Scars of early non-employment in a rigid labour market

C. Ghirelli

Discussion Paper 2015-8

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

IRES

UCL

Scars of early non-employment in a rigid labour market

Corinna Ghirelli*
IRES - Université catholique de Louvain

April 23, 2015

Abstract

This paper investigates whether the early experience of non-employment has a causal impact on workers' subsequent career. The analysis is based on a sample of low educated youth graduating between 1994 and 2002 in Flanders (Belgium), i.e. a rigid labour market. To correct for selective incidence of non-employment, we instrument early non-employment by the provincial unemployment rate at graduation. Since the instrument is clustered at the province-graduation year level and the number of clusters is small, inference is based on wild bootstrap methods. We find that one percentage point increase in the proportion of time spent in non-employment during the first two and a half years of the career decreases six years after graduation annual earnings from salaried employment by 10% and annual hours worked by 7% (unconditional effects).

Keywords: youth unemployment, scars, instrumental variable, wild bootstrap

JEL Classification: J31, J64

*Address: IRES, Université Catholique de Louvain, Place Montesquieu 3, B1348, Louvain-la-Neuve, Belgium. Email: Corinna.Ghirelli@uclouvain.be. Corinna Ghirelli is also associated to SHERPPA, Ghent University.

1 Introduction

High levels of youth unemployment are a great concern for policy makers, especially since the start of the Great Recession (Bell and Blanchflower, 2010). However, unemployment rates are typically higher for young workers than for older ones. This is understandable, since younger workers have the least experience and hence are often the easiest to remove. Moreover, they lose or leave a job more often than older workers, because job shopping helps them to find a good match (Gervais et al., 2014). For instance, Topel and Ward (1992) find that two third of job changes and wage growth occur in the first ten years of workers' career. This initial high turnover may involve also short spells in unemployment. Therefore, youth unemployment need not to be necessarily detrimental to workers' career, if this is part of the process to find stable employment.[1] Other views predict that the experience of youth unemployment may entail long-term penalties in terms of reduced wages and persistent unemployment. These results are explained by human capital loss (Pissarides, 1992), which may arise from the depreciation of existing capital - if the worker is not subject to any kind of training when not employed - as well as from forgone work experience (Ellwood, 1982). An alternative explanation comes from the signaling model, in which past unemployment records are interpreted by employers as signals of low productivity in a context of imperfect information (Lockwood, 1991).[2]

Therefore, whether the early experience of non-employment entails long-term repercussions on youth's career should be assessed empirically. This study addresses this research question for Flanders, the most prosperous of three Belgian regions.

Overall, the existing literature suggests that the consequences of experiencing youth unemployment on workers' subsequent career are not just temporary. However, each of these studies tends to explore these effects each time on a specific outcome. For instance, according to a number of studies, school-leavers experiencing early unemployment suffer from strong state dependence (Gregg, 2001 for UK, Schmillen and Umkehrer, 2013 for Germany, Cockx and Picchio, 2013 for Belgium).[3] Besides, the experience of youth unemployment is found to inflict persistent penalties on earnings (Gartell, 2009 for Sweden and Gregory and Jukes, 2001 for UK) and wages (Gregg and Tominey, 2005 for UK and Mroz and Savage, 2006 for US).

In this study, we contribute to this literature by evaluating the impact of the early experience

[1] According to the literature on job displacement, the costs of job loss for young workers are smaller and less persistent than for mature workers (Kletzer and Fairlie, 2003). This is because young workers, unlike mature workers, experience steep earnings growth in subsequent work experience, after a job loss. At the same time, earnings growth of displaced young workers is below the levels of young non-displaced workers: i.e. most of the earnings loss stems from foregone wage growth due to displacement.

[2] Kroft et al. (2013) send fictitious resumes to real job postings in US and find that the likelihood of receiving callbacks for interviews significantly decreases with the length of the unemployment spells, mostly in the first 8 months. This is evidence that employer screening plays a primary role in generating duration dependence.

[3] Based on the same data Cockx and Picchio (2012) find positive state dependence of employment irrespectively of the duration, suggesting that job experiences are a stepping-stone to stable employment.

of non-employment on a range of labour market outcomes: hours worked and earnings for salaried public and private sector employment, as well as indicators of salaried and self-employment. This gives us a comprehensive view of the long-term consequences of experiencing youth non-employment on workers' subsequent career. We exploit very rich data combining survey with administrative data which allow us to observe the aforementioned outcomes for a representative sample of school-leavers graduating in Flanders in the 1994-2002 period up to the first twelve years after graduation.

Belgium is an interesting case because it has one of the most rigid labour markets among the OECD countries (Kawaguchi and Murao, 2014).[4] A related literature suggests that young graduates are more exposed to adverse labour market conditions at graduation in a rigid labour market than in a flexible one, and that this exposure translates into more persistent penalties (Genda et al., 2010; Kawaguchi and Murao, 2014). Thus, we expect that, in Belgium, the early experience of non-employment inflicts large penalties on youth's subsequent career.

This view is confirmed by two existing studies on the long-term impact of unemployment for Belgium. Gangji and Plasman (2007) study the adverse effect of the incidence of unemployment on re-entry wages considering a representative sample of Belgian workers aged 18-64 in the 1994-2002 period.[5] They find that the incidence of unemployment is associated with a 5.1% penalty in hourly wages. Of course, this is an average of heterogenous effects comprising workers with different ages, while the incidence of unemployment for youth is likely to have different consequences than for mature workers (Kletzer and Fairlie, 2003). Besides, Cockx and Picchio (2013) focus on the long-run effects of youth long-term unemployment. They consider all Belgian school-graduates aged 18-25 who in 1998 were registered to the National Employment Office and remained unemployed for at least 9 months,[6] and follow them until 2002. They find evidence of strong negative duration dependence in the job finding probability: further prolonging the unemployment spell by one year reduces the probability to find a job in the following 2 years from 60% to 16% for men. Note, these results apply to the specific sub-sample of long-term unemployed youth. Compared to these two studies, we provide new evidence on a representative sample of Flemish youth, i.e. a less restrictive sample than long-term unemployed youth (as in the latter study) but more specific than the overall working-age population (as in the former one). Even if not directly comparable, our results are consistent with this evidence.

Moreover, Belgian labour market institutions differ for white and blue collar workers and this creates different sources of rigidities for these two groups of workers. White collar workers are sheltered by a very strict employment protection legislation (EPL), which represents the

[4] Kawaguchi and Murao (2014) construct a composite index which allows to rank most of the OECD countries according to their labour market rigidity: Belgium is at the top of this ranking.

[5] Similarly, Arulampalam (2001) finds that the incidence of unemployment reduces persistently subsequent wages in UK, based on a sample of UK workers aged 18-64 in 1991 (followed until 1997).

[6] In Belgium, before 2012 school-graduates were eligible to UI benefits if still unemployed 9 months after the registration to the National Employment Service.

main source of rigidity for these workers. In contrast, blue collars have quite loose EPL,[7] but are supported by a short-time work compensation program (STC) that subsidies temporary reduction of labour force during downturns. This introduces rigidity in the labour market as it strongly restrains blue-collar workers' mobility, thereby having similar consequences as EPL for white collars. Besides, the sectoral minimum wages are among the highest of OECD. Thus, together with a quite generous unemployment insurance (UI) system,[8] minimum wages and the STC program are more relevant sources of rigidity for low skilled workers.

The main identification problem of our research question is the presence of unobserved individual characteristics that may affect labour market performance as well as the (selective) incidence of early non-employment, thereby introducing endogeneity. We address this problem by means of an Instrumental Variable (IV) approach, where the provincial unemployment rate at graduation is used as instrument for early non-employment. This methodology has already been exploited in the literature for the UK and Germany (Gregg, 2001; Gregg and Tominey, 2005; Schmillen and Umkehrer, 2013).[9] The estimation strategy shared by these studies is based on the idea that the variation in the labour market conditions at school-leaving age is exogenous to the individual, and therefore generates a variation in the early experience of unemployment that is unrelated to the unobserved factors that may influence both early and adult performances. In particular, the identification relies on the identifying assumption that the scarring effects of graduating in downturns occur *entirely* through the reduced work experience early in the career (exclusion restriction). This is a strong assumption, especially because it rules out the possibility that the aforementioned long-term scars stem from other channels, such as the acceptance of low-paid jobs just after graduation.

Since the Belgian institutional setting generates different sources of rigidities for the white and blue collar workers, we would like to study these groups separately, but the choice between these worker categories is clearly endogenously related to the labour market conditions and hence may induce selectivity. Thus we distinguish between "low educated" new graduates (with at least secondary education) and "high educated" ones (with higher education).[10] In doing this we exploit two features of the data: first, there is a clear correspondence between low educated and blue collar workers and between high educated and white collar workers: within the first 6

[7]Note that since the beginning of 2014 a single employment contract has been introduced in Belgium, stipulating the same EPL for white and blue collar workers.

[8]The generosity of the UI system is mostly due to the absence of time limit on the payment of UB (and loose job search requirements), as opposed to many other countries. UI covers also unemployed school-graduates without employment records. Until 2012, school graduates below age 26 were entitled to UI 9 months after registration at the Public Employment Service. Since 2012 the waiting period has been extended from 9 to 12 months.

[9]A related study is Neumark (2002), who identifies the causality of early job stability on adult wages in the US with an IV approach, where indicators of job stability in the first 5 years since graduation are instrumented by labor market conditions faced in this early period.

[10]This solution is also adopted in Cockx and Ghirelli (2015) which exploit the same data of this study.

years after graduation, 70% of low educated are prevalently employed as blue collars, while this figure corresponds only to 14% for the high educated.[11] Second, the moment of graduation - and hence the educational attainment - is found to be unrelated to the business cycle.[12]

An additional contribution is that, compared to the aforementioned studies exploiting this IV approach, we can make the identification strategy more credible by relying on new evidence by Cockx and Ghirelli (2015) on the scarring effect of graduating in recessions for Flanders, based the same data of this study. According to this research, graduating in downturns inflicts a scar on earnings for both the low and the high educated Flemish youth, but the former are penalized in terms of annual hours worked and not of hourly wages, while the reverse occurs for the high educated: that is, the high educated downgrade to lower-paying jobs whereas low educated workers remain unemployed more often. Therefore, the aforementioned scar on earnings occurs through different channels depending on the educational level, i.e. (i) the loss of early work experience for the low educated and (ii) the acceptance of lower-paying jobs for high educated (ii). That is, evidence (i) for the low educated runs in favour of the identifying assumption required to investigate the long-term penalties from the early experience of non-employment by means of the aforementioned IV approach.[13] As a consequence, we focus on the low educated because the IV estimation strategy is most credible for this group.

The IV strategy relies on the exogenous variation of the provincial unemployment rate at graduation to identify the long-term effects of early non-employment. Since the estimation includes province fixed effects, the identification exploits the time variation of the provincial unemployment rate series in the 1994-2002 period. Although this is quite a short period, this variation is sufficient to identify the long-term effects of early non-employment. Note, by including province fixed effects, we account for spacial sorting, that is the possibility that individuals sort into provinces according to unobservable characteristics. Besides, in the paper we discuss the issue of commuting. Moreover, since the unemployment rate varies across 9 graduation years and 5 Flemish provinces, inference hinges at most on 45 clusters. This raises the possibility of underestimating standard errors due to few clusters. We tackle this problem applying wild bootstrap[14] to the IV approach, following Davidson and MacKinnon (2010).

The paper is organized as follows. The next section describes the data. In Section 3 we explain the estimation strategy: we discuss the IV approach, including the way in which we deal with the problem of inference with few clusters. Section 4 discusses the results and presents some sensitivity analyses. Section 5 concludes.

[11] These figures are reported in Table 3 of Appendix A.

[12] This evidence is relegated to Section S.5 of the Supplementary Online appendix of Cockx and Ghirelli (2015): http://users.ugent.be/~bcockx/Ascars.pdf.

[13] Evidence (ii) could be exploited to study if, for the high educated, the average hourly wages earned right after graduation has long-run consequences on the subsequent career using the same IV approach. We did report this exercise as the instrument had not sufficient explanatory power. The exercise is available upon request.

[14] Wild bootstrap preserves the group structure of the data (Cameron et al., 2008).

2 Data

The analysis is based on the Sonar survey database, a representative sample of three birth-cohorts of Flemish youth - born in 1976, 1978 or 1980, which were interviewed at age 23, 26 and 29.[15] The surveys register retrospectively and on monthly basis the most important activity of the respondents, among which education. Based on this information, graduation is identified to occur in the first month that education has been interrupted for more than 4 months. The surveys also contain control variables for the analysis, which are measured at age 17, such that they are predetermined at graduation: father's and mother's education (years of completed education since age 12), the type of educational program (general, technical, vocational, part-time vocational or apprenticeship) in which the individual is enrolled at age 17, and the number of repeated grades at age 17 since secondary education. From this database we calculate for each individual the number of completed years of education, i.e. the number of grades successfully passed from the start of secondary education until graduation. Based on the latter we divide the sample in low and high educated, with the former having completed at most secondary education and the latter having a higher level of education.[16] We make this distinction throughout the analysis, because, as mentioned in the Introduction, the Belgian institutional setting entails different sources of rigidities for blue and white collar workers and there is a clear correspondence between low and high educated, and blue and white collar workers, respectively.[17].

The original Sonar sample contains about 9000 individuals, 3000 for each birth cohort. We restrict it as follows. We exclude few observations (0.17% of the sample) who dropped out from schooling before the end of compulsory education, which in Belgium is set at age 18. We focus on men since female labour supply is different from male labour supply due to mothering.[18] To increase the homogeneity of the sample, we drop individuals who attended special needs and arts education, who were not Belgian or did not speak Flemish at home, or who did not reside in Flanders at graduation. We retain individuals graduating from age 18 and 24, as students graduating with more than 24 years old are less than 5% of the sample. After eliminating, in addition, individuals with missing or inconsistent values in variables, we are left with a final sample of 3586 low and high educated male youth. From this, we focus on 1902 low educated youth, who graduated in the period 1994-2002. Descriptive statistics of the final sample is in Appendix A.

The survey data are matched to administrative data of Belgian Social Insurance institutions centralized at the Cross Roads Bank of Social Security, which give us access to high quality infor-

[15] For more details, see Sonar (2004a, 2003, 2004b).

[16] Low educated are those with at most 6 years of completed education (7 years if enrolled in vocational track at age 17). High educated are those with higher years of completed education.

[17] See Table 3 in Appendix A. As discussed in the Introduction, the educational attainment is exogenous in our data, as opposed to the worker category which may be endogenously related to the business cycle.

