Evaluating the Italian training on the job contract (cfl)

G. Tattara[*] and M. Valentini[*]
tattara@unive.it and m.valentini@unive.it

Draft – Not be quoted

Abstract

The "Contratto Formazione e Lavoro" (CFL - training on the job contract) is a fixed term contract aiming at reducing unemployment and supporting the training of young Italian workers. Apprenticeship is the traditional training contract in Italy, but eligible individuals, till the last reform, are limited to 21 years of age. CFL was introduced 20 years ago and basically extended the training contract to 29 years of age, set a definite term, 24 months, for the training period, and provided an incentive for employers to hire young people under the CFL program reducing hiring and firing costs.

Using the introduction of the CFL contract as an exogenous innovation and exploiting the reform of the program over time (e.g. artisans were guaranteed higher fiscal benefits for a longer period in respect to non artisans) the effect of CFL on the Italian labour market is discussed and evaluated. The difference in difference technique together with the matching method are used in order to establish if CFL has improved Italian young hiring chances.

The CFL impact on the occupational level is *measured* using data obtained from the social security archives where this contract is clearly identified. The dataset covers all firms and all the population employed in the private sector across the years 1982-1997 for two North-East Italian provinces, Treviso and Vicenza, and allows to study the CFL impact at the very beginning, where its effect was reasonably more marked. Our estimates support a positive CFL effect on employment at the introduction of the program, while the subsequent 1991 reform bears no clear effect. In such a way our study questions some previous results bearing no evidence of a significant positive policy effect for this program.

1. Introduction

Program evaluation is defined as the use of research to measure the effects of a program in terms of its goals, outcomes, or criteria. Policy analysis is defined as the use of any evaluative research to improve or legitimate the practical implications of a policy-oriented program. Policy evaluation has not to be confused with cost-benefit analysis that considers the costs of the intervention and the received benefits.

Behind the seemingly simple question of whether a program works are a host of other more complex questions. For example, the first question is, what is a program supposed to do? It is often difficult to answer this question in a direct way and the most difficult part of the evaluation procedure is determining whether it is the program itself that is doing something. There may be other events or processes that are really causing the outcome, or preventing the hoped for outcome. However, due to the nature of the program, many evaluations cannot determine whether the program itself, or something else, is the 'cause'.

One main reason that evaluation cannot determine causation involves self-selection. That is, people select themselves to participate in a program. For example, in a jobs training program, some firms

[*] Dipartimento di Scienze Economiche, Università di Venezia, Cannaregio 873, Venezia, Italy.

decide to participate, and others, for whatever reason, do not participate. It may be that those who do participate are those who are the most determined to increase their employment, or who have the best support resources, thus allowing more people to participate and allowing them to find a job. The firms who participate are somehow different from those who don't participate, and it may be the difference, not the program, that leads to a successful outcome for the participants, that is, increasing employment.

If programs could, somehow, use random selection, then they could determine causation. That is, if a program could randomly assign firms to participate or to not participate in the program, then, theoretically, the group of firms who participate would be the same as the group who did not participate, and an evaluation could 'rule out' other causes. These are the main evaluation problems.

The paper evaluates a particular active labour marker policy, the “Contratto Formazione e Lavoro” (CFL - Training on the Job Contract). This program began in the mid eighties and lasted until the end of nineties. Its declared goal was to reduce youth unemployment and contemporaneously improve the skills of young workers. In order to spread CFL, policy makers granted a strong fiscal rebate to firms adhering to the program, and conceived CFL as a fixed term contract, so to avoid firing costs to the firms. The benefits connected to CFL varied through time and the category of the eligible also varied. In the present study CFL is considered only in relation of the goal of promoting young employment. Other possible benefits (education etc.) are not considered.

Section 2 describes CFL, its structure and the relevant innovations through the time. Section 3 introduces the evaluation problem and the methods used in this paper. Section 4 describes the data-set and investigates through both descriptive methods and econometric analysis the CFL impact. First we evaluate the effect of CFL at his introduction, the impact effect. Second, exploiting the variations in the eligible category through time, we try to identify if CFL had an effect or not. Section 5 concludes.

2. CFL setting

CFL is fixed term contract introduced by law 863/84, with a maximum duration of 24 months, not renewable[1]. At the beginning CFL eligible were young individuals aged between 15 and 29 years[2]. The law 451/94 made the boundaries looser and eligible age is between 16 and 32 years[3]. The CFL goal is to decrease the unemployment rate between young workers and, at the same time, to train them, easy their entrance into the labour market and make their career more stable and qualified.

In order to access to CFL firms must prepare a training project, which establishes, among other things, the timing and the pattern of the training activity. Each training project must be endorsed by the Regional Commission for the Employment.

According to the Law 863/84 eligible firms are public firm and private firms, which did not record massive firings, at the application date. Law 451/94 extended CFL to liberal professions, association and research centres[4] and created two kind of CFL aiming at: 1) providing new skills (intermediate and high skills); 2) easing the hiring process through a practical work experience that fits the worker skills and capacities to the production environment and to the organisational structure of the firm. CFL of the first type extend up to 24 months, CFL of the second type up to 12 months.

[1] Basically not renewable with the same training specification.
[2] 32 years old for south of Italy and region of north with an unemployment rate higher than national average.
[3] Starting date was 19/11/93 (D.L. n.462 of 18/11/93 and D.L. n. 32 of 17/1/94) , moreover regions of south of Italy until 31/12/97 were able to raise the maximum age of the eligible.
[4] The last was included by L.196/97

Since 1991 (L. 407/90) eligible firms are requested to have hired with a tenure contract, during the two preceding years at least 50% of the terminated CFL[5]. This percentage has been risen to 60% by L. 451/94.

Policy maker, in order to spread the new contract, provided rebates on social security contributions. At the beginning the same rebate as apprenticeship was allowed, i.e. the fiscal fee was limited to 5000 lire per week-per capita (almost 2 pounds). Later on, this allowance was limited to craftsmen. The law 291/88 allowed a rebate of 50% for non craftsman, rebate which has been decreased to 25% by Law 407/90 since 1/1/91 (for firms of trade and tourist sector with less than 15 employees the rebate was maintained at 40%) (Tab. 1). In the end, Law 451/94 allowed the rebate to be maintained one year after the contract transformation in case of CFL of the first type, while in case of CFL of the second type, the rebate is applied only if the contract is transformed. Southern and high unemployment regions enjoy a full rebate.

In the early eighties another important contract runs parallel to CFL, the apprenticeship contract. Like other apprenticeship contracts in Europe it aims at training young workers in order to become qualified workers. Eligible workers are young people, between 14 and 20 years of age (for artisan firms the upper limit was extended to 29 years of age by Law 56-28/2/87). The same Law56/87 allows all firms to hire directly their workers, without inquiring the Italian Ufficio di Collocamento. Previously firms, except artisan and firms with less than 10 employees, were allowed to hire directly only 25% of their apprentices. The length of apprenticeship may not be less than 18 months and more than five years; the preceding periods concerning the same activity with other employers may be cumulated. Social security contributions limited to 5000 lire per week extend one year after the confirmation.

Tab. 1: rebate on total social security contributions by category.

