INPS

Istituto Nazionale Previdenza Sociale

gennaio 2019 – numero 21

WorkINPS *Papers*

What are the returns to apprenticeships? Evidence from Italy

Luca Citino

ISSN 2532 -8565

In copertina: uno storico "Punto cliente" a Tuscania
INPS, Direzione generale, Archivio storico

I WORK*INPS* PAPER

Le basi dati amministrative dell' *INPS* rappresentano una fonte statistica unica per studiare scientificamente temi cruciali per l' economia italiana, la società e la politica economica: non solo il mercato del lavoro e i sistemi di protezione sociale, ma anche i nodi strutturali che impediscono all'Italia di crescere in modo adeguato. All' interno dell'Istituto, questi temi vengono studiati sia dai funzionari impiegati in attività di ricerca, sia dai *VisitInps Scholars*, ricercatori italiani e stranieri selezionati in base al loro curriculum vitae e al progetto di ricerca presentato.

I WORK*INPS* hanno lo scopo di diffondere i risultati delle ricerche svolte all' interno dell'Istituto a un più ampio numero possibile di ricercatori, studenti e policy makers. Questi saggi di ricerca rappresentano un prodotto di avanzamento intermedio rispetto alla pubblicazione scientifica finale, un processo che nelle scienze sociali può richiedere anche diversi anni. Il processo di pubblicazione scientifica finale sarà gestito dai singoli autori.

Pietro Garibaldi

What are the returns to apprenticeships? Evidence from Italy

Luca Citino

(London School of Economics)

INPS.0072.22/01/2019.0000056 E

What are the returns to apprenticeships? Evidence from Italy

Luca Citino*

(London School of Economics)

16th January 2019

Abstract

What are the returns to apprenticeships? Are these returns heterogeneous depending on the age at which one starts and the type of firm where one completes it? This paper tries to answer these questions by leveraging on novel administrative data from Italy on individual careers. We adopt a difference-in-difference methodology to compare the labor market outcomes of individuals starting an apprenticeship with those of similar individuals starting temporary contracts that, at least formally, do not provide formal training. Apprentices face significantly higher probabilities of transitioning to open-ended contracts and continuing their careers at the firms where they started their training. The payoffs are substantially higher for older individuals, and for individuals doing apprenticeships at bigger firms. We find workers' age and firm size to be complements in the determination of returns to apprenticeships.

*Contact:L.Citino@lse.ac.uk

Acknowledgement: I thank my advisors Jörn-Steffen Pischke and Stephen Machin for continued guidance and support. This project was made possible thanks to the VisitINPS initiative. I am thus very grateful to Massimo Antichi, Tito Boeri, Maria Cozzolino, Edoardo Di Porto, Pietro Garibaldi, Paolo Naticchioni and all the staff of Direzione Centrale Studi e Ricerche at INPS for making this programme possible and for their help with the data. This project has also benefited from discussions with Giulia Bovini, Anna D'Ambrosio, Alessandra Fenizia, Nicola Fontana, Xavier Jaravel, Matteo Paradisi, Alessandra Pasquini, Kilian Russ, Vincenzo Scrutinio, Martina Zanella and seminar participants at INPS. The findings and conclusions expressed are solely those of the author and do not represent the views of INPS. All remaining errors are my own.

1 Introduction

Apprenticeships are diffused in many European countries and constitute a middle-ground between high school and university education. Although there are differences across countries, apprenticeships usually consist of job contracts involving a combination of vocational education and training on the job in exchange for a salary and a final certification. In some countries firms are also granted reduced social security contributions or favorable taxation when hiring apprentices. Despite their diffusion, there is relatively little empirical evidence on returns to apprenticeships in terms of labor market outcomes. Part of this is due to the lack of high-quality administrative datasets that would allow researchers to track individual records over time, inside and outside apprenticeship programs. Also, it is hard to draw causal conclusions even from the best-quality datasets, because it is challenging to find an appropriate control group. Individuals who choose to get onto apprenticeships may be different from individuals who choose not to along important dimensions. This makes naive comparisons biased and prevents us from estimating the causal impact of apprenticeship programs.

This paper tries to address these issues by leveraging on high-quality administrative matched employer employee data from the Italian Social Security Institute (INPS). We have access to the full working history of the universe of individuals born in Italy in 1980 and 1981, regardless of whether they have been employees in the private sector, in the public sector, dependent self-employed (*parasubordinati*) or self-employed. Thanks to the richness of the data, it is easier for us to find an appropriate control group of individuals who have not started an apprenticeship during their youth but could have.

We define *returns to apprenticeships* as the extra gain coming from starting an apprenticeship compared to a temporary contract. Similarly to apprenticeships, temporary contracts also involve an employer-employee relationship but, at least formally, they don't require the firm to provide training.[2] Whether such training has any real content remains an empirical question. We know that firms may have scarce incentives to train if the human capital they need for production is general (Becker, 1962) and even more so if the labor market where they operate does not feature any frictions (Acemoglu and Pischke, 1999). In such cases, given the low enforcement level of apprenticeship contracts, firms may renege on the promise to provide training and the returns to apprenticeships would be close to zero. To the contrary, if firms need firm-specific human capital, or if labor market frictions are substantial, then apprenticeships should have a higher premium.

In addition to studying whether there are any returns to apprenticeship at all, we take advantage of the panel dimension of our data to trace the full path of returns at the quarterly frequency, from

[2] In our data we don't observe the contractual duration *ex ante*, but just the realized duration. As a consequence all our returns ought to be interpreted as returns to starting an apprenticeship. Since it is unheard that governments would force firms to retain apprentices, we think this parameter is indeed the policy relevant one.

the start of the contract up to six years afterwards. This allows us to appreciate the magnitude of the trade-offs between short-run investment costs and long-run benefits. We carry out our analysis separately by age at which individuals start either apprenticeships or temporary contracts, and by firm size. We suspect these to be important sources of heterogeneity when studying returns to apprenticeships.

To preview our results, we find that apprenticeships can yield positive returns in terms of future job stability, measured as the probability of working under an open-ended contract, and tenure at the initial firm. While the net returns are positive for individuals starting apprenticeships at all ages, we find the long-run returns to be positive and economically meaningful only for individuals starting apprenticeships after age 24. We can conclude that after an initial career boost, younger individuals initially enrolled into apprenticeships converge back to their initial career path. This distinction between older and younger apprentices is even more important when we consider contracts started at bigger firms: the gap between temporary contracts and apprenticeships widens with firm size, but only for older apprentices.

Thanks to the richness of our data, we are also able to document apprentices' transitions towards self-employment. Consistent with their improved opportunities as employees, relatively old apprentices are less likely to move into self-employment. The opposite is true for young apprentices, but this effect is quantitatively small. Overall, older individuals command higher returns than younger individuals in terms of labor market outcomes.[3]

This paper contributes to the literature documenting the effects of apprenticeships on labor market outcomes. Contrary to the literature on the returns to schooling there is not a consensus on what is the right control group for apprentices, and different studies have followed different routes. Fersterer et al. (2008) compare longer and shorter apprenticeships. For identification they exploit the unexpected closure of firms that employ apprentices at different tenure horizons. At such *intensive margin*, they find that an extra year into apprenticeship yields a 3.8% return in terms of higher earnings. Parey (2016) compares firm-sponsored training with school-based vocational education. He finds that the two tracks do not offer different returns, but that in the very short run firm-based apprenticeships provide stronger labor market attachment. Also he finds no effects on wages. Similarly Albanese et al. (2017) compares two apprenticeship tracks that co-existed in Italy in the early 2000s, one of which emphasized firm-sponsored training rather than school-based vocational education. In line with Parey (2016), they find that firm-sponsored training improved the prospects of young workers, increasing their probability of transitioning to open-ended contracts and raising their wage levels, especially in bigger firms.

Fewer papers have analyzed the returns to vocational education at the *extensive margin*, that is comparing apprenticeships to other types of temporary contracts that are not formally linked to the provision of training. This is the approach we follow in this paper. The studies closest

[3] In future versions of this work we also aim to look at the impact on earnings.

in spirit to ours therefore are Picchio and Staffolani (2013) and Berton et al. (2011). The first paper exploits age limits in the Italian apprenticeship system and use a regression discontinuity design to compare individuals who manage to get an apprenticeship just before age 30 and those who do not manage to do so. The authors find that, around age 30, individuals who start an apprenticeship are more likely to transition towards open-ended contracts, especially at the initial firm. The second paper uses a Multinomial Logit with individual fixed effects to study the transition matrices between different types of temporary contracts (including apprenticeships) and open-ended contracts. Both these papers try to address selection issues but must nevertheless limit their analysis to a short-run horizon We extend these analyses by characterizing the full age and time profile of returns to apprenticeships, and unfold important heterogeneity along the firm size dimension. We look into the very long run up to 6 years after the start of the contract. In addition to this, we are able to look at new outcomes that were unstudied before due to data limitation, such as the probability of entering self-employment.