[18] Scars of early labour market outcomes for women are equally interesting and left for future research.

mation on individuals' labour market outcomes for a sufficiently long time span after graduation. In particular, these data report quarterly information on the registration as self-employed, as well as earnings and time worked in dependent employment (for both public and private sector), between year 1998 and 2010. For salaried workers we construct log annual earnings and log annual hours worked. The log-transformation allows us to interpret the coefficients of interest as semi-elasticities. Note that we retain in the analysis also non-salaried employed by adding value one to the continuous variables before taking the logarithmic transformation.[19] Hence, non-salaried employed have zero log-earnings and zero log-hours worked. As a consequence our estimates on continuous outcomes are unconditional on being salaried employed. This rules out the problem of selectivity due to restricting to the salaried employed, i.e. a potentially positively selected group. Yet it introduces another complication since the distribution of the outcome variables has a mass point at zero. The fraction of corner solutions is however quite small, as only 15% of the low educated are censored at zero at the moment of the evaluation. As a consequence, we will rely on OLS.[20] In addition, we construct three employment indicators: salaried employment, defined by positive earnings from salaried employment; self-employment, based on the registration as self-employed for at least one day during the calendar year; overall employment, which is the sum of self- and salaried employment. Note that salaried employed who are also registered as self-employed in the same calendar year are considered self-employed.[21] These outcomes are measured 6 years after graduation. This choice is due to the availability of the administrative data (1998-2010) and by the fact that we want to measure the dependent variables as late as possible for all graduation cohorts. Since the low educated graduate in the period 1994-2002, the last graduation cohort is followed until potential experience 8. Later than that the sample gets smaller as the last graduation cohorts progressively drop out from the sample. However, the sample gets slightly smaller as from potential experience 7 due to some outliers. To exploit a maximum sample size, we evaluate the outcomes at potential experience 6.[22] Table 6 in Appendix A shows descriptive statistics of the outcome variables.

The administrative data also provide additional control variables measured at age 17: living in single parent household, not living together with either parents and the number of other household members by age class. Descriptive statistics of the control variables are reported in

[19] In a log-level regression, the coefficient of interest from regressing X on $log(Y)$ is interpreted as the % change in Y for a unit increase in X (semi-elasticity). In our case, the coefficient of interest is only "approximatively" a semi-elasticity, given that the dependent is $log(Y + 1)$. However, we verified that this interpretation is a good approximation (the exercise is available upon request).

[20] OLS estimation provides approximations of the unconditional effects, as it does not take into account the corner solutions at zero. In principle, Tobit models would be more appropriate.

[21] We choose this since the self-employed are very few relative to the salaried employed (83% of the sample are salaried employed, 7% are self-employed and 5% are both salaried and self-employed in the same year).

[22] Results are very similar when the outcomes are measured at potential experience 8 (despite the instrument becomes slightly weaker because of the smaller sample size). This exercise is available upon request.

Table 5 of Appendix A.[23] Finally, the administrative data give us access to yearly information on the province of residence between the year in which individuals turn age 18 and 2010.

From the year of graduation onwards, we associate each calendar year to a *potential* year of labour market experience,[24] which corresponds to zero in the year of graduation. Potential experience 0 lasts from the month subsequent to graduation until December of that calendar year. Therefore, its length (measured in months) is computed as $12 - month\_of\_graduation$: for a June graduate - which amounts to 90% of the sample - it lasts 6 months. All subsequent years of potential experience have a duration of 12 months. Our regressor of interest is a measure of the time spent in non-employment at potential experience 0-2, relative to the potential total hours if one would work full-time during the whole period. We express it as a proportion in order to take into account the fact that the reference period changes depending on the month of graduation, thereby ensuring that early non-employment is comparable across individuals. On average this period corresponds to 2.5 years after graduation (30 months), as 90% of the sample graduates in June. For simplicity hereafter we will refer to this reference period as to its average, i.e. 2.5 years after graduation.

This endogenous regressor can be measured precisely, by exploiting the administrative data on hours worked in salaried employment. However, the latter are available only since 1998, while students graduate since 1994 in our sample. Therefore, we base this variable on the administrative data for the individuals graduating since 1998 (68.5% of the sample), and exploit information from Sonar (survey data) whenever potential experience 0-2 occurs before 1998 (31.5% of sample). The reason why we combine administrative with survey data is to maximize the sample of study, thereby exploiting the variation of the instrument for the entire graduation period 1994-2002, rather than for the restricted period 1998-2002. Of course, the disadvantage is that the data on time worked in the Sonar database are less precise and hence we have to make some assumptions to convert this information into hours worked:[25] this certainly introduces measurement error in the endogenous regressor.

Briefly, the endogenous regressor is constructed as follows (for details, see Appendix B): first, define the reference period as the entire calendar year for potential experience 1 and 2, and the part of calendar year following the month of graduation for potential experience 0; sum up all hours worked including self-employment in the reference period (a);[26] compute the potential total hours if one would work full-time during the whole period (b); express early non-employment as $100 * (b - a)/b$. As already mentioned, the denominator takes into account the fact that the reference period changes depending on the month of graduation and ensures

[23] For details on the construction of the control and outcome variables, see Section S.1 of the Supplementary Online Appendix of Cockx and Ghirelli (2015): http://users.ugent.be/~bcockx/Ascars.pdf.

[24] This terminology is borrowed from the literature. "Potential" underlines that the variable counts all calendar years since graduation, as opposed to *actual* experience which endogenously considers only years of employment.

[25] These are imputed from monthly employment indicators assuming full-time employment (see Appendix B).

[26] For self-employed we assume the working regime of full-time salaried employed (see Appendix B).

comparability across individuals. This variable measures the intensity of early non-employment. It is equal to zero if in the reference period one has always worked as much as a full-time salaried employed, and above zero if one has worked less intensively than the full-time regime or if one has not worked for some time. Given the possibility of measurement error arising from the combination of survey and administrative data, Section 4.1 performs a sensitivity analysis for the restricted graduation period 1998-2002, where only administrative data are exploited to measure the endogenous regressor.

A final source of information is the Labour Force Survey (LFS), which provides long time series of the provincial unemployment rates for Flanders. In our analysis, we use the provincial unemployment rate (15-64) at graduation as instrument for early non-employment. Figure 2 in Appendix C plots this series from year 1993 until 2011. Note that the literature typically exploits more disaggregated unemployment rate series.[27] For Belgium, provincial unemployment rates are the most disaggregated data available for the period considered. The main drawback is that inference relies on too few clusters, as the identification of the effects of interest comes from the variation of the unemployment rates by provinces and years. To the extent that we tackle this problem with wild bootstrap (see Section 3.4), provincial data are not much of a limitation. In contrast, more aggregated series provide the advantage of reducing the problem of endogenous migration, that would arise if new graduates offset the long-term effects of early non-employment by moving or commuting into provinces where there are more job opportunities. Our data suggest that in Flanders less than 2% of individuals change province of residence in the 1998-2010 period. However, as Flanders is a relatively small region, people could commute to work across provinces. In this case, we would underestimate long-term effects of early non-employment.[28] However, the magnitude of the inter-provincial variation in the unemployment rate reported in Figure 2 demonstrates that mobility is limited and far from eliminates all inter-provincial variation. This is because the LFS series are based on the province of residence and not of job location: thus, if workers commute to avoid the adverse local labour market conditions, this evens out the provincial variation in the unemployment rate.

3 Estimation Strategy

We are interested in the causal relationship of the early experience of non-employment, say Y^0, on subsequent labour market outcomes of interest Y for the low educated. Namely, we want to estimate an equation of the following type, where X is a vector of control variables that will be

[27] At district level (Schmillen and Umkehrer, 2013) and by wards (Gregg, 2001; Gregg and Tominey, 2005).

[28] If anything, this should be more worrying for the high educated, since they are (1) less liquidity constrained because of high expected wages or better working conditions and (2) more mobile due to higher motivation to find jobs that meet their expectations about wages/job profiles.

defined below and ϵ is an idiosyncratic error term:

$$Y = aY^0 + bX + e \text{ with } e = (\theta + \epsilon) \tag{1}$$

The main identification problem is the presence of some factors θ, unobserved to the researcher, that may affect both early non-employment and subsequent labour market performances, thereby introducing endogeneity. Therefore, OLS estimates will be biased as a consequence of these omitted factors. We remove this bias by means of a two-stage least squares (2SLS) estimator, where the provincial unemployment rate at graduation is used as instrument (Z) for early non-employment.[29] In practice, the identification strategy relies on the variation of the provincial unemployment rate at graduation Z, which is exploited to generate an *exogenous* variation in early non-employment Y^0, which is then used to identify the causal relation of interest. In accordance with the traditional IV approach, we assume that the effect of interest is homogeneous.[30] In this framework, the 2SLS identifies the causal effect of interest under two conditions:

1. Z is uncorrelated with e. This implies that Z does not *directly* affect the outcome Y (exogeneity), and that any *indirect* effect of Z on Y occurs uniquely through the endogenous regressor Y^0 (exclusion restriction). This is an identifying assumption.

2. Z is correlated with Y^0, conditional on the controls X in (1) (strength). This condition can be tested by means of the F statistic of the excluded instrument in the first stage regression.

Note that in this framework the IV estimate refers to the entire population since the causal effect of interest is assumed to be homogeneous across individuals. The next section discusses in detail the identifying assumption 1. In particular, we will carefully examine which factors may violate the exclusion restriction and define the specification in such a way that the latter is most likely satisfied conditional on the covariates.

3.1 The Instrumental Variable Approach: identifying assumptions

Together, Conditions 1 and 2 above require that the instrument explains the endogenous regressor while being exogenous in Eq. (1). This has the following implications.

First, it amounts to rule out reverse causality between Z and Y^0, that is the unemployment rate at graduation affects early non-employment but not the other way around. If it were the case, Z would be not exogenous to Eq. (1). We exclude the possibility of reverse causality since

[29] A similar approach has been used by Neumark (2002), Schmillen and Umkehrer (2013), Gregg (2001) and Gregg and Tominey (2005).

[30] This is of course a restrictive assumption. Under heterogeneous effects, the IV estimator identifies a weighted average of local average treatment effects (LATEs) but this interpretation requires additional assumptions, Stable Unit Treatment Value Assumption (SUTVA) and Monotonicity, which are however very difficult to hold in this case since they rule out crowding-out effects in the labour market.

the instrument and the endogenous regressor are measured at the provincial and individual level, respectively, and an aggregate variable cannot be caused by an individual variable.

Second, the exogeneity assumption in Condition 1 requires that the unemployment rate Z cannot affect the unobserved composition of new graduates by year and province. If this were the case, the relation between the instrument and early non-employment would spuriously reflect changes in the composition of graduates rather than causality, which would introduce selectivity. To rule this out, one has to assume that students choose the moment of graduation independently of the business cycle (exogeneity of the timing of graduation), and that before graduation they do not move to provinces where the unemployment rate is lower relatively to others (exogeneity of place). We test the former condition in Section S.5 of the Supplementary Online Appendix of Cockx and Ghirelli (2015), and demonstrate that the duration between the end of compulsory education at the age of 18 and each year of potential graduation is unrelated to the provincial unemployment rate in those years.[31] As for mobility, almost nobody (0.44%) changes residence between the first year that our data inform about the place of living, i.e. on December 31 of the year in which the individual turns 17, and the year of graduation. Therefore, the issue can be safely ignored.[32]. On this basis, we argue that in our sample the choice of graduation is independent from the labour market conditions.

Third, the exclusion restriction in Condition 1 requires that the instrument is not correlated with any of the omitted factors in Eq. (1). This implies assuming that the scars of graduating in downturns for the low educated are determined *exclusively* by the proportion of time spent in early non-employment. This assumption is consistent with the evidence in Cockx and Ghirelli (2015), who find that the adverse labour market conditions at graduation inflict big initial penalties on earnings and hours worked of the low educated, and that these penalties fade away slowly. That is, the low educated who graduate in downturns experience longer periods of non-unemployment at the start of the career, and this has repercussions in the long-term. This persistence is exacerbated by the rigidities of the Belgian labour market. Such rigidities are due to the following institutions: the STC program which, by tying the employers to the employees in hard times, increases the expected hiring costs of blue collar workers during recession;[33] and

[31]In a discrete duration model, an indicator of graduating since age 17 is regressed on birth cohort dummies, individual characteristics and the province of living measured at age 17, the elapsed duration in education since age 17, and the unemployment rate in each potential year of graduation (interacted with the elapsed duration), testing whether the coefficients of latter interactions are jointly significantly different from zero. The test deals with selectivity induced by unobserved heterogeneity. It uses the same sample as this study. For details, see Section S.5 in http://users.ugent.be/~bcockx/Ascars.pdf.

[32]In principle, students may enroll in universities located in provinces where they expect to find more jobs in the future (endogenous commuting). However, this is not an issue, as the unemployment rate at graduation is measured in the province of residence and not in the province in which the university is located.

[33]Note that this scar may be nuanced by the extensive use of STC for blue collars: they will experience long periods of unemployment or reduced activity, but they are more likely to be called back.

the asymmetry in EPL - flexible for blue collar workers while rigid for white collar ones - which pushes the high educated to downgrade and hence increases the competition for low skilled positions.[34] The absence of a similar impact on wages is due to the presence of high minimum wages, which are likely to be binding for the low educated at the start of the career. This suggests that accepting lower paying jobs is not a relevant channel to explain the long-term penalties on earnings for the low educated. Thus, based on the aforementioned evidence, we argue that early non-employment is the relevant channel to explain long-term effects of adverse labour market conditions at graduation for the low educated.

Of course things can be a bit more blurry if we consider a wider definition of reservation wage which incorporates also the future wage growth linked with seniority in addition to the current wage. In this case, low educated graduating in a downturn may not only experience higher early non-employment, but could also accept lower-quality jobs, i.e. with less steep wage profile than the jobs accessed during a tight labour market. The unemployment rate at graduation would then entail a growing negative impact on subsequent wages as a consequence of accepting this initial job, and this would represent a violation of the exclusion restriction when wages are the outcome of interest. This possibility is discussed in Cockx and Ghirelli (2015), in which the unemployment rate at graduation is shown to have a negative impact on wages of the low educated starting from potential experience 6.[35] This evidence is compatible with the aforementioned hypothesis and hence represents a violation of the exclusion restriction when wages are evaluated. Thus, we do not consider subsequent wages as an outcome variable. In contrast, we restrict the analysis to hours worked and earnings, since the long-term penalties on the these outcomes are compatible with the idea that early non-employment is the main driver of these scars.

However, other channels may as well contribute to explain the long-term penalties of labour market conditions at graduation: these channels would invalidate the exclusion restriction if not included in the specification. An example is the persistence of the unemployment rate series. If the current unemployment rate affects the outcomes, the correlation between the unemployment rate at graduation and the current unemployment rate violates the assumption that the instrument affects the outcomes only through early non-employment. To prevent that, it is important to additionally control for the current unemployment rate, as typically done in the literature. However, this may not be sufficient, as in principle one should control for all unemployment rate series up to the moment of evaluation (Oreopoulos et al., 2012, 2008). To keep a parsimonious specification, we add the average unemployment rate between the end of

[34] The high educated graduating in downturns accept lower-paying jobs due to the high barriers to enter white collar positions. They may find it difficult to upgrade to a position matching their initial aspirations due to the (useless) accumulation of human capital specific to lower-paying jobs (Cockx and Ghirelli, 2015).

