Getting out of the starting gate on the right foot: employment effects of investment in human capital*

Agata Maida[a] and Daniela Sonedda[b,c†]

[a] *University of Milan.*
[b] *University of Piemonte Orientale.*
[c] *Centre For North South Economic Research (CRENoS), Cagliari, Italy.*

February 2019

Abstract

The technological progress and the globalisation process reshape the nature of jobs inducing a substantial drop in the incidence of permanent employment occupations. This paper estimates whether employers could be less reluctant to hire workers on a permanent basis in presence of a human capital investment which they partly finance. We find that the permanent employment rate of cohorts affected by law no. 92/2012 at the age threshold of 30 years increased by about 1% when compared to the permanent employment rate of similar untreated cohorts. This difference in discontinuity impact can be generated by the vocational apprenticeship labour contract only. After 36 months from the baseline, this positive effect persists and increases to about 5%. We interpret our results as evidence that a labour contract that invests in human capital serves as a stepping stone into permanent employment.

Keywords: Human Capital, Apprenticeship, Permanent employment

JEL Codes: J24, J41, J21

*We are indebted to Andrea Albanese, Lorenzo Cappellari, Roberta Gatti, Andrea Ichino, Deni Mazrekaj and Michele Pellizzari for helpful comments and discussions. We thank audiences at XXXIII AIEL Conference (Ancona) and 1st Bank of Italy - World Bank International Research Workshop "Building Human Capital for 21st Century Jobs" for useful comments. Daniela Sonedda thanks Giorgio Pedrazzi and the SuperComputing Applications and Innovation Department, CINECA for providing access to the supercomputer Marconi, ISCRA Class C Project n.EWEiHC and Cristiano Padrin for his excellent technical support. We are fully responsible for any errors.

†Corresponding author: Via Perrone 18, 28100 Novara, Italy; email: daniela.sonedda@uniupo.it

1 Introduction

A growing debate in labour economics has pointed to the relevance of creating not only jobs but good jobs in response to the surge of product market competition, related to the globalisation process, and in response to the development of advanced technological procedures that make traditional jobs obsolete. One expects that higher human capital accumulation helps in pursuing this goal. Is this the case? An important question is therefore which kind of labour contract is suited to create high quality jobs. Measuring and defining such jobs is a hard task. However, it is likely that employers are less reluctant to hire on a permanent basis if they perceive that the job created is potentially of good quality. One also expects that high quality jobs are associated to higher employee's productivity (and therefore higher wage rates and possibly higher firm's productivity) and higher permanent employment probability.

This paper contributes to the literature by analysing whether a labour contract that is committed to the provision of a costly human capital investment serves as a stepping stone into permanent employment. If this is the case, this contract is, possibly, creating not only jobs but also jobs of good quality. To the best of our knowledge, this is the first paper that tackles this empirical issue. A huge body of the literature focuses on the wage (earnings) impact of publicly funded training (see for instance Abadie, Angrist and Imbens 2002, Heckman, LaLonde and Smith 1999). The impact on the employment rate of publicly funded training programs is less documented, although it is crucial to bound the treatment effects on wages (Lee 2009).

Improving the employment outcomes of those facing significant employment barriers is surely an intended effect of government-supplied training programs. How individuals' employment prospects are affected by an initial human capital investment, partly financed by the firm, is less clear. In his seminal paper on investment in human capital Becker (1962) argues that employers do not provide general training because of their inability to capture any of the future returns of this investment. The developments of the literature have modelled and emphasised the role of the informational asymmetries for converting general into specific training. In fact, the current employer is better informed on its employees' abilities relative to other firms. This informational advantage creates ex-post monopsony power, and paves the way to firm sponsored training, even if the skills provided are general (Acemoglu and Pischke 1998, Acemoglu and Pischke 1999). On the top of asymmetric information, the commitment to the provision of training is a key mechanism that helps explaining why the apprenticeship labour contract works in a country rather than in another (Dustmann and Schönberg 2012).

In this paper we focus on Italian apprenticeships for two main reasons. The first follows an economic rationale. We expect that a commitment to invest in human capital can contribute to the creation of a good quality job and the consequent permanent employment position. If this is the case, apprenticeships have an advantage over the other labour contracts to lead to permanent employment. In fact, there are several channels through which human capital can be accumulated but the most relevant are education and vocational training. Apprenticeships are committed to provide vocational training programs to which they add general education outside the firm. The costs of this human capital investment are shared between the employer and the employees. The second reason follows a technical rationale. We start by assuming that the data generating process of the apprenticeship rate is based on the legal rule that job entry as apprentice is only available, albeit not mandatory, up to 29 years and 364 days of age.[1] This yields to a deterministic process of the apprenticeship rate on one side of the cutoff of 30 years. As a consequence the data generating process of the permanent employment rate exhibits a discontinuity around the cutoff of 30 years of age. This discontinuity in the permanent employment rate can depend on the apprenticeship labour contract only. There is no reason to observe such data generating process of the permanent employment rate in case of transitions from either unemployment or from a temporary labour contract. On the top of that, we expect that the introduction of law no. 92/2012 has exogenously changed this data generating process. In fact, the law explicitly aimed at encouraging apprenticeships as the main port of entry into permanent employment by implementing three measures. First, the enforcement of a mentoring scheme which is expected to increase the benefit of the apprenticeship training in terms of higher worker's productivity. Second, the introduction of a punishment on the

[1] This implies that we are considering apprenticeships as labour contracts committed to the provision of on the job training and of general education courses outside the firm. The role of apprenticeships as part of the vocational education and training system, alternative to a more academic education track, is here neglected.

firms which do not accomplish with the commitment of employing permanently a certain percentage of apprentices (excluding motivated lay-offs in the calculation of such percentage). Such firms cannot hire more than one apprentice in the future. This punishment increases the worker's value of apprenticeship and discourage the production-oriented, in favour of investment-oriented, usage of the contract. Third, labour costs of temporary contracts were increased while holding fixed the benefit of a tax rebate to hiring an apprentice. This setting allows us to design a difference in discontinuity regression model (Grembi, Nannicini and Troiano 2016). That is, the difference in the discontinuity around the cutoff of 30 years of age, generated by the labour market reform, creates a source of randomised variation.[2]

Whether the creation of jobs translates into the creation of good jobs relies on the dynamic pattern of earnings and permanent employment (from the employee perspective) and productivity (from the employer perspective). Evidence of dynamic treatment effects of job training for employed workers is much more limited than the huge literature on (dynamic) treatment effects of training for unemployed or dislocated workers. Recently, Rodríguez, Saltiel and Urzúa (2018) and Albanese, Cappellari and Leonardi (2017) provide evidence on this issue. However, there is no paper that addresses the issue of whether firms are less reluctant to hire workers on a permanent basis using a labour contract committed to a human capital investment rather than another labour contract. Looking at the dynamics of the impact on permanent employment of the initial human capital investment provides an important evidence on the main argument of the paper. The combination of a committed human capital investment in a open-ended contract (as apprenticeships) drives the screening-sorting process that lead to permanent employment, on the top of human capital accumulation. If this the case, the probability of creating a job of good quality increases. If this is true, the probability that this job match persists over time is higher than the same probability of other job matches created without the same commitment to the human capital investment. In the lack of medium-run effects on the permanent employment rate it would be hard to defend this interpretation even in presence of sizeable

[2]Our framework is slightly different from the one enlightened by (Grembi et al. 2016). While they use the difference in discontinuities to control for a confounding policy at the threshold, we are instead exploiting the difference in the same discontinuity policy rule, generated by a labour market reform, as a randomised source of variation to retrieve the effect of interest.

impacts at the baseline. We, therefore, retrieve the static ITT parameter at the baseline by comparing cohorts treated by the labour market reform to similar untreated individuals at the threshold of 30 years of age. We then extend our analysis to a dynamic setting up to three years after the introduction of the new regime (Cellini, Ferreira and Rothstein 2010).

Data are taken from a very rich dataset, "Campione Integrato delle Comunicazioni Obbligatorie" (CICO) by the Ministry of Labour and Social Policies. Our baseline sample is centered in a ± 24 months interval around June 2012 when law no. 92 was issued.

Our results suggest that at the threshold of 30 years of age the permanent employment rate of individuals affected by law no.92/2012 increased by 1% when compared to the permanent employment rate of similar untreated individuals. There is no evidence of a positive impact on the employment rate and on the employment rate at the firm or sector for which each individual last worked ruling out the possibility that the result is driven by conversions from temporary to permanent employment. This interpretation is further strengthen by our evidence on the dynamic effects. The dynamic ITT parameter on permanent employment after 36 months amounts to about 5%. In the same order of magnitude is the effect on permanent employment at the same firm (sector) for which each individual last worked. The dynamic impact on the employment probability is statistically different from zero but quite small (.01%) while the self-employment probability is slightly negatively affected (-.01%). Our findings support the view that a labour contract which invests in human capital serves as a stepping stone into permanent employment.

The rest of the paper is organised as follows. Section 2 outlines the setting. Section 3 describes the data and the preliminary analysis while section 4 illustrates the identification strategy. Results are reported in section 5. Finally, section 6 concludes.

2 The setting

2.1 Brief review of the literature

This paper contributes to an extensive literature that analyses the labour market outcomes of vocational training. The reduced form impact on earnings and employment probability of government-supplied training programs is widely documented (see also Heckman et al. 1999, Kluve 2010, Card, Kluve and Weber 2010, Card, Kluve and Weber 2018). A large number of these studies have identified the training effects through experimental settings.

In the US case, evidence has been produced from randomised evaluation of Job Training Partnership Act (JPTA) (see Bloom, Orr, Bell, Cave, Doolittle, Lin and Bos 1997, Heckman, Ichimura and Todd 1997, Heckman and Smith 2000), Job Corps, (see for instance Schochet, Burghardt and McConnell 2008, Flores-Lagunes, Gonzalez and Neumann 2010) and several other programs. For instance, Bloom et al. (1997) estimate the effects on the earnings of the publicly funded Job Training Partnership Act enacted in 1982, providing job-training services, job search assistance and basic education for those facing significant employment barriers. The authors followed out-of-school youths for 30 months after the treatment finding no earnings gains for program participants relative to the control group. Similarly, long term results by the US General Accounting Office (1996) show no significant difference in 5-years earnings and employment patterns between JPTA treated and control groups for both genders.[3]

The US Job Corps Program of 2005 consisted of remedial education, vocational training and job search assistance targeting individuals aged between 16 and 24 coming from a low-income family. Lee (2009) estimates the effects of Job Corps program disentangling the effect on the wage rate from the impact on the hours of work showing that the program also raised participants' wages consistently with human capital models. The author shows that, after 208 weeks since random assignment, the employment rate of those treated by the program is 17% higher than the employment rate of those untreated. Job Corps appears to have had significant short and medium-run effects on both genders[4] but not long-term impacts (5 to 8 years since random assignment) on earning and employment probability (Schochet et al. 2008).

In meta-analyses of evaluation of studies of active labor market policies in US and Europe, Card et al. (2010), Card et al. (2018) and Kluve (2010) find that on the job training has moderate effects on earnings and employment probability. Moreover, labour market policies are more effective for woman than for man. On the job training show

[3]Abadie et al. (2002) estimate the effects on the distribution of earnings of JTPA. The authors find that for women the impact is proportionally stronger at low quantiles of the earnings distribution, while for men only in the upper part of the distribution the positive effect of the program is statistically different from zero.

[4]This results are not valid for all the youths. For example, Flores-Lagunes et al. (2010) examine the impacts of Job Corps on Hispanics finding no earnings gains. The authors show that these results are related to the different local labor market conditions Hispanics face.

better labour market outcomes than classroom training and private sector programs are found to be more effective than public sector programs. Quasi experimental evaluation of youth training program in Latin-American countries generally find positive impacts of unemployed and youth training programs (Alzúa, Cruces and Lopez 2016, Attanasio, Kugler and Meghir 2011, Attanasio, Guarin, Medina and Meghir 2015, Card, Ibarraran, Regalia, Rosas-Shady and Soares 2011)

Recent developments of the literature has mainly pointed to the importance of providing dynamic treatment effects of training. In fact, estimates on the dynamic returns to training for unemployed workers are documented by several papers. Crépon, Ferracci, Jolivet and van den Berg (2009) estimate treatment effects of training program in a dynamic setting for the unemployed adult in France finding that training has little impact on unemployment duration. In a subsequent paper, Crépon, Ferracci and Fougere (2012) estimate the effects of duration of training program finding that longer training spells cause longer unemployment spells, but also more stable jobs. Studies on long-run effects of different types of government-sponsored training in West Germany (Lechner, Miquel and Wunsch 2011, Fitzenberger, Furdas and Sajons 2016) find positive and significant impact on the employment probability. Osikominu (2013), finds the longer program are effective in creating stable jobs.

The evidence on (dynamic) treatments effects of training for employees is instead rather limited. It is even more limited if the training is not publicly provided and funded. Recently, Rodríguez et al. (2018) and Albanese et al. (2017) estimate dynamic treatment effects of job training for employed workers. The former paper focuses only on earnings and considers a publicly funded training program in Chile. The latter paper focuses on both employment outcomes and earnings of apprentices before and after the reform introduced by law no 30/2003 in Italy. Our paper contributes to the ongoing debate by showing that a labour contract that invests in human capital, financed by both individuals and firms, has some advantages over the other labour contracts to create good jobs. Hence, this work also contributes to the empirical literature testing the port-of-entry hypothesis of different type of temporary contracts (Holmlund and Storrie 2002, Booth, Francesconi and Frank 2002, Heinrich, Mueser and Troske 2005, Ichino, Mealli and Nannicini 2008, Berton, Devicienti and Pacelli 2011). In what follows, we briefly describe the Italian institutional setting that helps explaining why Italy is an interesting context to retrieve the parameters of interest.

2.2 Institutional framework

On the verge of the new millennium the Italian labour market was characterised by several critical issues many of them involving young people. In order to tackle the high level of youth unemployment; the high percentage of young people not in employment, education or training; the low labour mobility; the low job-to-job mobility and the long-term unemployment, Italian governments have undertaken several liberalisation measures (law no. 196/1997, legislative decree no. 368/2001, legislative decree no. 276/2003 and law no. 183/2010) to achieve greater flexibility in employment. Following the reform introduced by legislative decree no. 276/2003, employment different from the standard open-ended contract amounted to 22 types of labour contracts corresponding to 48 atypical forms. The new regime of fixed term contracts generated productivity losses encouraging the substitution of temporary employees in favour of external staff. In contrast law no. 30/2003, which reformed the apprenticeship contract, generated an overall productivity enhancing effect increasing job turnover and substituting of external staff with firms' apprentices (Cappellari, Dell'Aringa and Leonardi 2012). All in all, two main problems characterise the Italian labour market. First, a strong divide between low-income temporary workers and those working on permanent basis leads to a dualisation of the labour market further enhanced by disguising salaried employment as various forms of self-employment. Second, transition to stable employment is difficult since employers are very reluctant to hire individuals on a permanent basis, especially in absence of human capital embodied in them.

One of the main aims of law no. 92/2012, was to fight the improper demand of atypical contracts encouraging the usage of the apprenticeship as the main port of entry into permanent employment. The pursuit of this goal resulted in strengthening the training provided by the apprenticeship labour contract while tightening the rules governing access to temporary contracts. The main policy interventions of the 2012 reform can be summarised as follows. Since 2013 employers hiring fixed-term workers were required to finance the new Social Insurance for Employment paying 1.4% of the individual's earnings. If this employment contract was converted into an open-ended one, or if the worker was hired within six months of the end of the fixed-term contract, the employer was eligible to be refunded for the contribution paid for an amount up to six months salary. Some of the existing rules governing apprenticeships were changed to ensure adequate training and education for potential employees and by discouraging the inappropriate use of this type of contract. A minimum period of six months was foreseen except for seasonal work, for which only vocational apprenticeship contracts are admissible. The law fixed also the maximum duration of the contract to three years (or five years for artisans jobs), of which out-of company training for basic and general skills and knowledge covers a maximum of 120 hours in total. In order to promote skilled employment, the reform established an adequate mentoring scheme to the apprentice. Employers with more than 10 employees were allowed to hire three apprentices for every two employees, compared to the previous ratio of one to one which still applied to employers with fewer than 10 employees. However, employers who, in the same job occupation of the apprentice do not employ qualified workers at all or who employ fewer than three, could hire no more than three apprentices. In order to enforce apprenticeships as a permanent labour contract the reform set new limitations to the inappropriate use of the contract on a temporary basis.[5] Employers with more than 10 employees could not hire more than one new apprentice at a time if the percentage of apprentices hired on open-ended contracts over the previous 36 months was less than 50% (30% for the first 36 months after the reform). These percentages exclude dismissals for justified reasons, for just cause, for resignation or for failure to pass the trial period.

