ISSN 2704-7407

UNIVERSITÀ CATTOLICA DEL SACRO CUORE

Dipartimento di Economia e Finanza

Working Paper Series

Education, health and health-related behaviors: Evidence from higher education expansion

Massimiliano Bratti, Elena Cottini, Paolo Ghinetti

Working Paper n. 114

February 2022

UNIVERSITÀ CATTOLICA del Sacro Cuore

Education, health and health-related behaviors: Evidence from higher education expansion

Massimiliano Bratti
Università degli Studi di Milano

Elena Cottini
Università Cattolica del Sacro Cuore

Paolo Ghinetti
Università del Piemonte Orientale

Working Paper n. 114
February 2022

Dipartimento di Economia e Finanza
Università Cattolica del Sacro Cuore
Largo Gemelli 1 - 20123 Milano – Italy
tel: +39.02.7234.2976 - fax: +39.02.7234.2781
e-mail: dip.economiaefinanza@unicatt.it

Education, health and health-related behaviors: Evidence from higher education expansion*

Massimiliano Bratti[1,4,5,6], Elena Cottini[2], and Paolo Ghinetti[3]

[1] *Univesità degli Studi di Milano*
[2] *Università Cattolica di Milano*
[3] *Università del Piemonte Orientale*
[4] *Institute of Labor Economics (IZA)*
[5] *Global Labor Organization (GLO)*
[6] *Centro Studi Luca d'Agaliano (LdA)*

*Contact information: Massimiliano Bratti, Department of Economics Management and Quantitative Methods (DEMM), Università degli Studi di Milano, via Conservatorio 7, Milan (MI), massimiliano.bratti@unimi.it; Elena Cottini, Department of Economics and Finance, Università Cattolica di Milano, Largo Gemelli 1, Milan (MI), Italy, elena.cottini@unicatt.it; Paolo Ghinetti, Dipartimento di Studi per l'Economia e l'Impresa, Università del Piemonte Orientale, via Perrone 18, Novara (NO), Italy, paolo.ghinetti@uniupo.it.

Abstract

This study throws light on the potential non-linear effects of education on individual health and health-related behaviors, finding a strong role for higher education. Using an instrumental variables (IVs) strategy, which leverages changes in within-province between-municipality college proximity across birth cohorts, we demonstrate that higher education affects individual health-related behavior. By contrast, IVs estimates based on a compulsory schooling age reform show mostly non-significant effects. Our results point to a complex link between education and health. On the one hand, higher education channels individuals into some healthy behaviors and better health outcomes namely healthy eating, more physical activity and a lower risk of obesity. On the other hand, it also appears to increase the prevalence of certain unhealthy behaviors, such as greater smoking and drinking prevalence and higher cigarettes consumption. Albeit effects are generally similar across genders, except in few cases (e.g. smoking behavior), our analysis highlights heterogeneous effects by age and helps explain potential differences in results reported in past quasi-experimental studies in which the cohorts affected by the educational reforms used for identification are observed at given ages and not over an individual's entire life-cycle.

Keywords: education; health; higher education expansion; health-related behavior.
JEL codes: I12, I24

1 Introduction

Education not only contributes to increasing individual labor incomes and labor market opportunities (Card, 2001), but also has important non-pecuniary returns. It affects individual health-related behavior and health status. The seminal work of Grossman (e.g. Grossman, 2006) gave some theoretical foundations to the potential role of educational investments on health. Grossman demonstrated how investments in education can change not only an individual's ability to produce a higher stock of health starting from the same inputs (*productive efficiency hypothesis*), but can also affect the choice of inputs into the health production function (*allocative efficiency hypothesis*). Yet, investigating these effects poses an important empirical challenge: unobserved individual characteristics such as the levels of innate ability or of the inter-temporal discount rates (i.e. an individual's future orientation) may affect investments in both education and health, making the correlation between the two outcomes only spurious (the *third variable hypothesis*; see Fuchs, 1982; Farrell and Fuchs, 1982).

Given this fundamental problem of identification, empirical researchers have resorted to quasi-natural experiments, typically reforms which have increased individuals' educational achievement, to investigate the health returns to education. Most studies leverage compulsory schooling laws (CSLs) in fuzzy Regression Discountinuity Designs (RDDs). The list is long — see Eide and Showalter (2011) and Galama et al. (2018) for some comprehensive reviews —, their results are mixed and often vary across health outcomes. Some studies are based on CSL reforms that had, in some cases, small effects on average educational achievement (Lleras-Muney, 2005). For the case of the US (Lleras-Muney, 2005), for instance, the result of a protective effect of education on mortality is not robust to the inclusion of state and birth cohort fixed effects or state-specific trends (Mazumder, 2008; Black et al., 2015), or effects are imprecisely estimated, leading to uncertain conclusions (Fletcher, 2015). Although in some cases low precision can be attributed to weak instruments (see also Albouy and Lequien, 2009), in others CSLs had large and significant effects on educational achievement, but the results remain mixed. Clark and Royer (2013) exploit two CLS reforms in the UK, of 1947 and 1972, the first of which affected a high share individuals (Oreopoulos, 2006). The authors use a convincing fuzzy RDD in which reform-eligible individuals are identified by month of birth (and not just by year of birth like in many other papers) but do not find any significant effect on a host of health outcomes, including mortality. CSL reforms in Sweden significantly improved educational achievement in the population. However, while evidence from more recent CSL reforms point to small and insignificant effects on mortality (Meghir et al., 2018; Lager and Torssander, 2012), a study using an older reform finds sizeable negative effects (Fischer et al., 2013). Importantly, the effect may also depend on the ages at which the affected individuals are observed in the data, and this may partly explain the different results, since mortality and health problems tend to emerge later in life. Consistent with this argument, focusing on older individuals (81-87 years old), van Kippersluis et al. (2011) find significant negative effects on mortality of educational improvements generated by a CSL in the Netherlands. Mixed results are not only observed for mortality but also for other health outcomes and behaviors. On a literature review on current obesity and current smoking, Galama et al. (2018) report non-significant effects on smoking from studies using the UK CSLs (Clark and Royer, 2013; Davies et al., 2016), with similar very weak evidence for obesity and just a couple of studies finding significant effects (Kemptner et al., 2011; James, 2015) that go in the expected direction, i.e. education reducing smoking or obesity.

Despite their differences, all of studies leveraging CSL reforms have three key features in common: (i) they focus on educational reforms that increased the education of potentially "low achievers", i.e. the marginal students who were retained in education only because of the extension of the compulsory school age obligation; (ii) they focus on relatively low levels of education, as compulsory schooling age generally falls within the primary, lower secondary — or for more recent reforms the initial years of upper secondary — schooling levels; (iii) they are not able to provide any evidence on the effects of tertiary education, unless the health returns to education are assumed to be linear, i.e. one more year of primary or secondary schooling has the same health return of one more year of tertiary education.

The last element is especially problematic since it makes these studies not very informative for current policy making in the presence of non-linear health returns to education. Indeed, in recent years secondary school attainment is almost universal in developed countries, which have tried to increase educational achievement in the population especially by improving participation in tertiary education. Hence, for policy makers in those countries, it is crucial to gain a better understanding of the potential health returns of the highest levels of education. Higher education may bring very different health returns compared to changes in CSLs. As stressed by Galama et al. (2018), part of the returns to increasing education may not simply stem from increasing its duration but from changes in the peer group or in the curricula and the skills acquired that transitioning from an educational level to the next entails (e.g. from lower to upper secondary, or from upper secondary to university). This seems to be confirmed by the studies which exploit reforms changing school tracks or other quasi-natural experiments that affect an individual's likelihood of completing higher education (de Walque, 2007; Kenkel et al., 2006; Grimard and Parent, 2007; Jürges et al., 2011), which tend to find stronger effects on health compared to those leveraging CSLs.

For this reason, it is key to provide further evidence on whether and to what extent expanding higher education and increasing educational achievement may bring improvements in individual health. In this paper, we focus on the Italian case. Italy is an interesting case study for three reasons. First, Italy's Higher Education (HE) system underwent a process of intense supply expansion, which was not equally spread across regions and birth cohorts. Overall, the number of universities increased from 39 in 1950, 51 in 1970, 62 in 1990 and 72 in 2010 (excluding on-line only universities and institutions offering only post-graduate programs). The proliferation of university campuses (i.e. local branches of HE institutions) was even larger. They were almost tripled in 60 years, from 44 in 1950 to 118 in 2010. Such new higher education premises resulted into higher college proximity and accessibility of higher education: of the about 8,000 Italian municipalities, those with at least one campus increased from 32 in 1950 (overall, the capitals of the 20 Italian regions, NUTS-2 level,[1] plus other smaller cities with a long-standing university tradition such as Pavia or Urbino) to 48 in 1970, 62 in 1990 and 96 in 2010 (basically, at least one campus in each capital of Italian provinces, NUTS-3 level). Further, each campus typically offers programs in several disciplines (see Section 4.2 below for additional details). This process was unrelated to local demand shocks or other factors which may be correlated with individual health, but was mainly driven by factors internal to university governance (Oppedisano, 2011).

Second, Italy is a country with low tertiary education achievement compared to other similarly developed countries. According to OECD statistics, Italy ranks 39th over 43 countries with a 28% of youth in the 25-34 age group having tertiary education, against the OECD average of 45%, and 33% for Germany, 50% for the US and 52% for the UK.[2] It is therefore a country for which there are large margins of improvement in terms of increasing tertiary education achievement, and for which the health returns to expanding higher education are also expected to be high.

Finally, in Italy college accessibility and distance represent important factors for the choice of attending higher education (Pigini and Staffolani, 2016; Bratti and Verzillo, 2019). This is important for our identification strategy, which is described below and leverages geographical variation in college proximity and its changes over time as a source of presumably exogenous variation in individuals' tertiary education achievement.

Our study uses several waves of a cross-section annual survey representative of the Ital-

[1] Italian regions (*regioni*) are level 2 Nomenclature of Territorial Units for Statistics' administrative units (NUTS), while Italian provinces (*province*) are level 3 NUTS administrative units and are comparable to US and UK counties.

[2] Source: https://data.oecd.org/eduatt/population-with-tertiary-education.htm, accessed April 17th, 2021.

ian population, the "Aspects of everyday life" (*Aspetti della vita quotidiana*, AVQ hereafter) administered by the Italian National Statistical Institute (ISTAT), which span the period 2001-2016. The survey is rich of information related to health-related behaviors (dietary habits, exercising, smoking and drinking), and also provides information on a limited number of health outcomes (BMI, obesity). Health-related behaviors are informative as they represent an important mediator of health outcomes, as predicted by Grossman's allocative efficiency hypothesis (Grossman, 2006) and shown by the past literature (Brunello et al., 2016).

Our paper makes three main contributions to the extant literature. First, unlike most previous literature based on CSLs, we focus on tertiary education and leverage variation in individual years of schooling that is generated by higher education expansion in Italy to address the endogeneity of education. This follows the idea that expansion of higher education, by increasing college proximity, reduces the cost of enrolling in higher education and attracts more individuals into tertiary education (Card, 1995; Currie and Moretti, 2003). Our paper demonstrates important causal effects of tertiary education on a wide range of health-related behaviors and on body weight-related health outcomes (BMI, obesity).

Second, we compare the health returns generated by increases in tertiary education attainment with those estimated leveraging a CSL that has been already exploited in the literature, namely the 1962 reform that raised compulsory schooling age from 11 up to 14 (Brandolini and Cipollone, 2002; Atella and Kopinska, 2014). We show that the returns estimated with college proximity and CSLs are very different, with gains in years of education induced by the latter generally producing no important effect on health-related behaviors and outcomes.

Third, we report evidence on age-specific health returns to tertiary education. Interestingly, this helps explaining potential differences in results reported in past studies in which the cohorts affected by the educational reforms used for identification are observed at given ages and not over the entire life-cycle.

The remainder of the paper is organized as follows. In the next section we briefly summarize the past literature on the effects of tertiary education on health outcomes and behavior. Section 3 discusses our empirical strategy and identification. The data and the main variables are described in Section 4. The main results of our analysis and some robustness checks are discussed in Section 5. Section 6 reports an analysis of effect heterogeneity by gender and age. Section 7 summarizes the main findings and draws conclusions.

2 Past literature on the effect of tertiary education on health and health-related behavior

The literature focusing on the effects of tertiary education on individual health outcomes and behaviors for the general population counts only a handful of studies. Our paper adds to this scant literature.[3]

Using a structural model and plausible exclusion restrictions (local unemployment, college present in county and local college tuition), Heckman et al. (2018) estimate the effect of college graduation over "some college" on smoking and "health limits work" in the US. In both cases the authors find strong causal effects. While effects on smoking do not appear to vary substantially by individual cognitive ability, those on "health limits work" increase in absolute value with the level of ability.

Kamhöfer et al. (2019) estimate the effects of tertiary education on physical and mental health leveraging the expansion of higher education (i.e. number of enrolled students over inhabitants at county level) in Germany. Importantly, they allow for differential effects according to individual-level unobservables, by providing marginal treatment effects. They report average positive effects on physical health but not on mental health, and that even when the average effect is positive, for 30%-40% of the population effects are null. This last result leads them to conclude that any further expansion of higher education is unlikely to bring substantial health returns in Germany.

Janke et al. (2020) leverage both the 1972 CSL reform and the expansion of higher education in the UK to estimate the causal effect of years of education on individual chronic diseases. While the CSL reform is used within a traditional fuzzy RDD framework like it is commonly done in the literature,[4] their specification for higher education expansion is less standard. In the latter, the effect of education is identified by the coefficient on the average years of schooling of different birth cohorts, after controlling *inter-alia* for age and period fixed effects, a cohort-specific unemployment rate and a cohort linear trend. As the authors acknowledge (p. 11), their estimates are valid under the assumption that the control variables included are able to purge out the cohort-specific educational achievement from any cohort-level unobservable also affecting health. For both reforms, the authors do not find statistically significant effects on the probability of having a chronic health condition, having a limiting chronic health condition and the number of chronic health conditions. On the contrary, they find evidence of a protective role of education on cardiovascular disease

[3] Other studies, exploiting CSLs, have been surveyed in the Introduction.