[35] In the benchmark model this growing negative impact is not statistically significant, but it is significant in the sensitivity analysis (see Table S.21 of the Supplementary Online Appendix of Cockx and Ghirelli, 2015).

the early period and the moment of evaluation - between potential experience 3 and 6.[36]

More generally, the problem of the persistence of the unemployment rates refers to the literature on wage determination, which investigates how the sequence of labour market conditions experienced by a worker affects current wages (Beaudry and DiNardo, 1991). According to this view, labour markets operate as spot markets if current wages are affected by current unemployment rates and not by past ones. In contrast, wages result from long-term implicit contracts if past unemployment rates explain current wages despite current ones: with costless mobility, the minimum unemployment rate since hiring should matter the most, as workers are able to renegotiate the wage once better labour market conditions arise; if instead mobility is costly, the unemployment rate at hiring should be the relevant one. Beaudry and DiNardo (1991) found that, once the minimum unemployment rate since hiring is included together with the unemployment rate at hiring, the former but not the latter significantly explains current wages. This is consistent with the idea that wages are negotiated according to long-term implicit contracts with renewals.[37] In this case, the exclusion restriction may be violated if the unemployment rate at graduation mistakenly picks up the effect of the minimum unemployment rate since hiring, because of the persistence of the unemployment rates. To prevent that, we include the minimum unemployment rate since graduation in the specification.[38]

Finally, other violations of the exclusion restriction may be due to, for instance, differences in institutions that could be correlated both with the unemployment rate at graduation and with the outcomes. We therefore include province fixed effects to ensure that permanent differences across provinces violate the exclusion restriction. Note, this also controls for spacial sorting, i.e. the possibility that individuals sort into different provinces based on unobservables. Similarly, we include province-specific time trends to capture whatever time-varying provincial heterogeneity, such as changes in legislations, that may be correlated with the instrument and the labour market outcomes at the moment of evaluation. In the next section we present the equation of interest in light of all these arguments.

3.2 The Equation of Interest

To avoid clutter, we state the following definitions: t is the observation period, which runs from graduation until the moment of evaluation of the outcomes of interest T, i.e. 6 years after graduation; t_0 is the time of measurement of predetermined individual controls, which

[36]To rule out multicollinearity, we run a sensitivity analysis only including current unemployment rate (see Table 9 in Appendix F). The stability of the results ensures that multicollinearity is not driving the results.

[37]Recently Hagedorn and Manovskii (2013) criticize this interpretation. They argue that wages are still determined by spot markets and not by long-term implicit contracts. They show that, once the current match quality is taken into account, past labour market conditions no longer play a role in the wage determination.

[38]Note that the hypothesis of long-term implicit contracts seems more likely for the high educated, for instance to capture returns in human capital accumulation. In contrast, recent evidence has shown that labour markets operate like spot markets for the low educated (Kilponen and Santavirta, 2010; Devereux and Hart, 2007).

corresponds to the year in which individuals are aged 17; t_1 is the time window in which we measure early non-employment, i.e. on average the first 2.5 years after graduation[39]. We estimate the following equation, where subscript i indicates the individual, g the graduation year and p the province of residence at graduation:

$$y_{igpT} = \alpha + \beta y^0_{igpt_1} + \gamma_1 UR_{pT} + \gamma_2 \overline{UR}_p + x'_{it_0}\delta + \zeta minUR_{pt} + \eta_p + \omega_p T + f(g) + e_{igpT}$$
$$\text{with } e_{igpT} = \theta_i + \epsilon_{ipgT} \tag{2}$$

- y_{igpT} represents the following outcomes of interest measured in T, i.e. 6 years after graduation: three indicators of salaried, self- and overall employment, as well as log hours worked and log earnings in salaried employment. Before taking the logarithm of continuous variables we add value one, so that non-salaried employed at the moment of evaluation are included with value of zero after the logarithmic transformation. Therefore, the effects on continuous outcomes are unconditional on being salaried employed. The reason why we do this is twofold: first, the instrument is not strong enough to estimate effects conditional on salaried employment, but it becomes relevant when non-salaried employed are also included in the sample.[40] Second, unconditional effects refer to the entire population of workers and avoid the problem of selectivity when focusing on the sub-population of salaried employed. We take the log of continuous outcomes to interpret the estimates as semi-elasticities.[41]

- $y^0_{igpt_1}$ is the endogenous regressor representing early non-employment: it is expressed as the percentage of time spent in non-employment in period t_1, relative to potential total hours if one would work full-time during the whole period.

- UR_{pT} is the current unemployment rate in the province of graduation, i.e. measured 6 years after graduation. It ensures that the exclusion restriction is not violated by the correlation between current local labour market conditions and local labour market conditions at graduation.

- $\overline{UR}_p$ is the time average of the unemployment rate in the period subsequent to the measurement of early non-employment, i.e. from potential experience 3 to 6. Together with UR_{pT} , it controls for the persistence of the unemployment rate series.

- $minUR_{pt}$ is the minimum unemployment rate in the province of residence of graduation over the entire period t. It controls for the possibility that wages are determined by

[39] This time window corresponds to potential experience 0-2, i.e. from the month after graduation until December of the second subsequent calendar year.

[40] In the first stage, the F statistic is about 4 in the conditional case and reaches 9 in the unconditional one.

[41] Despite the dependent is $log(Y+1)$ rather than $log(Y)$, we verified that interpreting the coefficient of interest as semi-elasticity is a good approximation. The exercise is available upon request.

long-term contracts which are renegotiated by the workers during upturns. Under this assumption, the persistence of the unemployment rate series (i.e. the correlation between $minUR_{pt}$ and the instrument) and the correlation between the outcomes and the minimum unemployment rate could violate the exclusion restriction.

- x_{it_0} is a set of individual control variables, predetermined since measured in t_0: birth cohort dummies, family composition, parental education, repeated years since secondary education as well as the educational track at age 17.

- η_p are fixed effects for the province of living at graduation: it controls for all differences across provinces that are constant over time, e.g. differences in institutions, or in the structure of the economy. Note, this controls for spatial sorting, i.e. the possibility that individuals with specific unobservables settle in specific provinces.

- $\omega_p T$ are provincial specific linear time trends, included because the unemployment rates exhibit differential downward time trends (see Figure 2 in Appendix C). More generally, it controls for any time-varying provincial heterogeneity, for instance for changes in the legislation or in the structure of the economy at the provincial level.

- $f(g)$ is a linear spline in the graduation year, which controls for aggregate shocks affecting all provinces over the graduation period. We impose a piece-wise linear specification because graduation year fixed effects absorb too much variation and as a consequence the instrument becomes weaker in the first stage. The spline is formulated as $f(g) = \alpha + \sum_{j=0}^{2} \beta_j.(g - 3j)\mathbf{1}[g \geq 3j]$ with $g = 1, .., 9$.

- ϵ_{igpT} is an i.i.d. error term, while θ_i represents unobserved individual factors correlated with $y^0_{igpt_1}$, thereby introducing endogeneity.

β is the coefficient of interest which represents the effect of one percentage point increase in the proportion of time spent in early non-employment on subsequent outcomes of interest (employment rates, hours worked and earnings) for the low educated: in presence of scarring we expect a negative β. The OLS estimate of β is biased due to the correlation between θ_i and $y^0_{igpt_1}$.

For all dependent variables, we estimate (2) by OLS or 2SLS. Thus, we estimate linear probability models for discrete labour market outcomes. For continuous variables we report unconditional OLS effects and hence do not take into account that these outcomes are left-censored at zero. However, we believe that this is a minor issue, since the fraction of corner solutions in the sample is quite small: only 15% of the low educated are not salaried employed at potential experience 6.[42] In any case, we provide heteroscedastic-robust standard errors in

[42] In Schmillen and Umkehrer (2013) the dependent variable, the number of days spent in unemployment in prime age, is censored at zero for almost 60% of cases. They use Tobit models.

all estimations to account for the fact that the dependent variables are dichotomous or censored at zero.

3.3 The Bias and Its Direction

The aforementioned bias can go in both directions. The latter depends on the sign of the relationship between the omitted factors and the outcome of interest and of the covariance between the omitted factor and early non-employment. Below we discuss four possible sources of bias and their corresponding sign.

- Ability and Motivation: everything else equal, more able and motivated individuals are more likely to perform well in the labour market at any point in time: therefore these factors are negatively correlated with early non-employment and positively correlated with the outcomes of interest. The overall bias is negative, so that OLS overestimate the (negative) scarring effect of early non-employment.

- Returns to job search: heterogeneous returns may arise because of differences in the search intensity or in the methods of search chosen. Ceteris paribus, individuals with higher returns search more and more successfully (also on-the-job), and therefore perform better in the labour market. Hence, the outcomes of interest are positively correlated with returns to search. At the same time, in the first phase of job shopping they may alternate jobs with short spells in non-employment, if they find it optimal to consume leisure when young and their opportunity cost is lower.[43] This may generate a positive correlation between returns to job search and early non-employment (Neumark, 2002). Under these assumptions, the bias is positive and OLS underestimate scarring.

- Liquidity constraints: individuals with high liquidity constraints have a low reservation wage. Thus, everything else equal, we expect liquidity constraints to be negatively correlated with early non-employment. At the same time, these individuals are likely to accept low quality jobs because of their low reservation wage, which is likely to translate into worse labour market performances over time. We therefore expect also a negative correlation with the outcomes of interest. The resulting bias is positive so that OLS underestimate scarring.

- Measurement error: as explained in Section 2, we introduce measurement error in the construction of early non-employment, because we use information from the Sonar database to impute hours worked in the first 2.5 years since graduation for students graduating before 1998, which are not observed in the administrative data. Measurement error in

[43] Neumark (2002) justifies this assumption as follows: in a standard life cycle utility-maximization model individuals are more likely to consume leisure at the point in the life cycle when their wages are low.

the endogenous regressor reduces OLS estimates towards zero (Hausman, 2001), thereby underestimating the scarring effect of early non-employment.[44]

To recapitulate, we expect OLS to overestimate the negative effect of early non-employment on the outcomes of interest, if the bias comes from ability. In contrast, the OLS estimate will be underestimated if the bias is due to returns to job search, measurement errors in the endogenous regressor or liquidity constraints. The literature is in favour of the latter hypothesis (Neumark, 2002; Schmillen and Umkehrer, 2013; Gregg and Tominey, 2005; Gregg, 2001).

3.4 Inference

It is well known that standard errors are underestimated in a micro-level regression with grouped covariates because it is assumed that each observation is independent of all others while the information of the grouped covariates is repeated within each cluster. Thus, correct inference requires to take into account that the independent information of the grouped covariates is at the group level by using cluster robust standard errors (Moulton, 1990; Angrist and Pischke, 2008, ch.8). In our 2SLS we use a grouped variable, the unemployment rate at graduation, as instrument for early non-employment that varies at the individual level. Therefore, the identification of causality comes from the time variation of the provincial unemployment rate at graduation, which is exploited to construct the fitted values of the first stage.[45]

The clustered estimator is consistent provided that the number of clusters is large enough, as consistency is determined by the law of large numbers. This is because, given the grouped structure of the data, the relevant unit are clusters and not observations. Since we consider the low educated graduating in the 5 Flemish provinces in the period 1994-2002, inference hinges on 44 clusters.[46] This raises the possibility of underestimating standard errors due to few clusters. Empiricists tend to agree that 50 clusters is enough when clusters have roughly the same size, but that a higher number of clusters is required when clusters are unbalanced (Cameron et al., 2008; MacKinnon and Webb, 2015). Applying the clustered estimator when clusters are too few is likely to worsen the bias, with cluster robust standard errors being even smaller than conventional ones. This is what we find by comparing conventional and cluster robust standard errors of 2SLS estimations (see Table 1), which suggests that we have too few clusters.

We tackle this problem with wild restricted efficient residual bootstrap (WRE bootstrap) proposed by Davidson and MacKinnon (2010), which are designed for 2SLS in context of het-

[44] In linear models, OLS estimates of y^0 are underestimated due to measurement errors and the bias can be eliminated with IV. However, this does not holds for non-linear models (Amemiya, 1985; Hausman et al., 1995).

[45] In 2SLS, the bias of the conventional variance estimator with grouped data is determined by the intra-class correlation of the second stage residuals (ρ_e) and by the intra-class correlation of the first stage fitted values (ρ_x). ρ_x is highest with grouped regressors in the first stage. As for OLS, $\rho_e > 0$ does not matter for standard errors as long as ρ_x is zero, but also a small ρ_e can give important bias with $\rho_x > 0$ (Angrist and Pischke, 2008, ch.8).

[46] Clusters are 44 since $g2002p3$ is empty. Table 4 in Appendix A shows the distribution across clusters.

eroscedasticity or clustered data. This procedure combines the restricted efficient residual bootstrap (RE bootstrap) proposed by Davidson and MacKinnon (2008) for the 2SLS, with wild bootstrap, which allows for intra-cluster correlation and heteroscedasticity (Cameron et al., 2008). For completeness, we apply wild bootstrap also to the t statistic of the instrument in the first stage, as well as to the t statistic of the regressor of interest when estimating (2) by OLS.[47] The bootstrap procedures are explained in detail in Appendix D.

Because of few clusters, also the F statistic of the first stage is overestimated. To adjust it, we exploit the fact that in case of one instrument the F statistic is the square of the t statistic of the instrument in the first stage: i.e. with G clusters, $F(1, G-1) = t^2(G-1)$. Therefore, the bootstrap F statistic is the critical value of the $F(1, G-1)$ distribution that corresponds to the bootstrap P-value of the t statistic of the instrument in the first stage.[48]

4 Results

Table 1 summarizes the results for the low educated from estimating Eq. (2) by OLS and 2SLS on alternative labour market outcomes, measured 6 years after graduation. As a matter of space, we report only the effects of interest, i.e. the effect of early non-employment in the structural equation (β in Eq. (2)) and the impact of the instrument in the first stage regression. The complete regressions are reported in Appendix E. Odds and even columns show heteroscedastic-robust and cluster robust standard errors, respectively. The former take into account that the dependent variables are dichotomous or censored at zero, while the latter allow for intra-cluster correlation induced by the fact that the instrument varies at the *gp* level. The fact that the 2SLS cluster robust standard errors are smaller than the 2SLS heteroscedastic-robust ones (columns 3 and 4 in Panel A) suggests that clustering is ineffective because of too few clusters. We ensure to make correct inference by bootstrapping the t statistic of the effects of interest and by reporting the corresponding P-value.

Panel B summarizes the results of the first stage regression. One percentage point (*pp*) increase in the unemployment rate at graduation increases early non-employment by 5 *pp*. For Flanders, the unemployment rate rises on average by 1.4 *pp* in the 1994-2010 period (and by 1.6 *pp* in the Great Recession in 2008). Thus, graduating in an average downturn increases the proportion of hours spent in non-employment early in the career by about 7% (1.4×5). We

[47] In the first stage the instrument is grouped, hence we need to cluster. In contrast, when estimating Eq. (2) by OLS the regressor of interest varies at the individual level: hence, clustering is not a major issue. For completeness we provide both heteroscedastic-robust and cluster-robust standard errors.