As established by law no. 30/2003 in Italy there are three types of apprenticeship labour contract: apprenticeship for vocational qualifications and diplomas, upper secondary education diplomas and high technical specialisation certificates (type 1); vocational apprenticeship (type 2); higher education and research apprenticeship (type 3). Type 1 is for those aged 15 to 25 and it can be included in the category of vocational education and training (VET) programmes at upper and post-secondary schooling levels. Type 2 is for those aged 18 to 29. Type 3 is for those aged 18 to 29 and includes two

[5]An employer who hires an apprentice benefits of a tax rebate whose amount depends on firm's size. Part of the costs of the human capital investment are burdened on the workers in the form of a lower initial wage.

sub-types: apprenticeship for higher education and training (i.e. university degrees, PhDs, and higher technical institute diplomas) and apprenticeship for research activities. In our empirical analysis we mainly refer to vocational apprenticeship and we do not consider apprenticeship for vocational qualifications and diplomas. Consequently, the mechanisms we underline to hold in our empirical evidence may differ from those that explain the usage of the apprenticeship labour contract (and related labour market outcomes) as vocational education and training system alternative to an academic track. In what follows, the apprenticeship labour contract is a committed on-the-job training program that is partly financed by the firm and that provides general skills and competencies in an open-ended contract.

3 The data and preliminary analysis

3.1 The data sample

In the estimation we make use of a very rich administrative dataset by the Ministry of Labour and Social Policies, *CICO* (the so-called *Comunicazioni Obbligatorie*). The database include, since 2009 to the second quarter of 2017, detailed information on the flow of different type of contracts, activated, transformed and dismissed, for dependent and independent workers for all sectors including the Agricultural sector and Public Administration. The relevant dates (day, month, year) of each event are available in the database together with the type of labour contract, the sector, the region of work and an anonymous identifier for both the firm and the worker involved and the type of benefit associated to the contract.[6] For each worker, we have information on the gender, the year of birth, the region of birth, citizenship, and education.[7] We keep missing information related to the educational level including an indicator which controls for this status. Finally, we further reconstruct the individual's past work experiences since 2009. The dataset is randomised

[6]We broadly reclassify benefits as no benefits, hiring incentives; reduction of social security contributions; benefits related to the apprenticeship labour contract; social insurance benefits.

[7]We classify education as primary school; lower secondary school; upper secondary school (vocational oriented); upper secondary (academic oriented); diploma degree; university diploma; university degree; bachelor (3 years degree, new university system introduced in 2001); university degree (4 or more years degree, old system); master degree; university degree (4 or more years, new system introduced in 2001); master degree (2 years which add to the bachelor, new system); post-graduate degree; master degree (first level, post-graduate); post-graduate diploma of specialisation and PhD.

on the basis of the worker's birth date (the 1st, the 9th, the 10th and the 11th of each month in a given year for each cohort of birth). The main drawback of the data is that they do not contain information on workers' earnings although it is recorded the first month salary for each job spell.

We start selecting an age interval from 15 to 40. To provide an idea of the representativeness of the sample, in the age range between 25 and 34 the number of the observations amounts roughly to 13% of the universe of Italian job flows.[8] For each year available (2009-2017) we keep all records of job spells and working status of all individuals month by month since the year when they have their first job spell recorded.[9] We exclude only job spells which are related to labour contracts which are usually not recorded through the main application form, *Unificato LAV*. More specifically, we exclude job spells of domestic workers hired by households which are directly recorded by the Italian Social Security Institute, INPS, since the employer is not a firm; observations regarding the agency labour contracts since usually these kind of contracts are registered by the agency through another application form *Unificato SOMM*; observations regarding job spells in the maritime sector which are recorded by a specific application form, *Gente di Mare*, and finally the dataset does not include all information that can be retrieved by the application which directly controls for entry and exit of firms, *Unificato Variazione Datori di Lavoro*. We exclude about 6% of the observations from 2009 to 2014 and about 5% from 2015 to 2017.[10] However, past work histories of individuals are calculated considering all the job episodes (i.e. including the observations described above). We merge this dataset with two databases: one which records self-employment activities and the other which accounts for independent job episodes in the professional orders. We are, therefore, able to consider the self-employment probability as an outcome of interest. Moreover, we exclude both types of independent work as a possible reason of being out of the dependent employment status which can only be due to either unemployment or being out of the labour force. However, we are not able to distinguish between these two statuses.

[8]The number for the universe has been taken from several issues of the Annual Report on Mandatory Communications by the Ministry of Labour and Social Policies.

[9]For instance, an individual recorded for the first time in 2010 could have experienced other job spells in 2008 which we are unable to track. We are, instead, able to know the starting date of all job spells which end since 2009 even if the starting date occurred several years before.

[10]For instance in 2014 in the age range 15-40 we exclude 892363 observations related to domestic job spells, 14488 related to the agency labour contracts and 845 to the maritime sector over 16071750 observations.

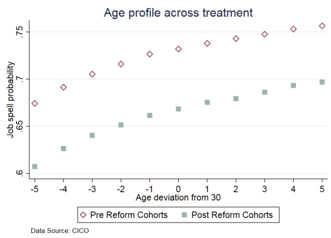
Our working sample is centered in a ± 24 months interval around June 2012 when law no. 92 was issued. This implies that at each age considered it is possible to identify those treated (from July 2012 to June 2014) and those untreated (from June 2010 to June 2012) by the reform gathering two and half affected and unaffected cohorts. Since we do not have information on the precise date of birth of the individual, we measure the age at the 31st December of the previous year to minimize measurement in its definition (we will come back to this issue in the next subsection). That is to say, for example, that in 2012 an individual is aged 29 with certainty if she is born in 1982 and she is turning to 30 in an unknown month during that year. We restrict our sample to an age interval of ± 5 years around 30. After this selection the sample includes 32,044,063 observations involving 1,114,731 workers and 738,111 firms. In the same age range, when we consider those who started either a job spell or a self-employment activity in a given year, the sample consists of 12,031,489 observations involving 744,282 individuals and 549,513 firms. When we restrict the age range to ± 1 (± 2) year(s) around the threshold of 30, we end up with 1,749,041 (3,512,232) observations gathering 164,395 (273,047) individuals and 145,791 (239,606) firms.

3.2 Graphical analysis

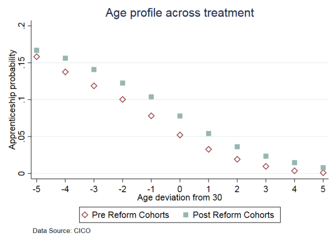
In Figure 1 we show the difference in discontinuity upon which the identification strategy relies. This difference in discontinuity is generated by law no 92/2012 around the age threshold of 30 years.

We consider three different outcomes in terms of employment: the two main targets of the law, the apprenticeship and the permanent employment probability, and the probability of a job spell of at least 15 days (employment probability). Since 2008 apprenticeship labour contracts in Italy are legally recognised as open-ended contracts while they were previously considered fixed-term. For this reason, we impute a value of 1 to the permanent employment indicator for apprentices. The first row of Figure 1 plots the age profile in the age interval 25-35, calculated as deviation from 30, including in the sample also the job spells which started in the previous year(s). The dots indicate averaged raw data.

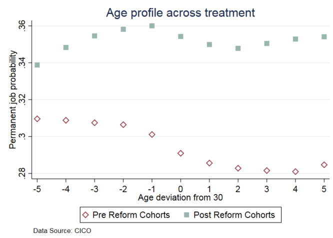
The employment probability (panel (a)) is increasing in age and higher for untreated



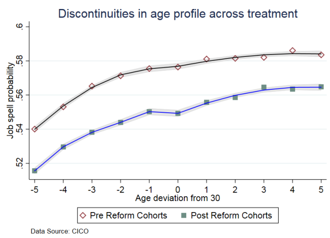
(a) Employment rate.



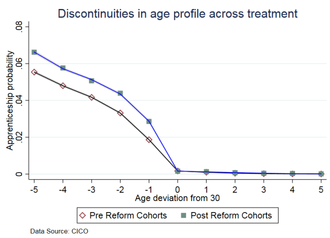
(b) Apprenticeship rate.



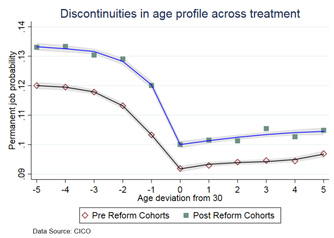
(c) Perm. Employment rate.



(d) Employment rate.



(e) Apprenticeship rate.



(f) Perm. Employment rate.

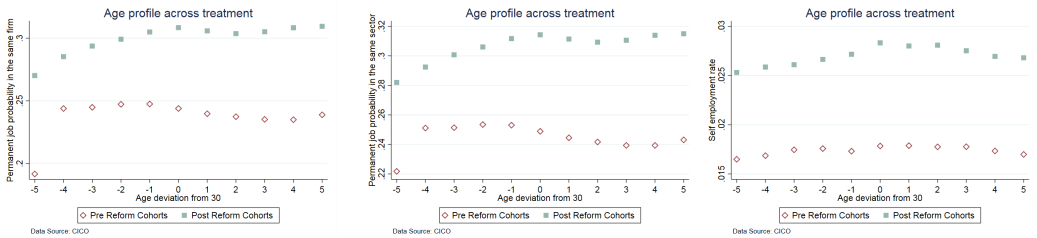
Notes: The dots indicate raw data while the line and the gray area refer to the parametric fit (third order polynomial in age) and its 99% confidence intervals. Heteroskedasticity robust standard errors.

Figure 1: Difference in discontinuity across contiguous cohorts generated by law n.92/2012 at the age cutoff

cohorts. In contrast, while approaching the threshold of thirty years of age, the permanent employment probability (panel (c)) reduces and this pattern holds for both treated and untreated cohorts. Moreover, at all ages, treated cohorts have a higher permanent employment probability. The apprenticeship labour contract lasts more than one year for a maximum of six years before the reform and 3 years after the reform.[11] Panel (b) of the Figure shows that when we do not condition the data on the year when the labour contract starts, the apprenticeship probability is continuous at the threshold albeit, as expected, decreasing in age because of the age limit. At all ages, the treated cohorts have a higher probability of apprenticeship providing suggestive evidence of a overall positive effect of the reform. This is further suggested by the second row of the Figure which plots the same outcomes when the sample is restricted to those who started a job spell (even of one day) or a self-employment activity in a given year of the sample period. The dots indicate averaged raw data while the line and the gray area refer to the parametric fit (third order polynomial in age) and its 99% confidence intervals. The parametric fit is es-

[11]The maximum duration of the contract could be extended by the collective agreements.

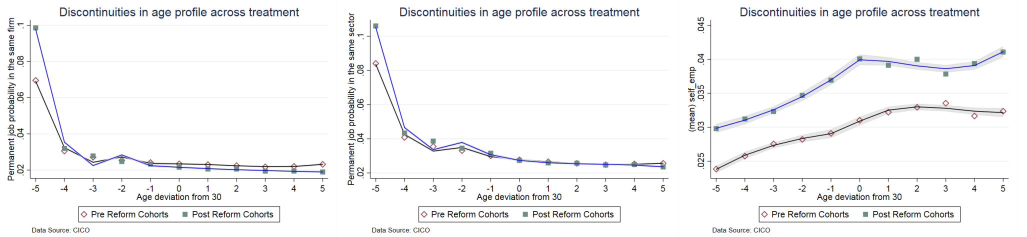
timated separately at the two sides of the age cutoff and in the two (before/after) regimes of the labour market reform. The permanent employment probability (panel (f)) clearly shows a discontinuity at the age threshold when job entries as apprentices are not possible. Instead, there is no clear evidence of a discontinuity in the employment probability (panel (d)). Overall the Figure is indicating that in the age range of ± 1 year averaged raw data are well centered into the polynomial fit only around the age threshold.



(a) Permanent employment in the same firm.

(b) Permanent employment in the same sector.

(c) Self employment.



(d) Permanent employment in the same firm.

(e) Permanent employment in the same sector.

(f) Self Employment.

Notes: The dots indicate raw data while the line and the gray area refer to the parametric fit (third order polynomial in age) and its 99% *confidence intervals. Heteroskedasticity robust standard errors.*

Figure 2: Difference in discontinuity across contiguous cohorts generated by law n.92/2012

To provide a comprehensive view on the employment effects of the law at the threshold, in Figure 2 we extend our analysis to two different types of permanent employment (conditioning on working at the firm or sector for which the individual last worked) and to an evaluation of the effect on the self-employment probability. The first row of Figure 2 displays higher permanent employment outcomes (panels (a) and (b)) at all ages for cohorts treated by the labour market reform. Nevertheless, when we restrict the sample to job entries in a given year (second row of Figure 2) any difference in discontinuity at the age threshold can be gauged. This rules out the possibility that the positive impact on the transition into permanent employment illustrated by Figure 1 can be associated to conversions from temporary to permanent employment within the same firm or sector. Panels (c) and (f) of the Figure do not detect any evidence of an impact on the self-employment probability.

3.3 Identification through a parametric functional form

Figures 1 and 2 pave the way to build an identification strategy which relies on the difference in discontinuity around the age threshold. However, to validate our analysis we need to discuss an important issue. We have information on the individual's year of birth only. As a consequence, we have to deal with a regression difference in discontinuity inference with discrete support and specification error. This latter issue could not be sorted out even with the precise worker's birth date. This is because only selected birth dates enter into the sample. Nevertheless, it is possible to calculate the probability that this error is positive for those who turn their 30th year of age and they are imputed an age of 29. For instance, in January the error is positive if the individual is born in that month. This occurs with probability $\frac{4}{12}$ (i.e. one over twelve months times the 4 possible dates) while the error is zero with probability $\frac{4*11}{12}$ if the worker is born in another month and so on. Since the randomisation process of the data does not change across repeated samples, if the employment rate, at a given age and for each month of birth, is constant across affected and unaffected cohorts, the difference in discontinuity at the threshold of 30 years cancels out these error terms. In such a case the specification error vanishes out since the same specification error prevails in the counterfactual world of contiguous similar, albeit untreated, cohorts. If, instead the employment rate, at a given age and for each month of birth, is not constant the difference in these errors terms between affected and unaffected cohorts might be random.

We follow Lee and Card (2008) who argue that discreteness in the treatment-determining covariate implies that the treatment is not identified without assuming a parametric functional form. We perform the goodness-of-fit F statistic,[12] suggested by the authors, to help us choose a reasonably accurate regression model. That is, we test whether the functional

[12]The F statistic is calculated as $\frac{\frac{ESS_R - ESS_{UR}}{G}}{\frac{ESS_{UR}}{N-J}} \overset{H_0}{\sim} F(\text{G,N-J})$ where $G = J - K$ are the number of restrictions, i.e. the difference between the number of parameters J in the unrestricted model and the number of parameters K in the restricted model.

Table 1: Functional form restrictions to flexible parametric specifications.

	Without DiD specification			DiD Model specification		
	[-1,1]	[-2,2]	[-3,3]	[-1,1]	[-2,2]	[-3,3]
First Order Polynomial						
FPFF	Polynomial degree zero	YES	YES	Polynomial degree zero	YES	YES
Second Order Polynomial						
FPFF	NO	NO	NO	NO	NO	NO
Third Order Polynomial						
FPFF	NO	NO	NO	NO	NO	NO
Fourth Order Polynomial						
FPFF	NO	NO	NO	NO	NO	NO

Notes: FPFF stands for a Flexible Parametric Functional Form that allows for all possible interaction terms.

form adopted is statistically equal to an unrestricted regression of the outcome on the full set of dummy variables (full set of dummies times the indicator function of the labour market reform for the difference in discontinuity model specification) for the possible values of age which define the age range. If the statistic exceeds the critical values CV, the null hypothesis is rejected suggesting that the polynomial function is too restrictive. The lower the value of the test than the critical value, the higher the confidence on the validity of the estimated effect. However, since we have information only on the year of birth, we are forced to use a bandwidth of one year of age. Consequently, around the age threshold, when using higher order polynomials, there are not enough degrees of freedom. In fact, the unrestricted model of the test can use a very limited number of age dummies. On the top of that, there is a strong degree of collinearity between the forcing variable age and the treatment indicator of being below or up the age cutoff. As a result, we cannot often adopt a flexible parametric functional form to approximate the true conditional expectation function. Table 1 summarises the functional form restrictions imposed by our data. In fact, the flexible parametric model specification that allows for all possible interaction terms (i.e. that allows for a different slope around the age cutoff and/or before/after the reform) can be used in the linear case only. Moreover, in the age range of ± 1 year around the threshold, the local linear regression collapses to a polynomial of degree zero in age.