[4] In which the reform provides an exogenous shifter for years of schooling.

and diabetes.

Cowan and Tefft (2020) exploits the higher education expansion at the state level (2- and 4-year higher education institutions per capita) in the US and find positive effects for the increased number of 2-year college on physical exercise and self-reported health, and negative effects on smoking. Interestingly, they only report intention-to-treat effects, that is they estimate the effect of college supply expansion and not of achieving higher education.[5]

Compared to the existing studies, like in Cowan and Tefft (2020) our identification strategy is based on higher education supply expansion, although we measure college proximity at a much higher level of granularity (municipality level),[6] and it is presumably more robust than that adopted by Janke et al. (2020), in which educational expansion is measured by average years of schooling by birth cohort, or by Kamhöfer et al. (2019) that use for identification the number of enrolled students divided by total inhabitants, in both cases potentially conflating education demand with supply factors. Like Janke et al. (2020), we provide evidence both from CSLs and from university education expansion, but we focus on health-related behaviors and body weight outcomes and not on chronic diseases.

3 Empirical strategy

We estimate the causal effect of education on health behaviors and outcomes using instrumental variables (IVs). Similarly to Janke et al. (2020), which provide evidence from quasi-natural experiments that increased educational attainment at different points in the distribution of years of education, we use two alternative identification strategies. In the first, we exploit reasonably exogenous variation in college proximity across birth cohorts using an original dataset on the supply of higher education in Italy at the municipality level Cottini et al. (2019). The high geographical granularity of the data enables us to build measures of college proximity at the individual level for several birth cohorts. Details on the HE institutional setting in Italy and on the construction of indicators of college proximity are given in Section 4.2. In the second identification strategy, the source of exogenous variation is a reform that increased compulsory schooling from age 11 to age 14 (from 5 to 8 years of schooling, corresponding to “middle school” or “lower secondary education”), starting

[5] Some studies have instead focused on specific sub-populations, such as those whose college achievement was increased by the Vietnam draft lottery in the US. These studies report negative effects on current smoking of the affected individuals but are hardly generalizable to the whole population (de Walque, 2007; Grimard and Parent, 2007). Currie and Moretti (2003) focus on mothers in the US and report that education reduces smoking of new mothers and improves child health.

[6] Moreover, we provide average treatment effects on the treated (based on two-stage least squares) while Cowan and Tefft (2020) only provide reduced form estimates.

to be effective since October 1963 (Law n. 1859, 31st December 1962) and that affected individuals born in 1949 or later.[7]

We present results based on compulsory schooling in addition to those obtained by college proximity for two reasons. First, they complement and facilitate the comparison with the existing evidence, which is based for the most part on CSLs. Second, they allow to investigate whether health returns are homogeneous across levels of education. If this were the case, using IVs strategies with instruments that, as in our case, act on compliers located at different points in the years of education's distribution (left tail for CSLs and right tail for college proximity), we should obtain similar results (Local Average Treatment Effects — LATE — interpretation).

In our baseline estimates, we use years rather than levels of education, as most studies that estimated causal effects of schooling did. This is the popular one-factor human capital model, where the marginal return from each additional year is constant. For the model identified by college proximity, we also estimate an additional specification with a dichotomous indicator for tertiary (versus upper secondary) education on the restricted sample of individuals with at least an upper secondary school diploma (i.e. excluding those with primary and lower secondary education), a specification that enables us to relax the linearity assumption.

In the baseline specification, we estimate the effect of educational attainment (i.e. years of schooling) on health-related behaviors and a limited number of health outcomes as follows:

$$y_{ipt} = \beta_0 + \beta_1 e_{ipt} + \beta_2 \mathbf{X}_{ipt} + D_{pt} + \epsilon_{ipt} \tag{1}$$

where y_{ipt} are health-related outcomes and behaviors of individual i surveyed at time t living in province p. e_{ipt} are individual's years of education or a college degree dummy, $\mathbf{X}_{ipt}$ individual controls, in particular a gender dummy, age and its square. D_{pt} are current province of residence by survey year fixed effects (FE).

When IVs are used (implemented using Two-Stage Least Squares, 2SLS), the first stage where college proximity dummies are used as instruments for years of education reads as:

$$e_{ipt} = \alpha_0 + \sum_k \alpha_{1k} Z_{iptk} + \alpha_2 \mathbf{X}_{ipt} + D_{pt} + u_{ipt} \tag{2}$$

where Z_{iptk} are dummies for different levels of proximity to the closest university premises

[7] Until 1963, only primary school was compulsory for children aged 6 to 11. The 1st October 1963, the leaving school age was raised to 14, thus adding three further years of compulsory lower secondary education. If the individual incurred in failures and repetitions, he or she was allowed to drop out at 14, even without finishing the junior high school. This reform has been already employed in the literature. See Brandolini and Cipollone (2002), Brunello et al. (2013), Atella and Kopinska (2014) for additional institutional details.

at age 19 (when the university choice is typically made) based on current municipality of residence. Geographical proximity is measured from the closest administrative unit where higher education is provided. k is ordered, increasing in distance and equal to: a) the municipality of residence; b) neighboring (i.e. bordering with that of residence) municipalities; c) the same province but not b); d) neighboring provinces; e) same region but not d); f) neighboring regions.[8] This means that the pattern of college proximity also varies across birth cohorts due to higher education expansion over time. For simplicity and to save on notation, we leave implicit this birth cohort dependence in Eq. (2).

Controlling for province by survey year fixed effects, which captures local time-varying factors that may affect both health status and behavior and educational attainment (e.g. current quantity or quality of health infrastructures), we *de facto* only exploit within-province variation in college proximity across municipalities and birth cohorts in each survey wave. A potential caveat in our analysis is that the survey data on health outcomes and behaviors only provide the current municipality of residence and not that at the time of making the college choice (age 19). This is not an uncommon issue in the literature, since information on individual birthplaces is often not publicly released (especially in individual health data) for confidentiality reasons (cf. Currie and Moretti, 2003). Yet, we asked to the data provider the access to an indicator for the individual currently residing in the municipality of birth and we are able to replicate the analysis on the sample of "stayers", potentially limiting measurement error and endogenous mobility concerns.[9] In our data, 42% of individuals currently reside in the municipality of birth.

Identification hinges on Z_{iptk} being exogenous, i.e. uncorrelated with individual i health or health-related behaviors, or factors affecting them. College proximity was originally proposed to estimate the returns to education (Card, 1995), and, with some variants in the literature, has been extensively used as a reasonably valid exclusion restriction in studies on the non-pecuniary returns to education (Currie and Moretti, 2003; Heckman et al., 2018; Kamhöfer et al., 2019). We discuss in more detail the exogeneity assumption in subsection 4.2.

We use Eq. (1) also to estimate the model using the CSL reform. Here, including age

[8] On the one hand, more precise distance indicators, e.g. based on travel times, are not available for the past (but can only be computed according to the current transportation network). On the other hand, proximity based on administrative units may be more informative than a measure based on geodesic distance, as the bus/rail transportation network is more likely to be integrated within the same administrative unit (e.g. within the same municipality vs. outside the municipality, within the same province vs. outside the province, etc.) and travel times to be non-linear in distance.

[9] E.g. healthier individuals might have moved to municipality with a higher HE supply to enroll in tertiary education.

and its square as well as province by survey year fixed effects in the estimated models allows us to control for secular changes in health and education, both in the first and second stage, and it is comparable to including a polynomial in birth cohort.[10] This is consistent with a fuzzy RDD approach, although, differently from Janke et al. (2020), we cannot use RDD, since in our case some cohorts are only partially treated by the CSL reform. This happens because exposure to treatment (the reform) is heterogeneous across different birth cohorts depending on the year of birth: cohorts born before 1949, in 1949, 1950 and 1951 or later are attributed 5, 6, 7, 8 years of compulsory schooling, respectively (our strategy is similar to Brunello et al., 2013).

The first stage is in this case:

$$e_{ipt} = \alpha_0 + \alpha_1 C_{ipt} + \alpha_2 X_{ipt} + D_{pt} + u_{ipt} \tag{3}$$

where C_{ipt} is the number of years of compulsory schooling of individual i that varies depending on year of birth (see Brunello et al., 2013). In analogy with the RDD literature, this model is estimated on a reduced sample including only birth cohorts around the first affected by the reform: we use a bandwidth of 10 years around the first affected cohort, considering only individuals born between 1940 and 1960.

4 Data and descriptive statistics

4.1 Individual health and education data

The sources of individual-level data used in our empirical analysis is ISTAT's AVQ survey for the period that goes from 2001 to 2016, except for 2004 for which the data are not available. The AVQ survey is a repeated cross section on a yearly nationally representative sample of the Italian population. It collects information on many aspects of everyday life at individual and household level. Together with basic demographic characteristics, for individuals aged at least 18, also information about self-reported height and weight, smoking, alcohol consumption, dietary regimes and physical activity are collected. Not all the outcomes are available in each year.

The set of health and health-related behaviors includes healthy habits (such as physical activity, consumption of fruit and vegetables, water, breakfast, consumption of junk food

[10] Indeed, birth year$_i$ = $t -$ age$_{it}$. After we include a polynomial in age and province-survey year FEs (which are collinear to a trend in survey year), the model does not allow for a polynomial in birth cohort too.

and gas beverages), body mass index (BMI),[11] alcohol consumption and smoking behavior.[12] A detailed description of all behaviors and health outcomes used in the empirical analysis and how they are constructed is reported in detail in the Appendix in Table A-1.

We use two measures for an individual's educational achievement. The first is years of schooling based on the highest attained degree. This is our benchmark, because years of education are generally used in the CSLs literature.[13] The second is a dichotomous indicator for holding a tertiary education degree (bachelor or higher).

Our sample consists of individuals aged between 25 and 85 years. We impose a lower cut-off on age because individuals younger than 25 are unlikely to have completed tertiary education in Italy, and a higher cut-off at 85 as differential mortality by education is likely to be an issue the older the individuals. Table 1 reports the description of sample summary statistics for the main dependent and independent variables used in our analysis. Half of the sample is female, and the average years of education are around 10, while only 10 percent of the sample achieved tertiary education. With respect to healthy habits, around 40 percent of individuals performs physical activity at least 1-2 times a week and eats fruits and vegetables at least once a day. 70 percent of individuals has breakfast every day, does not drink more than one glass of wine/beer per day, while nearly 90 percent does not eat junk food or drinks gas beverages. Half of the sample ever smoked, while one out of 4 currently smokes (with an average daily consumption of cigarettes of around 3). Only 10 percent of the sample is obese.

4.2 Geographical location of university premises and college proximity

The source of data that we use to construct our college proximity indicators is an original dataset on the History of Italian Universities (HIU), that contains detailed information on institutions providing higher education in Italy from 1861 (year of foundation of the Kingdom of Italy) to 2010 (see Cottini et al., 2019). College proximity is measured as the distance

[11] BMI that is defined as the weight in kilograms divided by the height in square meters. According to this index, the international standards classify individuals as obese for a BMI over 30.

[12] For alcohol and smoking consumption-related variables there are top coding limits in answering survey questions. For example, the maximum number of cigarettes smoked that can be reported is 99. For this reason, we operate a minimum trimming to the original data, excluding observations in the upper 0.1 percent of the distribution of such variables.

[13] The AVQ data contain information on individual schooling attainment expressed in terms of the highest degree completed. We converted that into years of schooling depending on the legal duration of each educational cycle. For example primary education corresponds to 5, lower secondary to 8, upper secondary to 13, post-Bologna reform bachelor degree to 16 years of education, etc.

Table 1: Summary statistics

Variables	**Mean**	**Std.Dev.**	**Min.**	**Max.**	**N**
Demographics					
Age	51.867	15.994	25	85	533577
Female	0.522	0.5	0	1	533577
Years of education	9.750	4.352	0	21	533577
Tertiary education	0.103	0.304	0	1	533577
Healthy habits					
Physical activity	0.424	0.494	0	1	531674
Fruit and vegetables	0.366	0.482	0	1	525357
Fruit and vegetables (pieces)	2.305	1.425	0	10	448181
Water	0.509	0.5	0	1	314050
Breakfast	0.697	0.460	0	1	525759
No junk food	0.860	0.347	0	1	445293
No gas beverages	0.895	0.307	0	1	513653
Weight					
Obesity	0.106	0.308	0	1	533577
Body Mass Index (BMI)	25.116	3.84	15.5	41.09	533577
Alcohol					
Light alcohol consumption	0.676	0.468	0	1	525303
Number units of alcohol a day	0.684	1.291	0	9	448161
Never consumes spirits	0.803	0.398	0	1	519443
Smoking					
Ever smoked	0.479	0.5	0	1	524602
Currently smokes	0.226	0.418	0	1	524602
N. of cigarettes	2.961	6.650	0	60	521260

from the closest university's administrative teaching unit, *Facoltà*, which is equivalent to University Schools or Departments in other countries. Data collection ends in 2010 because the Law 30/12/2010, n. 240 (the so called "Gelmini reform" after the name of the Public Education Minister in office at that time), removed the *Facoltà* from the governance of public universities.[14] HIU includes the study field (or college major) of the *Facoltà* (Economics, Humanities, Engineering, etc.),[15] the municipality where it is located, the year when the *Facoltà* was formally established as a provider of higher education.[16] We focus on HE institutions that deliver standard BSc education, those specialised in post-graduate education

[14] Since 2011, as in many other countries the governance is based on Departments, for both teaching and research activities. Before 2010, the Departments were responsible for research only.

[15] However, the information on college major is not provided in the AVQ data.