[48] We are aware of only one study by Baltagi et al. (2013) on the performance of wild bootstrap applied to the F statistic in context of heteroscedastic - but not clustered - data. Bootstrapping directly the F test in our wild bootstrap procedure did not always yield the expected results (sometimes, the bootstrap P-value of the F statistic was smaller than the P-value of the original F). For this reason, we relied on the bootstrap P-value of the t statistic of the instrument in the first stage.

report the original F statistic (10.51) as well as the bootstrap one (9.25), which accounts for the problem of few clusters. As expected, the former is overestimated. According to the Stock-Yogo critical values, this statistic indicates that the IV estimator of β over-rejects the null, as it leads to a rejection rate close to 15% when the true rejection rate is 5% (Stock and Yogo, 2005).[49] Thus, due to this test size, the IV estimates should be taken with caution.

The upper part of Panel A refers to the employment indicators. The sign of the estimates suggests that early non-employment has a positive impact on the probability to be self-employed and a negative impact on the salaried employment probability, but the size of both effects is very small. In contrast to OLS, 2SLS are not significant: this may be a consequence of a too small power of the test, because of the limited strength of the instrument. The null hypothesis of the exogeneity test is largely not rejected for all indicators, indicating that both estimators are consistent but the OLS is more efficient than the 2SLS one.[50] We therefore focus on OLS: for one *pp* increase in early non-employment, the probability to be salaried employed and overall employed decrease by 0.17% and 0.12%, respectively. These effects are statistically significant. Self-employment increases by 0.05%, but the impact is statistically insignificant.

More significant effects are shown in the bottom part of Panel A, which reports the unconditional effects of interest on continuous labour market outcomes. The null of the exogeneity test is rejected in all cases, meaning that the 2SLS estimator is consistent while the OLS one is not. A comparison between the estimates suggests that OLS underestimate the scarring effect of early non-employment, which is in line with the hypothesis that the bias is caused by returns to search, liquidity constraints or measurement errors in the endogenous regressor, and consistent with what found in the literature. The 2SLS results indicate that one *pp* rise in early non-employment reduces earnings and hours worked by 10% and 7%, respectively (column 4). Both estimates are highly significant (at 1% level). Note that, for the cluster robust case (column 2 and 4) the P-values are computed according to the $t(G-1)$ distribution (with G being the number of clusters), to make a conservative inference.[51] However, the P-values of column 4 may be still underestimated due to the small number of clusters. We tackle this computing the bootstrap P-value for the t statistic of β. The latter is higher than the P-value from cluster robust standard errors, but still lower than 0.05. Hence, despite the small number of clusters, the impact of early non-employment on continuous outcomes is significant.

[49] With one instrument, the critical value for maximal size test of 10% and 15% is 16.3 and 8.96.

[50] With clustered standard errors, the exogeneity test is defined as the difference of two Sargan-Hansen statistics: one for the equation with a smaller set of instruments where the suspect regressor is treated as endogenous, and one for the equation where the suspect regressor is treated as exogenous. Under the null that both sets of instruments are valid (i.e. the suspect regressor is exogenous), the statistic is distributed as $\chi^2(1)$. Note that this statistic is not corrected for the problem of few clusters. Hence, the P-value may be too small.

[51] In Stata this is automatically done in clustered OLS, but not in clustered 2SLS, where the P-values in the second stage are computed with the Normal distribution. In this case, the estimate remains significant at 1% level. Table 7 in Appendix E reports the clustered 2SLS with the significance level based on the Normal distribution.

These estimates suggest that the low educated who found it difficult to get a stable position at the start of the career, are still significantly penalized in terms of hours worked and earnings, 6 years after graduation. These scars may persist because of a number of channels: (i) signaling, that is the unemployment spells convey bad signals to the employers; (ii) human capital depreciation, which makes the unemployed less attractive to employers; (iii) processes of discouragement or habituation (Clark, 2003; Clark et al., 2001). The results on hours worked are not directly comparable but consistent with the existing literature, which reports persistent effects of early unemployment on subsequent unemployment. However, our results are somewhat larger. For UK new graduates aged 16-21, Gregg (2001) estimates that a 3-months increase in the unemployment duration before age 23 significantly increases the time out of work between age 28 and 33 by 2 months. Schmillen and Umkehrer (2013) focus on new graduates from the German apprenticeship program and find larger effects: one additional day of unemployment during the first 8 years since graduation increases unemployment in the following 16 years by 0.96 days - almost a one to one change. The larger scar estimated for Belgium may be explained by the Belgian institutional setting, which leads to high labour market rigidities, as argued in Section 3.1 (i.e. STC program for blue collar workers, the strict EPL for high collar workers and high minimum wages).[52]

Table 7 in Appendix E shows the entire OLS and 2SLS regressions. The individual controls show the expected signs: in the first stage regression, grade repetition in secondary education is positively associated with early non-employment, while technical, vocational and apprenticeship programs are associated with a lower proportion of time spent in early non-employment, compared to general education. This suggests that the former programs ease the transition from school to work. An interesting result refers to mother education, which in the first stage has a positive effect on early non-employment, whereas in the OLS regression of the structural Eq. (2) has a negative impact on hours worked, earnings and salaried employment as well as a positive effect on self-employment. This may capture the effect of unobserved liquidity constraints on time worked, so that less constrained individuals (associated to higher mother's education) spend more time in early non-employment and work less hours in salaried employment (with consequent lower earnings). In the same spirit, low educated individuals with low educated mothers are also more likely to opt for salaried employment (with expected stable income under long-term contracts), while they are more likely to engage in (riskier) self-employment if their mothers are high educated. These effects become insignificant in 2SLS, to the extent that the endogeneity problem due to omitted liquidity constraints in Eq. (2) is tackled by the IV approach.

[52] According to the Kawaguchi and Murao (2014)'s composite index of labour market rigidity, Belgium has the top of the ranking, while Germany is in fourth position and UK is among the last positions.

Table 1: Effect of Interest on Outcomes Measured 6 Years After Graduation for Low Educated

Panel A: Effect of early non-employment in the structural equation:					
		OLS		2SLS	
	standard errors†	robust	cluster $g*p$	robust	cluster $g*p$
outcomes:		(1)	(2)	(3)	(4)
salaried empl.	coeff	-0.00169***	-0.00169***	-0.00256	-0.00256
	se	(0.00034)	(0.00041)	(0.00375)	(0.00290)
	P-val§		0.00019		0.38202
	Bootstrap P-val‡		0		0.45646
	Exogeneity test P-val§§				0.767
self-empl.	coeff	0.00054*	0.00054	0.00248	0.00248
	se	(0.00030)	(0.00041)	(0.00338)	(0.00258)
	P-val		0.19177		0.34175
	Bootstrap P-val		0.18619		0.37437
	Exogeneity test P-val				0.438
overall empl.	coeff	-0.00115***	-0.00115***	-0.00008	-0.00008
	se	(0.00021)	(0.00025)	(0.00207)	(0.00151)
	P-val		0.00005		0.95655
	Bootstrap P-val		0		0.96697
	Exogeneity test P-val				0.467
log earnings	coeff	-0.0269***	-0.0269***	-0.1002**	-0.1002***
	se	(0.0033)	(0.0040)	(0.0419)	(0.0291)
	P-val		2.51E-08		0.0013
	Bootstrap P-val		0		0.0060
	Exogeneity test P-val				0.00970
log hours worked	coeff	-0.0203***	-0.0203***	-0.0723**	-0.0723***
	se	(0.0024)	(0.0029)	(0.0307)	(0.0207)
	P-val		9.35E-09		0.0011
	Bootstrap P-val		0		0.0060
	Exogeneity test P-val				0.0112

Panel B: Effect of the instrument in the first stage : OLS			
outcome:	*standard errors:*	robust	cluster $g*p$
early non-empl.	coeff	5.4615***	5.4615***
	se	(1.7273)	(1.6848)
	P-val		0.00230
	Bootstrap P-val		0.00400
	F stat		10.51
	Bootstrap F stat††		9.25

*** p<0.01, ** p<0.05, * p<0.1. Standard errors between parentheses. Panel A reports results from estimating β in Eq. (2) on outcomes measured at potential experience 6. β is the effect of one *pp* increase in $y^0_{igpt_1}$, i.e. the % of hours spent in non-employment in the first 2.5 years after graduation relative to potential total hours if one would work full-time during the whole period. For clustered standard errors, we report the P-value and the wild bootstrap P-value. Column 1-2 (3-4) show OLS (2SLS). In 2SLS the provincial unemployment rate at graduation is used as instrument for $y^0_{igpt_1}$. Panel B shows the effect of the instrument on $y^0_{igpt_1}$ in the first stage and the corresponding F statistic.

† Robust accounts for heteroscedasticity; clusters are defined by year g and province of residence at graduation p ($G = 44$ clusters).

§ The P-value from clustered standard errors is computed using the $t(G-1)$ distribution, with $G = 44$ (stars are reported accordingly).

‡ Bootstrap P-values are computed according to the wild bootstrap procedures explained in Appendix D for 999 repetitions.

§§ With clustered standard errors, this test is defined as the difference of two Sargan-Hansen statistics: one for the equation where $y^0_{igpt_1}$ is treated as endogenous and one for the equation where $y^0_{igpt_1}$ is treated as exogenous. Under the null that $y^0_{igpt_1}$ is exogenous, the statistic is distributed as $\chi^2(1)$.

†† Bootstrap F statistic is the F statistic corresponding to the bootstrap P-value of the t statistic of the instrument: we rely on the equivalence between F and t distribution: for $G = 44$, $t^2(G-1) = F(1, G-1)$.

4.1 Sensitivity Analysis for the Low Educated

As a first sensitivity analysis, we want to rule out that results in Table 1 are driven by multicollinearity, which may arise because in Eq. (2) we add many controls for the persistence of the unemployment rate (the current unemployment rate UR_{pT}, the average unemployment rate since the end of the early period and the moment of evaluation $\overline{UR}_p$ as well as the minimum unemployment rate since graduation $minUR_{pt}$). Therefore, to rule this possibility we re-run the model including only the current unemployment rate (thereby excluding $\overline{UR}_p$ and $minUR_{pt}$): the results of this restricted specification are reported in Table 9 of Appendix F. Note that the effect of interest should be interpreted as an average effect between the scarring of early non-employment due to adverse labour market conditions at graduation and the persistence of the unemployment rate between the end of the early period and T.[53] The stability of the results ensures that multicollinearity is not driving the results.

Next, we assess the impact of measurement error in the endogenous regressor, which arises from the combination of survey and administrative data. In fact, early non-employment is measured precisely for individuals graduating in the 1998-2002 period - by means of administrative data - but it is imputed for those graduating in the period 1994-1997, based on the Sonar database (see Appendix B for details). This allows us to maximize the variation of the instrument considering the entire graduation period 1994-2002, at the cost of introducing some measurement error in the endogenous regressor. Therefore, in this second sensitivity analysis we re-run the analysis for the low educated restricting the sample to graduation period 1998-2002, so that the endogenous regressor is measured uniquely by administrative data. Of course, clusters are drastically reduced from 44 to 24.[54] This is problematic not only because it exacerbates the problem of few clusters, but also because Eq. (2) contains too many parameters ($k = 30$) compared to the number of clusters, and as a consequence the rank condition in 2SLS is not satisfied.[55] Therefore, we need to reduce the parameters in Eq. (2).

We decide to exclude some of the non-significant individual controls and rather include in the specification all the aggregate regressors, which are very important to ensure the validity of the exclusion restriction.[56] In particular, we drop the following controls that are jointly not significant at the 5% level according to an F test in the first stage regression: dummy for living with single parent, dummy for not living with parents, number of household members aged 12-17, 18-29, 30-64, 65+ (we keep the number of household members aged 0-11 since it is significant); plus, we aggregate all educational tracks different from general education (technical, vocational, part-time education or apprenticeship) and include a dummy for general education

[53] This is because $\overline{UR}_p$ is significant in Eq. (2) while $minUR_{pt}$ is not: see Table 7 and 8 Appendix E.

[54] In principle, 5 graduation years times 5 provinces, i.e. $G = 25$. However, $g2002p3$ is empty.

[55] The variance-covariance matrix of moment conditions has size (30×30) and rank=24 (Baum et al., 2003).

[56] Compared to the aggregate regressors individual controls play a minor role, as they alleviate the problem of omitted individual characteristics, which is anyway tackled by the IV approach.

instead.[57] Therefore, the new regression includes the following individual controls: father and mother education, repeated grades at age 17, dummy for general education at age 17, number of household members aged 0-11 when the individual is aged 17, birth cohort dummies. Table 11 shows that results are robust to this alternative specification, as the OLS estimate of the endogenous regressor is very stable in the full specification (column 1 and 4) and in the restricted specification (column 2 and 5); both specifications consider the graduation period 1994-2002.

Panel B of Table 10 reports the first stage regression for the 1998-2002 period. This table should be compared to Table 1. First, the bootstrap F statistic is 3.6 for the graduation period 1998-2002 compared to 9.2 for the period 1994-2002: therefore, the instrument becomes weak by restricting to the former period. As expected, the increased discrepancy between the original F and the bootstrap F statistic in Table 10 compared to Table 1 shows that the few-clusters bias worsens a lot by shifting from 44 to 24 clusters. Second, the direct effect of the instrument on early non-employment for the 1998-2002 period doubles compared to the period 1994-2002. An explanation is that the former period focuses on the dot-com recession, whose effects are mitigated in considering a larger span. Given the low F statistic, 2SLS are not reliable.

However, we can focus on the OLS results in Table 11 to shed lights on the importance of the measurement error in the endogenous regressor. In principle, this should bias the OLS estimate for the period 1994-2002 (Column 1 and 4) towards zero. At the same time, the OLS estimate in Column 3 and 6 should not be affected by measurement error "by construction", since the endogenous regressor is entirely measured by administrative data for the 1998-2002 period. We therefore compare the first row across columns (1 with 3 and 4 with 6): for each outcome, the estimate based on the graduation period 1994-2002 is slightly smaller than the corresponding estimate for the period 1998-2002. This is consistent with the presence of measurement error in the endogenous regressor for the graduation period 1994-2002. However, this difference is small (0.2 *pp*), which suggests that overall the OLS estimates are quite close in the periods 1998-2002 and 1994-2002: as a consequence, we conclude that measurement error in early non-employment is not a major issue in the main results.

Of course, another explanation could be that the combination of bias from various sources (ability, returns to search, liquidity constraints) may differ in the period 1998-2002 - when measurement error is absent - compared to the period 1994-2002 - when measurement error is present - and yet yield the same net effect: however, this requires that these (fixed) omitted factors affect individual labour market performances differently in the 1994-1997 and 1998-2002 period, which is peculiar. We think that this last story is more difficult to be argued.

[57] The reference is an aggregated category for technical, vocational, part-time education and apprenticeship.

