Effects of Passive Smoking on Prenatal and Infant Development: Lessons from the Past*

Carlo Ciccarelli†
Università di Roma Tor Vergata

Gianni De Fraja‡
University of Nottingham
Università di Roma Tor Vergata and CEPR

Daniela Vuri§
Università di Roma Tor Vergata and IZA

April 6, 2021

Abstract

This paper studies the effect of passive smoking on child development. We use data from a time when the adverse effects of smoking on health were not known and when tobacco was not an inferior good. This allows us to disentangle the effect on foetuses and infants of smoking from that of other indicators of social and economic conditions. We exploit a set of unique longitudinal historical datasets defined at a detailed level of geographical disaggregation, namely the 69 Italian provinces. The datasets record precise information on the per capita consumption of tobacco products, the heights of twenty-year old conscripts in the second half of the 19th century Italy, and other relevant controls. We find a strong negative effect of smoking in the period immediately before and after birth on the height at age 20. Results are robust to changes in specification and consistent across the height distribution.

JEL Numbers: I12, J13, N33

Keywords: Passive Smoking, Stature, Nineteenth Century Italy, Infant Development.

---

*We are grateful to Lucia Pozzi and Brian A'Hearn, Franco Peracchi, and Giovanni Vecchi for sharing their data of Italian infant mortality and conscripts' stature distributions, respectively, and to Brian A'Hearn, Emilia Arcaleni, Federico Belotti, Frank Chaloupka, James Rockey, Jesse Matheson, John Komlos, Marco Rovinello, Roberto Basile, Stefano Fachin, Tommaso Proietti, and the editor Jörg Baten and two referees of this journal for useful suggestions.

†Università di Roma Tor Vergata, Dipartimento di Economia e Finanza, Via Columbia 2, I-00133 Rome, Italy; email: carlo.ciccarelli@uniroma2.it

‡Nottingham School of Economics, Sir Clive Granger Building, University Park, Nottingham, NG7 2RD, UK, Università di Roma Tor Vergata, Dipartimento di Economia e Finanza, Via Columbia 2, I-00133 Rome, Italy, and CEPR, 90-98 Goswell Street, London EC1V 7DB, UK; email: defraja@economia.uniroma2.it.

§Università di Roma Tor Vergata, Dipartimento di Economia e Finanza, Via Columbia 2, I-00133 Rome, Italy, and IZA, Berlin; email: daniela.vuri@uniroma2.it.

1 Introduction

Ireland banned indoor smoking in public spaces in 2004. This was at least partly due to the wide scientific consensus, despite the occasional dissenting voice (Enstrom et al., 2003), on the strength of the negative externality constituted by passive smoking (see USEPA, 1993; IARC, 2004; USDoH, 2006, or for recent reviews, Flouris et al. 2009; Banderali et al. 2015). Many countries, no doubt in consequence of its somewhat unexpected public popularity (Koh et al., 2007), have followed Ireland's lead and enacted legislation to reduce exposure to other people's smoking (WHO, 2007).

Our paper adds to the literature on the effect of passive smoking. We contribute to the specific area of research on the consequence of exposure of yet-to-be-born and very young children on their height as adult. The typical approach of research on the effects of passive smoking on children[1] is to analyse individuals explicitly recruited for the study, which therefore tend to limit the samples size and the time span between exposure and effects, and thus the external validity of the study. One of the earliest studies is Goldstein (1971), who recorded the smoking habit of pregnant women and measured their children at age seven. Goldstein finds an effect of about one centimetre for smoking ten or more cigarettes per day whilst pregnant, controlling for a host of other potential factors. Many subsequent studies (an earlier meta-analysis is Meredith, 1975) confirm this effect. Among more recent studies in Goldstein's footsteps are Berkey et al. (1984), Little et al. (1994), and Koshy et al. (2010). Görlitz and Tamm's (2020) findings that many German women stop smoking prior to becoming pregnant, and may restart only by the time their children reach adulthood suggest diffuse awareness of the health consequences of smoking around children.

[1] An extensive review is Charlton's (1994), its conclusions generally confirmed by the many subsequent studies, such as Aycicek et al. (2005) and Seyedzadeh et al. (2012).

In this paper, rather than relying on relatively small samples and relatively short time intervals between passive smoking and the measure of its effects, we use aggregate geographical data, from various separate historical longitudinal datasets for Italy, with annual data from 1875 to 1910. These allow us to link the height of 20-year old men to their likely past exposure to smoke, controlling for relevant co-variates. The use of aggregate data is fairly common for the study of the effects of tax or other policy variables on smoking (Baltagi and Levin, 1986; Peterson et al., 1992; Stehr, 2005) or other activities affecting health (Choudhury and Conway, 2020), though rarer for the analysis of the effect of smoking (but see Lightwood and Glantz (2016) and the literature cited there for recent examples).

A novel aspect of our approach is to use data which refers to a period, the second half of the 19th century, when the adverse effects of smoking, let alone passive smoking, were unknown, and to a country, Italy, where tobacco consumption increased with income. Both these features of our data help reduce the impact of two potentially confounding factors which nowadays may make it difficult to disentangle the direct effects of passive smoking from those of other unobservable behaviours which may also harm a foetus or an infant. Firstly, today's children growing up in family environments where adults exhibit disregard for their health by exposing them to passive smoking are more likely also to be suffering from other manifestations of the adults' negligence, such as inadequate medical care, deficient or unhealthy nutrition, insufficient exercise, and so on. Secondly, smoking is today an inferior good (Chaloupka and Warner, 2000; Franks et al., 2007; Gospodinov and Irvine, 2009). Thus, today's lower income parents smoke more on average than their better off peers, and because of their lower income, may also be less able to provide an adequate growing environment for their children.[2] In 19th century Italy, smoking in the presence of children was not seen as negligent,

[2]Further complexity may be added by non-linearities, whereby tobacco appears to be a normal good for low income households (Kenkel et al., 2014).

and tobacco was a normal good (Ciccarelli and De Fraja, 2014), and so these two confounding factors were absent in the period we study.

Another possible factor which could possible cast doubts on the causal link between passive smoking and infant development is smoking by women, which would cause nicotine and other harmful substances to be passed by the mother through the placenta in pregnancy or when breastfeeding. For cultural, rather than health reasons, smoking among 19th century Italian women was very unusual or even outlawed.[3] Our data reports separately consumption of snuff, tobacco powder which is consumed without smoke (Rogoziński (1990): chewing tobacco, another tobacco product which is not smoked, was absent in Italy at the time). These modes of tobacco consumptions are now known to be as harmful as cigarette smoking (England et al., 2003, among many others), but snuff (like chewing tobacco) does not otherwise affect people other than the consumers (Napierala et al., 2016). Thus our finding that consumption of snuff does not influence the adult height of the newborn is an indirect indication that women did not consume tobacco products while pregnant or lactating, and so we can attribute to passive smoking any developmental effect we find.

In detail, we link data on the annual consumption of dozens of tobacco products in each of the 69 Italian provinces to the average height of 20-year old Italian males recorded at their conscription medical visit, also at the province level. Controlling for changes in income, parental education, child mortality, and other potential unobservable environmental influences on child development, which we proxy with industrial output, urbanisation density, and infrastructure, we find that changes in the incidence of smoking in a province predict changes in the

[3] See, for example, Elliot (2001), or Brandt (2007, p 69ff): the "Torches of Freedom" Easter Sunday Parade which established smoking in public as an emblem of the emancipation of women happened only in 1929 (Amos and Haglund, 2000). The nearest we have to quantitative data on tobacco consumption by gender is a report on the link between smoking and public health, where women are simply excluded from the computation of the smoking population, on the grounds that they smoke "never or exceptionally rarely" Scalzi (1868, p 8, and p 19).

future average height of Italian conscripts. We estimate that a 10% reduction in the quantity of smoking tobacco consumed in a province translates into a cohort approximately 1.3 millimetres (one twentieth of an inch) taller twenty years later. This is a non-negligible effect, since it is concentrated on those exposed to secondary smoking, and it is robust to different specifications.[4]

We also find that potential exposure to passive smoking at later ages does not affect subsequent adult height. This is in line with the cited medical evidence, which identifies other negative health effects on toddlers and older children,[5] but concludes that passive smoking at later ages does not cause stunting. Thus our estimation of the effect of exposure to passive smoking beyond age two should be seen as a placebo.

The paper is organised as follows. Section 2 presents the data and the simple econometric specification, and Section 3 the results. A brief conclusion follows.

2 Data and econometric specification

The paper is built around the idea that passive smoking directly harm normal body development in very young children: if more babies in a sufficiently large population are exposed to more passive smoking, they will be, on average, shorter when they are adult. Height as young adults has a long standing pedigree of being used as a proxy for physical development from conception onwards: the remarkable drop in the average recorded height of French conscripts is recognised as one of the effects of the climatic change in the middle of the seventeenth century (Parker, 2013). Komlos (1994), Komlos (2007), Peracchi and Arcaleni (2011), and Hatton et al. (2016) among others, use height as a measure of the adverse long-term effects of

[4]Lack of data makes it impossible to replicate Case and Paxson's (2008) analysis of the effect of height on earnings: the importance of height and physical strength for developing countries would suggest possibly relevant distributional effects.

[5]From asthma (Evans et al., 1987) and other respiratory problems (Lebowitz et al., 1982; Sherrill et al., 1992) to blood pressure (Seyedzadeh et al., 2012) and cell and tissue damage (Kosecik et al., 2005).

disease and economic conditions in various European times and countries.

We can formalise this by positing that a adult's height is determined by exposure to passive smoking during their development. Formally, if $H_{i,t}$ is the height individual i born in period t will reach as a twenty year old, that is in year $t+20$, then

$$H_{i,t} = \beta \mathbf{S}_{i,t} + \gamma \mathbf{X}_{i,t} + \varepsilon_{i,t}. \tag{1}$$

In (1), $\mathbf{S}_{i,t}$ is a vector of the relevant measures of tobacco consumption by those in regular contact with individual i at date t, $\mathbf{X}_{i,t}$ are other time-varying environmental factors potentially affecting adult height, and $\varepsilon_{i,t}$ is an idiosyncratic error, which includes any relevant fixed effects.

We are able to draw from several datasets, which, despite the underdeveloped conditions of the Kingdom of Italy in the second half of the 19th century, are remarkably rich in detail. We do not have individual level data, and so the dependent variable in (1), $H_{i,t}$, is the average height, measured in year $t+20$, of the conscripts born in year t in each of the 69 administrative units, called provinces and indexed by i, Italy was divided in at the time.[6] This is based on the data organised by A'Hearn, Peracchi and Vecchi (2009), which in turn they obtain from military archival data collected at the medical examination for conscription. The data is an average of all the young men in the cohort, even those who did not meet the minimum height required for military service and hence were not conscripted (Arcaleni, 2012, p 1). We direct the reader to A'Hearn et al. (2009) for a detailed discussion of some of the data problems they identify and their careful solutions. Absenteeism due to desertion[7] was relatively rare after the first half of the 1870s, and internal

[6]Caserta and Naples are together in the original data on conscript height, and so we aggregate the rest of the data of the two provinces, into a single unit of observations.