Table 2 illustrates the F goodness-of-fit statistic applied to the permanent employment

Table 2: F goodness-of-fit statistic

	Without DiD specification			**DiD Model specification**		
	[-1,1]	[-2,2]	[-3,3]	[-1,1]	[-2,2]	[-3,3]
First Order Polynomial						
F			26.966			16.041
CV			4.605			3.319
p val	1.000	1	2.057×10^{-10}	1.000	1	1.569×10^{-9}
Second Order Polynomial						
F			59.765		0.493	26.024
CV			4.605		4.605	2.802
p val		1	4.035×10^{-24}		0.995	8.348×10^{-28}
Third Order Polynomial						
F			100.780		0.986	25.534
CV			6.635		6.635	3.017
p val		1	3.490×10^{-20}		0.995	5.236×10^{-22}
Fourth Order Polynomial						
F						4.685
CV						3.319
p val		1	1		0.995	0.067

Notes: The null hypothesis of the F goodness-of-fit statistic test is that the functional form adopted is statistically equal to an unrestricted regression of the outcome on the full set of dummy variables for the possible values of age (age times the indicator function for being treated by the labour market reform for the difference in discontinuity model specification) which define the age range reported in brackets. If the statistic exceeds the critical values CV, the null is rejected. The p-value refers to the p-value of an F test on the joint significance of the age dummies (age dummies times the labour market reform treatment indicator) in the auxiliary regression.

outcome.[13] When we use polynomial functions higher than 1 some restrictions are imposed on the data. It is mainly controlled for the polynomial in age without controlling for all possible interaction. These restrictions imply that the slopes around the age cutoff and before/after the reform are equal. The Table reports also the p-values of the F statistics on the joint significance of the age dummies (age dummies times the labour market reform treatment indicator) in the auxiliary regression. As expected, the two F tests provide the same information. However, the p-value on the joint significance of the age dummies is useful when there are not overidentifying restrictions to perform the F goodness-of-fit statistic. When this p-value is 1 or the value of the F goodness of fit statistic approaches 0, we are relatively confident that the estimated (difference in) discontinuity parameter approximates the (difference in) discontinuity in averaged raw data.

Moreover, Table 2 shows that a local linear regression could be adopted only when we

[13] Similar tests for all the other outcomes are reported in Appendix B.

restrict the sample to an age range of ± 1 and ± 2 year(s). The statistics are consistent with panel (f) of Figure 1 where in the age range of ± 3 years around the cutoff, raw data are outside the confidence intervals of the estimated parameter using a third order polynomial in age. In contrast, the statistic fails to reject the null hypothesis in the age range of ± 2 years around the cutoff using a first (or second or third) order polynomial in age while panel (f) of Figure 1 reveals that raw cell data at the age of -2 for cohorts treated by the 2012 reform, are not contained in the confidence intervals. We have not clustered standard errors, so far. Results do not change if we cluster standard errors by age, year of birth and region of birth to allow for possible within group correlation at this level. In what follows we will discuss this issue while presenting the difference in regression discontinuity design.

4 The difference in regression discontinuity design

4.1 Static model

Let's define n the fraction in the population at a given age a of those who start either working or a self-employment activity in a given year. The residual fraction $1 - n$ is made of those, aged a, who have started the current work in the previous year(s) and those who were unable or unwilling to start a job spell (even of one day) in the given year. Law no. 30/2003 fixed the maximum age at which job entry as vocational apprentice is possible at 29 years and 364 days. Law no. 92/2012, changed for this fraction n, the data generating process of the apprenticeship rate. If the apprenticeship labour contract has an advantage over the other contracts to lead to permanent employment, the introduction of the labour market reform has generated a randomised source of variation that allows a difference in regression discontinuity design. This is because there is not other reason to observe a discontinuity in age in the permanent employment probability if it is not related to the legal rule on the job entrance as an apprentice. This claim is supported by Figure 1. As discussed in the previous paragraph, our data limitations impose some restrictions on the regression model specification. However, the dataset is large enough to focus on the age range ± 1 year around the cutoff. As a result, the model specification is the following:

$$y_{it} = \alpha_0 + \alpha_1 r_{it} + \gamma_0 d_{it} + \gamma_1 d_{it} r_{it} + \epsilon_{it} \tag{1}$$

where d_{it} is an indicator function that takes the value of 1 if the individual i at time t is aged less than 30 and r_{it} is an indicator function if the individual i at time t (i.e. given her year of birth conditional on age) is treated by the 2012 reform.

Using the potential outcomes framework, in Appendix A1 we show that the parameter γ_1 identifies the difference in discontinuity at the age threshold. We claim that in the age range of ± 1 year around the cutoff, the general model specification collapses to equation 1 which corresponds to a polynomial of degree zero in age. This is because only locally the zero slope assumption holds. In fact, the parameter retrieved, γ_1, captures locally the difference in the jump of the intercept of the conditional expectation functions. As long as age is expressed in years in the age range of ± 1 (which amounts to comparing before and after the reform those age 29 with certainty and turning into 30 to those aged 30 with certainty and turning into 31), it is not possible to distinguish between the indicator function d_{it} and the forcing variable age.[14] Equation 1, therefore, imposes the restriction that the slope in age is constrained to be identical on both side of the cutoff and equal to zero. In Appendix A1 we generalise the linear to be estimated in age range larger than ± 1 year around the cutoff. As illustrated in Tables 1 and 2, this generalised model can be estimated using a first order polynomial in age only in the age range of ± 2 years around the threshold. However, will show that γ_1 estimated by equation 1 is consistent with the differences in means around the threshold in raw data as displayed in Figure 1.

The parameter γ_1 is the first parameter which we are able to retrieve, a static reduced-form Intention To Treat, ITT, parameter on permanent employment (and other definitions of employment outcomes), which is possibly the final target of the reform. Its interpretation simplifies to measuring to what extent, around the age threshold, the outcome of interest changes for cohorts treated by law no 92/2012 compared to similar individuals born in contiguous cohorts who reached the threshold age before the introduction of the law. In the next paragraph, we will show that this differential impact is compared across cohorts

[14]The former takes the value 1 if the latter takes the value -1 and the former takes the value 0 if the latter takes the value 0.

who are similar in terms of both average observable and possibly average unobservable characteristics. This is the consequence of the randomised variation generated by the 2012 reform around the legal age limit to job entries as apprentices that creates a discontinuity in the data generating process of the permanent employment rate.[15]

4.2 Estimation issues

Since the treatment of the apprenticeship labour contract is only available but not mandatory on one side of the threshold, the model matches up with a special case of fuzzy regression discontinuity design and it is identical to the sharp design in terms of necessary identification assumptions. Consider, instead, to disregard the discontinuity at the threshold age of 29 years and 364 days and design an identification strategy which corresponds to a regression discontinuity design that uses the reform to generate, at a given age and at a given point in time, a discontinuity across years of birth (Malamud and Pop-Eleches 2010). In fact at a given age, the cohort of birth randomly assigns the individual to the treatment of the reform. As a result, it is possible to exploit the variability across cohorts by considering exogenously defined groups exposed to different rules for obtaining an apprenticeship labour contract. In such a case, to estimate a causal parameter within the age range 15-29 requires functional form assumptions on how the effect evolves over age. This is not an easy task if the imposed identifying restrictions can not be rationalised on substantive knowledge about the selection process. Moreover, this is a fuzzy regression discontinuity design that requires a monotonicity assumption. Marginal benefits and costs (including the opportunity costs) of the apprenticeship labour contract are likely to vary over the age dimension. On the top of that, marginal costs and benefits are likely to be different between the two main types of the apprenticeship contract (VET apprenticeship program and vocational apprenticeship) that coexist in the age interval from 18 to 24 years (and 364 days). The difference in discontinuity design applied in the age range of ± 1 year around 30 allows us to overcome these issues. In presence of essential heterogeneity the optimal age at which entering into the labour market as vocational apprentice differs across individuals and firms. Nevertheless, under the difference in discontinuity design, it is not relevant if

[15] In the Appendix A2, we will discuss whether it is possible to retrieve the Average Treatment Effect, ATE parameter, in the population around the age threshold.

the age threshold is suboptimal. Both individuals and firms have some influence on the vocational apprenticeship probability since the age cutoff is known in advance by them. However, the design is valid if they are unable to precisely manipulate the age (whether optimal or not) at which this event may occur (if occurs). As a consequence, the variation in treatment around the threshold is randomised similarly to a randomised experiment (see Lee and Lemieux 2010). Ignorability, or unconfoundedness assumption is therefore satisfied. On the top of that, in the difference in discontinuity design, the imprecise control over the forcing variable does not substitute but rather complements the overlap condition because the comparison is between contiguous and, then similar, cohorts. This difference in discontinuity design can be conceived as a limiting case of the difference in differences estimator.[16] In fact, locally, the common trend assumption is trivially satisfied. However, since the source of variation is randomised at the threshold, the difference in discontinuity design does not need to condition on those exogenous variables that lead to differential trends as the difference in differences does.

We cluster standard errors by age, year of birth and region of birth. All possible employment outcomes of individuals, who are aged a at time t given their year of birth, tend to be correlated within the same region of birth. This is because these individuals share background characteristics (such as the quality of the educational achievements) and are exposed to the same local labour markets. As discussed by Angrist and Pischke (2009), heteroskedasticity rarely leads to dramatic changes in inference. In contrast, clustering can make all the difference. The general Moulton formula suggests that clustering has a bigger impact on standard errors when the correlation of regressors within groups is large. The regressor(s) of interest may also vary at the individual level and for different group sizes. The relevance of this assumption can be easily verified. If the regressors values are uncorrelated within groups, the grouped error structure does not matter for standard error. For this reason, for robustness check our main results will be replicated using a different group structure (age and year of birth only) and the heteroskedasticity-robust standard errors. As shown by Kolesár and Rothe (2018), the latter are more conservative when the window widths (i.e. the age range) is small and the forcing variable is discrete.

[16]We prove this remark in Appendix A1.

4.3 Dynamic model

We follow Cellini et al. (2010) to extend our analysis to a dynamic setting. We can write the outcome τ years later as:

$$y_{i,t+\tau} = \alpha_0 + \alpha_1 r_{it} + \gamma_0 d_{it} + \gamma_1 d_{it} r_{it} + \epsilon_{i,t+\tau} \tag{2}$$

The difference in discontinuities strategy ensures that the error term $\epsilon_{i,t+\tau}$ is independent of both treatment $(d_i; r_i)$ status. Nevertheless, it could reduce precision because it has an important component that varies at the individual level but it is fixed within individuals over time. Therefore, the effects retrieved are not precisely estimated. Our model specification accounts for (permanent) employment in-flows. After two years from the baseline all individuals are treated by the reform and what makes the difference is the age at which they are affected. Therefore, we expect that the impact fades away over time since those, who enter into the labour market thereafter, could be affected by the reform albeit at an age higher than 30 while the permanent employment status of those who benefitted from the labour market reform at the baseline does not change.

Moreover, equation 2 does not explicitly take into account the persistency in outcomes generated by the exogenous shock of the reform. To deal with this issue and to model how the effect of interest evolves over time, we estimate:

$$\begin{aligned} y_{i,t} = \alpha_0 + \alpha_1 r_{it} + \gamma_0 d_{it} + \gamma_1 d_{it} r_{it} + \\ + \phi_\tau \sum_{\tau=1}^{\tilde{\tau}} (\alpha_1 r_{i,t-\tau} + \gamma_0 d_{i,t-\tau} + \gamma_\tau^{TOT} d_{i,t-\tau} r_{i,t-\tau}) + \epsilon_{i,t} \end{aligned} \tag{3}$$

where ϕ_h measures the degree of time persistency in the output generated by the exogenous shock and γ_τ^{TOT} identifies the average Treatment effect On the Treated (TOT).

When the shock occurs therefore ϕ_h is equal to zero and the static ITT parameter corresponds to γ_1 described in equation 1. In fact, the dynamic ITT impact is equal to:

$$\gamma_{\tau}^{ITT} = \gamma_{\tau}^{TOT} + \sum_{h=1}^{\tau} \gamma_{\tau-h}^{TOT} \phi_h \tag{4}$$

We now present our main results.

5 Main results

5.1 Model validation

Observable and unobservable characteristics may be systematically related to the age at which the apprenticeship labour contract starts, if it starts. This because of firms' (people's) action taken to increase their probability of hiring (being hired as) apprentices. However, as long as individuals' control over this is imprecise, potential outcomes conditional on age are continuous. We follow Lee and Lemieux (2010) to provide graphical evidence on this issue.

We first test whether treated and untreated cohorts on the left and on the right of the age cutoff, on average, have statistically identical covariates. We claim that this balancing out of main observable characteristics contributes to provide suggestive evidence consistent with a randomised variation generated by the labour market reform at the age threshold. Table 3 includes the main time invariant characteristics (gender, region of birth, education and an indicator for missing information on education); predetermined variable (past experience and an indicator for missing information on past experience) and the time varying characteristics at the baseline (i.e. ± 24 months around June 2012). The latter sub-set of covariates includes the region of work, an indicator equal to 1 if the individual has switched sector of activity, an indicator of regional mobility and a bulk of dummy variables capturing the position of the job episode in the age specific distribution, at a given month and year, of some characteristics such as the number of multiple job spells; the number of job separations; the number of net job flows (hirings minus separations); the number of job episodes which benefitted from a reduction of labour costs or social insurance benefits. In raw data the difference in discontinuity of the covariates persists, albeit it is quite small. However, the final two columns of the Table show that when the polynomial function approximation is used, all differences in discontinuities between treated and untreated cohorts around the threshold vanish. Any difference in discontinuity is statistically significant implying that all baseline characteristics are balanced in the neighbourhood of ± 1 and ± 2 year(s) of the discontinuity age cutoff. For all covariates the zero order polynomial function in the age range of ± 1 year around the cutoff perfectly retrieves the difference in means at the threshold. Such goodness of fit does not hold in the age range of ± 2 years despite the (local linear) general model specification is applied.

Similarly, evidence on imprecise control over the forcing variable can be suggested by the density of the residuals V of the forcing variables (age a) regression conditional on observable characteristics X (grouped by gender, region of birth, sector of activity and level of education) and on the quartiles of the distribution of the residuals U of equation 1 augmented by our covariates (Lee and Lemieux 2010). In Appendix B we propose two different versions of this density plot.[17] In the first (Figure B1) we consider all individuals in the age range between 25 and 35 years of age who have started a job spell of at least one day in a given year of the sample period. As expected observable X and unobservable U characteristics affect the shape of the densities. However, while it is clear the discrete character of the forcing variable, the histograms indicate that observed and unobserved pre-determined characteristics have identical conditional distributions on either side of the age cutoff in the limit of the threshold. We then restrict the sample to the age range of ± 1 year around the threshold and to those who have an apprenticeship labour contract exploiting the fact that this labour contract can last more than one year. The purpose is to show that even in such selected sample there is not precise manipulation of the age at which this event occurs (i.e. the conditional distributions are never truncated). These model validation tests are re-assuring and make us confident that we are comparing similar individuals who differ only in terms of the random variation in the intention to assignment into treatment. Moreover, all this preliminary analysis confirms the irrelevance of including the baseline covariates in our regression model since they are not necessary to obtain consistent estimates of the treatment effect. However, we will add them to reduce

[17] According to the Bayes' Rule, $Pr[X = x, U = u|A = a] = Pr[a|X = x, U = u]\frac{Pr[X=x,U=u]}{Pr(a)}$, where Pr(a) and Pr[a|X=x, U=u] are marginal and conditional densities. When Pr[a|X=x, U=u] is identical on either side of the cutoff in the limit of the threshold, the distribution of U, V conditional on age will not be truncated in age. The graphs plot the density of V rather than the density of age but the two distributions are equivalent up to a translation shift.

sampling variability in the impact estimates. This is presented and discussed in the next sub-section.

5.2 Static estimates

Column (1) of Table 4 reports the unconditional estimates of the parameter γ_1 in equation 1 which matches the difference in discontinuity in raw data illustrated in Figure 1. As expected, the impact estimates on the apprenticeship and the permanent employment probability are somewhat insensitive to the inclusion of covariates in the regression model. The results are quite robust to various specifications. In fact, all the other columns add further baseline characteristics: time fixed effects (month and year dummies added in column 2); sector fixed effects (column 3); region of birth fixed effects (column 4). The inclusion of region of birth fixed effects is crucial to reduce sampling variability in the permanent employment regression model. This is because the environment where individuals were born and grown up matters and the correlation within group of the main regressors is large. After controlling for this sampling variability the estimated effect on the probability of permanent employment, which remains stable across all model specifications, is statically different from zero at the 1% level. This holds true when, in column (5), we include a polynomial of degree 1 in the employer identification code to control for firm fixed effects; in column (6) we include time invariant characteristics which are the real monthly earnings at the time of recruitment, gender, a dummy for missing information on education and past-experience and when, in column (7), we add time-varying baseline characteristics. The latter covariates are, in a given month and year, a dummy equal to 1 if the worker's educational level is higher than the 25th percentile of the education distribution conditional on age; a dummy equal to 1 if the worker's educational level is higher than the 75th percentile of the education distribution conditional on age; a dummy equal to 1 if the worker's past-experience is higher than the 75th percentile of the past-experience distribution conditional on age; a dummy equal to 1 for switching sector of activity; a dummy equal to 1 if the worker's number of multiple job spells is higher than the 25th percentile of the corresponding distribution conditional on age; a dummy equal to 1 if the job episode is associated to a number of job separations higher than the 25th percentile of the corresponding distribution conditional on age and region of birth; a dummy equal to 1 if the job episode is associated to a number of net flows (hirings minus separations) higher than the 25th percentile of the corresponding distribution conditional on age and region of birth; a dummy if the job episode benefits from a labour costs reduction higher than the 25th percentile of the corresponding distribution conditional on age, and finally a dummy if the job episode is covered by social insurance benefits higher than the 25th percentile of the corresponding distribution conditional on age.