5 Conclusions

This study investigates the causal impact of experiencing early non-employment on the subsequent labour market career of low educated youth graduating in the period 1994-2002 in Flanders, the most prosperous of the three Belgian region. The cases of Belgium is particularly interesting, because it is characterized by one of the most rigid labour markets in OECD. Besides, Belgian labour market institutions entail different sources of rigidities for blue and white collar workers. Cockx and Ghirelli (2015) provide evidence that, because of these differences, the adverse labour market conditions at graduation inflict different scars to the Belgian low and high educated new graduates. Due to the strict EPL of white collar workers, the high educated are forced to downgrade and as a consequence are trapped in lower-paying jobs. The low educated instead remain unemployed more often. This suggests that the scar from graduating in downturns occurs through the loss of early work experience for the low educated, and through the acceptance of lower-paying jobs for the high educated.

In this study, we define the empirical strategy in light of these results. We focus on the low educated and consider a measure of the early experience of non-employment as endogenous variable. The endogeneity of unobserved individual characteristics is tackled by an IV approach, where the provincial unemployment rate at graduation is used as instrument for early non-employment. This ensures that the exclusion restriction is most likely satisfied, based on the aforementioned results. Throughout the article we discuss the assumptions required by the IV estimator, their validity and the role they play in the identification of causality between early non-employment and subsequent labour market outcomes. The problem of few clusters is addressed by wild bootstrap methods.

We find that one *pp* increase in the time spent in non-employment in the first 2.5 years since graduation decreases annual earnings and hours worked from salaried employment by 10% and 7% respectively, 6 years after graduation. These effects are unconditional on being salaried employed. Provided that our identification strategy is correct, these effects are causal. These scars may be originated by the human capital depreciation occurring in the unemployment spell, by the foregone human capital that would have been accumulated in case of early work-experience, or because early non-employment is interpreted as a signal of low quality.

At a micro level, one can think of specific curative policies to reduce the impact of the early experience of non-employment, depending on the channel through which the scar materializes. If the first mechanism is the main driver of the scar, training schemes targeting young unemployed would be an appropriate cure. If instead the main cause of the scar is the loss of early work experience, policies that foster the integration of youth in the labor market should be advocated, such as wage subsidy programs. However, if the cause of the scar is rather the bad signals conveyed by the status of unemployment, it is likely that entering in subsidized programs would not improve the perception of the young unemployed by the employers. More reflection on the

actual channels should be needed to come up with a specific policy to target the Flemish youth.

From a more aggregate perspective, the acquisition of early work experience is clearly more difficult in a rigid labour market where hiring and firing workers is costly. Thus, the aforementioned scars could be diminished in a *flexicurity* system, where a flexible labour market is associated with a generous insurance system that provides for the unemployed. On the one hand, flexibility should be enhanced to encourage employers to take the entrepreneurial risks and create more jobs. This is because labor market rigidities may hamper productivity growth if they prevent workers reallocation or new hirings due to high expected future firing costs in future downswings. On the other hand, a generous unemployment insurance should compensate workers for bearing a higher risk of unemployment due to the increased flexibility in the labour market. Thus, in the context of flexicurity, low hiring costs should enhance workers reallocation and new hirings as well as decrease the expected duration of unemployment, thereby fostering early work experience. However, less employable individuals may be long-term unemployed. This group could be supported by active labor market programs as well as training schemes to improve their employability and help them to find a good match.

The Belgian low educated workers face a number of rigidities which restrain the reallocation or the new hirings of low educated workers: the STC program, by anchoring the blue collar employees to their employers in recessions, is an example. A second rigidity is represented by the very high Belgian minimum wages, which limit the absorbtion of low educated new graduates for whom minimum wages are binding. A third one is the asymmetry between the flexible EPL for blue collars and a rigid EPL for white collars, which characterized the Belgian labour market until 2013. Note that this controversial discrimination between blue and white collar workers has been removed since the beginning of 2014, as a single employment contract has been introduced, stipulating the same EPL for white and blue collar workers.

Acknowledgements

I wish to thank Bart Cockx, William Parienté, Matteo Picchio, Bruno Van der Linden, Dieter Verhaest, Florian Mayneris and the participants to the seminars at Ghent University (internal seminar), and the participants to the Doctoral Workshop at the Université Catholique de Louvain for all their valuable comments. I am grateful to the Cross Roads Bank for Social Security and the "Steunpunt Studie- en Schoolloopbanen" for the availability of the data. I acknowledge financial support from the Special Research Fund ("BOF") of Ghent University, as well as from the National Bank of Belgium (project code 05Y00110). I also acknowledge the financial support from the Université Catholique de Louvain ("WOLDEMP" project). The usual disclaimer applies.

APPENDIX

A Description of the Final Sample

In this study we consider almost the same sample as in Cockx and Ghirelli (2015). They consider 3514 individuals (comprising both low and high educated) while we consider 3586 individuals (see Table 2): i.e. we add 72 individuals (2% of the sample), including low educated graduating in 2002 and high educated graduating in 1995-96. For details in the construction of the control and outcome variables, see Section S.1, S.2 and S.3 of the Supplementary Online Appendix of Cockx and Ghirelli (2015): http://users.ugent.be/~bcockx/Ascars.pdf

Table 2: Dividing the Sample in Low and High Educated

completed education	low educated	high educated	Total
1	2		2
2	36		36
3	89		89
4	113		113
5	185		185
6	1,111		1,111
7	366	289	655
8		55	55
9		707	707
10		367	367
11		232	232
12		33	33
13		1	1
Total	1,902	1,684	3,586

"Completed education" refers to the number of years of education successfully attained from the beginning of secondary education, i.e. at age 12. Low educated are those who graduate with at most secondary education, which consists in 7 years of education in case of vocational track and 6 years for all other educational programs. High educated are those with higher than secondary education.

Table 3: Correspondence Between Low-High Educated and Blue-White Collar Workers

Function Undertaken 6 Years After Graduation†						
	Low educated			High educated		
	Freq.	Percent	Cum.	Freq.	Percent	Cum.
blue collar	1,193	62.72	62.72	184	10.93	10.93
white collar	390	20.5	83.23	1,193	70.84	81.77
functionary	68	3.58	86.8	105	6.24	88
missing	251	13.2	100	202	12	100
Total	1,902	100		1,684	100	
Prevalent Function Undertaken up to 6 Years After Graduation‡						
	Low educated			High educated		
	Freq.	Percent	Cum.	Freq.	Percent	Cum.
blue collar	1,346	70.77	70.77	235	13.95	13.95
white collar	401	21.08	91.85	1,337	79.39	93.35
functionary	53	2.79	94.64	36	2.14	95.49
missing	102	5.36	100	76	4.51	100
Total	1,902	100		1,684	100	

† It refers to the type of function undertaken at potential experience 6.

‡ It refers to the function that is undertaken more than 50% of the time from graduation up to potential experience 6. 70% of low educated are prevalently employed as blue collars while this figure is only 14% for the high educated. Thus, there is clear correspondence between low educated and blue collar workers and between high educated and white collar workers.

Table 4: Number of Individuals by Graduation Year and Province of Residence at Graduation

Low educated						
grad_year	prov1	prov2	prov3	prov4	prov5	Total
1994	30	9	31	48	25	143
1995	47	22	44	48	48	209
1996	84	45	65	85	38	317
1997	78	41	65	67	36	287
1998	111	46	78	90	61	386
1999	99	42	47	64	47	299
2000	56	18	30	28	31	163
2001	26	8	11	17	18	80
2002	10	3	0	2	3	18
Total	541	234	371	449	307	1902

The analysis considers the graduation period 1994-2002 for the low educated. The combination g2002 & prov3 is excluded since empty. Provinces are in the following order from 1 to 5: Antwerp, Flemish Brabant, Western Flanders, Eastern Flanders, Limburg. Each combination of graduation year and province of residence at graduation represents a cluster *gp* in the main analysis.

Table 5: Descriptive Statistics of Individual Control Variables: Low Educated

Variable	Obs	Mean	Std. Dev.	Min	Max	label
birth cohort76	1902	0.33	0.47	0	1	1 if born in 1976
birth cohort78	1902	0.33	0.47	0	1	1 if born in 1978
birth cohort80	1902	0.34	0.47	0	1	1 if born in 1980
live in single-parent	1902	0.12	0.33	0	1	1 if live with single parent at age17(Dec)
not live with parents	1902	0.06	0.24	0	1	1 if not live with either parents at age17(Dec)
HH members aged 0-11	1902	0.25	0.62	0	7	nr of other HH members aged0-11 at age17(Dec)
HH members aged 12-17	1902	0.51	0.69	0	7	nr of other HH members aged12-17 at age17(Dec)
HH members aged 18-29	1902	0.52	0.73	0	8	nr of other HH members aged18-29 at age17(Dec)
HH members aged 30-64	1902	1.89	0.40	0	5	nr of other HH members aged30-64 at age17(Dec)
HH members aged 65+	1902	0.04	0.21	0	2	nr of other HH members aged65+ at age17(Dec)
father education	1902	4.59	3.20	0	13	father completed education since age12
mother education	1902	4.21	3.06	0	13	mother completed education since age12
years of delay in sec.edu	1902	0.83	0.84	-1	4	years of delay at age17(Aug)
general education	1902	0.11	0.31	0	1	1 if general edu at age17(Aug)
technical education	1902	0.38	0.49	0	1	1 if technical edu at age17(Aug)
vocational education	1902	0.41	0.49	0	1	1 if vocational edu at age17(Aug)
apprenticeship/PT edu	1902	0.10	0.30	0	1	1 if apprenticeship/PT edu at age17(Aug)

Table 6: Descriptive Statistics of Outcomes of Interest: Low educated at potential experience 6

Variable	Obs	Mean	Std. Dev.	Min	Max	Label
Continuous outcomes in level						
earnings	1902	19277.29	10978.69	1	47551	yearly gross earnings from salaried empl.+1
hours worked	1902	1390.923	725.0779	1	2265	yearly hours worked in salaried empl.+1
Continuous outcomes in log						
log earnings	1902	8.42	3.59	0	10.77	log (yearly gross earnings from salaried empl.+1)
log hours worked	1902	6.20	2.67	0	7.73	log (yearly hours worked in salaried empl.+1)
Dichotomous outcomes						
self-empl.	1902	0.12	0.33	0	1	1 if only pos. earnings from salaried (& not self-empl)
salaried empl.	1902	0.84	0.37	0	1	1 if registered as self-empl.
overall empl.	1902	0.96	0.20	0	1	1 if pos.earnings from salaried or registered as self-empl.
Endogenous regressor						
early non-empl.	1902	30.60	29.65	0	100	% hours not worked relative to FT salaried empl.

The employment indicators are related as follows: $salaried + self = overall\,employment$. For continuous outcomes we add value one before taking the log, so that non-salaried employed at the moment of evaluation are included with outcomes=0 after the logarithmic transformation.

B Construction of the Endogenous Variable

We define potential experience as a variable counting each calendar year since graduation. Potential experience 0 corresponds to the year of graduation and runs from the month after graduation until December of that calendar year; therefore, it potentially lasts less than 12 months (6 months for a student graduating in June). Potential experience 1 runs from January until

December of the subsequent calendar year, thereby lasting 12 months. Subsequent potential experience years are defined similarly.

The regressor of interest is the percentage of hours spent in non-employment at potential experience 0-2, relative to potential total hours if one would work full-time during the whole period. We express everything in hours because this is the smallest unit of measurement used in the administrative data (it is used to measure time worked in part-time employment). The reference period is computed considering the entire calendar year for potential experience 1 and 2, and the part of the calendar year following the month of graduation for potential experience 0. That is, for one who graduated in June, the time spent working at potential experience 0-2 is divided by the total working hours in 30 months of full-time salaried employment. As already mentioned in Section 2 of the main text, for simplicity we refer to this reference period as "the first 2.5 years since graduation", i.e. the average reference period since most of the sample graduates in June.

Define a as the total hours worked (including self-employment) in the first 2.5 years from graduation and b the potential total hours if one wold work full-time during the whole period; then the regressor of interest is computed as $100 * (b - a)/b$. Below we explain in detail how these components are constructed.

1. Construct a according to the following steps.

    I. a is mostly based on the total hours worked in salaried employment and the date of registration and cancellation from the self-employment register from the Data Warehouse. Since hours worked are not available for self-employment, we assume that the latter work as much as a full-time salaried worker: i.e. 5 days per week and 8 hours per day until 2002, and 5 days per week and 7.6 hours per day from 2003 onwards. This is due to the introduction of a new law in Belgium that changed the daily working hours from 8 to 7.6 from the first of January 2003. Whenever one combines self-employment and salaried-employment in the same quarter we make the same assumption, so that the hours worked do not exceed the bounds.

    II. The construction of a requires an additional adjustment due to the limited availability of the administrative data, which cover the period 1998-2010. Since the sample contains 3 birth cohorts (1976, 1978, 1980) and that compulsory education ends at age 18 in Belgium, these data can be used in the following cases (68.5% of the final sample): all individuals born in 1980, those born in 1978 graduating at least at age 20 and those born in 1976 graduating at least at age 22. Figure 1 summarizes the availability of the data. To retain in the analysis also students born students born in 1976 graduating at age 18-21 as well as students born in 1978 graduating at age 18-19,[58] we exploit the monthly working status from the Sonar database and impute

[58] These cases correspond to 31.5% of the final sample.

the values of a following the procedure used for self-employed workers. That is, for each month in which individuals are working according to Sonar we attribute the working hours of a full-time salaried worker: i.e. 8 working hours per day until 2002 and 7.6 working hours per day strictly after 2002. For the entire observation period, we consider 21.6 working days per month (assuming 65 working days in a quarter gives 21.6 working days per month: $21.6 \times 3 = 65$).

2. Construct b. Recall that it is defined as the potential total hours if one wold work full-time during the first 2.5 years since graduation. As for a, we consider a full-time working regime of 5 days per week, 8 hours per day until 2002 as well as 7.6 hours per day strictly after 2002. This gives a total of 2080 annual working hours until 2002 (8 hours/day $\times$ 65 days/quarter $\times$ 4 quarters/year) and 1976 annual working hours from 2003 onwards (7.6 hours/day $\times$ 65 days/quarter $\times$ 4 quarters/year).

3. The regressor of interest is computed as $(b - a)/b * 100$ and hence ranges between $[0, 100]$. In some cases (10% of the final sample) this percentage is negative because of overtime work. Therefore, it is censored at 0.

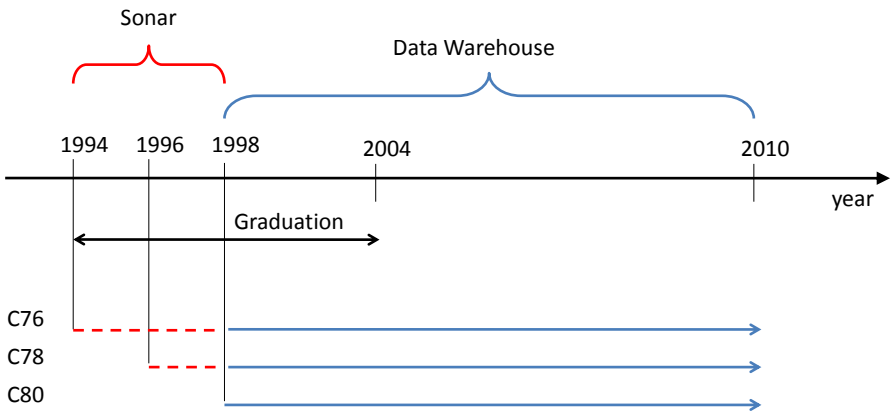

Figure 1: Availability of data for the construction of the main regressor. In 1998 birth cohorts 76, 78 and 80 are aged 22, 20 and 18, respectively. Birth cohort 76 and 78 turn age 18 in 1994 and 1996, respectively.