[16] HIU also includes name and municipality of each university, its governance structure (private or public). All changes over time, e.g. to the governance of the *Facoltà* or the university, or in the government assessments are recorded in the data.

or enrolling foreign students only are excluded from the sample also in light of the limited number of students enrolled and the nature of our analysis (e.g. the way college proximity is defined makes sense only for domestic students).

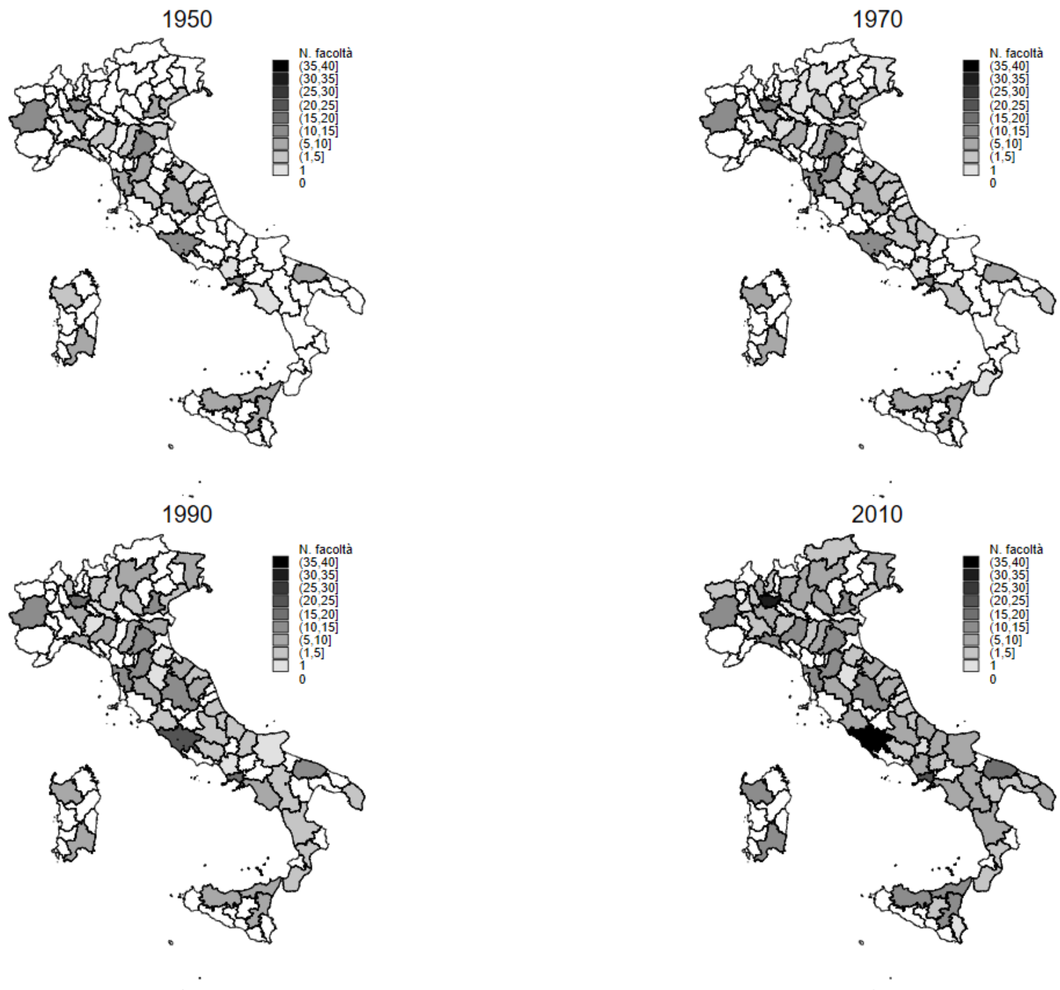
Our HIU sample includes 574 faculties (in 71 universities) registered in the Italian territory between 1861 and 2010. We use HIU to construct a panel dataset at the municipality level that, for each year, counts the number of *Facoltà* as a measure of HE supply. The data also record the province (NUTS-3) and the region (NUTS-2) where the municipality is located. We then use contiguity matrices of territorial units provided by ISTAT to match the municipality of residence of individuals in AVQ surveys with information on the geographical distribution of *Facoltà* when they were nineteen years old: whether one or more *Facoltà* were located in the same municipality of residence, or in a neighbor municipality (based on territorial borders), and so on, as described in Section 3. Given the sample selection we made on individual age (the oldest individuals are 85 years old), the relevant information on the territorial distribution and proximity of higher education providers starts from the mid-1930s.

Figure 1 presents an historical overview on the number of *Facoltà* at the province level every 20 years starting in 1950. We clearly see an expansion of tertiary education especially from the 70s, a pattern which reflects the development of tertiary education described in institutional reports and methodological studies, see e.g. OECD (2008), OECD (2014). Figure 2 shows instead the administrative borders of the municipalities in which the *Facoltà* are hosted.

As already mentioned, our instruments are defined in terms of proximity (six categories) of the municipality of residence from the nearest *Facoltà*. Considering the whole period, for about 21.8 percent of individuals the closest *Facoltà* was in the same municipality and for 11 percent in neighboring municipalities. For about 30 percent of individuals the nearest *Facoltà* was in the same province (outside the residence municipality and its neighbors), and for one third of them (33 percent) in neighboring provinces. Only for the remaining 4 percent of individuals the distance from the closest *Facoltà* was larger (same region or further away).

The main sources of identification for college proximity comes from within-province differences across municipalities (since we include province-by-year fixed effects) and variation over time in HE supply. As to the latter, one potential threat to identification may come from local-level unobservables that are correlated with both higher education provision and individual health, invalidating our strategy. In particular, the HE supply may increase more where individuals are healthier — since in those places the expected demand for higher education is higher, given that its returns are reaped over the entire life cycle —, for example

Figure 1: Expansion of Higher Education in Italy (N. *Facoltà*): Provinces

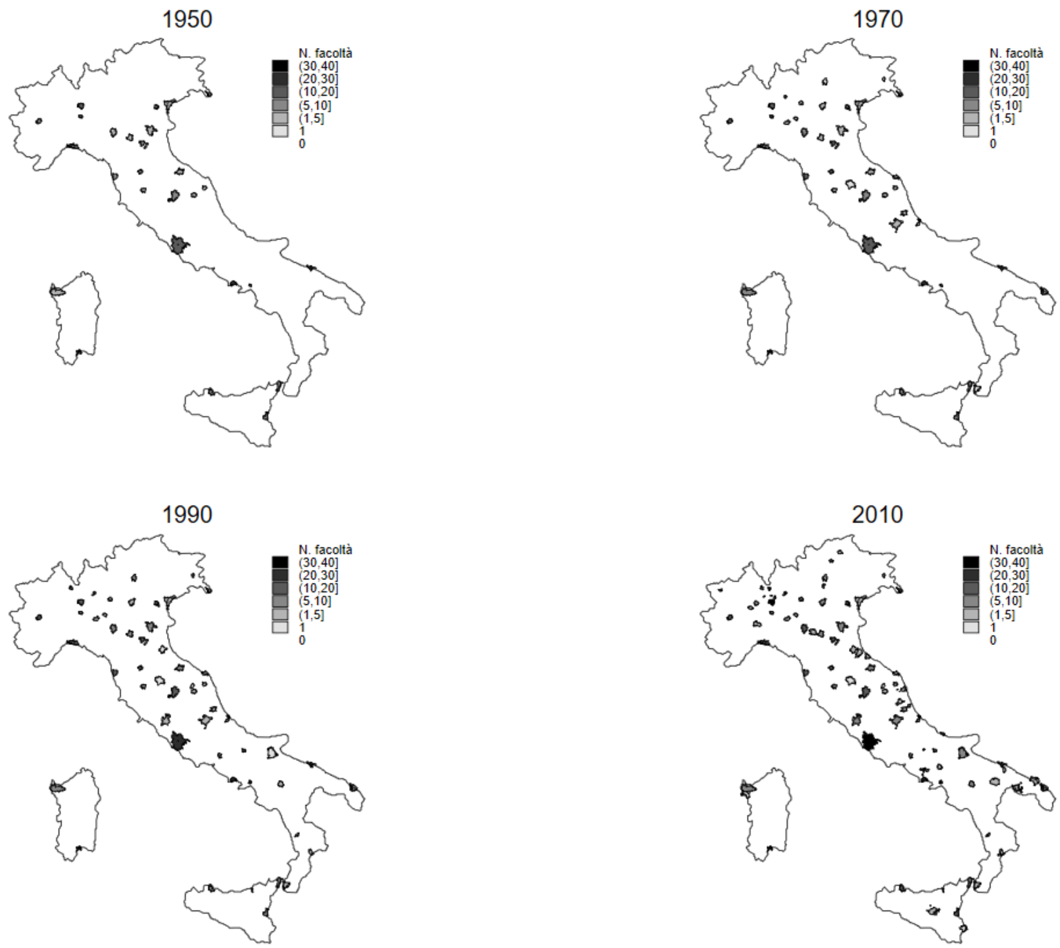


Note: The figures show the distribution of *Facoltà* by province (NUTS-3).

as a result of a better local health system. In this respect, we stress that this concern would be more relevant if we were to analyze health outcomes (self-reported health, incidence of given diseases) instead of individual health behaviors. Indeed, the latter are less likely to be correlated with the local availability and quality of health care services.

We provide nonetheless some evidence on the potential relevance of this threat by regressing the change in the number of *Facoltà* at the municipality level between 1992 and 2010 on life expectancy of males and females in 1992. Results are shown in Figure 3. The choice of the time span reflects data availability, from 1992 on life expectancy (by gender) and until 2010 for the HE supply. Estimates include province fixed effects. Results indicate that the life expectancy in the early '90s does not significantly predict the change in HE

Figure 2: Expansion of Higher Education in Italy (N. *Facoltà*): Municipalities



Note: Only the municipalities with at least one *Facoltà* are shown in the figure. Shades of grey represent the number of *Facoltà* in each municipality.

supply. This evidence seems to exclude that new HE institutions were created specially in municipalities with higher life expectancy, that is where a higher quality of health services or better health habits made individuals to live longer on average.

As to the instrument's relevance, Table A-3 in the Appendix reports the first-stage estimates using college proximity as an instrument, respectively on years of education (Panel (a)) and the individual college indicator (Panel (b)), and using the CSL reform to instrument years of schooling (Panel(c)). For the sake of brevity we report only the first-stage coefficients computed considering the outcome Physical Activity, the others are very similar.

Panel (a) shows that individuals for which the closest colleges are in the neighboring municipality or within the province (but not in neighboring municipalities) on average attain

1.3 and 1.6 less years of schooling, respectively, compared to those residing in a municipality with a college. Also the other categories have a penalty in terms of years of education although some coefficients are not precisely estimated. Panel (b) reports a similar negative effect of distance to college on the probability of completing a university education (in this case the sample only includes individuals with upper secondary education or higher). Individuals with a *Facoltà* in the municipality of residence are for instance 11.5 percentage points (pp) more likely to hold a university degree compared to those living in the same province but in municipalities without university premises. Finally, Panel (c) reports the effect of the CLS reform, which increased compulsory schooling by three years for the cohorts born after 1950. Interestingly, the first-stage coefficient indicates that individuals eligible for the reform achieved on average 0.15 more years of education. Although we might have expected larger effects, potentially a coefficient close to one in case of perfect compliance with the new school obligation, the smaller magnitude can be explained by the fact that compulsory schooling was not always enforced and that years of education are imputed based on completed educational qualifications (Brandolini and Cipollone, 2002). Many individuals who were forced to stay longer in education might have not obtained the corresponding qualification (i.e. a lower secondary education diploma). Finding first-stage coefficients lower than one, when using compulsory schooling years as an instrument for years of schooling is quite common in the literature using CSLs (Brunello et al., 2013), especially for reforms targeting earlier birth cohorts for which CSL enforcement was low. Yet the estimates for the Italian reform appear to be particularly low. Thus, it is important to bear in mind that we might have a similar issue as Lleras-Muney (2005): the effect of the CLS reform may not be large enough, and the small increase in education to be unlikely to have substantially changed individuals' health behaviors and outcomes.

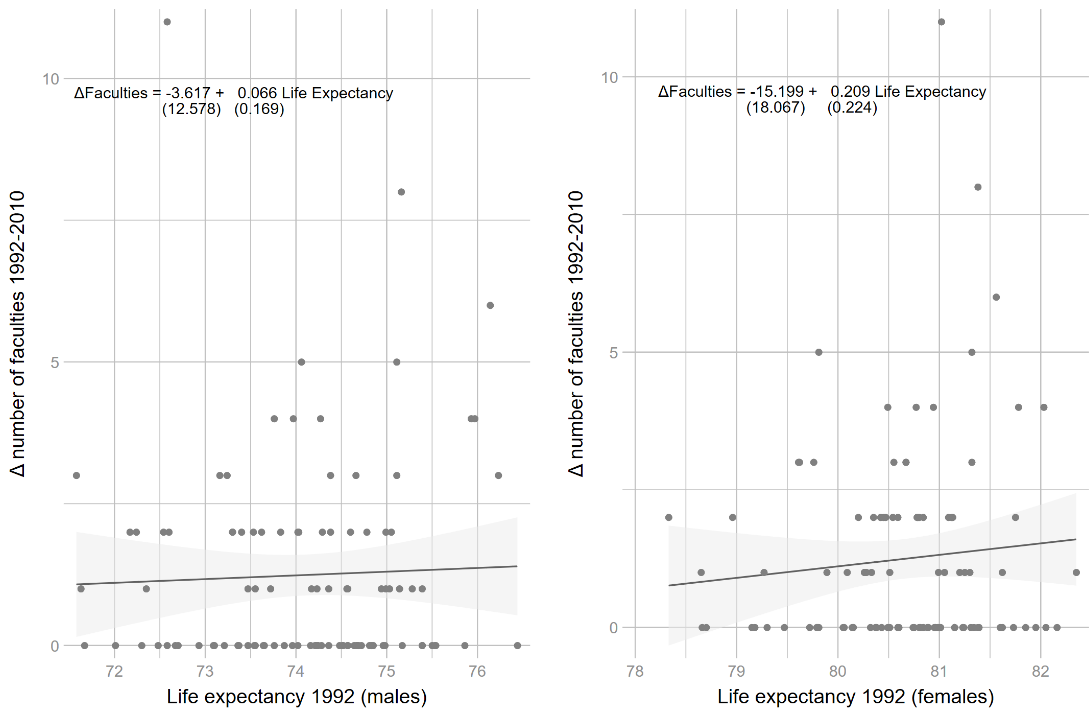
5 Second-stage results

5.1 Evidence from college proximity

For the model with years of schooling, Table 2 presents the OLS estimates, the IVs (2SLS) estimates using college proximity and their diagnostics. Results with the tertiary education dummy are shown in Table 3. All regressions control for age, age squared, gender and province by survey year fixed effects.