[7]The Ministry of war records that 3.5% of those summoned did not report for the military visit in 1875, 3.3% ten years later, but increasing to 6.1 in 1895, and 8.6 in 1905, presumably in consequence of the increase in emigration (Vercelli, 2019).

migration was limited and very close to the municipalities of birth.[8] External emigration was instead an important phenomenon in the period we consider,[9] and it accounted for most of the discrepancies between the conscription roll and those examined. Its effect on height is in general ambiguous as "emigrants can be selected either positively or negatively".[10] A further reason which allays a possible immigration bias in our results, is that there is no reason to think that intensity of smoking and propensity to emigrate from a province should be correlated. Boys are still growing in their late teens, which might affect the accuracy of the measurement of their adult height. Figure 1 reports the average age when the conscripts born in a given year where measured. It highlights in dark grey the subset of the period that we consider in our main regression, namely the conscripts born from 1875 to 1890; the lighter shade of grey are the years that we can include in our sample when we omit child mortality as an explanatory variable. As the figure indicates, these conscripts were measured when they were very close to 20 years of age. At any rate, A'Hearn et al. (2009) adjust carefully the figures, to calculate the theoretical height conscripts would have at age 20, to ensure comparability across years,[11] of those born in each of the calendar years considered, irrespective of the

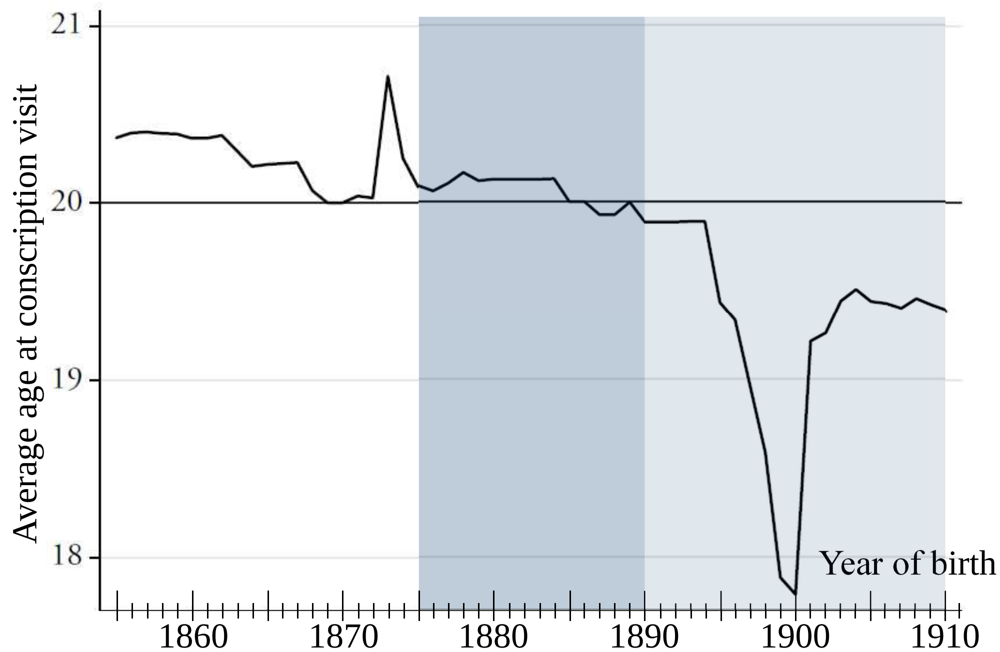
[8]Less than 10% of citizens of all ages were living in a commune different from that of birth, given the likely pattern of local migration, it seems natural that children were probably underrepresented in those moving far (Golini, 1976).

[9]"An estimated 14 million left Italy between 1871 and 1914" (Hatton and Williamson, 1998, p 95), though many came back soon. To put things in perspective, the Italian population in 1913 was around 35 million.

[10]As explained by A'Hearn et al. (2009, p 6). The limited evidence (Danubio et al., 2005) while suggesting positive selection, which would reduce the average height of the conscripts, relies on a small sample and lacks geographic details. Spitzer and Zimran (2018) also find a positive selection, at least at province level, and note the uneven distribution across regions and social classes of those landing in the US, but also point out that those liable to military service could not legally emigrate. In addition, we include province and year fixed effects, which should account for systematic difference in space and in time.

[11]This adjustment is necessary as teenage men are still growing so a lower height recorded for, say, the 1900 cohort would be rightly attributed to the fact that when those born in 1900 were measured they were more that two years younger that those born in 1880. They were of course conscripted at such a young age because of the dire need to replete the army in the final months of the first world war. We have not found any plausible explanation for the 1873 spike, which is anyway outside our period of interest.

Figure 1:
Average conscript age. 1855-1911



Note: National average age of the conscripts at the military visit. The darker grey area includes the years included in the main regression the light area the one in column (M4) in Table 3. Source: A'Hearn et al. (2009).

time when they were measured.

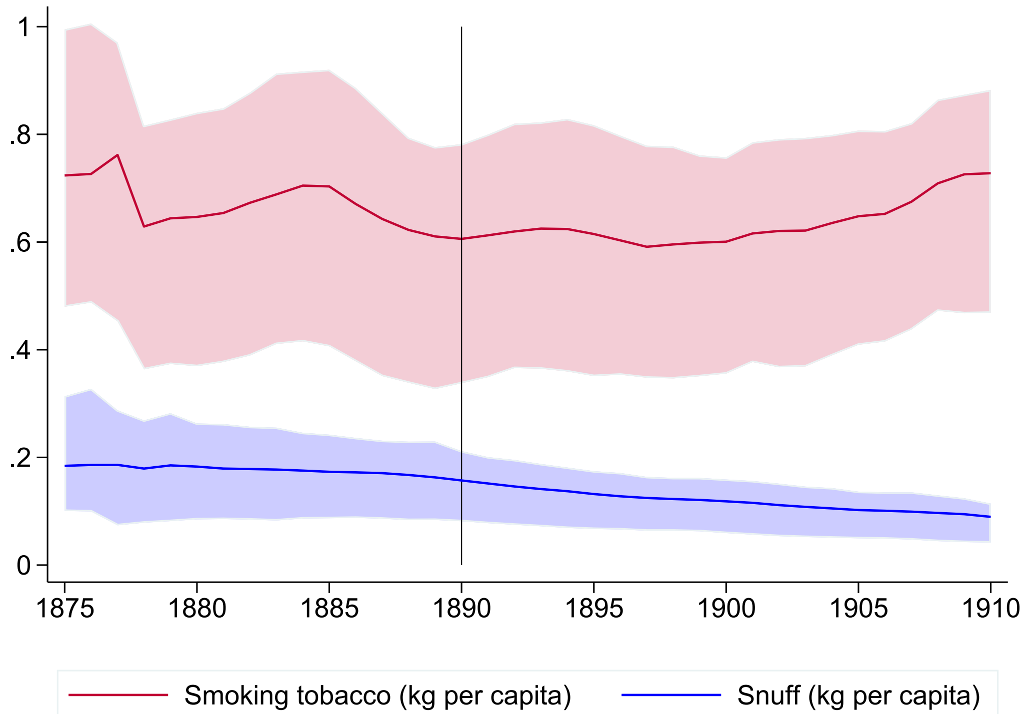
We proxy exposure to passive smoking during pregnancy and infancy in province i in year t with the per capita quantity of smoking tobacco sold in the province. This is collected in a dataset described in detail in Ciccarelli and De Fraja (2014). It reports the quantities sold of each individual tobacco product on sale in each province in Italy in each year, for around 40 years. Following that paper, we group all products into two categories, which we include separately in the regression: cigars, cigarettes,[12] and loose leaf or fine-cut tobacco, for rolling or use in pipes, denoted by $C_{i,t}$, and snuff, powder which is sniffed and not burnt, denoted by $F_{i,t}$. Formally, the vector $\mathbf{S}_{i,t}$ in (1) is given by $\mathbf{S}_{i,t} = (C_{i,t}, F_{i,t})$. Figure 2

[12]Cigarettes sales increased sharply in the period we consider, mainly due to the greater availability which followed the purchase of Bonsack cigarette rolling machines (Hannah, 2006, pp. 64–67), by the Italian manufacturers, at the expense of cigars sales.

illustrates the per capita consumption of the two product types, measured in kg, with the provincial differences highlighted in the coloured band which includes the middle half in the rank of the provinces, namely from the 25th to the 75th percentile, evidencing considerable national variation. We include the two types of products as separate regressors, because, in view of their respective modes of consumption, they are likely to differ in the extent the harmful components of the tobacco consumed by the primary user affect those around them. Spending time in a smoky environment harms physical development, whereas snuff can harm a person other than the primary user only through placental exchange and lactation. Given that the few women which used snuff were well past the child bearing age (Elliot, 2001), we expect the coefficient for $C_{i,t}$ and $F_{i,t}$ to differ, only the former indicating a significant effect. This is indeed what we find. The national average of between half and one kilogramme of tobacco per adult male per year corresponds to around three-to-five modern cigarettes per adult male per day: a rather small amount given that women rarely smoked in the period and men may not have always smoked in the presence of newborns. On average Italian households in 1901 had about four to five members present, with little variation across regions (Barbagli, 1984, p. 117, Livi, 1915, p. 33) which suggest only one adult male. Of course both smoking and the reduction in height are measured across the provincial population, and without specific information about individual level data it is impossible to distinguish between diffuse but small consumption of tobacco which affects each infant by a relatively small amount, and consumption of tobacco concentrated in very few households, whose infants are heavily harmed by passive smoking: if the relationship between passive smoking and height is linear, both cases would determine the same average consumption and the same average loss in height; and if not, they could not be distinguished.

Though constrained by the data availability, the vector of time-province specific

Figure 2:
Tobacco consumption in Italy



Note: The graph illustrates the per capita consumption of smoking tobacco and snuff in Italy from 1875 to 1910. Both are measured in kg per year. The coloured bands include the 25th and 75th centile of the province distribution of the year. Source Ciccarelli and De Fraja (2014).

controls, $\mathbf{X}_{i,t}$, does contain several of the variables the literature has identified as affecting child development. In detail, we include parental education, today an important determinant of both smoking (de Walque, 2007; Farrell and Fuchs, 1982) and health outcomes (Lindeboom et al., 2009; Shrestha, 2020). We proxy parental education with the percentage of the population who are literate, $\ell_{i,t}$, recorded in the census years, 1871, 1881, 1901, and 1911 (there was no population census in 1891), interpolating it for the non-census years.[13] Figure A1(b) highlight the very high education inequality across the country, also confirmed by (A'Hearn et al., 2011) analysis of primary and secondary enrolment rates at the regional level.

[13] Alternative measures, such as the literacy rate for the population over a certain age, or of one sex only, are all highly correlated with our measure.

Average real per capita income, $Y_{i,t}$, is a proxy for the province average socio-economic conditions, which in general affects child development. It is in turn proxied by certain categories of taxation, which, as argued in Ciccarelli and De Fraja (2014) and Brunetti et al. (2011), are very pro-cyclical, and, in addition, correlate strongly with independent measures of *regional* income in the period (Brunetti et al., 2011). As with literacy, there is considerable income inequality across the country (see Figure A1(c)). Note, however, that at 0.37, the correlation between income and literacy is lower than perhaps one might expect. An important determinant of children's hight is environmental pollution (Bailey et al., 2018; Rosales-Rueda and Triyana, 2019; Bobak et al., 2004). We do not have direct data for the period, but we can include three proxies: the industrial output, measured as the manufacturing value added, $O_{i,t}$, (obtained from Ciccarelli and Fenoaltea, 2013) the extent of the railway infrastructure,[14] $R_{i,t}$, following the literature which uses it to measure changes in the level of economic development.[15], and a measure of the rate of urbanisation, $U_{i,t}$, calculated as the proportion of the population of province i in year t who was living in a urban centre with more than 20,000 inhabitants (available in Giusti, 1913). We also interact the latter with consumption of smoking tobacco. The idea is that in a province with a high rate of urban living, an equal amount of smoking is more likely to cause pregnant mothers and newborns to be exposed to passive smoking, than one where dwelling are less cramped and rural life allows a greater extent of life in the open air. While we might expect a high level of urbanisation to increase average height, due to improved access to infant education and medical care, we should also expect that a higher rate of urban living should exacerbate the detrimental effects of passive smoking: thus we expect

[14]Calculated from (Ciccarelli and Groote, 2017); see `data.mendeley.com/datasets/rjwccjgy7z/1` which provides a historical geodatabase of the Italian railway network for the years 1839-1913.

[15]See Groote et al. (2009) for a historical example and Jedwab and Moradi (2016) for a recent one in developing countries.

the estimated coefficient for $U_{i,t} \times C_{i,t}$ to be negative.