C LFS - Provincial unemployment rate

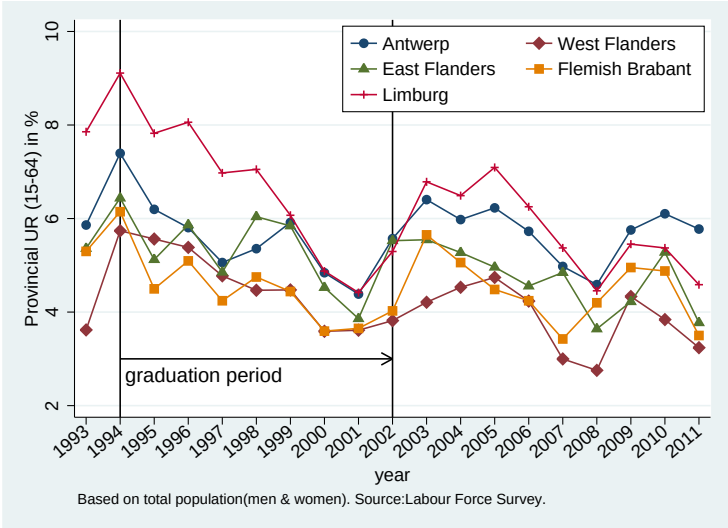

Figure 2: Provincial unemployment rates (15-64) for Flanders: graduation period for the low educated ranges 1994-2002. For details on the series, see Appendix S.1.5 of the Supplementary Online Appendix of Cockx and Ghirelli (2015).

D Bootstrap Procedure

The basic idea of bootstrap testing is to compare the observed value of some test statistic with the empirical distribution of B bootstrap test statistics computed on as many pseudo-samples, where B is the number of bootstrap replications. We use a *Wild Restricted Efficient Residual Bootstrap* (WRE Bootstrap) proposed by Davidson and MacKinnon (2010). It is the wild version of the Restricted Efficient Residual bootstrap designed for 2SLS by Davidson and MacKinnon (2008).

Few words on the terminology (which will become clearer below): *Residual* means that the objects to sample in generating the pseudo-samples are the residuals.[59] *Wild* refers to a procedure that creates pseudo-samples based on $residuals*1$ with probability 0.5 and $residuals*(-1)$ with probability 0.5, with this assignment at the cluster level. This allows to preserve the intra-cluster correlation. *Efficient* means that the first stage of 2SLS is efficiently estimated in case of weak instruments. *Restricted* means that the null hypothesis of interest is imposed on the data generating process (DGP): this enhances efficiency in the procedure. Consider the

[59] Alternatively, one can sample pairs $[y\,X]$ of data.

following system of equation, which is a simplified version of the 2SLS model of interest:

$$y_{igp} = \beta y^0_{igp} + x'_i\delta + e_{igp} \tag{A.1}$$

$$y^0_{igp} = \pi Z_{gp} + x'_i\delta + u_{igp} \tag{A.2}$$

Eq. (A.1) is the structural equation where individual labour market outcomes y_{igp} are regressed on early non-employment y^0_{igp} and individual controls (for simplicity we omit the grouped covariates and the time subscripts in Eq. 2), and Eq. (A.2) is the first stage regression where the endogenous explanatory variable is regressed on the grouped instrument Z_{gp} and all exogenous regressors x_i. The fact that the instrument is grouped requires cluster robust standard errors in 2SLS. We are interested in bootstrapping the t statistic of y^0_{igp}, i.e. $t(\hat{\beta}, \beta_0) = \frac{(\hat{\beta}-\beta_0)}{se(\hat{\beta})}$. Call $\hat{\tau}$ the observed value of this statistic. The bootstrap procedure will generate an empirical distribution of B bootstrap test statistics τ^*, with B being the number of repetitions, and where these statistics are generated using the bootstrap DGP which imposes the null hypothesis that is tested. In practice this is implemented as follows:

1. Estimate the system in Eq. (A.1)-(A.2) by 2SLS with cluster robust standard errors and obtain the t statistic $\hat{\tau}$.

2. Estimate the restricted version of Eq. (A.1) by OLS imposing the null hypothesis $\beta = 0$ (with conventional standard errors). Predict the residuals $\tilde{e}_{igp}$ and the fitted values $\tilde{y}_{igp}$. (*Restricted*)

3. Estimate Eq. (A.2) including $\tilde{e}_{igp}$ as additional control, i.e.: $y^0_{igp} = \pi Z_{gp} + x'_i\delta + \gamma\tilde{e}_{igp} + residuals$. Compute the residuals $\tilde{u}_{igp} = residuals + \hat{\gamma}\tilde{e}_{igp}$. This allows the residuals of the first stage not to be too small in case of weak instrument (*Efficient*). Accordingly, compute the fitted values $\tilde{y}^0_{igp} = \hat{\pi}Z_{gp} + x'_i\hat{\delta}$.

4. At the cluster level, multiply the residuals $\tilde{u}_{igp}$ and $\tilde{e}_{igp}$ by a random variable ν^*, where $\nu^* = 1$ and $\nu^* = -1$ with probability 1/2, respectively. Note that the same ν^* is applied to both residuals: this preserves the correlation across Eq. (A.1) and (A.2).[60] (*Wild*)

5. Construct $y^*_{igp} = \tilde{y}_{igp} + \nu^*\tilde{e}_{igp}$ and $y^{0*}_{igp} = \tilde{y}^0_{igp} + \nu^*\tilde{u}_{igp}$.

6. Estimate Eq. (A.1)-(A.2) by 2SLS where y_{igp} is replaced by y^*_{igp} and y^0_{igp} by y^{0*}_{igp}, with cluster robust standard errors and obtain the t statistic τ^*.

7. Repeat steps 4-6 B times, where B is the number of repetitions, so to get an empirical distribution of τ^*_j for $j = 1, ..., B$.

8. Calculate the bootstrap P-value as $p^*(\hat{\tau}) = 2min(\frac{1}{B}\sum_1^B \mathbf{1}[\tau^*_j < \hat{\tau}], \frac{1}{B}\sum_1^B \mathbf{1}[\tau^*_j > \hat{\tau}])$.

[60] Here we use the Rademacher weights, which have been shown to work well when the residuals are not too asymmetric. Other weights can be used.

Below we describe the simpler procedure to compute wild bootstrap in the OLS case. In the main analysis, we apply the latter to the OLS estimations of the structural equation and to the first stage. Below we take the first stage as example: we are interested in estimating Eq. (A.2) by OLS and then in bootstrapping the t statistic of the instrument, $t(\hat{\pi}, \pi_0) = \frac{(\hat{\pi}-\pi_0)}{se(\hat{\pi})}$. The procedure is reported below:

1. Estimate Eq. (A.2) by OLS with cluster robust standard errors and obtain the t statistic $\hat{\tau}$.

2. Re-estimate Eq. (A.2) imposing the null hypothesis $\pi = 0$ (with conventional standard errors). Predict the residuals $\tilde{u}_{igp}$ and the fitted values $\tilde{y}^0_{igp}$. (*Restricted*)

3. At the cluster level, multiply the residuals $\tilde{u}_{igp}$ by a random variable ν^*, where $\nu^* = 1$ and $\nu^* = -1$ with probability 1/2, respectively.

4. Construct $y^{0*}_{igp} = \tilde{y}^0_{igp} + \nu^* \tilde{u}_{igp}$.

5. Estimate Eq. (A.2) by OLS where y^0_{igp} is replaced by y^{0*}_{igp}, with cluster robust standard errors and obtain the t statistic τ^*.

6. Repeat steps 3-5 B times, where B is the number of repetitions, so to get an empirical distribution of τ^*_j for $j = 1, ..., B$.

7. Calculate the bootstrap P-value as $p^*(\hat{\tau}) = 2min(\frac{1}{B}\sum_1^B \mathbf{1}[\tau^*_j < \hat{\tau}], \frac{1}{B}\sum_1^B \mathbf{1}[\tau^*_j > \hat{\tau}])$.

E Complete Results for the Low Educated

Table 7: Complete Estimations on Continuous Outcomes for the Low Educated

	second stage on continuous outcomes				first stage
Outcomes‡	log earnings		log hours worked		early non-empl.
	OLS	2SLS	OLS	2SLS	OLS
clustered standard errors:	$g*p$	$g*p$	$g*p$	$g*p$	$g*p$
	(1)	(2)	(3)	(4)	(5)
UR_grad					5.4615***
					(1.6848)
early non-empl.	-0.0269***	-0.1002***	-0.0203***	-0.0723***	
	(0.0040)	(0.0291)	(0.0029)	(0.0207)	
UR_pe6	0.4823**	0.4172**	0.3341**	0.2880**	0.3779
	(0.1904)	(0.1919)	(0.1382)	(0.1392)	(1.6071)
lin_grad_year	-0.1588	0.0469	-0.1139	0.0319	3.5329
	(0.2945)	(0.2995)	(0.2134)	(0.2191)	(2.2946)
lin_grad_year|trend>3	0.9033**	0.6407	0.6676**	0.4814	-4.4281
	(0.4003)	(0.4242)	(0.2915)	(0.3109)	(2.8479)
lin_grad_year|trend>6	-0.4604*	-0.4394	-0.3335*	-0.3186	3.9169
	(0.2582)	(0.2852)	(0.1866)	(0.2036)	(3.1378)
d_province2	-1.3812*	-1.2603	-0.9479*	-0.8622	3.6574
	(0.7107)	(0.9544)	(0.5041)	(0.6684)	(8.6769)
d_province3	-2.1753***	-2.9226***	-1.5774***	-2.1070***	-11.7452
	(0.6066)	(0.9315)	(0.4344)	(0.6703)	(8.9621)
d_province4	0.0841	0.1824	0.0740	0.1436	0.7495
	(0.4801)	(0.4983)	(0.3504)	(0.3570)	(4.7080)
d_province5	0.8990	0.9210	0.6674	0.6829	-15.1893**
	(0.6099)	(0.6721)	(0.4495)	(0.4866)	(6.6042)
lin_calend_year_prov2	0.0498	0.0410	0.0349	0.0287	0.0774
	(0.1386)	(0.1413)	(0.1019)	(0.1024)	(0.9931)
lin_calend_year_prov3	0.0723	0.1638	0.0588	0.1236	2.2624
	(0.1136)	(0.1441)	(0.0811)	(0.1040)	(1.7325)
lin_calend_year_prov4	-0.0946	-0.1524	-0.0684	-0.1094	-1.0198
	(0.1267)	(0.1477)	(0.0927)	(0.1051)	(1.2618)
lin_calend_year_prov5	-0.0813	-0.0760	-0.0593	-0.0555	1.8419
	(0.1067)	(0.1194)	(0.0795)	(0.0868)	(1.1761)
avg_UR_pe3-6	-1.2836***	-1.4348**	-0.9652***	-1.0724***	-0.3037
	(0.4509)	(0.5863)	(0.3262)	(0.4140)	(4.4775)
min_UR_pe0-6	-0.5718	-0.4848	-0.3128	-0.2511	-6.4742
	(0.6219)	(0.6613)	(0.4391)	(0.4664)	(7.1097)
birth cohort76	0.7043	-0.3300	0.5421	-0.1909	-13.2641***
	(0.5632)	(0.7686)	(0.4125)	(0.5583)	(3.6802)
birth cohort78	0.4261	-0.0504	0.3188	-0.0190	-5.9156**
	(0.3677)	(0.4687)	(0.2691)	(0.3409)	(2.5815)
live in single-parent	0.3453	0.7897	0.2731	0.5880	6.1392
	(0.4632)	(0.5207)	(0.3466)	(0.3840)	(4.1354)
not live with parents	0.4272*	0.5203	0.3476*	0.4136	1.0913
	(0.2431)	(0.3501)	(0.1792)	(0.2530)	(2.5611)
HH members aged 0-11	-0.0177	0.0846	-0.0462	0.0263	1.3007
	(0.1123)	(0.1472)	(0.0878)	(0.1069)	(1.1570)

Continued on next page

Table 7 – continued from previous page

Outcomes‡	log earnings		log hours worked		early non-empl.
	OLS	2SLS	OLS	2SLS	OLS
	(1)	(2)	(3)	(4)	(5)
HH members aged 12-17	0.1680	0.1467	0.1261	0.1110	-0.2429
	(0.1177)	(0.1473)	(0.0883)	(0.1087)	(0.9030)
HH members aged 18-29	0.0112	0.2054	0.0073	0.1449	2.6587**
	(0.1164)	(0.1561)	(0.0870)	(0.1134)	(1.0291)
HH members aged 30-64	-0.0334	-0.0801	-0.0112	-0.0443	-0.4836
	(0.4065)	(0.4287)	(0.3022)	(0.3129)	(3.2574)
HH members aged 65+	-0.0680	0.0308	-0.0289	0.0411	1.2468
	(0.3639)	(0.4196)	(0.2683)	(0.3079)	(3.1042)
father education	0.0011	0.0294	0.0000	0.0201	0.3816
	(0.0255)	(0.0330)	(0.0188)	(0.0240)	(0.2531)
mother education	-0.1053***	-0.0485	-0.0765***	-0.0363	0.7800**
	(0.0356)	(0.0432)	(0.0269)	(0.0317)	(0.2925)
years of delay in sec. edu	-0.0545	0.3485*	-0.0423	0.2433	5.4123***
	(0.1112)	(0.2094)	(0.0808)	(0.1510)	(1.1119)
technical edu	0.4431	-0.4557	0.3563	-0.2807	-11.8663***
	(0.3274)	(0.4733)	(0.2470)	(0.3399)	(2.9533)
vocational edu	0.4572	-0.3520	0.3675*	-0.2059	-10.6291***
	(0.2772)	(0.4516)	(0.2101)	(0.3233)	(3.3428)
apprenticeship/PT edu	-0.1879	-1.1341*	-0.1093	-0.7798*	-12.5300**
	(0.4530)	(0.6234)	(0.3375)	(0.4500)	(4.7050)
Constant	15.0564***	18.1335***	10.7790***	12.9598***	24.5592
	(3.7529)	(5.2571)	(2.6949)	(3.7277)	(49.3992)
Observations	1,902	1,902	1,902	1,902	1,902
R-squared	0.0895	-0.2401	0.0895	-0.2099	0.1070
F stat of first step§					10.51
Exogeneity test P-val†		0.00970		0.0112	

*** p<0.01, ** p<0.05, * p<0.1. Standard errors between parentheses. Columns 1-4 report the results from estimating Eq. (2) by OLS (odds columns) and 2SLS (even columns). Column 5 reports OLS results from estimating the first stage. All estimations report cluster robust standard errors by graduation year g and province of residence at graduation p ($G = 44$). Column 5 reports the F statistic of the first stage and even columns report the exogeneity test for early non-employment ($y^0_{igpt_1}$).