Year	South, artisans, high unemployment area	Trade-tourism firms with less than 15 workers	All other firms
1/5/84-31/5/88	About 98%	About 98%	About 98%
1/6/88-23/11/90	About 98%	50%	50%
24/11/90-31/12/90	About 98%	50%	25%
1/1/91-…	About 98%	40%	25%

Last but not least the CFL entitles firms to hire directly without recourse to the Ufficio di Collocamento (the unemployment lists), an advantage that has been extended to all people hired according to the law 223/91.

3. Evaluation problem

Let D_i be a dummy variable representing the treatment status, assuming value 1 if individual i has been treated and 0 otherwise. Let Y_{0i} be the outcome variable for individual i when it is not treated, and Y_{1i} is the variable for the treated. The evaluation problem accounts to a missing data problem, because the two outcomes, Y_{0i} and Y_{1i}, are not observed together for any i. What is observable is:

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i} \tag{1}$$

where X_i represents the individual characteristics.

[5] The percentage of hirings does not take into account lay offs and firings with good cause.

The impact of the program for the *i-th* subject is $Y_{1i} - Y_{0i}$. The observables are Y_i, D_i, X_i. In the simplest of all the worlds, treatment can be assumed to be constant across individuals.
$\gamma = Y_{1i} - Y_{0i}$, for any *i,* and in such a simplified context the only source of bias is the selection bias.
Given different treatments across individuals, $\gamma_i = Y_{1i} - Y_{0i}$, treatment variations might depend on systematic differences of observables and/or of unobservable characteristics and the calculation of the relevant policy impact is not possible.

The economic literature has stated several conditions under which some specific caracteristic of the $Y_{1i} - Y_{0i}$ distribution can be identified (Heckman et al. ,1999; Blundell, Costas Dias, 2002). Broadly speaking any parameter which depends on the joint distribution of Y_{1i} and Y_{0i} cannot be estimated. But some parameters can be assumed to depend on the marginal distribution of Y_{1i} and Y_{0i} [6]:

1) The average treatment effect (*ATE*). The parameter $E(Y_1 - Y_0)$ or its conditional version $E(Y_1 - Y_0 | X)$. This is the parameter of interest if the program is assigned to all individuals or is assigned randomly to the population;
2) The average treatment effect on treated (*ATT*). The parameter $E(Y_1 - Y_0 | D=1)$ or its conditional version $E(Y_1 - Y_0 | D=1, X)$. This is the parameter of interest if the program is voluntary or if one wishes to evaluate the program effect only on participants;
3) The average treatment effect on untreated (*ATNT*). The parameter $E(Y_1 - Y_0 | D=0)$ or its conditional version $E(Y_1 - Y_0 | D=0, X)$. This is the parameter of interest if one wants to evaluate the impact of extending the program to a wider population.

As

$$E(Y_1 - Y_0 \mid D = 1) = E(Y_1 \mid D = 1) - E(Y_0 \mid D = 1) \qquad (2)$$

The first term on the right-hand side can be computed directly from the data, whereas the second term needs to be estimated under appropriate assumptions.

The "naïve" estimator: $E(Y_1 \mid D = 1) - E(Y_0 \mid D = 0)$, assumes $E(Y_0 \mid D = 1) = E(Y_0 \mid D = 0)$ which is reasonable only in particular circumstances, such as a random experiment where the two populations (treated and untreated) are really the same. Contrary the estimator is biased any time different populations with different distributions are compared.

$$\begin{aligned} E(Y_1 \mid D = 1) - E(Y_0 \mid D = 0) &= [E(Y_1 \mid D = 1) - E(Y_0 \mid D = 1)] + \\ &+ [E(Y_0 \mid D = 1) - E(Y_0 \mid D = 0)] \end{aligned} \qquad (3)$$

In other words, the naive estimator is made up by two components, respectively the true *ATT* component and the bias. As shown by Heckman et al. (1998) the bias can be split in three parts: the first component is due to the *non overlapping support* (the two population have completely different characteristics, *X*), the second to a different distribution of *X*, within the two populations, the third is due to differences in outcomes that remain even after controlling for the first two biases. The latter is the selection bias and is due to the selection of the unobservable. Heckman et al. (1997) have shown that if the data are administered by the same questionnaire, and participants and controls reside in the same local labour market, then the latter bias component is the least important, so that once the common support and the miss-weighting issues have been correctly addressed, most of the bias disappears.

[6] The parameters considered are three if we exclude local average treatment effect (LATE), see Imbens and Angrist (1994).

3.1 Differences in differences

Suppose a panel data made by two populations. The first population is made up by not treated. The second population is made by individuals who switch the treatment status over time, so that two outcomes can be observed for the same people, albeit not simultaneously but at different point in time.

Label period $t=t^2$ as the period when the reform takes place and $t=t^1$ as the earlier period. Every individual has $D_{it^1}=0$, while not treated are identified by $D_{it^2}=0$ and treated by $D_{it^2}=1$ (a change in status). Write

$$\begin{aligned}&E(Y_{1t^2}-Y_{0t^2}\mid D=1)=E(Y_{1t^2}-Y_{0t^1}\mid D_{t^2}=1)-E(Y_{0t^2}-Y_{0t^1}\mid D_{t^2}=1)=\\&=E(Y_{1t^2}-Y_{0t^1}\mid D_{t^2}=1)-E(Y_{0t^2}-Y_{0t^1}\mid D_{t^2}=0)+\\&-[E(Y_{0t^2}-Y_{0t^1}\mid D_{t^2}=1)-E(Y_{0t^2}-Y_{0t^1}\mid D_{t^2}=0)]\end{aligned} \tag{4}$$

The first two terms are observables and ATT can be identified only when the expression within square brackets is zero, hence the identifying assumption of the Differences in Differences (*DID*) estimator is

$$E(Y_{0t^2}-Y_{0t^1}\mid D_{t^2}=1)=E(Y_{0t^2}-Y_{0t^1}\mid D_{t^2}=0) \tag{5}$$

The meaning of expression (5) is that a difference between participants and not participants in the trend outcome is not permitted. In other words if treated and not treated are systematically different, this difference cannot change over time.

Two similar groups can possibly share the same trend, but when they have different observable characteristics changing over time, the DID required assumptions become rather implausible. The problem can be bypassed in two ways. First, running a regression to control for the observable characteristics; in such a way the identifying assumption (5) can be written on the errors terms (Bell et al., 1999). Second, matching treated and not treated observations which have the same pre-treatment observable characteristics, and running a *DID* estimator over the matched observations. In such a case people who have the same characteristics are very likely to have the same trend.

In the first case label C the comparison group and the target group T (treated group). Write the relationship between Y and the individual characteristics for the treated as[7]

$$Y_{it}=\gamma_{it}D_{it}+\beta' X_{it}+u_{it} \tag{6}$$

DID has the form:

$$\gamma^{DID}=(\widetilde{Y}^T_{t^2}-\widetilde{Y}^T_{t^1})-(\widetilde{Y}^C_{t^2}-\widetilde{Y}^C_{t^1}) \tag{7}$$

where γ is the ATT effect, $\widetilde{Y}^T_{t^j}=E\left[Y_{it}-\beta' X_{it}\mid t=t^j, g=T\right]$ and g denotes whether a firm is part of the T or C group. This would be a consistent estimator of γ if the non observables satisfy

$$E(u_{it}\mid t=s, i\in g, X_{is})=\varepsilon_g+km_s \quad \forall\, g,\, s \tag{8}$$

[7] Note we allow heterogeneous treatment effect.