For healthy habits, benchmark OLS estimates in Table 2 report a negative association with years of education in the case of water consumption (−0.002), while all other indicators

Figure 3: Instrument check: growth of HE supply vs, life expectancy



Note: Cross-plot of changes in HE supply expansion (N. of *Facoltà*) at the municipality level between 1992 and 2010 and life expectancy in 1992. The coefficient estimates reported in the graphs refer to regressions also including province (NUTS-3) fixed effects. Data on life expectancy are drawn from Health for All, ISTAT.

show a positive and statistically significant coefficient. Years of education are negatively associated with both obesity and body mass index (-0.7 pp and -0.139, respectively, for an additional year of education) and also with the amount of alcohol consumed. By contrast, the probability of never consuming alcohol decreases with education. Years of education are negatively associated with current smoking. Since OLS results are likely to suffer from the biases originated by the endogeneity of education, in the remainder of this section we focus on the IVs estimates.

Overall, IVs results in Table 2 show larger coefficients compared to OLS and only in few cases the statistical significance is lost. As for healthy habits, the estimates suggest that one additional year of education contributes to increasing the likelihood of physical activity by 3 pp, of having breakfast by 1.3 pp, of eating fruits and vegetables by 1.4 pp (with an increase of 0.074 in the number of pieces, corresponding to a 3% increase with respect to the average consumption, i.e. 2.3 pieces), of not eating junk food by 0.9 pp and of not drinking gas beverages by 0.8 pp. The literature brings no consensus on whether

Table 2: Baseline Estimates, OLS and IVs estimates (years of education)

	OLS		IV					
					Diagnostics			
Outcome	Coef	s.e.	Coef	s.e.	Wald 1st stage	Hansen J-stat	pvalue	N
Healthy habits								
Physical activity	0.023***	(0.00)	0.030***	(0.01)	21.724	3.274	0.513	531,674
Fruit and Vegetables	0.003***	(0.00)	0.014**	(0.01)	21.191	3.791	0.435	525,357
Fruit and vegetables (pieces)	0.032***	(0,00)	0.074***	(0.02)	20.958	4.832	0.305	448,181
Water	-0.002***	(0.00)	-0.012	(0.01)	19.109	10.632	0.031	314,050
Breakfast	0.011***	(0.00)	0.013***	(0.00)	21.829	5.168	0.271	525,759
No junk food	0.003***	(0.00)	0.009**	(0.00)	20.896	6.134	0.189	445,293
No gas beverages	0.004***	(0.00)	0.008***	(0.00)	21.532	4.637	0.327	513,653
Weight								
Obesity	-0.007***	(0.00)	-0.014***	(0.00)	21.208	3.927	0.416	533,577
Body Mass Index	-0.139***	(0.00)	-0.299***	(0.00)	21.202	5.176	0.270	533,577
Alcohol								
Light alcohol consumption	-0.002***	(0.00)	-0.003	(0.00)	21.131	5.062	0.281	525,303
N. alcohol units a day	-0.011***	(0.09)	-0.007	(0.01)	20.889	4.190	0.381	448,161
Never consumes spirits	-0.006***	(0.00)	-0.022***	(0.00)	21.211	3.809	0.432	519,443
Smoking								
Ever smoked	-0.001	(0.02)	0.008*	(0.00)	21.133	4.332	0.363	524,602
Currently smokes	-0.006***	(0.00)	0.009***	(0.00)	21.133	7.533	0.110	524,602
N. cigarettes	-0.133***	(0.01)	0.144**	(0.06)	21.184	2.719	0.606	521,260

Note: All regressions control for age, age squared, gender and province by survey year fixed effects. Standard errors clustered by province are reported in parentheses.Significance levels: $*$: 10% $**$: 5% $***$: 1% .

education improves healthy habits. Clark and Royer (2013), for instance, do not find any effect in terms of regular physical activity or dietary habits. Conversely, Park and Kang (2008) reports that high school education is associated with a higher probability of physical exercise. Atella and Kopinska (2014) leveraging the 1962 CSL reform in Italy found that completing lower secondary education reduced individual BMI, and calories consumption, and promoted physical activity in Italy. However, these studies have leveraged reforms or educational expansions affecting educational levels lower than tertiary, and are not strictly comparable to our estimates based on college proximity.

In terms of body weight, one additional year of education reduces the chance of obesity by 1.4 pp, and BMI by 0.3, corresponding to a 1.2% decrease (as the average BMI is 25.1). These effects are precisely estimated and significant at the 1 percent level. They may be local and relate to individuals in the upper part of the years of schooling distribution (i.e. the compliers with college proximity), but their magnitude is not trivial, also in policy terms, and twice as large as the OLS estimates. They are also consistent with the international literature that finds that both BMI and obesity are negatively affected by education. Our findings add to Brunello et al. (2013) and Kemptner et al. (2011), which focus on CSL reforms, and confirm that also tertiary education has positive health returns in terms of lowering the risk of obesity.

As to unhealthy habits, no effect is found on alcohol consumption, with the (important) exception of drinking super alcoholics as one additional year of education increases its consumption by 2.2 pp. A striking difference with the OLS results is that smoking behavior is positively affected by education: one additional year of schooling increases the probability of being a current smoker and of having ever smoked (smoking extensive margins) by 0.8 pp and 0.9 pp, respectively, and the number of cigarettes smoked daily by 0.14. Our results are in contrast to those of other studies that have exploited quasi-natural experiments focusing on very specific sub-populations, which found negative effects of education on smoking, such as de Walque (2007) and Grimard and Parent (2007) who leverage the Vietnam draft lottery.[17] Currie and Moretti (2003) exploit college expansion and report negative effects on maternal smoking during pregnancy. However, pregnancy is a very special period in which individuals are likely to temporarily avoid several behaviors that may be harmful to the unborns' health, including smoking. To the best of our knowledge, to date, only Cowan and Tefft (2020) report evidence of negative effects of education on smoking in the general population when exploiting college expansion.

[17] However, as argued by Cawley et al. (2017), the exclusion restriction may be invalid since exposure to the Vietnam war might have increased smoking initiation and translated into a worse smoking addiction.

Findings referring to lower levels of education are instead more mixed. Kemptner et al. (2011) and Clark and Royer (2013), for instance, exploiting CSL reforms do not find any effect on smoking. Conversely, Kenkel et al. (2006) reports that the returns to high school completion may include less smoking but that health returns to GED receipt are much smaller; using longitudinal data on a sample of twins, Koning et al. (2015) finds that one additional year of education reduces the duration of smoking by 9 months but has no effect on the decision to start smoking.

There might be several behavioral explanations for the counter-intuitive effects of education on (especially) smoking in cross-sectional studies. First, smoking (also in moderate quantities) is more likely to produce addiction than other health-related behaviors. This may induce bounded rational decisions and time-varying preferences, e.g. due to self-control problems, time inconsistency, hyperbolic discounting, see Frederick et al. (2002) and Gruber and Mullainathan (2006) that interact with education and health-related outcomes in ways not captured by our instruments, especially in a cross-sectional context where it is not possible to fully take account of the dynamics of smoking habits and of time-varying unobservables driving individual preferences. Moreover, highly educated individuals may be exposed to higher levels of job-related stress, which positively affect smoking.[18]

Table 3 shows OLS and IVs estimates when we restrict the sample to individuals with at least upper secondary education. The main independent variable of interest is now a dichotomous indicator for holding a tertiary education degree. OLS estimates are qualitatively equivalent, in terms of sign and statistical significance as those shown in Table 2 on the whole sample. Of course, given the different scale used to measure education, magnitudes differ. For example, having a tertiary education as compared to upper secondary education reduces BMI by 0.6 points (i.e. a -3% change), i.e. roughly the coefficient reported on Table 2 (-0.14) multiplied by 3-4 years (the additional years of education of tertiary educated vs. upper secondary educated individuals).[19]

The IVs estimates too are generally in line with those shown in Table 2 as to the signs and statistical significance of the effects. For example, holding a tertiary degree increases the probability of performing physical activity at least 1-2 times a week by 28 pp, the consumption of fruit and vegetables by almost one piece (corresponding to a 40% percent increase), reduces the BMI by 3 points (-12%) and the risk of obesity by 12 pp. Tertiary education increases the probability of being currently a smoker by 11 pp and smoking intensity by 1.8

[18] Kamhöfer et al. (2019) report that this may also be an explanation for the absence of positive effects of education on mental health.

[19] Most tertiary educated individuals have undergraduate degrees, whose typical length was four years before the Bologna reform and three years after the reform.

Table 3: OLS and IVs estimates, tertiary vs. upper secondary education

	OLS		IV					
					Diagnostics			
Outcome	Coef	s.e.	Coef	s.e.	Wald 1st stage	Hansen J-stat	pvalue	N
Healthy habits								
Physical activity	0.086***	(0.00)	0.276***	(0.07)	54.342	3.426	0.489	203,074
Fruit and Vegetables	0.024***	(0.00)	0.230***	(0.07)	54.782	2.360	0.670	200,615
Fruit and vegetables (pieces)	0.150***	(0.01)	0.954***	(0.31)	63.359	4.344	0.361	175,220
Water	-0.012**	(0.00)	-0.002	(0.07)	49.929	3.729	0.444	118,423
Breakfast	0.047***	(0.00)	0.122	(0.08)	54.741	4.989	0.288	200,689
No junk food	0.035***	(0.00)	0.102*	(0.05)	63.431	3.936	0.415	174,380
No gas beverages	0.025***	(0.00)	0.069*	(0.04)	54.436	2.358	0.670	196,606
Weight								
Obesity	-0.022***	(0.00)	-0.119***	(0.03)	56.553	3.423	0.490	203,796
Body Mass Index	-0.591***	(0.02)	-3.004***	(0.52)	56.766	5.369	0.251	203,796
Alcohol								
Light alcohol consumption	-0.013***	(0.00)	-0.106**	(0.05)	55.061	4.582	0.333	200,646
N. alcohol units a day	-0.032***	(0.00)	0.111	(0.16)	63.194	5.101	0.277	175,265
Never consumes spirits	-0.029***	(0.00)	-0.238***	(0.05)	55.005	4.238	0.375	198,825
Smoking								
Ever smoked	-0.065***	(0.00)	0.009	(0.06)	54.443	5.862	0.210	200,310
Currently smokes	-0.050***	(0.00)	0.107***	(0.03)	54.443	1.816	0.770	200,310
N. cigarettes	-0.875***	(0.05)	1.778***	(0.48)	53.464	0.321	0.988	198,441

Note: All regressions control for age, age squared, gender and province by survey year fixed effects. Standard errors clustered by province are reported in parentheses.Significance levels: $*$: 10% $**$: 5% $***$: 1% .

cigarettes (i.e. a 60% increase). Unlike when considering years of schooling, a significant negative effect on light alcohol consumption is found for holding a tertiary education degree (about -11 pp).

All in all, the evidence presented so far points to ambiguous effects of education on health-related behaviors, with positive effects on healthy dietary and exercise habits that also bring positive effects on body weight-related outcomes, but some negative effects in terms of higher incidence of unhealthy behaviors related to drinking and smoking.

5.2 Evidence from compulsory schooling laws

Results for the IVs estimates using CSLs as an instrument are shown in Table 4. For ease of comparison, we also report the OLS estimates, that are identical to those in Table 2.

The estimates show a lack of statistically significant effects except in two cases: one additional year of education decreases the probability of drinking gas beverages by 1.3 pp, while — counter-intuitively — decreases the consumption of vegetables by 0.043 pieces (-4%). The IVs diagnostics are generally worse than those using college-proximity. $F-$statistics are generally larger than 10, the rule of thumb often used to detect a weak instrument problem (Staiger and Stock, 1997), but not very far from it.

In Table 5, we investigate whether the different results using college-proximity and CSLs may be due to the differences in the sample composition in terms of birth cohorts. If we restrict the IVs estimation using college proximity to the same cohorts used in Table 4, we observe that the majority of coefficients still maintain the same sign and level of statistical significance. Interestingly enough, the effects estimated on smoking are even larger and more precisely estimated.