The paper is structured as follows. In Section 2 we describe how apprenticeships are regulated in Italy and the data we employ for our analysis. In Section 3 we present our identification strategy and regression framework. In Section 4 and Section 5 we present our main findings and in Section 6 we conclude.

2 Institutional framework and Data

2.1 Apprenticeships in Italy

The Italian apprenticeship system is made of three separate programmes, with different rules: (1) *Apprendistato per l'espletamento del diritto/dovere di istruzione* (apprenticeship for the completion of compulsory schooling), performed during upper secondary education for individuals up to age 18 (2) *Apprendistato professionalizzante* (apprenticeship for a professional qualification), performed after the completion of secondary education, for individuals aged 18-29 and (3) *Apprendistato di alta formazione e ricerca* (apprenticeship for high-skill qualifications and research), still oriented to individuals between 18 and 29, but who are enrolled in or have already earned a university degree and would like to carry out a thesis or a research project within a firm. In our analysis we require individuals to be at least 22 when doing their apprenticeships, so this excludes type (1) apprenticeships by construction.[4] On the other hand, in the data we do not have information needed to distinguish apprenticeships of type (2) from those of type (3) before 2007, so in what follows type (2) and type (3) are pooled together. The vast majority of apprenticeships in Italy are of type (2).

[4]Our analysis excludes individuals younger than 22 at start of the contract in order to have sufficient information on the pre-event working history.

In terms of contractual obligations, apprenticeships are job contracts signed between firms and workers, mostly limited to the private sector, for which workers and firm regularly pay social security contributions and work accidents insurance. It must be stressed that the formal training content of apprenticeships is quite low. The minimum number of training hours that the firm must provide is 120 per year, split in the following way: 65% are dedicated to occupation-specific training and 35% are dedicated to general training (job safety, psychology of labor and team working). In exchange for training, firms obtain a reduction in social security contributions. The latter amounts to 10% of apprentices' gross earnings, compared to 27% for open-ended and temporary contracts. Also, firms can pay apprentices a lower wage, up to two levels below what a qualified worker would get, according to the corresponding collective bargaining agreement (CBA). At the end of the programme the workers receive a certification which is recognized by firms applying the same CBA. This implies a worker cannot be trained twice for the same occupation in the same CBA. Eligibility on the side of firms is linked to the presence of a *mentor*. The mentor must attend preparatory training and cannot train more than 5 individuals at each point in time. The law sets ceilings in apprenticeship use: they can never be more than the number of qualified workers in the firm (however if firm size is less than 3 the firm can hire up to 3 apprentices). Eligibility on the side of workers is exclusively age-dependent. Recent reforms raised the age limits (measured on the day of hiring).[5] A more complete description of the Italian apprenticeship contract and its recent reforms can be found in Albanese et al. (2017).

2.2 Data and descriptive statistics

We use rich and novel administrative data on careers at the individual level made available by the Italian Social Security Institute (INPS).

Matched employer-employee data: our primary source is a matched employer-employee dataset covering all job spells in non-agricultural firms with at least one employee. The dataset spans the whole time period 1983-2017. The public sector and firms with no employees are not included. The data records the presence of job spells at the monthly frequency, which gives us the advantage of detecting nonemployment or job spells that last less than a year. Thanks to this we are able to trace career dynamics at a very fine level. In each month we observe at which firm(s) the worker is employed, the type of contract(s) the worker has (open-ended, temporary, dependent self-employed, self-employed), the type of work-time arrangement (full time or part time) and a job ladder code (apprentice, blue collar, white collar, supervisor or manager). Absent any change in the aforementioned characteristics, we observe one earning record per year for each worker. In case a worker has a contractual change during the year (e.g. becomes a white collar worker) we see two separate earning records. This allows us to precisely separate earning

[5] The 1997 (*Treu* reform): from age 20 to age 24 (but 27 in regions entitled to EU structural funds - i.e. the South - and age 29 in artisan firms). The 2003 (*Biagi* reform): from age 24 to age 29 in all firms in all regions.

records which belong to different contract characteristics, different firms and different years. For each individual we also observe a series of basic socio-demographic characteristics such as gender, year of birth and place of birth. Given the nature of the dataset, we are also able to build the total firm size in every year, and therefore check whether individuals starting apprenticeships in bigger firms obtain higher returns.

Dependent self-employment spells: starting from 1996, we also have information on dependent self-employment. The latter is a form of work where workers are formally self-employed but *de facto* employees. This dataset also has a matched employer-employee structure. For each job spell we observe unique worker and firm identifiers, the beginning and end date of the spell, the type of contract and the overall compensation received for the job in every year. Given that firm and worker identifiers are the same across datasets we are able to merge this information with the matched employer-employee dataset.

Contribution Histories: for a subset of individuals in the matched employer-employee dataset we were able to obtain further information on their full contribution history, including spells in the public sector, and as self-employed. This allows us to build more precise measures of work experience and investigate whether apprenticeships have an impact on the probability of entering self-employment. Unfortunately this dataset does not contain a firm identifier. We obtained such information for the universe of individuals born in Italy between 1980 and 1981, that is our main sample of interest.

Sample selection: we select all individuals born in 1980 and 1981. We choose to focus our attention on these two cohorts because in the INPS data the information on whether an individual works in a open-ended or temporary contract is only available from 1998 onwards (approximately when our individuals leave school). On the other hand we don't choose cohorts younger than 1981 to have a long enough period to observe them in the post event years. For each individual we consider either the first temporary contract or first apprenticeship. We drop all individuals who do an apprenticeship and a temporary contract at the same age; and drop all temporary contract workers who did an apprenticeship beforehand. Contrary to this, we allow apprentices to have had temporary contracts before. We leave unrestricted the possibility that individuals in temporary contracts embark onto apprenticeships later on (or vice-versa). We end up with a sample of 319,476 individuals who either start an apprenticeship or a temporary contract at some point during the age range 22-29.

Descriptive statistics for our sample are available in Table 1. Characteristics are measured two years before the start of the contract (either temporary or apprenticeship). The starkest difference is in overall experience, and the size of the firm where the contract starts. Despite having entered in the labor market at roughly the same age as individuals in the control group, apprentices have considerably less experience. Also, they are much more likely to have started their contract in a small firm. This implies that apprentices are negatively selected, and that a naive comparison

would bias returns downwards. Other than this, the composition of their careers in terms of contract characteristics and the probability of being a female are roughly similar.

3 Estimating returns to apprenticeships

In this section we present the empirical strategy we use to identify returns to apprenticeships. We define returns to apprenticeship as the extra gain in labor market outcomes an individual obtains from starting an apprenticeship relative to another type of temporary contract that does not oblige the firm to provide training. We measure returns to apprenticeships at different time horizons after the start of the contract.

In comparing individuals in apprenticeships and in temporary contracts, one challenge arises: individuals starting temporary contracts may be on different career trajectories even before starting their respective contracts. This would make a naive comparison between the two groups meaningless. In order to solve this problem we implement a two-step procedure that we detail in the two following subsections. First we use Coarsened Exact Matching (CEM) (Iacus et al., 2012) to find a subsample of individuals starting temporary contracts that are observationally similar to yet-to-be apprentices 2 years before the start of the contract. Secondly, we run a staggered difference-in-differences (DiD) specification where we normalize event time to zero and compare individuals careers before and after the start of the contract.

The combination of CEM and DiD is now standard in the literature, and has been used to study a variety of topics in Labour and Public Economics (Azoulay et al., 2010; Jäger, 2016; Jaravel et al., 2018)

3.1 Coarsened Exact Matching

The first step in the empirical strategy consists of purging the sample from control individuals that are too dissimilar from the apprentices under study in their careers up to 2 years before the start of the contract. We employ a Coarsened Exact Matching procedure (CEM) (Iacus et al., 2012) at the individual level, where we match on a combination of gender, age at entry in the labor market (in years), 6 months bins of overall labor market experience, deciles of average weekly wage as an employee and of the average daily wage as dependent self-employed, 10 percentage point bins of the share of experience with a full-time contract, and similarly for the share of experience as a blue collar and the share of experience as a dependent self-employed. Once individuals are grouped in *strata* according to the aforementioned characteristics, we drop all strata where there are either only treated or only control individuals. We are left with a sample of 230,382 individuals. Descriptives on the matched sample are found in Table 2.

Importantly, we do not match on characteristics in the time period from 2 years to 1 quarter before the event. Instead we use this time frame to test for the presence of differential pre-trends in the outcome variables across the two groups. Once this is verified we employ a staggered difference-in-differences methodology, where we compare, quarter by quarter, the career evolution of individuals starting a temporary contract with those starting an apprenticeship.

The identification assumption is that apprentices would have followed the same *trend* in outcomes as individuals starting temporary contracts, had they started temporary contracts instead. In the next subsection we detail the regression specification that we run in order to retrieve returns to apprenticeships.