Pozzi (2000) reports detailed data on natality and child mortality by province and by year in the second half of the 19th century Italy. We use these data as an attempt to control for the possible indirect effects of passive smoking on average adult height through infant mortality. Recent evidence for developed countries (Anderson and Cook, 1997) suggests that passive smoking by members of the family, controlling for pre-natal smoking by the mother, roughly doubles the likelihood of sudden infant death syndrome (SIDS, or "cot-death"). To the extent that this likelihood and the potential height reached by the infant as an adult are correlated, SIDS has a sample selection effect, and thus influences the average height of the cohort through a separate channel from the stunting and scarring that we aim to uncover in this paper. If the children more likely to succumb to cot-death are weaker than the average and therefore less likely to develop fully, then the selection effect of SIDS would prevent children destined to remain short from being measured at age 20, and hence *increase* the average height of those reaching the age of conscription. Bozzoli, Deaton and Quintana-Domeque (2009, p. 655-657) construct a model linking child mortality, whether due to passive smoking, especially severe winters, or other causes, and adult height. Their theoretical model is simulated, and, in line with their empirical evidence, suggests that the selection effect is stronger than the stunting effect when mortality is high. This is consistent with their observation that "Africans are [nowadays] relatively tall in spite of extremely unfavorable income and disease environments".[16] Given that infant mortality was very high in 19th century Italy (see Figure A1(a)), it seems plausible that the selection effect of child mortality did countervail the stunting

[16]A recent review of the literature on native African heights during the colonialisation period is Baten and Maravall (2021). Hatton (2011) examines historical town-level panel data on the heights of school children, and find no evidence for the selection effect, but some support for the stunting effect. Overall average height increased in England and Wales by about half a centimetre per decade in the first half of the twentieth century (p. 951).

effect of passive smoking: in this case our estimated coefficients would be lower bounds of the direct effect of passive smoking on adult height. To account for the possible selection effect of child mortality on average height, we include the post natal mortality rate, $m_{i,t}$, measured by the proportion of children born who die within one year of birth, in our main regression.[17]

Finally, we attempt to control for the parents' height by including in the regression the average height of the individuals who were 20 years before their military examination. While the age at marriage of men was around 27 (Livi Bacci, 1977, Table 2.22, p 100), and thus fathers had a considerably higher average age, as only a minority of the conscripts were first born, increasing the gap too much reduces the number of available observations. We take 20 as a compromise between not including parental age and including a realistic proxy for it.

To sum up, we use standard panel data analysis to estimate the following detailed specification of (1):

$$\begin{aligned} H_{i,t} =& \beta_C C_{i,t} + \beta_F F_{i,t} + \beta_\ell \ell_{i,t} + \beta_Y Y_{i,t} + \beta_R R_{i,t} + \beta_O O_{i,t} + \beta_U U_{i,t} + \beta_m m_{i,t} \\ & + \beta_H H_{i,t-20} + \alpha_i + \tau_t + \varepsilon_{i,t} \qquad t = 1, \ldots, T, \quad i = 1, \ldots, 68 \end{aligned} \tag{2}$$

In (2) $H_{i,t}$ is the average height at age 20 of individuals born in year t in province i. $C_{i,t}$ and $F_{i,t}$ are the log of the per capita consumption of tobacco in province i in year t. The controls $\ell_{i,t}$, $Y_{i,t}$, $R_{i,t}$, $O_{i,t}$, $U_{i,t}$, and $m_{i,t}$ measure the literacy rate, the log of (the proxy for) per capita income, the rail network, the industry value added, the rate of urbanisation, and the one-year mortality rate, all in province i in year t. The variable $H_{i,t-20}$ proxies the height of the conscripts' fathers. The province fixed effects, α_i

[17] This variable can be split in its two components of those dying in the first month and those who die after the first month, but before their first birthday. Replacing our measure with these variables, separately or in conjunction, does not alter the results, and neither does considering the mortality of boys only. We have also run the regressions with infant mortality in year $t+1$, to account for the possibility that a child death is due to event during the the first year of life, but happens a few months later. The results change at most marginally.

Table 1:
Summary Statistics.

	1875-1890				1875-1910			
	Mean	St. Dev	Min	Max	Mean	St. Dev	Min	Max
Smoking tobacco	0.66	0.33	0.20	1.90	0.64	0.31	0.20	1.90
Snuff	0.20	0.15	0.03	0.74	0.16	0.13	0.01	0.74
Average height	163.48	1.86	158.87	170.61	163.96	2.00	158.87	170.61
Child mortality	202.25	30.84	119.52	325.30				
Per capita income	5.04	3.16	1.82	33.74	5.83	3.46	1.82	33.74
Literacy rate	0.32	0.15	0.10	0.70	0.38	0.17	0.10	0.82
Railway network	145.70	100.80	0.00	526.37	192.14	133.35	0.00	782.49
Industrial output	31.39	27.98	3.8	203.2	33.17	30.43	3.8	242.2
Urbanisation rate	0.13	0.147	0	0.668	0.136	0.15	0	0.687
Observations	1074				2434			

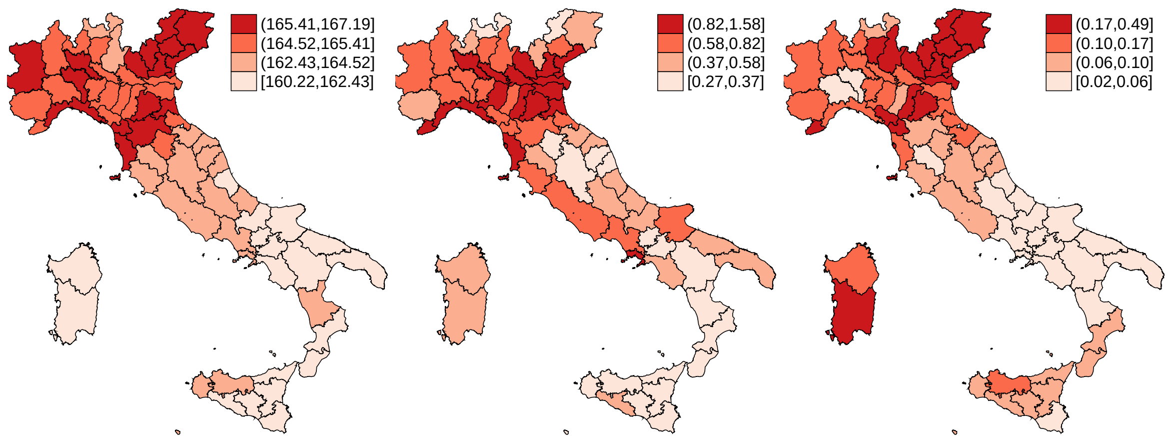
Notes: The period covered by the data is 1875-1910, except for child mortality, which ends in 1890. Smoking tobacco and snuff are the per capita annual consumption, in kg. Average height is the theoretical height in centimetres at age 20, computed by A'Hearn et al. (2009); child mortality is the percentage of live born children who died before the age of one, reported by Pozzi (2000). It is available only up to year 1890. The per capita income at 1911 prices is derived in Ciccarelli and De Fraja (2014), literacy rate is obtained by interpolation from census data, and the railway network is the end of the year total endowment of railway lines in the province, in km, obtained from Ciccarelli and Groote (2017). Industrial output is measured in million lira of value added in manufacturing at 1911 prices (Ciccarelli and Fenoaltea, 2013), and the urbanisation rate is the share of the population who was living in urban centres with at least 20,000 inhabitants, obtained from Giusti (1913).

account for time invariant factors, from genetic make-up to different geographical weather patterns, affecting provinces differently: these vary considerably across the national territory as shown in Figure 3(a). The variables τ_t, year fixed effects (or in a robustness test a time trend) account for changes in smoking habits and other countrywide factors affecting children's development over time. The appendix also reports a regression with province time trend.

Our analysis is constrained by the availability of data. Data on the conscripts height is available for those born from 1855 to 1910; data on the sales of tobacco products is instead available for all provinces from 1871 to 1910 (from 1877 for the seven Sicilian provinces), and data on mortality from 1872 to 1890. To include data on parental height and on child mortality we therefore need to restrict our data to

Figure 3:
Provincial disparities 1875-1890

3(a): Height age 20 3(b): Smoking tobacco 3(c): Snuff



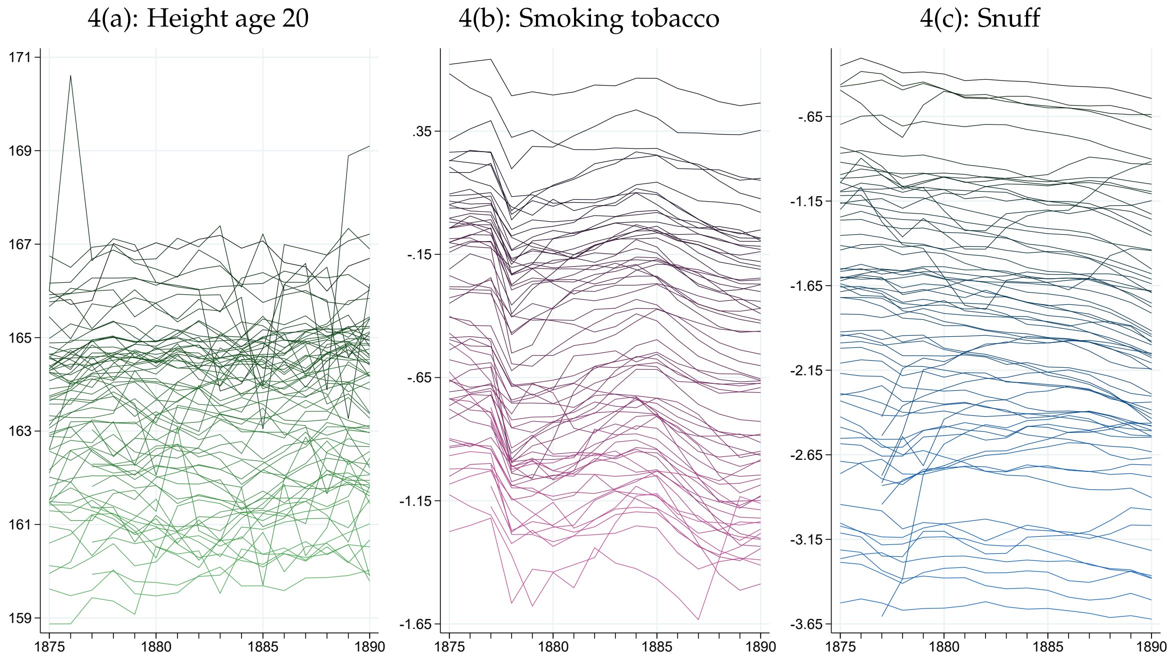
Note: Choropleth maps of the main variables included in the regressions. The maps report provincial averages of the variables evaluated over the period 1875-1890. The units of measurement are given in the footnote to Table 1. The maps for the control variables are in Figure A1. Sources: Tobacco from Ciccarelli and De Fraja (2014), heights from A'Hearn et al. (2009).

the those born between 1875 and 1890 inclusive, although we drop child mortality to extend the period in the robustness regression we report in the fourth column of Table 3.[18]

A child born on 1st January 1880 would be exposed while in uterus to tobacco smoked in her province in the year 1879. Children born later in the year would be exposed to a combination of the tobacco smoked in the later months on 1879 and in the months of 1880 before her birthday. And all children born from October would be exposed exclusively to tobacco smoked in 1880. If births, survivals, and exposure to passive smoking were all uniform throughout the year, the weights of years -21 and -20 would be approximately $\frac{1}{3}$ and $\frac{2}{3}$. But survival was certainly affected by the month of birth, as was the pattern of exposure to passive smoking, likely to more intense in the winter months, when more time is spent indoors, by active and

[18]We also note that the Sicilian provinces only entered in the smoking data in 1877, thus the panel is unbalanced prior to that year.