The above condition allows for a macro trend effect, however the macro effect is assumed to be the same across the target and the comparison group, as pointed out in (5). In order to make this assumption more reliable the *DID* regression is run (after adjusting) within some cells described by X, in order to reduce the unobserved heterogeneity (Cochrane, 1968).

If each of the two groups is allowed to respond differently to the business cycle effect:

$$E(u_{it} \mid t = s, i \in g, X_{is}) = \varepsilon_g + k_g m_s \quad \forall\, g, s \tag{9}$$

where the K_g acknowledges the differential macro effect across the two groups. The *DID* estimators consistently estimates γ (Bell, et al., 1999):

$$p \lim \hat{\gamma}^{DID} = \gamma + (k_T - k_C)(m_{t^2} - m_{t^1}) \tag{10}$$

The *true* effect of the programme is identified when $k_T = k_C$.

Now let us consider a different time interval, t^3 and t^4, over which occurred a macro trend that matches the term $(k_T - k_C)(m_{t^2} - m_{t^1})$. Then running "*DID* of *DID*":

$$\gamma^{DID-DID} = \left[(\tilde{Y}^T_{t^2} - \tilde{Y}^T_{t^1}) - (\tilde{Y}^C_{t^2} - \tilde{Y}^C_{t^1})\right] - \left[(\tilde{Y}^T_{t^4} - \tilde{Y}^T_{t^3}) - (\tilde{Y}^C_{t^4} - \tilde{Y}^C_{t^3})\right] \tag{11}$$

one gets consistent estimates for γ.

3.2 Regression adjustment

The assumption is that only observable variables are able to affect the outcome; in other words there is not a selection on the non observables. Label u_{Di} as non observed heterogeneity. Selection on the observables can be written as

$$E(u_{0i} \mid D_i, X_i) = E(u_{0i} \mid X_i) = 0 \tag{12}$$

The meaning of (12) is that, once X_i has been taken into account, the error term (on average) is independent of the treatment status, i.e. u_i is a random sample. Hence the naïve estimator on the residuals can be used:

$$Y_i = \gamma D_i + \beta' X_i + u_i \tag{13}$$

Under the abovementioned hypothesis and under the classical OLS assumption, $\gamma = ATT$. Indeed (12) can be written as

$$Y_{0i} = \beta' X_{0i} + u_{0i} \tag{14}$$

$$Y_{1i} = \gamma + \beta' X_{1i} + u_{1i} \tag{15}$$

If (12) holds:

$$E(Y_{0i} \mid D_i = 1, X_i) = E(Y_{0i} \mid D_i = 0, X_i) = E(Y_{0i} \mid X_i) = \beta' X \tag{16}$$

Now $ATT(X) = E(Y_1 - Y_0 \mid D=1, X)$ and substituting the above equation (16), $ATT(X) = \gamma$.

This method can be improved in order to allow for the heterogeneous impact and to account for the misspecification issues. The method is straightforward although it relies on very strong assumptions. Indeed using a linear forecast, the regression adjustment overcomes the common support problem, and taking into account the *X's* , is free from misweighting.

3.3 Matching

Matching selects the untreated observations which are as close as possible to the treated ones in terms of observable characteristics. In other words matching "mimics" experimental data using observational data.

Matching is based on the unconditional independence assumption[8] (*CIA*):

$$Y_0 \perp\!\!\!\perp D | X \qquad (17)$$

Conditioning on X, Y and D are independent. A weaker version of *CIA* is required, $E(Y_0 \mid D = 1, X) = E(Y_0 \mid D = 0, X)$. Given *X*, on average, the non-treated outcomes are what the treated outcomes would have been if they had not been treated.

Moreover matching assumes

$$0 < \Pr(D = 1 \mid X) < 1 \qquad (18)$$

for all *X*. Then *ATT* can be defined for all values of *X*. The meaning of (18) is that for each *X* both treated and untreated observations are available. This hypothesis is quite stringent in relation to the available observations, and is usually replaced by the common support assumption, $\overline{X} = X_1 \cap X_0$. Under this assumption, ATT can be identified as:

$$\begin{aligned} &E(Y_0 \mid D = 1, X \in \overline{X}) = E[E(Y_0 \mid D = 1, X \in \overline{X}) \mid D = 1] = \\ &E[E(Y_0 \mid D = 0, X \in \overline{X}) \mid D = 1] = \int_{\overline{X}} E(Y_0 \mid D = 0, X) f(X \mid D = 1) dX \end{aligned} \qquad (19)$$

Matching requires for each treated observation, a not treated observation sharing the same individual characteristics.

Matching (19) requires to match on a high dimensional *X*, i.e. high dimension density. Rosenbaum and Rubin (1983) have shown that if *CIA* holds, then it is valid for any balancing score[9], *b(X)*. Defining the propensity score, *p(X)*, as the conditional probability of assignment to treatment

$$p(X) = (D = 1 \mid X) \qquad (20)$$

The propensity score is the coarsest balancing score and *X* is the finest. Putting altogether we can write

$$E(Y_0 \mid D = 1, p(X)) = E(Y_0 \mid D = 0, p(X)) \qquad (21)$$

Hence matching on propensity score bears a *ATT* unbiased estimate. A powerful result because only one dimensional density is needed in order to get the appropriate estimation.

[8] This assumption is called by Rosenbaum and Rubin: strongly ignorable treatment assignment.
[9] A balancing score, *b(X)*, is a function of the observed covariates *X* such that the conditional distribution of *X* given *b(X)* is the same for treated and control units: $X \perp\!\!\!\perp D | b(X)$. The most trivial balancing score is *X*.

In general after having run logit, probit or semiparametric estimation on pre-treatment variables *X,* the fitted values, *p(X)*, are used in order to match treated and control units. Following Heckman et al. (1997) the form of the matching estimator can be cast in the following framework

$$A\hat{T}T = \sum_{i\in T} \omega(i)[Q_{1i} - \sum_{j\in C} W(i,j)Q_{0j}] \quad \text{for } X \in \overline{X} \tag{22}$$

where Q_{1i} is function of the treatment outcome, and Q_{0i} is function of the comparison group outcome, *W(i,j)* is a weight with $\sum_{j\in C} W(i,j) = 1$ and *W(i)* is a weight that accounts for heteroscedasticity and scale.

The notation is general so to allow regression adjustments of *Y*, or *DID* estimates. Matches for each participant are constructed by taking weighted averages over the comparison group members. Matching can be performed within various strata to recover estimates relative to different subset of the population of interest (Heckman et al., 1997, Rosenbaum and Rubin, 1984). Matching estimators differ in the weight attached to the members of the comparison group (for a survey see Heckman et al.,1999).

Label *C(pi)* as the set of untreated neighbours of treated *i,* which has a propensity score estimated value of p_i. The nearest-neighbour matching estimator sets $Q_{1i}= Y_{1i}$, $Q_{0i}= Y_{0i}$, $\omega(i) = 1/N_T$, and $C(i) = \min_j \|p_i - p_j\|$, where $\|...\|$ is a norm.

C(pi) is a singleton set, except for ties that are broken by a random draw, hence $W(i,j) = 1$ if $j \in C(p_i)$ and zero otherwise. [10] Two version of this method are available: with replacement or without replacement (reusing or not reusing X_j for other matches).