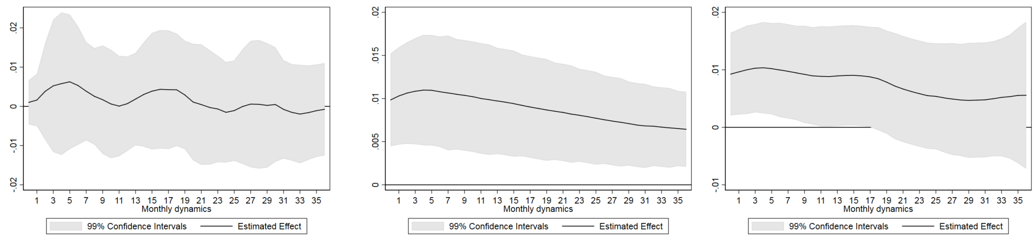
Table 4 corroborates our graphical analysis. At the age cutoff, the permanent employment rates of treated cohorts increases by about one percentage point above the permanent employment rates of similar untreated cohorts. This positive impact of the labour market reform is entirely due to the apprenticeship labour contract. In fact, the magnitude of the estimated coefficients in the permanent employment and the apprenticeship probability almost coincides. Moreover, the lack of any statistically different from zero effect on the permanent employment rates at the firm or sector for which each worker last worked, rules out the possibility that the estimated reduced form effect can be associated to transitions from temporary to permanent employment. The coefficients are precisely estimated with small and, from column 4 onwards, quite stable standard errors. In Appendix B we include a Table (Table B6) showing the estimated effects of the baseline covariates whose inclusion do not affect our main results.

5.3 Dynamic estimates

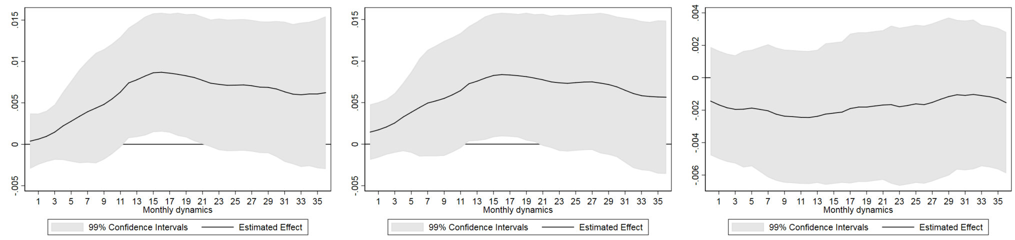
The apprenticeship labour contract can be conveyed as the main port of entry into a permanent labour contract through the analysis presented in the previous section. Nevertheless, it is important to show to what extent its positive impact persists over time.

Figure 3 depicts the dynamic effects estimated by equation 2. After 36 months from the baseline the impact of law no. 92/2012 on the apprenticeship rate is positive albeit, as expected, decreasing over time since the duration of this type of labour contract is fixed. However, the estimated effect is always statistically different from zero suggesting that the labour market reform succeeded in encouraging the use of the apprenticeship labour contract with respect to the previous regime. Without controlling for the persistency in

Figure 3: Difference in discontinuity: dynamic effect up to 36 months without controlling for persistency in outcomes.



(a) Employment rate. (b) Apprenticeship rate. (c) Perm. Employment rate.

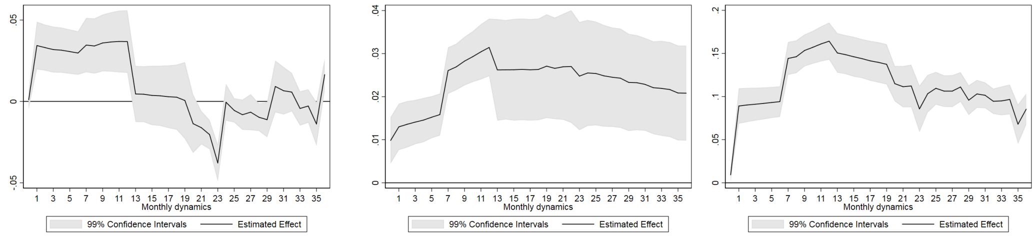


(d) Perm. Emp: same firm. (e) Perm. Emp: same sector. (f) Self Employment rate.

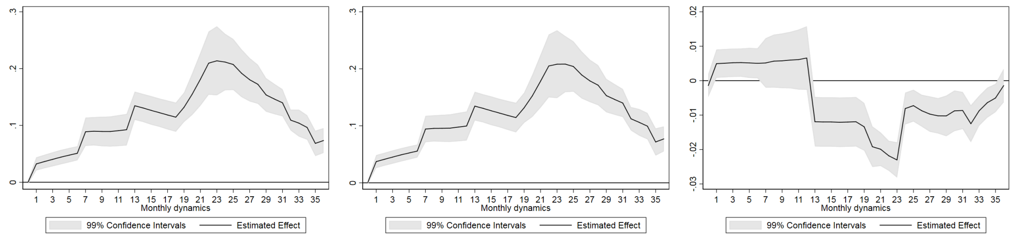
Notes: See notes in Table 4. The gray area indicates 99% confidence intervals.

the outcome generated by the exogenous shock of the reform, Panel (c) shows that after 36 months from the baseline, the permanent employment probability of cohorts who are treated by the labour market reform at the age threshold are still 0.6% higher than the permanent employment rates of similar untreated cohorts. The t-statistics fails to reject the null hypothesis of a coefficient equal to zero at 10% level from the 23th month onwards. On the one hand, this could be related to the inefficiency of equation 2 since standard errors increase over time because the error term has a large fixed component. On the other hand, once the apprenticeship labour contract is not longer available the impact as estimated by equation 2 is expected to vanish over time. Once we depart from the baseline, the impact of the labour market reform at the age threshold on the permanent employment rates at the firm or sector for which each worker last (month) worked is positive and statistically different from zero (panels (d) and (e)). This positive effect reaches a maximum just after 12 months and then it stabilises at about 0.6 percentage point (statistically different from zero at 10% level after 36 months, 33 for the equation for the permanent employment rate at the same sector). The other two panels ((a) and (f)) confirm that the labour market reform has not generated around the age threshold any differential impact on the

Figure 4: Difference in discontinuities: TOT parameters up to 36 months.



(a) Employment rate. (b) Apprenticeship rate. (c) Perm. Employment rate.



(d) Perm. Emp: same firm. (e) Perm. Emp: same sector. (f) Self Employment rate.

Notes: See notes in Table 4. The gray area indicates 99% confidence intervals.

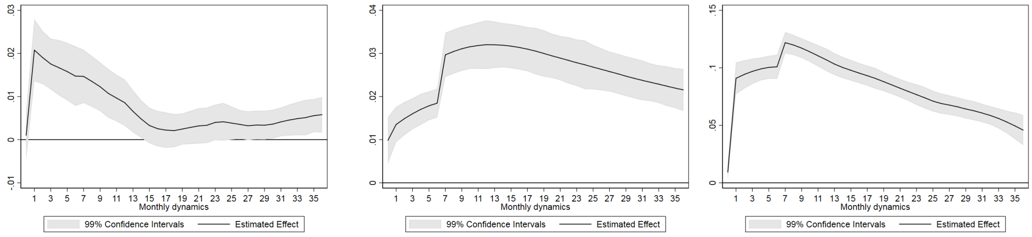
employment and self-employment probability.

5.3.1 Dynamic TOT and ITT parameters

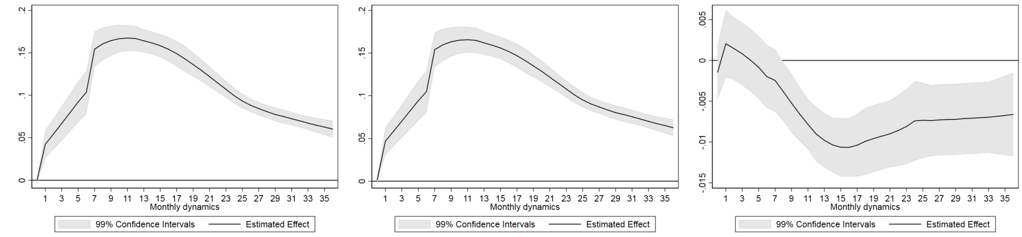
We complete our analysis on the employment effects by estimating equation 4 which allows us to retrieve the dynamic TOT and ITT parameters adding to the baseline static model the persistency in outcomes generated by the labour market reform at the age cutoff. Figures 4 and 5 show that once persistency is controlled for, the dynamic effect is stronger than the impact at the baseline. The coefficients are precisely estimated as suggested by the narrow 99% confidence intervals. As expected, the ITT effect is smoother than the TOT impact which measures how each additional month contributes to the overall ITT effect given the dynamic pattern up to that point.

After 36 months from the baseline, the permanent employment rate (panel c) of individuals affected by the labour market reform at the age threshold increased by 5% when compared to the permanent employment rate of similar untreated cohorts. The impact on the permanent employment rate at the firm (and consequently at the same sector) for which each worker last (month) worked (panels d and e) is in the same order of magnitude. We observe also a positive impact in a range of 2% for the apprenticeship rate

Figure 5: Difference in discontinuities, ITT parameters up to 36 months.



(a) Employment rate. (b) Apprenticeship rate. (c) Perm. Employment rate.



(d) Perm. Emp: same firm. (e) Perm. Emp: same sector. (f) Self Employment rate.

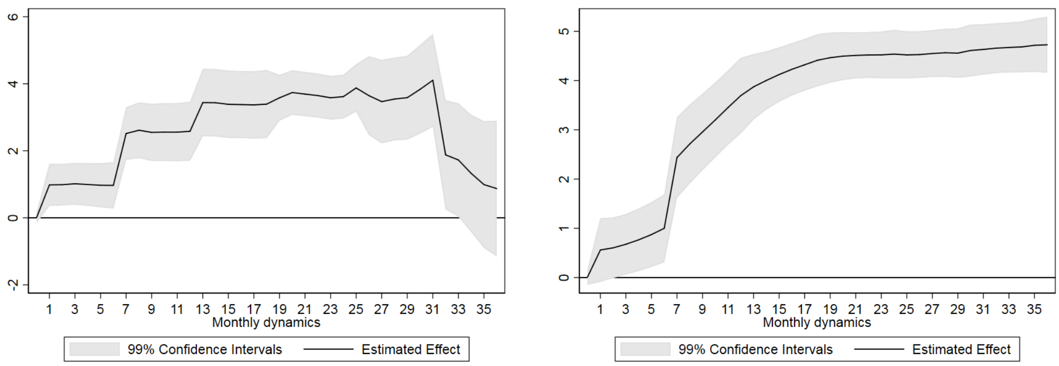
Notes: See notes in Table 4. The gray area indicates 99% confidence intervals.

(panel b). After two years from the baseline both our sources of treatment are vanished: the apprenticeship labour contract cannot longer be signed and all individuals who enter into the labour market are affected by the 2012 reform. Nevertheless, the initial advantage remains and amplifies because of the persistency in outcomes induced by the labour market reform at the age cutoff. In contrast, the impact on the employment rate (panel a) is much smaller and not significantly different from zero in several months of the period considered. The dynamic ITT effect on the self-employment probability (panel f) is negative and statistically different from zero after 6 months from the baseline, albeit it is quite small.

5.4 Impact on tenure

To strengthen our results we consider job tenure as an outcome. In Appendix B, Table B8 reports the difference in discontinuity instantaneous impact on tenure, which is not statistically different from zero. The coefficient is not stable across model specifications and standard errors get smaller as we add further covariates. We claim that this evidence corroborates our interpretation that the estimated instantaneous impact on the permanent employment probability cannot be driven by conversions of temporary employment to permanent status otherwise job tenure would have increased.

Figure 6: Difference in discontinuity, dynamic parameters on tenure.



(a) Treatment on the Treated effect. (b) Intention to Treat effect.

Notes: See notes in Table 4. The gray area indicates 99% confidence intervals.

The following Figure 6 illustrates instead the dynamic TOT and ITT parameters on tenure. The dynamic ITT parameters are precisely estimated and statistically different from zero. After 36 months from the baseline, the tenure of individuals treated by the labour market reform at the age cutoff is about 5 months longer than the tenure of similar untreated individuals. Our findings suggest that a labour contract which provides investment in human capital serves as a stepping stone into permanent employment. In fact, firms are less reluctant to hire on a permanent basis if they partly finance the human capital investment. Moreover, because of the benefits associated to the human capital investment and the cost incurred by both parties, it is more likely that the job match persists over time. In absence of data on earnings, this evidence suggests that the quality of the job positions created by the apprenticeship labour contract might be higher than the quality of job positions created by other labour contracts.

5.5 Robustness

In this section, we consider several robustness checks.

5.5.1 Static placebo estimates

First, we implement a placebo estimate on pairs of months from the pre-reform period of March 2011 to March 2012. These months have been selected to center data ($\pm$ 6 months) around September 2011 when the so-called *Testo Unico per l'Apprendistato*, legislative decree no.167/2011, was issued setting common rules for the apprenticeship labour contract at national level.[18]

We consider a range of ±6 months since on the 5th of April 2012 the government announced that the labour market reform (i.e. law no 92 issued then in June) would have changed the employment protection legislation. Therefore, the aim of this exercise is twofold. First, we produce placebo estimates. Second, we verify whether the effects presented in Table 4 might be due to the legislative decree no.167/2011 rather than to law n.92/2012. Estimates for the various model specifications are presented in Table 5: they are small and insignificant with standard errors of the same magnitude as those in Table 4. We cannot present dynamic estimates for this placebo sub-sample. This is because the comparison between dynamic pattern of treated and untreated individuals is now mainly made within cohort of birth. As long as we have not any information on the month of birth we cannot precisely control for the differences in current and past outcomes between affected and unaffected individuals.

5.5.2 Sample centered ±12 months around June 2012

As a second robustness check we consider the sub-sample centered ±12 months around June 2012 (i.e. from June 2011 to June 2013). All the findings are reported in Appendix C. Table $C1$ shows that covariates are balanced out. Estimates are consistent with our main findings. The static effect is positive and statistically significant although slightly smaller, 0.7%. The dynamic TOT parameter is in the same order of magnitude of the impact estimated using our working sample, albeit it is more noisily estimated in the proximity of the announcement of a hiring incentive on permanent labour contracts. Nevertheless, the dynamic ITT parameters are robust to those reported in Figure 5. This is expected. In the working sample the level effect of this policy could be singled out, by a time fixed effect.[19] However, as any other covariate the exclusion/inclusion of time fixed effects does not impact on consistency of the estimates in (a difference in) regression discontinuity

[18]Earlier than September 2011 the rules governing the apprenticeship labour contract were heterogeneous and fragmented since they were established by the regional governments or by the sectorial national collective agreements. (see Cappellari et al. 2012)

[19]This is to separate out within treated cohorts, individuals affected by both policies from those affected only by law no 92/2012.

design (Lee and Lemieux 2010).

On the top of that, this robustness check suggests that our evidence is not affected by another potential confounding factor. This factor could be even more problematic than the introduction of the hiring incentives that took place a couple of months after the baseline period. Legislative decree no 76/2013, instead, was issued in June 2013 and fixed a limited amount of public expenditures to encourage firms to hire individuals younger than 30 years of age on a permanent basis. The benefit applied also to conversions from temporary to permanent labour contracts. These public resources were administered by the regional governments and there were additional requirements to the age. Individuals could benefit of this hiring incentive because of their age and because of having been unemployed in the previous six months or because their educational level was lower than the upper secondary degree or because they lived with a dependent family member. The combination of different timing responses of the regional governments in a context of heterogeneous regional labour markets and the presence of further requirements, which have shrunk the number of potential recipients, undo the potential impact of this policy intervention in our working sample. In fact, we can rule out the possibility that our results are driven by transitions to permanent employment from either unemployment or temporary labour contracts. Moreover, this robustness check displays static and dynamic permanent employment effects of similar size for the two samples.