Figure 4:
Provincial breakdown 1875-1890



Note: Annual values, by province, of the dependent variable, average height of the conscripts, and of the per capita tobacco consumption, smoking tobacco and snuff (in logs). Sources: Tobacco from Ciccarelli and De Fraja (2014), heights from A'Hearn et al. (2009).

passive smokers alike. In addition, it is not necessarily the case that the potential harm of passive exposure to tobacco products is constant throughout the pregnancy. For all these reasons, it seems difficult and arbitrary to construct a realistic weighted average of the amount of tobacco smoked in years $t-1$ and t of those born in year t, and we simply take the amount of tobacco smoked in year t as the exposure to passive smoking during both pregnancy and neo-natal development.

Table 1 reports summary statistics for our datasets. Figure 3 illustrates differences across the country in the dependent variable, average male height at age 20, and the two main independent ones, consumption of cigars and cigarettes, and of snuff. The corresponding maps for the controls are in Figure A1 in the online appendix.

There may be a concern of a lack of sufficient variation both across provinces and within each province. To allay these possible concerns, we include Figure 4, which shows the time series of height and tobacco consumption in the 68 provinces, with the intensity of the colour in each panel determined by the province's average value of the variable across the period. While the provincial time series have a somewhat common pattern over time, and the overall level differs from province to province, they also cross each other repeatedly over time. Overall, the variability in the data helps identify the effect of interest.[19]

3 Empirical results

Table 2 reports the main results from the estimation of (1) in the time period 1875-1890. In the first column, we include only the variables measuring tobacco consumption of smoking tobacco and of snuff, both in kg per capita, with a time trend and province fixed effect. The sign for the consumption of tobacco is negative and statistically significantly different from zero for smoking tobacco, and not statistically significantly different from zero for snuff. Model (M2) in the second column confirms these coefficients when "biology" variables are added, with the plausible result that taller fathers beget taller children, and the suggestion in the negative sign of adverse factors in the year of birth affecting both the likelihood of a child dying and his future physical development. We further add, in the third column, socio-economic and environmental variables such as the province's average income, literacy rate, and our proxies for exposure to pollution. These are not significant, with the exception of railway infrastructure, and the coefficients for the smoking variables do not change qualitatively. The literacy ratio has very limited time series variation within province (see Figure 2 in Ciccarelli and Elhorst, 2018): this is likely to explain its lack of significance (excluding it

[19]The spikes in 1876 and at the end of the period occur in the small province of Lucca, and could well be due to measurement errors. Our robustness tests will ensure that these plays no role in determining our results.

from the estimations does not alter any of the results). The estimates from this last regression imply that a 10% change in per-capita consumption of smoking tobacco (cigars and cut-tobacco) is associated with a reduction in future heights of about 1.3 millimetres.

The last two columns of the table consider a different time frame for the effects of smoking to influence a child's development. In model (M4), we change the time structure of the model. We assume that the dependent variable is affected by exposure to passive smoking, consumption of snuff, and all the other socio-demographic influences experienced when the child is in his first year of age. This is considered still an age when passive smoking can affect adult height (Banderali et al., 2015; Little et al., 1994; Aycicek et al., 2005): indeed, the results of this estimation confirms this to be the case. Formally, we estimate the following model:

$$\begin{aligned} H_{i,t} =& \beta_C C_{i,t+1} + \beta_F F_{i,t+1} + \beta_\ell \ell_{i,t+1} + \beta_Y Y_{i,t+1} + \beta_R R_{i,t+1} + \beta_m m_{i,t} \\ &+ \beta_H H_{i,t-20} + \alpha_i + \tau_t + \varepsilon_{i,t} \qquad t = 1, \ldots, T, \quad i = 1, \ldots, 68 \end{aligned} \tag{3}$$

In the last column of the main table, we run the same regression as in the previous column, but taking the values of the environmental and socio-demographic variables at $t + 6$ instead of $t + 1$. This gives the interpretation of the coefficients as the effect on adult height of passive smoking and of the other covariate on six year old boys. The literature provides evidence (Charlton, 1994) of many ways in which children's health can be affected by passive smoking, but, beyond a certain age when children are exposed, lower adult height is not one of them. Model (M5) should therefore be seen as a placebo regression, and we would expect no effect on height: indeed, we find none. Tables A1 and A2 in the online appendix report the regression results for ages 2-5, both on their own and in combination. In no case is potential exposure to passive smoking at age two or higher affecting adult height.

Having established the negative effect of passive smoking around the time of a

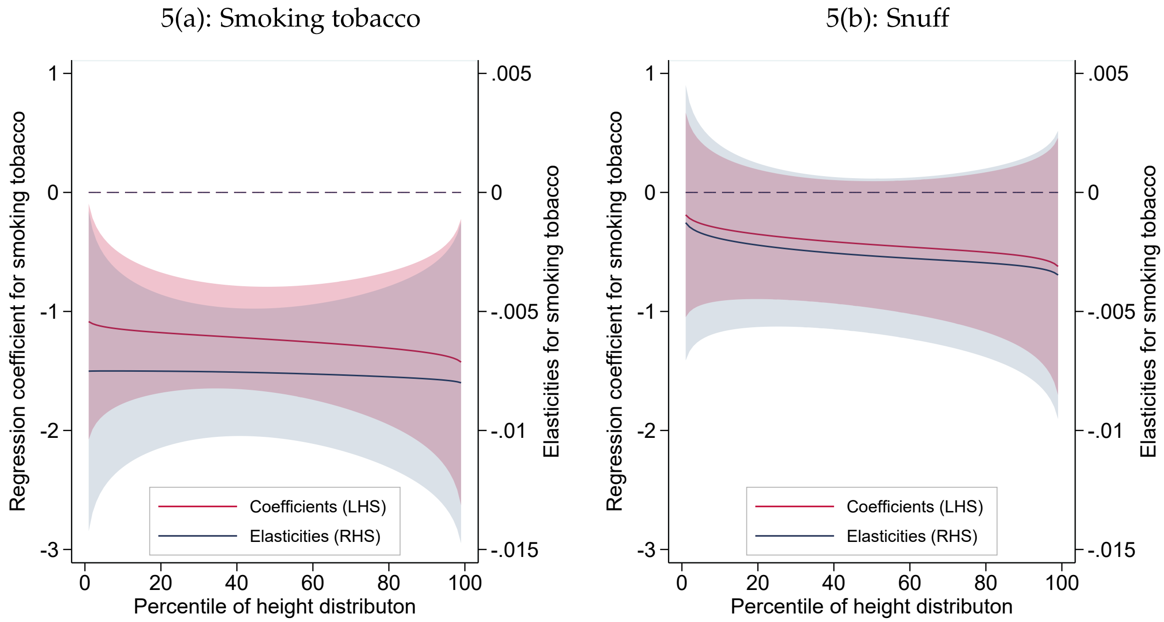
Table 2:
Height and consumption of tobacco: main regression results.

	(M1) f. effects only	(M2) health contr.	(M3) main regr.	(M4) effect on infants	(M5) effect on 6yr olds
Smoking tobacco	-1.108***	-1.075***	-1.293***	-1.033**	-0.741
	(0.354)	(0.357)	(0.335)	(0.417)	(0.477)
Snuff	-0.291	-0.324	-0.400	-0.255	0.156
	(0.255)	(0.245)	(0.282)	(0.337)	(0.335)
Father's height		0.114**	0.125***	0.109**	0.131***
		(0.047)	(0.046)	(0.045)	(0.045)
Child mortality		-0.180***	-0.180***	-0.171***	-0.168**
		(0.064)	(0.062)	(0.060)	(0.068)
Per capita income			-0.034	-0.202	-0.475
			(0.237)	(0.271)	(0.349)
Literacy rate			4.102	3.244	0.419
			(3.788)	(4.156)	(3.896)
Railways endowment			0.134**	0.062	-0.005
			(0.062)	(0.115)	(0.105)
Industrial output			0.006	0.005	0.001
			(0.005)	(0.004)	(0.004)
Urbanisation rate			0.102	0.114	0.123*
			(0.070)	(0.070)	(0.066)
Observations	1074	1074	1074	1074	1074

Notes: All regressions include the constant term and provincial fixed effects. Robust standard errors in parenthesis, clustered at the province level (68 clusters). *,**,*** denote significance at the 10, 5, and 1 percent level. The dependent variable is the height at age 20 of the Italian conscripts born in the years 1875-1890 in the 68 Italian provinces. The panel is unbalanced as the Sicilian provinces are included from 1877. The measurement of the variables are explained in the note to Table 1. All regressions include year and province fixed effects. In the various columns, (M1) contains no covariates, only year fixed effects. (M2) adds "genetic" covariates, and (M3) includes all covariates, it is our main regression. A t-test of the null hypothesis that the coefficients for smoking tobacco and snuff in (M3) are equal returns a p-value of 0.0657. In (M4) and (M5) the specification is the same as in (M3), but all independent variables are measured one and six years ahead, respectively. Further forward lags are in Table A1 in the online appendix.

child's birth, and the lack of effect of snuff consumption, we can delve deeper into the analysis by investigating if the effects differ across the height distribution. The results are reported in Figure 5. On both panels, the LHS axis measures the coefficients and the RHS the elasticity of height with respect to tobacco consumption, on

Figure 5:
Effect of tobacco consumption across the height distribution



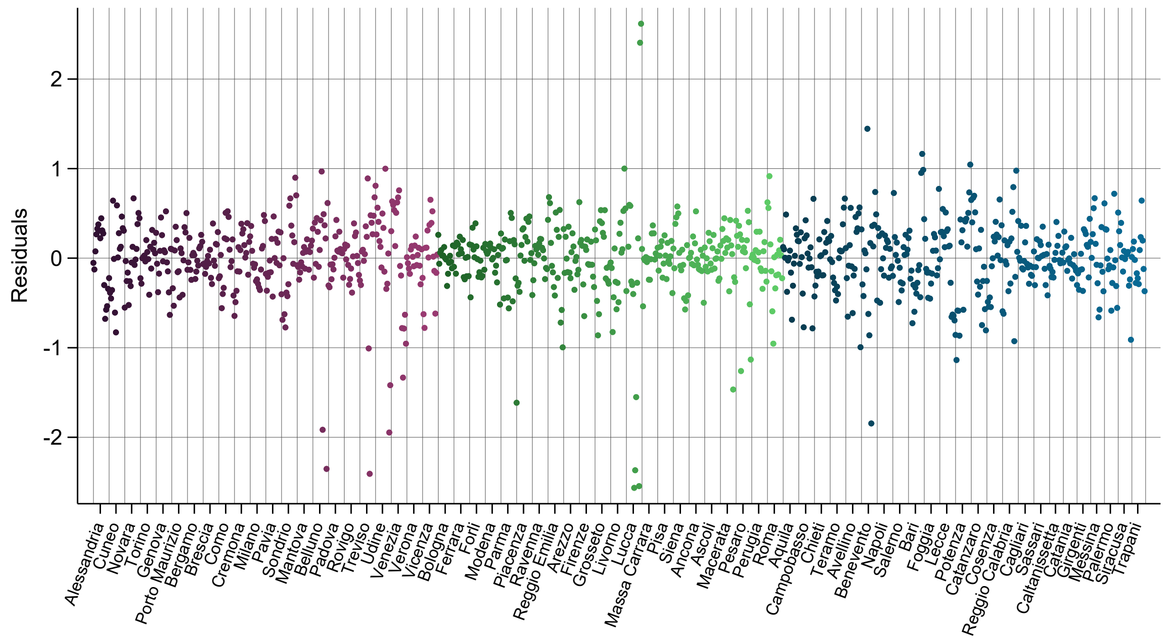
Note: The graph illustrates the coefficient and the confidence intervals for smoking tobacco obtained from estimating (2) with each percentile of the height distribution instead of the average as dependent variable.

panel 5(a) for smoking tobacco and on panel 5(b) for snuff. To obtain the values we have run (2) replacing the average height of the provincial cohort as the dependent variable with the *i*-th percentile of the height distribution.[20] The pictures confirms that the effect is similar for all heights. The shaded areas are the 95% confidence intervals.