3.2 Staggered Difference in Differences

In line with the ideas described in the previous section, we present the regression that we run to measure returns to apprenticeships. We run equations as in 1, separately by age in years, at the quarterly frequency:

$$\begin{aligned} Y_{it} &= \alpha_i + \eta_t + \theta_a + \zeta_y + \sum_{k=-12}^{d_M} \beta_k \times \mathbf{1}(\text{distance} = k) \\ &+ \sum_{k=-12}^{d_M} \beta_k^T \times \mathbf{1}(\text{distance} = k) \times \text{Apprentice}_i + \epsilon_{it}. \end{aligned} \tag{1}$$

where Y_{it} is a labor market outcome measured in year $\times$ quarter t; α_i are individual fixed effects, which control for any time-invariant unobserved heterogeneity at the worker level, and η_t are year $\times$ quarter fixed effects, which control for time-varying unobservables that are common across the two groups. We also include age in quarters fixed effects (θ_a) to control for life-cycle patterns even within age in years and year of birth dummies to control for cohort patterns.[6] Given that both our treatment and control group are assigned to a job contract at distance time $k = 0$, we include both a set of distance-to-event dummies that are common to both groups i.e. $\mathbf{1}(\text{distance} = k)$, and a set of distance-to-event dummies interacted with treatment i.e. $\mathbf{1}(\text{distance} = k) \times \text{Apprentice}_i$. Distance time runs from -8 quarters till an upper bound d_M that is age specific. This specification is very similar to Jaravel et al. (2018) and addresses the presence of time-varying differences around the start of a new job contract. This is akin to a tenure profile.

Regressions are weighted according to the procedure detailed in Iacus et al. (2012). We call equation 1 a staggered difference in difference (DiD) specification. Our identification rests on the assumption that, conditional on a rich set of covariates, and time-invariant unobservable characteristics at the individual

[6] Given that the regression is run separately by age in years, θ_a and ζ_y do not affect the results in a substantive manner.

level, the treatment assignment rule is as good as randomly assigned. In other words we assume that there is no selection into treatment based on the *trend* of the outcome variable.

The coefficients of interest are the β_k^T, for $k \neq -1$. Due to multicollinearity issues we omit both $\mathbf{1}(\text{distance } = -1) \times \text{Apprentice}_i$ and $\mathbf{1}(\text{distance } = -1)$. All coefficients β_k^T must thus be interpreted as changes in the difference across the two groups with respect to any pre-existing difference at distance $k = -1$ (one quarter before event). It follows that $\beta_k^T = \beta_k = 0 \forall k < 0$ implies the absence of differential trends in outcome variables before the start of the treatment.

Once the assumption of zero pre-trends is graphically verified, we run the following specification in order to increase power:

$$\begin{aligned} Y_{it} &= \tilde{\alpha}_i + \tilde{\eta}_t + \tilde{\theta}_a + \tilde{\zeta}_y + \sum_{k=0}^{d_M} \tilde{\beta}_k \times \mathbf{1}(\text{distance } = k) \\ &+ \sum_{k=0}^{d_M} \tilde{\beta}_k^T \times \mathbf{1}(\text{distance } = k) \times \text{Apprentice}_i + \tilde{\epsilon}_{it}. \end{aligned} \tag{2}$$

which is very similar to equation 1 but the indices of summation start from 0 instead of -12. Effectively we are imposing that $\tilde{\beta}_k^T = \tilde{\beta}_k = 0 \; \forall k < 0$, and gain statistical power as a consequence. As in Borusyak and Jaravel (2016), we call this specification semi-dynamic. In the semi-dynamic specification, causal effects are to be interpreted as differences with respect to the average difference of the outcome between the two groups in the pre-event period.

4 Graphical evidence of returns to apprenticeships

Figures 1, 2, 3, 4 report the coefficients β_k^T and associated 95% confidence intervals from specifications 1 (black dots) and similarly coefficients $\tilde{\beta}_k^T$ from specification 2 (blue dots), for different outcome variables: the probability that the worker still works at the same firm as in $k = 0$, that she obtains a open-ended contract at such firm, that she obtains a open-ended contract at any firm, and that she is self-employed. Each subfigure refers to a particular age group. Coefficients are plotted at a quarterly frequency. The number of post-event quarters differs by age group and corresponds to the maximum number of periods available for that age group. Given that all individuals are either born in 1980 or 1981, younger individuals are observed for a greater number of periods.

The point estimates for β_{-1}^T are set to 0 and all other coefficients are interpreted as differences from period -1. Pre-trends are quantitatively small in many outcomes except for the probability of having an open-ended contract in any firm, but only at younger ages. Yet-to-be apprentices have a higher probability to have transitioned from such types of contracts even before the event. In future extensions of this work we plan on finding techniques to reduce the amount of pre-trends.[7]

[7] This is not unexpected. One of the matching variables is indeed the share of experience spent in open-ended

While such figures are useful for grasping the extent to which outcomes display differential pre-trends across groups, they are not the best way to explore heterogeneity by age. We now overlay these graphs so that it is clear what is the common period we observe for all ages and returns are easily comparable. The vertical lines in these graphs indicate the number of quarters during which we are able to observe all age groups. For this latter set of graphs we always employ the semi-dynamic specification. Formally, in this case the $\tilde{\beta}_k^T$ have to be interpreted as differences from the *average* difference between the two groups in the pre-event period. All coefficients in the post-event period are therefore simply shifted by a constant value with respect to the staggered DiD. If outcomes are not trending strongly in the pre-period, the semi-dynamic specification and the staggered one should yield similar results.

In Figure 5 we study whether apprentices are more likely to still work at the same employer as in distance time $k = 0$. We observe two facts: first, apprenticeships constitute more stable jobs. Compared to temporary contracts, the difference in the probability of retention is as high as 34 p.p. after one year, and then slowly decays over time. Second, there is a clear positive age gradient, whereby older individual obtain higher returns. Six years into the apprenticeship route, individuals aged 29 have a higher probability to still work at the initial firm by 18 p.p., compared to the control group. The same number is only 8 p.p. for individuals aged 22. Although our information is incomplete for horizons longer than 6 years, we can still notice a general trend whereby returns decay and then plateau. We can speculate that, while the very long-run return (at around 12 years) is close to zero for very young individuals, it would still be high for older individuals.

In order to check whether these longer tenures also reflects the fact that workers get promoted at the initial firm, in Figure 6 we look at the probability that an apprentice still works at the same employer and has an open-ended contract. After an initial hollow lasting around 3 years, where apprentices are "locked-in" their training contract, we see that apprenticeships yield higher returns for a very long period of time. After 6 years, the difference in probability peaks at 8.7 p.p. for younger individuals, and can reach up to 21.2 p.p. for older individuals.

In Figure 7 we explore whether individuals starting apprenticeships have higher probabilities of transitioning to open-ended contracts in any firm. We observe a qualitatively similar picture. Again there is a stark division between negative short-run returns and (weakly) positive long-run returns. After approximately 12 quarters, when many apprenticeships end, there is a jump in the probability to have an open-ended contract. The return after 4 years is 7.2 p.p. for younger individuals, although such probability differential converges to zero in the very long-run. For older individuals returns get as high as 21 p.p. over the same time horizon. Contrary to what we observed for other outcomes, these returns seem to be approximately stable over time.

A slightly different pattern emerges when looking at the probability of being self-employed. Looking at Figure 8 we see that, overall, apprentices have lower probabilities to transition to self-employment for as long as six years after the start of the apprenticeship. However, older individuals are significantly less

contracts. However for individuals who start apprenticeships at age e.g. 21 or 22, this is measured at age 19 and 20 respectively. It is still possible that individuals have a similar share when leaving school, and then they diverge in the time period from 8 to 1 quarter before the event. When studying individuals starting apprenticeships at age e.g. 25, we have more pre-event experience to match on, and therefore the power of our controls is higher.

likely to work as self-employed. This is the mirror image of the fact older individuals kept working for their initial firm. This is an important finding because it highlights the fact that not necessarily younger apprentices are faring worse in the labor market. They may be simply transitioning to other forms of employment. It has to be said though that while this effect is statistically significant, it is quantitatively negligible, as the difference is only 2 percentage points and only materializes after 10 years. It's interesting to note that the transition towards self-employment does not occur sharply after 3 years, but manifests itself smoothly.

5 Are returns higher in big firms?

In this section we replicate our entire analysis but restricting the sample to individuals who either started an apprenticeship or a temporary contract in a "big" firm. We label firms as big if they are above the median of the firm size distribution in that year. Figures 9, 10, 11, 12 thus have exactly the same structure as Figures in Section 4. Although the general picture is qualitatively similar, there are important differences in the economic magnitudes involved.

In Figure 13 we look at whether individuals still work at the initial firm as in period $k = 0$. While for younger individuals, aged 22 and 23, the size of the firm is not a relevant factor in determining returns to apprenticeships, the same cannot be said for individuals aged 24 and older. Compared to individuals starting temporary contracts, older apprentices have a higher probability of still working at the same firm in the order of 15-29 p.p.. The same figure was approximately 8-18 p.p. in the overall sample. To the extent that bigger firms have greater market power in the labor market or employ technologies that require firm-specific human capital, they may be more able to retain workers at the end of the training than the average firm (Acemoglu and Pischke, 1999). It is still not entirely clear is why this should be different between younger and older apprentices. We intend to investigate this in future versions of this work.