‡ Continuous outcomes are measured at potential experience 6; early non-emp-loyment is measured in the first 2.5 years after gradaution. For continuous outcomes we add value one before taking the logarithm, so that those who are not salaried employed at the moment of evaluation are included with outcomes equal to zero after the logarithmic transformation.

§ This statistic is not corrected for the problem of few clusters. The corrected value resulting from the bootstrap procedure is 9.25 (see table 1).

† With clustered standard errors, the exogeneity test is defined as the difference of two Sargan-Hansen statistics: one for the equation where y^0_{t1} is treated as endogenous and one for the equation where $y^0_{igpt_1}$ is treated as exogenous. Under the null that $y^0_{igpt_1}$ is exogenous, the statistic is distributed as $\chi^2(1)$. This statistic is not corrected for the problem of few clusters.

Table 8: Complete Estimations on Discrete Outcomes for the Low Educated

	second stage					
Outcomes:‡	salaried empl.		self-empl.		overall empl.	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
clustered standard errors:	$g*p$	$g*p$	$g*p$	$g*p$	$g*p$	$g*p$
	(1)	(2)	(3)	(4)	(5)	(6)
early non-empl.	-0.00169***	-0.00256	0.00054	0.00248	-0.00115***	-0.00008
	(0.00041)	(0.00290)	(0.00041)	(0.00258)	(0.00025)	(0.00151)
UR_pe6	0.00481	0.00404	0.01081	0.01253	0.01562	0.01657*
	(0.02385)	(0.02209)	(0.01947)	(0.01765)	(0.01020)	(0.00999)
lin_grad_year	-0.00253	-0.00009	-0.01367	-0.01911	-0.01621	-0.01920
	(0.03120)	(0.03028)	(0.02544)	(0.02343)	(0.01397)	(0.01340)
lin_grad_year|trend>3	0.02993	0.02680	0.00338	0.01032	0.03331*	0.03712**
	(0.04020)	(0.04043)	(0.03029)	(0.02903)	(0.01832)	(0.01813)
lin_grad_year|trend>6	-0.01720	-0.01695	0.01884	0.01829	0.00164	0.00134
	(0.02568)	(0.02477)	(0.02503)	(0.02252)	(0.01457)	(0.01375)
d_province2	-0.22288**	-0.22144**	0.17666*	0.17347**	-0.04621	-0.04797
	(0.09844)	(0.09548)	(0.09049)	(0.08719)	(0.04174)	(0.03996)
d_province3	-0.17150***	-0.18038***	0.13845**	0.15820**	-0.03304	-0.02218
	(0.06193)	(0.06407)	(0.06588)	(0.06577)	(0.03720)	(0.04285)
d_province4	-0.11409**	-0.11292**	0.10736**	0.10477**	-0.00673	-0.00816
	(0.04892)	(0.04910)	(0.04459)	(0.04794)	(0.02450)	(0.02611)
d_province5	0.04132	0.04158	-0.06171	-0.06229	-0.02039	-0.02071
	(0.06183)	(0.06032)	(0.04557)	(0.04483)	(0.03418)	(0.03449)
lin_calend_year_prov2	0.01976	0.01966	-0.01544	-0.01521	0.00432	0.00445
	(0.01597)	(0.01524)	(0.01253)	(0.01214)	(0.00565)	(0.00651)
lin_calend_year_prov3	-0.00496	-0.00387	0.00501	0.00259	0.00005	-0.00128
	(0.01207)	(0.01140)	(0.01565)	(0.01376)	(0.00716)	(0.00690)
lin_calend_year_prov4	0.01299	0.01231	-0.01369	-0.01217	-0.00070	0.00014
	(0.01118)	(0.01157)	(0.00964)	(0.01055)	(0.00642)	(0.00679)
lin_calend_year_prov5	-0.00608	-0.00602	0.01656	0.01642*	0.01048**	0.01041**
	(0.01074)	(0.01027)	(0.01012)	(0.00967)	(0.00471)	(0.00491)
avg_UR_pe3-6	-0.08309**	-0.08489**	0.03333	0.03733	-0.04976*	-0.04756*
	(0.04062)	(0.04005)	(0.03418)	(0.03230)	(0.02622)	(0.02516)
min_UR_pe0-6	-0.01193	-0.01089	0.00452	0.00222	-0.00741	-0.00867
	(0.06247)	(0.05970)	(0.05849)	(0.05672)	(0.02769)	(0.02854)
birth cohort76	0.00488	-0.00741	0.01471	0.04204	0.01959	0.03462
	(0.04397)	(0.06402)	(0.03588)	(0.05434)	(0.02535)	(0.03517)
birth cohort78	0.02309	0.01743	-0.00086	0.01173	0.02223	0.02916
	(0.03666)	(0.04407)	(0.03059)	(0.03645)	(0.01745)	(0.02152)
live in single-parent	-0.00309	0.00220	-0.03561	-0.04736	-0.03870	-0.04516
	(0.05525)	(0.05471)	(0.05233)	(0.05472)	(0.03217)	(0.03371)
not live with parents	0.02521	0.02631	-0.01502	-0.01748	0.01019	0.00883
	(0.02900)	(0.02995)	(0.02753)	(0.03031)	(0.01581)	(0.01592)
HH members aged 0-11	-0.00246	-0.00124	0.00596	0.00325	0.00350	0.00201
	(0.01217)	(0.01251)	(0.01277)	(0.01265)	(0.00574)	(0.00627)
HH members aged 12-17	0.02291*	0.02265*	-0.01252	-0.01196	0.01039	0.01070
	(0.01185)	(0.01173)	(0.01041)	(0.01069)	(0.00709)	(0.00683)
HH members aged 18-29	0.00009	0.00240	0.00374	-0.00139	0.00384	0.00101
	(0.01152)	(0.01355)	(0.01034)	(0.01150)	(0.00604)	(0.00688)
HH members aged 30-64	-0.01697	-0.01752	-0.02276	-0.02152	-0.03972	-0.03904

Continued on next page

Table 8 – continued from previous page

	salaried empl.		self-empl.		overall empl.	
	(1)	(2)	(3)	(4)	(5)	(6)
	(0.05212)	(0.05087)	(0.04874)	(0.04896)	(0.02847)	(0.02961)
HH members aged 65+	-0.02001	-0.01884	0.00730	0.00469	-0.01271	-0.01414
	(0.04165)	(0.04136)	(0.03573)	(0.03540)	(0.02182)	(0.02205)
father education	-0.00132	-0.00099	0.00045	-0.00030	-0.00087	-0.00128
	(0.00198)	(0.00226)	(0.00202)	(0.00213)	(0.00160)	(0.00170)
mother education	-0.00971***	-0.00904**	0.00937***	0.00787**	-0.00034	-0.00116
	(0.00341)	(0.00417)	(0.00294)	(0.00327)	(0.00181)	(0.00210)
years of delay in sec.edu	0.00440	0.00919	-0.02391**	-0.03456*	-0.01952***	-0.02537*
	(0.01275)	(0.02067)	(0.01070)	(0.01764)	(0.00702)	(0.01381)
technical edu	0.01038	-0.00031	0.02957	0.05333	0.03995***	0.05301**
	(0.03694)	(0.04459)	(0.03180)	(0.04181)	(0.01452)	(0.02678)
vocational edu	0.00818	-0.00144	0.03776	0.05914	0.04594***	0.05770**
	(0.03164)	(0.04096)	(0.02700)	(0.04065)	(0.01546)	(0.02569)
apprenticeship/PT edu	-0.06240	-0.07366	0.08533*	0.11033**	0.02292	0.03667
	(0.05336)	(0.06364)	(0.04311)	(0.05434)	(0.02842)	(0.03967)
Constant	1.41749***	1.45409***	-0.15519	-0.23651	1.26230***	1.21757***
	(0.35242)	(0.37325)	(0.32471)	(0.34533)	(0.20910)	(0.22952)
Observations	1,902	1,902	1,902	1,902	1,902	1,902
R-squared	0.04170	0.03730	0.03098	0.00332	0.05818	0.03540
Exogeneity test P-val†		0.767		0.438		0.467

*** p<0.01, ** p<0.05, * p<0.1. Standard errors between parentheses. Columns 1-6 report the results from estimating Eq. (2) by OLS (odds columns) and 2SLS (even columns). The first stage regression is reported in Table 7 (Column 5). All estimations report cluster robust standard errors by graduation year g and province of residence at graduation p ($G = 44$). Even columns report the exogeneity test for early non-employment ($y^0_{igpt_1}$).

‡ The discrete outcomes are measured at potential experience 6.

† With clustered standard errors, the exogeneity test is defined as the difference of two Sargan-Hansen statistics: one for the equation where $y^0_{igpt_1}$ is treated as endogenous and one for the equation where $y^0_{igpt_1}$ is treated as exogenous. Under the null that $y^0_{igpt_1}$ is exogenous, the statistic is distributed as $\chi^2(1)$. This statistic is not corrected for the problem of few clusters.

F Sensitivity Analysis For the Low Educated

Table 9: Effect of Interest Excluding $\overline{UR}_p$ and $minUR_{pt}$ from the Specification

Panel A: Effect of early non-employment in the structural equation:

		OLS		2SLS	
	standard errors†	robust	cluster $g * p$	robust	cluster $g * p$
outcomes:		(1)	(2)	(3)	(4)
salaried employment	coeff	-0.00169***	-0.00169***	-0.00199	-0.00199
	se	(0.00034)	(0.00041)	(0.00387)	(0.00348)
	P-val§		0.00019		0.57020
	Bootstrap P-val‡		0.00000		0.62462
	Exogeneity test P-val§§			0.937	0.931

Continued on next page

Table 9 – continued from previous page

		OLS		2SLS	
	standard errors†	robust	cluster $g*p$	robust	cluster $g*p$
outcomes:		(1)	(2)	(3)	(4)
self-employment	coeff	0.00054*	0.00054	0.00219	0.00219
	se	(0.00030)	(0.00041)	(0.00340)	(0.00313)
	P-val		0.19253		0.48668
	Bootstrap P-val		0.19219		0.52853
	Exogeneity test P-val		0.619		0.587
overall employment	coeff	-0.00115***	-0.00115***	0.00020	0.00020
	se	(0.00021)	(0.00025)	(0.00229)	(0.00164)
	P-val		0.00005		0.90232
	Bootstrap P-val		0.00000		0.91291
	Exogeneity test P-val			0.540	0.386
log earnings	coeff	-0.0269***	-0.0269***	-0.0947**	-0.0947***
	se	(0.0033)	(0.0040)	(0.0447)	(0.0354)
	P-val		2.78E-08		0.01051
	Bootstrap P-val		0		0.03003
	Exogeneity test P-val				0.0361
log hours worked	coeff	-0.0203***	-0.0203***	-0.0666**	-0.0666***
	se	(0.0024)	(0.0029)	(0.0326)	(0.0251)
	P-val		1.05E-08		0.01105
	Bootstrap P-val		0		0.03403
	Exogeneity test P-val				0.0481

Panel B: Effect of the instrument from the first stage (OLS)

outcome‡‡*:*	*standard errors:*	robust	cluster (g*p)
early non-empl.	coeff	5.0319***	5.0319***
	se	(1.6519)	(1.7139)
	P-val		0.0053
	Bootstrap P-val		0.0120
	F stat		8.620
	Bootstrap F stat††		5.84

*** p<0.01, ** p<0.05, * p<0.1. Standard errors between parentheses. Panel A reports results from estimating β in Eq. (2) on outcomes measured at potential experience 6, excluding $\overline{UR}_p$ and $minUR_{pt}$. β is the effect of one *pp* increase in $y^0_{igpt_1}$, i.e. the % of hours spent in non-employment in the first 2.5 years after graduation relative to potential total hours if one would work full-time during the whole period. For clustered standard errors, we report the P-value and the wild bootstrap P-value. Column 1-2 (3-4) show OLS (2SLS). In 2SLS the provincial unemployment rate at graduation is used as instrument for $y^0_{igpt_1}$. Panel B reports the effect of the instrument on $y^0_{igpt_1}$ in the first stage and the F statistic. For continuous outcomes we add value one before taking the log, so that non-salaried employed at the moment of evaluation are included with outcomes equal to zero after the logarithmic transformation.

† Robust indicates heteroscedastic-robust standard errors; clusters are defined by year g and province of residence at graduation p (G=44 clusters).

§ The P-value from clustered standard errors is computed using the $t(G-1)$ distribution, with G=44 (stars are reported accordingly).

‡ Bootstrap P-values are computed according to the wild bootstrap procedures explained in Appendix D for 999 repetitions.

§§ With clustered standard errors, this test is defined as the difference of two Sargan-Hansen statistics: one for the equation where $y^0_{igpt_1}$ is treated as endogenous and one for the equation where $y^0_{igpt_1}$ is treated as exogenous. Under the null that $y^0_{igpt_1}$ is exogenous, the statistic is distributed as $\chi^2(1)$.

†† Bootstrap F statistic is the F statistic corresponding to the Bootstrap P-value of the t statistic of the instrument: we rely on the equivalence between F and t distribution: with $G = 44$, $t^2(G-1) = F(1, G-1)$.

Table 10: Effect of Interest for Graduation Period 1998-2002

Panel A: Effect of early non-employment in the structural equation:					
		OLS		2SLS	
	standard errors†	robust	cluster $g*p$	robust	cluster $g*p$
Continuous outcomes:‡‡		(1)	(2)	(3)	(4)
log earnings	coeff	-0.0287***	-0.0287***	-0.1406**	-0.1406**
	se	(0.0045)	(0.0049)	(0.0666)	(0.0597)
	P-val§		6.09E-06		0.027298462
	Bootstrap P-val‡		0		0.082082082
	Exogeneity test P-val§§				0.0306
log hours worked	coeff	-0.0215***	-0.0215***	-0.1032**	-0.1032**
	se	(0.0034)	(0.0035)	(0.0492)	(0.0439)
	P-val		3.23E-06		0.027651497
	Bootstrap P-val		0		0.078078078
	Exogeneity test P-val				0.0318
Panel B: Effect of the instrument in the first stage : OLS					
outcome:	*standard errors:*	robust	cluster (g*p)		
early non-empl.(% hours)	coeff	11.9484***	11.9484***		
	se	(3.4994)	(3.4918)		
	P-val	0.00233			
	Bootstrap P-val	0.07007			
	F stat	11.70921			
	Bootstrap F stat††	3.60923			

*** p<0.01, ** p<0.05, * p<0.1. Standard errors between parentheses. Panel A reports results from estimating β in Eq. (2) on continuous outcomes measured at potential experience 6. β is the effect of one *pp* increase in $y^0_{igpt_1}$, i.e. the % of hours spent in non-employment in the first 2.5 years after graduation relative to potential total hours if one would work full-time during the whole period. For clustered standard errors, we report the P-value and the wild bootstrap P-value. Column 1-2 (3-4) show OLS (2SLS). In 2SLS the provincial unemployment rate at graduation is used as instrument for $y^0_{igpt_1}$. Panel B shows the effect of the instrument on $y^0_{igpt_1}$ in the first stage and the corresponding F statistic.

