.034)	(0.021)

Note: all controls included; regional clustered standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table B4 - Does it matter who pays?

	Model 2	with Worker pays
training	0.15	0.12
	(0.12)	(0.16)
South	-0.0039	0.0018
	(0.31)	(0.30)
training x South	0.43**	0.57**
	(0.17)	(0.22)
Worker pays		0.17
		(0.30)
Worker pays x South		-0.58
		(0.40)
Constant	-4.91	-4.84
	(5.46)	(5.39)
Observations	3983	3983
Pseudo R-squared	0.152	0.153
Effects of training on probability of remaining employed, at means of covariates		
North	0.022	0.018
	(0.017)	(0.023)
South	0.10***	0.12
	(0.014)	(0.022)
South - North difference	0.081***	0.10
	(0.023)	(0.031)

Note: all controls included; regional clustered standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table B5 - Two definitions of training: conservative definition (*training* in main paper) vs. *all training* (includes short-term forms: conferences, seminars, workshops, fairs and other)

	training	all training
training	0.15 (0.12)	
training_all		0.23* (0.11)
training x South	0.43** (0.17)	
training_all x South		0.30* (0.14)
South	-0.0039 (0.31)	-0.012 (0.31)
medium-educated	0.26** (0.093)	0.25** (0.094)
high-educated	0.83*** (0.15)	0.80*** (0.15)
Observations	3983	3983
Pseudo R^2	0.152	0.153
North	0.022 (0.017)	0.033* (0.015)
South	0.10*** (0.014)	0.099*** (0.018)
South-North difference	0.080*** (0.022)	0.068*** (0.018)

Note: all controls included; regional clustered standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001