As we already said, the lower precision of the IVs estimates using CSLs may be partly due to the small increase in education induced by the 1962 compulsory schooling age reform. For this reason, in Table A-4 in the Appendix we follow a strategy similar to Atella and Kopinska (2014), who use the CSLs instrument focusing on the population of individuals with at most lower secondary schooling and estimate the return to lower secondary schooling versus primary schooling. In this case, the first-stage $F-$statistics improve as we should expect from the fact that the 1962 reform is likely to have increased lower secondary education attainment and not higher levels of education of marginal students. This reduces the concerns related to weak instruments. However, unlike Atella and Kopinska (2014) we find that education does not affect health-related outcomes. The different results may be explained by the fact that we consider a longer period, i.e. 2001-2016 in our study vs. 2005-2009 in their study, and that

Table 4: OLS and IVs estimates, compulsory schooling laws (CSLs)

	OLS		IV				
					Diagnostics		
Outcome	Coef	s.e.	Coef	s.e.	Wald 1st stage	pvalue	N
Healthy habits							
Physical activity	0.021***	(0.00)	0.014	(0.02)	10.54	0.001	194,609
Fruit and vegetables	0.002***	(0.00)	0.013	(0.02)	10.557	0.001	192,500
Fruit and vegetables (pieces)	0.032***	(0,00)	0.052	(0.05)	11.539	0.001	162,630
Water	-0.003***	(0.00)	0.006	(0.02)	11.891	0.001	115,479
Breakfast	0.010***	(0.00)	-0.002	(0.01)	10.564	0.001	192,663
No junk food	0.001***	(0.00)	0.018	(0.01)	11.717	0.001	161,345
No gas beverages	0.002***	(0.00)	0.013**	(0.01)	10.850	0.001	186,927
Weight							
Obesity	-0.008***	(0.00)	-0.018	(0.02)	10.499	0.007	195,242
Body Mass Index	-0.131***	(0.00)	0.287	(0.22)	10.499	0.001	192,216
Alcohol							
Light alcohol consumption	0.000	(0.00)	0.004	(0.01)	10.579	0.001	192,383
N. alcohol units a day	-0.014***	(0.00)	0.026	(0.04)	11.552	0.003	162,540
Never consumes spirits	-0.006***	(0.00)	-0.015	(0.01)	10.450	0.007	189,923
Smoking							
Ever smoked	0.006***	(0.00)	0.032	(0.02)	10.595	0.007	192,216
Currently smokes	0.000	(0.00)	0.006	(0.02)	10.595	0.007	192,216
N. cigarettes	-0.061***	(0.02)	0.279	(0.25)	11.717	0.006	161,345

Note: All regressions control for age, age squared, gender and province by survey year fixed effects. Standard errors clustered by province are reported in parentheses.Significance levels: $*$: 10% $**$: 5% $***$: 1% .

Table 5: OLS and IVs estimates, years of education, same sample as CSLs estimates

	OLS		IV					
					Diagnostics			
Outcome	Coef	s.e.	Coef	s.e.	Wald 1st stage	Hansen J-stat	pvalue	N
Healthy habits								
Physical activity	0.021***	(0.00)	0.025***	(0.01)	19.287	4.601	0.330	194,609
Fruit and vegetables	0.002***	(0.00)	0.006	(0.01)	18.896	8.211	0.084	192,500
Fruit and vegetables (pieces)	0.032***	(0.00)	0.039**	(0.02)	19.888	3.173	0.506	162,630
Water	-0.003***	(0.00)	-0.014**	(0.01)	20.098	2.083	0.720	115,479
Breakfast	0.010***	(0.00)	0.013***	(0.00)	19.127	4.472	0.368	192,663
No junk food	0.001***	(0.00)	0.004	(0.00)	19.878	1.987	0.738	161,345
No gas beverages	0.003***	(0.00)	0.005**	(0.00)	19.477	4.995	0.288	186,927
Weight								
Obesity	-0.008***	(0.00)	-0.016***	(0.00)	19.256	7.809	0.099	195,242
Body Mass Index	-0.131***	(0.00)	-0.289***	(0.04)	19.256	7.247	0.123	195,242
Alcohol								
Light alcohol consumption	-0.001***	(0.00)	-0.001	(0.00)	18.997	10.749	0.033	192,383
N. alcohol units a day	-0.014***	(0.09)	-0.005	(0.00)	20.542	6.331	0.176	162,540
Never consumes spirits	-0.006***	(0.00)	-0.020***	(0.00)	18.879	5.930	0.204	189,923
Smoking								
Ever smoked	-0.006***	(0.02)	0.018***	(0.00)	18.997	6.284	0.179	192,216
Currently smokes	-0.001	(0.00)	0.017***	(0.00)	18.997	1.552	0.817	192,216
N. cigarettes	-0.061***	(0.01)	0.282***	(0.00)	18.874	1.044	0.903	190,594

Note: All regressions control for age, age squared, gender and province by survey year fixed effects. Standard errors clustered by province are reported in parentheses.Significance levels: $*$: 10% $**$: 5% $***$: 1% .

they include in their specification a quadratic polynomial in age and region indicators but not AVQ survey waves indicators, while, as explained in Section 3, by including a quadratic polynomial in age and province-survey year fixed effects, we also control *de facto* for birth cohort effects as well as for survey year.

All in all, a comparison between the effects of education estimated leveraging quasi-experimental variation induced by CSLs vs. tertiary education expansion points to the presence of possible non-linearities in the health returns to education estimated using IVs, with larger effects at higher levels of education. These non-linearities may also be induced by complementarity between individuals' innate ability and education, as recently suggested by Heckman et al. (2018). To put in other words, the LATE estimated on low ability individuals (IVs estimates based on CSLs) may be lower that those estimated on high ability individuals (IVs estimates based on college proximity).

5.3 Robustness checks

In this section we adopt as a benchmark the specification of Table 2 and perform a number of robustness checks to the main analysis that address potential empirical concerns.

First, we refine the analysis to account for mobility across the Italian territory over the life cycle. This addresses measurement issues owing to the fact that the municipality of residence may not be a good proxy of where the individuals lived when they were nineteen years old (at the time of the university enrolment choice). The data provide no specific information about the latter, but in more recent waves (since 2008) they record whether the municipality of birth is the same as that of residence[20]. This happens in about the 42 percent of cases. Table 6 presents estimates on this sample of "stayers". Overall, results are robust and the magnitudes and significance levels for most coefficients remain very similar as to our baseline analysis, supporting our earlier findings. Worth noting are the larger effects estimated for smoking in this sample.

Second, we change the age threshold that we assume for the relevance of college proximity. It may be argued that the choice in terms of tertiary education already takes place when the individual chooses the secondary high school track, which is not taken into account by our instrumental variable if we record the supply of higher education when the individual is 19 years of age. We then construct an alternative instrumental variable that records proximity from the closest university settlement at age 14 instead of 19. Results are in Table A-5 in the

[20] Kemptner et al. (2011) perform a similar robustness check to reduce this potential source of attenuation bias.

Appendix, and, also in this case, are virtually unchanged as compared to the main analysis of Table 2.

Although the frequency of statistically significant effects of years of schooling on health behaviors and outcomes is well above the share one should expect by pure chance, we corrected the $p-$values for multiple testing because we consider several outcome variables. The $p-$values adjusted for multiple hypothesis testing using the Bonferroni-Holm method in Figure A-1 in the Appendix indicate that most coefficients maintain conventional statistical significance levels. Only the coefficients for junk food, vegetable and fruits (extensive margin) and ever smoked lose statistical significance after the correction.

6 Effect heterogeneity

6.1 By gender

Figure 4 presents IVs estimates of Eq. (1) separately by gender.[21] In the left subfigure we show point estimates for outcomes defined as dummy variables. When we consider healthy habits (such as physical activity, consumption of fruit and vegetables, having breakfast, avoiding junk food) the estimated effects are generally consistent across genders, in both sign and magnitude (except for the probability of drinking at list 1.5 liters of water). A protective effect of education is found for both genders in terms of super-alcoholic drinks consumption. As for obesity, IVs estimates show a similar negative coefficient on education for both genders. In contrast with the literature based on CSLs, we do not find evidence in support of larger effects for females (Brunello et al., 2013). Education increases especially women's current smoking behavior, both the probability of being a current smoker and of ever being a smoker. Some gender differences also emerge for drinking. One additional year of education does not increase the probability of light alcohol consumption for males (coefficient is 0.007) but decreases it for females by 1.3 pp.

The left subfigure shows estimates for continuous outcomes. We do no find statistical significant results in terms of consumption of fruit and vegetables and alcohol consumption for both genders. One additional year of education significantly decreases BMI for both males and females. Interestingly, some differences are found in terms of smoking behavior: one additional year of education increases the average number of cigarettes smoked for females only (by 0.23).

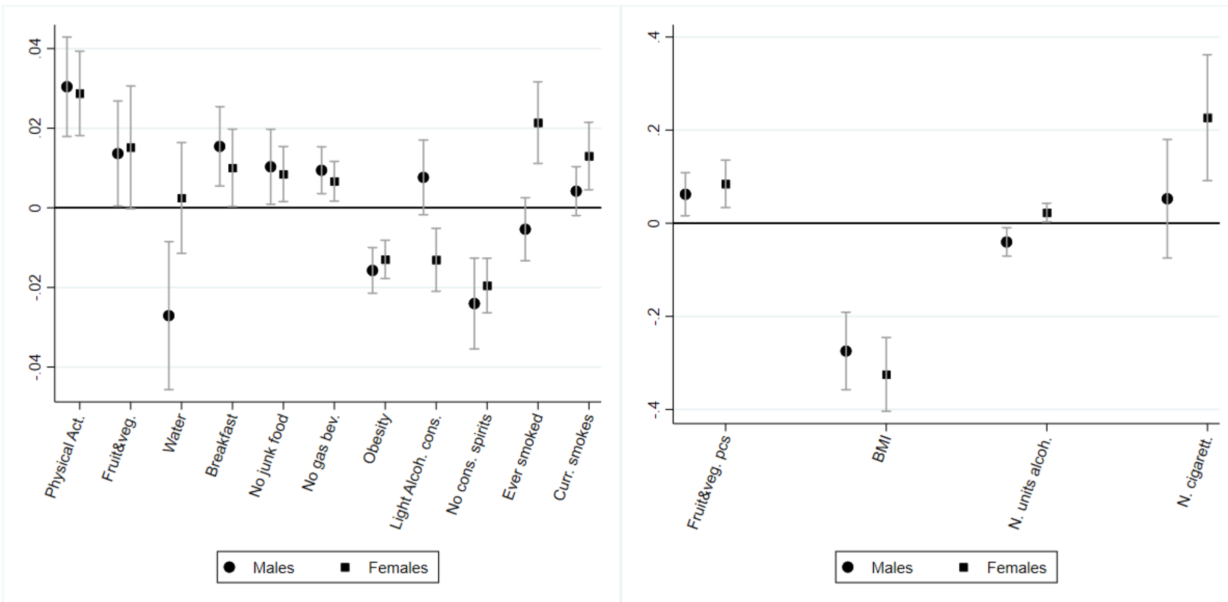
[21] First-stage coefficients are reported in Table A-3.

Table 6: Robustness 1, same municipality of birth and residence OLS and IVs estimates (years of education)

	OLS		IV					
					Diagnostics			
Outcome	Coef	s.e.	Coef	s.e.	Wald 1st stage	Hansen J-stat	pvalue	N
Healthy habits								
Physical activity	0.023***	(0.00)	0.028***	(0.01)	21.777	5.085	0.279	143,902
Fruit and vegetables	0.003***	(0.00)	0.006	(0.01)	22.097	3.767	0.439	142,556
Fruit and vegetables (pieces)	0.035***	(0.00)	0.044*	(0.03)	22.097	6.075	0.194	142,556
Water	0.001	(0.00)	0.018	(0.01)	18.092	4.060	0.398	73,597
Breakfast	0.010***	(0.00)	0.013***	(0.00)	22.341	3.777	0.398	142,685
No junk food	0.003***	(0.00)	0.008	(0.01)	22.196	5.555	0.235	141,668
No gas beverages	0.004***	(0.00)	0.004	(0.00)	21.927	4.441	0.350	139,552
Weight								
Obesity	-0.007***	(0.00)	-0.011***	(0.00)	21.822	3.807	0.433	144,295
Body Mass Index	-0.140***	(0.00)	-0.258***	(0.00)	21.822	3.122	0.538	144,295
Alcohol								
Light alcohol consumption	-0.002***	(0.00)	-0.007	(0.00)	21.950	6.866	0.143	142,602
N. alcohol units a day	-0.015***	(0.09)	0.021	(0.01)	21.905	4.484	0.674	142,602
Never consumes spirits	-0.006***	(0.00)	-0.018***	(0.00)	21.918	7.749	0.101	140,730
Smoking								
Ever smoked	-0.001	(0.02)	0.021***	(0.00)	22.014	3.084	0.544	142,404
Currently smokes	-0.006***	(0.00)	0.016***	(0.00)	22.014	4.789	0.310	142,404
N. cigarettes	-0.142***	(0.01)	0.256***	(0.08)	22.265	3.156	0.532	141,642

Note: All regressions control for age, age squared, gender and province by survey year fixed effects. Standard errors clustered by province are reported in parentheses.Significance levels: $*$: 10% $**$: 5% $***$: 1% .