Nearest-neighbour matching use all the treated group, without caring about the effective distance between individuals *i* and *j,* but distance has to be set to the minimum. In order to avoid matching too different individuals, a pre-specified tolerance (calliper) can be imposed (Cochrane and Rubin, 1993). Calliper matching sets requires $C(i) = \min_j \|p_i - p_j\| < \tau$ where τ is the tolerance level.

Matching can be easily extended to the 'smoothed matching'. For example, the n-nearest neighbour matching considers as counterfactuals the *n* closest control observations for each treated unit, instead the radius matching considers all the untreated observations with have a propensity score falling within a radius τ from p_i: $C(i) = \|p_i - p_j\| < \tau$.

In the last two methods *W(i,j)*=$1/N_i^C$ if $j \in C(p_i)$ and zero otherwise (N_i^C is the number of control unit matched with *i*-th treated observation).

Through the process of choosing and re-weighting observations, matching corrects for the first two sources of bias, and selection on not observables is assumed to be zero. Compared regression matching is non parametric and therefore avoid inconsistency due to misspecification; estimating over the common support only comparables individuals considered.

Heckman, Ichimura and Todd (1998) combine matching method and regression adjustment on *X*. Their method extends the classical matching by utilizing information on the functional form of the outcome equations. Their estimator can be seen as a compromise between the fully nonparametric approach and a parametric model.

Regression-adjusted matching is performed by the following procedure. Assume a conventional econometric model for outcomes in the non-treated state that is additively separable in observables and not observables (14); estimate the component of $E(Y_0 \mid X, D = 0) = X\beta_0 + E(U_0 \mid X, D = 0)$ using partially linear regression methods. Before

[10] In this case matching corresponds to weighted least square.

estimating ATT by matching methods (22), $X\hat{\beta}_0$ are removed from Y_0 and Y_1 by setting $Q_{1j} = (Y_{1i} - X_i\hat{\beta}_0)$ and $Q_{0j} = (Y_{0j} - X_j\hat{\beta}_0)$.

ATT is obtained by substituting $Q_{1i} = (Y_{1it^2} - Y_{0it^1})$ and $Q_{0j} = (Y_{0jt^2} - Y_{0jt^1})$ in (22), or alternatively making the difference between the before and the after treatment matching estimates. Obviously the *CIA* assumption needs to hold not on levels but on differences:

$$E(Y_{0t^2} - Y_{0t^1} \mid X, D = 1) = E(Y_{0t^2} - Y_{0t^1} \mid X, D = 0) \tag{23}$$

and under additive separability of the errors and index sufficiency the condition becomes

$$E(u_{0t^2} - u_{0t^1} \mid p(X), D = 1) = E(u_{0t^2} - u_{0t^1} \mid p(X), D = 0) \tag{24}$$

From the econometric point of view matching "is an attractive estimator because it permits selection to be based on potential programme outcomes and allows for selection on unobservables" (Heckman et al., 1997).

4. Main results

4.1 Data

The longitudinal panel used in this research is constructed from the administrative records of the Italian Social Security System (Inps). It refers to the entire population of employee and workers in two provinces, Treviso and Vicenza, of an Italian region, Veneto. The database covers each single plant and each single individual employed in the private sector (no state and local government, with few exceptions) except for farm workers and people receiving no salary.

Inps data include register-based information on all establishments and employees that have been hired by those establishments for at least one day during the period of observation, independent of the workers place of residence.[11] The unit of observation is the employer-day; such information is used to build a monthly history of the working life of each employee. Employers are identified by their identification number, which changes if ownership, in a strict sense, changes. This has been amended and any time more than 50% of all employees are taken over by the new legal employer, the employment spell is said to be continuing. Similarly, if there are short breaks in the employment spell, as long as the worker continues at the old employer, his spell is considered uninterrupted.[12]

Data include all individual employment spells with an employer, of whatever duration, and this probably results in a lot of very short spells. Although short spells characterize the average job, they are concentrated at workers young age, while long spells characterize the mature worker current experience. We keep all employment size in our data set, because our territory is

[11] The entire working life for all employee that have worked at least one day in Treviso and Vicenza, has been reconstructed, considering the occupational spell out of Treviso and Vicenza as well.

[12] A 'cleaned' social security archive has been used. The engagements/separations and the creations/destructions that are due to a change in the unit that pays the social security contribution not matched by a corresponding change of the working population assessed at the establishment level are defined as 'spurious' and have been deleted. This has lead to a reduction of 9% of total engagements and separations in manufacturing. This procedure is common practice among people working with social security data.

characterized by a multitude of very small units (establishments with ≤ 5 employee account for almost 12% of the total manufacturing employment).[13]

Veneto labour market has been characterized since almost a decade by almost full employment and by a positive rate of job creation in manufacturing, before a negative national rate. It is a dynamic territory based on manufacturing, with a large population of small firms; the average establishment size is 14 employees. The stock of manufacturing workers in the two Veneto provinces of Treviso and Vicenza has varied between 194.000 employees at the early eighties and 233.000 employees in 1996, with a yearly positive average rate of variation of 1.4%. The average rate of growth in employment is the result of a marked increase of white collars and women.

Our database has records on establishment and worker from 1975 to 1997. Employers are classified in the three-digit ATECO 1981 standard classification. The period of time covered by the database allows us to discuss the role of the CFL, which is clearly identified in the social security archives, from the very beginning.

4.2 Descriptive analysis

One of the CFL goals was to reduce the unemployment rate among young workers. In order to assess the impact of CFL on unemployment over time the employment rate by age cohort is analyzed.[14] The study is restricted to the employment in private sector (without farm and public sector employees), because of our dataset limitations and because CFL was a program typically addressed to private employees.[15]

Our basic dependent variable is the employment rate and not the absolute number of employees, to get rid of the main demographic changes that have affected the Veneto labour market through time. The baby boom in the early sixties and a subsequent dramatic decline in birth rate imply complex effects on labour supply. The growth in average life expectancy and the rapid increase in school attendance determined a delay in the entrance in the labour market – and this might influence the composition of the employee stock according to the age (Canu, Tattara, 2004).

[13] The absolute importance of small establishments makes the comparison with other countries doubtful; for example in our territory the percentage of employment in establishments with ≥ 100 employee is 27% while in Denmark is more than 40% and is still larger in the United States.

[14] We know that two concepts are quite different and not mathematically related, but we are convinced that in any case if CFL rose employment rate it would be a good result, all things being equal.

[15] We know on average self-employed, farmer and public sector workers cover more than one third of all employed, of course, this rate increases over ages. Furthermore we are conscious that studying private employment instead of "total", we take the risk to consider an increase in the employees due to drops in school attendance or self-employment.
We believe this kind of changes are really rare.