A similar pattern emerges when looking at conversions to open-ended contracts, both at the initial firm and in any firm (Figures 14 and 15). The gap between temporary contracts and apprenticeship does not widen with the size of the firm for younger individuals, but it does so for older individuals. After 6 years, apprentices aged 28-29 have substantially higher probabilities to have been converted to open-ended contracts at the same firm (+38 p.p.) and at any firm (+30 p.p).

In Figure 16 we investigate what happens to the probability of being self-employed. This outcome is the only one to display a qualitatively different trend compared to the whole sample of firms. All individuals, irrespective of their age at the start of the contract, have a lower probability of becoming self-employed. This is interesting: while older individuals have better prospects as employees, and therefore we would expect them to have lower probabilities to become self-employed, the contrary should be happening for younger individuals. This is however not the case: younger individuals do not compensate with self-employment the low returns that they get as employees. We will explore reasons behind this pattern in future versions of our work.

6 Concluding remarks

In this paper we have studied the returns to apprenticeships for two cohorts of Italian individuals. We employed a combination of matching and difference-in-differences techniques to trace the whole age and time profile of returns. We find that apprenticeships are indeed a superior port of entry into stabler job contracts, especially for older individuals and in bigger firms. Apprentices are significantly more likely to transition to open-ended contracts and to keep working at their initial employer. Young apprentices are slightly more likely to become self-employed while the opposite is true for older apprentices.

7 Tables

Table 1: Summary statistics - entire sample

	Temporary		Apprentices	
Variable	**Mean**	**SD**	**Mean**	**SD**
Overall experience (months)	17.656	28.907	9.92	17.442
Female	0.479	0.5	0.478	0.5
Never worked	0.491	0.5	0.512	0.5
Share of experience as blue collar	0.174	0.337	0.222	0.389
Share of experience with an open-ended contract	0.276	0.409	0.164	0.334
Share of experience with a full time contract	0.28	0.424	0.243	0.4
Share of experience as dependent self-employed	0.023	0.127	0.028	0.197
Age of entry into the labour market	21.472	3.89	21.351	2.817
Age at first apprenticeship	NA	NA	23.8	2.328
Age at first temporary contract	25.03	2.153	24.55	4.74
Share starting contract in big firm	0.615	0.487	0.249	0.433

Notes: This table provides summary statistics for our full sample, inclusive of 216,815 temporary contracts and 102,661 apprentices. All characteristics are measured 8 quarters before the start of the contract. Sample selection is described in more detail in Section 2.2.

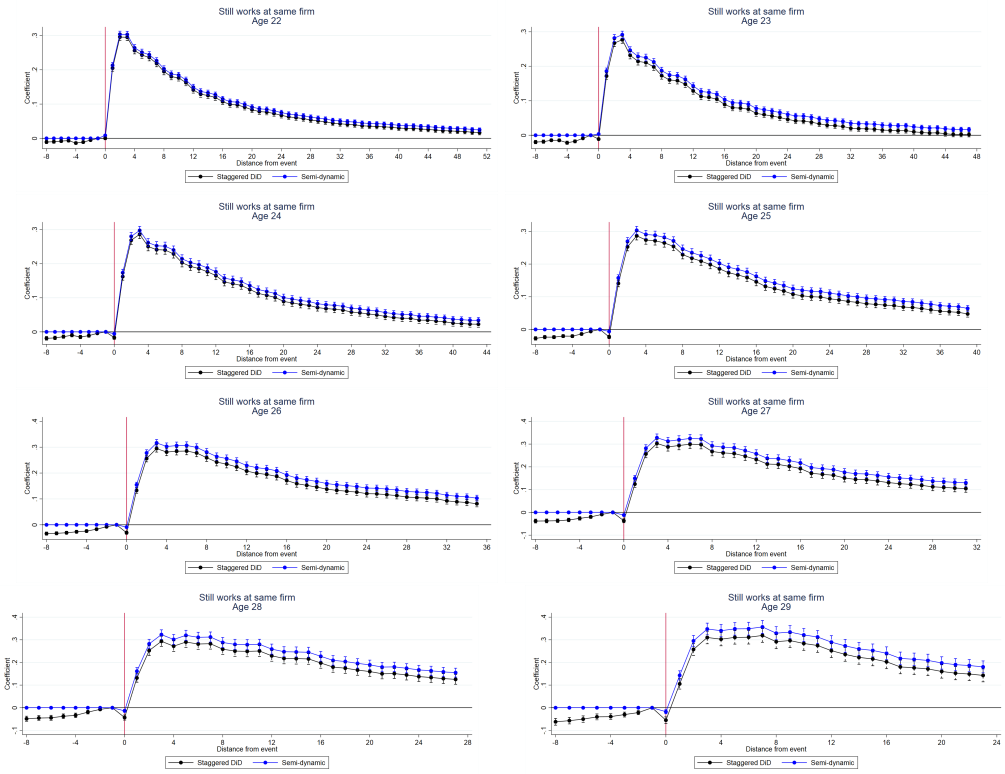
Table 2: Summary statistics - matched sample

	Temporary		Apprentices	
Variable	**Mean**	**SD**	**Mean**	**SD**
Overall experience (months)	5.09	12.675	4.681	11.72
Female	0.504	0.5	0.485	0.5
Never worked	0.699	0.458	0.673	0.469
Share of experience as blue collar	0.101	0.298	0.145	0.348
Share of experience with an open-ended contract	0.126	0.327	0.116	0.314
Share of experience with a full time contract	0.124	0.328	0.155	0.359
Share of experience as dependent self-employed	0.013	0.112	0.019	0.188
Age of entry into the labour market	23.128	3.013	22.071	2.624
Age at first apprenticeship	NA	NA	23.615	2.044
Age at first temporary contract	24.72	2.06	26.04	4.537
Share starting contract in big firm	0.667	0.471	0.233	0.423

Notes: This table provides summary statistics for our matched sample, inclusive of 152,239 temporary contracts and 78,143 apprentices. All characteristics are measured 8 quarters before the start of the contract. Sample selection is described in more detail in Section 2.2.