Following Kelly's suggestion, (2020) we present in Figure 6 a scattergram of the residuals. The vertical axis in the graph reports the residuals from the main regression, Column (M3) in Table 2. The horizontal axis order the provinces geo-

[20]Following the standard assumption that human height is normally distributed in a homogeneous population (Tanner, 1978), we used the provincial mean and provincial standard deviation in the data used by A'Hearn et al. (2009) to determine the height at each percentile in each province. However, the results illustrated in Figure 5 should be interpreted as only suggestive, given that A'Hearn et al. (2009) highlight some divergence from normality in the Italian conscripts' height distribution (see in particular their Figure 3, p. 9).

Figure 6:
Scattergram of the regression residuals



Note: The vertical axis in the graph reports the residuals from the main regression, Column (M3) in Table 2. Data for each province are ordered by year. The first groups of provinces, with red dots of decreasing darkeness, are those in the north of Italy, followed by the green ones, the centre, and the blue ones, the south and islands.

graphically: the dots between two vertical lines are the residuals for all the years from 1875 to 1890 (from 1877 to 1890 for the seven Sicilian provinces) for the corresponding province. The first groups of provinces, with red dots of decreasing darkeness, are those in the regions in north of Italy, followed by the green ones, the centre, and the blues one, the south and islands. We have excluded the residual for Lucca in 1876, which has a value of 4, to avoid squeezing too much the rest of the observations. The scattergram in Figure 6 suggests that there are no substantial differences in the distribution of the residuals across space.[21] It also shows a few

[21]Besides visual inspection, we run (Jan Ditzen's Stata `xtcd2` command)) the Pesaran (2015) CD-test of the hypothesis of weak against strong cross-sectional dependence in the residuals of the main regression, Column (M3) in Table 2. Strong cross sectional dependence in the error term can lead to endogeneity and therefore to inconsistent estimates. Reassuringly, the average pairwise sample correlation in the residuals, which represents the building block of Pesaran's (2015) CD-test, is very low (indeed about zero) and the result of the test suggests that strong-cross sectional dependence

cases with a “poor fit”, such as the 24 out of 1074 observations where the residual exceeds 1 in absolute value.

Kelly (2020) also warns that results might be driven by a handful of outliers. To allay these concerns we run “leave-one-out regressions” and report the results in Section A.4 of the online appendix, Figure A3: this confirms that the few outliers do not drive our results. A complementary perspective is in Figure A4 in the same section of the appendix, which depicts box plots by province for the residuals of the main regression.

Table 3 reports further robustness tests. In column (M1) in Table 3 we replace heights with their natural logarithm, confirms that there are no qualitative differences in the regression results once elasticities are converted into centimetres. The same is true for the next column, model (M2), where we calculate spatial HAC errors following Conley (1999), with a 200km distance cut-off (further “spatial” robustness tests, following Cameron and Miller (2015) and Kelly (2020) are reported in Table A3 in the online appendix). Model (M3) replaces the year fixed effects with time trend: as can be seen, there are limited changes in the coefficients. In column (M4), we extend the period to all the years for which we have data, namely up to 1910. This comes at the sacrifice of the inclusion of child mortality as an explanatory variable, as the data is not available beyond 1890, but again confirms the result in the considerably longer dataset. In column (M5), we include in the regression the interaction term between the consumption of smoking tobacco and the urbanisation rate, defined above as the proportion of the population who reside in urban centres with a population exceeding 20,000. As explained above, one would expect the effect of this term to be negative, that is the urban environment worsen the effect of smoking tobacco on future adult height. While the estimated coefficient in column (M5) is negative as expected, it is not statistically different from 0: thus

does not affect our results (the CD statistic is -1.434, with associated p-value of 0.152, for the null hypothesis of weak cross-sectional dependence.)

it does not support our hypothesis. In the last column of this table, we replace all variables with their three-year moving average. The only qualitative difference with the main regression in column (M3) of Table 2 is the loss of significance for the child mortality, but the coefficient for the tobacco variables are unaltered. This suggest that the results are not driven by spikes in odd years or by the potentially arbitrary long-term time structure that we have posited.

We report further robustness test in the online appendix. In addition to those mentioned with regard to the geographical dispersion of the residuals, Table A1 confirms that exposure to passive smoking beyond a certain age does not affect adult height, and as also mentioned, Section A.4 examines whether individual provinces are outliers driving the results on their own. Finally, in Table A2 we include province fixed effects, we weight the regressions with provincial population, averaged across the period, and we include the first 4 year of life together and as an average.

4 Conclusion

This simple paper exploits a number of rich historical datasets collecting Italian provincial data from 1855 to 1910. We linked these datasets to determine the effect of being exposed to passive smoking during one's mother's pregnancy and one's first year of life on one's height as adult. We detect a non-negligible effect, even though, of course, the aggregate provincial data does not permit to distinguish among individuals subject to different levels of exposure to second-hand smoking. A 50% increase in the consumption of smoking tobacco in a province would cause the average height of the boys born in that province of over 6 millimetres. In addition we modify our main regression in several ways, and conclude that the result is robust, and does not depend on the details of the regression.

Today the harmful effects of passive smoking are well understood. A positive link found in contemporary studies between exposure to passive smoking and

Table 3:
Height and consumption of tobacco: additional regression results

	(M1) log height	(M2) HAC std. errors	(M3) time trend	(M4) long sample	(M5) urbanis. interact.	(M6) 3 year mov. avg.
Smoking tobacco	-0.008***	-1.293***	-1.116***	-1.072***	-1.141***	-1.144***
	(0.002)	(0.231)	(0.221)	(0.332)	(0.392)	(0.345)
Smoking tobacco×					-0.011	
urbanisation rate					(0.014)	
Snuff	-0.002	-0.400**	-0.495*	-0.087	-0.426	-0.275
	(0.002)	(0.189)	(0.268)	(0.286)	(0.284)	(0.250)
Father's height	0.001***	0.125***	0.101**	0.038	0.121**	0.136***
	(0.000)	(0.042)	(0.044)	(0.079)	(0.047)	(0.038)
Child mortality	-0.001***	-0.002**	-0.190***		-0.184***	-0.059
	(0.000)	(0.001)	(0.051)		(0.063)	(0.062)
Per capita income	-0.000	-.034	-0.067	-0.075	-0.052	-0.064
	(0.001)	(0.182)	(0.178)	(0.235)	(0.227)	(0.228)
Literacy rate	0.025	4.102	3.029	1.833	3.902	3.709
	(0.023)	(2.624)	(3.694)	(2.992)	(3.719)	(3.597)
Railways endowment	0.001**	0.134***	0.141**	0.129*	0.139**	0.092
	(0.000)	(0.046)	(0.061)	(0.068)	(0.062)	(0.070)
Industrial output	0.000	0.006*	0.006	0.005	0.006	0.005
	(0.000)	(0.003)	(0.005)	(0.004)	(0.005)	(0.004)
Urbanisation rate	0.001	0.102*	0.110	0.104	0.100	0.107
	(0.000)	(0.051)	(0.049)	(0.065)	(0.071)	(0.067)
Observations	1074	1074	1074	2434	1074	1074

Notes: *,**,*** denote significance at the 10, 5, and 1 percent level. Data and explanatory variables as in Table 2. The dependent variable is the logarithm of the average height in (M1). In (M2), standard errors are spatially clustered, and in (M3), we replace time fixed effect with a linear time trend. In (M4), we extend the sample to 1910, at the cost of having to omit the child mortality, for which we have no provincial data after 1890. In the last two columns we include an urbanisation×passive smoking interaction term, and we smooth the variables by taking their three year moving average.

adverse consequences may be biased by other variables which correlate both to this exposure and to the adverse effects, such as parental disregard for the health of their children and the household's socio-economic conditions and exposure to environmental pollution. Here we study a place and a time when passive smoking was not considered harmful, and, given that Italy did not begin its process of industrialisation until the late 19th century, general environmental pollution was unlikely to exert adverse effects. These considerations and the good quality of the data suggest that our findings provide an accurate and unbiased measure of the effects of passive smoking on children's early development.

References

A'Hearn, Brian, Claudio Auria, and Giovanni Vecchi, "Istruzione," in Giovanni Vecchi, ed., *In ricchezza e in povertà. Il benessere degli italiani dall'Unità a oggi*, Bologna: il Mulino, 2011, pp. 159–206.

—, Franco Peracchi, and Giovanni Vecchi, "Height and the normal distribution: Evidence from Italian military data," *Demography*, 2009, *46* (1), 1–25.

Amos, Amanda and Margaretha Haglund, "From social taboo to 'Torch of Freedom': The marketing of cigarettes to women," *Tobacco Control*, 2000, *9* (1), 3–8.

Anderson, H Ross and Derek G Cook, "Passive smoking and sudden infant death syndrome: Review of the epidemiological evidence," *Thorax*, 1997, *52* (11), 1003–1009.

Arcaleni, Emilia, "Secular trend and regional differences in the stature of Italians," *Journal of Anthropological Sciences*, 2012, *90*, 233–237.

Aycicek, Ali, Ozcan Erel, and Abdurrahim Kocyigit, "Increased oxidative stress in infants exposed to passive smoking," *European Journal of Pediatrics*, 2005, *164* (12), 775–778.

Bailey, Roy E, Timothy J Hatton, and Kris Inwood, "Atmospheric pollution, health, and height in late nineteenth century Britain," *The Journal of Economic History*, 2018, *78* (4), 1210–1247.

Baltagi, Badi H and Dan Levin, "Estimating dynamic demand for cigarettes using panel data: the effects of bootlegging, taxation and advertising reconsidered," *The Review of Economics and Statistics*, 1986, pp. 148–155.

Banderali, G, A Martelli, M Landi, F Moretti, F Betti, G Radaelli, C Lassandro, and E Verduci, "Short and long term health effects of parental tobacco smoking during pregnancy and lactation: A descriptive review," *Journal of Translational Medicine*, 2015, *13* (1), 327.

Barbagli, Marzio, *Sotto lo stesso tetto: Mutamenti della famiglia in Italia dal XV al XX secolo*, il Mulino, 1984.

Baten, Jörg and Laura Maravall, "The influence of colonialism on Africa's welfare: An anthropometric study," *Journal of Comparative Economics*, 2021, p. forthcoming.

Berkey, Catherine S, James H Ware, Frank E Speizer, and Benjamin G Ferris Jr, "Passive smoking and height growth of preadolescent children," *International Journal of Epidemiology*, 1984, *13* (4), 454–458.

Bobak, Martin, Marcus Richards, and Michael Wadsworth, "Relation between children's height and outdoor air pollution from coal-burning sources in the British 1946 birth cohort," *International Archives of Occupational and Environmental Health*, 2004, *77* (6), 383–386.

Bozzoli, Carlos, Angus Deaton, and Climent Quintana-Domeque, "Adult height and childhood disease," *Demography*, 2009, *46* (4), 647–669.

Brandt, Allan M, *The Cigarette C: The Rise Fall, and Deadly Persistence of the Product That Defined America*, Basics Books, 2007.

Brunetti, Alessandro, Emanuele Felice, and Giovanni Vecchi, "Reddito," in Giovanni Vecchi, ed., *In ricchezza e in povertà. Il benessere degli italiani dall'Unità a oggi*, Bologna: il Mulino, 2011, pp. 209–234.

Cameron, A Colin and Douglas L Miller, "A practitioner's guide to cluster-robust inference," *Journal of Human Resources*, 2015, *50* (2), 317–372.

Case, Anne and Christina Paxson, "Stature and status: Height, ability, and labor market outcomes," *Journal of Political Economy*, 2008, *116* (3), 499–532.

Chaloupka, Frank J and Kenneth E Warner, "The economics of smoking," in Anthony J Culyer and Joseph P Newhouse, eds., *Handbook of Health Economics. Volume 1, Part B*, Amsterdam: North-Holland, 2000, pp. 1539–1627.

Charlton, Anne, "Children and passive smoking: A review," *Journal of Family Practice*, 1994, *38* (3), 267–278.

Choudhury, Rebecca Sen and Karen Smith Conway, "The effect of tobacco policies on youth physical activity," *Economics & Human Biology*, 2020, *38*, 100872.

Ciccarelli, Carlo and Gianni De Fraja, "The demand for tobacco in post-unification Italy," *Cliometrica*, 2014, *8* (2), 145–171.