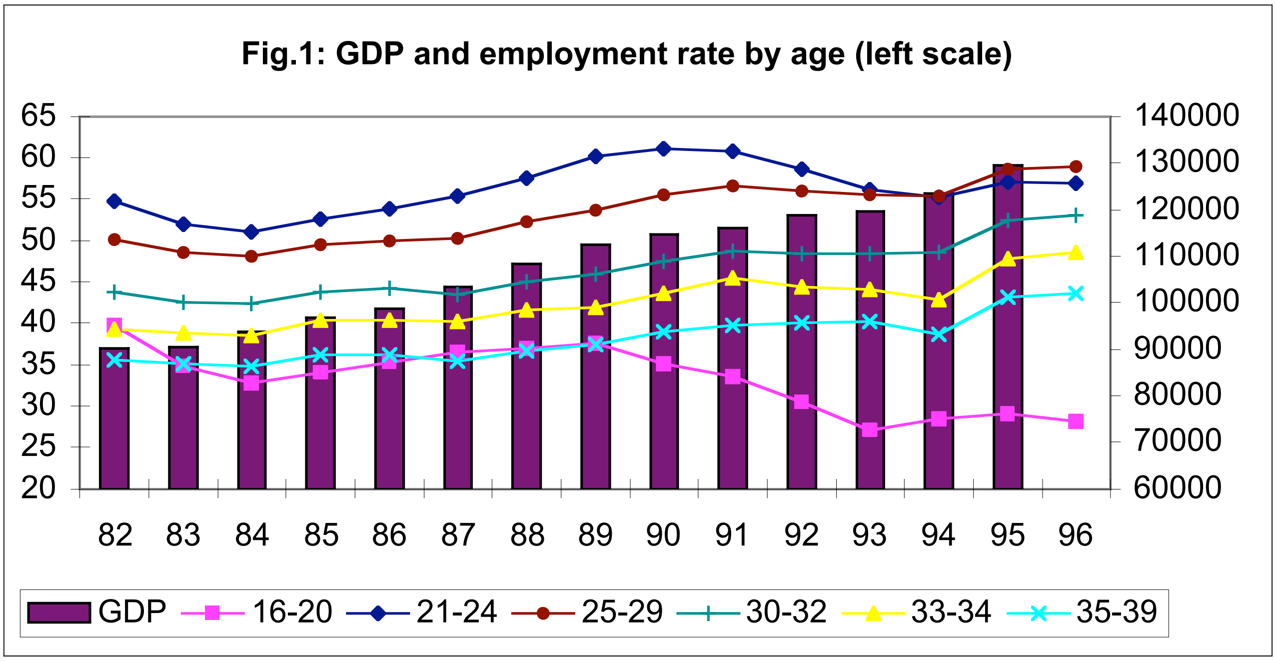
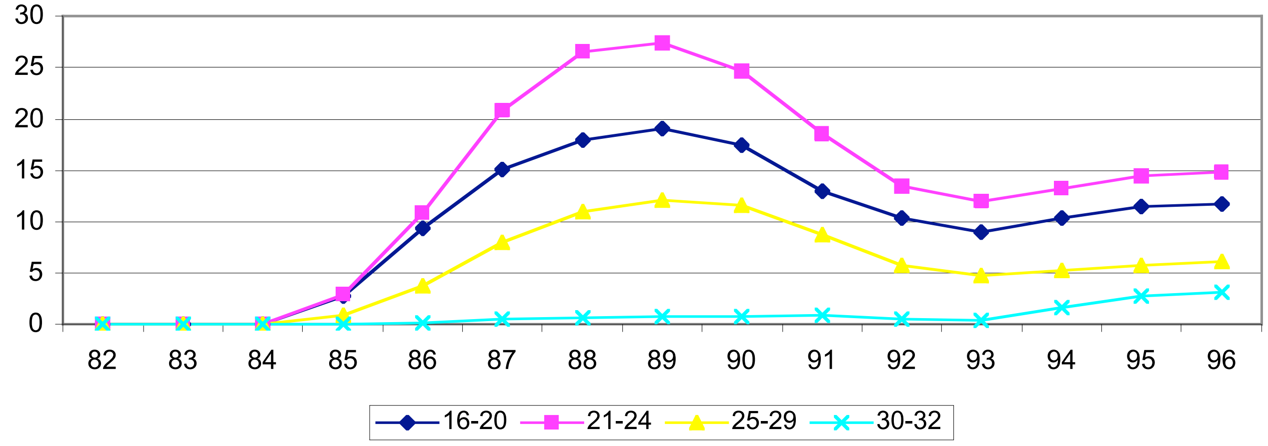


Figure 1 depicts the employment rate of workers aged between 21 and 29 years (the CFL eligible workers), showing a positive increase of almost the employment rate by age cohorts due to the CFL introduction and its subsequent diffusion till the early nineties, when much of the CFL effectiveness was reduced by the severe benefit cuts. The rate of employment for CFL eligible grows faster in respect to other workers. Moreover the same rate for workers younger than 30 years of age slows down around 1989, just when CFL setting changed, and decreases in the first nineties when business cycle declined. Eventually in Treviso e Vicenza, CFL seems to have had an impact, but its effect looks quite variable according to the different age groups.[16] Indeed, the group 21-24 seems more sensible to changes in the program, even if this cohort was eligible, after 1988, to the apprenticeship contract and this might have raised the employment rate as well. On the contrary the 25-29 age group seems to follow more closely the business cycle than the institutional reforms. Evidence is mixed and a lot of heterogeneity enters these data. For example the increase in the younger worker stock might depend mainly by firms that did not use the CFL program, by the diffusion of apprenticeship and so on. Additionally figure 1 implicitly assumes that workers with the same age but in different points in time have the same "economic behaviour" while an attentive birth cohort analysis is really required.

Figure 2 depicts the number of CFL over time. Indeed when the cycle is high, 1989-1991, CFL increases more slowly or diminishes due to the contract reform; the decrease in 1992-1993 essentially depends on the international economic slowdown (we remind that in 1992 Italy left the SME); during 1994-1996 CFL grows again because of the recovery of the economy. Again only the joint effect of the CFL reform and of the growth of the economy is apparent. Indeed looking at the not eligible workers of the age cohort immediately subsequent to eligibility (34-35 years old), the employment rate seems to have grown, more or less, at the same rate of the growth rate of the eligible workers (30-32 year old). Clearly the CFL has been used as a buffer (Pacelli, 2002), and it is not positively clear if CFL have increased the Italian employment rate, or it CFL is just "another name" for people who, in any case, would have been hired or who were already employed.

[16] We set workers in those groups in order to avoid interferences with apprenticeship contract, respect different Italian school attendance (around 20 high school terminates, and around 25 people graduate), distinguish possible different skills and work experiences and, finally, to account for extension of CFL to older workers.

Fig. 2: Share of CFL on the employment stock by age



The basic question is if there was a substitution effect between eligible and not eligible workers. Looking at the above picture it seems likely that a slight substitution effect has taken place, but the problem is difficult to assess. Indeed during the end of eighties, when the economic cycle was high, the employment stock grew, but the eligible increase has been stronger. Was there a substitution effect? According to the neo-classical theory, fiscal rebate on young workers could induce firms to replace capital with labour, and particularly, but not only,[17] with young workers. Carrying on the same idea an opposite effect is expected when the CFL rebate diminished, but again from the chart 2 there is not a clear replacement effect, because the different pattern are due to the joint effect, the CFL's reform and the economic cycle (there was a slump in 1992 and after)[18] Indeed the decrease for the 21-24 cohort may depend by the fact that more young people are achieving their degree in most recent years, deferring the entrance in the labour market.