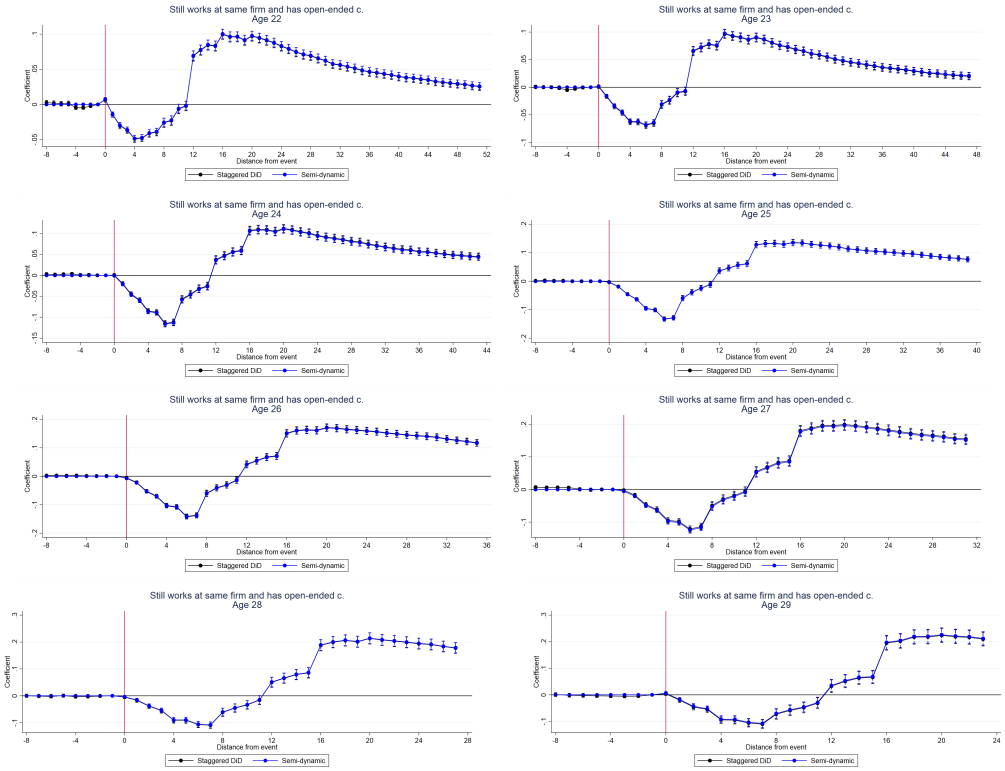
8 Figures

Figure 1: Still works at initial firm



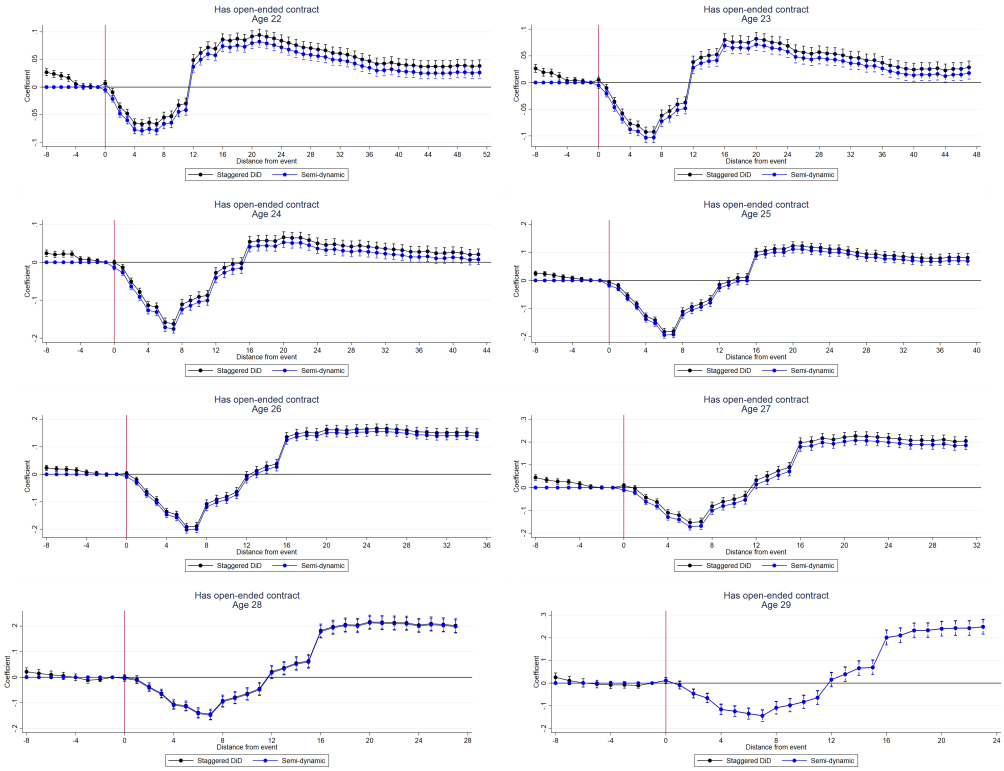
Note: The figure plots coefficients β_k^T (in black) coefficients $\tilde{\beta}_k^T$ (in blue) with associated 95% confidence intervals from specifications 1 and 2 respectively. Time is in quarters.

Figure 2: Has an open-ended contract (initial firm)



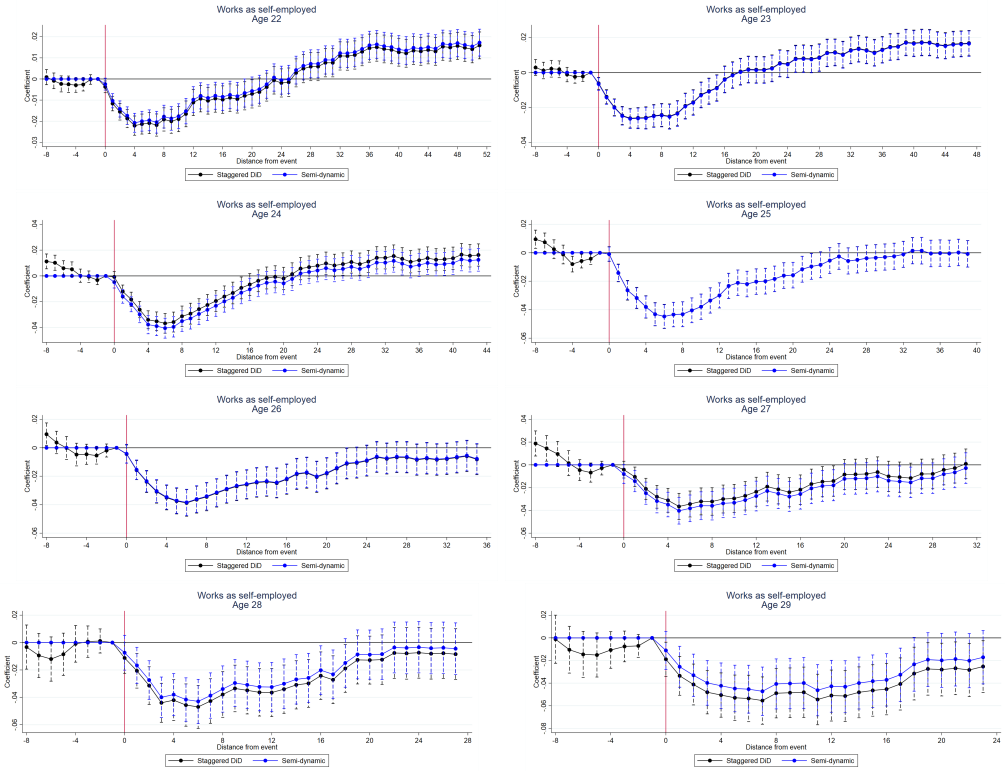
Note: The figure plots coefficients β_k^T (in black) coefficients $\tilde{\beta}_k^T$ (in blue) with associated 95% confidence intervals from specifications 1 and 2 respectively. Time is in quarters.

Figure 3: Has an open-ended contract (any firm)



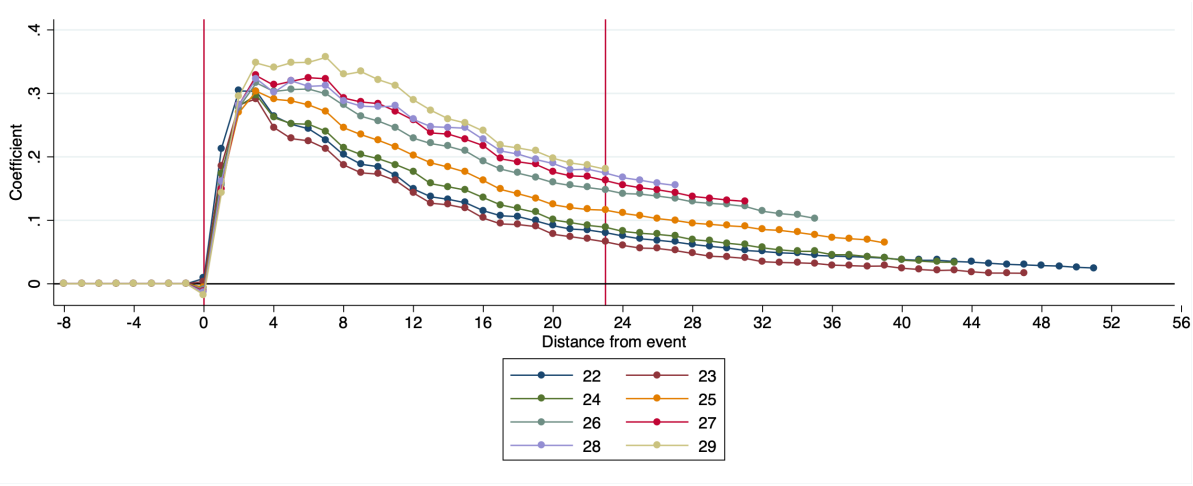
Note: The figure plots coefficients β_k^T (in black) coefficients $\tilde{\beta}_k^T$ (in blue) with associated 95% confidence intervals from specifications 1 and 2 respectively. Time is in quarters.

Figure 4: Works as a self-employed



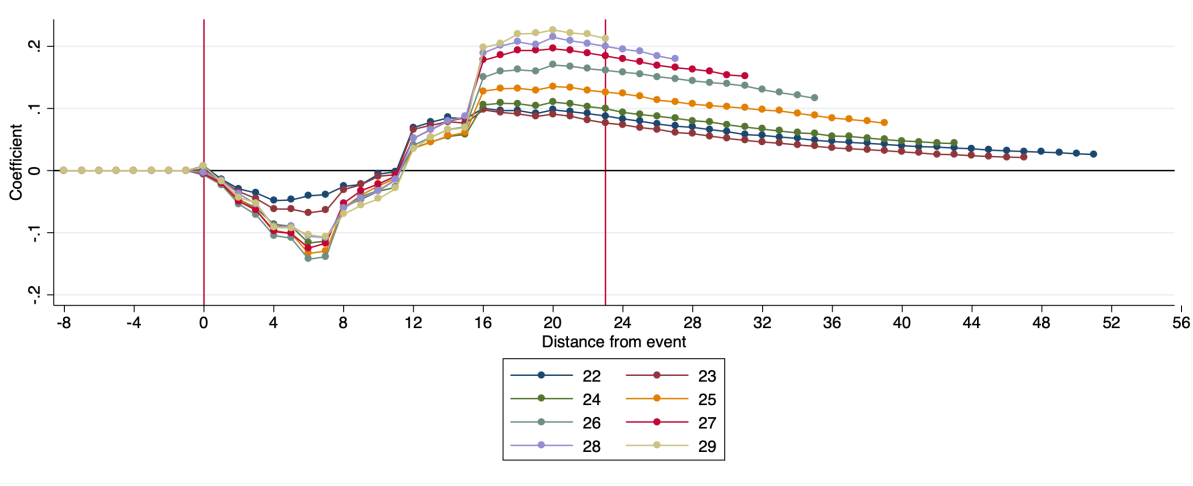
Note: The figure plots coefficients β_k^T (in black) coefficients $\tilde{\beta}_k^T$ (in blue) with associated 95% confidence intervals from specifications 1 and 2 respectively. Time is in quarters.