Figure 4: IVs estimates (years of education) by gender



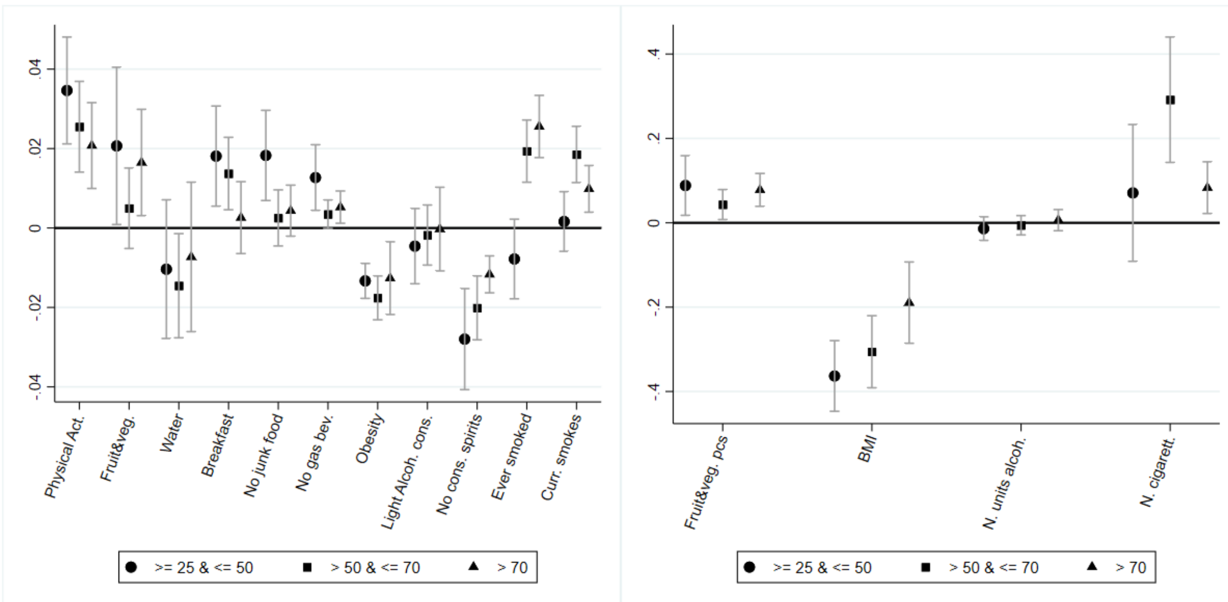
Note: All regressions control for age, age squared, gender and province by survey year fixed effects. Standard errors are clustered by province. 95% confidence interval are shown.

6.2 By age

Previous studies have generally not investigated differences in the health returns to education by age. The use of pooled cross sections of a survey spanning about 15 years enables us to investigate heterogeneity by age in potential effects. Figure 5 plots the IVs estimates on years of schooling based on college-proximity obtained splitting the total sample in three age brackets, 25-50, 50-70, 70+.[22] It must be kept in mind that especially the estimates for the oldest age group could be potentially affected by differential mortality related to education (van Kippersluis et al., 2011). Yet, with this caveat in mind, Figure 5 gives some interesting insights. First, for some outcomes we find very similar effects across ages: physical activity, water consumption, obesity, light alcohol consumption, fruit and vegetables pieces and number of alcoholic units. Some age heterogeneity emerges for some healthy dietary habits. Effects on having breakfast are not observed for the oldest group, while effects on consumption of gas beverages and junk food are larger for the youngest age group. Quite

[22] First-stage coefficients are available upon request.

Figure 5: IVs estimates (years of education) by age groups



Note: All regressions control for gender and province by survey year fixed effects. Standard errors are clustered by province. 95% confidence intervals are shown.

interestingly, effects on smoking are specular: effects on ever smoking and current smoking are not observed for the youngest group, while those on the number of cigarettes are larger for the 50-70 age group. The effects on no consumption of spirit and BMI are generally decreasing in magnitude by age, with significant differences between the youngest and the oldest age group.

All in all, this subsection helps rationalize the differences in results that studies might find depending on the timing of the reforms they use and the affected cohorts. On the one hand, cross-sectional studies generally provide effects at a given age (or on individuals in a narrow age bracket) at which the affected cohort(s) is (are) observed. On the other hand, even studies using longitudinal data or pooled cross-sections in which the affected cohorts are observed at different ages, if age-specific effects are not explicitly modeled estimate average effects over the entire life-cycle, which may however sharply differ across ages.[23]

[23] Clark and Royer (2013), for instance, use the 1986-96 waves of the General Household Survey and the 1991-2004 waves of the Health Survey for England but do not model age-specific effects.

7 Concluding remarks

Despite the literature investigating the health returns to education is already vast, most studies have focused on evidence coming from reforms increasing the age of compulsory schooling (CSLs). Results are generally mixed, and vary across health outcomes and countries.

By contrast, the literature leveraging higher education expansion or focusing on tertiary education is much thinner. This study contributes to the extant literature in a number of ways. First, for the first time, we provide evidence on the health returns to education generated by higher education expansion in Italy. Expansion of higher education was heterogeneous across time (birth cohorts) and space, and was presumably unrelated to unobservable health-related factors. Second, we contrast the IVs results obtained with college proximity with those originating from the 1962 reform that increased compulsory schooling age by 3 years. In both cases, we use the same data (ISTAT's "Aspects of everyday life", AVQ) and a similar empirical specification which enables us to compare the college-proximity and the CSLs IVs estimates. The comparison between IVs estimates generated by different complying sub-populations (marginal students for CSLs and abler students for college proximity) enables us to throw light on potential non-linearities in the health returns to education. Last, thanks to the longevity of the AVQ survey, we are able to investigate heterogeneous effects by age. Indeed, differences in health-related behaviors and outcomes across individuals with different levels of education may differ by age.

Our analysis uncovers important non-linearities in the effects of education on health. Indeed, effects on healthy and unhealthy behaviors and health outcomes are found for highly educated individuals when we exploit changes in college-proximity for identification but the same are not found for estimates based on CSLs, which apply to lower-secondary educated individuals.

However, effects on behaviors do not always go in the expected direction. While education generally induces the adoption of good dietary and exercise habits, it also increases the incidence of some unhealthy behaviors related to smoking and drinking.

Our effect heterogeneity analysis highlights similar effects across genders, except in a few cases. For instance, we find that the positive effects of education on smoking found in the pooled sample is mainly driven by females. Moreover, our study shows some evidence that the effects of education may be sensitive to the age at which they are measured. In particular, differences in healthy dietary habits (e.g. no consumption of gas beverages or junk food), are larger in the age group 25-50 and that on BMI is decreasing in magnitude by age. These findings suggest that younger individuals may abstain from these behaviors not only on the basis of pure health concerns but also because they impact on their physical appearance (Averett et al., 2008) and labor market outcomes (Brunello and d'Hombres, 2007; Atella et al., 2008). By contrast, effects on smoking and drinking are generally larger at older ages. Although the cross-sectional data we use do not enable us to observe the dynamics of these behaviors over time, the latter results could be related to some individuals quitting smoking or drinking (or drinking/smoking less) depending on the incidence of some given diseases and the worsening of their health conditions (Clark and Etilé, 2002) while ageing, which may differ across individuals with different levels of education. An alternative explanation could be that the mediating effects of these behaviors on health are larger for older individuals (Brunello et al., 2016). Exploring these hypotheses using longitudinal data could be a promising avenue for future research.

Acknowledgments

We thank participants in the 2020 annual conference of the Italian Association of Labour Economists (AIEL), and seminar participants at the Catholic University of Milan, Lille University and Marche Polytechnic University for useful comments. We acknowledge funding from the Catholic University of Milan D 3.2 strategic research grant "Fight against poverty: empirical research in support of public policies". All elaborations in this paper use restrict-access data from the survey *ISTAT, Aspetti della vita quotidiana* and were run at the Italian National Statistical Institute's "Laboratorio per l'Analisi dei Dati ELEmentari" (ADELE) (Lombardy branch). We gratefully acknowledge ISTAT personnel that made this project possible. Special acknowledgments go to Lorena Viviano, Giusy Del Vicario and Luigi Di Gennaro. ISTAT does not bear any responsiblity for the content of this paper. The usual disclaimers apply.

References

Albouy, V. and L. Lequien (2009). Does compulsory education lower mortality. *Journal of Health Economics 28*(1), 155–168.

Atella, V. and J. Kopinska (2014). Body weight, eating patterns, and physical activity: The role of education. *Demography 51*(4), 1225–1249.

Atella, V., N. Pace, and D. Vuri (2008). Are employers discriminating with respect to weight?: European evidence using quantile regression. *Economics & Human Biology 6*(3), 305–329.

Averett, S. L., A. Sikora, and L. M. Argys (2008). For better or worse: Relationship status and body mass index. *Economics & Human Biology 6*(3), 330–349.

Black, D. A., Y.-C. Hsu, and L. J. Taylor (2015). The effect of early-life education on later-life mortality. *Journal of Health Economics 44*, 1–9.

Brandolini, A. and P. Cipollone (2002). Return to education in italy 1992-1997. *Bank of Italy, Research Department, working paper*.

Bratti, M. and S. Verzillo (2019). The gravity of quality: Research quality and the attractiveness of universities in italy. *Regional Studies 53*(10), 1385–1396.

Brunello, G. and B. d'Hombres (2007). Does body weight affect wages?: Evidence from europe. *Economics & Human Biology 5*(1), 1–19.

Brunello, G., D. Fabbri, and M. Fort (2013). The causal effect of education on body mass: Evidence from Europe. *Journal of Labor Economics 31*(1), 195–223.

Brunello, G., M. Fort, N. Schneeweis, and R. Winter-Ebmer (2016). The causal effect of education on health: What is the role of health behaviors? *Health Economics 25*(3), 314–336.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling. In C. Louis, E. K. Grant, and R. Swindinsky (Eds.), *Aspects of Labour Economics: Essays in Honour of John Vanderkamp*. University of Toronto Press.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127–1160.

Cawley, J., D. De Walque, and D. Grossman (2017). The Effect of Stress on Later-Life Health: Evidence from the Vietnam Draft. *National Bureau of Economic Research Woring Paper* (23334).

Clark, A. and F. Etilé (2002). Do health changes affect smoking? Evidence from British panel data. *Journal of Health Economics 21*(4), 533–562.

Clark, D. and H. Royer (2013). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review 103*(6), 2087–2120.

Cottini, E., P. Ghinetti, and S. Moriconi (2019). Higher education supply, neighbourhood effects and economic welfare. *CESifo Working Paper Series* (7483).

Cowan, B. W. and N. Tefft (2020). College access and adult health. *National Bureau of Economic Research Working Paper* (26685).

Currie, J. and E. Moretti (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics 118*(4), 1495–1532.

Davies, N. M., M. Dickson, G. D. Smith, G. van den Berg, and F. Windmeijer (2016). The causal effects of education on health, mortality, cognition, well-being, and income in the UK Biobank. *bioRxiv*, 074815.

de Walque, D. (2007). Does education affect smoking behaviors? Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics 26*(5), 877–95.

Eide, E. and M. Showalter (2011). Estimating the relation between health and education: What do we know and what do we need to know? *Economics of Education Review 30*(5), 778–791.

Farrell, P. and V. Fuchs (1982). Schooling and health: The cigarette connection. *Journal of Health Economics 1*(3), 217–230.

Fischer, M., M. Karlsson, and T. Nilsson (2013). Effects of compulsory schooling on mortality: Evidence from Sweden. *International Journal of Environmental Research and Public Health 10*(8), 3596–3618.

Fletcher, J. M. (2015). New evidence of the effects of education on health in the US: Compulsory schooling laws revisited. *Social Science & Medicine 127*, 101–107.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002, June). Time discounting and time preference: A critical review. *Journal of Economic Literature 40*(2), 351–401.

Fuchs, V. (1982). Time preference and health: An explanatory study. In V. Fuchs (Ed.), *Economic Aspects of Health*, pp. 93–120. University of Chicago Press.

Galama, T. J., A. Lleras-Muney, and H. Van Kippersluis (2018). The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. *National Bureau of Economic Research Working Paper* (24225).

Grimard, F. and D. Parent (2007). Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking? *Journal of Health Economics 26*(5), 896–892.

Grossman, M. (2006). Education and nonmarket outcomes. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Chapter 10, pp. 577–633. Amsterdam: Elsevier.

Gruber, J. and S. Mullainathan (2006). *Do Cigarette Taxes Make Smokers Happier?*, pp. 109–146. London: Palgrave Macmillan UK.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy 126*(S1), S197–S246.

James, J. (2015). Health and education expansion. *Economics of Education Review 49*, 193–215.

Janke, K., D. W. Johnston, C. Propper, and M. A. Shields (2020). The causal effect of education on chronic health conditions in the UK. *Journal of Health Economics 70*, 102252.

Jürges, H., S. Reinhold, and M. Salm (2011). Does schooling affect health behavior? Evidence from the educational expansion in Western Germany. *Economics of Education Review 30*(5), 862–872.

Kamhöfer, D. A., H. Schmitz, and M. Westphal (2019). Heterogeneity in marginal non-monetary returns to higher education. *Journal of the European Economic Association 17*(1), 205–244.

Kemptner, D., H. Jürges, and S. Reinhold (2011). Changes in compulsory schooling and the causal effect of education on health. *Journal of Health Economics 30*(2), 340–354.

Kenkel, D., D. Lillard, and A. Mathios (2006). The roles of high school completion and GED receipt in smoking and obesity. *Journal of Labor Economics 24*(3), 635–660.

Koning, P., D. Webbink, and N. G. Martin (2015). The effect of education on smoking behavior: new evidence from smoking durations of a sample of twins. *Empirical Economics 48*(4), 1479–1497.

Lager, A. C. J. and J. Torssander (2012). Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. *Proceedings of the National Academy of Sciences 109*(22), 8461–8466.

Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies 72*(1), 189–221.

Mazumder, B. (2008). Does education improve health? A reexamination of the evidence from compulsory schooling laws. *Economic Perspectives 32*, 2–16.

Meghir, C., M. Palme, and E. Simeonova (2018). Education and mortality: Evidence from a social experiment. *American Economic Journal: Applied Economics 10*(2), 234–56.

OECD (2008). *OECD Review of Tertiary Education. Final Report, OECD, Paris.* OECD Paris.

OECD (2014). *The State of Higher Education 2014. A. Glass, (eds.). The OECD Higher Education Programme (IMHE), Paris.* OECD Paris.

Oppedisano, V. (2011). The (adverse) effects of expanding higher education: Evidence from Italy. *Economics of Education Review 30*(5), 997–1008.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review 96*(1), 152–175.

Park, C. and C. Kang (2008). Does education induce healthy lifestyle? *Journal of Health Economics 27*(6), 1516–1531.