5.5.3 Different clustering and heteroskedasticity robust standard errors

Kolesár and Rothe (2018) argue that confidence intervals based on standard errors clustered by the running variable do not guard against model misspecification in regression discontinuity inference with discrete support. They vary the accuracy of the fitted specification for small and moderate number of discrete values of the running variable showing that confidence intervals based on heteroskedasticity-robust (EHW) standard errors are larger than confidence intervals based on standard errors clustered by the running variable. Moreover, EHW confidence intervals perform well and have good coverage. Our preliminary analysis, discussed in section 3.3, shows that in the neighbourhood of the age cutoff, in the age range of ±1 year around the threshold, the fitted specification is very good albeit, clearly, the window width is the smallest possible in our data. To make a comparison across confidence intervals based on different standard errors we replicate our analysis using heteroskedasticity robust standard errors and a clustering based on age and year of birth rather than age, year of birth and region of birth. Clustering on age and year of birth amounts to clustering on the two running variables in the difference in regression discontinuity design (Lee and Card 2008). Results reported in Appendix C (sub-section C1.2) illustrates that, in our data, clustering makes all the difference. Considering, for instance, static model estimation of equation 1, standard errors drop by about 11% when moving from clustering by age, year of birth and region of birth to heteroskedasticity-robust standard errors and by about 4% when moving from clustering by age, year of birth and region of birth to clustering by age and year of birth. This is because possibly in our data the group structure matters. According to the Moulton formula, if the difference in discontinuity regressor values were uncorrelated within groups, the grouped errors structure would not matter for standard errors. This issue is even more relevant for the dynamic estimates when the regressor of interest is mainly fixed within groups, (see also Angrist and Pischke 2009). In all cases, our estimates are the most conservative. Therefore, we rule out the possibility that t-tests are failing to reject the null hypothesis of a statistically different from zero coefficient much more often than what is meant by confidence intervals with the nominal level 99% percent.

5.5.4 Extending the age range to ± 2 years around the threshold

As a final robustness check we extend the age range to ± 2 years around the threshold. As discussed in section 3.3 the main drawback to carry out this analysis relies on functional form restrictions imposed by the discreteness of the age variable. We start by adopting a first order polynomial in age model specification (see Appendix A1). Table C5, reported in Appendix C, shows that the difference in discontinuity parameter on the apprenticeship probability it is statistically different from zero at 5% level and coincides (up to 4 decimal places) with that displayed in Table 4. The Table illustrates further that the difference in discontinuity coefficient on the permanent employment probability is slightly smaller (at the fourth decimal places, 8 instead of 9) when compared to what reported in Table 4. However, it is not longer significant at conventional levels. This is because standard errors are inflated by the strong degree of collinearity between the difference in discontinuity term and its interaction with age. These two terms are identical apart from one cohort of birth in each side of the threshold. All in all, it is reassuring that the difference in discontinuity coefficients on both the apprenticeship and permanent employment probability are stable when we extend the analysis to a larger age range. This evidence indicates that the polynomial of degree zero model specification estimated in the neighbouring of the age cutoff is identifying an effect that can be generalised to higher age ranges.[20] In Appendix C we report also Table C6 and Table C7. The former Table shows that, for both the apprenticeship and permanent employment probability equations, the t-test on the interaction term $a * r * d$ is not statistically different from zero. As illustrated in Appendix A1, this amounts to say that there is not statistical difference in the slope coefficients around the cutoff between treated cohorts and those untreated by the labour market reform ($[(\beta_{1p} - \beta_{0p}) - (\beta_{1b} - \beta_{0b})] = 0$). Put it differently, the data seems to support, for the age range ± 2 years around the threshold, the stability bias hypothesis. Table C7 replicates the analysis for the restricted model where we exclude the interaction term $a * r * d$. The difference in discontinuity parameters are stable and standard errors are smaller. The impact on permanent employment probability is statistically different from zero at 10% level. If we restrict the model excluding also the interaction between $a * r$, the difference in discontinuity parameters are stable and standard errors reduce further still. In such a case the effect on both apprenticeship and permanent employment probability is significant at 1% level (Table C8).

We do not estimate the dynamic treatment effects using the age range of ± 2 years around the threshold (o larger than 2). This is because we are not confident on the model specification given the restrictions imposed by the discreteness of the age variable. Dynamic treatment effects rely on the persistency in the outcome generated by the reform at the age cutoff. That is, they depend upon the labour market history of the individual onwards. If the model specification is unable to control for workers' histories, dynamic treatment effects are biased and inconsistent.

How general are, therefore, our results? Our estimated intention to treat parameters are a weighted average of the intention to treat effects in the population where the weights

[20]The difference in discontinuity parameter is instead much less stable on all the other employment outcomes which, as the graphical analysis has displayed, do not exhibit a discontinuity at the age threshold.

are directly proportional to the ex ante likelihood (that depends upon observable and unobservable characteristics) that an individual's realisation of the apprenticeship labour contract event will be close to the threshold before or after the labour market reform. The weights maybe relatively similar across individuals, in which case the estimated weighted average intention to treat effects are close to the overall effects. While it is not possible to know how different these two parameters are, it remains the case the our difference in discontinuity impacts are averaged over a population that is larger than what a purely cutoff interpretation would suggest (Lee and Lemieux 2010).

6 Conclusions

Globalisation and technological progress creates a continuous process of labour market adaptation. New challenges and concerns are raised about job quality and stability. This paper estimates whether employers could be less reluctant to hire workers on a permanent basis in presence of a human capital investment which they partly finance. We retrieve a static and a dynamic ITT parameter by comparing cohorts treated by law no.92/2012 to similar untreated individuals at the threshold of 30 years of age above which job entries as vocational apprentices cannot occur. Our results suggest that at the threshold, the permanent employment rate of individuals affected by law no.92/2012 increased by about 1% when compared to the permanent employment rate of similar untreated individuals. This result is due the vocational apprenticeship labour contract. In fact, at the baseline, there is no evidence instead of a positive impact on job tenure, and on the permanent employment rate at the firm or sector which each individual last worked. This rules out the possibility that the result is driven by conversions from temporary to permanent employment. Moreover, the employment rate of affected cohorts is not statistically different from the employment rate of similar unaffected cohorts. If indeed vocational apprenticeship has an advantage over the other labour contracts to create jobs of good quality, the positive effects estimated at the baseline might persists several months after. Our ITT dynamic parameter indicates that after 36 months from the baseline the permanent employment rate of individuals affected by the labour market reform at the age threshold increased by about 5% when compared to the permanent employment rate of similar untreated individuals. The impact on the permanent employment rate at the firm (and consequently at the sector) which each individual last worked is in the same order of magnitude. Five months is the tenure advantage of treated over untreated cohorts at the age cutoff after 36 months from the baseline. We interpret our findings as evidence that a labour contract that invests in costly human capital serves as a stepping stone into permanent employment.

References

Abadie, A., Angrist, J. and Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings, *Econometrica* **70**(1): 91–117.

Acemoglu, D. and Pischke, J.-S. (1998). Why do firms train? Theory and Evidence, *The Quarterly Journal of Economics* **113**(1): 79–119.

Acemoglu, D. and Pischke, J.-S. (1999). Beyond Becker: Training in imperfect labour markets, *Economic Journal* **109**(453): F112–42.

Albanese, A., Cappellari, L. and Leonardi, M. (2017). The effects of youth labor market reforms: Evidence from italian apprenticeships, *IZA Discussion Papers 10766*, Institute for the Study of Labor (IZA).

Alzúa, M. l., Cruces, G. and Lopez, C. (2016). Long-run effects of youth training programs: Experimental evidence from argentina, *Economic Inquiry* **54**(4): 1839–1859.

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*, 1 edn, Princeton University Press.

Attanasio, O., Guarin, A., Medina, C. and Meghir, C. (2015). Long term impacts of vouchers for vocational training: Experimental evidence for colombia, *Borradores de economia*, Banco de la Republica de Colombia.

Attanasio, O., Kugler, A. and Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in colombia: Evidence from a randomized trial, *American Economic Journal: Applied Economics* **3**(3): 188–220.

Becker, G. (1962). Investment in human capital: A theoretical analysis, *Journal of Political Economy* **70**.

Berton, F., Devicienti, F. and Pacelli, L. (2011). Are temporary jobs a port of entry into permanent employment?: Evidence from matched employer-employee, *International Journal of Manpower* **32**(8): 879–899.

Bloom, H. S., Orr, L. L., Bell, S. H., Cave, G., Doolittle, F., Lin, W. and Bos, J. M. (1997). The benefits and costs of jtpa title ii-a programs: Key findings from the national job training partnership act study, *Journal of Human Resources* **32**(3): 549–576.

Booth, A., Francesconi, M. and Frank, J. (2002). Temporary jobs: Stepping stones or dead ends?, *Economic Journal* **112**(480): F189–F213.

Cappellari, L., Dell'Aringa, C. and Leonardi, M. (2012). Temporary employment, job flows and productivity: A tale of two reforms, *Economic Journal* **122**(562): F188–F215.

Card, D., Ibarraran, P., Regalia, F., Rosas-Shady, D. and Soares, Y. (2011). The labor market impacts of youth training in the dominican republic, *Journal of Labor Economics* **29**(2): 267 – 300.

Card, D., Kluve, J. and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis, *Economic Journal* **120**(548): F452–F477.

Card, D., Kluve, J. and Weber, A. (2018). What works? a meta analysis of recent active labor market program evaluations, *Journal of the European Economic Association* **16**(3): 894–931.

Cellini, S. R., Ferreira, F. and Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design, *The Quarterly Journal of Economics* **125**(1): 215–261.

Crépon, B., Ferracci, M. and Fougere, D. (2012). Training the unemployed in france: How does it affect unemployment duration and recurrence?, *Annals of Economics and Statistics* **107**(108): 175–199.

Crépon, B., Ferracci, M., Jolivet, G. and van den Berg, G. (2009). Active labor market policy effects in a dynamic setting, *Journal of the European Economic Association* **7**(2-3): 595–605.

Dustmann, C. and Schönberg, U. (2012). What makes firm-based vocational training schemes successful? the role of commitment, *American Economic Journal: Applied Economics* **4**(2): 36–61.

Fitzenberger, B., Furdas, M. and Sajons, C. (2016). End-of-year spending and the long-run employment effects of training programs for the unemployed, *IZA Discussion Papers 10441*, Institute for the Study of Labor (IZA).

Flores-Lagunes, A., Gonzalez, A. and Neumann, T. (2010). Learning but not earning? the impact of job corps training on hispanic youth, *Economic Inquiry* **48**(3): 651–667.

Grembi, V., Nannicini, T. and Troiano, U. (2016). Do fiscal rules matter?, *American Economic Journal: Applied Economics* **8**(3): 1–30.

Heckman, J., Ichimura, H. and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme, *Review of Economic Studies* **64**(4): 605–654.

Heckman, J., LaLonde, R. and Smith, J. (1999). The economics and econometrics of active labor market programs, *in* O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, 1 edn, Vol. 3, Part A, Elsevier, chapter 31, pp. 1865–2097.

Heckman, J. and Smith, J. (2000). The sensitivity of experimental impact estimates (evidence from the national jtpa study), *Youth Employment and Joblessness in Advanced Countries*, National Bureau of Economic Research, Inc, pp. 331–356.

Heinrich, C., Mueser, P. and Troske, K. (2005). Welfare to temporary work: Implications for labor market outcomes, *The Review of Economics and Statistics* **87**(1): 154–173.

Holmlund, B. and Storrie, D. (2002). Temporary work in turbulent times: The swedish experience, *Economic Journal* **112**(480): F245–F269.

Ichino, A., Mealli, F. and Nannicini, T. (2008). From temporary help jobs to permanent employment: what can we learn from matching estimators and their sensitivity?, *Journal of Applied Econometrics* **23**(3): 305–327.

Kluve, J. (2010). The effectiveness of european active labor market programs, *Labour Economics* **17**(6): 904–918.

Kolesár, M. and Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable, *American Economic Review* **108**(8): 2277–2304.

Lechner, M., Miquel, R. and Wunsch, C. (2011). Long run effects of public sector sponsored training in West Germany, *Journal of the European Economic Association* **9**(4): 742–784.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects, *Review of Economic Studies* **76**(3): 1071–1102.

Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error, *Journal of Econometrics* **142**(2): 655–674.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics, *Journal of Economic Literature* **48**(2): 281–355.

Malamud, O. and Pop-Eleches, C. (2010). General education versus vocational training: Evidence from an economy in transition, *The Review of Economics and Statistics* **92**(1): 43–60.

Office, G. A. (1996). Job training partnership act: Longterm earnings and employment outcomes., *Report Washington, DC:*, General Accounting Office.

Osikominu, A. (2013). Quick job entry or long-term human capital development? the dynamic effects of alternative training schemes, *Review of Economic Studies* **80**(1): 313–342.

Rodríguez, J., Saltiel, F. and Urzúa, S. (2018). Dynamic treatment effects of job training. Mimeo.

Schochet, P. Z., Burghardt, J. and McConnell, S. (2008). Does job corps work? impact findings from the national job corps study, *American Economic Review* **98**(5): 1864–86.

A1 Difference in discontinuity design in a potential outcomes framework

Consider the fraction, n, in the population that starts a job spell (even of one) in a given time period, b, before the introduction of law no.92/2012. Consider then an outcome, y, and let's express quantities in terms of expected potential outcomes, μ, given the individual's year of birth, tb, on the basis of which she is aged, a_i, in a given time period.

$$E[y_{1ib}|a_{ib}, tb_{ib}] = \mu_{1b}$$
$$E[y_{0ib}|a_{ib}, tb_{ib}] = \mu_{0b}$$

where 1 and 0 refer to the left and right side of the age threshold level, 30. In fact:

$$\mu_{1b} = \alpha_0 + \gamma_0 + \beta_{1b} a_{it}$$
$$\mu_{0b} = \alpha_0 + \beta_{0b} a_{it}$$

Other covariates are not included for the sake of simplicity. The two sides of the cutoff identify the treatment and control states since the intention to assignment into treatment (i.e. the vocational apprenticeship labour contract) is based on the following selection rule:

$$d_{it} = \begin{cases} 1 & \text{if } a_{it} < a_m \\ 0 & \text{if } a_{it} \geq a_m \end{cases}$$

where a_m is the age cutoff.

We start from the usual definition of the observed outcome for this n fraction in the population whose age is close to the threshold. We take expectations under the assumption that the intention to assignment into treatment is locally randomised (i.e. independence assumption: $y_{1ib} \perp d_{it}$ and $y_{0ib} \perp d_{it}$, i.e. the law could have established another age threshold.)

$$y_{ib} = y_{1ib} d_{it} + y_{0ib}(1 - d_{it})$$
$$E[y_{ib}|d_{it}, a_{it}, tb_{it}] = E[y_{ib}|a_{it}, tb_{it}] = E[y_{0ib}|a_{it}, tb_{it}] + \{E[y_{1ib}|a_{it}, tb_{it}] - E[y_{0ib}|a_{it}, tb_{it}]\} d_{it}$$
$$E[y_{ib}|a_{it}, tb_{it}] = \mu_{0b} + [\mu_{1b} - \mu_{0b}] d_{it}$$

In absence of a confounding policy, the difference $(\mu_{1b} - \mu_{0b})$ can be estimated by γ_0 in the following (first order polynomial in age) regression discontinuity model in a given period t before the labour market reform:

$$y_{it} = \alpha_0 + \beta_{0b} a_{it} + \gamma_0 d_{it} + (\beta_{1b} - \beta_{0b}) d_{it} a_{it} + \epsilon_{it}$$

It is possible to extend this model specification to higher order polynomials in age, a_{it}, by introducing squared, cubic etc terms in age and in the interaction term $d_{it} a_{it}$.

Consider now that law n. 92/2012 introduced another source of randomised variation based on this discontinuity rule:

$$r_{it} = \begin{cases} 1 & \text{if } tb_{it} \leq tb_c \\ 0 & \text{if } tb_{it} > tb_c \end{cases}$$

where tb_c corresponds to the year of birth that is aged a_{it} in June 2012. Observed outcome, centered around the labour market reform at the age cutoff in a time period t, amounts to:

$$y_{it} = y_{ip} r_{it} + y_{ib}(1 - r_{it})$$

where p denotes post labour market reform.