_ **and J Paul Elhorst**, "A dynamic spatial econometric diffusion model with common factors: The rise and spread of cigarette consumption in Italy," *Regional Science and Urban Economics*, 2018, *72*, 131–142.

_ **and Peter Groote**, "Railway endowment in Italy's provinces, 1839-1913," *Rivista di storia economica*, 2017, *33* (1), 45–88.

_ **and Stefano Fenoaltea**, "Through the magnifying glass: Provincial aspects of industrial growth in post-unification Italy," *The Economic History Review*, 2013, *66* (1), 57–85.

Conley, Timothy G, "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Danubio, Maria Enrica, Elisa Amicone, and Rita Vargiu, "Height and BMI of Italian immigrants to the USA, 1908–1970," *Economics & Human Biology*, 2005, *3* (1), 33–43.

de Walque, Damien, "Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education," *Journal of Health Economics*, 2007, *26* (5), 877–895.

Elliot, Rosemary Elizabeth, "'Destructive but sweet': Cigarette smoking among women 1890-1990." PhD dissertation, University of Glasgow 2001.

England, Lucinda J, Richard J Levine, James L Mills, Mark A Klebanoff, F Yu Kai, and Sven Cnattingius, "Adverse pregnancy outcomes in snuff users," *American Journal of Obstetrics and Gynecology*, 2003, *189* (4), 939–943.

Enstrom, James E, Geoffrey C Kabat, and Davey Smith, "Environmental tobacco smoke and tobacco related mortality in a prospective study of Californians, 1960-98," *British Medical Journal*, 2003, *326* (7398), 1057.

Evans, David, Moshe J Levison, Charles H Feldman, Noreen M Clark, Yvonne Wasilewski, Bruce Levin, and Robert B Mellins, "The impact of passive smoking on emergency room visits of urban children with asthma," *American Review of Respiratory Disease*, 1987, *135* (3), 567–572.

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

Flouris, Andreas D, Constantine I Vardavas, Giorgos S Metsios, Aristidis M Tsatsakis, and Yiannis Koutedakis, "Biological evidence for the acute health effects of secondhand smoke exposure," *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 2009, *298* (1), L3–L12.

Franks, Peter, Anthony F Jerant, J Paul Leigh, Dennis Lee, Alan Chiem, Ilene Lewis, and Sandy Lee, "Cigarette prices, smoking, and the poor: Implications of recent trends," *American Journal of Public Health*, 2007, *97* (10), 1873–1877.

Giusti, Ugo, *L'addensamento e l'affollamento nei centri urbani italiani al 10 giugno 1911*, Alfani e Venturi, Firenze, 1913.

Goldstein, H, "Factors influencing the height of seven year old children: Results from the National Child Development Study," *Human Biology*, 1971, pp. 92–111.

Golini, Antonio, "Le migrazioni interne," in Ruggiero Romano and Corrado Vivanti, eds., *Storia d'Italia. VI. Atlante*, Turin: Giulio Einaudi Editore, 1976, pp. 723–730.

Görlitz, Katja and Marcus Tamm, "Parenthood and smoking," *Economics & Human Biology*, 2020, *38*, 100874.

Gospodinov, Nikolay and Ian Irvine, "Tobacco taxes and regressivity," *Journal of Health Economics*, 2009, *28* (2), 375–384.

Groote, Peter D, J Paul Elhorst, and PG Tassenaar, "Standard of living effects due to infrastructure improvements in the 19th century," *Social Science Computer Review*, 2009, *27* (3), 380–389.

Hannah, Leslie, "The Whig fable of American tobacco, 1895–1913," *The Journal of Economic History*, 2006, *66* (1), 42–73.

Hatton, Timothy J, "Infant mortality and the health of survivors: Britain, 1910–50," *Economic History Review*, 2011, *64* (3), 951–972.

— **and Jeffrey G Williamson**, *The Age of Mass Migration: Causes and Economic Impact*, Oxford University Press, 1998.

— **, Roy E Bailey, and Kris Inwood**, "Health, height and the household at the turn of the 20th century," *Economic History Review*, 2016, *69* (1), 35–53.

IARC, International Agency for Research on Cancer, *Tobacco smoke and involuntary smoking*, Vol. 83, "World Health Organization, Lyon, France", 2004.

Jedwab, Remi and Alexander Moradi, "The permanent effects of transportation revolutions in poor countries: Evidence from Africa," *The Review of Economics and Statistics*, 2016, *98* (2), 268–284.

Kelly, Morgan, "Understanding Persistence," Technical Report, DP15246, CEPR, London 2020.

Kenkel, Donald S, Maximilian D Schmeiser, and Carly Urban, "Is smoking inferior? Evidence from variation in the earned income tax credit," *Journal of Human Resources*, 2014, *49* (4), 1094–1120.

Koh, Howard K, Luk X Joossens, and Gregory N Connolly, "Making smoking history worldwide," *New England Journal of Medicine*, 2007, *356* (15), 1496–1498.

Komlos, John, *Stature, Living Standards, and Economic Development: Essays in Anthropometric History*, University of Chicago Press, 1994.

— , "Anthropometric evidence on economic growth, biological well-being and regional convergence in the Habsburg Monarchy, c. 1850–1910," *Cliometrica*, 2007, *1* (3), 211–237.

Kosecik, Mustafa, Ozcan Erel, Eylem Sevinc, and Sahabettin Selek, "Increased oxidative stress in children exposed to passive smoking," *International Journal of Cardiology*, 2005, *100* (1), 61–64.

Koshy, Gibby, Ali Delpisheh, and Bernard J Brabin, "Dose response association of pregnancy cigarette smoke exposure, childhood stature, overweight and obesity," *The European Journal of Public Health*, 2010, *21* (3), 286–291.

Lebowitz, Michael D, David B Armet, and Ronald Knudson, "The effect of passive smoking on pulmonary function in children," *Environment International*, 1982, *8* (1-6), 371–373.

Lightwood, James and Stanton A Glantz, "Smoking behavior and healthcare expenditure in the United States, 1992–2009: Panel data estimates," *PLoS medicine*, 2016, *13* (5), e1002020.

Lindeboom, Maarten, Ana Llena-Nozal, and Bas van Der Klaauw, "Parental education and child health: Evidence from a schooling reform," *Journal of Health Economics*, 2009, *28* (1), 109–131.

Little, Ruth E, M Dow Lambert III, Bonnie Worthington-Roberts, and Cynthia H Ervin, "Maternal smoking during lactation: Relation to infant size at one year of age," *American Journal of Epidemiology*, 1994, *140* (6), 544–554.

Livi Bacci, Massimo, *A History of Italian Fertility during the Last Two Centuries* Office of Population Research Series, Princeton University Press, 1977.

Livi, Livio, *La composizione delle famiglie*, Tipografia M. Ricci, Florence, 1915.

Meredith, Howard V, "Relation between tobacco smoking of pregnant women and body size of their progeny: A compilation and synthesis of published studies," *Human Biology*, 1975, pp. 451–472.

Napierala, Marta, Jan Mazela, T Allen Merritt, and Ewa Florek, "Tobacco smoking and breastfeeding: Effect on the lactation process, breast milk composition and infant development. A critical review," *Environmental Research*, 2016, *151*, 321–338.

Parker, Geoffrey, *Global Crisis: War, Climate Change and Catastrophe in the Seventeenth Century*, Yale University Press, 2013.

Peracchi, Franco and Emilia Arcaleni, "Early-life environment, height and BMI of young men in Italy," *Economics & Human Biology*, 2011, *9* (3), 251–264.

Pesaran, M Hashem, "Testing weak cross-sectional dependence in large panels," *Econometric reviews*, 2015, *34* (6-10), 1089–1117.

Peterson, Dan E, Scott L Zeger, Patrick L Remington, and Henry A Anderson, "The effect of state cigarette tax increases on cigarette sales, 1955 to 1988," *American Journal of Public Health*, 1992, *82* (1), 94–96.

Pozzi, Lucia, *La lotta per la vita. Evoluzione e geografia della sopravvivenza in Italia fra '800 e '900*, Editrice Universitaria Udinese, Udine, 2000.

Rogoziński, Jan, *Smokeless Tobacco in the Western World, 1550-1950*, Praeger, 1990.

Rosales-Rueda, Maria and Margaret Triyana, "The persistent effects of early-life exposure to air pollution: Evidence from the Indonesian forest fires," *Journal of Human Resources*, 2019, *54* (4), 1037–1080.

Scalzi, Francesco, *Il consumo del tabacco da fumo in Roma in attinenza alla salute publica*, Stabilimento Tipografico di Via Corso, Roma, 1868.

Seyedzadeh, Abolhassan, Forough Hashemi, and Akram Soleimani, "Relationship between blood pressure and passive smoking in elementary school children," *Iranian Journal of Pediatrics*, 2012, *22* (3), 351.

Sherrill, Duane L, Fernando D Martinez, Michael D Lebowitz, M David Holdaway, Erin M Flannery, G Peter Herbison, Warren R Stanton, Phil A Silva, and Malcolm R Sears, "Longitudinal effects of passive smoking on pulmonary function in New Zealand children," *American Journal of Respiratory and Critical Care Medicine*, 1992, *145* (5), 1136–1141.

Shrestha, Vinish, "Maternal education and infant health gradient: New answers to old questions," *Economics & Human Biology*, 2020, *39*, 100894.

Spitzer, Yannay and Ariell Zimran, "Migrant self-selection: Anthropometric evidence from the mass migration of Italians to the United States, 1907–1925," *Journal of Development Economics*, 2018, *134*, 226–247.

Stehr, Mark, "Cigarette tax avoidance and evasion," *Journal of Health Economics*, 2005, *24* (2), 277–297.

Tanner, James M, *Fetus into man: Physical growth from conception to maturity*, Harvard University Press, 1978.

USDoH, "The health consequences of involuntary exposure to tobacco smoke: A report of the Surgeon General," 2006.

USEPA, *Respiratory Health Effects of Passive Smoking: Lung Cancer & Other Disorders*, Environmental Protection Agency, DIANE Publishing, 1993.

Vercelli, Claudio, *Soldati. Storia dell'esercito italiano*, Laterza, 2019.

WHO, "Protection from exposure to second-hand tobacco smoke: Policy recommendations," World Health Organization, Geneva 2007.

A Appendix: not for publication

A.1 Introduction

In this appendix we present some additional results and details of the data. We begin in Figure A1 with a set of choropleth maps illustrating the geographical distribution of the control variables included in the main regression (column (M3) in Table 2). These are are not reported in the main text for brevity. As we are averaging across years, the height of the fathers would look identical to Figure 3(a), and so we do not include it.

A.2 Further regressions

Tables A1 and A2 report the results of further regressions.