In 1994 the employment rate diminished for all cohorts, while the CFL quota increased and this can support the thesis that the CFL was used as a buffer, meaning that without CFL the employment rate would have diminished even more; in 1994 the CFL eligibility extended to other categories of firms, the age limit shifted from 29 to 32 years of age. Looking at the 33-35 age cohort, the employment rate diminished by 1%, while the same rate for the 30-32 cohort stayed constant and the CFL share for the same group grew at about 1%. But, of course, this is not a clear sign of the CFL effectiveness, because data include many heterogeneous elements, as already noticed.

Let us now examine the different behaviour for artisans and non-artisan firms (Figure 3-6).

Both type of firms, until 1989, used CFL. The CFL share of the employment stock was around 30%, but after 1989 the number of CFL diminished until 1993, bringing the share around 10-15%; since 1994 artisan firms started again to use CFL, while other firms were much less interested in the contract, and indeed the CFL quota on the worker stock was on average less than 10%.

[17] Because sometimes older worker are irreplaceable or they have to work together with younger worker, such as skill worker and apprentice.

[18] We remind CFL is fixed term contract.

Fig. 3: Employment rate by age by non artisan firms

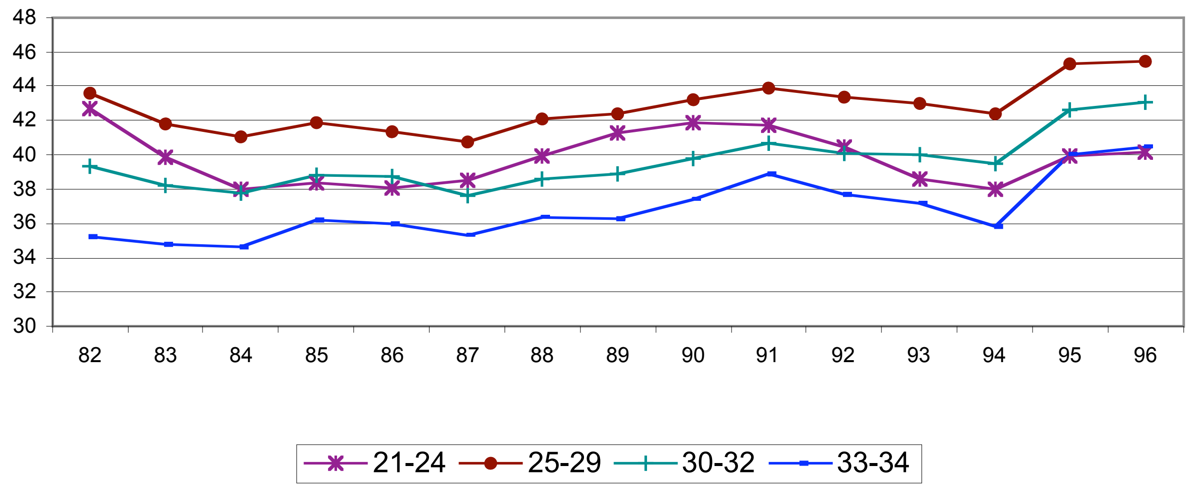
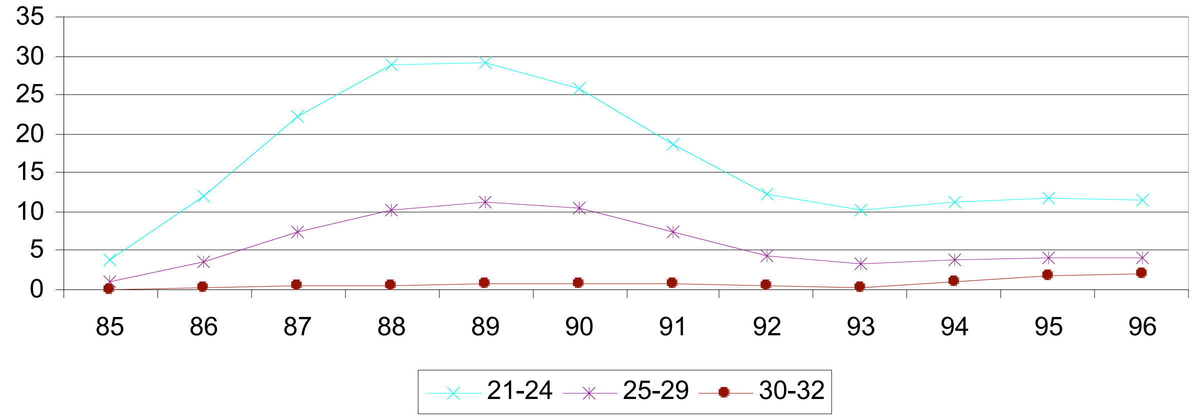
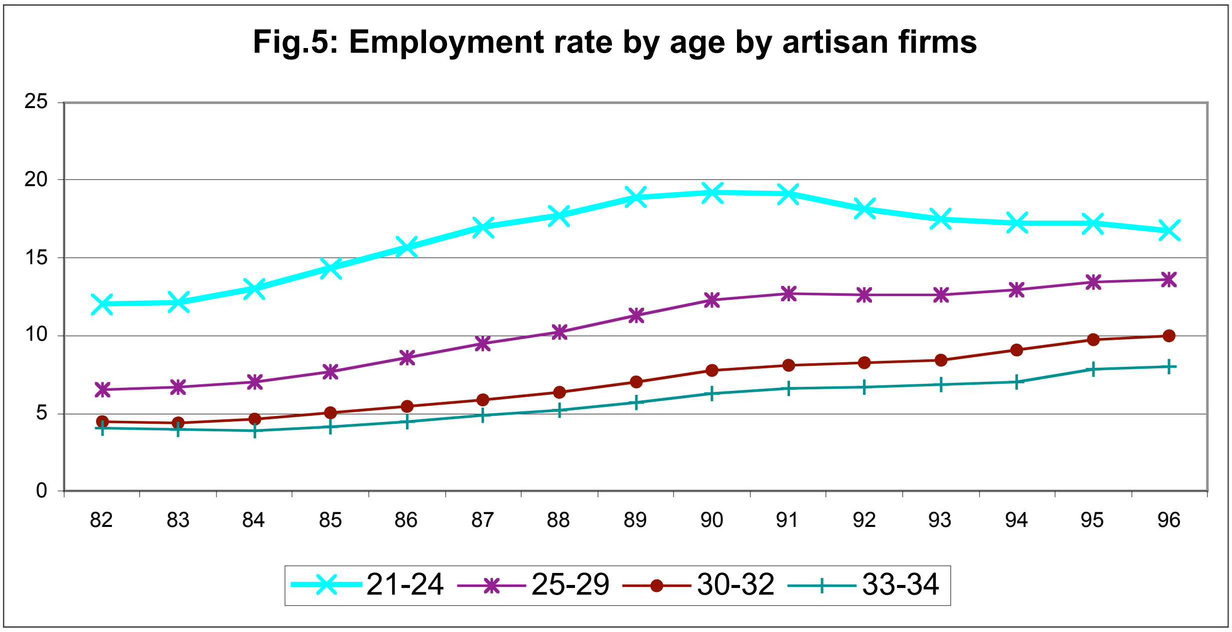