Figure 5: Still works at initial firm



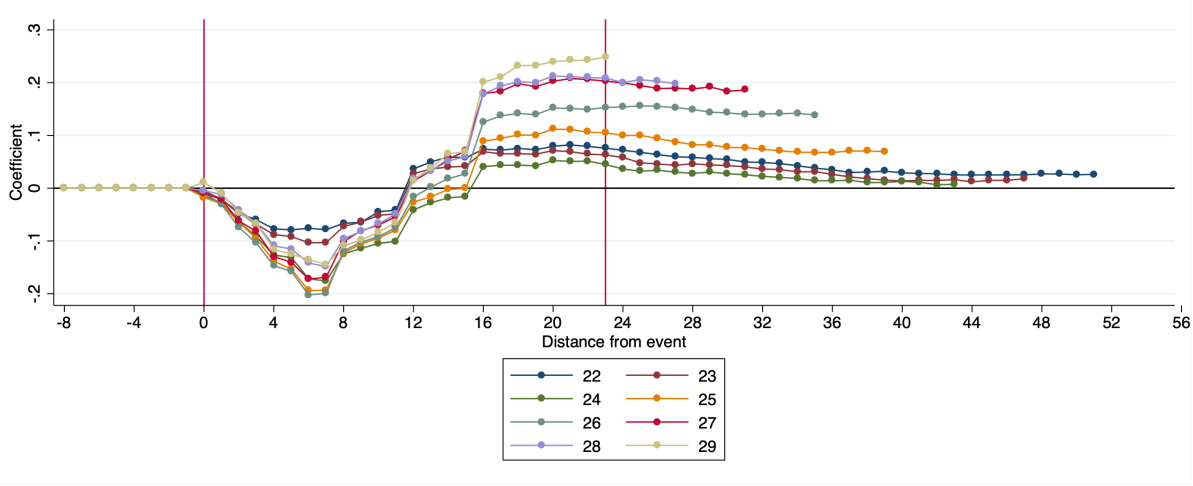
Note: The figure plots, for every quarter k, coefficients $\tilde{\beta}_k^T$ from specification 2, separately by age group.

Figure 6: Has an open-ended contract (initial firm)



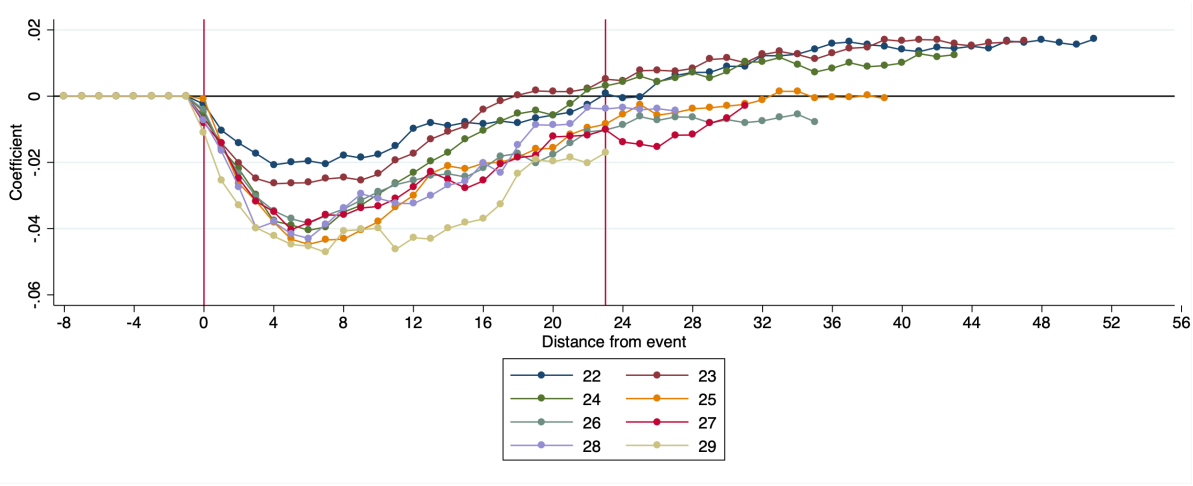
Note: The figure plots, for every quarter k, coefficients $\tilde{\beta}_k^T$ from specification 2, separately by age group.

Figure 7: Has an open ended contract (any firm)



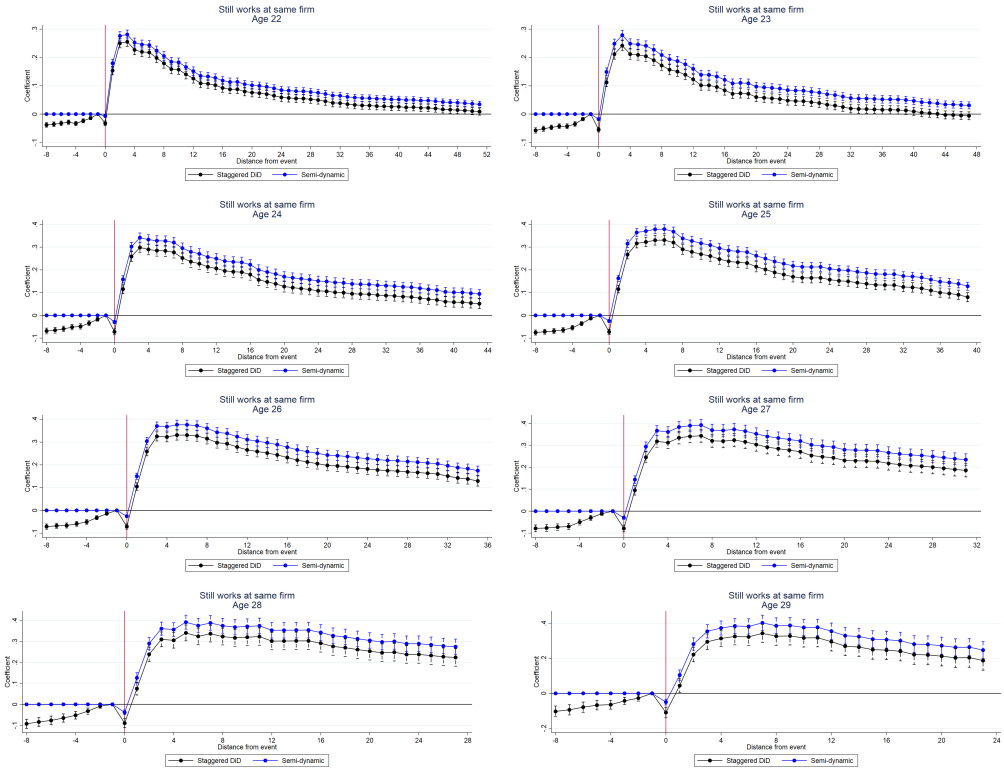
Note: The figure plots, for every quarter k, coefficients $\tilde{\beta}_k^T$ from specification 2, separately by age group.

Figure 8: Works as self-employed



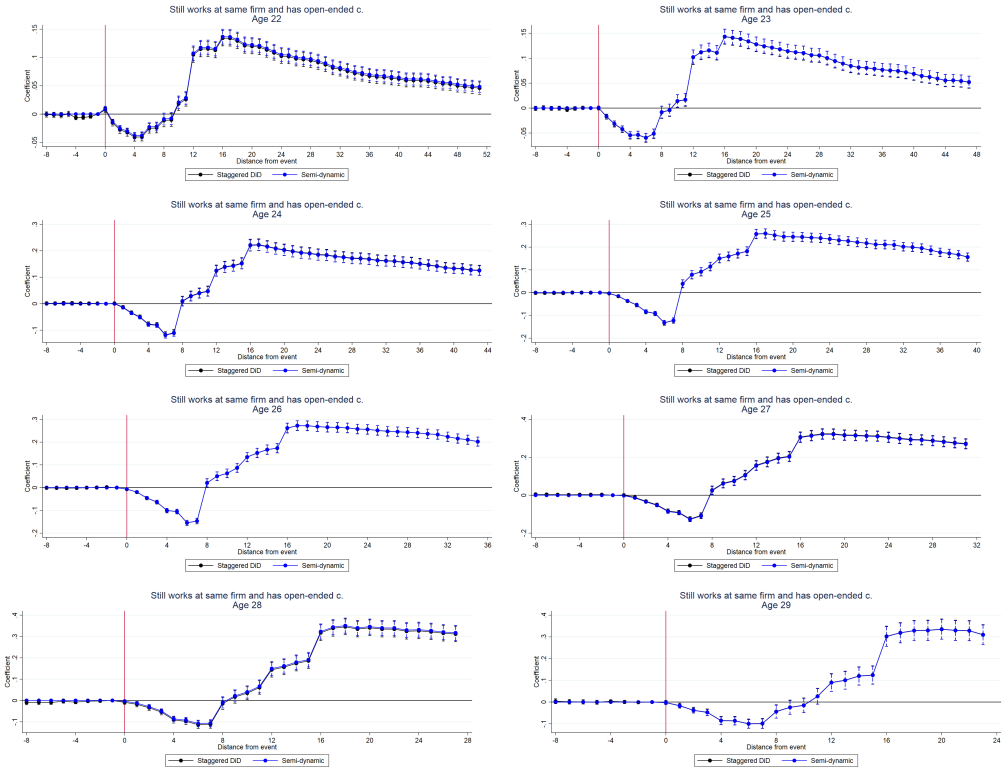
Note: The figure plots, for every quarter k, coefficients $\tilde{\beta}_k^T$ from specification 2, separately by age group.