† Robust indicates heteroscedastic-robust standard errors; clusters are defined by year g and province of residence at graduation p (G=24 clusters).

‡‡ For continuous outcomes we add value one before taking the log, so that non-salaried employed at the moment of evaluation are included with outcomes equal to after the logarithmic transformation.

§ The P-value from clustered standard errors is computed using the $t(G-1)$ distribution, with G=24 (stars are reported accordingly).

‡ Bootstrap P-values are computed according to the wild bootstrap procedures explained in Appendix D for 999 repetitions.

§§ With clustered standard errors, this test is defined as the difference of two Sargan-Hansen statistics: one for the equation where $y^0_{igpt_1}$ is treated as endogenous and one for the equation where $y^0_{igpt_1}$ is treated as exogenous. Under the null that $y^0_{igpt_1}$ is exogenous, the statistic is distributed as $\chi^2(1)$.

†† Bootstrap F statistic is the F statistic corresponding to the bootstrap P-value of the t statistic of the instrument: we rely on the equivalence between F and t distribution: $t^2(G-1) = F(1, G-1)$, with $G = 24$.

Table 11: Complete OLS: Period 1994-2002 vs 1998-2002; Full vs Restricted Specification

Continuous outcomes[‡‡]:	log earnings			log hours worked		
	g94-02[§]		g98-02[†]	g94-02		g98-02
	full spec.[§§]	restricted spec.[††]		full spec.	restricted spec.	
	(1)	(2)	(3)	(4)	(5)	(6)
cluster	$g*p$	$g*p$	$g*p$	$g*p$	$g*p$	$g*p$
early non-empl (% hours)	-0.027***	-0.027***	-0.029***	-0.020***	-0.020***	-0.022***
	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.004)
UR_pe6	0.482**	0.617***	0.594	0.334**	0.432***	0.445
	(0.190)	(0.179)	(0.544)	(0.138)	(0.130)	(0.393)
lin_grad_year	-0.159	-0.114	0.811*	-0.114	-0.082	0.635*
	(0.294)	(0.284)	(0.460)	(0.213)	(0.205)	(0.333)
lin_grad_year|trend>3	0.903**	0.661*	-1.032	0.668**	0.492*	-0.795
	(0.400)	(0.360)	(0.701)	(0.292)	(0.260)	(0.503)
lin_grad_year|trend>6	-0.460*			-0.334*		
	(0.258)			(0.187)		
avg_UR_pe3-6	-1.284***	-0.826**	-1.879	-0.965***	-0.634**	-1.481
	(0.451)	(0.394)	(1.373)	(0.326)	(0.289)	(0.993)
min_UR_pe0-6	-0.572	-0.692	-3.140***	-0.313	-0.398	-2.204**
	(0.622)	(0.590)	(1.118)	(0.439)	(0.419)	(0.791)
d_province2	-1.381*	-0.913	-4.627***	-0.948*	-0.612	-3.214**
	(0.711)	(0.704)	(1.597)	(0.504)	(0.503)	(1.169)
d_province3	-2.175***	-1.756***	-4.127**	-1.577***	-1.275***	-3.024**
	(0.607)	(0.547)	(1.740)	(0.434)	(0.392)	(1.238)
d_province4	0.084	-0.109	-1.617	0.074	-0.069	-1.118
	(0.480)	(0.456)	(1.380)	(0.350)	(0.333)	(0.984)
d_province5	0.899	0.741	1.125	0.667	0.552	0.743
	(0.610)	(0.609)	(1.902)	(0.450)	(0.451)	(1.397)
lin_calend_year_prov2	0.050	0.075	0.126	0.035	0.055	0.064
	(0.139)	(0.144)	(0.307)	(0.102)	(0.106)	(0.228)
lin_calend_year_prov3	0.072	0.147	-0.135	0.059	0.114	-0.105
	(0.114)	(0.104)	(0.137)	(0.081)	(0.075)	(0.094)
lin_calend_year_prov4	-0.095	-0.007	-0.120	-0.068	-0.004	-0.102
	(0.127)	(0.109)	(0.304)	(0.093)	(0.080)	(0.220)
lin_calend_year_prov5	-0.081	-0.102	-0.078	-0.059	-0.075	-0.041
	(0.107)	(0.107)	(0.357)	(0.079)	(0.080)	(0.263)
birth cohort76	0.704	0.813	0.816	0.542	0.622	0.625
	(0.563)	(0.510)	(0.599)	(0.413)	(0.373)	(0.437)
birth cohort78	0.426	0.487	0.471	0.319	0.364	0.353
	(0.368)	(0.350)	(0.353)	(0.269)	(0.255)	(0.256)
HH members aged 0-11	-0.018	0.016	0.005	-0.046	-0.021	-0.027
	(0.112)	(0.116)	(0.136)	(0.088)	(0.090)	(0.103)
father education	0.001	-0.002	-0.017	0.000	-0.003	-0.014
	(0.026)	(0.025)	(0.041)	(0.019)	(0.019)	(0.031)
mother education	-0.105***	-0.100***	-0.103**	-0.077***	-0.073***	-0.074**
	(0.036)	(0.035)	(0.045)	(0.027)	(0.026)	(0.033)
years of delay in sec. edu	-0.054	-0.143	-0.190	-0.042	-0.106	-0.140
	(0.111)	(0.096)	(0.166)	(0.081)	(0.071)	(0.121)
general edu		-0.449	-0.161		-0.361	-0.145
		(0.299)	(0.353)		(0.226)	(0.265)

Continued on next page

Table 11 – continued from previous page

Continuous outcomes‡‡:	log earnings			log hours worked		
	g94-02§		g98-02†	g94-02		g98-02
	full spec.§§	restricted spec.††		full spec.	restricted spec.	
	(1)	(2)	(3)	(4)	(5)	(6)
live in single-parent	0.345			0.273		
	(0.463)			(0.347)		
not live with parents	0.427*			0.348*		
	(0.243)			(0.179)		
HH members aged 12-17	0.168			0.126		
	(0.118)			(0.088)		
HH members aged 18-29	0.011			0.007		
	(0.116)			(0.087)		
HH members aged 30-64	-0.033			-0.011		
	(0.407)			(0.302)		
HH members aged 65+	-0.068			-0.029		
	(0.364)			(0.268)		
technical edu	0.443			0.356		
	(0.327)			(0.247)		
vocational edu	0.457			0.368*		
	(0.277)			(0.210)		
apprenticeship/PT edu	-0.188			-0.109		
	(0.453)			(0.337)		
Constant	15.056***	12.908***	29.987***	10.779***	9.282***	22.069***
	(3.753)	(3.432)	(7.716)	(2.695)	(2.454)	(5.520)
Observations	1,902	1,902	946	1,902	1,902	946
R-squared	0.090	0.084	0.097	0.089	0.084	0.098

*** p<0.01, ** p<0.05, * p<0.1. Standard errors between parentheses. Columns 1 and 4 estimate Eq. (2) by OLS considering the graduation period 1994-2002: they are equivalent to the estimations reported in columns 1 and 3 of Table 7. Columns 2 and 5 estimate the restricted specification discussed for the sensitivity exercise in Section 4.1, based on the graduation period 1994-2002. Columns 3 and 6 estimate the same restricted specification on the graduation period 1998-2002, which is used in the second sensitivity analysis. Standard errors are clustered by graduation year g and province of living at graduation p.

‡‡ For continuous outcomes we add value one before taking the log, so that non-salaried employed at the moment of evaluation are included with outcomes equal to zero after the logarithmic transformation.

§ graduation period 1994-2002 considered.

† graduation period 1998-2002 considered: for this reason, the third graduation year spline $lin\_grad\_year|trend > 6$ is omitted.

§§ The full specification corresponds to Eq. (2).

†† The restricted specification is the one used in the second sensitivity analysis in Section 4.1.

References

Amemiya, Y. (1985). Instrumental variable estimator for the nonlinear errors-in-variables model. *Journal of Econometrics 28*(3), 273–289.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Arulampalam, W. (2001). Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages. *The Economic Journal 111*(475), 585–606.

Baltagi, B., C. Kao, and S. Na (2013). Testing for cross-sectional dependence in a panel factor model using the wild bootstrap F test. *Statistical Papers 54*(4), 1067–1094.

Baum, C. F., M. E. Schaffer, and S. Stillman (2003). Instrumental variables and GMM: Estimation and testing. *Stata Journal 3*(1), 1–31(31).

Beaudry, P. and J. DiNardo (1991). The Effect of Implicit Contracts on the Movement of Wages Over the Business Cycle: Evidence from Micro Data. *Journal of Political Economy 99*(4), 665–688.

Bell, D. N. and D. G. Blanchflower (2010). Youth Unemployment: Déjà Vu? IZA Discussion Paper 4705, IZA, Bonn.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics 90*(3), 414–427.

Clark, A. E. (2003). Unemployment as a Social Norm: Psychological Evidence from Panel Data. *Journal of Labor Economics 21*(2), 323–351.

Clark, A. E., Y. Georgellis, and P. Sanfey (2001). Scarring: The Psychological Impact of Past Unemployment. *Economica 68*(270), 221–241.

Cockx, B. and C. Ghirelli (2015). Scars of Recessions in a Rigid Labor Market. IZA Discussion Paper 8889, IZA, Bonn. Supplementary Online Appendix: http://users.ugent.be/~bcockx/Ascars.pdf.

Cockx, B. and M. Picchio (2012). Are short-lived jobs stepping stones to long-lasting jobs? *Oxford Bulletin of Economics and Statistics 74*(5), 646–675.

Cockx, B. and M. Picchio (2013). Scarring effects of remaining unemployed for long-term unemployed school-leavers. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 176*(4), 951–980.

Davidson, R. and J. MacKinnon (2010). Wild Bootstrap Tests for IV Regression. *Journal of Business & Economic Statistics 28*(1), 128–144.

Davidson, R. and J. G. MacKinnon (2008). Bootstrap inference in a linear equation estimated by instrumental variables. *Econometrics Journal 11*(3), 443–477.

Devereux, P. J. and R. A. Hart (2007). The Spot Market Matters: Evidence on Implicit Contracts from Britain. *Scottish Journal of Political Economy 54*(5), 661–683.

Ellwood, D. T. (1982). Teenage Unemployment: Permanent Scars or Temporary Blemishes? In R. B. Freeman and e. David A. Wise (Eds.), *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*, pp. 349–390. Chicago: University of Chicago Press.

Gangji, A. V. and R. Plasman (2007). The Matthew effect of unemployment: how does it affect wages in Belgium. DULBEA Working Papers 07-19.RS, Université Libre de Bruxelles (ULB).

Gartell, M. (2009). Unemployment and subsequent earnings for Swedish college graduates. A study of scarring effects. Arbetsrapport 2009:2, Institute for Futures Studies, Stockholm.

Genda, Y., A. Kondo, and S. Ohta (2010). Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States. *Journal of Human Resources 45*(1), 157–196.

Gervais, M., N. Jaimovich, H. E. Siu, and Y. Yedid-Levi (2014). What Should I Be When I Grow Up? Occupations and Employment over the Life Cycle and Business Cycle. NBER Working Paper Series 20628, NBER, MA.

Gregg, P. (2001). The Impact of Youth Unemployment on Adult Unemployment in the NCDS. *The Economic Journal 111*(475), 626–653.

Gregg, P. and E. Tominey (2005). The wage scar from male youth unemployment. *Labour Economics 12*(4), 487–509.

Gregory, M. and R. Jukes (2001). Unemployment and Subsequent Earnings: Estimating Scarring among British Men 1984–94. *The Economic Journal 111*(475), 607–625.

Hagedorn, M. and I. Manovskii (2013). Job Selection and Wages over the Business Cycle. *American Economic Review 103*(2), 771–803.

Hausman, J. (2001). Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left. *Journal of Economic Perspectives 15*(4), 57–67.

Hausman, J. A., W. K. Newey, and J. L. Powell (1995). Nonlinear errors in variables Estimation of some Engel curves. *Journal of Econometrics 65*(1), 205–233.

Kawaguchi, D. and T. Murao (2014). Labor Market Institutions and Long-Term Effects of Youth Unemployment. IZA Discussion Paper 8156, IZA, Bonn.

Kilponen, J. and T. Santavirta (2010). New Evidence on Implicit Contracts from Linked Employer–Employee Data. *Scandinavian Journal of Economics 112*(4), 864–883.

Kletzer, L. G. and R. W. Fairlie (2003). The long-term costs of job displacement for young adult workers. *Industrial and Labor Relations Review 56*(4), 682–698.

Kroft, K., F. Lange, and M. J. Notowidigdo (2013). Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment. *The Quarterly Journal of Economics*.

Lockwood, B. (1991). Information Externalities in the Labour Market and the Duration of Unemployment. *Review of Economic Studies 58*(4), 733–53.

MacKinnon, J. G. and M. D. Webb (2015). Wild Bootstrap Inference for Wildly Different Cluster Sizes. Queens Economics Department Working Paper No. 1314, Queen's University.

Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *The Review of Economics and Statistics 72*(2), 334.

Mroz, T. A. and T. H. Savage (2006). The Long–Term Effects of Youth Unemployment. *The Journal of Human Resources 41*(2), 259–293.

Neumark, D. (2002). Youth Labor Markets in the United States: Shopping Around vs. Staying Put. *Review of Economics and Statistics 84*(3), 462–482.

Oreopoulos, P., T. von Wachter, and A. Heisz (2012). The Short- and Long-Term Career Effects of Graduating in a Recession. *American Economic Journal: Applied Economics 4*(1), 1–29.

Oreopoulos, P., T. v. Wachter, and A. Heisz (2008). The Short- and Long-Term Career Effects of Graduating in a Recession: Hysteresis and Heterogeneity in the Market for College Graduates. IZA Discussion Paper 3578, IZA, Bonn.

Pissarides, C. A. (1992). Loss of Skill During Unemployment and the Persistence of Employment Shocks. *The Quarterly Journal of Economics 107*(4), 1371–1391.

Schmillen, A. and M. Umkehrer (2013). The scars of youth: effects of early-career unemployment on future unemployment experience. IAB Discussion Paper 201306, IAB, Nürnberg.

Sonar (2003). Hoe maken de jongeren de overgang van school naar werk? Basisrapportering Cohorte 1978 (eerste golf). Technical report, Leuven: Steunpunt WAV.

Sonar (2004a). Hoe maken de jongeren de overgang van school naar werk? Basisrapportering Cohorte 1976 (tweede golf). Technical report, Leuven: Steunpunt WAV.

Sonar (2004b). Hoe maken de jongeren de overgang van school naar werk? Cohorte 1980 (eerste golf). Technisch rapport. Technical report, Leuven: Steunpunt WAV.

Stock, J. and M. Yogo (2005). Testing for Weak Instruments in Linear IV Regression. In D. W. Andrews (Ed.), *Identification and Inference for Econometric Models*, pp. 80–108. New York: Cambridge University Press.

Topel, R. H. and M. P. Ward (1992). Job Mobility and the Careers of Young Men. *The Quarterly Journal of Economics 107*(2), 439–79.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2015/3082/8