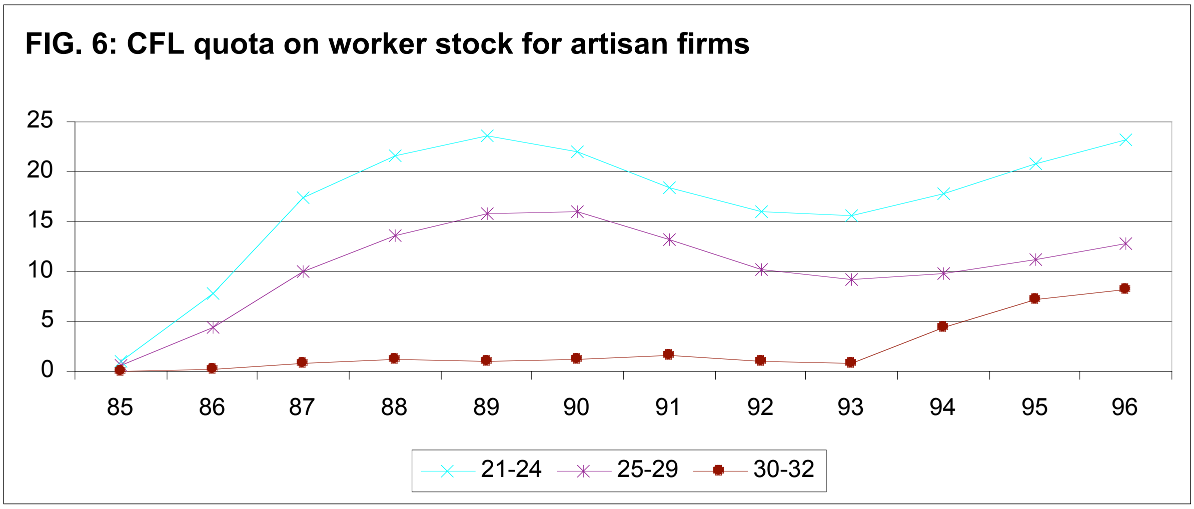
FIG. 4: CFL quota on worker stock for non artisan firms



CFL had few consequences on non artisan firms: a substitution between contracts can be depicted. In fact, looking at the years 1988-91, the worker stock grows, whereas the CFL stock diminishes due to the CFL rebate reform; indeed the stock level is sensible to the cycle, while the CFL level is sensible to both the cycle and the change of the law. After 1993, with the recovery, the employee stock has grown, while the CFL quota stays constant, as firms have now a little benefit from this contract. On the contrary for artisan firms until 1990, the CFL quota and the employee stock go hand in hand. During the period 1990-93 the employee level is almost constant (decreasing for the 21-24 age group), while the CFL quota diminishes. A possible reflection of the 1991 CFL program,[19] but in 1990, especially for the 21-24 age cohort, this result embeds the increased rate of university attendance. Since 1994 the CFL quota starts to grow again accompanied by the growth of the employee stock, at least for workers older than 24 years.

[19] We remind that for craftsmen the fiscal rebate was stationary, but in order to keep on hiring through CFL firms had to hire at least 50% of CFL. This amounts to a higher firing cost.





Taking the 30-32 cohort, indeed both workers employed in artisan and non artisan firms increase, while the CFL share grows only for artisans. A possible sign of the different behaviour of firms enjoying different benefits.

Summarizing the descriptive evidence, we speculate that the CFL program is likely to have raised employment at its very beginning. Evidence related to the CFL reform is rather mixed.This provisional conclusion needs a proper test.

4.3 Econometric analysis

A common assumption underlying the estimation strategy is the general equilibrium hypothesis, i.e., the program must not affect the control group. In actuality this assumption looks reasonable, although CFL was a widely diffused practice, a large amount of young workers was available for hiring by the non treated. Moreover we assume that CFL did not affect the standard economic behaviour of not treated firms.

The panel dimension allowed by our data is exploited, because panel data provide more credible outcomes compared to repeated cross sectional data. The panel excludes firms that are born or have died in the years under consideration so the possible stock variation due to the to natural process of firm's birth and death is ruled out. Were it included, the CFL impact might embed the firm's birth/death effect.

Furthermore we used panel dimension in another sense. The eligible cohorts with is made by employees born between 1962 and 1965 and we split the CFL exposed age group (15-29) in two cohorts: the 21-24 and the 25-29 cohort, in order to distinguish younger workers from more skilled workers. Our endogenous variable is the number of employee exposed to the CFL program and not the number of CFL workers: the latter implies a lot of variation in the dependent variables. Finally we use monthly level stocks and not yearly stocks, because seasonality for young workers is quite strong, and since it has a cyclical pattern, yearly data could mix the CFL effect with the seasonality effect.

The cell is based on size and sector. The first cell clusters firms under 50 employees belonging to metal manufacturing (Ateco 81 3), the second cell firms under 50 employees belonging to other manufacturing (Ateco 81 4). In order to compare firms with the same characteristics we use the p-score matching method. We run a logit estimation on the following "pre-treatment" explanatory variables, assumed exogenous: size, sector, industrial area dummies, firm age, men quota, blue-collar quota (as proxy for more capital or labour intensive firms), apprenticeship quota (as proxy for firm inclination to use fixed term contract and training contracts). The specification that maximizes the overall classification rate is chosen.[20]

On average the most powerful predictors of the treatment effect are the dimension and the sector. Calliper nearest-neighbour matching with a tolerance under 5% imposes a common support and at the same time excludes less than the 5% of the treated population.[21] Seasonality is dealt by matching firms with the p-score as close as possible, so to get comparable firms, and taking the differences between observations which belong to the same firm. Finally, of course, we took year averages of monthly differences and we got the *ATT* comparing the two averages (tables 2-4).

To evaluate the initial impact,[22] treated firms are compared with not treated firms within the same cell in 1986 and 1985. Results are reported in table 2 and 3. Cell 1 γ coefficient (artisan firms, belonging ateco3= metal manufacture, size <=30) means that the employee stock for treated firms, for the age cohort 21-24, increases about 0.6 in comparison with not treated firms, i.e. every 10 treated, firms hired about 6 young workers in addition to those hired by comparable non treated firms.[23] Cell 2 (Ateco 4= other manufacture) results are very similar, with a $\gamma = 0.5$.

The result is stronger considering the non artisan firms (table 2). In such a case the cell has been extended by increasing the firm size (to account for the different dimension of the average non artisan firm) and this makes the two coefficients not immediately comparable (the coefficient needs rescaling according to the average size) but non artisan firms had in fact an additional benefit in respect to artisan firms, and this makes the larger impact a reasonable outcome. Non artisan firms, before 1990, through CFL were able to hire their workers directly, without the intermediation of Ufficio di Collocamento, and this facility provided a sensible additional benefit.

[20] Estimate results are available on request. The balancing property of propensity score has been tested using the procedure described in Becker and Ichino (2002). In the future also we would like to account for earning differences among firms, and dimensional and compositional firm variation which affect both outcome and treatment status.

[21] In some specification when there (were a lot of ties) we imposed to use not only the nearest neighbour but the 5 nearest. In this way we improved precision of estimates but variance increased. The tolerance and the number of comparison were chosen in order to minimise the difference in terms of *Y* between the two groups during pre-treatment period, when reforms did not take place yet.

[22] We remind CFL started in 1984, but because of trivial number of firms which used this program we are able to evaluate it since 1985.

[23] Standard errors are calculated through one hundred bootstrap replication. The asymptotic distribution of the matching estimator is unknown, but sample variance has been calculated, see for example Becker and Ichino (2002).