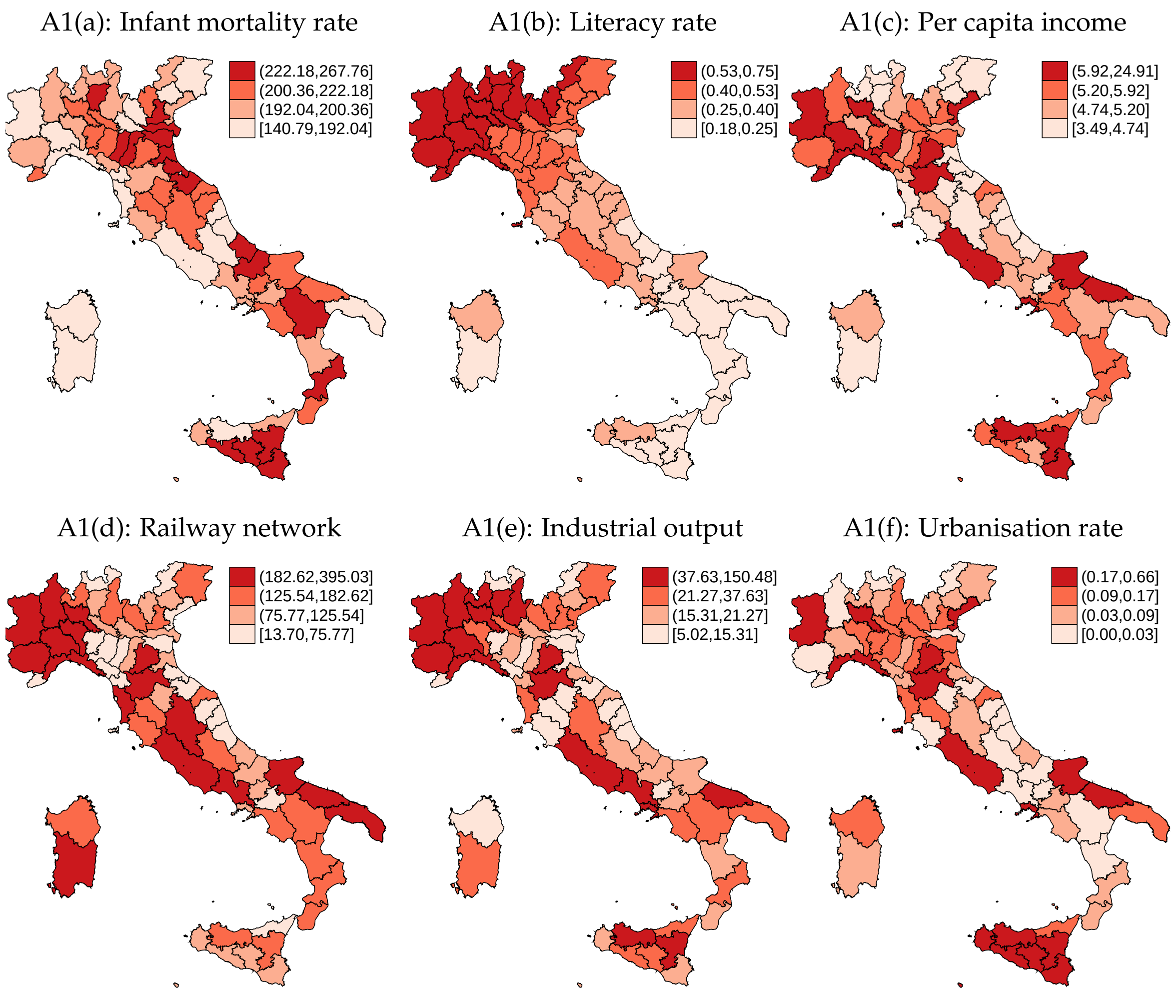
In Table A1, we complete the analysis of columns (M4) and (M5) in Table 2, by considering the effect of being subject to passive smoking at ages two, three, four, and five. The results at each of these ages are similar to those reported for age six in column (M5) of Table 2, namely there there is no effect on future height.

Consider now the next set of regressions, reported in Table A2. In Column (M1), the main regression, equation (2), which gives model (M3) in Table 2, is run by replacing the year fixed effects with a province time trend. Allowing this introduces significance to the coefficient for snuff. In (M2), we weight the coefficient using population (calculated as the period average): here there are small qualitative changes. In the next two column we considered the effects of smoking over a four year period from birth to age three: in Column (M3) the four years are added separately, whereas in Column (M4) as the average of the four years. Results do not change qualitatively.

A.3 Urbanisation rates: sources and methods

Nineteenth century Italy was at the beginning of the demographic transition. The population was young and growing at fast rates, both in urban centres and in the countryside. The urbanisation rate was also rising. To account for these demographic factors we included the urbanisation rate variable in our empirical models.

Figure A1:
Further provincial disparities 1875-1890



Note: Choropleth maps of the main variables included in the regressions. The maps report provincial averages of the variables evaluated over the period 1875-1890. The units of measurement and the source of the various data are given in the footnote to Table 1.

Table A1:
Height and consumption of tobacco: effects at older ages

	(M1) effect on 2 year olds	(M2) effect on 3 year olds	(M3) effect on 4 year olds	(M4) effect on 5 year olds
Smoking tobacco	-0.660	-0.552	-0.582	-0.572
	(0.446)	(0.402)	(0.384)	(0.366)
Snuff	-0.265	-0.190	0.039	0.192
	(0.367)	(0.360)	(0.337)	(0.321)
Father's height	0.114***	0.119***	0.116**	0.117**
	(0.043)	(0.043)	(0.045)	(0.047)
Child mortality	-0.167**	-0.180***	-0.172**	-0.178**
	(0.065)	(0.067)	(0.067)	(0.067)
Per capita income	-0.487	-0.493	-0.425	-0.301
	(0.316)	(0.356)	(0.330)	(0.286)
Literacy rate	1.164	0.932	1.284	1.764
	(4.318)	(4.223)	(4.225)	(4.164)
Railways endowment	0.068	0.072	0.045	0.015
	(0.115)	(0.120)	(0.117)	(0.119)
Industrial output	0.004	0.004	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.004)
Urbanisation rate	0.135*	0.140**	0.138**	0.137**
	(0.070)	(0.070)	(0.067)	(0.067)
Observations	1074	1074	1074	1074

Notes: *,**,*** denote significance at the 10, 5, and 1 percent level. Data and explanatory variables as in Table 2. The environmental independent variables (tobacco consumption, income, literacy, and environmental variables) in each column are measured when the cohort was age 2 to 5.

The urbanisation rate represents in particular one of the provinces' environmental control variables used in various specifications of our regression model.

The urbanization rates are defined as the ratio of the population living in an urban center (numerator) over total provincial population (denominator). The data for 1871, 1881, 1901, and 1911 on the population of urban centres are from Giusti (1913). The data on total provincial population for the same years are from the population censuses. Intercensal urbanisation rates were obtained by linear

Table A2:
Height and consumption of tobacco: further robustness checks

	(M1) province time trend	(M2) population weighting	(M3) up to 3 years forward	(M4) average $\{t, t+3\}$
Smoking tobacco t	-0.680***	-1.142***	-1.288***	
	(0.206)	(0.343)	(0.335)	
Smoking tobacco $t+1$			-0.121	
			(0.344)	
Smoking tobacco $t+2$			0.574	
			(0.486)	
Smoking tobacco $t+3$			-0.661*	
			(0.377)	
Avg smoking tobacco $\{t, t+3\}$				-1.278***
				(0.465)
Snuff at t	-0.775***	-0.481*	-0.502	
	(0.264)	(0.287)	(0.352)	
Snuff at $t+1$			0.236	
			(0.429)	
Snuff at $t+2$			-0.570	
			(0.772)	
Snuff at $t+3$			0.520	
			(0.675)	
Avg snuff $\{t, t+3\}$				-0.391
				(0.386)
Father's height	-0.016	0.152***	0.124**	0.114**
	(0.058)	(0.039)	(0.047)	(0.047)
Child mortality	-0.213***	-0.194***	-0.175***	-0.161**
	(0.054)	(0.067)	(0.062)	(0.062)
Per capita income	-0.030	-0.053	-0.012	-0.134
	(0.186)	(0.252)	(0.243)	(0.240)
Literacy rate	22.965	1.940	4.386	3.313
	(15.546)	(3.694)	(3.871)	(3.868)
Railways endowment	0.193**	0.131**	0.139**	0.129*
	(0.076)	(0.057)	(0.062)	(0.066)
Industrial output	0.016	0.010***	0.006	0.006
	(0.024)	(0.003)	(0.005)	(0.005)
Urbanisation rate	0.169	0.055	0.101	0.103
	(0.132)	(0.064)	(0.070)	(0.070)
Observations	1074	1074	1074	1074

Notes: *,**,*** denote significance at the 10, 5, and 1 percent level. Data and explanatory variables as in Table 2. See Section A.4 for details of the regressions.

interpolation. To define urbanisation rates we used the same threshold used by Giusti, that is, we included in the calculations of the numerator only municipalities with at least 20,000 inhabitants in June 1911. Figure A2 reports the territorial distribution, in February 1901, of Italian municipalities with a population of at least 20,000 inhabitants in their urban centre, when the population census was taken. Milan and Naples are the only cities with a population of half million or above. Turin, Genoa, Florence, and Rome are also big cities by the standard of the time. In the north, the urban network is formed by cities aligned through the horizontal west-east axis from Venice to Turin passing through Milan, and the vertical axis from Milan to Bologna and to Ancona on the Adriatic coast. In continental South, Naples is instead essentially the only genuine city, surrounded by several large towns which are separate administrative municipalities: this determines a high value of the population density in the province. At the opposite end, Sardinia is clearly a very rural region, with the only sizeable centres being the two provincial capital cities of Cagliari and Sassari.

A.4 Robustness checks

We have also attempted to ascertain whether one particular province was particularly influential in driving the results. We have done so by running repeatedly the specification in the main regression (equation (2), and Table 2 model (M3)) excluding one province at the time. Figure A3 reports the coefficient of the smoking tobacco obtained in the 68 estimations, and suggested that no province is crucial in obtaining our results. Figure A4 reports box plots, as explained in the note.

The same message is conveyed by the plot of the residuals against the fitted values shown in Figure A5: the whole sample in panel (a), and the central area which excludes the residual that exceed one in absolute value in panel (b) on the RHS. The pattern of the residuals highlight the presence of an outlier, the Rovigo province, which at the time was largely agrarian, and whose predicted values are to the left of the range. It also shows a distribution of fitted values with two local modes, with most provinces concentrating around 163cm with some with higher averages. The latter do not appear to be concentrated in some geographical areas: towards the RHS of the panels, there are dots of all colours. By the same token, the residuals we obtain appear to come from a distribution of the errors that is similar across provinces. Thus heteroskedasticity does not appear to be a problem at first sight. Similarly, the pattern of the residuals does not indicate that assumption of a linear fit leads to a misspecified model. These residuals obviously average 0 in every province; their provincial standard deviation has a mean of 0.372 and a standard deviation of 0.251, and ranges from 0.133 to 0.885, excluding the value of 1.85 of the Lucca province, an outlier which, as mentioned, may be due to measurement errors. In any case, since we calculate robust standard errors, the inference obtained from our regression would be valid even if heteroskedasticity were present. .

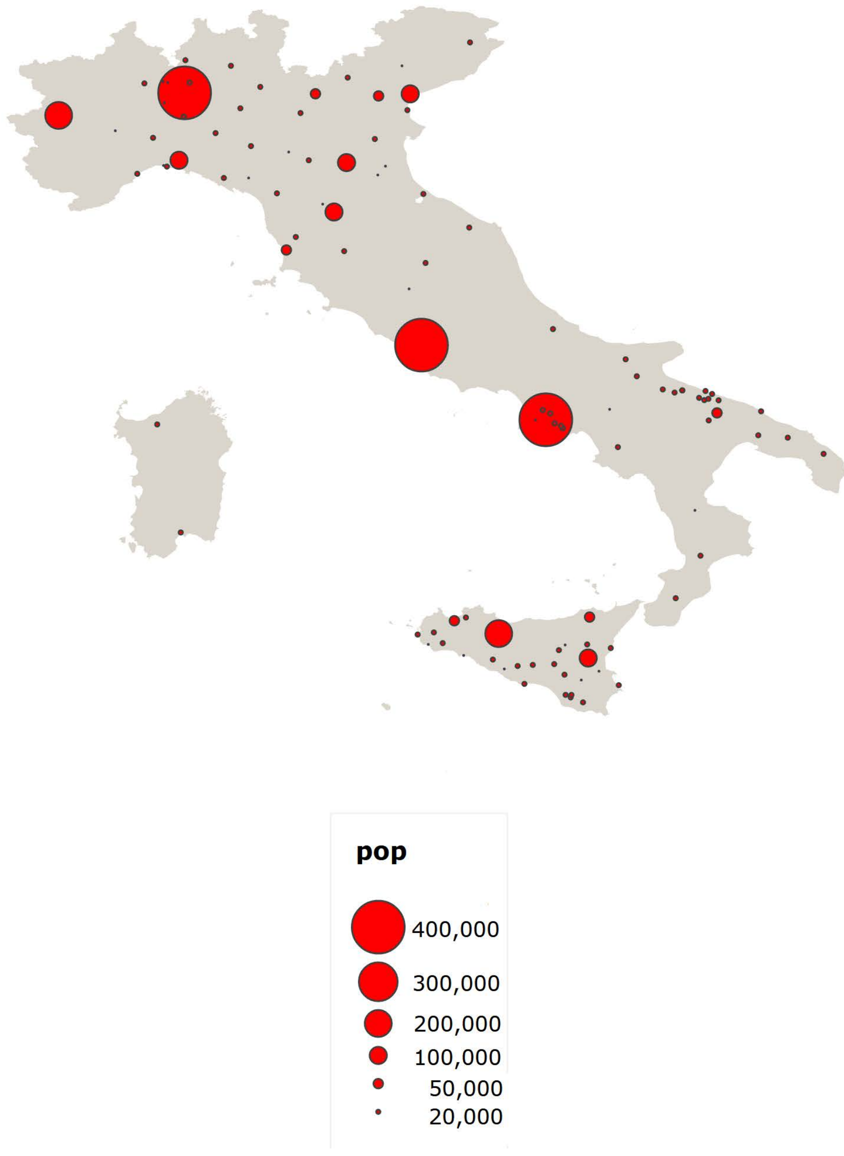
Finally, in the last table, Table A3, we report the results of some robustness with regard to possible spatial correlation of the residuals we have carried out. In detail, following Kelly (2020) we vary both the type of weight given to distance, and the distance itself. In addition, following Cameron and Miller (2015), we consider a higher level of geographical clustering than the province that we have considered in the main regression, namely the regional level: Italy at the time had 16 regions (three of which had only one province each).