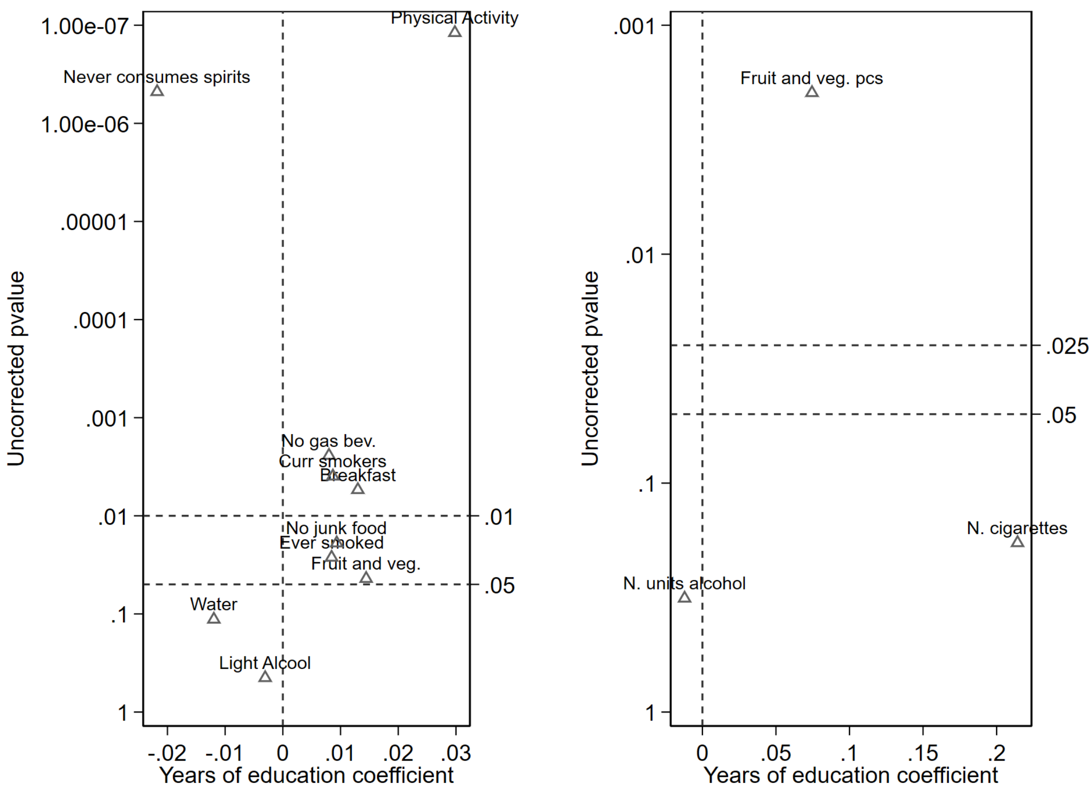
Pigini, C. and S. Staffolani (2016). Beyond participation: do the cost and quality of higher education shape the enrollment composition? The case of Italy. *Higher education 71*(1), 119–142.

Staiger, D. and J. H. Stock (1997). Instrumental variables with weak instruments. *Econometrica 65*(3), 557–86.

van Kippersluis, H., H. O'Donnell, and H. van Doorslaer (2011). Long-run returns to education. Does schooling lead to an extended old age? *Journal of Human Resources 46*(4), 695–721.

Appendix A Supplementary tables and figures

Figure A-1: Multiple testing using Bonferroni correction



Note. The Figures show the corrected $p-$values using Bonferroni correction.

Table A-1: Outcome variables' definition

Variable	Definition
Healthy habits	
Physical activity	Performs physical activity at least 1-2 times a week
Fruit and vegetables	Eats fruit and vegetables at least once a day
Fruit and vegetables (pieces)	N. of pieces of fruit and vegetables consumed per day
Water	Drinks at least 1.5 litres of water per day
Breakfast	Eats something for breakfast
No junk food	Eats less than once a week or never eat sweets and salty snacks
No gas beverages	Consumes less than 1-2 glasses of gas beverages (excluding water) per day
Weight	
Obese	Indicator for BMI ≥ 30
Body Mass Index (BMI)	BMI (= weight in Kg / height^2 in cm)
Alcohol	
Light alcohol consumption	Does not drink more than one glass of wine / beer per day
N. Alcohol units a day	Glasses of alcoholic beverages consumed daily
Never consumes spirits	Never consumes spirits
Smoking	
Ever smoked	Ever smoked
Current smoker	Current smoker
N. of cigarettes	Average daily consumption of cigarettes
Instruments	
Location nearest HE (at least 1 Facoltà:	
Same municipality	Nearest Facoltà is in the Municipality of residence
Neighbouring municipalities	Nearest Facoltà is in neighbouring Municipalities of that of residence
Same province	Nearest Facoltà is in the Province of residence
Neighbouring provinces	Nearest Facoltà is in neighbouring Provinces of that of residence
Same region	Nearest Facoltà is in the Region of residence
Neighbouring regions	Nearest Facoltà is in one neighbouring Region of that of residence or even further

Table A-2: Outcomes variables by survey coverage

Outcome	Available years	N
Healthy habits		
Physical activity	2001-2016	531,674
Fruit and vegetables	2001-2016	525,357
Fruit and vegetables (pieces)	2003-2016	448,181
Water	2003-2012	314,050
Breakfast	2001-2016	525,759
No junk food	2003-2016	445,293
No gas beverages	2001-2016	513,653
Weight		
Obesity	2001-2016	533,577
Body Mass Index	2001-2016	533,577
Alcohol		
Light alcohol consumption	2001-2016	525,303
N. alcohol units a day	2003-2016	448,161
Never consumes spirits	2001-2016	519,443
Smoking		
Ever smoked	2001-2016	524,602
Currently smokes	2001-2016	524,602
N. cigarettes	2001-2016	521,260

Note: AVQ was not administered in 2004.

Table A-3: First-stage estimates for the outcome Physical Activity

	All		Male		Female	
	Coef (1)	s.e.	Coef (2)	s.e.	Coef (3)	s.e.
Instrument: Location nearest Facoltà (college proximity)						
Panel (a). Years of education:						
Neighb. municipality	-1.282***	(0.154)	-1.297***	(0.17)	-1.28***	(0.144)
Same province	-1.600***	(0.164)	-1.643***	(0.176)	-1.573***	(0.153)
Neighb. province	-1.492***	(0.154)	-1.537***	(0.168)	-1.472***	(0.146)
Same region	-1.298***	(0.214)	-1.459***	(0.247)	-1.162***	(0.228)
Neighb. region	-1.937***	(0.298)	-1.990***	(0.309)	-1.909***	(0.302)
N. observations	531,674		254,061		277,612	
Panel (b). College:						
Neighb. municipality	-0.115***	(0.008)				
Same province	-0.131***	(0.010)				
Neighb. province	-0.108***	(0.008)				
Same region	-0.087***	(0.027)				
Neighb. region	-0.093***	(0.019)				
N. observations	203,074					
Instrument: CSLs						
Panel (c). Years of compulsory schooling:						
	0.149***	(0.034)				
N. observations	194,609					

Note: All regressions control for age, age squared, gender and province by survey fixed effects. Panel a), column (1) shows the first stage for Table 2 while column (2) and (3) for Figure 4; Panel b), column (1) shows the first stage for Table 3; Panel c), column (1) shows the first stage for Table 4. Standard errors (reported in parentheses) are clustered by province in panels a) and b) and by year of birth in panel c). First-stage coefficients are reported only for the Physical Activity Outcome, they vary across the samples used for the different outcomes (depending on their time availability), but they are very similar. Significance levels: $*$: 10% $**$: 5% $***$:1%.

Table A-4: OLS and IVs estimates, compulsory schooling law, low secondary vs primary — Atella and Kopinska (2014) approach

	OLS		IV				
					Diagnostics		
Outcome	Coef	s.e.	Coef	s.e.	Wald 1st stage	pvalue	N
Healthy habits							
Physical activity	0.079***	(0.00)	-0.180	(0.20)	13.75	0.006	115,612
Fruit and vegetables	0.009**	(0.00)	0.137	(0.20)	14.427	0.005	114,282
Fruit and vegetables (pieces)	0.148***	(0.01)	-0.611	(0.45)	14.354	0.006	96,696
Water	-0.007	(0.01)	0.125	(0.18)	8.606	0.003	68,863
Breakfast	0.037***	(0.00)	-0.078	(0.15)	15.323	0.004	114,392
No junk food	0.001	(0.00)	0.141	(0.16)	15.238	0.005	95,835
No gas beverages	0.004*	(0.00)	0.137	(0.11)	16.629	0.003	110,962
Weight							
Obesity	-0.045***	(0.00)	-0.180	(0.16)	14.567	0.005	115,995
Body Mass Index	-0.633***	(0.00)	-2.773*	(1.54)	14.567	0.005	115,995
Alcohol							
Light alcohol consumption	0.002	(0.00)	-0.003	(0.14)	15.847	0.004	114,180
N. alcohol units a day	-0.075***	(0.01)	-0.147	(0.40)	15.958	0.006	96,611
Never consumes spirits	-0.025***	(0.00)	-0.019	(0.14)	16.654	0.004	112,599
Smoking							
Ever smoked	0.046***	(0.00)	0.107	(0.25)	16.106	0.004	114,144
Currently smokes	0.001	(0.00)	0.001	(0.15)	16.106	0.004	114,144
N. cigarettes	-0.181**	(0.07)	0.400	(2.46)	15.968	0.004	113,557

Note:All regressions control for age, age squared, gender and province by survey year fixed effects. Standard errors clustered by year of birth are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1% .

Table A-5: IVs estimates with college proximity at age 14 (years of education)

	OLS		IV					
					Diagnostics			
Outcome	Coef	s.e.	Coef	s.e.	Wald 1st stage	Hansen J-stat	pvalue	N
Healthy habits								
Physical activity	0.023***	(0.00)	0.030***	(0.01)	21.145	2.996	0.559	533,213
Fruit and vegetables	0.003***	(0.00)	0.014*	(0.01)	21.173	2.732	0.604	526,881
Fruit and vegetables (pieces)	0.032***	(0,00)	0.074***	(0.02)	19.432	3.866	0.424	449,614
Water	-0.002***	(0.00)	0.012	(0.01)	19.187	8.829	0.066	315,099
Breakfast	0.011***	(0.00)	0.013**	(0.00)	21.299	2.445	0.654	527,284
No junk food	0.003***	(0.00)	0.010**	(0.00)	19.556	6.699	0.153	446,673
No gas beverages	0.004***	(0.00)	0.008***	(0.00)	21.957	4.785	0.310	515,033
Weight								
Obesity	-0.007***	(0.00)	-0.015***	(0.00)	21.229	1.947	0.746	535,120
Body Mass Index	-0.139***	(0.00)	-0.299***	(0.00)	21.714	5.246	0.263	535,120
Alcohol								
Light alcohol consumption	-0.002***	(0.00)	-0.003	(0.00)	21.483	4.891	0.299	526,844
N. alcohol units a day	-0.015***	(0.09)	0.013	(0.01)	19.617	5.465	0.243	449,611
Never consumes spirits	-0.006***	(0.00)	-0.022***	(0.00)	21.722	5.318	0.256	520,897
Smoking								
Ever smoked	-0.001	(0.02)	0.008**	(0.00)	21.278	4.983	0.289	526,117
Currently smokes	-0.006***	(0.00)	0.009***	(0.00)	21.278	6.831	0.145	526,117
N. cigarettes	-0.137***	(0.01)	0.142**	(0.06)	21.323	4.191	0.381	522,755

Note: All regressions control for age, age squared, gender and province by survey year fixed effects. Standard errors clustered by province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Working Paper del Dipartimento di Economia e Finanza

1. L. Colombo, H. Dawid, *Strategic Location Choice under Dynamic Oligopolistic Competition and Spillovers*, novembre 2013.
2. M. Bordignon, M. Gamalerio, G. Turati, *Decentralization, Vertical Fiscal Imbalance, and Political Selection*, novembre 2013.
3. M. Guerini, *Is the Friedman Rule Stabilizing? Some Unpleasant Results in a Heterogeneous Expectations Framework*, novembre 2013.
4. E. Brenna, C. Di Novi, *Is caring for elderly parents detrimental to women's mental health? The influence of the European North-South gradient*, novembre 2013.
5. F. Sobbrio, *Citizen-Editors' Endogenous Information Acquisition and News Accuracy*, novembre 2013.
6. P. Bingley, L. Cappellari, *Correlation of Brothers Earnings and Intergenerational Transmission*, novembre 2013.
7. T. Assenza, W. A. Brock, C. H. Hommes, *Animal Spirits, Heterogeneous Expectations and the Emergence of Booms and Busts*, dicembre 2013.
8. D. Parisi, *Is There Room for 'Fear' as a Human Passion in the Work by Adam Smith?*, gennaio 2014.
9. E. Brenna, F. Spandonaro, *Does federalism induce patients' mobility across regions? Evidence from the Italian experience*, febbraio 2014.
10. A. Monticini, F. Ravazzolo, *Forecasting the intraday market price of money*, febbraio 2014.
11. Tiziana Assenza, Jakob Grazzini, Cars Hommes, Domenico Massaro, *PQ Strategies in Monopolistic Competition: Some Insights from the Lab*, marzo 2014.
12. R. Davidson, A. Monticini, *Heteroskedasticity-and-Autocorrelation-Consistent Bootstrapping*, marzo 2014.
13. C. Lucifora, S. Moriconi, *Policy Myopia and Labour Market Institutions*, giugno 2014.
14. N. Pecora, A. Spelta, *Shareholding Network in the Euro Area Banking Market*, giugno 2014.
15. G. Mazzolini, *The economic consequences of accidents at work*, giugno 2014.
16. M. Ambrosanio, P. Balduzzi, M. Bordignon, *Economic crisis and fiscal federalism in Italy*, settembre 2014.
17. P. Bingley, L. Cappellari, K. Tatsiramos, *Family, Community and Long-Term Earnings Inequality*, ottobre 2014.
18. S. Frazzoni, M. L. Mancusi, Z. Rotondi, M. Sobrero, A. Vezzulli, *Innovation and export in SMEs: the role of relationship banking*, novembre 2014.
19. H. Gnutzmann, *Price Discrimination in Asymmetric Industries: Implications for Competition and Welfare*, novembre 2014.
20. A. Baglioni, A. Boitani, M. Bordignon, *Labor mobility and fiscal policy in a currency union*, novembre 2014.
21. C. Nielsen, *Rational Overconfidence and Social Security*, dicembre 2014.
22. M. Kurz, M. Motolese, G. Piccillo, H. Wu, *Monetary Policy with Diverse Private Expectations*, febbraio 2015.
23. S. Piccolo, P. Tedeschi, G. Ursino, *How Limiting Deceptive Practices Harms Consumers*, maggio 2015.
24. A.K.S. Chand, S. Currarini, G. Ursino, *Cheap Talk with Correlated Signals*, maggio 2015.
25. S. Piccolo, P. Tedeschi, G. Ursino, *Deceptive Advertising with Rational Buyers*, giugno 2015.