Figure 9: Still works at initial firm - contract started in a big firm



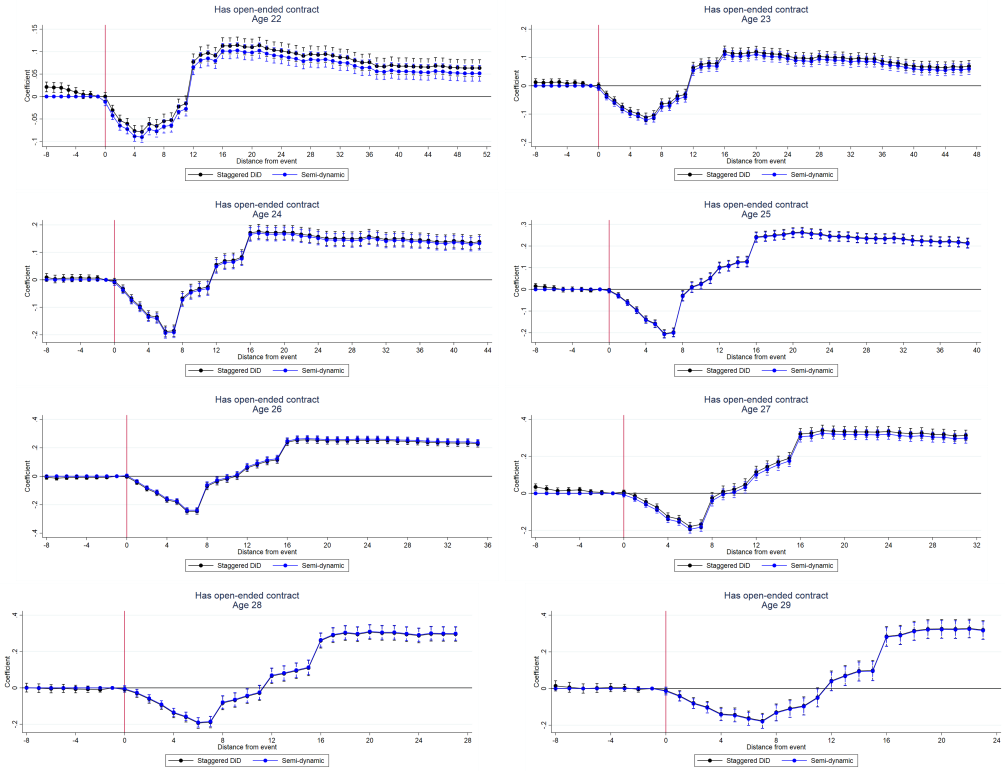
Note: The figure plots coefficients β_k^T (in black) coefficients $\tilde{\beta}_k^T$ (in blue) with associated 95% confidence intervals from specifications 1 and 2 respectively. Time is in quarters. Big firms are those in the upper half of the firm-size distribution in the year when the contract starts ($k = 0$).

Figure 10: Has an open-ended contract (initial firm) - contract started in a big firm



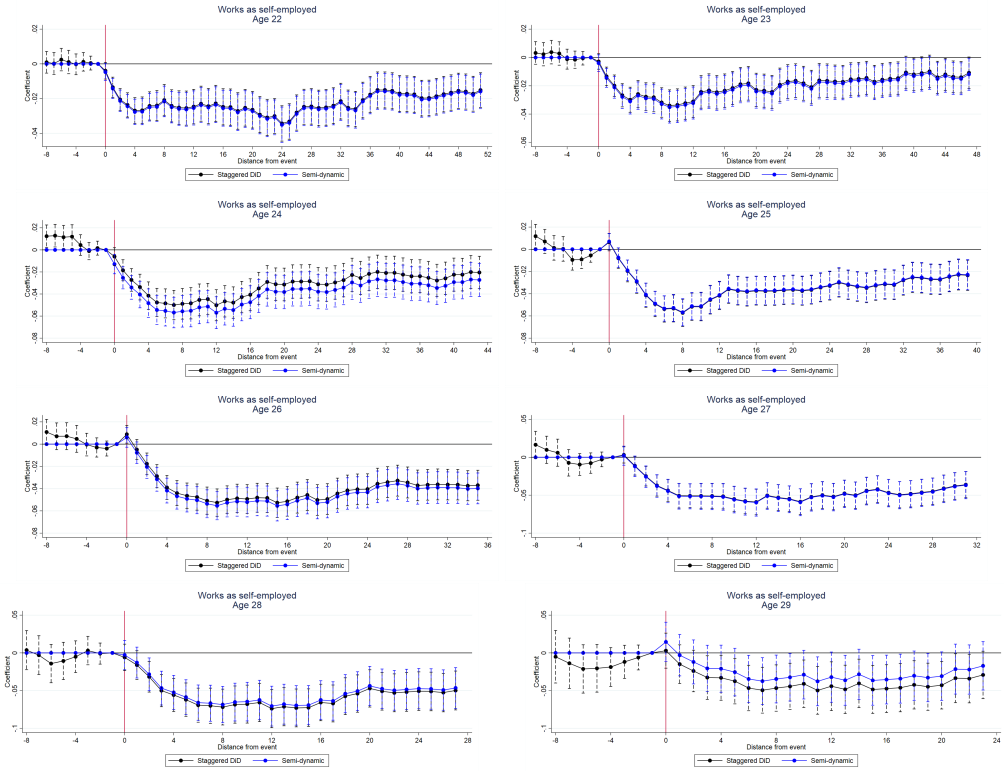
Note: The figure plots coefficients β_k^T (in black) coefficients $\tilde{\beta}_k^T$ (in blue) with associated 95% confidence intervals from specifications 1 and 2 respectively. Time is in quarters. Big firms are those in the upper half of the firm-size distribution in the year when the contract starts ($k = 0$).

Figure 11: Has an open-ended contract (any firm) - contract started in a big firm



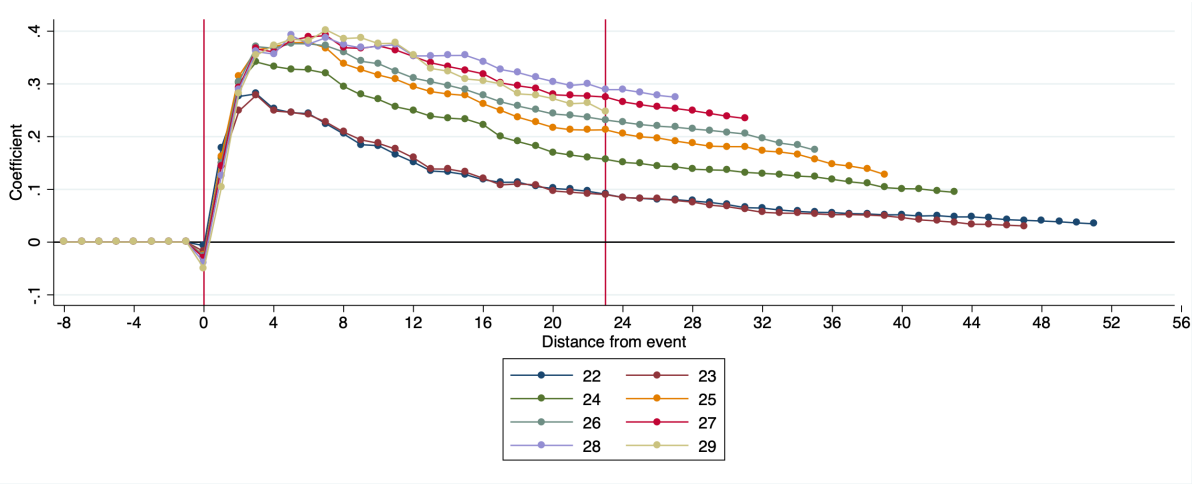
Note: The figure plots coefficients β_k^T (in black) coefficients $\tilde{\beta}_k^T$ (in blue) with associated 95% confidence intervals from specifications 1 and 2 respectively. Time is in quarters. Big firms are those in the upper half of the firm-size distribution in the year when the contract starts ($k = 0$).