Let's define potential outcomes around the age threshold in the post reform period:

$$E[y_{1ip}|a_{ip}, tb_{ip}] = \mu_{1p}$$
$$E[y_{0ip}|a_{ip}, tb_{ip}] = \mu_{0p}$$

where, as above, 1 and 0 refer to the left and right side of the age cutoff. In fact:

$$\mu_{1p} = \alpha_0 + \gamma_0 + \gamma_1 + \alpha_1 + \beta_{1p} a_{it}$$
$$\mu_{0p} = \alpha_0 + \alpha_1 + \beta_{0p} a_{it}$$

Then we take expectations of observed outcome y_{it} under the assumption that the intention to assignment into treatment of the labour market reform at the age cutoff is locally randomised (i.e. independence assumption: $y_{1ip} \perp r_{it}$ and $y_{0ip} \perp r_{it}$ and $y_{1ib} \perp r_{it}$ and $y_{0ib} \perp r_{it}$):[21]

$$E[y_{it}|d_{it}, r_{it}, a_{it}, tb_{it}] = E[y_{it}|a_{it}, tb_{it}] = E[y_{ib}|a_{it}, tb_{it}] + \{E[y_{ip}|a_{it}, tb_{it}] - E[y_{ib}|a_{it}, tb_{it}]\} r_{it}$$
$$E[y_{it}|a_{it}, tb_{it}] = \mu_{0b} + (\mu_{1b} - \mu_{0b}) d_{it} + (\mu_{0p} - \mu_{0b}) r_{it} + [(\mu_{1p} - \mu_{0p}) - (\mu_{1b} - \mu_{0b})] r_{it} d_{it}$$

$$E[y_{it}|a_{it}, tb_{it}] = \alpha_0 + \beta_{0b} a_{it} + \gamma_0 d_{it} + (\beta_{1b} - \beta_{0b}) a_{it} d_{it} + \alpha_1 r_{it} + (\beta_{0p} - \beta_{0b}) a_{it} r_{it} + \quad \text{(A7)}$$
$$+ \gamma_1 r_{it} d_{it} + [(\beta_{1p} - \beta_{0p}) - (\beta_{1b} - \beta_{0b})] a_{it} r_{it} d_{it} \quad \text{(A8)}$$

Equation A8 corresponds to the first order polynomial in age of the difference in regression discontinuity design. The model could be extended to second, third etc order polynomials by augmenting the model specification by the squared, cubic, etc..., age terms and the interaction in age terms. As discussed in the main text, this regression model cannot be estimated in the closest (given our data) proximity of the age cutoff. In fact, in the age range ± 1 year around the threshold the indicator d_{it} is perfectly multicollinear with the age variable a_{it}. This implies a zero slope assumption ($\beta_{1p} = \beta_{0p} = \beta_{1b} = \beta_{0b}$=0) that in the proximity of the threshold is not a strong assumption as it is once we move

[21] The independent assumption at the age threshold level holds also in the post-reform period, $y_{1ip} \perp d_{it}$ and $y_{0ip} \perp d_{it}$

far away of the age cutoff. As a result, the regression model collapses to a zero order polynomial in age. In fact:

$$E[y_{it}|a_{it}, tb_{it}] = \alpha_0 + \gamma_0 d_{it} + \alpha_1 r_{it} + \gamma_1 r_{it} d_{it} \tag{A9a}$$

$$y_{it} = \alpha_0 + \alpha_1 r_{it} + \gamma_0 d_{it} + \gamma_1 d_{it} r_{it} + \epsilon_{it} \tag{A9b}$$

Equation A9b correspond to equation 1 reported in the main text.

Therefore, in the age range of ± 1 year around the threshold, because of the zero slope assumption, the stability bias assumption trivially holds and the difference in discontinuity parameter coincides with the difference in differences parameter. In fact:

$$\begin{aligned}\gamma_1 = (E[y|d=1, r=1] - E[y|d=0, r=1]) - (E[y|d=1, r=0] - E[y|d=0, r=0]) = \\ = [(\mu_{1p} - \mu_{0p}) - (\mu_{1b} - \mu_{0b})]\end{aligned}$$

Rearranging the terms we have the usual expression for the difference in differences parameter:

$$\gamma_1 = (E[y|d=1, r=1] - E[y|d=1, r=0]) - (E[y|d=0, r=1] - E[y|d=0, r=0])$$

In general, in all the other age ranges, the two parameters differ since they are equal only if some restrictions on a flexible model specification (which could also allow for higher order polynomials in age) are imposed.

A2 Average treatment effect of the population around the age threshold

Consider the fraction $1 - n$ of those who are unable to have a job spell, even of one day, in a given year or are permanently working without changing status or job. For them, before (0) and post (1) the introduction of law no.92/2012, on the left 1 and on the right 0 of the threshold, the potential outcomes are the following:

$$E[y_{10}|a_{it}, tb_{it}] = \theta_{10}$$
$$E[y_{00}|a_{it}, tb_{it}] = \theta_{00}$$

$$E[y_{11}|a_{it}, tb_{it}] = \theta_{11}$$
$$E[y_{01}|a_{it}, tb_{it}] = \theta_{01}$$

For this $1-n$ fraction of individuals the ITT parameter at the age threshold is $[\theta_{1t} - \theta_{0t}]$ where $t = 0$ in the before reform period and $t = 1$ in the post-reform period. The difference in discontinuity ITT parameter amounts to $[(\theta_{11} - \theta_{01}) - (\theta_{10} - \theta_{00})]$ under the crucial assumption that the fraction $1 - n$ is constant across pre and post-reform periods.

It is possible to define three ATE parameters:

$$ATE_b = n[\mu_{10} - \mu_{00}] + (1 - n)[\theta_{10} - \theta_{00}] \text{(A13a)}$$
$$ATE_p = n[\mu_{11} - \mu_{01}] + (1 - n)[\theta_{11} - \theta_{01}] \text{(A13b)}$$
$$ATE_b - ATE_p = n[(\mu_{11} - \mu_{01}) - (\mu_{10} - \mu_{00})] + (1 - n)[(\theta_{11} - \theta_{01}) - (\theta_{11} - \theta_{01})] \text{(A13c)}$$

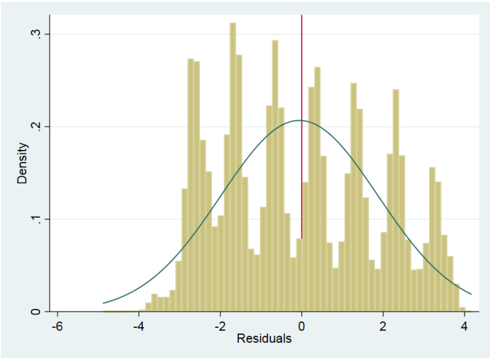
Equation A13a identifies the ATE at the threshold before the reform; equation A13b identifies the ATE at the threshold after the reform and EquationATE3 identifies the ATE in difference in discontinuity.

We assume that $\theta_{11} = \theta_{01} = \theta_{11} = \theta_{01} = 0$ which is a plausible assumption although for different reasons across the two types of individuals who belong to the $1 - n$ fraction. On the one hand, it is unlikely that those who already work permanently switch to the apprenticeship labour contract. Moreover, this event would have been observed in the data and the individual would have belonged to the fraction n in the population. On the other hand, a positive potential outcome associated to a labour contract which implies costly (for both parties involved in the job match) human capital accumulation is an unlikely event for those who are unable or who are unwilling to have a job spell even of one day in a given year. For these individuals we cannot exclude that they are out of the labour force.

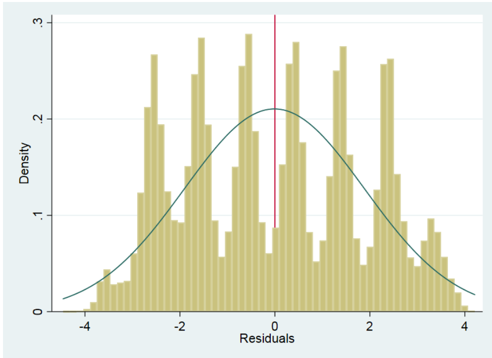
Around the age threshold, the difference in discontinuity ATE static parameter therefore amounts to $n[(\mu_{11} - \mu_{01}) - (\mu_{10} - \mu_{00})]$, that is the fraction n times the ITT static parameter. It is crucial that the source of the randomised variation (i.e. law no.92/2012) has not any impact on the selection into employment (i.e. the employment probability) in a given year at a given age but only on the probability of the apprenticeship labour contract around the threshold. This hypothesis is supported by the data.

A3 Additional Tables and Figures

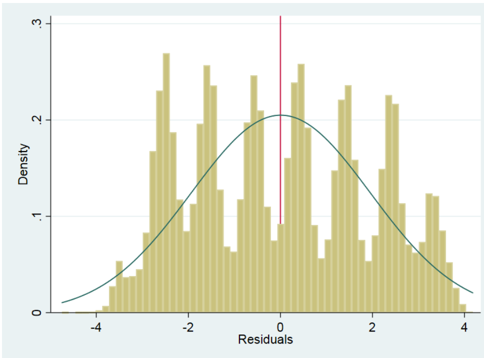
Figure B1: Density of the residuals of age conditional on observable characteristics and quartile of the distributions of residuals of static main regression of permanent employment probability.



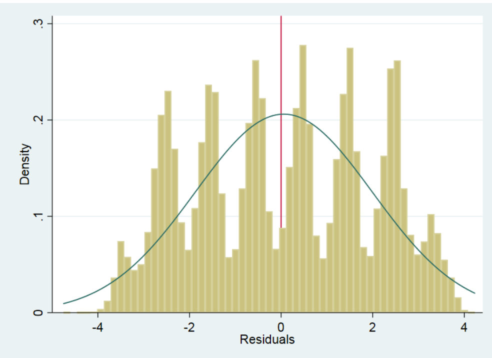
(a) First quartile



(b) Second quartile



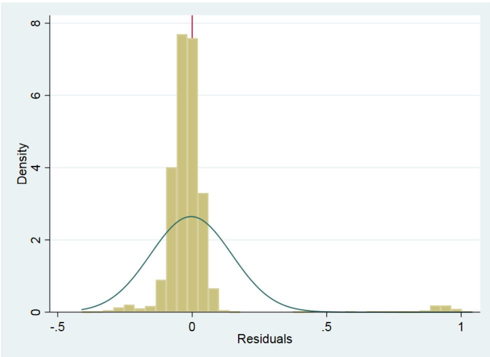
(c) Third quartile



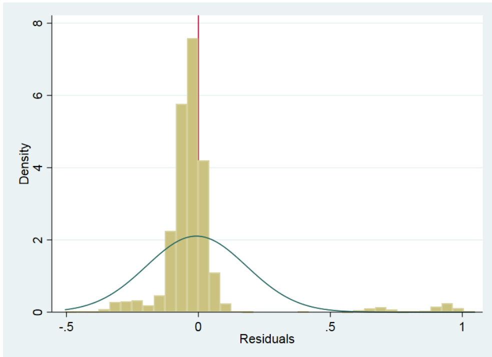
(d) Fourth quartile

Notes: Density of the residuals V of the forcing variables (age a) regression conditional on observable characteristics X (grouped by gender, region of birth, sector of activity and level of education) and on the quartiles of the distribution of the residuals U of equation 1 augmented by our covariates.

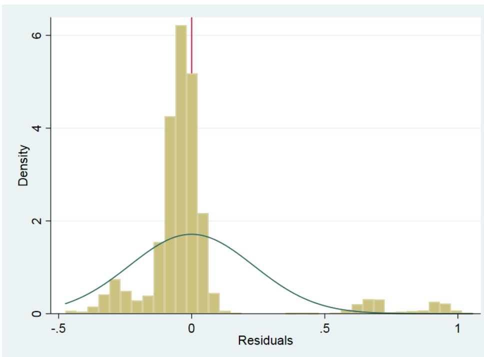
Figure B2: Density of the residuals of age conditional on observable characteristics and quartile of the distributions of residuals of static main regression of permanent employment probability. Only selected sample of those who are apprentices in the age range of ±1 year around the threshold.



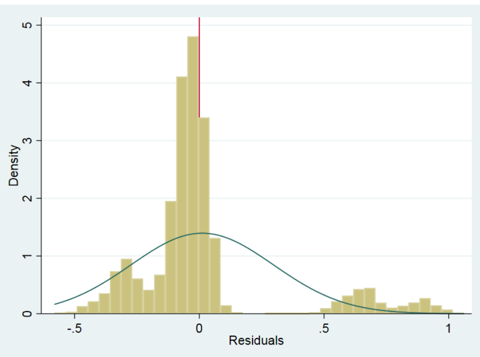
(a) First quartile



(b) Second quartile



(c) Third quartile



(d) Fourth quartile

Notes: Density of the residuals V of the forcing variables (age a) regression conditional on observable characteristics X (grouped by gender, region of birth, sector of activity and level of education) and on the quartiles of the distribution of the residuals U of equation 1 augmented by our covariates.

Table 3: Main observable characteristics: difference in discontinuity

	Main Sample			
	Raw data		Polynomial fit	
	[-1,1]	[-2,2]	[-1,1]	[-2,2]
	DiD (Std. Dev.)	DiD (Std. Dev.)	DiD (Std. Err.)	DiD (Std. Err.)
Gender	-0.004^{***}	-0.005^{***}	-0.004	0.001
	0.001	0.001	0.047	0.080
Region of birth	-0.486^{***}	-0.415^{***}	-0.486	-1.574
	0.063	0.044	22.547	38.557
Education	-0.201^{***}	-0.215^{***}	-0.201	-0.149
	0.035	0.025	6.162	10.646
Missing education	-0.002^{***}	0.001^{**}	-0.002	-0.009
	0.001	0.001	0.115	0.197
Past experience (in days)	-72.251^{***}	-161.334^{***}	-72.251	85.876
	1.075	0.747	66.093	103.424
Missing past exp.	0.002^{***}	0.015^{***}	0.002	-0.022
	0.001	0.001	0.020	0.034
Region of work	-0.027^{***}	-0.010^{*}	-0.027	-0.034
	0.008	0.006	0.797	1.388
Switching sector	-0.002^{***}	-0.003^{***}	-0.002	0.003
	0.001	0.000	0.020	0.036
Regional mobility	-0.008^{***}	-0.008^{***}	-0.008	-0.017
	0.001	0.001	0.194	0.333
Higher 25 per. monthly job spells	-0.004^{***}	-0.002^{***}	-0.004	-0.006
	0.001	0.001	0.036	0.061
Higher 25 per. monthly sep. flows	-0.002^{***}	-0.002^{***}	-0.002	-0.001
	0.000	0.000	0.009	0.015
Higher 25 per. monthly net job flows	-0.001^{*}	-0.003^{***}	-0.001	0.002
	0.001	0.000	0.009	0.014
Higher than 25 perc. costs reduction	0.001^{***}	0.000	0.001	0.002
	0.000	0.000	0.010	0.017
Higher than 25 perc. soc. insurance benefits	0.001^{***}	-0.000	0.001	0.001
	0.000	0.000	0.001	0.001

Notes: The polynomial fit corresponds to a zero (first) order polynomial in age when the age range is $\pm 1(2)$. Each variable, defined as higher than the 25th percentile, is a dummy variable which is equal to 1 if the job episode sits in a percentile higher than the 25th of the age specific distribution of each covariate in a given month and year.

Table 4: Static model estimates

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment prob.	.00189	.0015	.00212	.00176	.00164	.00192	.00101
	.038	.0127	.0106	.0052	.0052	.0054	.0022
Apprenticeship prob.	.01003***	.01004***	.01004***	.00998***	.00998***	.00993***	.00984***
	.0031	.0031	.0028	.0021	.0021	.0021	.0021
Perm. Employment prob.	.00861	.00855	.00905	.00869***	.00866***	.00835***	.00925***
	.0138	.0127	.0087	.0032	.0032	.0032	.0028
Perm. Empl. prob same firm	.00099	.00095	.00106	.00109	.00109	.00088	.00038
	.0025	.0017	.002	.0014	.0014	.0014	.0013
Perm. Empl. prob. same sector	.00183	.0018	.00193	.00195	.00195	.00172	.00145
	.003	.002	.0021	.0015	.0015	.0014	.0013
Self employment	-.00121	-.00116	-.00137	-.00151	-.00149	-.00143	-.00144
	.003	.0028	.0026	.0013	.0013	.0013	.0013
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

Notes: Time invariant characteristics correspond to the real monthly earnings at the time of recruitment, gender, a dummy for missing information on education and past experience. Time-varying baseline characteristics include , in a given month and year, a dummy equal to 1 if the worker's educational level is higher than the 25th percentile of the education distribution conditional on age; a dummy equal to 1 if the worker's educational level is higher than the 75th percentile of the education distribution conditional on age; a dummy equal to 1 if the worker's past-experience is higher than the 75th percentile of the past-experience distribution conditional on age; a dummy equal to 1 for changing sector; a dummy equal to 1 if the worker's number of multiple job spells is higher than the 25th percentile of the corresponding distribution conditional on age; a dummy equal to 1 if the job episode is associated to a number of job separations higher than the 25th percentile of the corresponding distribution conditional on age and region of birth; a dummy equal to 1 if the job episode is associated to a number of net flows (hirings minus separations) higher than the 25th percentile of the corresponding distribution conditional on age and region of birth; a dummy if the job episode benefitted of a labour costs reduction higher than the 25th percentile of the corresponding distribution conditional on age, and finally a dummy if the job episode benefitted of social insurance benefits higher than the 25th percentile of the corresponding distribution conditional on age. Standard errors clustered at year of birth, age and region of birth level.