Table A3:
Spatial autocorrelation: robustness results

	(M1) Bartlett 200km	(M2) Uniform 200km	(M3) Bartlett 100km	(M4) Bartlett 150km	(M5) Regional clustering
Smoking tobacco	–1.293***	-1.293***	-1.293***	-1.293***	-1.293***
	(0.231)	(0.236)	(0.231)	(0.230)	(0.262)
Snuff	-0.400**	-0.400*	-0.400**	-0.400**	-0.400
	(0.189)	(0.211)	(0.182)	(0.183)	(0.334)
Father's height	0.125***	0.125***	0.125***	0.125***	0.125***
	(0.042)	(0.038)	(0.042)	(0.042)	(0.036)
Child mortality	-0.002**	-0.002**	-0.002**	-0.002**	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Per capita income	-0.034	-0.034	-0.034	-0.034	-0.034
	(0.182)	(0.178)	(0.182)	(0.182)	(0.299)
Literacy rate	4.102	4.102	4.102	4.102	4.102
	(2.624)	(2.838)	(2.556)	(2.573)	(3.358)
Railways endowment	0.134***	0.134***	0.134***	0.134***	0.134**
	(0.046)	(0.042)	(0.048)	(0.047)	(0.048)
Industrial output	0.006*	0.006*	0.006*	0.006*	0.006
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Urbanisation rate	0.102**	0.102*	0.102**	0.102**	0.102
	(0.051)	(0.052)	(0.050)	(0.050)	(0.076)
Observations	1074	1074	1074	1074	1074

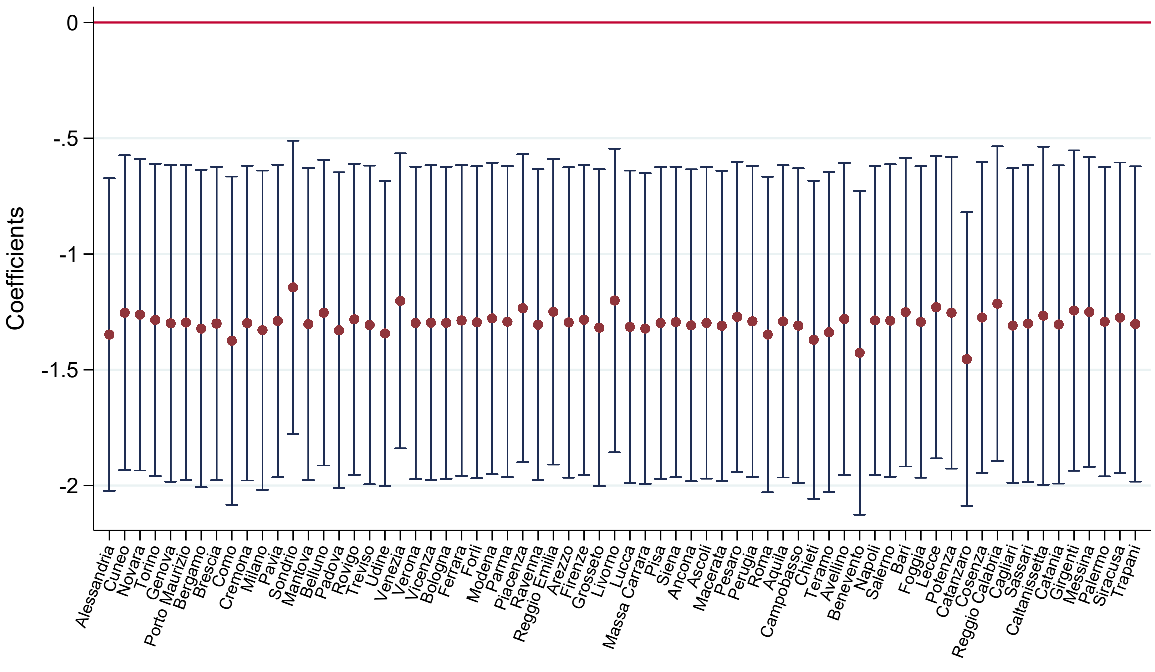
Notes: *,**,*** denote significance at the 10, 5, and 1 percent level. Data and explanatory variables as in Table 2. The estimated coefficients are identical in all the columns, as they differ only in the spatial clusturing. Column (M1) reproduces Column (M2) in Table 3, where the kernel is Bartlett with a maximum distance of 200km. In Column (M2) the kernel is uniform rather than Bartlett. The next two columns shrink the kernel by shortening the maximum distance. Finally in column (M5), the clustering is at the regional level.

Figure A2:
Urban centres in 1901 Italy



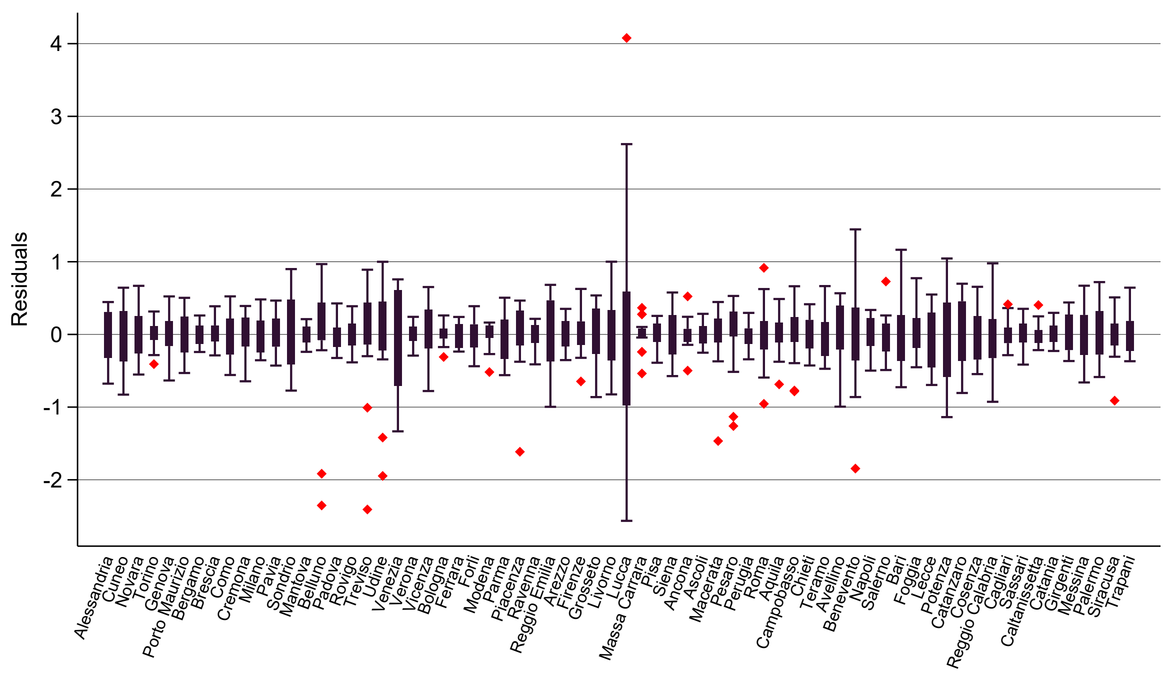
Note: Urban centres with at least 20,000 inhabitants in 1911; different size circles denote centres with population exceeding the corresponding value. Source: authors' elaboration from data in Giusti (1913).

Figure A3:
Results from leave-one-out regressions



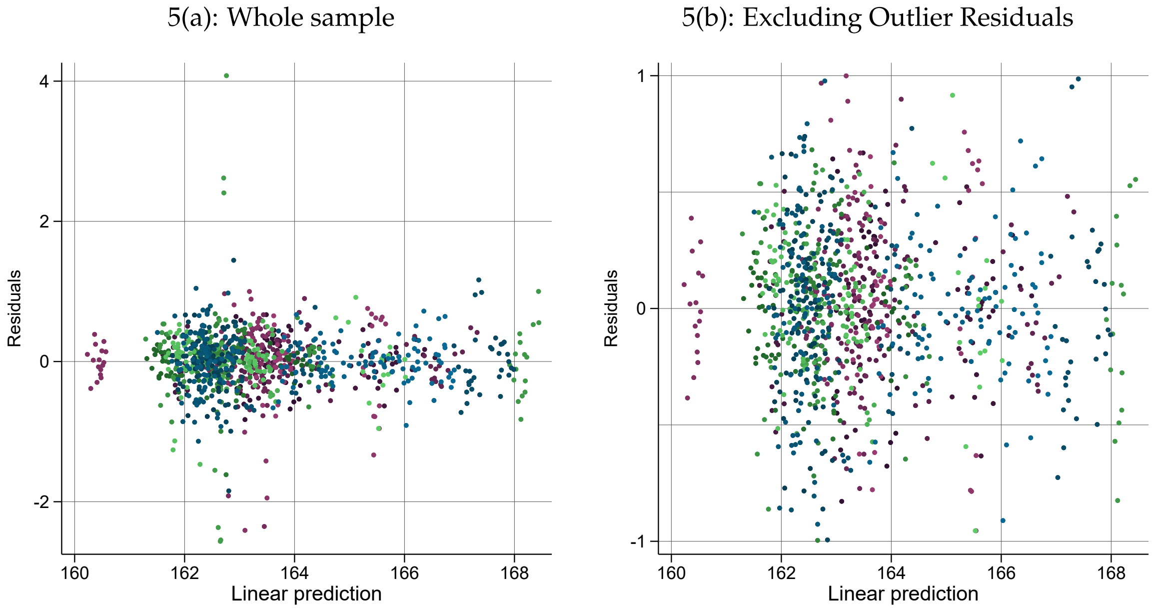
Note: The figure depicts as the red dots the coefficient for smoking tobacco obtained by running the main regression (the specification of (2) in column (M3) of Table 2) excluding the province on the horizontal axis, with the vertical thin bars measuring the 95% confidence interval. Recall that two provinces (Naples and Caserta), are aggregated into one.

Figure A4:
Province box plots from the regression residuals



Note: Box plots from the residuals in each province from the main regression, Column (M3) in Table 2. The red diamonds are the outliers, the box include the values from the 25th and the 75th percentile, and the lines the adjacent values. Provinces are ordered in the same way as in Figure 6.

Figure A5:
Residual vs Fit plot



Note: The scattergrams map the residuals from the main regression, Column (M3) in Table 2, against the fitted values: for every observation on the LHS, and only the observations where the residual is less than 1 in absolute value on the RHS. The colour of each dot is the same for all the observations of the same province, and is also the same as the corresponding dots in Figure 6.