26. S. Piccolo, E. Tarantino, G. Ursino, *The Value of Transparency in Multidivisional Firms*, giugno 2015.
27. G. Ursino, *Supply Chain Control: a Theory of Vertical Integration*, giugno 2015.
28. I. Aldasoro, D. Delli Gatti, E. Faia, *Bank Networks: Contagion, Systemic Risk and Prudential Policy*, luglio 2015.
29. S. Moriconi, G. Peri, *Country-Specific Preferences and Employment Rates in Europe*, settembre 2015.
30. R. Crinò, L. Ogliari, *Financial Frictions, Product Quality, and International Trade*, settembre 2015.
31. J. Grazzini, A. Spelta, *An empirical analysis of the global input-output network and its evolution*, ottobre 2015.
32. L. Cappellari, A. Di Paolo, *Bilingual Schooling and Earnings: Evidence from a Language-in-Education Reform*, novembre 2015.
33. A. Litina, S. Moriconi, S. Zanaj, *The Cultural Transmission of Environmental Preferences: Evidence from International Migration*, novembre 2015.
34. S. Moriconi, P. M. Picard, S. Zanaj, *Commodity Taxation and Regulatory Competition*, novembre 2015.
35. M. Bordignon, V. Grembi, S. Piazza, *Who do you blame in local finance? An analysis of municipal financing in Italy*, dicembre 2015.
36. A. Spelta, *A unified view of systemic risk: detecting SIFIs and forecasting the financial cycle via EWSs*, gennaio 2016.
37. N. Pecora, A. Spelta, *Discovering SIFIs in interbank communities*, febbraio 2016.
38. M. Botta, L. Colombo, *Macroeconomic and Institutional Determinants of Capital Structure Decisions*, aprile 2016.
39. A. Gamba, G. Immordino, S. Piccolo, *Organized Crime and the Bright Side of Subversion of Law*, maggio 2016.
40. L. Corno, N. Hildebrandt, A. Voena, *Weather Shocks, Age of Marriage and the Direction of Marriage Payments*, maggio 2016.
41. A. Spelta, *Stock prices prediction via tensor decomposition and links forecast*, maggio 2016.
42. T. Assenza, D. Delli Gatti, J. Grazzini, G. Ricchiuti, *Heterogeneous Firms and International Trade: The role of productivity and financial fragility*, giugno 2016.
43. S. Moriconi, *Taxation, industry integration and production efficiency*, giugno 2016.
44. L. Fiorito, C. Orsi, *Survival Value and a Robust, Practical, Joyless Individualism: Thomas Nixon Carver, Social Justice, and Eugenics*, luglio 2016.
45. E. Cottini, P. Ghinetti, *Employment insecurity and employees' health in Denmark*, settembre 2016.
46. G. Cecere, N. Corrocher, M. L. Mancusi, *Financial constraints and public funding for eco-innovation: Empirical evidence on European SMEs*, settembre 2016.
47. E. Brenna, L. Gitto, *Financing elderly care in Italy and Europe. Is there a common vision?*, settembre 2016.
48. D. G. C. Britto, *Unemployment Insurance and the Duration of Employment: Theory and Evidence from a Regression Kink Design*, settembre 2016.
49. E. Caroli, C.Lucifora, D. Vigani, *Is there a Retirement-Health Care utilization puzzle? Evidence from SHARE data in Europe*, ottobre 2016.
50. G. Femminis, *From simple growth to numerical simulations: A primer in dynamic programming*, ottobre 2016.
51. C. Lucifora, M. Tonello, *Monitoring and sanctioning cheating at school: What works? Evidence from a national evaluation program*, ottobre 2016.

52. A. Baglioni, M. Esposito, *Modigliani-Miller Doesn't Hold in a "Bailinable" World: A New Capital Structure to Reduce the Banks' Funding Cost*, novembre 2016.
53. L. Cappellari, P. Castelnovo, D. Checchi, M. Leonardi, *Skilled or educated? Educational reforms, human capital and earnings*, novembre 2016.
54. D. Britto, S. Fiorin, *Corruption and Legislature Size: Evidence from Brazil*, dicembre 2016.
55. F. Andreoli, E. Peluso, *So close yet so unequal: Reconsidering spatial inequality in U.S. cities*, febbraio 2017.
56. E. Cottini, P. Ghinetti, *Is it the way you live or the job you have? Health effects of lifestyles and working conditions*, marzo 2017.
57. A. Albanese, L. Cappellari, M. Leonardi, *The Effects of Youth Labor Market Reforms: Evidence from Italian Apprenticeships*; maggio 2017.
58. S. Perdichizzi, *Estimating Fiscal multipliers in the Eurozone. A Nonlinear Panel Data Approach*, maggio 2017.
59. S. Perdichizzi, *The impact of ECBs conventional and unconventional monetary policies on European banking indexes returns*, maggio 2017.
60. E. Brenna, *Healthcare tax credits: financial help to taxpayers or support to higher income and better educated patients? Evidence from Italy*, giugno 2017.
61. G. Gokmen, T. Nannicini, M. G. Onorato, C. Papageorgiou, *Policies in Hard Times: Assessing the Impact of Financial Crises on Structural Reforms*, settembre 2017.
62. M. Tettamanzi, *E Many Pluribus Unum: A Behavioural Macro-Economic Agent Based Model*, novembre 2017.
63. A. Boitani, C. Punzo, *Banks' leverage behaviour in a two-agent New Keynesian model*, gennaio 2018.
64. M. Bertoni, G. Brunello, L. Cappellari, *Parents, Siblings and Schoolmates. The Effects of Family-School Interactions on Educational Achievement and Long-term Labor Market Outcomes*, gennaio 2018.
65. G. P. Barbetta, G. Sorrenti, G. Turati, *Multigrading and Child Achievement*, gennaio 2018.
66. S. Gagliarducci, M. G. Onorato, F. Sobbrio, G. Tabellini, *War of the Waves: Radio and Resistance During World War II*, febbraio 2018.
67. P. Bingley, L. Cappellari, *Workers, Firms and Life-Cycle Wage Dynamics*, marzo 2018.
68. A. Boitani, S. Perdichizzi, *Public Expenditure Multipliers in recessions. Evidence from the Eurozone*, marzo 2018.
69. M. Le Moglie, G. Turati, *Electoral Cycle Bias in the Media Coverage of Corruption News*, aprile 2018.
70. R. Davidson, A. Monticini, *Improvements in Bootstrap Inference*, aprile 2018.
71. R. Crinò, G. Immordino, S. Piccolo, *Fighting Mobile Crime*, giugno 2018.
72. R. Caminal, L. Cappellari, A. Di Paolo, *Linguistic skills and the intergenerational transmission of language*, agosto 2018.
73. E. Brenna, L. Gitto, *Adult education, the use of Information and Communication Technologies and the impact on quality of life: a case study*, settembre 2018.
74. M. Bordignon, Y. Deng, J. Huang, J. Yang, *Plunging into the Sea: Ideological Change, Institutional Environments and Private Entrepreneurship in China*, settembre 2018.
75. M. Bordignon, D. Xiang, L. Zhan, *Predicting the Effects of a Sugar Sweetened Beverage Tax in a Household Production Model*, settembre 2018.
76. C. Punzo, L. Rossi, *The Redistributive Effects of a Money-Financed Fiscal Stimulus*, gennaio 2019.
77. A. Baglioni, L. Colombo, P. Rossi, *Debt restructuring with multiple bank relationships*, gennaio 2019.

78. E. Cottini, P. Ghinetti, S. Moriconi, *Higher Education Supply, Neighbourhood effects and Economic Welfare*, febbraio 2019.
79. S. Della Lena, F. Panebianco, *Cultural Transmission with Incomplete Information: Parental Perceived Efficacy and Group Misrepresentation,* marzo 2019.
80. T. Colussi, Ingo E. Isphording, Nico Pestel, *Minority Salience and Political Extremism,* marzo 2019.
81. G. P. Barbetta, P. Canino, S. Cima, *Let's tweet again? The impact of social networks on literature achievement in high school students: Evidence from a randomized controlled trial*, maggio 2019.
82. Y. Brilli, C. Lucifora, A. Russo, M. Tonello, *Vaccination take-up and health: evidence from a flu vaccination program for the elderly*, giugno 2019.
83. C. Di Novi, M. Piacenza, S. Robone, G. Turati, *Does fiscal decentralization affect regional disparities in health? Quasi-experimental evidence from Italy,* luglio 2019.
84. L. Abrardi, L. Colombo, P. Tedeschi, *The Gains of Ignoring Risk: Insurance with Better Informed Principals,* luglio 2019.
85. A. Garnero, C. Lucifora, *Turning a Blind Eye? Compliance to Minimum Wages and Employment*, gennaio 2020.
86. M. Bordignon, M. Gamalerio, E. Slerca, G. Turati, *Stop invasion! The electoral tipping point in anti-immigrant voting,* marzo 2020.
87. D. Vigani, C. Lucifora, *Losing control? Unions' Representativeness, "Pirate" Collective Agreements and Wages*, marzo 2020.
88. S. L. Comi, E. Cottini, C. Lucifora, *The effect of retirement on social relationships: new evidence from SHARE*, maggio 2020.
89. A. Boitani, S. Perdichizzi, C. Punzo, *Nonlinearities and expenditure multipliers in the Eurozone*, giugno 2020.
90. R. A. Ramos, F. Bassi, D. Lang, *Bet against the trend and cash in profits*, ottobre 2020.
91. F. Bassi, *Chronic Excess Capacity and Unemployment Hysteresis in EU Countries. A Structural Approach*, ottobre 2020.
92. M. Bordignon, T. Colussi, *Dancing with the Populist. New Parties, Electoral Rules and Italian Municipal Elections*, ottobre 2020.
93. E. Cottini, C. Lucifora, G. Turati, D. Vigani, *Children Use of Emergency Care: Differences Between Natives and Migrants in Italy,* ottobre 2020.
94. B. Fanfani, *Tastes for Discrimination in Monopsonistic Labour Markets*, ottobre 2020.
95. B. Fanfani, *The Employment Effects of Collective Bargaining*, ottobre 2020.
96. O. Giuntella, J. Lonsky, F. Mazzonna, L. Stella, *Immigration Policy and Immigrants' Sleep. Evidence from DACA,* dicembre 2020.
97. E. Cottini, P. Ghinetti, E. Iossa, P. Sacco, *Stress and Incentives at Work*, gennaio 2021.
98. L. Pieroni, M. R. Roig, L. Salmasi, *Italy: immigration and the evolution of populism,* gennaio 2021.
99. L. Corno, E. La Ferrara, A. Voena, *Female Genital Cutting and the Slave Trade*, febbraio 2021.
100. O. Giuntella, L. Rotunno, L. Stella, *Trade Shocks, Fertility, and Marital Behavior* , marzo 2021.
101. P. Bingley, L. Cappellari, K. Tatsiramos, *Parental Assortative Mating and the Intergenerational Transmission of Human Capital*, aprile 2021.
102. F. Devicienti, B. Fanfani, *Firms' Margins of Adjustment to Wage Growth. The Case of Italian Collective Bargaining*; aprile 2021.
103. C. Lucifora, A. Russo, D. Vigani, *Does prescribing appropriateness reduce health expenditures? Main e_ects and unintended outcomes*, maggio 2021.

104. T. Colussi, *The Political Effects of Threats to the Nation: Evidence from the Cuban Missile Crisis*, giugno 2021.
105. M. Bordignon, N. Gatti, M. G. Onorato, *Getting closer or falling apart? Euro countries after the Euro crisis*, giugno 2021.
106. E. Battistin, M. Ovidi, *Rising Stars*, giugno 2021.
107. D. Checchi, A. Fenizia, C. Lucifora, PUBLIC SECTOR JOBS: *Working in the public sector in Europe and the US,* giugno 2021.
108. K. Aktas, G. Argentin, G. P. Barbetta, G. Barbieri, L. V. A. Colombo, *High School Choices by Immigrant Students in Italy: Evidence from Administrative Data*, luglio 2021.
109. B. Fanfani, C. Lucifora, D. Vigani, *Employer Association in Italy. Trends and Economic Outcomes*, luglio 2021.
110. F. Bassi, A. Boitani, *Monetary and macroprudential policy: The multiplier effects of cooperation,* settembre 2021.
111. S. Basiglio, A. Foresta, G. Turati, *Impatience and crime. Evidence from the NLSY97,* settembre 2021.
112. A. Baglioni, A. Monticini, D. Peel, *The Impact of the ECB Banking Supervision Announcements on the EU Stock Market*, novembre 2021.
113. E. Facchetti, L. Neri, M. Ovidi, *Should you Meet The Parents? The impact of information on non-test score attributes on school choice*, dicembre 2021.
114. M. Bratti, E. Cottini, P. Ghinetti, *Education, health and health-related behaviors: Evidence from higher education expansion*, febbraio 2022.