Table 5: Placebo: static model

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment prob.	-.00233	.00085	.00035	.00059	.00055	.00049	.00025
	.0657	.0157	.0168	.0177	.0177	.0174	.0058
Apprenticeship prob.	-.00155	.00154	.00154	.00153	.00153	.00155	.00153
	.0021	.0019	.0017	.0014	.0014	.0014	.0014
Perm. Employment prob.	-.00237	.00202	.00166	.00184	.00182	.00177	.00163
	.0201	.0061	.0031	.004	.004	.004	.0029
Perm. Empl. prob. same firm	-.00013	.00006	.00001	.00002	.00002	.00002	.00013
	.0042	.0013	.0011	.0011	.0011	.0011	.0014
Perm. Empl. prob. same sector	-.00125	.00115	.00108	.00107	.00107	.00107	.00113
	.0052	.0015	.0013	.0011	.0011	.001	.0014
Self employment	.00253	.00251	.00231	.0022	.00219	.00222	.00219
	.0025	.0029	.0026	.0016	.0016	.0016	.0014
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

Notes: The placebo sample comprises -6/+6 months around September 2011. See notes in Table 4.

Table B1: Goodness of fit F statistic for the polynomial functional form: apprenticeship probability

	Without DiD specification			**DiD Model specification**		
	[-1,1]	[-2,2]	[-3,3]	[-1,1]	[-2,2]	[-3,3]
First Order Polynomial						
F			177.777			103.287
CV			4.605			3.319
p val	1.000	1	1.068×10^{-74}	1	1	9.588×10^{-82}
Second Order Polynomial						
F			467.835		2.278	170.616
CV			4.605		4.605	2.802
p val		1	5.075×10^{-200}		0.718	1.368×10^{-212}
Third Order Polynomial						
F			772.452		4.556	163.480
CV			6.635		6.635	3.017
p val		1	1.033×10^{-164}		0.718	3.399×10^{-168}
Fourth Order Polynomial						
F		0				3.803
CV						3.319
p val		1	1		0.718	0.172

Notes: The null hypothesis of the F goodness-of-fit statistic test is that the functional form adopted is statistically equal to an unrestricted regression of the outcome on the full set of dummy variables for the possible values of age (age times the indicator function for being treated by the labour market reform for the difference in discontinuity model specification) which define the age range reported in brackets. If the statistic exceeds the critical values CV, the null is rejected. The p-value refers to the p-value of an F test on the joint significance of the age dummies (age dummies times the labour market reform treatment indicator) in the auxiliary regression.

Table B2: Goodness of fit F statistic for the polynomial functional form: employment probability

	Without DiD specification			**DiD Model specification**		
	[-1,1]	[-2,2]	[-3,3]	[-1,1]	[-2,2]	[-3,3]
First Order Polynomial						
F			4.347			2.615
CV			4.605			3.319
p val	1.000	1.000	0.120	1.000	1	0.494
Second Order Polynomial						
F			2.824		1.397	2.710
CV			4.605		4.605	2.802
p val		1.000	0.337		0.901	0.135
Third Order Polynomial						
F			2.913		2.795	2.652
CV			6.635		6.635	3.017
p val		1.000	0.713		0.901	0.277
Fourth Order Polynomial						
F						2.485
CV						3.319
p val		1.000	1		0.901	0.542

Notes: The null hypothesis of the F goodness-of-fit statistic test is that the functional form adopted is statistically equal to an unrestricted regression of the outcome on the full set of dummy variables for the possible values of age (age times the indicator function for being treated by the labour market reform for the difference in discontinuity model specification) which define the age range reported in brackets. If the statistic exceeds the critical values CV, the null is rejected. The p-value refers to the p-value of an F test on the joint significance of the age dummies (age dummies times the labour market reform treatment indicator) in the auxiliary regression.

Table B3: Goodness of fit F statistic for the polynomial functional form: permanent employment probability, same firm

	Without DiD specification			**DiD Model specification**		
	[-1,1]	[-2,2]	[-3,3]	[-1,1]	[-2,2]	[-3,3]
First Order Polynomial						
F			1.261			4.301
CV			4.605			3.319
p val	1.000	1.000	0.779	1.000	1	0.103
Second Order Polynomial						
F			4.325		1.055	4.911
CV			4.605		4.605	2.802
p val		1.000	0.126		0.955	0.002
Third Order Polynomial						
F			6.574		2.110	5.910
CV			6.635		6.635	3.017
p val		1.000	0.248		0.955	0.002
Fourth Order Polynomial						
F						6.113
CV						3.319
p val		1.000	1		0.955	0.011

Notes: The null hypothesis of the F goodness-of-fit statistic test is that the functional form adopted is statistically equal to an unrestricted regression of the outcome on the full set of dummy variables for the possible values of age (age times the indicator function for being treated by the labour market reform for the difference in discontinuity model specification) which define the age range reported in brackets. If the statistic exceeds the critical values CV, the null is rejected. The p-value refers to the p-value of an F test on the joint significance of the age dummies (age dummies times the labour market reform treatment indicator) in the auxiliary regression.

Table B4: Goodness of fit F statistic for the polynomial functional form: permanent employment probability, same sector

	Without DiD specification			**DiD Model specification**		
	[-1,1]	[-2,2]	[-3,3]	[-1,1]	[-2,2]	[-3,3]
First Order Polynomial						
F			1.653			2.605
CV			4.605			3.319
p val	1.000	1	0.651	1.000	1	0.488
Second Order Polynomial						
F			6.831		0.595	4.133
CV			4.605		4.605	2.802
p val		1	0.019		0.992	0.010
Third Order Polynomial						
F		0	10.607		1.190	4.773
CV			6.635		6.635	3.017
p val		1	0.060		0.992	0.013
Fourth Order Polynomial						
F		0				3.645
CV						3.319
p val		1	1		0.992	0.203

Notes: The null hypothesis of the F goodness-of-fit statistic test is that the functional form adopted is statistically equal to an unrestricted regression of the outcome on the full set of dummy variables for the possible values of age (age times the indicator function for being treated by the labour market reform for the difference in discontinuity model specification) which define the age range reported in brackets. If the statistic exceeds the critical values CV, the null is rejected. The p-value refers to the p-value of an F test on the joint significance of the age dummies (age dummies times the labour market reform treatment indicator) in the auxiliary regression.

Table B5: Goodness of fit F statistic for the polynomial functional form: self-employment probability

	Without DiD specification			**DiD Model specification**		
	[-1,1]	[-2,2]	[-3,3]	[-1,1]	[-2,2]	[-3,3]
First Order Polynomial						
F			0.664			1.857
CV			4.605			3.319
p val	1.000	1.000	0.937	1	1	0.767
Second Order Polynomial						
F			1.927		9.886	7.873
CV			4.605		4.605	2.802
p val		1.000	0.561		0.006	1.986×10^{-6}
Third Order Polynomial						
F			3.489		19.772	9.156
CV			6.635		6.635	3.017
p val		1.000	0.641		0.006	3.734×10^{-6}
Fourth Order Polynomial						
F						10.852
CV						3.319
p val		1.000	1		0.006	9.110×10^{-6}

Notes: The null hypothesis of the F goodness-of-fit statistic test is that the functional form adopted is statistically equal to an unrestricted regression of the outcome on the full set of dummy variables for the possible values of age (age times the indicator function for being treated by the labour market reform for the difference in discontinuity model specification) which define the age range reported in brackets. If the statistic exceeds the critical values CV, the null is rejected. The p-value refers to the p-value of an F test on the joint significance of the age dummies (age dummies times the labour market reform treatment indicator) in the auxiliary regression.

Table B6: Main estimates, covariates: static model

	Emp.	**Appr.**	**Perm. Emp.**	**Same Firm Perm.**	**Same Sect. Perm.**	**Self Emp.**
Model specification in Table 4	(7)	(7)	(7)	(7)	(7)	(7)
Gender	-0.008***	0.002***	-0.009***	-0.005***	-0.005***	-0.016***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Missing education	0.005***	-0.001*	-0.001	-0.001**	-0.002**	-0.006***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Real first earnings	-0.000***	-0.000***	0.000***	0.000***	0.000***	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing first earnings	0.026***	-0.017***	-0.005	0.009***	0.007***	0.023***
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)
Missing past exp.	-0.042***	0.004***	0.031***	0.069***	0.069***	0.017***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
Higher education than 25	0.009***	0.007***	0.026***	0.003***	0.005***	0.001
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Higher education than 75	0.026***	0.004***	0.019***	0.001	0.002*	-0.020***
	(0.002)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
Higher past experience than 75	0.072***	-0.002***	0.046***	-0.010***	-0.003***	-0.009***
	(0.002)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)
Changing sector	-0.006***	0.003***	0.014***	-0.006***	-0.017***	-0.004***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Higher 25 per. monthly job spells	-0.794***	-0.022***	-0.095***	-0.034***	-0.028***	0.067***
	(0.003)	(0.002)	(0.004)	(0.001)	(0.001)	(0.002)
Higher 25 per. monthly sep. flows	-0.252***	-0.019***	-0.137***	-0.033***	-0.037***	0.012***
	(0.005)	(0.002)	(0.004)	(0.001)	(0.001)	(0.001)
Higher 25 per. monthly net job flows	0.147***	-0.004***	-0.004**	-0.004***	-0.000	-0.007***
	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Higher than 25 perc. costs reduction	0.106***	-0.026***	0.105***	0.049***	0.055***	-0.001
	(0.003)	(0.003)	(0.007)	(0.004)	(0.005)	(0.001)
Higher than 25 perc. soc. insurance benefits	0.058***	-0.023***	0.066***	-0.009*	-0.004	-0.010***
	(0.008)	(0.004)	(0.019)	(0.005)	(0.007)	(0.001)

Notes: See notes in Table 4.

Table B7: Placebo estimates, covariates: static model

	Emp.	**Appr.**	**Perm. Emp.**	**Same Firm Perm.**	**Same Sect. Perm.**	**Self Emp.**
Model specification	(7)	(7)	(7)	(7)	(7)	(7)
Gender	-0.007***	0.001	-0.010***	-0.005***	-0.006***	-0.015***
	(0.002)	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)
Missing education	0.003*	-0.001	-0.003	-0.002*	-0.003	-0.006***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Real first earnings	-0.000***	-0.000***	0.000***	0.000**	0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing first earnings	0.023***	-0.012***	-0.002	0.007***	0.005**	0.024***
	(0.003)	(0.003)	(0.007)	(0.002)	(0.003)	(0.003)
Missing past exp.	-0.044***	0.005***	0.031***	0.067***	0.066***	0.016***
	(0.003)	(0.001)	(0.003)	(0.004)	(0.004)	(0.004)
haeduc	0.009***	0.004***	0.026***	0.001	0.003*	0.003*
	(0.002)	(0.001)	(0.004)	(0.002)	(0.002)	(0.002)
mhaeduc	0.032***	0.003*	0.029***	0.006**	0.008***	-0.019***
	(0.003)	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)
Higher past experience than 75	0.077***	-0.002*	0.055***	-0.009***	-0.003**	-0.008***
	(0.004)	(0.001)	(0.006)	(0.001)	(0.001)	(0.001)
Changing sector	-0.006***	0.004***	0.015***	-0.006***	-0.016***	-0.002
	(0.002)	(0.001)	(0.003)	(0.001)	(0.002)	(0.002)
Higher 25 per. monthly job spells	-0.790***	-0.018***	-0.083***	-0.033***	-0.027***	0.063***
	(0.007)	(0.003)	(0.007)	(0.002)	(0.002)	(0.003)
Higher 25 per. monthly sep. flows	-0.251***	-0.016***	-0.128***	-0.032***	-0.035***	0.011***
	(0.009)	(0.003)	(0.007)	(0.001)	(0.001)	(0.001)
Higher 25 per. monthly net job flows	0.146***	-0.003**	-0.003	-0.003***	0.000	-0.007***
	(0.006)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)
Higher 25 perc. hiring incentive	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)
Higher than 25 perc. costs reduction	0.111***	-0.021***	0.091***	0.043***	0.051***	-0.001
	(0.005)	(0.004)	(0.010)	(0.007)	(0.008)	(0.002)
Higher than 25 perc. soc. insurance benefits	0.045**	-0.022***	0.097***	0.003	0.011	-0.010***
	(0.018)	(0.004)	(0.031)	(0.011)	(0.018)	(0.002)

Notes: The placebo sample comprises -6/+6 months around the Testo Unico Apprendistato. See notes in Table 4.

Table B8: Static model estimates on job tenure

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	-.03531	-.03043	-.02645	-.03209	-.03414	-.02732	.01317
	.238	.1341	.1527	.0783	.078	.0781	.0629
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

Notes: See notes in Table 4.

C1 Robustness

C1.1 Data centered ± 12 months around June 2012

Table C1: Main observable characteristics: difference in discontinuity in the robustness sample

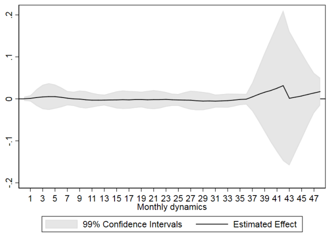
	Main Sample			
	Raw data		Polynomial fit	
	[-1,1]	[-2,2]	[-1,1]	[-2,2]
	DiD (Std. Dev.)	DiD (Std. Dev.)	DiD (Std. Err.)	DiD (Std. Err.)
Gender	-0.003^{***}	-0.003^{***}	-0.003	0.000
	0.001	0.001	0.045	0.076
Region of birth	-0.121	-0.288^{***}	-0.121	-0.610
	0.088	0.062	21.570	36.781
Education	-0.221^{***}	-0.089^{**}	-0.221	-0.306
	0.049	0.035	5.862	10.092
Missing education	0.001	0.001^{*}	0.001	-0.002
	0.001	0.001	0.108	0.186
Past experience	-13.522^{***}	-72.597^{***}	-13.522	92.317
	1.525	1.068	63.633	104.256
Missing past exp.	-0.004^{***}	0.006^{***}	-0.004	-0.020
	0.001	0.001	0.018	0.031
Region of work	-0.000	-0.008	-0.000	0.012
	0.012	0.008	0.759	1.324
Changing sector	0.001	-0.001^{*}	0.001	0.006
	0.001	0.001	0.019	0.034
Regional mobility	-0.003^{**}	-0.004^{***}	-0.003	-0.006
	0.001	0.001	0.186	0.318
Higher 25 per. monthly job spells	-0.002	0.001	-0.002	-0.006
	0.001	0.001	0.066	0.112
Higher 25 per. monthly sep. flows	-0.001	-0.000	-0.001	0.001
	0.001	0.000	0.016	0.028
Higher 25 per. monthly net job flows	-0.001	-0.003^{***}	-0.001	0.002
	0.001	0.001	0.016	0.027
Higher than 25 perc. costs reduction	0.001^{**}	-0.001^{**}	0.001	0.001
	0.000	0.000	0.011	0.019
Higher than 25 perc. soc. insurance benefits	0.000^{***}	-0.000^{***}	0.000	0.001
	0.000	0.000	0.001	0.001

Notes: The polynomial fit corresponds to a zero (first) order polynomial in age when the age range is $\pm 1(2)$. Each variable, defined as higher than the 25th percentile, is a dummy variable which is equal to 1 if the job episode sits in a percentile higher than the 25th of the age specific distribution of each covariate in a given month and year. See also notes in Table 4.

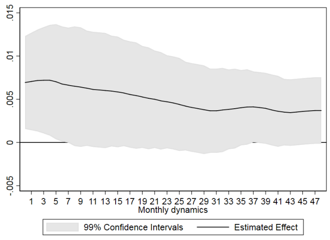
Table C2: Robustness sample: static model estimates

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment prob.	.00144	.00003	.00077	.00032	.00018	-.0006	.00035
	.0696	.0143	.012	.0072	.0072	.0074	.0019
Apprenticeship prob.	.00685**	.00685**	.00693**	.00689***	.00689***	.00692***	.00694***
	.003	.003	.0027	.0021	.0021	.0021	.0021
Perm. Employment prob.	.00707	.00674	.00776	.00738	.00733	.00707	.00753**
	.0182	.0138	.0102	.0047	.0047	.0047	.0037
Perm. Empl. prob. same firm	.00056	.00049	.0007	.0007	.0007	.00093	.00075
	.0038	.0013	.0015	.001	.001	.001	.0011
Perm. Empl. prob. same sector	.00157	.00148	.00175	.00172	.00173	.00194	.00183
	.0048	.0017	.0017	.0012	.0012	.0012	.0013
Self employment	-.00053	-.00049	-.00059	-.00075	-.00072	-.00059	-.00068
	.0033	.0028	.0026	.0015	.0015	.0015	.0014
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

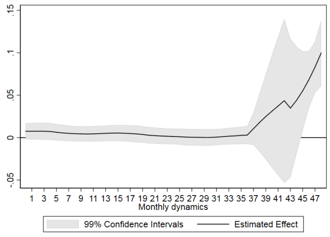
Notes: See notes in Table 4.