Figure 12: Works as a self-employed - contract started in a big firm



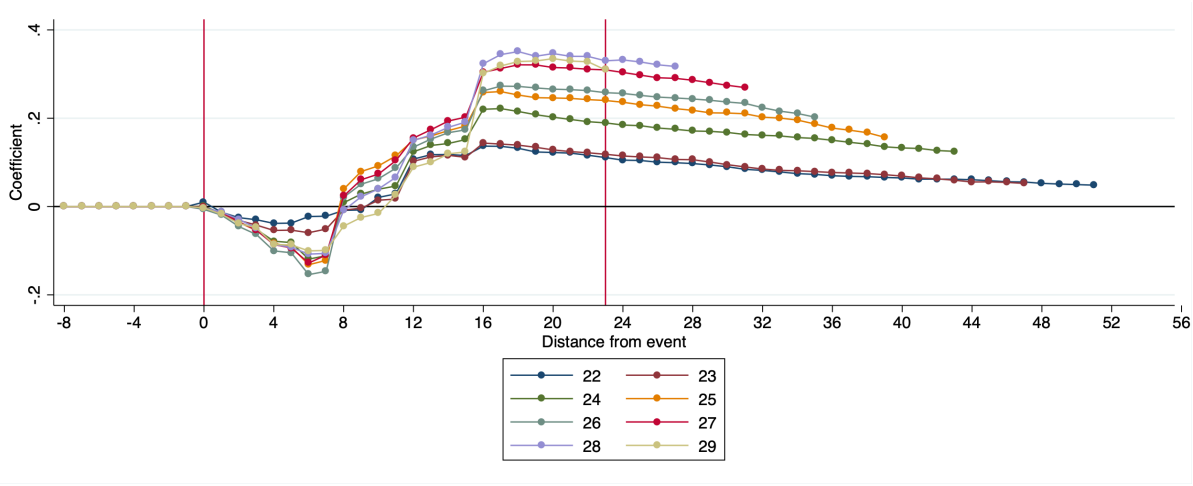
Note: The figure plots coefficients β_k^T (in black) coefficients $\tilde{\beta}_k^T$ (in blue) with associated 95% confidence intervals from specifications 1 and 2 respectively. Time is in quarters. Big firms are those in the upper half of the firm-size distribution in the year when the contract starts ($k = 0$).

Figure 13: Still works at initial firm - contract started in a big firm



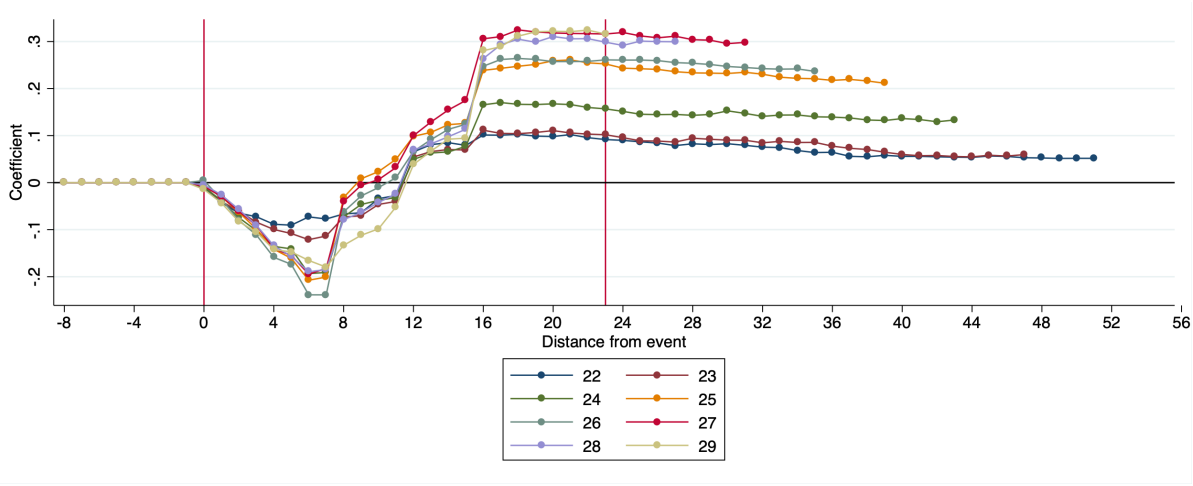
Note: The figure plots, for every quarter k, coefficients $\tilde{\beta}_k^T$ from specification 2, separately by age group. Big firms are those in the upper half of the firm-size distribution in the year when the contract starts ($k = 0$).

Figure 14: Has an open-ended contract (initial firm) - contract started in a big firm



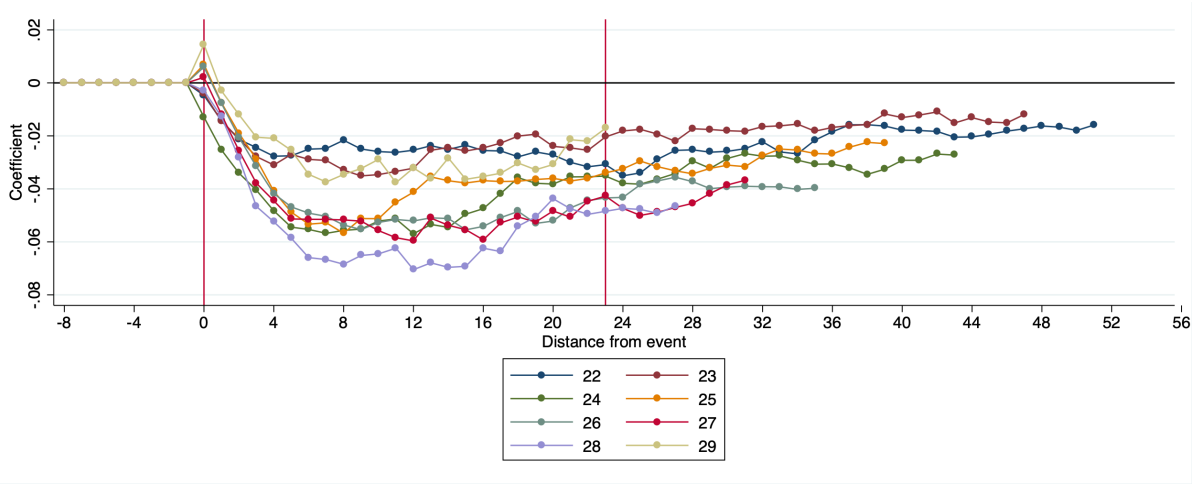
Note: The figure plots, for every quarter k, coefficients $\tilde{\beta}_k^T$ from specification 2, separately by age group. Big firms are those in the upper half of the firm-size distribution in the year when the contract starts ($k = 0$).

Figure 15: Has an open ended contract (any firm) - contract started in a big firm



Note: The figure plots, for every quarter k, coefficients $\tilde{\beta}_k^T$ from specification 2, separately by age group. Big firms are those in the upper half of the firm-size distribution in the year when the contract starts ($k = 0$).

Figure 16: Works as self-employed - contract started in a big firm



Note: The figure plots, for every quarter k, coefficients $\tilde{\beta}_k^T$ from specification 2, separately by age group. Big firms are those in the upper half of the firm-size distribution in the year when the contract starts ($k = 0$).

References

Daron Acemoglu and Jörn-Steffen Pischke. The structure of wages and investment in general training. *Journal of political economy*, 107(3):539–572, 1999.

Andrea Albanese, Lorenzo Cappellari, and Marco Leonardi. The Effects of Youth Labor Market Reforms: Evidence from Italian Apprenticeships. IZA Discussion Papers 10766, Institute for the Study of Labor (IZA), May 2017. URL https://ideas.repec.org/p/iza/izadps/dp10766.html.

Pierre Azoulay, Joshua S Graff Zivin, and Jialan Wang. Superstar extinction. *The Quarterly Journal of Economics*, 125(2):549–589, 2010.

Gary Becker. Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70, 1962. URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:70:y:1962:p:9.

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

Kirill Borusyak and Xavier Jaravel. Revisiting event study designs. 2016.

Josef Fersterer, Jörn-Steffen Pischke, and Rudolf Winter-Ebmer. Returns to apprenticeship training in austria: Evidence from failed firms. *Scandinavian journal of economics*, 110(4):733–753, 2008.

Stefano M Iacus, Gary King, and Giuseppe Porro. Causal inference without balance checking: Coarsened exact matching. *Political analysis*, 20(1):1–24, 2012.

Simon Jäger. How substitutable are workers? evidence from worker deaths. 2016.

Xavier Jaravel, Neviana Petkova, and Alex Bell. Team-specific capital and innovation. *American Economic Review*, 108(4-5):1034–73, 2018.

Matthias Parey. Vocational schooling versus apprenticeship training. evidence from vacancy data. 2016.

Matteo Picchio and Stefano Staffolani. Does apprenticeship improve job opportunities? a regression discontinuity approach. *Empirical Economics*, pages 1–38, 2013.