Tab. 2: Evaluation of CFL for 21-24 age group in 1986

	Artisan		Non artisan	
	Ateco 3	Ateco 4	Ateco 3	Ateco 4
γ^{DID} (std. err.)	0.5906373 (0.0321291)	0.5129629 (0.0373477)	0.7572222 (0.0470121)	0.9467909 (0.0465071)

Table 3 shows that for the 25-29 cohort, the CFL impact is somewhat weaker but significant. For the artisan cell within Ateco 3 we are not able to identify the impact, due to the scarcity of treated firms[24] (31 firms, less than 3% of firms within the cell). On the contrary for Ateco 4 the treated number is a bit higher (almost 5%), and the γ coefficient estimate is 0.3, which means on average each 10 treated firms hired 3 workers more than the untreated. For non artisan firms we get stronger outcomes, around 0.7 both for Ateco 3 and 4. The results are quite similar to those presented in table 2, although of reduced in magnitude. We interpret these results as the fact that the program was effective for the two cohorts, i.e. for all the relevant ages affected by the program.

Tab. 3: Evaluation of CFL for 25-29 age group in 1986

	Artisan		Non artisan	
	Ateco 3	Ateco 4	Ateco 3	Ateco 4
γ^{DID} (std. err.)	-	0.308641975 (0.0383495)	0.7021605 (0.0741631)	0.748611111 (0.1813863)

Coming to the second CFL reform, which took place in January 1st, 1991, the same technique is applied, although treated and not treated firms (non artisan and artisan) are 'exogenously' selected. The estimation strategy is based on the following argument: in 1991 the fiscal rebate diminished by 50% only for non artisans, hence *DID* is used, utilizing artisans as the control group, in order to account for the observable heterogeneity, and the matching procedure is applied as well. Now firm size is below 20 employees (the average boundary for artisan firms) and the sector must belong to Ateco 3 or 4. Table 4 reports the results of equation (22), based on the usual explanatory variables. In all cases but one the coefficient not significant; only for the cohort 21-24 within Ateco 4 the impact is positive but limited in magnitude.

Lack of impact means that the program did not work, due to several causes. The lack of a positive impact might be the result of a labour demand positively depending on fiscal benefits, but discontinuously or/and lagged: until the firms do not jump the treshold they have a stable labour demand (and the reduced benefit is too small to make the jump), or more simply it needs some time so that firms to change their stock. Of course these hypothesis need to be tested. The anomalous result for cohort 21-24, Ateco 4 can be due to the intertemporal heterogeneity and needs a different analysis[25].

Tab. 4: Evaluation of CFL for in 1991

---

[24] The main problem concern the logit estimation, since with so few treated firms we were not able to get reliable estimates, at least in terms of standard deviation.

[25] Heckman et al. (1997) suggest to apply the matching method together with regression adjustment. More precisely one has to run regression on control group in the second year, and remove the fitted effects from both groups ($Y_{it^2} - X_{it^2}\beta_{0t^2}$). Blundell et al. (2003) observe that there are two assignments that are non-random: the first is to eligible population and the other assignment is to the relevant time period (before and after reform); "for the evaluation to make sense with heterogeneous effects, we must guarantee that the distribution of the relevant observables characteristics is the same in the four cells defined by eligibility and time", they suggest to use two propensity score: one for eligibility and one for time period.

	Ateco 3		Ateco 4	
	21-24 years old	25-29 years old	21-24 years old	25-29 years old
$\gamma^{DID\text{-}DID}$ (std. err.)	0. 004357298** (0. 0370467)	0.004493464** (0.0450935)	0.188461538 (0. 0324727)	0. 023450586** (0.0371799)

Note: ** not significant at 10%.

Conclusions

Matching-*DID* estimate results suggest that the CFL program had a positive impact on the young people employment rate. At the introduction, the CFL impact was positive, during the mid eighties, and was quite homogenous between the different age groups, but heterogeneous among the different cells: artisans and non artisans had a different impact factor, larger for non artisans that got the bigger benefits (the fiscal benefit and the benefit due to the Collocamento avoidance). In such a way our study differs and supplements previous works devised to assess the CFL impact and bearing no evidence of a significant policy impact (Contini et al. 2002). Moreover during the first nineties the CFL reform is characterised by a reduced benefit in respect to the initial benefit for non artisan firms. Our estimate exploits the difference between artisan and non artisan firms in order to asses the impact of the reform, and does not provide a significant and reliable change in the treated set in relation to the not treated. These results are provisional.

References

Becker, S. O., Ichino, A. (2002). *Estimation of averages treatment effects based on propensity scores*. Stata Journal.

Bell, B., Blundell, R., Van Reenen, J. (1999). *Getting the unemployment back to work: the role of targeted wage subsidies*. International Tax and public Finance, 6: 339-360.

Blundell, R., Costas Dias, M. (2002). *Alternative Approaches to Evaluation in Empirical Microeconomics.* Cemmap W.P. 10/02.

Blundell, R., Costas Dias, M., Meghir, C., Van Reenen, J. (2003). *Evaluating the employment impact of a mandatory job search program*. Cemmap W.P.

Canu, R, Tattara, G. (2004). *Quando le farfalle mettono le ali. Riflessioni sull'ingresso delle donne nel lavoro dipendente*. WP.n.53 9/2003. Gruppo MIUR Dinamiche e persistenze nel mercato del lavoro italiano ed effetti di politiche.

Cochrane, W. (1968). *The effectiveness of adjustment by subclassification in removing bias in observational studies*. Biometrics: 295-313.

Cochrane,W., Rubin, D. (1973). *Controlling bias in observational studies*. Sankyha, Vol. 35: 417-446.

Contini, B., Cornaglia, F., Malpede, C., Rettore, E. (2002). *Measuring the impact of the Italian CFL programme on the job opportunities for the youth*. WP.n.43 9/2002. Gruppo MIUR Dinamiche e persistenze nel mercato del lavoro italiano ed effetti di politiche.

Heckman, J., Ichimura, H., Smith, J., Todd, P. (1998). *Characterizing selection bias using experimental data.* Econometrica, Vol. 66, n.5: 1017-1098.

Heckman, J., Ichimura, H., Todd, P. (1997). *Matching as an econometric evaluation estimator: evidence from evaluating a job training programme*. Review of Economic Studies, Vol. 64, n.: 605-654.

Heckman, J., Ichimura, H., Todd, P. (1998). *Matching as an econometric evaluation estimator*. Review of Economic Studies, Vol. 64, n.: 605-654.

Heckman, J., LaLonde, R., Smith, J. (1999). *The Economics and Econometrics of Active labour Market Programs*. In A. Ashenfelter and D. Card, eds., Handbook of labour Economics, Vol. 3 Amsterdam: Elsevier Science.

Imbens, G., Angrist, J. (1994). *Identification and Estimation of Local Average Treatment Effects. Econometrica*, Vol. 62, n. 2: 467-75.

Pacelli, L. (2002). *Fixed term contract, social security rebates and labour demand in Italy*. W.P. LABORatorio Riccardo Revelli.

Rosenbaum, P.R., Rubin D.B. (1983). *The central role of the propensity score in observational studies for causal effects*. Biometrika, Vol. 70, n.1: 41-55.

Rosenbaum, P.R., Rubin D.B. (1984). *Reducing bias in observational studies using subclassification on the propensity score*. Journal of American Statistical Association, Vol. 79, n.387: 516-524.