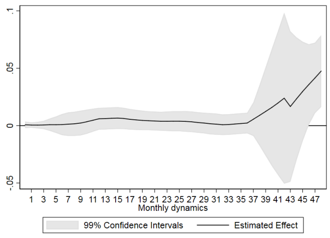
(a) Employment rate.



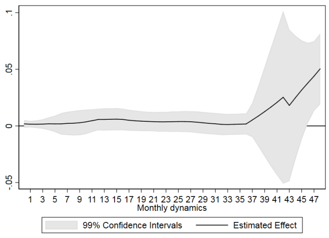
(b) Apprenticeship rate.



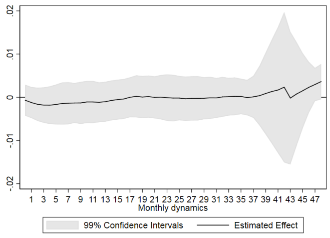
(c) Perm. Employment rate.



(d) Perm. Emp: same firm.

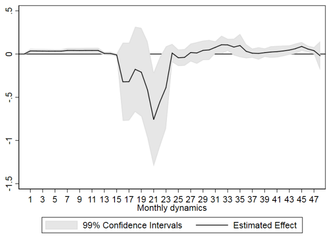


(e) Perm. Emp: same sector.

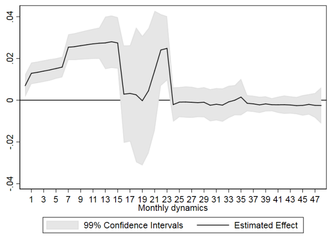


(f) Self Employment rate.

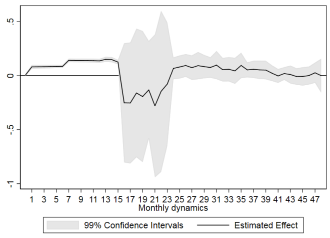
Figure C1: Robustness sample: difference in discontinuity, dynamic effect up to 36 months.



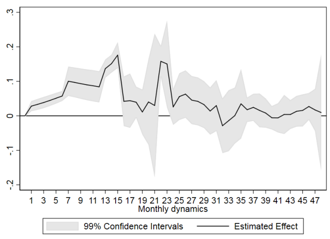
(a) Employment rate.



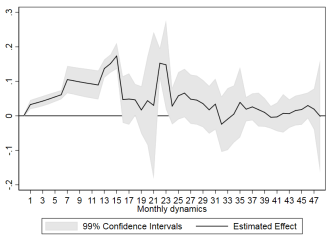
(b) Apprenticeship rate.



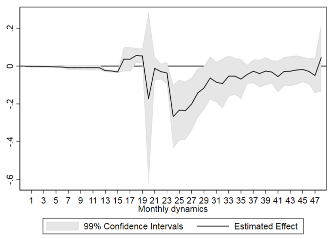
(c) Perm. Employment rate.



(d) Perm. Emp: same firm.

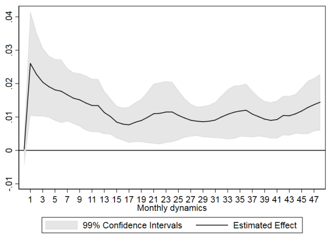


(e) Perm. Emp: same sector.

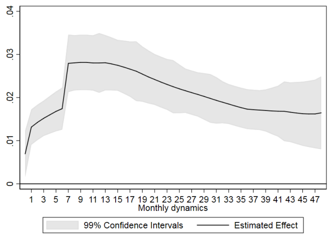


(f) Self Employment rate.

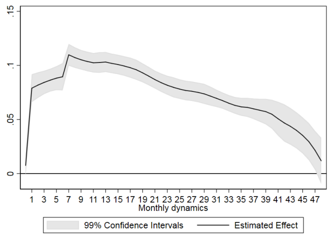
Figure C2: Robustness sample: difference in discontinuity, TOT parameters up to 47 months.



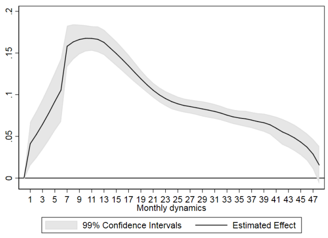
(a) Employment rate.



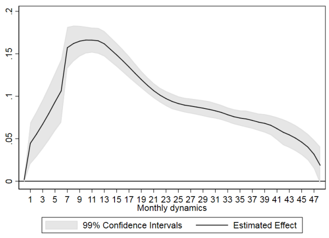
(b) Apprenticeship rate.



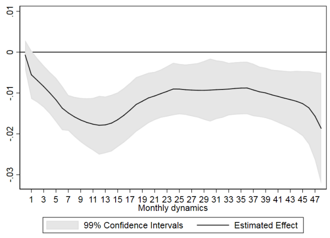
(c) Perm. Employment rate.



(d) Perm. Emp: same firm.



(e) Perm. Emp: same sector.



(f) Self Employment rate.

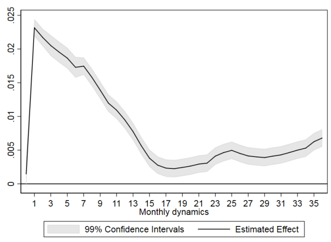
Figure C3: Robustness sample: difference in discontinuity, ITT parameters up to 47 months.

C1.2 Heteroskedasticity robust standard errors

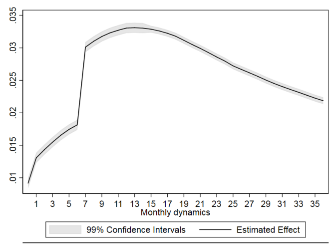
Table C3: Static model estimates

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment prob.	.00189	.0015	.00212	.00176	.00164	.00192	.00101
	.0015	.0014	.0014	.0014	.0014	.0014	.0008
Apprenticeship prob.	.01003***	.01004***	.01004***	.00998***	.00998***	.00993***	.00984***
	.0003	.0003	.0003	.0003	.0003	.0003	.0003
Perm. Employment prob.	.00861***	.00855***	.00905***	.00869***	.00866***	.00835***	.00925***
	.0009	.0009	.0009	.0009	.0009	.0009	.0009
Perm. Empl. prob. same firm	.00099**	.00095*	.00106**	.00109**	.00109**	.00088**	.00038
	.0005	.0005	.0005	.0005	.0005	.0004	.0004
Perm. Empl. prob. same sector	.00183***	.0018***	.00193***	.00195***	.00195***	.00172***	.00145***
	.0005	.0005	.0005	.0005	.0005	.0005	.0005
Self employment	-.00121**	-.00116*	-.00137**	-.00151**	-.00149**	-.00143***	-.00144***
	.0006	.0006	.0006	.0006	.0006	.0005	.0005
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

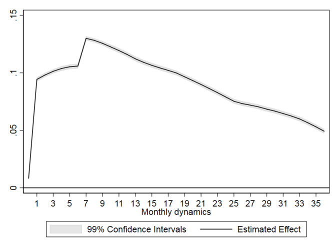
Notes: See notes in Table 4. Heteroskedasticity robust standard errors



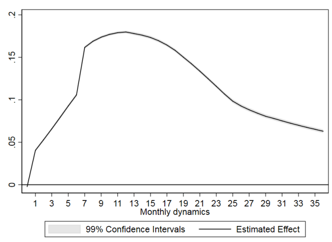
(a) Employment rate.



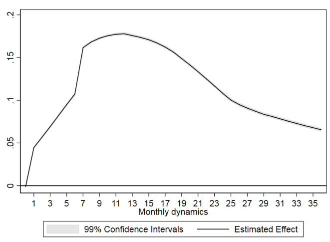
(b) Apprenticeship rate.



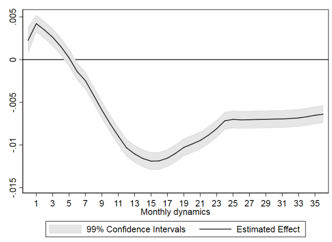
(c) Perm. Employment rate.



(d) Perm. Emp: same firm.



(e) Perm. Emp: same sector.



(f) Self Employment rate.

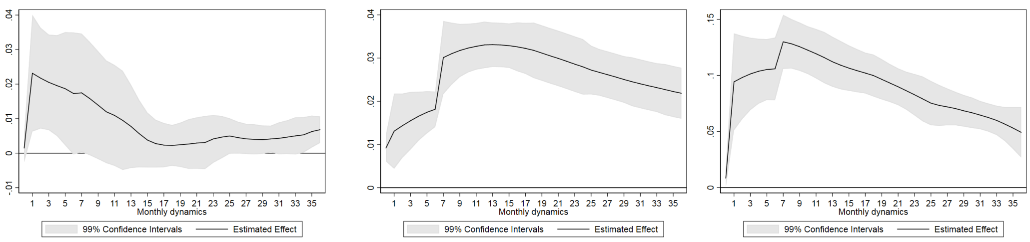
Figure C4: Heteroskedasticity robust standard errors: difference in discontinuity, ITT parameters up to 36 months.

C1.3 Clustering standard errors by age and year of birth

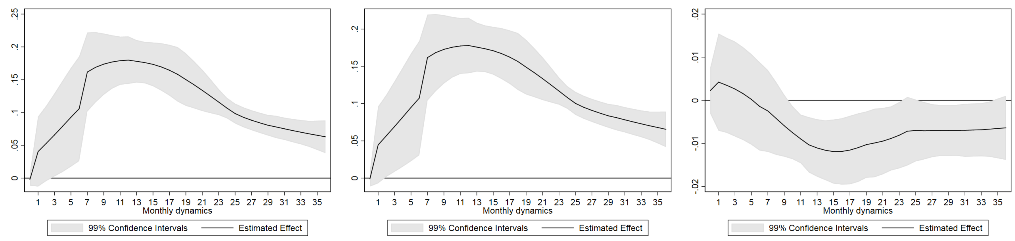
Table C4: Static model estimates

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment prob.	.00189	.0015	.00212	.00176	.00164	.00192	.00101
	.1067	.0016	.0018	.0018	.0017	.0018	.0016
Apprenticeship prob.	.01003***	.01004***	.01004***	.00998***	.00998***	.00993***	.00984***
	.0014	.0009	.001	.001	.001	.001	.001
Perm. Employment prob.	.00861	.00855***	.00905***	.00869***	.00866***	.00835***	.00925***
	.0202	.0007	.0006	.0005	.0005	.0005	.001
Perm. Empl. prob. same firm	.00099	.00095**	.00106***	.00109***	.00109***	.00088***	.00038
	.0055	.0004	.0003	.0003	.0003	.0002	.0003
Perm. Empl. prob. same sector	.00183	.0018***	.00193***	.00195***	.00195***	.00172***	.00145***
	.0065	.0003	.0003	.0003	.0003	.0004	.0004
Self employment	-.00121	-.00116*	-.00137*	-.00151**	-.00149**	-.00143*	-.00144
	.0045	.0007	.0007	.0007	.0007	.0008	.0009
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

Notes: See notes in Table 4. Standard errors clustered at age and year of birth level.



(a) Employment rate. (b) Apprenticeship rate. (c) Perm. Employment rate.



(d) Perm. Emp: same firm. (e) Perm. Emp: same sector. (f) Self Employment rate.

Figure C5: Standard errors clustered at age and year of birth level: difference in discontinuity, ITT parameters up to 36 months.

C1.4 Extending the age range to ± 2 years around the threshold.

Table C5: Static model estimates: range ± 2 years around the threshold

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment prob.	.00396	.00319	.00377	.0033	.00347	-.00048	-.00365
	.0653	.022	.0184	.009	.009	.0091	.0037
Apprenticeship prob.	.00923	.00923	.00916	.00907**	.00907**	.00911**	.00928**
	.0072	.0072	.0063	.0043	.0043	.0042	.0042
Perm. Employment prob.	.00952	.00937	.00961	.0091	.00915	.009	.00808
	.0254	.0235	.0161	.0059	.006	.006	.0052
Perm. Empl. prob. same firm	.00075	.00071	.0008	.00087	.00086	.00235	.00228
	.0041	.0027	.0032	.0022	.0022	.0022	.0021
Perm. Empl. prob. same sector	.00193	.00188	.00199	.00205	.00204	.00352	.00331
	.0052	.0035	.0036	.0024	.0024	.0024	.0023
Self employment	.00015	.00019	-.00016	-.00032	-.00035	.00028	.00039
	.0049	.0046	.0044	.0024	.0024	.0023	.0023
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

Notes: See notes in Table 4.

Table C6: Range ± 2 years around the threshold: interaction terms

	All sample			
	age	age*d	age*r	age*d*r
Employment prob.	.0033**	.0033*	.0002	-.0049*
	.0014	. .0020	.0020	.0030
Apprenticeship prob.	-.0006	-.0135***	.00043	-.0009
	.0014	.0018	.0019	.0033
Perm. Employment prob.	.0037**	-.0085***	.0001	-.0013
	.0017	.0028	.0027	.0040
Perm. Empl. prob. same firm	-.0003	-.0006	-.0001	.0025
	.0010	.0013	.0013	.0018
Perm. Empl. prob. same sector	-.0007	-.0014	-.0004	.0022
	.0010	.0015	.0013	.0020
Self employment	.0014	-.00002	-.0017	.0035*
	.0009	.0012	.0015	.0020
Time fixed effect	YES	YES	YES	YES
Sector fixed effect	YES	YES	YES	YES
Region of birth fixed effect	YES	YES	YES	YES
Firm fixed effect	YES	YES	YES	YES
Time invariant covariates	YES	YES	YES	YES
Time varying covariates	YES	YES	YES	YES

Notes: See notes in Table 4.

Table C7: Range ±2 years around the threshold: restricted model I

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment prob.	.0038	.00366	.00431	.00394	.00403	.00179	-.00128
	.06	.0199	.0167	.008	.0081	.0082	.0032
Apprenticeship prob.	.00981*	.00982*	.00978**	.00974***	.00973***	.00971***	.00974***
	.0051	.0051	.0045	.0036	.0036	.0035	.0035
Perm. Employment prob.	.00929	.00925	.00968	.00937*	.0094*	.00898*	.00869*
	.0221	.0203	.0138	.0054	.0054	.0054	.0046
Perm. Empl. prob. same firm	.00056	.00056	.00067	.00069	.00069	.00143	.00109
	.004	.0027	.0031	.0023	.0022	.0022	.0021
Perm. Empl. prob. same sector	.00171	.00171	.00184	.00186	.00186	.00257	.00225
	.0048	.0032	.0033	.0024	.0024	.0024	.0022
Self employment	-.00153	-.00152	-.00179	-.00188	-.0019	-.00144	-.0013
	.0048	.0043	.0042	.0023	.0022	.0022	.0022
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

Notes: See notes in Table 4.

Table C8: Range ±2 years around the threshold: restricted model II

	All sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment prob.	-.00003	-.00063	-.00011	-.00052	-.00085	.00195	.00326**
	.0269	.0089	.0075	.0037	.0037	.0037	.0015
Apprenticeship prob.	.01023***	.01025***	.01028***	.01019***	.0102***	.0101***	.00982***
	.0024	.0024	.0021	.0017	.0017	.0017	.0017
Perm. Employment prob.	.00794	.00786	.00839	.00795***	.00785***	.00767***	.00977***
	.0102	.0092	.0062	.0023	.0023	.0023	.002
Perm. Empl. prob. same firm	.00141	.00134	.00145	.00146	.00147	.00033	-.00029
	.0019	.0012	.0013	.001	.0009	.0009	.0009
Perm. Empl. prob. same sector	.00196	.00189	.00203	.00202**	.00203**	.00088	.0007
	.0023	.0014	.0015	.001	.001	.001	.001
Self employment	-.00084	-.00079	-.00096	-.00108	-.00103	-.00136	-.00154
	.002	.0019	.0018	.001	.001	.001	.001
Time fixed effect	NO	YES	YES	YES	YES	YES	YES
Sector fixed effect	NO	NO	YES	YES	YES	YES	YES
Region of birth fixed effect	NO	NO	NO	YES	YES	YES	YES
Firm fixed effect	NO	NO	NO	NO	YES	YES	YES
Time invariant covariates	NO	NO	NO	NO	NO	YES	YES
Time varying covariates	NO	NO	NO	NO	NO	NO	YES

Notes: See notes in Table 4.